8

Multilayer Film Applications

8.1 *Multilayer Dielectric Structures at Oblique Incidence*

Using the matching and propagation matrices for transverse fields that we discussed in Sec. 7.3, we derive here the layer recursions for multiple dielectric slabs at oblique incidence.

Fig. 8.1.1 shows such a multilayer structure. The layer recursions relate the various field quantities, such as the electric fields and the reflection responses, at the *left* of each interface.
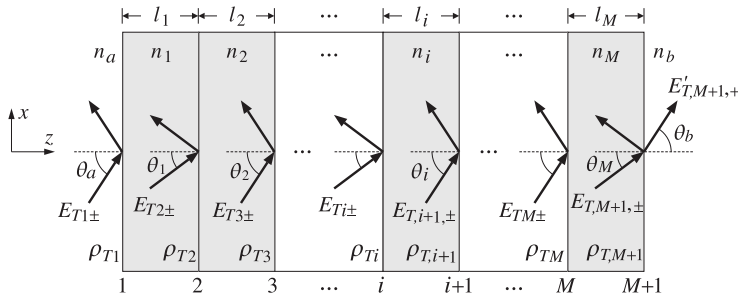

Fig. 8.1.1 Oblique incidence on multilayer dielectric structure.

We assume that there are no incident fields from the right side of the structure. The reflection/refraction angles in each medium are related to each other by Snel's law applied to each of the $M + 1$ interfaces:

$$n_a \sin \theta_a = n_i \sin \theta_i = n_b \sin \theta_b \ , \quad i = 1, 2, \dots, M \qquad (8.1.1)$$

It is convenient also to define by Eq. (7.3.8) the propagation phases or *phase thicknesses* for each of the M layers, that is, the quantities $\delta_i = k_{zi} l_i$. Using $k_{zi} = k_0 n_i \cos \theta_i$, where k_0 is the free-space wavenumber, $k_0 = \omega/c_0 = 2\pi f/c_0 = 2\pi/\lambda$, we have for $i = 1, 2, \dots, M$:

$$\delta_i = \frac{\omega}{c_0} n_i l_i \cos \theta_i = \frac{2\pi}{\lambda} n_i l_i \cos \theta_i = \frac{2\pi}{\lambda} l_i n_i \sqrt{1 - \frac{n_a^2 \sin^2 \theta_a}{n_i^2}} \qquad (8.1.2)$$

where we used Eq. (8.1.1) to write $\cos \theta_i = \sqrt{1 - \sin^2 \theta_i} = \sqrt{1 - n_a^2 \sin^2 \theta_a/n_i^2}$. The transverse reflection coefficients at the $M + 1$ interfaces are defined as in Eq. (6.1.1):

$$\rho_{Ti} = \frac{n_{T,i-1} - n_{Ti}}{n_{T,i-1} + n_{Ti}} \ , \quad i = 1, 2, \dots, M + 1 \qquad (8.1.3)$$

where we set $n_{T0} = n_{Ta}$, as in Sec. 6.1. and $n_{T,M+1} = n_{Tb}$. The transverse refractive indices are defined in each medium by Eq. (7.2.13):

$$n_{Ti} = \begin{cases} \dfrac{n_i}{\cos \theta_i} \ , & \text{TM polarization} \\ n_i \cos \theta_i \ , & \text{TE polarization} \end{cases} \ , \quad i = a, 1, 2, \dots, M, b \qquad (8.1.4)$$

To obtain the layer recursions for the electric fields, we apply the propagation matrix (7.3.5) to the fields at the left of interface $i + 1$ and propagate them to the right of the interface i, and then, apply a matching matrix (7.3.11) to pass to the left of that interface:

$$\begin{bmatrix} E_{Ti+} \\ E_{Ti-} \end{bmatrix} = \frac{1}{\tau_{Ti}} \begin{bmatrix} 1 & \rho_{Ti} \\ \rho_{Ti} & 1 \end{bmatrix} \begin{bmatrix} e^{j\delta_i} & 0 \\ 0 & e^{-j\delta_i} \end{bmatrix} \begin{bmatrix} E_{T,i+1,+} \\ E_{T,i+1,-} \end{bmatrix}$$

Multiplying the matrix factors, we obtain:

$$\begin{bmatrix} E_{Ti+} \\ E_{Ti-} \end{bmatrix} = \frac{1}{\tau_{Ti}} \begin{bmatrix} e^{j\delta_i} & \rho_{Ti} e^{-j\delta_i} \\ \rho_{Ti} e^{j\delta_i} & e^{-j\delta_i} \end{bmatrix} \begin{bmatrix} E_{T,i+1,+} \\ E_{T,i+1,-} \end{bmatrix}, \quad i = M, M - 1, \dots, 1 \qquad (8.1.5)$$

This is identical to Eqs. (6.1.2) with the substitutions $k_i l_i \to \delta_i$ and $\rho_i \to \rho_{Ti}$. The recursion is initialized at the left of the $(M + 1)$st interface by performing an additional matching to pass to the right of that interface:

$$\begin{bmatrix} E_{T,M+1,+} \\ E_{T,M+1,-} \end{bmatrix} = \frac{1}{\tau_{T,M+1}} \begin{bmatrix} 1 & \rho_{T,M+1} \\ \rho_{T,M+1} & 1 \end{bmatrix} \begin{bmatrix} E'_{T.M+1,+} \\ 0 \end{bmatrix} \qquad (8.1.6)$$

It follows now from Eq. (8.1.5) that the reflection responses, $\Gamma_{Ti} = E_{Ti-}/E_{Ti+}$, will satisfy the identical recursions as Eq. (6.1.5):

$$\Gamma_{Ti} = \frac{\rho_{Ti} + \Gamma_{T,i+1} e^{-2j\delta_i}}{1 + \rho_{Ti} \Gamma_{T,i+1} e^{-2j\delta_i}} \ , \quad i = M, M - 1, \dots, 1 \qquad (8.1.7)$$

and initialized at $\Gamma_{T,M+1} = \rho_{T,M+1}$. Similarly, we obtain the following recursions for the total transverse electric and magnetic fields at each interface (they are continuous across each interface):

$$\begin{bmatrix} E_{Ti} \\ H_{Ti} \end{bmatrix} = \begin{bmatrix} \cos \delta_i & j\eta_{Ti} \sin \delta_i \\ j\eta_{Ti}^{-1} \sin \delta_i & \cos \delta_i \end{bmatrix} \begin{bmatrix} E_{T,i+1} \\ H_{T,i+1} \end{bmatrix}, \quad i = M, M - 1, \dots, 1 \qquad (8.1.8)$$

where η_{Ti} are the transverse characteristic impedances defined by Eq. (7.2.12) and related to the refractive indices by $\eta_{Ti} = \eta_0/n_{Ti}$. The wave impedances, $Z_{Ti} = E_{Ti}/H_{Ti}$, satisfy the following recursions initialized by $Z_{T,M+1} = \eta_{Tb}$:

$$\boxed{Z_{Ti} = \eta_{Ti}\frac{Z_{T,i+1} + j\eta_{Ti}\tan\delta_i}{\eta_{Ti} + jZ_{T,i+1}\tan\delta_i}}, \quad i = M, M-1, \dots, 1 \tag{8.1.9}$$

The MATLAB function `multidiel` that was introduced in Sec. 6.1 can also be used in the oblique case with two extra input arguments: the incidence angle from the left and the polarization type, TE or TM. Its full usage is as follows:

```
[Gamma1,Z1] = multidiel(n,L,lambda,theta,pol);     % multilayer dielectric structure
```

where `theta` is the angle $\theta = \theta_a$ and `pol` is one of the strings `'te'` or `'tm'`. If the angle and polarization arguments are omitted, the function defaults to normal incidence for which TE and TM are the same. The other parameters have the same meaning as in Sec. 6.1.

In using this function, it is convenient to normalize the wavelength λ and the optical lengths $n_i l_i$ of the layers to some reference wavelength λ_0. The frequency f will be normalized to the corresponding reference frequency $f_0 = c_0/\lambda_0$.

Defining the normalized thicknesses $L_i = n_i l_i/\lambda_0$, so that $n_i l_i = L_i \lambda_0$, and noting that $\lambda_0/\lambda = f/f_0$, we may write the phase thicknesses (8.1.2) in the normalized form:

$$\boxed{\delta_i = 2\pi\frac{\lambda_0}{\lambda}L_i\cos\theta_i = 2\pi\frac{f}{f_0}L_i\cos\theta_i}, \quad i = 1, 2, \dots, M \tag{8.1.10}$$

Typically, but not necessarily, the L_i are chosen to be quarter-wavelength long at λ_0, that is, $L_i = 1/4$. This way the same multilayer design can be applied equally well at microwave or at optical frequencies. Once the wavelength scale λ_0 is chosen, the physical lengths of the layers l_i can be obtained from $l_i = L_i\lambda_0/n_i$.

8.2 Lossy Multilayer Structures

The `multidiel` function can be revised to handle lossy media. The reflection response of the multilayer structure is still computed from Eq. (8.1.7) but with some changes. In Sec. 7.7 we discussed the general case when either one or both of the incident and transmitted media are lossy.

In the notation of Fig. 8.1.1, we may assume that the incident medium n_a is lossless and all the other ones, n_i, $i = 1, 2, \dots, M, b$, are lossy (and nonmagnetic). To implement `multidiel`, one needs to know the real and imaginary parts of n_i as functions of frequency, that is, $n_i(\omega) = n_{Ri}(\omega) - jn_{Ii}(\omega)$, or equivalently, the complex dielectric constants of the lossy media:

$$\epsilon_i(\omega) = \epsilon_{Ri}(\omega) - j\epsilon_{Ii}(\omega), \quad i = 1, 2, \dots, M, b$$

$$n_i(\omega) = \sqrt{\frac{\epsilon_i(\omega)}{\epsilon_0}} = \sqrt{\frac{\epsilon_{Ri}(\omega) - j\epsilon_{Ii}(\omega)}{\epsilon_0}} = n_{Ri}(\omega) - jn_{Ii}(\omega) \tag{8.2.1}$$

Snel's law given in Eq. (8.1.1) remains valid, except now the angles θ_i and θ_b are complex valued because n_i, n_b are. One can still define the transverse refractive indices n_{Ti} through Eq. (8.1.4) using the complex-valued n_i, and $\cos\theta_i$ given by:

$$\cos\theta_i = \sqrt{1 - \sin^2\theta_i} = \sqrt{1 - \frac{n_a^2\sin^2\theta_a}{n_i^2}}, \quad i = a, 1, 2\dots, M, b \tag{8.2.2}$$

The reflection coefficients defined in Eq. (8.1.3) are equivalent to those given in Eq. (7.7.2) for the case of arbitrary incident and transmitted media.

The phase thicknesses δ_i now become complex-valued and are given by $\delta_i = k_{zi}l_i$, where k_{zi} is computed as follows. From Snel's law we have $k_{xi} = k_{xa} = \omega\sqrt{\mu_0\epsilon_0}n_a\sin\theta_a = k_0 n_a \sin\theta_a$, where $k_0 = \omega\sqrt{\mu_0\epsilon_0} = \omega/c_0$ is the free-space wave number. Then,

$$k_{zi} = \sqrt{\omega^2\mu_0\epsilon_i - k_{xi}^2} = \frac{\omega}{c_0}\sqrt{n_i^2 - n_a^2\sin^2\theta_a}, \quad i = a, 1, \dots, M, b \tag{8.2.3}$$

Thus, the complex phase thicknesses are given by:

$$\delta_i = k_{zi}l_i = \frac{\omega l_i}{c_0}\sqrt{n_i^2 - n_a^2\sin^2\theta_a}, \quad i = 1, 2, \dots, M \tag{8.2.4}$$

Writing $c_0 = f_0\lambda_0$ for some reference frequency and wavelength, we may re-express (8.2.4) in terms of the normalized frequency and normalized physical lengths:

$$\delta_i = k_{zi}l_i = 2\pi\frac{f}{f_0}\frac{l_i}{\lambda_0}\sqrt{n_i^2 - n_a^2\sin^2\theta_a}, \quad i = 1, 2, \dots, M \tag{8.2.5}$$

To summarize, given the complex $n_i(\omega)$ as in Eq. (8.2.1) at each desired value of ω, we calculate $\cos\theta_i$ from Eq. (8.2.2), n_{Ti} and ρ_{Ti} from Eqs. (8.1.4) and (8.1.3), and thicknesses δ_i from Eq. (8.2.5). Then, we use (8.1.7) to calculate the reflection response. The MATLAB function `multidiel2` implements these steps, with usage:

```
[Gamma1,Z1] = multidiel2(n,l,f,theta,pol);     % lossy multilayer structure
```

Once Γ_1 is determined, one may calculate the power entering each layer as well as the power lost within each layer. The time-averaged power per unit area entering the ith layer is the z-component of the Poynting vector, which is given in terms of the transverse E, H fields as follows:

$$\mathcal{P}_i = \frac{1}{2}\text{Re}\left(E_{Ti}H_{Ti}^*\right), \quad i = 1, 2, \dots, M \tag{8.2.6}$$

The power absorbed within the ith layer is equal to the difference of the power entering the layer and the power leaving it:

$$\mathcal{P}_i^{\text{loss}} = \mathcal{P}_i - \mathcal{P}_{i+1}, \quad i = 1, 2, \dots, M \tag{8.2.7}$$

The transverse fields can be calculated by inverting the recursion (8.1.8), that is,

$$\begin{bmatrix} E_{T,i+1} \\ H_{T,i+1} \end{bmatrix} = \begin{bmatrix} \cos\delta_i & -j\eta_{Ti}\sin\delta_i \\ -j\eta_{Ti}^{-1}\sin\delta_i & \cos\delta_i \end{bmatrix} \begin{bmatrix} E_{Ti} \\ H_{Ti} \end{bmatrix}, \quad i = 1, 2, \dots, M \tag{8.2.8}$$

The recursion is initialized with the fields E_{T1}, H_{T1} at the first interface. These can be calculated with the help of Γ_1:

$$E_{T1} = E_{T1+} + E_{T1-} = E_{T1+} (1 + \Gamma_1)$$

$$H_{T1} = \frac{1}{\eta_{Ta}} (E_{T1+} - E_{T1-}) = \frac{1}{\eta_{Ta}} E_{T1+} (1 - \Gamma_1)$$

(8.2.9)

where $\eta_{Ta} = \eta_0/n_{Ta}$. The field E_{T1+} is the transverse component of the incident field. If we denote the total incident field by E_{in}, then E_{T1+} will be given by:

$$E_{T1+} = \begin{cases} E_{in}, & \text{TE case} \\ E_{in} \cos \theta_a, & \text{TM case} \end{cases}$$

(8.2.10)

The total incident power (along the direction of the incident wave vector), its z-component, and the power entering the first layer will be given as follows (in both the TE and TM cases):

$$\mathcal{P}_{in} = \frac{1}{2\eta_a} |E_{in}|^2, \quad \mathcal{P}_{in,z} = \mathcal{P}_{in} \cos \theta_a, \quad \mathcal{P}_1 = \mathcal{P}_{in,z} (1 - |\Gamma_1|^2)$$

(8.2.11)

where $\eta_a = \eta_0/n_a$. Thus, one can start with $E_{in} = \sqrt{2\eta_a \mathcal{P}_{in}}$, if the incident power is known.

8.3 Single Dielectric Slab

Many features of oblique incidence on multilayer slabs can be clarified by studying the single-slab case, shown in Fig. 8.3.1. Assuming that the media to the left and right are the same, $n_a = n_b$, it follows that $\theta_b = \theta_a$ and also that $\rho_{T1} = -\rho_{T2}$. Moreover, Snel's law implies $n_a \sin \theta_a = n_1 \sin \theta_1$.
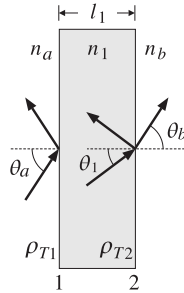

Fig. 8.3.1 Oblique incidence on single dielectric slab.

Because there are no incident fields from the right, the reflection response at the left of interface-2 is: $\Gamma_{T2} = \rho_{T2} = -\rho_{T1}$. It follows from Eq. (8.1.7) that the reflection response at the left of interface-1 will be:

$$\Gamma_{T1} = \frac{\rho_{T1} + \rho_{T2} e^{-2j\delta_1}}{1 + \rho_{T1} \rho_{T2} e^{-2j\delta_1}} = \frac{\rho_{T1} (1 - e^{-2j\delta_1})}{1 - \rho_{T1}^2 e^{-2j\delta_1}}$$

(8.3.1)

These are analogous to Eqs. (5.4.6) and (5.5.4), i.e., the normal and oblique incidence cases differ only in the definitions of the reflection coefficients. According to Eq. (8.1.10), the phase thickness can be written in the following normalized form, where $L_1 = n_1 l_1 / \lambda_0$:

$$\delta_1 = 2\pi \frac{\lambda_0}{\lambda} L_1 \cos \theta_1 = 2\pi \frac{f}{f_0} L_1 \cos \theta_1 = \pi \frac{f}{f_1}$$

(8.3.2)

$$f_1 = \frac{f_0}{2L_1 \cos \theta_1}$$

(8.3.3)

At frequencies that are integral multiples of f_1, $f = mf_1$, the reflection response vanishes because $2\delta_1 = 2\pi(mf_1)/f_1 = 2\pi m$ and $e^{-2j\delta_1} = 1$. Similarly, at the half-integral multiples, $f = (m + 0.5)f_1$, the response is maximum because $e^{-2j\delta_1} = -1$.

Because f_1 depends inversely on $\cos \theta_1$, then as the angle of incidence θ_a increases, $\cos \theta_1$ will decrease and f_1 will shift towards higher frequencies. The maximum shift will occur when θ_1 reaches its maximum refraction value $\theta_{1c} = \text{asin}(n_a/n_1)$ (assuming $n_a < n_1$.)

Similar shifts occur for the 3-dB width of the reflection response notches. By the same calculation that led to Eq. (5.5.9), we find for the 3-dB width with respect to the variable δ_1:

$$\tan\left(\frac{\Delta\delta_1}{2}\right) = \frac{1 - \rho_{T1}^2}{1 + \rho_{T1}^2}$$

Setting $\Delta\delta_1 = \pi\Delta f/f_1$, we solve for the 3-dB width in frequency:

$$\Delta f = \frac{2f_1}{\pi} \text{atan}\left(\frac{1 - \rho_{T1}^2}{1 + \rho_{T1}^2}\right)$$

(8.3.4)

The left/right bandedge frequencies are $f_1 \pm \Delta f/2$. The dependence of Δf on the incidence angle θ_a is more complicated here because ρ_{T1} also depends on it.

In fact, as θ_a tends to its grazing value $\theta_a \rightarrow 90^\circ$, the reflection coefficients for either polarization have the limit $|\rho_{T1}| \rightarrow 1$, resulting in zero bandwidth Δf. On the other hand, at the Brewster angle, $\theta_{aB} = \text{atan}(n_1/n_a)$, the TM reflection coefficient vanishes, resulting in maximum bandwidth. Indeed, because $\text{atan}(1) = \pi/4$, we have $\Delta f_{max} = 2f_1 \text{atan}(1)/\pi = f_1/2$.

Fig. 8.3.2 illustrates some of these properties. The refractive indices were $n_a = n_b = 1$ and $n_1 = 1.5$. The optical length of the slab was taken to be half-wavelength at the reference wavelength λ_0, so that $n_1 l_1 = 0.5\lambda_0$, or, $L_1 = 0.5$.

The graphs show the TE and TM reflectances $|\Gamma_{T1}(f)|^2$ as functions of frequency for the angles of incidence $\theta_1 = 75^\circ$ and $\theta_a = 85^\circ$. The normal incidence case is also included for comparison.

The corresponding refracted angles were $\theta_1 = \text{asin}(n_a \text{asin}(\theta_a)/n_1) = 40.09^\circ$ and $\theta_1 = 41.62^\circ$. Note that the maximum refracted angle is $\theta_{1c} = 41.81^\circ$, and the Brewster angle, $\theta_{aB} = 56.31^\circ$.
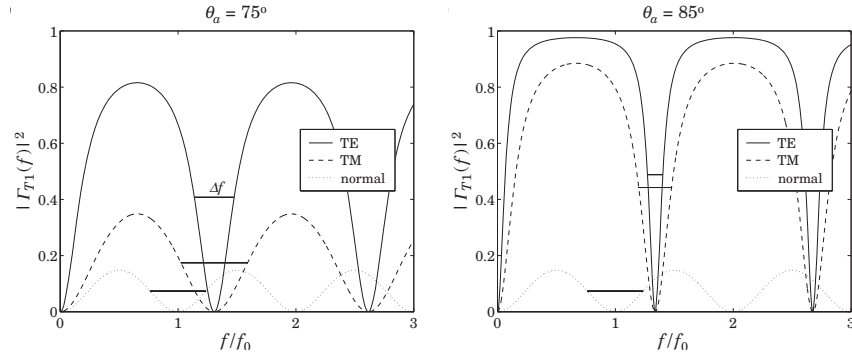
Fig. 8.3.2 TE and TM reflectances of half-wavelength slab.

The notch frequencies were $f_1 = f_0/(2L_1 \cos\theta_1) = 1.31f_0$ and $f_1 = 1.34f_0$ for the angles $\theta_a = 75^\circ$ and 85°. At normal incidence we have $f_1 = f_0/(2L_1) = f_0$, because $L_1 = 0.5$.

The graphs also show the 3-dB widths of the notches, calculated from Eq. (8.3.4). The reflection responses were computed with the help of the function `multidiel` with the typical MATLAB code:

```
na = 1; nb = 1;
n1 = 1.5; L1 = 0.5;

f = linspace(0,3,401);
theta = 75;

G0 = abs(multidiel([na,n1,nb], L1, 1./f)).^2;
Ge = abs(multidiel([na,n1,nb], L1, 1./f, theta, 'te')).^2;
Gm = abs(multidiel([na,n1,nb], L1, 1./f, theta, 'tm')).^2;
```

The shifting of the notch frequencies and the narrowing of the notch widths is evident from the graphs. Had we chosen $\theta_a = \theta_{aB} = 56.31^\circ$, the TM response would have been identically zero because of the factor ρ_{T1} in Eq. (8.3.1).

The single-slab case is essentially a simplified version of a Fabry-Perot interferometer [638], used as a spectrum analyzer. At multiples of f_1, there are narrow transmittance bands. Because f_1 depends on $f_0/\cos\theta_1$, the interferometer serves to separate different frequencies f_0 in the input by mapping them onto different angles θ_1.

Next, we look at three further applications of the single-slab case: (a) frustrated internal reflection, (b) surface plasmon resonance, and (c) the perfect lens property of negative-index media.

8.4 Frustrated Total Internal Reflection

As we discussed in Sec. 7.5, when a wave is incident at an angle greater than the total internal reflection (TIR) angle from an optically denser medium n_a onto a rarer medium

n_b, with $n_a > n_b$, then there is 100 percent reflection. The transmitted field into the rarer medium n_b is evanescent, decaying exponentially with distance.

However, if an object or another medium is brought near the interface from the n_b side, the evanescent field is "frustrated" and can couple into a propagating wave. For example, if another semi-infinite medium n_a is brought close to the interface, then the evanescent field can "tunnel" through to the other side, emerging as an attenuated version of the incident wave. This effect is referred to as "frustrated" total internal reflection.

Fig. 8.4.1 shows how this may be realized with two 45° prisms separated by a small air gap. With $n_a = 1.5$ and $n_b = 1$, the TIR angle is $\theta_c = \mathrm{asin}(n_b/n_a) = 41.8^\circ$, therefore, $\theta = 45^\circ > \theta_c$. The transmitted fields into the air gap reach the next prism with an attenuated magnitude and get refracted into a propagating wave that emerges at the same angle θ.
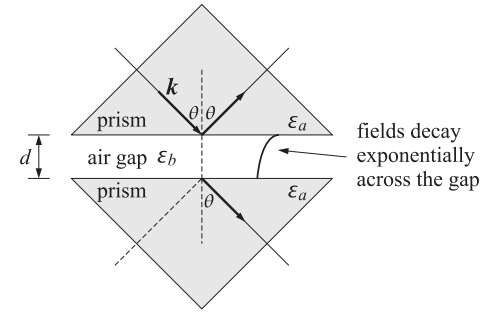

Fig. 8.4.1 Frustrated total internal reflection between two prisms separated by an air gap.

Fig. 8.4.2 shows an equivalent problem of two identical semi-infinite media n_a, separated by a medium n_b of length d. Let $\varepsilon_a = n_a^2$, $\varepsilon_b = n_b^2$ be the relative dielectric constants. The components of the wavevectors in media n_a and n_b are:

$$k_x = k_0 n_a \sin\theta, \quad k_0 = \frac{\omega}{c_0}$$

$$k_{za} = \sqrt{k_0^2 n_a^2 - k_x^2} = k_0 n_a \cos\theta \tag{8.4.1}$$

$$k_{zb} = \begin{cases} k_0\sqrt{n_b^2 - n_a^2 \sin^2\theta}, & \text{if } \theta \le \theta_c \\ -jk_0\sqrt{n_a^2 \sin^2\theta - n_b^2} = -j\alpha_{zb}, & \text{if } \theta \ge \theta_c \end{cases}$$

where $\sin\theta_c = n_b/n_a$. Because of Snel's law, the k_x component is preserved across the interfaces. If $\theta > \theta_c$, then k_{zb} is pure imaginary, that is, evanescent. The transverse reflection and transmission responses are:

$$\Gamma = \frac{\rho_a + \rho_b e^{-2jk_{zb}d}}{1 + \rho_a\rho_b e^{-2jk_{zb}d}} = \frac{\rho_a(1 - e^{-2jk_{zb}d})}{1 - \rho_a^2 e^{-2jk_{zb}d}}$$

$$T = \frac{\tau_a\tau_b e^{-jk_{zb}d}}{1 + \rho_a\rho_b e^{-2jk_{zb}d}} = \frac{(1 - \rho_a^2)e^{-jk_{zb}d}}{1 - \rho_a^2 e^{-2jk_{zb}d}} \tag{8.4.2}$$
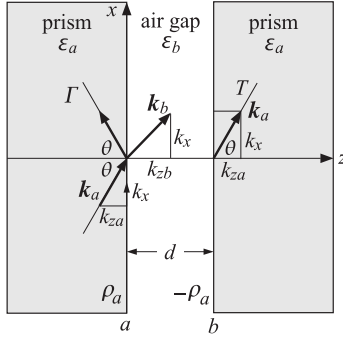
Fig. 8.4.2 Frustrated total internal reflection.

where ρ_a, ρ_b are the transverse reflection coefficients at the a, b interfaces and $\tau_a = 1 + \rho_a$ and $\tau_b = 1 + \rho_b$ are the transmission coefficients, and we used the fact that $\rho_b = -\rho_a$ because the media to the left and right of the slab are the same. For the two polarizations, ρ_a is given in terms of the above wavevector components as follows:

$$\rho_a^{TE} = \frac{k_{za} - k_{zb}}{k_{za} + k_{zb}}, \quad \rho_a^{TM} = \frac{k_{zb}\varepsilon_a - k_{za}\varepsilon_b}{k_{zb}\varepsilon_a + k_{za}\varepsilon_b} \tag{8.4.3}$$

For $\theta \le \theta_c$, the coefficients ρ_a are real-valued, and for $\theta \ge \theta_c$, they are unimodular, $|\rho_a| = 1$, given explicitly by

$$\rho_a^{TE} = \frac{n_a \cos\theta + j\sqrt{n_a^2 \sin^2\theta - n_b^2}}{n_a \cos\theta - j\sqrt{n_a^2 \sin^2\theta - n_b^2}}, \quad \rho_a^{TM} = \frac{-jn_a\sqrt{n_a^2 \sin^2\theta - n_b^2} - n_b^2 \cos\theta}{-jn_a\sqrt{n_a^2 \sin^2\theta - n_b^2} + n_b^2 \cos\theta} \tag{8.4.4}$$

For all angles, it can be shown that $1 - |\Gamma|^2 = |T|^2$, which represents the amount of power that enters perpendicularly into interface a and exits from interface b. For the TIR case, Γ, T simplify into:

$$\Gamma = \frac{\rho_a(1 - e^{-2\alpha_{zb}d})}{1 - \rho_a^2 e^{-2\alpha_{zb}d}}, \quad T = \frac{(1 - \rho_a^2)e^{-\alpha_{zb}d}}{1 - \rho_a^2 e^{-2\alpha_{zb}d}}, \quad \alpha_{zb} = \frac{2\pi}{\lambda_0}\sqrt{n_a^2 \sin^2\theta - n_b^2} \tag{8.4.5}$$

where we defined the free-space wavelength through $k_0 = 2\pi/\lambda_0$. Setting $\rho_a = e^{j\phi_a}$, the magnitude responses are given by:

$$|\Gamma|^2 = \frac{\sinh^2(\alpha_{zb}d)}{\sinh^2(\alpha_{zb}d) + \sin^2\phi_a}, \quad |T|^2 = \frac{\sin^2\phi_a}{\sinh^2(\alpha_{zb}d) + \sin^2\phi_a} \tag{8.4.6}$$

For a prism with $n_a = 1.5$ and an air gap $n_b = 1$, Fig. 8.4.3 shows a plot of Eqs. (8.4.5) versus the distance d at the incidence angle $\theta = 45°$. The reflectance becomes almost 100 percent for thickness of a few wavelengths.

Fig. 8.4.4 shows the reflectance versus angle over $0 \le \theta \le 90°$ for the thicknesses $d = 0.4\lambda_0$ and $d = 0.5\lambda_0$. The TM reflection response vanishes at the Brewster angle $\theta_B = \text{atan}(n_b/n_a) = 33.69°$.
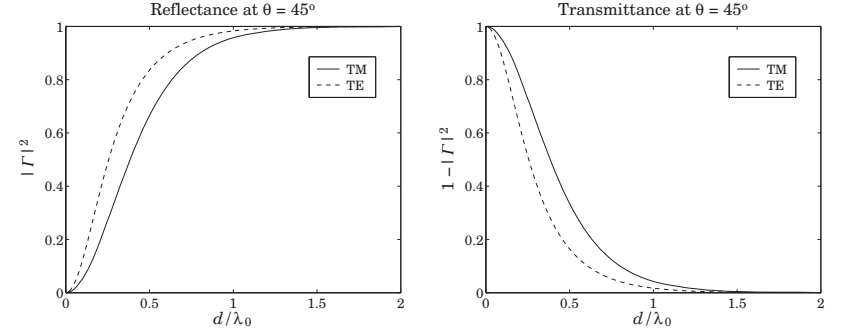
Fig. 8.4.3 Reflectance and transmittance versus thickness d.

The case $d = 0.5\lambda_0$ was chosen because the slab becomes a half-wavelength slab at normal incidence, that is, $k_{zb}d = 2\pi/2$ at $\theta = 0°$, resulting in the vanishing of Γ as can be seen from Eq. (8.4.2).

The half-wavelength condition, and the corresponding vanishing of Γ, can be required at any desired angle $\theta_0 < \theta_c$, by demanding that $k_{zb}d = 2\pi/2$ at that angle, which fixes the separation d:

$$k_{zb}d = \pi \quad \Rightarrow \quad \frac{2\pi d}{\lambda_0}\sqrt{n_b^2 - n_a^2 \sin^2\theta_0} = \pi \quad \Rightarrow \quad d = \frac{\lambda_0}{2\sqrt{n_b^2 - n_a^2 \sin^2\theta_0}}$$

Fig. 8.4.5 depicts the case $\theta_0 = 20°$, which fixes the separation to be $d = 0.5825\lambda_0$.
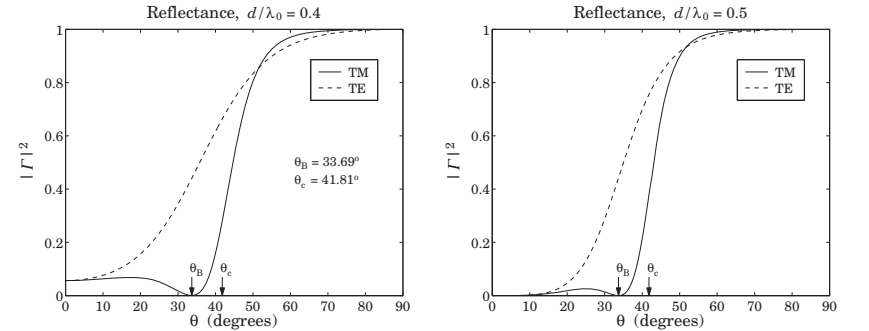

Fig. 8.4.4 Reflectance versus angle of incidence.

The fields within the air gap can be determined using the layer recursions (8.1.5). Let E_{a+} be the incident transverse field at the left side of the interface a, and $E_\pm$ the transverse fields at the right side. Using Eq. (8.1.5) and (8.1.6), we find for the TIR case:

$$E_+ = \frac{(1 + \rho_a)E_{a+}}{1 - \rho_a^2 e^{-2\alpha_{zb}d}}, \quad E_- = \frac{-\rho_a e^{-2\alpha_{zb}d}(1 + \rho_a)E_{a+}}{1 - \rho_a^2 e^{-2\alpha_{zb}d}} \tag{8.4.7}$$
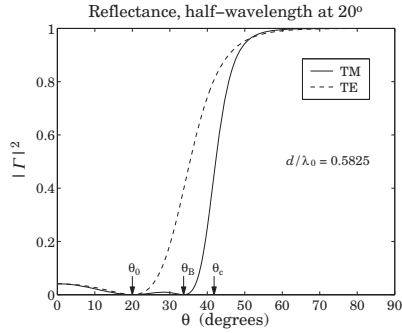
Reflectance, half–wavelength at 20°

Fig. 8.4.5   Reflectance vanishes at $\theta_0 = 20°$.

The transverse electric field within the air gap will be then $E_T(z) = E_+e^{-\alpha_{zb}z} + E_-e^{\alpha_{zb}z}$, and similarly for the magnetic field. Using (8.4.7) we find:

$$E_T(z) = \left[\frac{1 + \rho_a}{1 - \rho_a^2 e^{-2\alpha_{zb}d}}\right][e^{-\alpha_{zb}z} - \rho_a e^{-2\alpha_{zb}d}e^{\alpha_{zb}z}]E_{a+}$$

$$H_T(z) = \left[\frac{1 - \rho_a}{1 - \rho_a^2 e^{-2\alpha_{zb}d}}\right][e^{-\alpha_{zb}z} + \rho_a e^{-2\alpha_{zb}d}e^{\alpha_{zb}z}]\frac{E_{a+}}{\eta_{aT}}$$

(8.4.8)

where η_{aT} is the transverse impedance of medium n_a, that is, with $\eta_a = \eta_0/n_a$:

$$\eta_{aT} = \begin{cases} \eta_a \cos\theta_a, & \text{TM, or parallel polarization} \\ \eta_a/\cos\theta_a, & \text{TE, or perpendicular polarization} \end{cases}$$

It is straightforward to verify that the transfer of power across the gap is independent of the distance z and given by

$$\mathcal{P}_z(z) = \frac{1}{2}\operatorname{Re}[E_T(z)H_T^*(z)] = (1 - |\Gamma|^2)\frac{|E_{a+}|^2}{2\eta_{aT}}$$

Frustrated total internal reflection has several applications [556–592], such as internal reflection spectroscopy, sensors, fingerprint identification, surface plasmon resonance, and high resolution microscopy. In many of these applications, the air gap is replaced by another, possibly lossy, medium. The above formulation remains valid with the replacement $\varepsilon_b = n_b^2 \to \varepsilon_b = \varepsilon_{br} - j\varepsilon_{bi}$, where the imaginary part ε_{ri} characterizes the losses.

8.5   Surface Plasmon Resonance

We saw in Sec. 7.11 that surface plasmons are TM waves that can exist at an interface between air and metal, and that their wavenumber k_x of propagation along the interface is larger that its free-space value at the same frequency. Therefore, such plasmons cannot couple directly to plane waves incident on the interface.

However, if the incident TM plane wave is from a dielectric and from an angle that is greater than the angle of total internal reflection, then the corresponding wavenumber will be greater than its vacuum value and it could excite a plasmon wave along the interface. Fig. 8.5.1 depicts two possible configurations of how this can be accomplished.
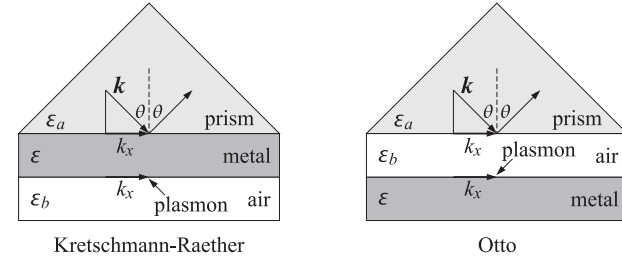

Fig. 8.5.1   Kretschmann-Raether and Otto configurations.

In the so-called Kretschmann-Raether configuration [595,598], a thin metal film of thickness of a fraction of a wavelength is sandwiched between a prism and air and the incident wave is from the prism side. In the Otto configuration [596], there is an air gap between the prism and the metal. The two cases are similar, but we will consider in greater detail the Kretschmann-Raether configuration, which is depicted in more detail in Fig. 8.5.2.
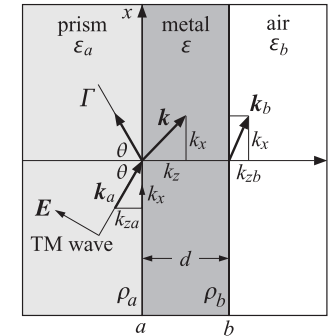

Fig. 8.5.2   Surface plasmon resonance excitation by total internal reflection.

The relative dielectric constant ε_a and refractive index n_a of the prism are related by $\varepsilon_a = n_a^2$. The air side has $\varepsilon_b = n_b^2 = 1$, but any other lossless dielectric will do as long as it satisfies $n_b < n_a$. The TIR angle is $\sin\theta_c = n_b/n_a$, and the angle of incidence from the prism side is assumed to be $\theta \geq \theta_c$ so that[†]

$$k_x = k_0 n_a \sin\theta \geq k_0 n_b, \quad k_0 = \frac{\omega}{c_0} \tag{8.5.1}$$

[†] The geometrical picture in Fig. 8.5.2 is not valid for $\theta \geq \theta_c$ because the wavevectors are complex-valued.

Because of Snel's law, the k_x component of the wavevector along the interface is preserved across the media. The z-components in the prism and air sides are given by:

$$k_{za} = \sqrt{k_0^2 n_a^2 - k_x^2} = k_0 n_a \cos\theta$$

$$k_{zb} = -j\alpha_{zb} = -j\sqrt{k_x^2 - k_0^2 n_b^2} = -jk_0\sqrt{n_a^2 \sin^2\theta - n_b^2} \tag{8.5.2}$$

where k_{zb} is pure imaginary because of the TIR assumption. Therefore, the transmitted wave into the ε_b medium attenuates exponentially like $e^{-jk_{zb}z} = e^{-\alpha_{zb}z}$.

For the metal layer, we assume that its relative dielectric constant is $\varepsilon = -\varepsilon_r - j\varepsilon_i$, with a negative real part ($\varepsilon_r > 0$) and a small negative imaginary part ($0 < \varepsilon_i \ll \varepsilon_r$) that represents losses. Moreover, in order for a surface plasmon wave to be supported on the ε–ε_b interface, we must further assume that $\varepsilon_r > \varepsilon_b$. The k_z component within the metal will be complex-valued with a dominant imaginary part:

$$k_z = -j\sqrt{k_x^2 - k_0^2\varepsilon} = -j\sqrt{k_x^2 + k_0^2(\varepsilon_r + j\varepsilon_i)} = -jk_0\sqrt{n_a^2\sin^2\theta + \varepsilon_r + j\varepsilon_i} \tag{8.5.3}$$

If there is a surface plasmon wave on the ε–ε_b interface, then as we saw in Sec. 7.7, it will be characterized by the specific values of k_x, k_z, k_{zb}:

$$k_{x0} = \beta_{x0} - j\alpha_{x0} = k_0\sqrt{\frac{\varepsilon\varepsilon_b}{\varepsilon + \varepsilon_b}}, \quad k_{z0} = -\frac{k_0\varepsilon}{\sqrt{\varepsilon + \varepsilon_b}}, \quad k_{zb0} = \frac{k_0\varepsilon_b}{\sqrt{\varepsilon + \varepsilon_b}} \tag{8.5.4}$$

Using Eq. (7.11.10), we have approximately to lowest order in ε_i:

$$\beta_{x0} = k_0\sqrt{\frac{\varepsilon_r\varepsilon_b}{\varepsilon_r - \varepsilon_b}}, \quad \alpha_{x0} = k_0\left(\frac{\varepsilon_r\varepsilon_b}{\varepsilon_r - \varepsilon_b}\right)^{3/2}\frac{\varepsilon_i}{2\varepsilon_r^2} \tag{8.5.5}$$

and similarly for k_{z0}, which has a small real part and a dominant imaginary part:

$$k_{z0} = \beta_{z0} - j\alpha_{z0}, \quad \alpha_{z0} = \frac{k_0\varepsilon_r}{\sqrt{\varepsilon_r - \varepsilon_b}}, \quad \beta_{z0} = \frac{k_0(\varepsilon_r - 2\varepsilon_b)\varepsilon_i}{(\varepsilon_r - \varepsilon_b)^{3/2}} \tag{8.5.6}$$

If the incidence angle θ is such that k_x is near the real-part of k_{x0}, that is, $k_x = k_0 n_a \sin\theta = \beta_{x0}$, then a resonance takes place exciting the surface plasmon wave. Because of the finite thickness d of the metal layer and the assumed losses ε_i, the actual resonance condition is not $k_x = \beta_{x0}$, but is modified by a small shift: $k_x = \beta_{x0} + \bar\beta_{x0}$, to be determined shortly.

At the resonance angle there is a sharp drop of the reflection response measured at the prism side. Let ρ_a, ρ_b denote the TM reflection coefficients at the ε_a–ε and ε–ε_b interfaces, as shown in Fig. 8.5.2. The corresponding TM reflection response of the structure will be given by:

$$\Gamma = \frac{\rho_a + \rho_b e^{-2jk_z d}}{1 + \rho_a\rho_b e^{-2jk_z d}} = \frac{\rho_a + \rho_b e^{-2\alpha_z d}e^{-2j\beta_z d}}{1 + \rho_a\rho_b e^{-2\alpha_z d}e^{-2j\beta_z d}} \tag{8.5.7}$$

where d is the thickness of the metal layer and $k_z = \beta_z - j\alpha_z$ is given by Eq. (8.5.3). The TM reflection coefficients are given by:

$$\rho_a = \frac{k_z\varepsilon_a - k_{za}\varepsilon}{k_z\varepsilon_a + k_{za}\varepsilon}, \quad \rho_b = \frac{k_{zb}\varepsilon - k_z\varepsilon_b}{k_{zb}\varepsilon + k_z\varepsilon_b} \tag{8.5.8}$$

where k_{za}, k_{zb} are given by (8.5.2). Explicitly, we have for $\theta \geq \theta_c$:

$$\rho_a = -\frac{\varepsilon\sqrt{k_0^2\varepsilon_a - k_x^2} + j\varepsilon_a\sqrt{k_x^2 - k_0^2\varepsilon}}{\varepsilon\sqrt{k_0^2\varepsilon_a - k_x^2} - j\varepsilon_a\sqrt{k_x^2 - k_0^2\varepsilon}} = -\frac{\varepsilon\cos\theta + jn_a\sqrt{\varepsilon_a\sin^2\theta - \varepsilon}}{\varepsilon\cos\theta - jn_a\sqrt{\varepsilon_a\sin^2\theta - \varepsilon}}$$

$$\rho_b = \frac{\varepsilon\sqrt{k_x^2 - k_0^2\varepsilon_b} - \varepsilon_b\sqrt{k_x^2 - k_0^2\varepsilon}}{\varepsilon\sqrt{k_x^2 - k_0^2\varepsilon_b} + \varepsilon_b\sqrt{k_x^2 - k_0^2\varepsilon}} = \frac{\varepsilon\sqrt{\varepsilon_a\sin^2\theta - \varepsilon_b} - \varepsilon_b\sqrt{\varepsilon_a\sin^2\theta - \varepsilon}}{\varepsilon\sqrt{\varepsilon_a\sin^2\theta - \varepsilon_b} + \varepsilon_b\sqrt{\varepsilon_a\sin^2\theta - \varepsilon}} \tag{8.5.9}$$

We note that for the plasmon resonance to be excited through such a configuration, the metal must be assumed to be slightly lossy, that is, $\varepsilon_i \neq 0$. If we assume that it is lossless with a negative real part, $\varepsilon = -\varepsilon_r$, then, ρ_a becomes a unimodular complex number, $|\rho_a| = 1$, for all angles θ, while ρ_b remains real-valued for $\theta \geq \theta_c$, and also k_z is pure imaginary, $\beta_z = 0$. Hence, it follows that:

$$|\Gamma|^2 = \frac{|\rho_a|^2 + 2\,\text{Re}(\rho_a)\rho_b e^{-2\alpha_z d} + \rho_b^2 e^{-4\alpha_z d}}{1 + 2\,\text{Re}(\rho_a)\rho_b e^{-2\alpha_z d} + |\rho_a|^2\rho_b^2 e^{-4\alpha_z d}} = 1$$

Thus, it remains flat for $\theta \geq \theta_c$. For $\theta \leq \theta_c$, ρ_a is still unimodular, and ρ_b also becomes unimodular, $|\rho_b| = 1$. Setting $\rho_a = e^{j\phi_a}$ and $\rho_b = e^{j\phi_b}$, we find for $\theta \leq \theta_c$:

$$|\Gamma|^2 = \left|\frac{e^{j\phi_a} + e^{j\phi_b}e^{-2\alpha_z d}}{1 + e^{j\phi_a}e^{j\phi_b}e^{-2\alpha_z d}}\right|^2 = \frac{1 + 2\cos(\phi_a - \phi_b)e^{-2\alpha_z d} + e^{-4\alpha_z d}}{1 + 2\cos(\phi_a + \phi_b)e^{-2\alpha_z d} + e^{-4\alpha_z d}} \tag{8.5.10}$$

which remains almost flat, exhibiting a slight variation with the angle for $\theta \leq \theta_c$.

As an example, consider a quartz prism with $n_a = 1.5$, coated with a silver film of thickness of $d = 50$ nm, and air on the other side $\varepsilon_b = 1$. The relative refractive index of the metal is taken to be $\varepsilon = -16 - 0.5j$ at the free-space wavelength of $\lambda_0 = 632$ nm. The corresponding free-space wave number is $k_0 = 2\pi/\lambda_0 = 9.94$ rad/μm.

Fig. 8.5.3 shows the TM reflection response (8.5.7) versus angle. The TIR angle is $\theta_c = \text{asin}(n_b/n_a) = 41.81°$. The plasmon resonance occurs at the angle $\theta_{\text{res}} = 43.58°$. The graph on the right shows an expanded view over the angle range $41° \leq \theta \leq 45°$. Both angles θ_c and θ_{res} are indicated on the graphs as black dots.

The computation can be carried out with the help of the MATLAB function `multidiel1.m`, or alternatively `multidiel.m`, with the sample code:

```
na = 1.5;  ea = na^2;              % prism side
er = 16;   ei = 0.5; ep = -er-j*ei;  % silver layer
nb = 1;    eb = nb^2;              % air side
d = 50; la0 = 632;                 % in units of nanometers

th = linspace(0,89,8901);          % incident angle in degrees

n1 = sqrte(ep);                    % evanescent SQRT, needed if εi = 0
L1 = n1*d/la0;                     % complex optical length in units of λ0
n = [na, n1, nb];                  % input to multidiel1

for i=1:length(th),                % TM reflectance
   Ga(i) = abs(multidiel1(n, L1, 1, th(i), 'tm')).^2;   % at λ/λ0 = 1
end

plot(th,Ga);
```
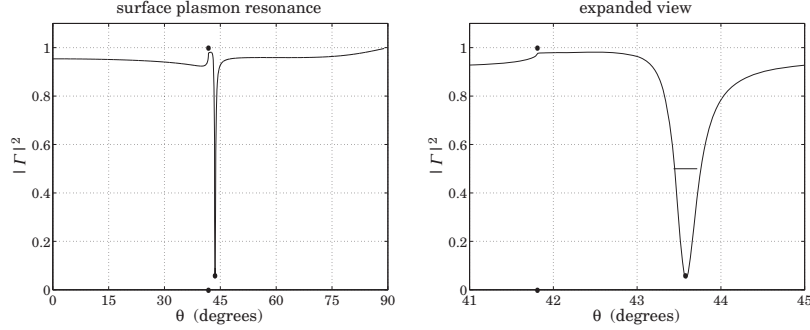
surface plasmon resonance

expanded view

Fig. 8.5.3 Surface plasmon resonance.

Fig. 8.5.4 shows the reflection response when the metal is assumed to be lossless with $\varepsilon = -16$, all the other parameters being the same. As expected, there is no resonance and the reflectance stays flat for $\theta \geq \theta_c$, with mild variation for $\theta < \theta_c$.
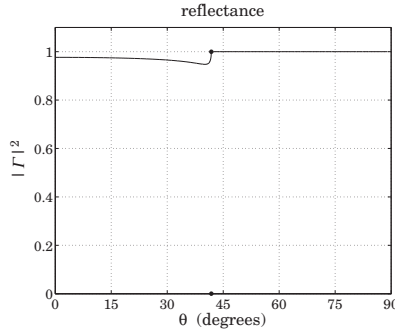
reflectance

Fig. 8.5.4 Absence of resonance when metal is assumed to be lossless.

Let E_{a+}, E_{a-} be the forward and backward transverse electric fields at the left side of interface a. The fields at the right side of the interface can be obtained by inverting the matching matrix:

$$
\begin{bmatrix} E_{a+} \\ E_{a-} \end{bmatrix} = \frac{1}{1 + \rho_a} \begin{bmatrix} 1 & \rho_a \\ \rho_a & 1 \end{bmatrix} \begin{bmatrix} E_+ \\ E_- \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} E_+ \\ E_- \end{bmatrix} = \frac{1}{1 - \rho_a} \begin{bmatrix} 1 & -\rho_a \\ -\rho_a & 1 \end{bmatrix} \begin{bmatrix} E_{a+} \\ E_{a-} \end{bmatrix}
$$

Setting $E_{a-} = \Gamma E_{a+}$, with Γ given by Eq. (8.5.7), we obtain:

$$
E_+ = \frac{1 - \rho_a \Gamma}{1 - \rho_a} E_{a+} = \frac{(1 + \rho_a) E_{a+}}{1 + \rho_a \rho_b e^{-2jk_z d}}
$$

$$
E_- = \frac{-\rho_a + \Gamma}{1 - \rho_a} E_{a+} = \frac{\rho_b e^{-2jk_z d} (1 + \rho_a) E_{a+}}{1 + \rho_a \rho_b e^{-2jk_z d}}
$$

The transverse electric and magnetic fields within the metal layer will be given by:

$$
E_T(z) = E_+ e^{-jk_z z} + E_- e^{jk_z z}, \quad H_T(z) = \frac{1}{\eta_T} \left[E_+ e^{-jk_z z} - E_- e^{jk_z z} \right]
$$

Using the relationship $\eta_T / \eta_{aT} = (1 + \rho_a)/(1 - \rho_a)$, we have:

$$
E_T(z) = \left[\frac{1 + \rho_a}{1 + \rho_a \rho_b e^{-2jk_z d}} \right] \left[e^{-jk_z z} + \rho_b e^{-2jk_z d} e^{jk_z z} \right] E_{a+}
$$

$$
H_T(z) = \left[\frac{1 - \rho_a}{1 + \rho_a \rho_b e^{-2jk_z d}} \right] \left[e^{-jk_z z} - \rho_b e^{-2jk_z d} e^{jk_z z} \right] \frac{E_{a+}}{\eta_{aT}}
$$

(8.5.11)

where $\eta_{aT} = \eta_a \cos \theta$ is the TM characteristic impedance of the prism. The power flow within the metal strip is described by the z-component of the Poynting vector:

$$
\mathcal{P}(z) = \frac{1}{2} \text{Re} \left[E_T(z) H_T^*(z) \right]
$$

(8.5.12)

The power entering the conductor at interface a is:

$$
\mathcal{P}_{\text{in}} = (1 - |\Gamma|^2) \frac{|E_{a+}|^2}{2 \eta_{aT}} = \frac{1}{2} \text{Re} \left[E_T(z) H_T^*(z) \right] \bigg|_{z=0}
$$

(8.5.13)

Fig. 8.5.5 shows a plot of the quantity $\mathcal{P}(z)/\mathcal{P}_{\text{in}}$ versus distance within the metal, $0 \leq z \leq d$, at the resonant angle of incidence $\theta = \theta_{\text{res}}$. Because the fields are evanescent in the right medium n_b, the power vanishes at interface b, that is, at $z = d$. The reflectance at the resonance angle is $|\Gamma|^2 = 0.05$, and therefore, the fraction of the incident power that enters the metal layer and is absorbed by it is $1 - |\Gamma|^2 = 0.95$.

power flow versus distance
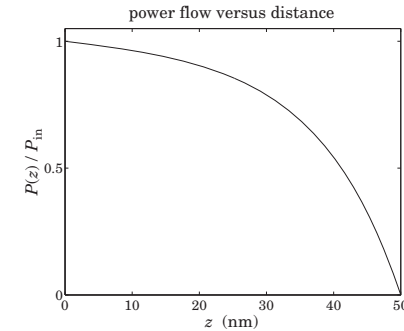
Fig. 8.5.5 Power flow within metal layer at the resonance angle $\theta_{\text{res}} = 43.58^\circ$.

The angle width of the resonance of Fig. 8.5.3, measured at the 3-dB level $|\Gamma|^2 = 1/2$, is very narrow, $\Delta \theta = 0.282^\circ$. The width $\Delta \theta$, as well as the resonance angle θ_{res}, and the optimum metal film thickness d, can be estimated by the following approximate procedure.

To understand the resonance property, we look at the behavior of Γ in the neighborhood of the plasmon wavenumber $k_x = k_{x0}$ given by (8.5.4). At this value, the TM

reflection coefficient at the ε-ε_b interface develops a pole, $\rho_b = \infty$, which is equivalent to the condition $k_{zb0}\varepsilon + k_{z0}\varepsilon_b = 0$, with k_{zb0}, k_{z0} defined by Eq. (8.5.4).

In the neighborhood of this pole, $k_x \simeq k_{x0}$, ρ_b will be given by $\rho_b \simeq K_0/(k_x - k_{x0})$, where K_0 is the residue of the pole. It can be determined by:

$$K_0 = \lim_{k_x \to k_{x0}} (k_x - k_{x0})\rho_b = \lim_{k_x \to k_{x0}} (k_x - k_{x0}) \frac{k_{zb}\varepsilon - k_z\varepsilon_b}{k_{zb}\varepsilon + k_z\varepsilon_b} = \frac{k_{zb}\varepsilon - k_z\varepsilon_b}{\dfrac{d}{dk_x}(k_{zb}\varepsilon + k_z\varepsilon_b)}\Bigg|_{k_x = k_{x0}}$$

The derivative dk_z/dk_x can be determined by differentiating $k_z^2 + k_x^2 = k_0^2\varepsilon$, that is, $k_z\,dk_z + k_x\,dk_x = 0$, which gives $dk_z/dk_x = -k_x/k_z$, and similarly for dk_{zb}/dk_x. It follows that:

$$K_0 = \frac{k_{zb0}\varepsilon - k_{z0}\varepsilon_b}{-\dfrac{k_{x0}}{k_{zb0}}\varepsilon - \dfrac{k_{x0}}{k_{z0}}\varepsilon_b}$$

Inserting k_{x0}, k_{z0}, k_{zb0} from Eq. (8.5.4), we obtain:

$$K_0 = k_0 \left(\frac{2}{\varepsilon_b - \varepsilon}\right)\left(\frac{\varepsilon\varepsilon_b}{\varepsilon + \varepsilon_b}\right)^{3/2} \tag{8.5.14}$$

The reflection response can then be approximated near $k_x \simeq k_{x0}$ by

$$\Gamma \simeq \frac{\rho_a + \dfrac{K_0}{k_x - k_{x0}}e^{-2jk_zd}}{1 + \rho_a\dfrac{K_0}{k_x - k_{x0}}e^{-2jk_zd}}$$

The quantities ρ_a and e^{-2jk_zd} can also be replaced by their values at k_{x0}, k_{z0}, k_{zb0}, thus obtaining:

$$\Gamma = \rho_{a0}\frac{k_x - k_{x0} + \rho_{a0}^{-1}K_0 e^{-2jk_{z0}d}}{k_x - k_{x0} + \rho_{a0}K_0 e^{-2jk_{z0}d}} \tag{8.5.15}$$

where

$$\rho_{a0} = \frac{k_{z0}\varepsilon_a - k_{za0}\varepsilon}{k_{z0}\varepsilon_a + k_{za0}\varepsilon} = \frac{\varepsilon_a + \sqrt{\varepsilon(\varepsilon_a - \varepsilon_b) + \varepsilon_a\varepsilon_b}}{\varepsilon_a - \sqrt{\varepsilon(\varepsilon_a - \varepsilon_b) + \varepsilon_a\varepsilon_b}}$$

which was obtained using $k_{za0} = \sqrt{k_0^2\varepsilon_a - k_{x0}^2}$ and Eqs. (8.5.4). Replacing $\varepsilon = -\varepsilon_r - j\varepsilon_i$, we may also write:

$$\rho_{a0} = \frac{\varepsilon_a + j\sqrt{(\varepsilon_r + j\varepsilon_i)(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}}{\varepsilon_a - j\sqrt{(\varepsilon_r + j\varepsilon_i)(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}} \equiv -b_0 + ja_0 \tag{8.5.16}$$

which serves as the definition of b_0, a_0. We also write:

$$\rho_{a0}^{-1} = \frac{\varepsilon_a - j\sqrt{(\varepsilon_r + j\varepsilon_i)(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}}{\varepsilon_a + j\sqrt{(\varepsilon_r + j\varepsilon_i)(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}} = -\frac{b_0 + ja_0}{b_0^2 + a_0^2} \equiv -b_1 - ja_1 \tag{8.5.17}$$

We define also the wavenumber shifts that appear in the denominator and numerator of (8.5.15) as follows:

$$\bar{k}_{x0} = -\rho_{a0}K_0 e^{-2jk_{z0}d} = (b_0 - ja_0)K_0 e^{-2jk_{z0}d} \equiv \bar{\beta}_{x0} - j\bar{\alpha}_{x0}$$
$$\bar{k}_{x1} = -\rho_{a0}^{-1}K_0 e^{-2jk_{z0}d} = (b_1 + ja_1)K_0 e^{-2jk_{z0}d} \equiv \bar{\beta}_{x1} + j\bar{\alpha}_{x1} \tag{8.5.18}$$

Then, Eq. (8.5.15) becomes, replacing $k_{x0} = \beta_{x0} - j\alpha_{x0}$

$$\Gamma = \rho_{a0}\frac{k_x - k_{x0} - \bar{k}_{x1}}{k_x - k_{x0} - \bar{k}_{x0}} = \rho_{a0}\frac{(k_x - \beta_{x0} - \bar{\beta}_{x1}) + j(\alpha_{x0} - \bar{\alpha}_{x1})}{(k_x - \beta_{x0} - \bar{\beta}_{x0}) + j(\alpha_{x0} + \bar{\alpha}_{x0})} \tag{8.5.19}$$

resulting in the reflectance:

$$|\Gamma|^2 = |\rho_{a0}|^2 \frac{(k_x - \beta_{x0} - \bar{\beta}_{x1})^2 + (\alpha_{x0} - \bar{\alpha}_{x1})^2}{(k_x - \beta_{x0} - \bar{\beta}_{x0})^2 + (\alpha_{x0} + \bar{\alpha}_{x0})^2} \tag{8.5.20}$$

The shifted resonance wavenumber is determined from the denominator of (8.5.19), that is, $k_{x,\text{res}} = \beta_{x0} + \bar{\beta}_{x0}$. The resonance angle is determined by the matching condition:

$$\boxed{k_x = k_0 n_a \sin\theta_{\text{res}} = k_{x,\text{res}} = \beta_{x0} + \bar{\beta}_{x0}} \tag{8.5.21}$$

The minimum value of $|\Gamma|^2$ at resonance is obtained by setting $k_x = \beta_{x0} + \bar{\beta}_{x0}$:

$$|\Gamma|^2_{\text{min}} = |\rho_{a0}|^2 \frac{(\bar{\beta}_{x0} - \bar{\beta}_{x1})^2 + (\alpha_{x0} - \bar{\alpha}_{x1})^2}{(\alpha_{x0} + \bar{\alpha}_{x0})^2} \tag{8.5.22}$$

We will see below that $\bar{\beta}_{x0}$ and $\bar{\beta}_{x1}$ are approximately equal, and so are $\bar{\alpha}_{x0}$ and $\bar{\alpha}_{x1}$. The optimum thickness for the metal layer is obtained by minimizing the numerator of $|\Gamma|^2_{\text{min}}$ by imposing the condition $\alpha_{x0} = \bar{\alpha}_{x1}$. This condition can be solved for d.

The angle width is obtained by solving for the left and right bandedge wavenumbers, say $k_{x,\pm}$, from the 3-dB condition:

$$|\Gamma|^2 = |\rho_{a0}|^2 \frac{(k_x - \beta_{x0} - \bar{\beta}_{x1})^2 + (\alpha_{x0} - \bar{\alpha}_{x1})^2}{(k_x - \beta_{x0} - \bar{\beta}_{x0})^2 + (\alpha_{x0} + \bar{\alpha}_{x0})^2} = \frac{1}{2} \tag{8.5.23}$$

and then obtaining the left/right 3-dB angles by solving $k_0 n_a \sin\theta_\pm = k_{x,\pm}$.

Although Eqs. (8.5.16)–(8.5.23) can be easily implemented numerically, they are unnecessarily complicated. A further simplification can be made by replacing the quantities K_0, ρ_{a0}, and k_{z0} by their lossless values obtained by setting $\varepsilon_i = 0$. This makes ρ_{a0} a unimodular complex number so that $\rho_{a0}^{-1} = \rho_{a0}^*$. We have then the approximations:

$$K_0 = k_0 \left(\frac{2}{\varepsilon_r + \varepsilon_b}\right)\left(\frac{\varepsilon_r\varepsilon_b}{\varepsilon_r - \varepsilon_b}\right)^{3/2}$$

$$\rho_{a0} = \frac{\varepsilon_a + j\sqrt{\varepsilon_r(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}}{\varepsilon_a - j\sqrt{\varepsilon_r(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}} \equiv -b_0 + ja_0, \quad \rho_{a0}^{-1} = -b_0 - ja_0 \tag{8.5.24}$$

$$k_{z0} = -j\alpha_{z0}, \quad \alpha_{z0} = \frac{k_0\varepsilon_r}{\sqrt{\varepsilon_r - \varepsilon_b}}$$

so that

$$b_0 = \frac{\varepsilon_r(\varepsilon_a - \varepsilon_b) - \varepsilon_a(\varepsilon_a + \varepsilon_b)}{(\varepsilon_a - \varepsilon_b)(\varepsilon_r + \varepsilon_b)}, \quad a_0 = \frac{2\varepsilon_a\sqrt{\varepsilon_r(\varepsilon_a - \varepsilon_b) - \varepsilon_a\varepsilon_b}}{(\varepsilon_a - \varepsilon_b)(\varepsilon_r + \varepsilon_b)} \tag{8.5.25}$$

The wavenumber shifts (8.5.18) then become:

$$\bar{k}_{x0} = (b_0 - ja_0)K_0 e^{-2\alpha_{z0}d} = \bar{\beta}_{x0} - j\bar{\alpha}_{x0}$$
$$\bar{k}_{x1} = (b_0 + ja_0)K_0 e^{-2\alpha_{z0}d} = \bar{\beta}_{x0} + j\bar{\alpha}_{x0} = \bar{k}_{x0}^* \tag{8.5.26}$$

with

$$\bar{\beta}_{x0} = b_0 K_0 e^{-2\alpha_{z0}d}, \quad \bar{\alpha}_{x0} = a_0 K_0 e^{-2\alpha_{z0}d} \tag{8.5.27}$$

Then, the reflectance becomes in the neighborhood of the resonance:

$$|\Gamma|^2 = \frac{(k_x - \beta_{x0} - \bar{\beta}_{x0})^2 + (\alpha_{x0} - \bar{\alpha}_{x0})^2}{(k_x - \beta_{x0} - \bar{\beta}_{x0})^2 + (\alpha_{x0} + \bar{\alpha}_{x0})^2} \tag{8.5.28}$$

with a minimum value:

$$|\Gamma|^2_{\min} = \frac{(\alpha_{x0} - \bar{\alpha}_{x0})^2}{(\alpha_{x0} + \bar{\alpha}_{x0})^2} \tag{8.5.29}$$

In this approximation, the resonance angle is determined from:

$$k_0 n_a \sin\theta_{\mathrm{res}} = k_{\mathrm{res}} = \beta_{x0} + \bar{\beta}_{x0} = k_0\sqrt{\frac{\varepsilon_r \varepsilon_b}{\varepsilon_r - \varepsilon_b}} + b_0 K_0 e^{-2\alpha_{z0}d} \tag{8.5.30}$$

Since the second term on the right-hand side represents a small correction, a necessary condition that such a resonance angle would exist is obtained by setting $\theta_{\mathrm{res}} = 90^\circ$ and ignoring the second term:

$$n_a > \sqrt{\frac{\varepsilon_r \varepsilon_b}{\varepsilon_r - \varepsilon_b}} \equiv n_a^{\min} \tag{8.5.31}$$

For example, for the parameters of Fig. 8.5.3, the minimum acceptable refractive index n_a would be $n_a^{\min} = 1.033$. Thus, using a glass prism with $n_a = 1.5$ is more than adequate. If the right medium is water instead of air with $n_b = 1.33$, then $n_a^{\min} = 1.41$, which comes close to the prism choice. The 3-dB angles are obtained by solving

$$|\Gamma|^2 = \frac{(k_x - k_{\mathrm{res}})^2 + (\alpha_{x0} - \bar{\alpha}_{x0})^2}{(k_x - k_{\mathrm{res}})^2 + (\alpha_{x0} + \bar{\alpha}_{x0})^2} = \frac{1}{2}$$

with solution $k_{x,\pm} = k_{\mathrm{res}} \pm \sqrt{6\alpha_{x0}\bar{\alpha}_{x0} - \alpha_{x0}^2 - \bar{\alpha}_{x0}^2}$, or

$$k_0 n_a \sin\theta_\pm = k_0 n_a \sin\theta_{\mathrm{res}} \pm \sqrt{6\alpha_{x0}\bar{\alpha}_{x0} - \alpha_{x0}^2 - \bar{\alpha}_{x0}^2} \tag{8.5.32}$$

The angle width shown on Fig. 8.5.3 was calculated by $\Delta\theta = \theta_+ - \theta_-$ using (8.5.32). The optimum thickness d_{opt} is obtained from the condition $\alpha_{x0} = \bar{\alpha}_{x0}$, which drives $|\Gamma|^2_{\min}$ to zero. This condition requires that $\alpha_{x0} = a_0 K_0 e^{-2\alpha_{z0}d}$, with solution:

$$d_{\mathrm{opt}} = \frac{1}{2\alpha_{z0}} \ln\left(\frac{a_0 K_0}{\alpha_{x0}}\right) = \frac{\lambda_0}{4\pi} \frac{\sqrt{\varepsilon_r - \varepsilon_b}}{\varepsilon_r} \ln\left(\frac{4a_0 \varepsilon_r^2}{\varepsilon_i(\varepsilon_r + \varepsilon_b)}\right) \tag{8.5.33}$$

where we replaced α_{x0} from Eq. (8.5.5). For the same parameters of Fig. 8.5.3, we calculate the optimum thickness to be $d_{\mathrm{opt}} = 56.08$ nm, resulting in the new resonance angle of $\theta_{\mathrm{res}} = 43.55^\circ$, and angle-width $\Delta\theta = 0.227^\circ$. Fig. 8.5.6 shows the reflectance in this case. The above approximations for the angle-width are not perfect, but they are adequate.

One of the current uses of surface plasmon resonance is the detection of the presence of chemical and biological agents. This application makes use of the fact that the
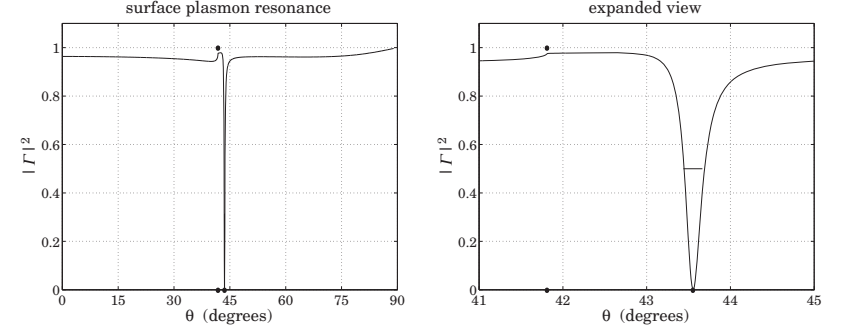
Fig. 8.5.6 Surface plasmon resonance at the optimum thickness $d = d_{\mathrm{opt}}$.
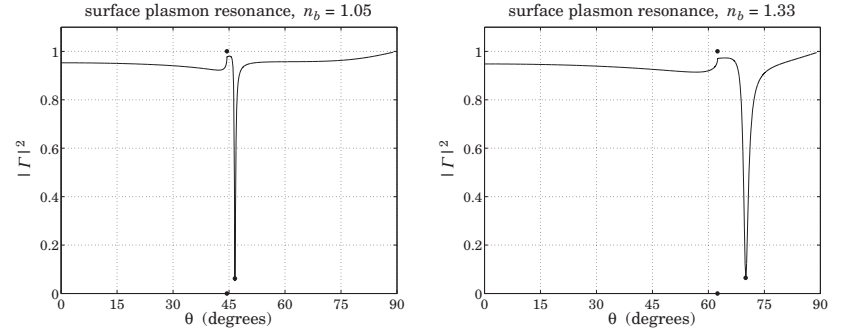

Fig. 8.5.7 Shift of the resonance angle with the refractive index n_b.

resonance angle θ_{res} is very sensitive to the dielectric constant of the medium n_b. For example, Fig. 8.5.7 shows the shift in the resonance angle for the two cases $n_b = 1.05$ and $n_b = 1.33$ (water). Using the same data as Fig. 8.5.3, the corresponding angles and widths were $\theta_{\mathrm{res}} = 46.57^\circ$, $\Delta\theta = 0.349^\circ$ and $\theta_{\mathrm{res}} = 70^\circ$, $\Delta\theta = 1.531^\circ$, respectively.

A number of applications of surface plasmons were mentioned in Sec. 7.11, such as nanophotonics and biosensors. The reader is referred to [593–631] for further reading.

8.6 Perfect Lens in Negative-Index Media

The perfect lens property of negative-index media was originally discussed by Veselago [391], who showed that a slab with $\epsilon = -\epsilon_0$ and $\mu = -\mu_0$, and hence with refractive index $n = -1$, can focus perfectly a point-source of light. More recently, Pendry [398] showed that such a slab can also amplify the evanescent waves from an object, and completely restore the object's spatial frequencies on the other side of the slab. The possibility of overcoming the diffraction limit and improving resolution with such a lens has generated a huge interest in the literature [391–473].

Fig. 8.6.1 shows the perfect lens property. Consider a ray emanating from an object at distance z_0 to the left of the slab ($z = -z_0$). Assuming vacuum on either side of the slab ($n_a = n_b = 1$), Snel's law, implies that the angle of incidence will be equal to the angle of refraction, bending in the same direction of the normal as the incident ray. Indeed, because $n_a = 1$ and $n = -1$, we have:

$$n_a \sin \theta_a = n \sin \theta \quad \Rightarrow \quad \sin \theta_a = -\sin \theta \quad \Rightarrow \quad \theta_a = -\theta$$
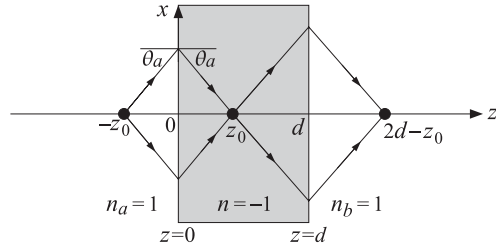

Fig. 8.6.1 Perfect lens property of a negative-index medium with $n = -1$.

Moreover, $\eta = \sqrt{\mu/\epsilon} = \sqrt{\mu_0/\epsilon_0} = \eta_0$ and the slab is matched to the vacuum. Therefore, there will be no reflected ray at the left and the right interfaces. Indeed, the TE and TM reflection coefficients at the left interface vanish at any angle, for example, we have for the TM case, noting that $\cos \theta = \cos(-\theta_a) = \cos \theta_a$:

$$\rho_{TM} = \frac{\eta \cos \theta - \eta_0 \cos \theta_a}{\eta \cos \theta + \eta_0 \cos \theta_a} = \frac{\cos \theta - \cos \theta_a}{\cos \theta + \cos \theta_a} = 0$$

Assuming that $z_0 < d$, where d is the slab thickness, it can be seen from the geometry of Fig. 8.6.1 that the refracted rays will refocus at the point $z = z_0$ within the slab and then continue on to the right interface and refocus again at a distance $d - z_0$ from the slab, that is, at coordinate $z = 2d - z_0$.

Next, we examine the field solutions inside and outside the slab for propagating and for evanescent waves. For the TM case, the electric field will have the following form within the three regions of $z \leq 0$, $0 \leq z \leq d$, and $z \geq d$:

$$E = \begin{cases} \left[E_0 \left(\hat{\mathbf{x}} - \dfrac{k_x}{k_z} \hat{\mathbf{z}} \right) e^{-jk_z z} + E_0 \Gamma \left(\hat{\mathbf{x}} + \dfrac{k_x}{k_z} \hat{\mathbf{z}} \right) e^{jk_z z} \right] e^{-jk_x x}, & \text{for } z \leq 0 \\[2ex] \left[A_+ \left(\hat{\mathbf{x}} - \dfrac{k_x}{k_z'} \hat{\mathbf{z}} \right) e^{-jk_z' z} + A_- \left(\hat{\mathbf{x}} + \dfrac{k_x}{k_z'} \hat{\mathbf{z}} \right) e^{jk_z' z} \right] e^{-jk_x x}, & \text{for } 0 \leq z \leq d \quad (8.6.1) \\[2ex] E_0 T \left(\hat{\mathbf{x}} - \dfrac{k_x}{k_z} \hat{\mathbf{z}} \right) e^{-jk_z(z-d)} e^{-jk_x x}, & \text{for } z \geq d \end{cases}$$

where Γ, T denote the overall transverse reflection and transmission coefficients, and A_+, A_-, the transverse fields on the right-side of the left interface (i.e., at $z = 0+$). The

corresponding magnetic field is:

$$H = \begin{cases} \hat{\mathbf{y}} E_0 \left(\dfrac{\omega \epsilon}{k_z} \right) [e^{-jk_z z} - \Gamma e^{jk_z z}] e^{-jk_x x}, & \text{for } z \leq 0 \\[2ex] \hat{\mathbf{y}} \left(\dfrac{\omega \epsilon'}{k_z'} \right) [A_+ e^{-jk_z' z} - A_- e^{jk_z' z}] e^{-jk_x x}, & \text{for } 0 \leq z \leq d \\[2ex] \hat{\mathbf{y}} E_0 T \left(\dfrac{\omega \epsilon}{k_z} \right) e^{-jk_z(z-d)} e^{-jk_x x}, & \text{for } z \geq d \end{cases} \quad (8.6.2)$$

where k_x is preserved across the interfaces, and k_z, k_z' must satisfy:

$$k_x^2 + k_z^2 = \omega^2 \mu_0 \epsilon_0, \quad k_x^2 + k_z'^2 = \omega^2 \mu \epsilon \quad (8.6.3)$$

Thus, $k_z = \pm \sqrt{\omega^2 \mu_0 \epsilon_0 - k_x^2}$ and $k_z' = \pm \sqrt{\omega^2 \mu \epsilon - k_x^2}$. The choice of square root signs is discussed below. To include evanescent waves, we will define k_z by means of the evanescent square root, setting $k_0 = \omega \sqrt{\mu_0 \epsilon_0}$:

$$k_z = \text{sqrte}(k_0^2 - k_x^2) = \begin{cases} \sqrt{k_0^2 - k_x^2}, & \text{if } k_x^2 \leq k_0^2 \\[1ex] -j\sqrt{k_x^2 - k_0^2}, & \text{if } k_x^2 \geq k_0^2 \end{cases} \quad (8.6.4)$$

We saw in Sec. 7.16 that for a single interface between a positive- and a negative-index medium, and for propagating waves, we must have $k_z > 0$ and $k_z' < 0$ in order for the power transmitted into the negative-index medium to flow away from the interface. But in the case of a slab within which one could have both forward and backward waves, the choice of the sign of k_z' is not immediately obvious. In fact, it turns out that the field solution remains *invariant* under the substitution $k_z' \to -k_z'$, and therefore, one could choose either sign for k_z'. In particular, we could select it to be given also by its evanescent square root, where $n^2 = \epsilon \mu / \epsilon_0 \mu_0$:

$$k_z' = \text{sqrte}(k_0^2 n^2 - k_x^2) = \begin{cases} \sqrt{k_0^2 n^2 - k_x^2}, & \text{if } k_x^2 \leq k_0^2 n^2 \\[1ex] -j\sqrt{k_x^2 - k_0^2 n^2}, & \text{if } k_x^2 \geq k_0^2 n^2 \end{cases} \quad (8.6.5)$$

By matching the boundary conditions at the two interfaces $z = 0$ and $z = d$, the parameters $\Gamma, A_\pm, T$ are obtained from the usual transfer matrices (see Sec. 8.1):

$$\begin{bmatrix} E_0 \\ E_0 \Gamma \end{bmatrix} = \frac{1}{1 + \rho_{TM}} \begin{bmatrix} 1 & \rho_{TM} \\ \rho_{TM} & 1 \end{bmatrix} \begin{bmatrix} A_+ \\ A_- \end{bmatrix}$$

$$\begin{bmatrix} A_+ \\ A_- \end{bmatrix} = \begin{bmatrix} e^{jk_z' d} & 0 \\ 0 & e^{-jk_z' d} \end{bmatrix} \frac{1}{1 - \rho_{TM}} \begin{bmatrix} 1 & -\rho_{TM} \\ -\rho_{TM} & 1 \end{bmatrix} \begin{bmatrix} E_0 T \\ 0 \end{bmatrix} \quad (8.6.6)$$

where,

$$\rho_{TM} = \frac{k_z' \epsilon - k_z \epsilon'}{k_z' \epsilon + k_z \epsilon'} = \frac{\zeta_{TM} - 1}{\zeta_{TM} + 1}, \quad \zeta_{TM} = \frac{\eta_{TM}'}{\eta_{TM}} = \frac{k_z' \epsilon}{k_z \epsilon'} \quad (8.6.7)$$

where ζ_{TM} is a normalized characteristic impedance. The solution of Eqs. (8.6.6) is then,

$$\Gamma = \frac{\rho_{TM}(1 - e^{-2jk_z' d})}{1 - \rho_{TM}^2 e^{-2jk_z' d}} = \frac{(\zeta_{TM}^2 - 1)(1 - e^{-2jk_z' d})}{(\zeta_{TM} + 1)^2 - (\zeta_{TM} - 1)^2 e^{-2jk_z' d}}$$

$$T = \frac{(1 - \rho_{TM}^2) e^{-jk_z' d}}{1 - \rho_{TM}^2 e^{-2jk_z' d}} = \frac{4\zeta_{TM}}{(\zeta_{TM} + 1)^2 e^{jk_z' d} - (\zeta_{TM} - 1)^2 e^{-jk_z' d}} \quad (8.6.8)$$

Similarly, the coefficients $A_\pm$ are found to be:

$$A_+ = \frac{1 - \rho_{TM}\Gamma}{1 - \rho_{TM}} = \frac{1}{2}\left[1 + \zeta_{TM} + (1 - \zeta_{TM})\Gamma\right]E_0$$

$$A_- = \frac{-\rho_{TM} + \Gamma}{1 - \rho_{TM}} = \frac{1}{2}\left[1 - \zeta_{TM} + (1 + \zeta_{TM})\Gamma\right]E_0$$

$$(8.6.9)$$

The TE case is obtained from the TM case by a duality transformation, that is, by the replacements, $\boldsymbol{E} \to \boldsymbol{H}$, $\boldsymbol{H} \to -\boldsymbol{E}$, $\epsilon \to \mu$, $\epsilon' \to \mu'$, and $\rho_{TM} \to \rho_{TE}$, where

$$\rho_{TE} = \frac{k_z\mu' - k_z'\mu}{k_z\mu' + k_z'\mu} = \frac{\zeta_{TE} - 1}{\zeta_{TE} + 1}, \quad \zeta_{TE} = \frac{\eta_{TE}'}{\eta_{TE}} = \frac{k_z\mu'}{k_z'\mu}$$

The invariance under the transformation $k_z' \to -k_z'$ follows from these solutions. For example, noting that $\zeta_{TM} \to -\zeta_{TM}$ under this transformation, we have:

$$\Gamma(-k_z') = \frac{(\zeta_{TM}^2 - 1)(1 - e^{2jk_z'd})}{(-\zeta_{TM} + 1)^2 - (-\zeta_{TM} - 1)^2 e^{2jk_z'd}} = \frac{(\zeta_{TM}^2 - 1)(1 - e^{-2jk_z'd})}{(\zeta_{TM} + 1)^2 - (\zeta_{TM} - 1)^2 e^{-2jk_z'd}} = \Gamma(k_z')$$

Similarly, we find $T(-k_z') = T(k_z')$ and $A_\pm(-k_z') = A_\mp(k_z')$. These imply that the field solutions remain invariant. For example, the electric field inside the slab will be:

$$\boldsymbol{E}(z, -k_z') = \left[A_+(-k_z')\left(\hat{\boldsymbol{x}} - \frac{k_x}{-k_z'}\hat{\boldsymbol{z}}\right)e^{-j(-k_z')z} + A_-(-k_z')\left(\hat{\boldsymbol{x}} + \frac{k_x}{-k_z'}\hat{\boldsymbol{z}}\right)e^{j(-k_z')z}\right]e^{-jk_xx}$$

$$= \left[A_+(k_z')\left(\hat{\boldsymbol{x}} - \frac{k_x}{k_z'}\hat{\boldsymbol{z}}\right)e^{-jk_z'z} + A_-(k_z')\left(\hat{\boldsymbol{x}} + \frac{k_x}{k_z'}\hat{\boldsymbol{z}}\right)e^{jk_z'z}\right]e^{-jk_xx} = \boldsymbol{E}(z, +k_z')$$

Similarly, we have for the magnetic field inside the slab:

$$\boldsymbol{H}(z, -k_z') = \hat{\boldsymbol{y}}\left(\frac{\omega\epsilon'}{-k_z'}\right)\left[A_+(-k_z')e^{-j(-k_z')z} - A_-(-k_z')e^{j(-k_z')z}\right]e^{-jk_xx}$$

$$= \hat{\boldsymbol{y}}\left(\frac{\omega\epsilon'}{k_z'}\right)\left[A_+(k_z')e^{-jk_z'z} - A_-(k_z')e^{jk_z'z}\right]e^{-jk_xx} = \boldsymbol{H}(z, +k_z')$$

Next, we apply these results to the case $\mu = -\mu_0$ and $\epsilon = -\epsilon_0$, having $n = -1$. It follows from Eq. (8.6.5) that $k_z' = \mp k_z$ with k_z given by (8.6.4). In this case, $\zeta_{TM} = k_z'\epsilon/k_z\epsilon' = -k_z'/k_z = \pm 1$. Then, Eq. (8.6.8) implies that $\Gamma = 0$ for either choice of sign. Similarly, we have $T = e^{jk_zd}$, again for either sign of ζ_{TM}:

$$T = e^{jk_zd} = \begin{cases} e^{jk_zd}, & \text{if } k_x^2 \le k_0^2, \quad k_z = \sqrt{k_0^2 - k_x^2} \\ e^{\alpha_zd}, & \text{if } k_x^2 \ge k_0^2, \quad k_z = -j\sqrt{k_x^2 - k_0^2} \equiv -j\alpha_z \end{cases}$$

$$(8.6.10)$$

Thus, the negative-index medium amplifies the transmitted evanescent waves, which was Pendry's observation [398]. The two choices for k_z' lead to the $A_\pm$ coefficients:

$$\begin{array}{llll} k_z' = -k_z & \Rightarrow & \zeta_{TM} = +1 & \Rightarrow & A_+ = E_0, \quad A_- = 0 \\ k_z' = +k_z & \Rightarrow & \zeta_{TM} = -1 & \Rightarrow & A_+ = 0, \quad A_- = E_0 \end{array}$$

$$(8.6.11)$$

For either choice, the field solutions are the same. Indeed, inserting either set of A_+, A_- into Eqs. (8.6.1) and (8.6.2), and using (8.6.10), we find:

$$\boldsymbol{E} = \begin{cases} E_0\left(\hat{\boldsymbol{x}} - \frac{k_x}{k_z}\hat{\boldsymbol{z}}\right)e^{-jk_zz}e^{-jk_xx}, & \text{for } z \le 0 \\ E_0\left(\hat{\boldsymbol{x}} + \frac{k_x}{k_z}\hat{\boldsymbol{z}}\right)e^{jk_zz}e^{-jk_xx}, & \text{for } 0 \le z \le d \\ E_0\left(\hat{\boldsymbol{x}} - \frac{k_x}{k_z}\hat{\boldsymbol{z}}\right)e^{-jk_z(z-2d)}e^{-jk_xx}, & \text{for } z \ge d \end{cases}$$

$$(8.6.12)$$

and the corresponding magnetic field:

$$\boldsymbol{H} = \begin{cases} \hat{\boldsymbol{y}}E_0\left(\frac{\omega\epsilon_0}{k_z}\right)e^{-jk_zz}e^{-jk_xx}, & \text{for } z \le 0 \\ \hat{\boldsymbol{y}}E_0\left(\frac{\omega\epsilon_0}{k_z}\right)e^{jk_zz}e^{-jk_xx}, & \text{for } 0 \le z \le d \\ \hat{\boldsymbol{y}}E_0\left(\frac{\omega\epsilon_0}{k_z}\right)e^{-jk_z(z-2d)}e^{-jk_xx}, & \text{for } z \ge d \end{cases}$$

$$(8.6.13)$$

The solution effectively corresponds to the choice $k_z' = -k_z$ and is valid for both propagating and evanescent waves with k_z given by (8.6.4). In Eq. (8.6.12) the constant E_0 refers to the value of the transverse electric field at $z = 0$. Changing the reference point to $z = -z_0$ at the left of the slab as shown in Fig. 8.6.1, amounts to replacing $E_0 \to E_0e^{-jk_zz_0}$. Then, (8.6.12) reads:

$$\boldsymbol{E} = \begin{cases} E_0\left(\hat{\boldsymbol{x}} - \frac{k_x}{k_z}\hat{\boldsymbol{z}}\right)e^{-jk_z(z+z_0)}e^{-jk_xx}, & \text{for } -z_0 \le z \le 0 \\ E_0\left(\hat{\boldsymbol{x}} + \frac{k_x}{k_z}\hat{\boldsymbol{z}}\right)e^{jk_z(z-z_0)}e^{-jk_xx}, & \text{for } 0 \le z \le d \\ E_0\left(\hat{\boldsymbol{x}} - \frac{k_x}{k_z}\hat{\boldsymbol{z}}\right)e^{-jk_z(z-2d+z_0)}e^{-jk_xx}, & \text{for } z \ge d \end{cases}$$

$$(8.6.14)$$

Setting $k_z = -\alpha_z$ as in (8.6.4), we find the evanescent fields:

$$\boldsymbol{E} = \begin{cases} E_0\left(\hat{\boldsymbol{x}} - \frac{k_x}{-j\alpha_z}\hat{\boldsymbol{z}}\right)e^{-\alpha_z(z+z_0)}e^{-jk_xx}, & \text{for } -z_0 \le z \le 0 \\ E_0\left(\hat{\boldsymbol{x}} + \frac{k_x}{-j\alpha_z}\hat{\boldsymbol{z}}\right)e^{\alpha_z(z-z_0)}e^{-jk_xx}, & \text{for } 0 \le z \le d \\ E_0\left(\hat{\boldsymbol{x}} - \frac{k_x}{-j\alpha_z}\hat{\boldsymbol{z}}\right)e^{-\alpha_z(z-2d+z_0)}e^{-jk_xx}, & \text{for } z \ge d \end{cases}$$

$$(8.6.15)$$

The field is amplified inside the slab. The propagation factors along the z-direction agree at the points $z = -z_0$, $z = z_0$, and $z = 2d - z_0$,

$$e^{-jk_z(z+z_0)}\Big|_{z=-z_0} = e^{jk_z(z-z_0)}\Big|_{z=z_0} = e^{-jk_z(z-2d+z_0)}\Big|_{z=2d-z_0} = 1$$

$$e^{-\alpha_z(z+z_0)}\Big|_{z=-z_0} = e^{\alpha_z(z-z_0)}\Big|_{z=z_0} = e^{-\alpha_z(z-2d+z_0)}\Big|_{z=2d-z_0} = 1$$

$$(8.6.16)$$

which imply the complete restoration of the source at the focal points inside and to the right of the slab:

$$E_x(x,z)\big|_{z=-z_0} = E_x(x,z)\big|_{z=z_0} = E_x(x,z)\big|_{z=2d-z_0} \tag{8.6.17}$$

Fig. 8.6.2 shows a plot of the evanescent component $E_x(z)$ of Eq. (8.6.15) versus distance z inside and outside the slab.
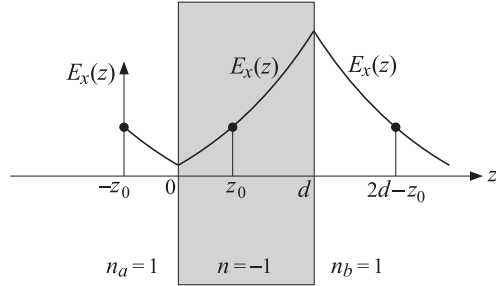

Fig. 8.6.2 Evanescent wave amplification inside a negative-index medium.

Using the plane-wave spectrum representation of Sec. 19.2, a more general (single-frequency) solution can be built by superposition of the plane waves (8.6.14) and (8.6.15). If the field at the image plane $z = -z_0$ has the general representation:

$$E(x,-z_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(k_x) \left(\hat{\mathbf{x}} - \frac{k_x}{k_z}\hat{\mathbf{z}}\right) e^{-jk_x x} \, dk_x \tag{8.6.18}$$

where the integral over k_x includes both propagating and evanescent modes and k_z is given by (8.6.4), then, then field in the three regions to the left of, inside, and to the right of the slab will have the form:

$$E(x,z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(k_x) \left(\hat{\mathbf{x}} - \frac{k_x}{k_z}\hat{\mathbf{z}}\right) e^{-jk_z(z+z_0)} e^{-jk_x x} \, dk_x, \qquad \text{for} \quad -z_0 \le z \le 0$$

$$E(x,z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(k_x) \left(\hat{\mathbf{x}} + \frac{k_x}{k_z}\hat{\mathbf{z}}\right) e^{jk_z(z-z_0)} e^{-jk_x x} \, dk_x, \qquad \text{for} \quad 0 \le z \le d$$

$$E(x,z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(k_x) \left(\hat{\mathbf{x}} - \frac{k_x}{k_z}\hat{\mathbf{z}}\right) e^{-jk_z(z-2d+z_0)} e^{-jk_x x} \, dk_x, \qquad \text{for} \quad z \ge d$$

It is evident that Eq. (8.6.17) is still satisfied, showing the perfect reconstruction of the object field at the two image planes.

The perfect lens property is highly sensitive to the deviations from the ideal values of $\epsilon = -\epsilon_0$ and $\mu = -\mu_0$, and to the presence of losses. Fig. 8.6.3 plots the transmittance in dB, that is, the quantity $10\log_{10}|Te^{-jk_z d}|^2$ versus k_x, with T computed from Eq. (8.6.8) for different values of ϵ, μ and for $d = 0.2\lambda = 0.2(2\pi/k_0)$. In the ideal case, because of the result (8.6.10), we have $|Te^{-jk_z d}| = 1$ for both propagating and evanescent values of k_x, that is, the transmittance is flat (at 0 dB) for all k_x.
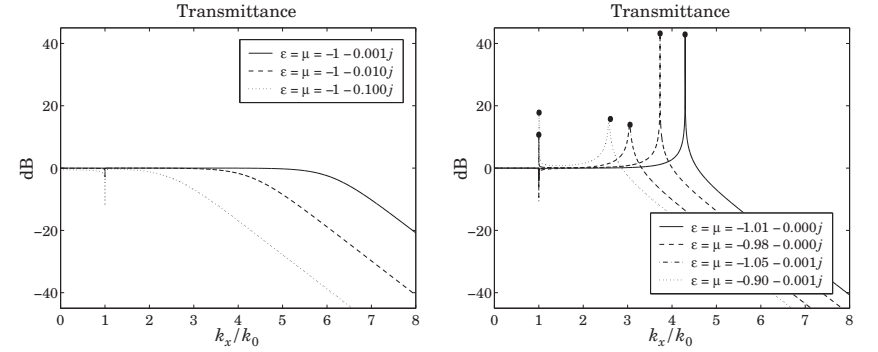
Fig. 8.6.3 Transmittance under non-ideal conditions (ϵ, μ are in units of ϵ_0, μ_0).

The left graph shows the effect of losses while keeping the real parts of ϵ, μ at the ideal values $-\epsilon_0, -\mu_0$. In the presence of losses, the transmittance acts like a lowpass filter in the spatial frequency k_x.

The right graph shows the effect of the deviation of the real parts of ϵ, μ from the ideal values. If the real parts deviate, even slightly, from $-\epsilon_0, -\mu_0$, the transmittance develops resonance peaks, which are related to the excitation of surface plasmons at the two interfaces of the slab [407,408]. The peaks are due to the poles of the denominator of T in Eq. (8.6.8), that is, the roots of

$$1 - \rho_{TM}^2 e^{-2jk_z' d} = 0 \quad \Rightarrow \quad e^{2jk_z' d} = \rho_{TM}^2 \quad \Rightarrow \quad e^{jk_z' d} = \pm\rho_{TM}$$

For evanescent k_x, we may replace $k_z = -j\alpha_z$ and $k_z' = -j\alpha_z'$, where $\alpha_z = \sqrt{k_x^2 - k_0^2}$ and $\alpha_z' = \sqrt{k_x^2 - k_0^2 n^2}$, and obtain the conditions:

$$e^{\alpha_z' d} = \pm\rho_{TM} = \pm\frac{\alpha_z'\epsilon_0 - \alpha_z\epsilon}{\alpha_z'\epsilon_0 + \alpha_z\epsilon} \tag{8.6.19}$$

These are equivalent to [407,408]:

$$\tanh\left(\frac{\alpha_z' d}{2}\right) = -\frac{\alpha_z\epsilon}{\alpha_z'\epsilon_0}, \quad \tanh\left(\frac{\alpha_z' d}{2}\right) = -\frac{\alpha_z'\epsilon_0}{\alpha_z\epsilon} \tag{8.6.20}$$

For $k_x \gg k_0$, we may replace $\alpha_z = \alpha_z' \simeq k_x$ in (8.6.19) in order to get en estimate of the resonant k_x:

$$e^{k_{x,\text{res}} d} = \pm\frac{\epsilon_0 - \epsilon}{\epsilon_0 + \epsilon} \quad \Rightarrow \quad e^{\text{Re}(k_{x,\text{res}})d} = \left|\frac{\epsilon_0 - \epsilon}{\epsilon_0 + \epsilon}\right| \quad \Rightarrow \quad \text{Re}(k_{x,\text{res}}) = \frac{1}{d}\ln\left|\frac{\epsilon_0 - \epsilon}{\epsilon_0 + \epsilon}\right| \tag{8.6.21}$$

and for the TE case, we must replace ϵs by μs. The value $k_x = \text{Re}(k_{x,\text{res}})$ represents the highest achievable resolution by the slab, with the smallest resolvable transverse distance being of the order of $\Delta x = 1/\text{Re}(k_{x,\text{res}})$.

If ϵ is real-valued and near $-\epsilon_0$, then, $k_{x,\text{res}}$ is real and there will be an infinite resonance peak at $k_x = k_{x,\text{res}}$. This is seen in the above figure in the first two cases of

$\epsilon/\epsilon_0 = \mu/\mu_0 = -1.01$ and $\epsilon/\epsilon_0 = \mu/\mu_0 = -0.98$ (the apparent finite height of these two peaks is due to the finite grid of k_x values in the graph.)

The last two cases have complex-valued ϵ, μ with a small imaginary part, with the resulting peaks being finite. In all cases, the peak locations $k_x = \text{Re}(k_{x,\text{res}})$—obtained by solving Eqs. (8.6.20) numerically for $k_{x,\text{res}}$—are indicated in the graphs by bullets placed at the peak maxima. The numerical solutions were obtained by the following iterative procedures, initialized at the approximate (complex-valued) solution of (8.6.21):

$$\text{initialize:}\quad k_x = \frac{1}{d}\ln\left(\frac{\epsilon_0 - \epsilon}{\epsilon_0 + \epsilon}\right)$$

for $i = 1, 2, \ldots, N_{\text{iter}}$, do:

$$\alpha_z' = \sqrt{k_x^2 - k_0^2 n^2}$$

$$\alpha_z = -\frac{\epsilon_0}{\epsilon}\alpha_z'\tanh\left(\frac{\alpha_z' d}{2}\right)$$

$$k_x = \sqrt{\alpha_z^2 + k_0^2}$$

or,

$$k_x = \frac{1}{d}\ln\left(-\frac{\epsilon_0 - \epsilon}{\epsilon_0 + \epsilon}\right), \quad \alpha_z' = \sqrt{k_x^2 - k_0^2 n^2}$$

for $i = 1, 2, \ldots, N_{\text{iter}}$, do:

$$\alpha_z = \sqrt{k_x^2 - k_0^2}$$

$$\alpha_z' = -\frac{\epsilon}{\epsilon_0}\alpha_z\tanh\left(\frac{\alpha_z' d}{2}\right)$$

$$k_x = \sqrt{\alpha_z'^2 + k_0^2 n^2}$$

The number of iterations was typically $N_{\text{iter}} = 30$. Both graphs of Fig. 8.6.3 also show dips at $k_x = k_0$. These are due to the zeros of the transmittance T arising from the numerator factor $(1 - \rho_{TM}^2)$ in (8.6.10). At $k_x = k_0$, we have $\alpha_z = 0$ and $\rho_{TM} = 1$, causing a zero in T. In addition to the zero at $k_x = k_0$, it is possible to also have poles in the vicinity of k_0, as indicated by the peaks and bullets in the graph. Fig. 8.6.4 shows an expanded view of the structure of T near k_0, with the k_x restricted in the narrow interval: $0.99 k_0 \leq k_x \leq 1.01 k_0$.
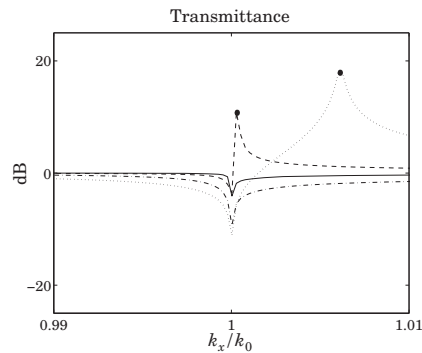

Fig. 8.6.4 Expanded view of the zero/pole behavior in the vicinity of $k_x = k_0$.

For last two cases depicted on this graph that have $|n^2| = |\epsilon\mu|/\epsilon_0\mu_0 \lesssim 1$, an approximate calculation of the pole locations near k_0 is as follows. Since $\alpha_z = \sqrt{k_x^2 - k_0^2}$ is small, and $\alpha_z' = \sqrt{\alpha_z^2 + k_0(1 - n^2)}$, we have to first order in α_z, $\alpha_z' \simeq k_0^2\sqrt{1 - n^2} \equiv \alpha_{z0}'$,

which is itself small. Then, we apply Eq. (8.6.21) to get α_z and from it, the resonant $k_{x,\text{res}}$:

$$\alpha_z = -\frac{\epsilon_0}{\epsilon}\alpha_{z0}'\tanh\left(\frac{\alpha_{z0}' d}{2}\right) \quad \Rightarrow \quad k_{x,\text{res}} = \sqrt{\alpha_z^2 + k_0^2}$$

8.7 Antireflection Coatings at Oblique Incidence

Antireflection coatings are typically designed for normal incidence and then used over a limited range of oblique incidence, such as up to about 30°. As the angle of incidence increases, the antireflection band shifts towards lower wavelengths or higher frequencies. Any designed reflection zeros at normal incidence are no longer zeros at oblique incidence.

If a particular angle of incidence is preferred, it is possible to design the antireflection coating to match that angle. However, like the case of normal design, the effectiveness of this method will be over an angular width of approximately 30° about the preferred angle.

To appreciate the effects of oblique incidence, we look at the angular behavior of our normal-incidence designs presented in Figs. 6.2.1 and 6.2.3.

The first example was a two-layer design with refractive indices $n_a = 1$ (air), $n_1 = 1.38$ (magnesium fluoride), $n_2 = 2.45$ (bismuth oxide), and $n_b = 1.5$ (glass). The designed normalized optical lengths of the layers were $L_1 = 0.3294$ and $L_2 = 0.0453$ at $\lambda_0 = 550$ nm.

Fig. 8.7.1 shows the TE and TM reflectances $|\Gamma_{T1}(\lambda)|^2$ as functions of λ, for the incidence angles $\theta = 0°, 20°, 30°, 40°$.
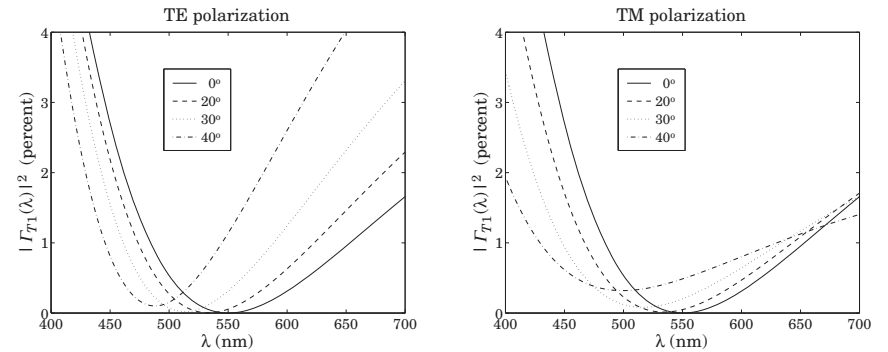

Fig. 8.7.1 Two-layer antireflection coating at oblique incidence.

We note the shifting of the responses towards lower wavelengths. The responses are fairly acceptable up to about 20°–30°. The typical MATLAB code used to generate these graphs was:

```
n = [1, 1.38, 2.45, 1.5]; L = [0.3294, 0.0453];
la0 = 550; la = linspace(400,700,101); pol='te';
```

```
G0  = abs(multidiel(n, L, la/la0)).^2 * 100;
G20 = abs(multidiel(n, L, la/la0, 20, pol)).^2 * 100;
G30 = abs(multidiel(n, L, la/la0, 30, pol)).^2 * 100;
G40 = abs(multidiel(n, L, la/la0, 40, pol)).^2 * 100;

plot(la, [G0; G20; G30; G40]);
```

As we mentioned above, the design can be matched at a particular angle of incidence. As an example, we choose $\theta_a = 30^o$ and redesign the two-layer structure.

The design equations are still (6.2.2) and (6.2.1), but with the replacement of n_i, ρ_i by their transverse values n_{Ti}, ρ_{Ti}, and the replacement of k_1l_1, k_2l_2 by the phase thicknesses at $\lambda = \lambda_0$, that is, $\delta_1 = 2\pi L_1 \cos\theta_1$ and $\delta_2 = 2\pi L_2 \cos\theta_2$. Moreover, we must choose to match the design either for TE or TM polarization.

Fig. 8.7.2 illustrates such a design. The upper left graph shows the TE reflectance matched at 30°. The designed optical thicknesses are in this case, $L_1 = 0.3509$ and $L_2 = 0.0528$. The upper right graph shows the corresponding TM reflectance, which cannot be matched simultaneously with the TE case.

The lower graphs show the same design, but now the TM reflectance is matched at 30°. The designed lengths were $L_1 = 0.3554$ and $L_2 = 0.0386$.
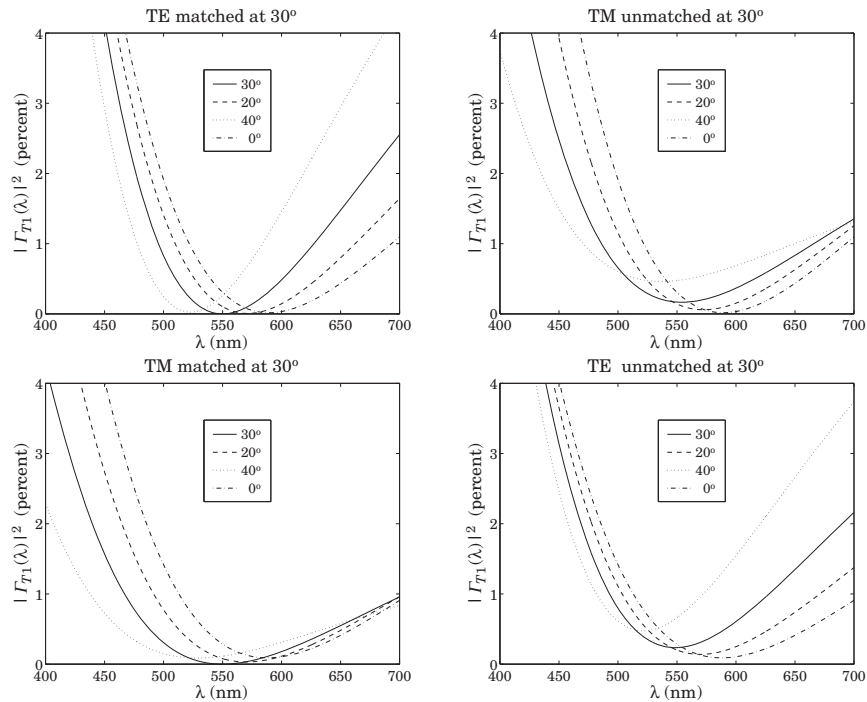

Fig. 8.7.2 Two-layer antireflection coating matched at 30 degrees.

The design steps are as follows. First, we calculate the refraction angles in all media from Eq. (8.1.1), $\theta_i = \text{asin}(n_a \sin\theta_a/n_i)$, for $i = a, 1, 2, b$. Then, assuming TE polarization, we calculate the TE refractive indices for all media $n_{Ti} = n_i \cos\theta_i$, $i = a, 1, 2, b$.

Then, we calculate the transverse reflection coefficients ρ_{Ti} from Eq. (8.1.3) and use them to solve Eq. (6.2.2) and (6.2.1) for the phase thicknesses δ_1, δ_2. Finally, we calculate the normalized optical lengths from $L_i = \delta_i/(2\pi \cos\theta_i)$, $i = 1, 2$. The following MATLAB code illustrates these steps:

```
n = [1, 1.38, 2.45, 1.5]; na = 1;
tha = 30;   thi = asin(na*sin(pi*tha/180)./n);

nt = n.*cos(thi);                    % for TM use nt = n./cos(thi)
r = n2r(nt);

c = sqrt((r(1)^2*(1-r(2)*r(3))^2 - (r(2)-r(3))^2)/(4*r(2)*r(3)*(1-r(1)^2)));
de2 = acos(c);
G2 = (r(2)+r(3)*exp(-2*j*de2))/(1 + r(2)*r(3)*exp(-2*j*de2));
de1 = (angle(G2) - pi - angle(r(1)))/2;
if de1 <0, de1 = de1 + 2*pi; end

L = [de1,de2]/2/pi;
L = L./cos(thi(2:3));

la0 = 550; la = linspace(400,700,401);

G30 = abs(multidiel(n, L, la/la0, 30, 'te')).^2 * 100;
G20 = abs(multidiel(n, L, la/la0, 20, 'te')).^2 * 100;
G40 = abs(multidiel(n, L, la/la0, 40, 'te')).^2 * 100;
G0  = abs(multidiel(n, L, la/la0)).^2 * 100;

plot(la, [G30; G20; G40; G0]);
```

Our second example in Fig. 6.2.3 was a quarter-half-quarter 3-layer design with refractive indices $n_1 = 1$ (air), $n_1 = 1.38$ (magnesium fluoride), $n_2 = 2.2$ (zirconium oxide), $n_3 = 1.63$ (cerium fluoride), and $n_b = 1.5$ (glass). The optical lengths of the layers were $L_1 = L_3 = 0.25$ and $L_2 = 0.5$.

Fig. 8.7.3 shows the TE and TM reflectances $|\Gamma_{T1}(\lambda)|^2$ as functions of λ, for the incidence angles $\theta = 0^o, 20^o, 30^o, 40^o$.

The responses are fairly acceptable up to about 20°–30°, but are shifted towards lower wavelengths. The typical MATLAB code used to generate these graphs was:

```
n = [1, 1.38, 2.2, 1.63, 1.5]; L = [0.25, 0.50, 0.25];

la0 = 550; la = linspace(400,700,401);

G0  = abs(multidiel(n, L, la/la0)).^2 * 100;
G20 = abs(multidiel(n, L, la/la0, 20, 'te')).^2 * 100;
G30 = abs(multidiel(n, L, la/la0, 30, 'te')).^2 * 100;
G40 = abs(multidiel(n, L, la/la0, 40, 'te')).^2 * 100;

plot(la, [G0; G20; G30; G40]);
```
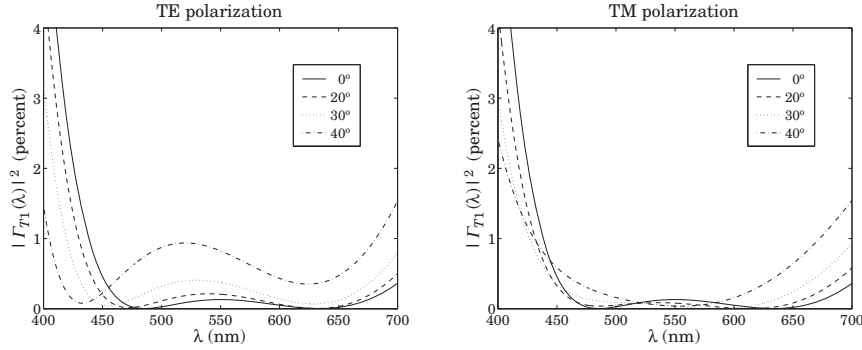
Fig. 8.7.3   Three-layer antireflection coating at oblique incidence.

8.8 Omnidirectional Dielectric Mirrors

Until recently, it was generally thought that it was impossible to have an omnidirectional dielectric mirror, that is, a mirror that is perfectly reflecting at all angles of incidence and for both TE and TM polarizations. However, such mirrors are possible and have recently been manufactured [777,778] and the conditions for their existence clarified [777–781].

We consider the same dielectric mirror structure of Sec. 6.3, consisting of alternating layers of high and low index. Fig. 8.8.1 shows such a structure under oblique incidence. There are N bilayers and a total of $M = 2N + 1$ single layers, starting and ending with a high-index layer.
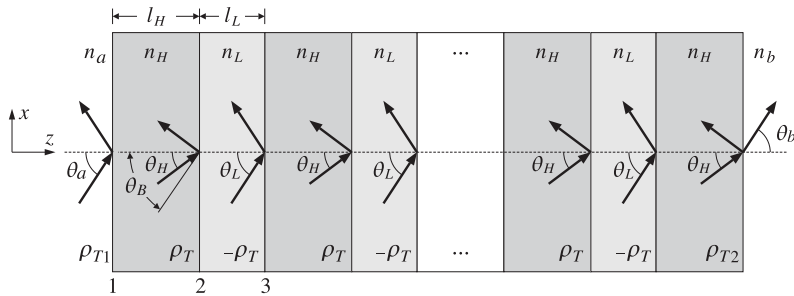

Fig. 8.8.1   Dielectric mirror at oblique incidence.

The incidence angles on each interface are related by Snel's law:

$$n_a \sin \theta_a = n_H \sin \theta_H = n_L \sin \theta_L = n_b \sin \theta_b \qquad (8.8.1)$$

The phase thicknesses within the high- and low-index layers are in normalized form:

$$\delta_H = 2\pi \frac{f}{f_0} L_H \cos \theta_H , \quad \delta_L = 2\pi \frac{f}{f_0} L_L \cos \theta_L \qquad (8.8.2)$$

where $L_H = n_H l_H / \lambda_0$, $L_L = n_L l_L / \lambda_0$ are the optical thicknesses normalized to some λ_0, and $f_0 = c_0 / \lambda_0$. Note also, $\cos \theta_i = \sqrt{1 - n_a^2 \sin^2 \theta_a / n_i^2}$, $i = H, L$.

A necessary (but not sufficient) condition for omnidirectional reflectivity for both polarizations is that the maximum angle of refraction $\theta_{H,\max}$ inside the first layer be less than the Brewster angle θ_B of the second interface, that is, the high-low interface, so that the Brewster angle can never be accessed by a wave incident on the first interface. If this condition is not satisfied, a TM wave would not be reflected at the second and all subsequent interfaces and will transmit through the structure.

Because $\sin \theta_{H,\max} = n_a / n_H$ and $\tan \theta_B = n_L / n_H$, or, $\sin \theta_B = n_L / \sqrt{n_H^2 + n_L^2}$, the condition $\theta_{H,\max} < \theta_B$, or the equivalent condition $\sin \theta_{H,\max} < \sin \theta_B$, can be written as $n_a / n_H < n_L / \sqrt{n_H^2 + n_L^2}$, or

$$n_a < \frac{n_H n_L}{\sqrt{n_H^2 + n_L^2}} \qquad (8.8.3)$$

We note that the exact opposite of this condition is required in the design of multi-layer Brewster polarizing beam splitters, discussed in the next section.

In addition to condition (8.8.3), in order to achieve omnidirectional reflectivity we must require that the high-reflectance bands have a *common* overlapping region for all incidence angles and for both polarizations.

To determine these bands, we note that the entire discussion of Sec. 6.3 carries through unchanged, provided we use the transverse reflection coefficients and transverse refractive indices. For example, the transverse version of the bilayer transition matrix of Eq. (6.3.5) will be:

$$F_T = \frac{1}{1 - \rho_T^2} \left[\begin{array}{cc} e^{j(\delta_H + \delta_L)} - \rho_T^2 e^{j(\delta_H - \delta_L)} & -2j\rho_T e^{-j\delta_H} \sin \delta_L \\ 2j\rho_T e^{j\delta_H} \sin \delta_L & e^{-j(\delta_H + \delta_L)} - \rho_T^2 e^{-j(\delta_H - \delta_L)} \end{array} \right] \qquad (8.8.4)$$

where $\rho_T = (n_{HT} - n_{LT}) / (n_{HT} + n_{LT})$ and:

$$n_{HT} = \left\{ \begin{array}{l} \dfrac{n_H}{\cos \theta_H} \\ n_H \cos \theta_H \end{array} \right. \qquad n_{LT} = \left\{ \begin{array}{ll} \dfrac{n_L}{\cos \theta_L} & \text{(TM polarization)} \\ n_L \cos \theta_L & \text{(TE polarization)} \end{array} \right. \qquad (8.8.5)$$

Explicitly, we have for the two polarizations:

$$\rho_{TM} = \frac{n_H \cos \theta_L - n_L \cos \theta_H}{n_H \cos \theta_L + n_L \cos \theta_H}, \quad \rho_{TE} = \frac{n_H \cos \theta_H - n_L \cos \theta_L}{n_H \cos \theta_H + n_L \cos \theta_L} \qquad (8.8.6)$$

The trace of F_T is as in Eq. (6.3.13):

$$a = \frac{\cos(\delta_H + \delta_L) - \rho_T^2 \cos(\delta_H - \delta_L)}{1 - \rho_T^2} \qquad (8.8.7)$$

The eigenvalues of the matrix F_T are $\lambda_\pm = e^{\pm jKl}$, where $K = \text{acos}(a)/l$ and $l = l_H + l_L$. The condition $a = -1$ determines the bandedge frequencies of the high-reflectance bands. As in Eq. (6.3.16), this condition is equivalent to:

$$\cos^2\left(\frac{\delta_H + \delta_L}{2}\right) = \rho_T^2 \cos^2\left(\frac{\delta_H - \delta_L}{2}\right) \tag{8.8.8}$$

Defining the quantities $L_\pm = L_H \cos\theta_H \pm L_L \cos\theta_L$ and the normalized frequency $F = f/f_0$, we may write:

$$\frac{\delta_H \pm \delta_L}{2} = \pi\frac{f}{f_0}(L_H\cos\theta_H \pm L_L\cos\theta_L) = \pi F L_\pm \tag{8.8.9}$$

Then, taking square roots of Eq. (8.8.8), we have:

$$\cos(\pi F L_+) = \pm|\rho_T|\cos(\pi F L_-)$$

The plus sign gives the left bandedge, $F_1 = f_1/f_0$, and the minus sign, the right bandedge, $F_2 = f_2/f_0$. Thus, F_1, F_2 are the solutions of the equations:

$$\begin{aligned}\cos(\pi F_1 L_+) &= |\rho_T|\cos(\pi F_1 L_-)\\ \cos(\pi F_2 L_+) &= -|\rho_T|\cos(\pi F_2 L_-)\end{aligned} \tag{8.8.10}$$

The bandwidth and center frequency of the reflecting band are:

$$\frac{\Delta f}{f_0} = \Delta F = F_2 - F_1, \qquad \frac{f_c}{f_0} = F_c = \frac{F_1 + F_2}{2} \tag{8.8.11}$$

The corresponding bandwidth in wavelengths is defined in terms of the left and right bandedge wavelengths:

$$\lambda_1 = \frac{\lambda_0}{F_2} = \frac{c_0}{f_2}, \qquad \lambda_2 = \frac{\lambda_0}{F_1} = \frac{c_0}{f_1}, \qquad \Delta\lambda = \lambda_2 - \lambda_1 \tag{8.8.12}$$

An approximate solution of Eq. (8.8.10) can be obtained by setting $L_- = 0$ in the right-hand sides of Eq. (8.8.10):

$$\cos(\pi F_1 L_+) = |\rho_T|, \quad \cos(\pi F_2 L_+) = -|\rho_T| \tag{8.8.13}$$

with solutions:

$$F_1 = \frac{\text{acos}(|\rho_T|)}{\pi L_+}, \quad F_2 = \frac{\text{acos}(-|\rho_T|)}{\pi L_+} \tag{8.8.14}$$

Using the trigonometric identities $\text{acos}(\pm|\rho_T|) = \pi/2 \mp \text{asin}(|\rho_T|)$, we obtain the bandwidth and center frequency:

$$\Delta f = f_2 - f_1 = \frac{2f_0\,\text{asin}(|\rho_T|)}{\pi L_+}, \qquad f_c = \frac{f_1 + f_2}{2} = \frac{f_0}{2L_+} \tag{8.8.15}$$

It follows that the center wavelength will be $\lambda_c = c_0/f_c = 2L_+\lambda_0$ or,

$$\lambda_c = 2L_+\lambda_0 = 2(l_H n_H \cos\theta_H + l_L n_L \cos\theta_L) \tag{8.8.16}$$

At normal incidence, we have $\lambda_c = 2(l_H n_H + l_L n_L)$. For quarter-wavelength designs at λ_0 at normal incidence, we have $L_+ = 1/4 + 1/4 = 1/2$, so that $\lambda_c = \lambda_0$.

The accuracy of the approximate solution (8.8.14) depends on the ratio $d = L_-/L_+$. Even if at normal incidence the layers were quarter-wavelength with $L_H = L_L = 0.25$, the equality of L_H and L_L will no longer be true at other angles of incidence. In fact, the quantity d is an increasing function of θ_a. For larger values of d, the exact solution of (8.8.10) can be obtained by the following iteration:

$$\begin{aligned}&\text{initialize with } F_1 = F_2 = 0,\\ &\text{for } i = 0, 1, \ldots, N_{\text{iter}}, \text{ do:}\\ &\quad F_1 = \frac{1}{\pi L_+}\text{acos}(|\rho_T|\cos(\pi F_1 L_-))\\ &\quad F_2 = \frac{1}{\pi L_+}\text{acos}(-|\rho_T|\cos(\pi F_2 L_-))\end{aligned} \tag{8.8.17}$$

Evidently, the $i = 0$ iteration gives the zeroth-order solution (8.8.14). The iteration converges extremely fast, requiring only 3–4 iterations N_{iter}. The MATLAB function omniband implements this algorithm. It has usage:

```
[F1,F2] = omniband(na,nH,nL,LH,LL,theta,pol,Niter)   % bandedge frequencies
[F1,F2] = omniband(na,nH,nL,LH,LL,theta,pol)         % equivalent to Niter = 0
```

where theta is the incidence angle in degrees, pol is one of the strings 'te' or 'tm' for TE or TM polarization, and Niter is the desired number of iterations. If this argument is omitted, only the $i = 0$ iteration is carried out.

It is straightforward but tedious to verify the following facts about the above solutions. First, f_1, f_2 are *increasing* functions of θ_a for both TE and TM polarizations. Thus, the center frequency of the band $f_c = (f_1 + f_2)/2$ shifts towards higher frequencies with increasing angle θ_a. The corresponding wavelength intervals will shift towards lower wavelengths.

Second, the bandwidth $\Delta f = f_2 - f_1$ is an increasing function of θ_a for TE, and a decreasing one for TM polarization. Thus, as θ_a increases, the reflecting band for TE *expands* and that of TM *shrinks*, while their (slightly different) centers f_c shift upwards.

In order to achieve omnidirectional reflectivity, the TE and TM bands must have a common overlapping intersection for all angles of incidence. Because the TM band is always narrower than the TE band, it will determine the final common omnidirectional band.

The worst case of overlap is for the TM band at 90º angle of incidence, which must overlap with the TM/TE band at 0º. The left bandedge of this TM band, $f_{1,TM}(90°)$, must be *less* than the right bandedge of the 0º band, $f_2(0°)$. This is a sufficient condition for omnidirectional reflectivity.

Thus, the minimum band shared by all angles of incidence and both polarizations will be $[f_{1,TM}(90°), f_2(0°)]$, having width:

$$\boxed{\Delta f_{\min} = f_2(0^o) - f_{1,TM}(90^o)} \qquad \text{(minimum omnidirectional bandwidth)} \qquad (8.8.18)$$

In a more restricted sense, the common reflecting band for both polarizations and for angles up to a given θ_a will be $[f_{1,TM}(\theta_a), f_{2,TM}(0^o)]$ and the corresponding bandwidth:

$$\Delta f(\theta_a) = f_2(0^o) - f_{1,TM}(\theta_a) \qquad (8.8.19)$$

In addition to computing the bandwidths of either the TM or the TE bands at any angle of incidence, the function omniband can also compute the above common bandwidths. If the parameter pol is equal to 'tem', then F_1, F_2 are those of Eqs. (8.8.18) and (8.8.19). Its extended usage is as follows:

```
[F1,F2] = omniband(na,nH,nL,LH,LL,theta,'tem')      % Eq. (8.8.19)
[F1,F2] = omniband(na,nH,nL,LH,LL,90,'tem')         % Eq. (8.8.18)
[F1,F2] = omniband(na,nH,nL,LH,LL)                  % Eq. (8.8.18)
```

Next, we discuss some simulation examples that will help clarify the above remarks.

Example 8.8.1: The first example is the angular dependence of Example 6.3.2. In order to flatten out and sharpen the edges of the reflecting bands, we use $N = 30$ bilayers. Fig. 8.8.2 shows the TE and TM reflectances $|\Gamma_{T1}(\lambda)|^2$ as functions of the free-space wavelength λ, for the two angles of incidence $\theta_a = 45^o$ and 80°.

Fig. 8.8.3 depicts the reflectances as functions of frequency f. The refractive indices were $n_a = 1$, $n_H = 2.32$, $n_L = 1.38$, $n_b = 1.52$, and the bilayers were quarter-wavelength $L_H = L_L = 0.25$ at the normalization wavelength $\lambda_0 = 500$ nm.

The necessary condition (8.8.3) is satisfied and we find for the maximum angle of refraction and the Brewster angle: $\theta_{H,\max} = 25.53^o$ and $\theta_B = 30.75^o$ Thus, we have $\theta_{H,\max} < \theta_B$.
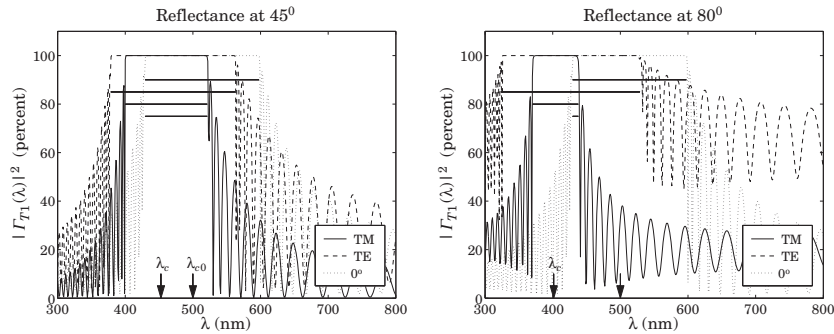

Fig. 8.8.2 TM and TE reflectances for $n_H = 2.32$, $n_L = 1.38$.

On each graph, we have indicated the corresponding bandwidth intervals calculated with omniband. The indicated intervals are for 0° incidence, for TE and TM, and for the common band Eq. (8.8.19) at θ_a. We observe the shifting of the bands towards higher frequencies,
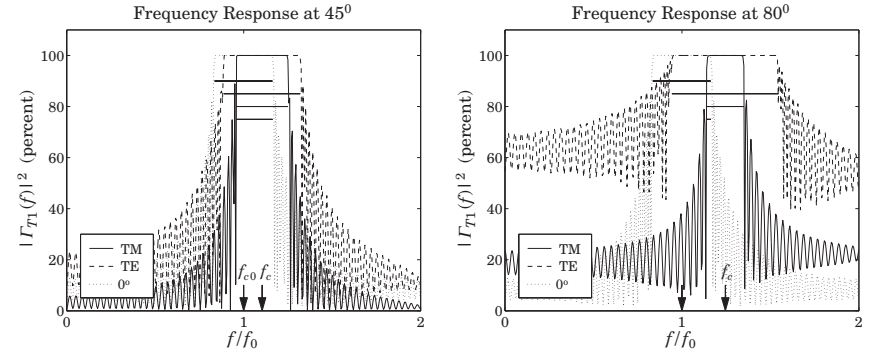
Fig. 8.8.3 TM and TE frequency responses for $n_H = 2.32$, $n_L = 1.38$.

or lower wavelengths, and the shrinking of the TM and expanding of the TE bands, and the shrinking of the common band.

At 45°, there is still sufficient overlap, but at 80°, the TM band has shifted almost to the end of the 0° band, resulting in an extremely narrow common band.

The arrows labeled f_{c0} and f_c represent the (TM) band center frequencies at 0° and 45° or 80°. The calculated bandedges corresponding to 90° incidence were $\lambda_1 = \lambda_0/F_{2,TM}(0^o) = 429.73$ nm and $\lambda_2 = \lambda_0/F_{1,TM}(90^o) = 432.16$ nm, with bandwidth $\Delta\lambda = \lambda_2 - \lambda_1 = 2.43$ nm. Thus, this structure does exhibit omnidirectional reflectivity, albeit over a very narrow band. The MATLAB code used to generate these graphs was:

```
na = 1; nb = 1.52; nH = 2.32; nL = 1.38;
LH = 0.25; LL = 0.25;

la0 = 500;
la = linspace(300,800,501);

th = 45; N = 30;
n = [na, nH, repmat([nL,nH], 1, N), nb];
L = [LH, repmat([LL,LH], 1, N)];
Ge = 100*abs(multidiel(n,L,la/la0, th, 'te')).^2;
Gm = 100*abs(multidiel(n,L,la/la0, th, 'tm')).^2;
G0 = 100*abs(multidiel(n,L,la/la0)).^2;

plot(la,Gm, la,Ge, la,G0);

[F10,F20] = omniband(na,nH,nL,LH,LL, 0, 'te');
[F1e,F2e] = omniband(na,nH,nL,LH,LL, th,'te');
[F1m,F2m] = omniband(na,nH,nL,LH,LL, th,'tm');
[F1,F2] =  omniband(na,nH,nL,LH,LL, th,'tem');
```

Because the reflectivity bands shrink with decreasing ratio n_H/n_L, if we were to slightly decrease n_H, then the TM band could be made to shift beyond the end of the 0° band and there would be no common overlapping reflecting band for all angles. We can observe this behavior in Fig. 8.8.4, which has $n_H = 2$, with all the other parameters kept the same.
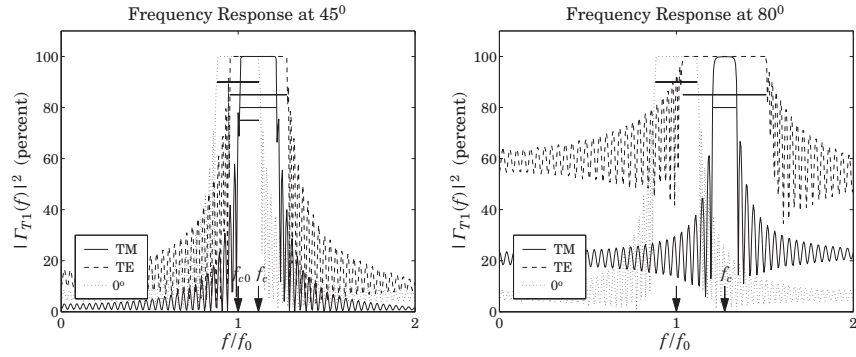
Frequency Response at 45^0

Frequency Response at 80^0

Fig. 8.8.4 TM and TE reflectances for $n_H = 2$, $n_L = 1.38$.

At 45° there is a common overlap, but at 80°, the TM band has already moved beyond the 0° band, while the TE band still overlaps with the latter. This example has no omnidirectional reflectivity, although the necessary condition (8.8.3) is still satisfied with $\theta_{H,\max} = 30°$ and $\theta_B = 34.61°$.

On the other hand, if we were to increase n_H, all the bands will widen, and so will the final common band, resulting in an omnidirectional mirror of wider bandwidth. Fig. 8.8.5 shows the case of $n_H = 3$, exhibiting a substantial overlap and omnidirectional behavior.
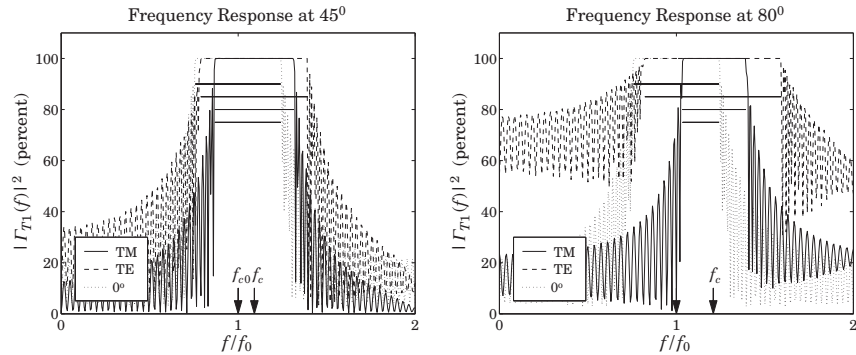
Frequency Response at 45^0

Frequency Response at 80^0

Fig. 8.8.5 TM and TE reflectances for $n_H = 3$, $n_L = 1.38$.

The minimum band (8.8.18) was $[F_1, F_2] = [1.0465, 1.2412]$ corresponding to the wavelength bandedges $\lambda_1 = \lambda_0/F_2 = 402.84$ nm and $\lambda_2 = \lambda_0/F_1 = 477.79$ nm with a width of $\Delta\lambda = \lambda_2 - \lambda_1 = 74.95$ nm, a substantial difference from that of Fig. 8.8.2. The bandedges were computed with $N_{\text{it}} = 0$ in Eq. (8.8.17); with $N_{\text{it}} = 3$, we obtain the more accurate values: $[F_1, F_2] = [1.0505, 1.2412]$.

To illustrate the dependence of the TE and TM bandwidths on the incident angle θ_a, we have calculated and plotted the normalized bandedge frequencies $F_1(\theta_a)$, $F_2(\theta_a)$ for the

range of angles $0 \le \theta_a \le 90°$ for both polarizations. The left graph of Fig. 8.8.6 shows the case $n_H = 3$, $n_L = 1.38$, and the right graph, the case $n_H = 2$, $n_L = 1.38$.
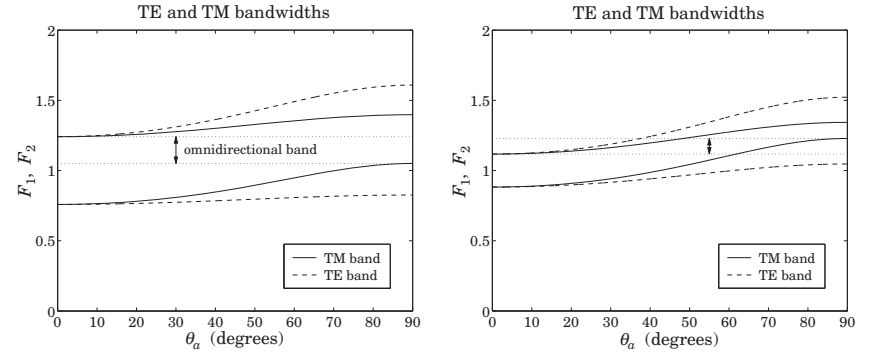
TE and TM bandwidths

TE and TM bandwidths

Fig. 8.8.6 TM/TE bandgaps versus angle for $n_H = 3$, $n_L = 1.38$ and $n_H = 2$, $n_L = 1.38$.

We note that the TE band widens with increasing angle, whereas the TM band narrows. At the same time, the band centers move toward higher frequencies. In the left graph, there is a common band shared by both polarizations and all angles, that is, the band defined by $F_2(0°)$, and $F_{1,\text{TM}}(90°)$. For the right graph, the bandedge $F_{1,\text{TM}}(\theta_a)$ increases beyond $F_2(0°)$ for angles θ_a greater than about $61.8°$, and therefore, there is no omnidirectional band. The calculations of $F_1(\theta_a)$, $F_2(\theta_a)$ were done with `omniband` with $N_{\text{iter}} = 3$. □

Example 8.8.2: In Fig. 8.8.7, we study the effect of changing the optical lengths of the bilayers from quarter-wavelength to $L_H = 0.3$ and $L_L = 0.1$. The main result is to narrow the bands. This example, also illustrates the use of the iteration (8.8.17). The approximate solution (8.8.15) and exact solutions for the 80° bandedge frequencies are obtained from the two MATLAB calls:

```
[F1,F2] = omniband(na,nH,nL,LH,LL,80,'tem',0);
[F1,F2] = omniband(na,nH,nL,LH,LL,80,'tem',3);
```

with results $[F_1, F_2] = [1.0933, 1.3891]$ and $[F_1, F_2] = [1.1315, 1.3266]$, respectively. Three iterations produce an excellent approximation to the exact solution. □

Example 8.8.3: Here, we revisit Example 6.3.3, whose parameters correspond to the recently constructed omnidirectional infrared mirror [777]. Fig. 8.8.8 shows the reflectances as functions of wavelength and frequency at $\theta_a = 45°$ and 80° for both TE and TM polarizations. At both angles of incidence there is a wide overlap, essentially over the desired 10–15 μm band.

The structure consisted of nine alternating layers of Tellurium ($n_H = 4.6$) and Polystyrene ($n_L = 1.6$) on a NaCl substrate ($n_b = 1.48$.) The physical lengths were $l_H = 0.8$ and $l_L = 1.6$ μm. The normalizing wavelength was $\lambda_0 = 12.5$ μm. The optical thicknesses in units of λ_0 were $L_H = 0.2944$ and $L_L = 0.2112$.

The bandedges at 0° were $[F_1, F_2] = [0.6764, 1.2875]$ with center frequency $F_{c0} = 0.9819$, corresponding to wavelength $\lambda_{c0} = \lambda_0/F_{c0} = 12.73$ μm. Similarly, at 45°, the band centers
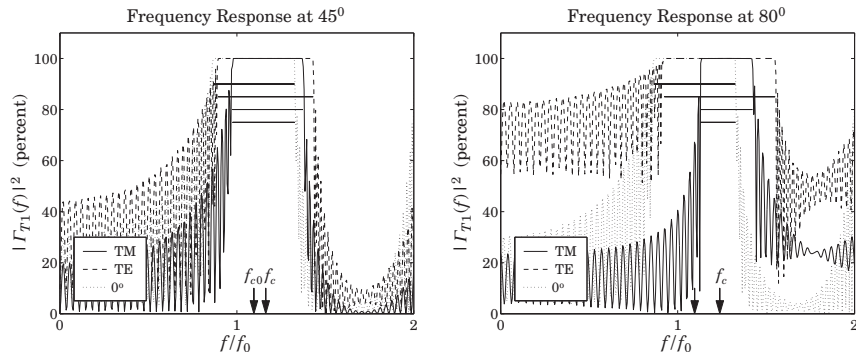
Fig. 8.8.7 Unequal length layers $L_H = 0.30$, $L_L = 0.15$.

for TE and TM polarizations were $F_{c,TE} = 1.0272$ and $F_{c,TM} = 1.0313$, resulting in the wavelengths $\lambda_{c,TE} = 12.17$ and $\lambda_{c,TM} = 12.12$ μm (shown on the graphs are the TE centers
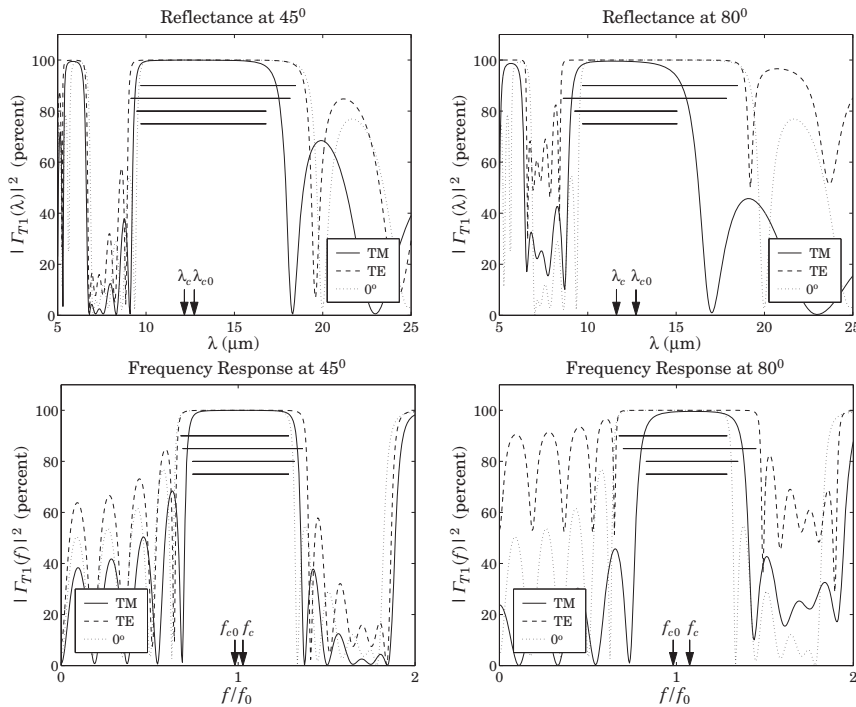

Fig. 8.8.8 Nine-layer Te/PS omnidirectional mirror over the infrared.

only.)

The final bandedges of the common reflecting band computed from Eq. (8.8.18) were $[F_1, F_2] = [0.8207, 1.2875]$, resulting in the wavelength bandedges $\lambda_1 = \lambda_0/F_2 = 9.71$ and $\lambda_2 = \lambda_0/F_1 = 14.95$ μm, with a width of $\Delta\lambda = \lambda_2 - \lambda_1 = 5.24$ μm and band center $(\lambda_1 + \lambda_2)/2 = 12.33$ μm (the approximation (8.8.15) gives 5.67 and 12.4 μm.) The graphs were generated by the following MATLAB code:

```
la0 = 12.5; la = linspace(5,25,401);
na = 1; nb = 1.48; nH = 4.6; nL = 1.6;
lH = 0.8; lL = 1.65;  LH = nH*lH/la0; LL = nL*lL/la0;

th = 45;
N = 4;
n = [na, nH, repmat([nL,nH], 1, N), nb];
L = [LH, repmat([LL,LH], 1, N)];
Ge = 100*abs(multidiel(n,L,la/la0, th, 'te')).^2;
Gm = 100*abs(multidiel(n,L,la/la0, th, 'tm')).^2;
G0 = 100*abs(multidiel(n,L,la/la0)).^2;

plot(la,Gm, la,Ge, la,G0);

Ni = 5;
[F10,F20] = omniband(na,nH,nL,LH,LL, 0, 'te', Ni);    % band at 0°
[F1e,F2e] = omniband(na,nH,nL,LH,LL, th,'te', Ni);    % TE band
[F1m,F2m] = omniband(na,nH,nL,LH,LL, th,'tm', Ni);    % TM band
[F1,F2]   = omniband(na,nH,nL,LH,LL, th,'tem',Ni);    % Eq. (8.8.19)
[F1,F2]   = omniband(na,nH,nL,LH,LL, 90,'tem',Ni);    % Eq. (8.8.18)
```

Finally, Fig. 8.8.9 shows the same example with the number of bilayers doubled to $N = 8$. The mirror bands are flatter and sharper, but the widths are the same. ☐
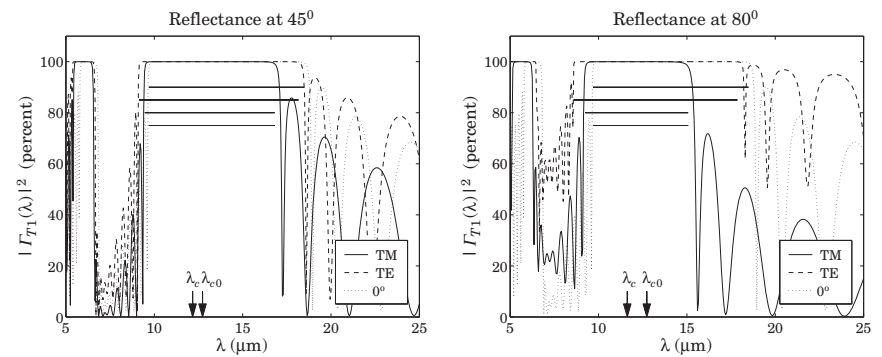

Fig. 8.8.9 Omnidirectional mirror with $N = 8$.

Example 8.8.4: The last example has parameters corresponding to the recently constructed omnidirectional reflector over the visible range [778]. The refractive indices were $n_a = 1$, $n_H = 2.6$ (ZnSe), $n_L = 1.34$ (Na$_3$AlF$_6$ cryolite), and $n_b = 1.5$ (glass substrate.) The layer

lengths were $l_H = l_L = 90$ nm. There were $N = 9$ bilayers or $2N + 1 = 19$ layers, starting and ending with n_H.

With these values, the maximum angle of refraction is $\theta_{H,\max} = 22.27^o$ and is less than the Brewster angle $\theta_B = 27.27^o$.

The normalizing wavelength was taken to be $\lambda_0 = 620$ nm. Then, the corresponding optical lengths were $L_L = n_L l_L/\lambda_0 = 0.1945$ and $L_H = n_H l_H/\lambda_0 = 0.3774$. The overall minimum omnidirectional band is $[\lambda_1, \lambda_2] = [605.42, 646.88]$ nm. It was computed by the MATLAB call to omniband with $N_i = 5$ iterations:

```
[F1,F2] = omniband(na,nH,nL,LH,LL,90,'tem',Ni);
la1 = la0/F2; la2 = la0/F1;
```

(The values of λ_1, λ_2 do not depend on the choice of λ_0.) Fig. 8.8.10 shows the reflectance at 45° and 80°. The upper panel of graphs has $N = 9$ bilayers as in [778]. The lower panel has $N = 18$ bilayers or 38 layers, and has more well-defined band gaps. The two arrows in the figures correspond to the values of λ_1, λ_2 of the minimum omnidirectional band. □
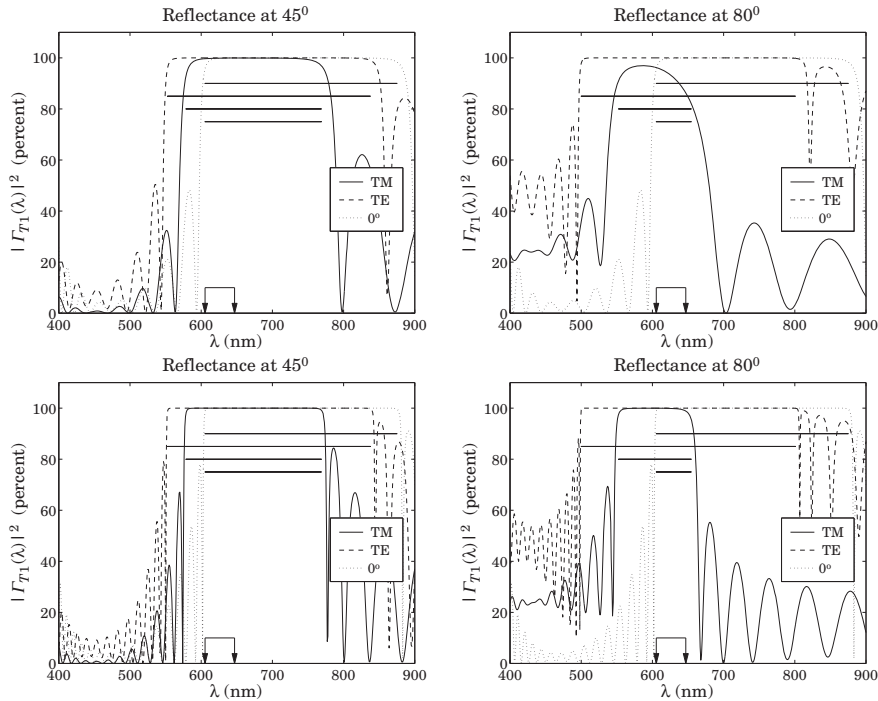

Fig. 8.8.10 Omnidirectional mirror over visible band.

8.9 Polarizing Beam Splitters

The objective of an omnidirectional mirror is to achieve high reflectivity for both polarizations. However, in polarizers, we are interested in separating the TE and TM polarizations. This can be accomplished with a periodic bilayer structure of the type shown in Fig. 8.8.1, which is highly reflecting only for TE and highly transmitting for TM polarizations. This is the principle of the so-called MacNeille polarizers [640,644,647,666,669,684–690].

If the angle of incidence θ_a is chosen such that the angle of refraction in the first high-index layer is equal to the Brewster angle of the high-low interface, then the TM component will not be reflected at the bilayer interfaces and will transmit through. The design condition is $\theta_H = \theta_B$, or $\sin \theta_H = \sin \theta_B$, which gives:

$$n_a \sin \theta_a = n_H \sin \theta_H = n_H \sin \theta_B = \frac{n_H n_L}{\sqrt{n_H^2 + n_L^2}} \tag{8.9.1}$$

This condition can be solved either for the angle θ_a or for the index n_a of the incident medium:

$$\sin \theta_a = \frac{n_H n_L}{n_a \sqrt{n_H^2 + n_L^2}} \quad \text{or,} \quad n_a = \frac{n_H n_L}{\sin \theta_a \sqrt{n_H^2 + n_L^2}} \tag{8.9.2}$$

In either case, the feasibility of this approach requires the opposite of the condition (8.8.3), that is,

$$n_a > \frac{n_H n_L}{\sqrt{n_H^2 + n_L^2}} \tag{8.9.3}$$

If the angle θ_a is set equal to the convenient value of 45°, then, condition Eq. (8.9.2) fixes the value of the refractive index n_a to be given by:

$$n_a = \frac{\sqrt{2} n_H n_L}{\sqrt{n_H^2 + n_L^2}} \tag{8.9.4}$$

Fig. 8.9.1 depicts such a multilayer structure sandwiched between two glass prisms with 45° angles. The thin films are deposited along the hypotenuse of each prism and the prisms are then cemented together. The incident, reflected, and transmitted beams are perpendicular to the prism sides.

Not many combinations of available materials satisfy condition (8.9.4). One possible solution is Banning's [647] with $n_H = 2.3$ (zinc sulfide), $n_L = 1.25$ (cryolite), and $n_a = 1.5532$. Another solution is given in Clapham, et al, [669], with $n_H = 2.04$ (zirconium oxide), $n_L = 1.385$ (magnesium fluoride), and $n_a = 1.6205$ (a form of dense flint glass.)

Fig. 8.9.2 shows the TE and TM reflectances of the case $n_H = 2.3$ and $n_L = 1.25$. The incident and output media had $n_a = n_b = 1.5532$. The maximum reflectivity for the TE component is 99.99%, while that of the TM component is 3% (note the different vertical scales in the two graphs.)
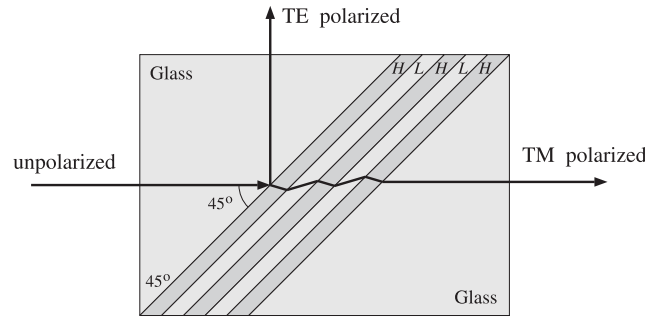
Fig. 8.9.1 Polarizing beam splitter.

The number of bilayers was $N = 5$ and the center frequency of the TE band was chosen to correspond to a wavelength of $\lambda_c = 500$ nm. To achieve this, the normalizing wavelength was required to be $\lambda_0 = 718.38$ nm. The layer lengths were quarter-wavelengths at λ_0. The TE bandwidth calculated with `omniband` is also shown.

The Brewster angles inside the high- and low-index layers are $\theta_H = 28.52^\circ$ and $\theta_L = 61.48^\circ$. As expected, they satisfy $\theta_H + \theta_L = 90^\circ$.
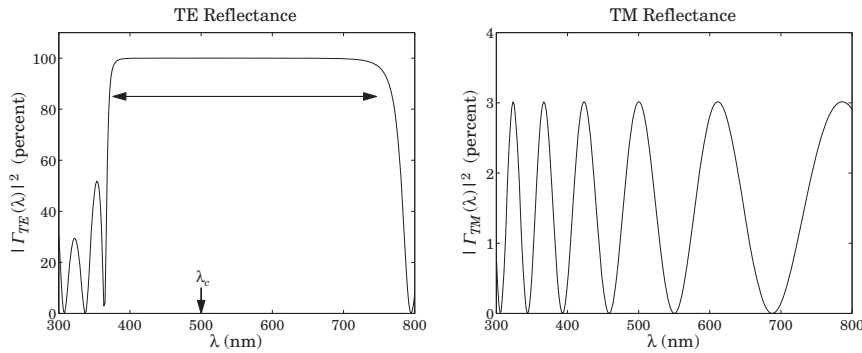

Fig. 8.9.2 Polarizer with $n_H = 2.3$ and $n_L = 1.25$.

Fig. 8.9.3 shows the second case having $n_H = 2.04$, $n_L = 1.385$, $n_a = n_b = 1.6205$. The normalizing wavelength was $\lambda_0 = 716.27$ nm in order to give $\lambda_c = 500$ nm. This case achieves a maximum TE reflectivity of 99.89% and TM reflectivity of only 0.53%. The typical MATLAB code generating these examples was:

```
nH = 2.3; nL = 1.25;
LH = 0.25; LL = 0.25;

na = nH*nL/sqrt(nH^2+nL^2)/sin(pi/4); nb=na;

[f1e,f2e] = omniband(na,nH,nL,LH,LL,th,'te',5);
```
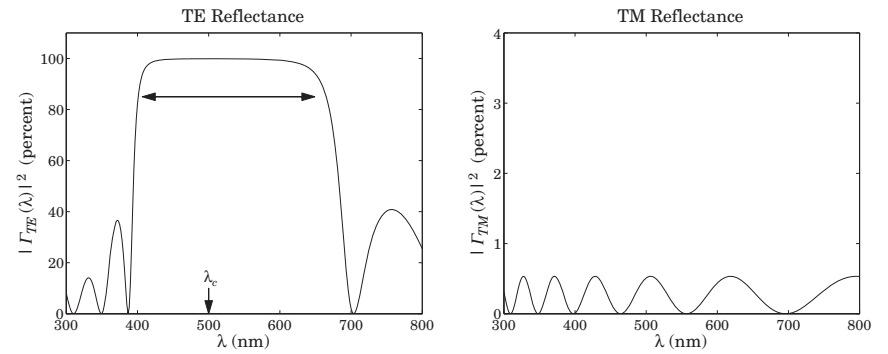
Fig. 8.9.3 Polarizer with $n_H = 2.04$ and $n_L = 1.385$.

```
lac = 500;
la0 = lac*(f1e+f2e)/2;          because λc = λ0/Fc

la = linspace(300,800,301);

N = 5;
n = [na, nH, repmat([nL,nH], 1, N), nb];
L = [LH, repmat([LL,LH], 1, N)];
Ge = 100*abs(multidiel(n,L,la/la0, th, 'te')).^2;
Gm = 100*abs(multidiel(n,L,la/la0, th, 'tm')).^2;

plot(la,Ge);
```

8.10 Reflection and Refraction in Birefringent Media

Uniform plane wave propagation in biaxial media was discussed in Sec. 4.6. We found that there is an effective refractive index N such that $k = Nk_0 = N\omega/c_0$. The index N, given by Eq. (4.6.8), depends on the polarization of the fields and the direction of the wave vector. The expressions for the TE and TM fields were given in Eqs. (4.6.18) and (4.6.27).

Here, we discuss how such fields get reflected and refracted at planar interfaces between biaxial media. Further discussion can be found in [638,57] and [698–719].

Fig. 7.1.1 depicts the TM and TE cases, with the understanding that the left and right biaxial media are described by the triplets of principal indices $\mathbf{n} = [n_1, n_2, n_3]$ and $\mathbf{n}' = [n_1', n_2', n_3']$, and that the E-fields are not perpendicular to the corresponding wave vectors in the TM case. The principal indices are aligned along the xyz axes, the xy-plane is the interface plane, and the xz-plane is the plane of incidence.

The boundary conditions require the matching of the electric field components that are tangential to the interface, that is, the components E_x in the TM case or E_y in TE. It proves convenient, therefore, to re-express Eq. (4.6.27) directly in terms of the E_x component and Eq. (4.6.18) in terms of E_y.

For the TM case, we write $E = \hat{x} E_x + \hat{z} E_z = E_x(\hat{x} - \hat{z}\tan\bar{\theta})$, for the electric field of the left-incident field, where we used $E_z = -E_x\tan\bar{\theta}$. Similarly, for the magnetic field we have from Eq. (4.6.26):

$$H = \frac{N}{\eta_0}\hat{y}(E_x\cos\theta - E_z\sin\theta) = \frac{N}{\eta_0}\hat{y}E_x\cos\theta\left(1 - \frac{E_z}{E_x}\tan\theta\right)$$

$$= \frac{N}{\eta_0}\hat{y}E_x\cos\theta\left(1 + \frac{n_1^2}{n_3^2}\tan^2\theta\right) = \frac{N}{\eta_0}\hat{y}E_x\cos\theta\left(\frac{n_3^2\cos^2\theta + n_1^2\sin^2\theta}{n_3^2\cos^2\theta}\right)$$

$$= \frac{N}{\eta_0}\hat{y}E_x\cos\theta\left(\frac{n_3^2 n_1^2}{N^2 n_3^2\cos^2\theta}\right) = \frac{E_x}{\eta_0}\frac{n_1^2}{N\cos\theta}\hat{y}$$

where we replaced $E_z/E_x = -\tan\bar{\theta} = -(n_1^2/n_3^2)\tan\theta$ and used Eq. (4.7.10). Thus,

$$\boxed{\begin{aligned} E(r) &= E_x\left(\hat{x} - \hat{z}\frac{n_1^2}{n_3^2}\tan\theta\right)e^{-j\mathbf{k}\cdot\mathbf{r}} \\ H(r) &= \frac{E_x}{\eta_0}\frac{n_1^2}{N\cos\theta}\hat{y}e^{-j\mathbf{k}\cdot\mathbf{r}} = \frac{E_x}{\eta_{TM}}\hat{y}e^{-j\mathbf{k}\cdot\mathbf{r}} \end{aligned}} \quad \text{(TM)} \quad (8.10.1)$$

Similarly, we may rewrite the TE case of Eq. (4.6.18) in the form:

$$\boxed{\begin{aligned} E(r) &= E_y\hat{y}e^{-j\mathbf{k}\cdot\mathbf{r}} \\ H(r) &= \frac{E_y}{\eta_0}n_2\cos\theta(-\hat{x} + \hat{z}\tan\theta)e^{-j\mathbf{k}\cdot\mathbf{r}} = \frac{E_y}{\eta_{TE}}(-\hat{x} + \hat{z}\tan\theta)e^{-j\mathbf{k}\cdot\mathbf{r}} \end{aligned}} \quad \text{(TE)} \quad (8.10.2)$$

The propagation phase factors are:

$$\boxed{e^{-j\mathbf{k}\cdot\mathbf{r}} = e^{-jk_0 x N\sin\theta - jk_0 z N\cos\theta}} \quad \text{(TM and TE propagation factors)} \quad (8.10.3)$$

Unlike the isotropic case, the phase factors are different in the TM and TE cases because the value of N is different, as given by Eq. (4.6.8), or,

$$N = \begin{cases} \dfrac{n_1 n_3}{\sqrt{n_1^2\sin^2\theta + n_3^2\cos^2\theta}}, & \text{(TM or p-polarization)} \\ n_2, & \text{(TE or s-polarization)} \end{cases} \quad (8.10.4)$$

In Eqs. (8.10.1) and (8.10.2), the effective *transverse impedances* are defined by $\eta_{TM} = E_x/H_y$ and $\eta_{TE} = -E_y/H_x$, and are given as follows:

$$\boxed{\eta_{TM} = \eta_0\frac{N\cos\theta}{n_1^2}, \quad \eta_{TE} = \frac{\eta_0}{n_2\cos\theta}} \quad \text{(transverse impedances)} \quad (8.10.5)$$

Defining the TM and TE effective *transverse refractive indices* through $\eta_{TM} = \eta_0/n_{TM}$ and $\eta_{TE} = \eta_0/n_{TE}$, we have:

$$\boxed{\begin{aligned} n_{TM} &= \frac{n_1^2}{N\cos\theta} = \frac{n_1 n_3}{\sqrt{n_3^2 - N^2\sin^2\theta}} \\ n_{TE} &= n_2\cos\theta \end{aligned}} \quad \text{(transverse refractive indices)} \quad (8.10.6)$$

where we used Eq. (4.6.23) for the TM case, that is,

$$\boxed{N\cos\theta = \frac{n_1}{n_3}\sqrt{n_3^2 - N^2\sin^2\theta}} \quad (8.10.7)$$

In the isotropic case, $N = n_1 = n_2 = n_3 = n$, Eqs. (8.10.6) reduce to Eq. (7.2.13). Next, we discuss the TM and TE reflection and refraction problems of Fig. 7.1.1.

Assuming that the interface is at $z = 0$, the equality of the total tangential electric fields (E_x component for TM and E_y for TE), implies as in Sec. 7.1 that the propagation phase factors must match at all values of x:

$$e^{-jk_{x+}x} = e^{-jk_{x-}x} = e^{-jk'_{x+}x} = e^{-jk'_{x-}x}$$

which requires that $k_{x+} = k_{x-} = k'_{x+} = k'_{x-}$, or, because $k_x = k\sin\theta = Nk_0\sin\theta$:

$$N\sin\theta_+ = N\sin\theta_- = N'\sin\theta'_+ = N'\sin\theta'_-$$

This implies Snel's law of reflection, that is, $\theta_+ = \theta_- \equiv \theta$ and $\theta'_+ = \theta'_- \equiv \theta'$, and Snel's law of refraction,

$$\boxed{N\sin\theta = N'\sin\theta'} \quad \text{(Snel's law for birefringent media)} \quad (8.10.8)$$

Thus, Snel's law is essentially the same as in the isotropic case, provided one uses the effective refractive index N. Because N depends on the polarization, there will be two different refraction angles[†] for the same angle of incidence. In particular, Eq. (8.10.8) can be written explicitly in the two polarization cases:

$$\frac{n_1 n_3\sin\theta}{\sqrt{n_1^2\sin^2\theta + n_3^2\cos^2\theta}} = \frac{n'_1 n'_3\sin\theta'}{\sqrt{n_1'^2\sin^2\theta' + n_3'^2\cos^2\theta'}} \quad \text{(TM)} \quad (8.10.9a)$$

$$n_2\sin\theta = n'_2\sin\theta' \quad \text{(TE)} \quad (8.10.9b)$$

Both expressions reduce to Eq. (7.1.6) in the isotropic case. The explicit solutions of Eq. (8.10.9a) for $\sin\theta'$ and $\sin\theta$ are:

$$\begin{aligned} \sin\theta' &= \frac{n_1 n_3 n'_3\sin\theta}{\sqrt{\left[n_1'^2 n_3'^2(n_1^2 - n_3^2) - n_1^2 n_3^2(n_1'^2 - n_3'^2)\right]\sin^2\theta + n_1'^2 n_3'^2 n_3^2}} \\ \sin\theta &= \frac{n'_1 n'_3 n_3\sin\theta'}{\sqrt{\left[n_1^2 n_3^2(n_1'^2 - n_3'^2) - n_1'^2 n_3'^2(n_1^2 - n_3^2)\right]\sin^2\theta' + n_1^2 n_3^2 n_3'^2}} \end{aligned} \quad (8.10.10)$$

[†] Hence, the name birefringent.

The MATLAB function snel, solves Eqs. (8.10.9) for θ' given the angle of incidence θ and the polarization type. It works for any type of medium, isotropic, uniaxial, or biaxial. It has usage:

```
thb = snel(na,nb,tha,pol);        % refraction angle from Snel's law
```

The refractive index inputs na, nb may be entered as 1-, 2-, or 3-dimensional column or row vectors, for example, $\mathbf{n}_a = [n_a]$ (isotropic), $\mathbf{n}_a = [n_{ao}, n_{ae}]$ (uniaxial), or $\mathbf{n}_a = [n_{a1}, n_{a2}, n_{a3}]$ (biaxial).

Next, we discuss the propagation and matching of the transverse fields. All the results of Sec. 7.3 translate verbatim to the birefringent case, provided one uses the proper transverse refractive indices according to Eq. (8.10.6).

In particular, the propagation equations (7.3.5)–(7.3.7) for the transverse fields, for the transverse reflection coefficients Γ_T, and for the transverse wave impedances Z_T, remain unchanged.

The *phase thickness* δ_z for propagating along z by a distance l also remains the same as Eq. (7.3.8), except that the index N must be used in the optical length, and therefore, δ_z depends on the polarization:

$$\delta_z = k_z l = k l \cos\theta = N k_0 l \cos\theta = \frac{2\pi}{\lambda} l N \cos\theta \qquad (8.10.11)$$

Using Eq. (8.10.7), we have explicitly:

$$\delta_z = \frac{2\pi}{\lambda} l \frac{n_1}{n_3} \sqrt{n_3^2 - N^2 \sin^2\theta}, \qquad \text{(TM)} \qquad (8.10.12a)$$

$$\delta_z = \frac{2\pi}{\lambda} l n_2 \cos\theta, \qquad \text{(TE)} \qquad (8.10.12b)$$

The transverse matching matrix (7.3.11) and Fresnel reflection coefficients (7.3.12) remain the same. Explicitly, we have in the TM and TE cases:

$$\rho_{TM} = \frac{n_{TM} - n'_{TM}}{n_{TM} + n'_{TM}} = \frac{\dfrac{n_1^2}{N\cos\theta} - \dfrac{n_1'^2}{N'\cos\theta'}}{\dfrac{n_1^2}{N\cos\theta} + \dfrac{n_1'^2}{N'\cos\theta'}}$$

$$\rho_{TE} = \frac{n_{TE} - n'_{TE}}{n_{TE} + n'_{TE}} = \frac{n_2 \cos\theta - n'_2 \cos\theta'}{n_2 \cos\theta + n'_2 \cos\theta'} \qquad (8.10.13)$$

Using Eq. (8.10.6) and the TM and TE Snel's laws, Eqs. (8.10.9), we may rewrite the reflection coefficients in terms of the angle θ only:

$$\rho_{TM} = \frac{n_1 n_3 \sqrt{n_3'^2 - N^2 \sin^2\theta} - n'_1 n'_3 \sqrt{n_3^2 - N^2 \sin^2\theta}}{n_1 n_3 \sqrt{n_3'^2 - N^2 \sin^2\theta} + n'_1 n'_3 \sqrt{n_3^2 - N^2 \sin^2\theta}}$$

$$\rho_{TE} = \frac{n_2 \cos\theta - \sqrt{n_2'^2 - n_2^2 \sin^2\theta}}{n_2 \cos\theta + \sqrt{n_2'^2 - n_2^2 \sin^2\theta}} \qquad (8.10.14)$$

The quantity $N^2 \sin^2\theta$ can be expressed directly in terms of θ and the refractive indices of the incident medium. Using Eq. (8.10.4), we have:

$$N^2 \sin^2\theta = \frac{n_1^2 n_3^2 \sin^2\theta}{n_1^2 \sin^2\theta + n_3^2 \cos^2\theta} \qquad (8.10.15)$$

The TE reflection coefficient behaves like the TE isotropic case. The TM coefficient exhibits a much more complicated behavior. If $n_1 = n'_1$ but $n_3 \neq n'_3$, it behaves like the TM isotropic case. If $n_3 = n'_3$ but $n_1 \neq n'_1$, the square-root factors cancel and it becomes independent of θ:

$$\rho_{TM} = \frac{n_1 - n'_1}{n_1 + n'_1} \qquad (8.10.16)$$

Another interesting case is when both media are uniaxial and $n'_3 = n_1$ and $n'_1 = n_3$, that is, the refractive index vectors are $\mathbf{n} = [n_1, n_1, n_3]$ and $\mathbf{n}' = [n_3, n_3, n_1]$. It is straightforward to show in this case that $\rho_{TM} = \rho_{TE}$ at all angles of incidence. Multilayer films made from alternating such materials exhibit similar TM and TE optical properties [698].

The MATLAB function fresnel can evaluate Eqs. (8.10.14) at any range of incident angles θ. The function determines internally whether the media are isotropic, uniaxial, or biaxial.

8.11 Brewster and Critical Angles in Birefringent Media

The maximum angle of refraction, critical angle of incidence, and Brewster angle, have their counterparts in birefringent media.

It is straightforward to verify that θ' is an increasing function of θ in Eq. (8.10.9). The maximum angle of refraction θ'_c is obtained by setting $\theta = 90^\circ$ in Eq. (8.10.9).

For the TE case, we obtain $\sin\theta'_c = n_2/n'_2$. As in the isotropic case of Eq. (7.5.2), this requires that $n_2 < n'_2$, that is, the incident medium is less dense than the transmitted medium, with respect to the index n_2. For the TM case, we obtain from Eq. (8.10.9a):

$$\sin\theta'_c = \frac{n_3 n'_3}{\sqrt{n_3^2 n_3'^2 + n_1'^2 (n_3'^2 - n_3^2)}} \qquad \text{(maximum TM refraction angle)} \qquad (8.11.1)$$

This requires that $n_3 < n'_3$. On the other hand, if $n_3 > n'_3$, we obtain the critical angle of incidence θ_c that corresponds to $\theta' = 90^\circ$ in Eq. (8.10.10):

$$\sin\theta_c = \frac{n_3 n'_3}{\sqrt{n_3^2 n_3'^2 + n_1^2 (n_3^2 - n_3'^2)}} \qquad \text{(critical TM angle)} \qquad (8.11.2)$$

whereas for the TE case, we have $\sin\theta_c = n'_2/n_2$, which requires $n_2 > n'_2$.

In the isotropic case, a Brewster angle always exists at which the TM reflection coefficient vanishes, $\rho_{TM} = 0$. In the birefringent case, the Brewster angle does not necessarily exist, as is the case of Eq. (8.10.16), and it can also have the value zero, or even be imaginary.

The Brewster angle condition $\rho_{TM} = 0$ is equivalent to the equality of the transverse refractive indices $n_{TM} = n'_{TM}$. Using Eq. (8.10.6), we obtain:

$$n_{TM} = n'_{TM} \quad \Rightarrow \quad \frac{n_1 n_3}{\sqrt{n_3^2 - N^2 \sin^2 \theta}} = \frac{n'_1 n'_3}{\sqrt{n_3'^2 - N^2 \sin^2 \theta}} \qquad (8.11.3)$$

where $N^2 \sin^2 \theta$ is given by Eq. (8.10.15). Solving for θ, we obtain the expression for the Brewster angle from the left medium:

$$\boxed{\tan \theta_B = \frac{n_3 n'_3}{n_1^2} \sqrt{\frac{n_1^2 - n_1'^2}{n_3^2 - n_3'^2}}} \qquad \text{(Brewster angle)} \qquad (8.11.4)$$

Working instead with $N' \sin \theta' = N \sin \theta$, we obtain the Brewster angle from the right medium, interchanging the roles of the primed and unprimed quantities:

$$\boxed{\tan \theta'_B = \frac{n_3 n'_3}{n_1'^2} \sqrt{\frac{n_1^2 - n_1'^2}{n_3^2 - n_3'^2}}} \qquad \text{(Brewster angle)} \qquad (8.11.5)$$

Eqs. (8.11.4) and (8.11.5) reduce to Eqs. (7.6.2) and (7.6.3) in the isotropic case. It is evident from Eq. (8.11.4) that θ_B is a real angle only if the quantity under the square root is non-negative, that is, only if $n_1 > n'_1$ and $n_3 > n'_3$, or if $n_1 < n'_1$ and $n_3 < n'_3$. Otherwise, θ_B is imaginary. In the special case, $n_1 = n'_1$ but $n_3 \neq n'_3$, the Brewster angle vanishes. If $n_3 = n'_3$, the Brewster angle does not exist, since then ρ_{TM} is given by Eq. (8.10.16) and cannot vanish.

The MATLAB function `brewster` computes the Brewster angle θ_B, as well as the critical angles θ_c and θ'_c. For birefringent media the critical angles depend on the polarization. Its usage is as follows:

```
[thB,thc] = brewster(na,nb)          % isotropic case
[thB,thcTE,thcTM] = brewster(na,nb)  % birefringent case
```

In multilayer systems, it is convenient to know if the Brewster angle of an internal interface is accessible from the incident medium. Using Snel's law we have in this case $N_a \sin \theta_a = N \sin \theta$, where θ_a is the incident angle and N_a the effective index of the incident medium. It is simpler, then, to solve Eq. (8.11.3) directly for θ_a:

$$\boxed{N_a^2 \sin^2 \theta_a = N^2 \sin^2 \theta_B = \frac{n_3^2 n_3'^2 (n_1^2 - n_1'^2)}{n_1^2 n_3^2 - n_3'^2 n_1'^2}} \qquad (8.11.6)$$

Example 8.11.1: To illustrate the variety of possible Brewster angle values, we consider the following birefringent cases:

$$
\begin{array}{lll}
(a) & \mathbf{n} = [1.63, 1.63, 1.5], & \mathbf{n}' = [1.63, 1.63, 1.63] \\
(b) & \mathbf{n} = [1.54, 1.54, 1.63], & \mathbf{n}' = [1.5, 1.5, 1.5] \\
(c) & \mathbf{n} = [1.8, 1.8, 1.5], & \mathbf{n}' = [1.5, 1.5, 1.5] \\
(d) & \mathbf{n} = [1.8, 1.8, 1.5], & \mathbf{n}' = [1.56, 1.56, 1.56]
\end{array}
$$

These cases were discussed in [698]. The corresponding materials are: (a) birefringent polyester and isotropic polyester, (b) syndiotactic polystyrene and polymethylmethacrylate (PMMA), (c) birefringent polyester and PMMA, and (d) birefringent polyester and isotropic polyester.

Because $n_1 = n'_1$ in case (a), the Brewster angle will be zero, $\theta_B = 0^\circ$. In case (b), we calculate $\theta_B = 29.4^\circ$. Because $n_2 > n'_2$ and $n_3 > n'_3$, there will be both TE and TM critical angles of reflection: $\theta_{c,TE} = 76.9^\circ$ and $\theta_{c,TM} = 68.1^\circ$.

In case (c), the Brewster angle does not exist because $n_3 = n'_3$, and in fact, the TM reflection coefficient is independent of the incident angle as in Eq. (8.10.16). The corresponding critical angles of reflection are: $\theta_{c,TE} = 56.4^\circ$ and $\theta_{c,TM} = 90^\circ$.

Finally, in case (d), because $n_2 > n'_2$ but $n_3 < n'_3$, the Brewster angle will be imaginary, and there will be a TE critical angle of reflection and a TM maximum angle of refraction: $\theta_{c,TE} = 60.1^\circ$ and $\theta'_{c,TM} = 74.1^\circ$.

Fig. 8.11.1 shows the TM and TE reflection coefficients $|\rho_{TM}(\theta)|$ of Eq. (8.10.14) versus θ in the range $0 \leq \theta \leq 90^\circ$.
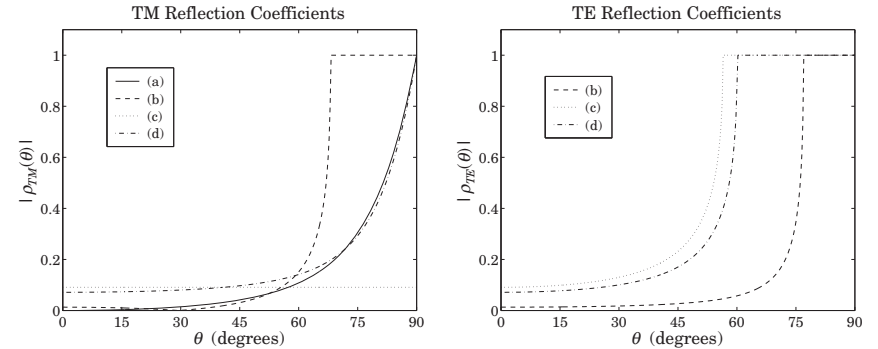
Fig. 8.11.1 TM and TE birefringent Fresnel reflection coefficients versus incident angle.

The TE coefficient in case (a) is not plotted because it is identically zero. In order to expand the vertical scales, Fig. 8.11.2 shows the TM reflectances normalized by their values at $\theta = 0^\circ$, that is, it plots the quantities $|\rho_{TM}(\theta)/\rho_{TM}(0^\circ)|^2$. Because in case (a) $\rho_{TM}(0^\circ) = 0$, we have plotted instead the scaled-up quantity $|100\rho_{TM}(\theta)|^2$.

The typical MATLAB code used to compute the critical angles and generate these graphs was:

```
th = linspace(0,90,361);              % θ at 1/4° intervals

na = [1.63,1.63,1.5]; nb = [1.63,1.63,1.63];   % note the variety of
```
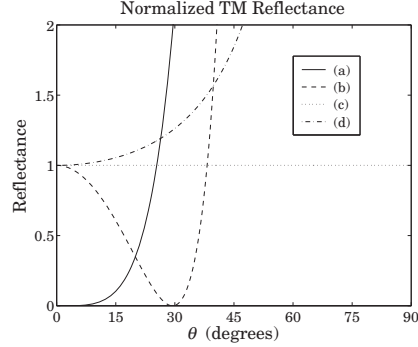
Fig. 8.11.2 TM reflectances normalized at normal incidence.

```
[rte1,rtm1] = fresnel(na,nb,th);          % equivalent ways of
[thb1,thcTE1,thcTM1] = brewster(na,nb);   % entering na and nb

na = [1.54,1.63];
nb = [1.5, 1.5];                          % FRESNEL and BREWSTER
[rte2,rtm2] = fresnel(na,nb,th);          % internally extend
[thb2,thcTE2,thcTM2] = brewster(na,nb);   % na,nb into 3-d arrays

na = [1.8, 1.5];                          % same as na=[1.8,1.8,1.5]
nb = 1.5;                                 % and nb=[1.5,1.5,1.5]
[rte3,rtm3] = fresnel(na,nb,th);
[thb3,thcTE3,thcTM3] = brewster(na,nb);   % in this case, θ_B = []

na = [1.8,1.5];
nb = 1.56;
[rte4,rtm4] = fresnel(na,nb,th);
[thb4,thcTE4,thcTM4] = brewster(na,nb);

plot(th, abs([rtm1; rtm2; rtm3; rtm4]));
```

We note four striking properties of the birefringent cases that have no counterparts for isotropic materials: (i) The Brewster angle can be zero, (ii) the Brewster angle may not exist, (iii) the Brewster angle may be imaginary with the TE and TM reflection coefficients both increasing monotonically with the incident angle, and (iv) there may be total internal reflection in one polarization but not in the other.

8.12 *Multilayer Birefringent Structures*

With some redefinitions, all the results of Sec. 8.1 on multilayer dielectric structures translate essentially unchanged to the birefringent case.

We assume the same M-layer configuration shown in Fig. 8.1.1, where now each layer is a biaxial material. The orthogonal optic axes of all the layers are assumed to be aligned with the xyz film axes. The xz-plane is the plane of incidence, the layer interfaces are parallel to the xy-plane, and the layers are arranged along the z-axis.

The ith layer is described by the triplet of refractive indices $\mathbf{n}_i = [n_{i1}, n_{i2}, n_{i3}]$, $i = 1, 2, \ldots, M$. The incident and exit media a, b may also be birefringent with $\mathbf{n}_a = [n_{a1}, n_{a2}, n_{a3}]$ and $\mathbf{n}_b = [n_{b1}, n_{b2}, n_{b3}]$, although in our examples, we will assume that they are isotropic.

The reflection/refraction angles in each layer depend on the assumed polarization and are related to each other by the birefringent version of Snel's law, Eq. (8.10.8):

$$\boxed{N_a \sin \theta_a = N_i \sin \theta_i = N_b \sin \theta_b}, \quad i = 1, 2 \ldots, M \tag{8.12.1}$$

where N_a, N_i, N_b are the effective refractive indices given by Eq. (8.10.4). The phase thickness of the ith layer depends on the polarization:

$$\delta_i = \frac{2\pi}{\lambda} l_i N_i \cos \theta_i = \begin{cases} \dfrac{2\pi}{\lambda} l_i n_{i1} \sqrt{1 - \dfrac{N_a^2 \sin^2 \theta_a}{n_{i3}^2}}, & \text{(TM)} \\[3mm] \dfrac{2\pi}{\lambda} l_i n_{i2} \sqrt{1 - \dfrac{N_a^2 \sin^2 \theta_a}{n_{i2}^2}}, & \text{(TE)} \end{cases} \tag{8.12.2}$$

where we used Eq. (8.10.7) and Snel's law to write in the TM and TE cases:

$$N_i \cos \theta_i = \begin{cases} \dfrac{n_{i1}}{n_{i3}} \sqrt{n_{i3}^2 - N_i^2 \sin^2 \theta_i} = n_{i1} \sqrt{1 - \dfrac{N_i^2 \sin^2 \theta_i}{n_{i3}^2}} = n_{i1} \sqrt{1 - \dfrac{N_a^2 \sin^2 \theta_a}{n_{i3}^2}} \\[4mm] n_{i2} \cos \theta_i = n_{i2} \sqrt{1 - \sin^2 \theta_i} = n_{i2} \sqrt{1 - \dfrac{N_a^2 \sin^2 \theta_a}{n_{i2}^2}} \end{cases}$$

To use a unified notation for the TM and TE cases, we define the layer optical lengths at normal-incidence, normalized by a fixed free-space wavelength λ_0:

$$L_i = \begin{cases} \dfrac{l_i n_{i1}}{\lambda_0}, & \text{(TM)} \\[3mm] \dfrac{l_i n_{i2}}{\lambda_0}, & \text{(TE)} \end{cases}, \quad i = 1, 2, \ldots, M \tag{8.12.3}$$

We define also the cosine coefficients c_i, which represent $\cos \theta_i$ in the TE birefringent case and in the TM isotropic case:

$$c_i = \begin{cases} \sqrt{1 - \dfrac{N_a^2 \sin^2 \theta_a}{n_{i3}^2}}, & \text{(TM)} \\[3mm] \sqrt{1 - \dfrac{N_a^2 \sin^2 \theta_a}{n_{i2}^2}}, & \text{(TE)} \end{cases}, \quad i = 1, 2, \ldots, M \tag{8.12.4}$$

At normal incidence the cosine factors are unity, $c_i = 1$. With these definitions, Eq. (8.12.2) can be written compactly in the form:

$$\boxed{\delta_i = 2\pi \frac{\lambda_0}{\lambda} L_i c_i = 2\pi \frac{f}{f_0} L_i c_i}, \quad i = 1, 2, \ldots, M \tag{8.12.5}$$

where λ is the operating free-space wavelength and $f = c_0/\lambda$, $f_0 = c_0/\lambda_0$. This is the birefringent version of Eq. (8.1.10). A typical design might use quarter-wave layers, $L_i = 1/4$, at λ_0 and at normal incidence.

The reflection coefficients ρ_{Ti} at the interfaces are given by Eq. (8.1.3), but now the transverse refractive indices are defined by the birefringent version of Eq. (8.1.4):

$$n_{Ti} = \begin{cases} \dfrac{n_{i1}^2}{N_i \cos\theta_i} = \dfrac{n_{i1}n_{i3}}{\sqrt{n_{i3}^2 - N_a^2 \sin^2\theta_a}}, & \text{(TM)} \\[2ex] n_{i2}\cos\theta_i = \sqrt{n_{i2}^2 - N_a^2 \sin^2\theta_a}, & \text{(TE)} \end{cases} \quad , \quad i = a, 1, 2, \ldots, M, b \quad (8.12.6)$$

With the above redefinitions, the propagation and matching equations (8.1.5)–(8.1.9) remain unchanged. The MATLAB function `multidiel` can also be used in the birefringent case to compute the frequency reflection response of a multilayer structure. Its usage is still:

```
[Gamma1,Z1] = multidiel(n,L,lambda,theta,pol);     % birefringent multilayer structure
```

where the input n is a $1 \times (M+2)$ vector of refractive indices in the isotropic case, or a $3 \times (M+2)$ matrix, where each column represents the triplet of birefringent indices of each medium. For uniaxial materials, n may be entered as a $2 \times (M+2)$ matrix.

8.13 Giant Birefringent Optics

The results of Sec. 8.8 can be applied almost verbatim to the birefringent case. In Fig. 8.8.1, we assume that the high and low alternating layers are birefringent, described by the triplet indices $\mathbf{n}_H = [n_{H1}, n_{H2}, n_{H3}]$ and $\mathbf{n}_L = [n_{L1}, n_{L2}, n_{L3}]$. The entry and exit media may also be assumed to be birefringent. Then, Snel's laws give:

$$N_a \sin\theta_a = N_H \sin\theta_H = N_L \sin\theta_L = N_b \sin\theta_b \qquad (8.13.1)$$

The phase thicknesses δ_H and δ_L within the high and low index layers are:

$$\delta_H = 2\pi \frac{f}{f_0} L_H c_H, \quad \delta_L = 2\pi \frac{f}{f_0} L_L c_L \qquad (8.13.2)$$

where L_H, c_H and L_L, c_L are defined by Eqs. (8.12.3) and (8.12.4) for $i = H, L$. The effective transverse refractive indices within the high and low index layers are given by Eq. (8.12.6), again with $i = H, L$.

The alternating reflection coefficient ρ_T between the high/low interfaces is given by Eq. (8.10.14), with the quantity $N^2 \sin^2\theta$ replaced by $N_a^2 \sin^2\theta_a$ by Snel's law:

$$\rho_{TM} = \frac{n_{H1}n_{H3}\sqrt{n_{L3}^2 - N_a^2 \sin^2\theta_a} - n_{L1}n_{L3}\sqrt{n_{H3}^2 - N_a^2 \sin^2\theta_a}}{n_{H1}n_{H3}\sqrt{n_{L3}^2 - N_a^2 \sin^2\theta_a} + n_{L1}n_{L3}\sqrt{n_{H3}^2 - N_a^2 \sin^2\theta_a}}$$

$$\rho_{TE} = \frac{\sqrt{n_{H2}^2 - N_a^2 \sin^2\theta_a} - \sqrt{n_{L2}^2 - N_a^2 \sin^2\theta_a}}{\sqrt{n_{H2}^2 - N_a^2 \sin^2\theta_a} + \sqrt{n_{L2}^2 - N_a^2 \sin^2\theta_a}}$$

$$(8.13.3)$$

The multilayer structure will exhibit reflection bands whose bandedges can be calculated from Eqs. (8.8.7)–(8.8.17), with the redefinition $L_{\pm} = L_H c_H \pm L_L c_L$. The MATLAB function `omniband2` calculates the bandedges. It has usage:

```
[F1,F2] = omniband2(na,nH,nL,LH,LL,th,pol,N);
```

where `pol` is one of the strings `'te'` or `'tm'` for TE or TM polarization, and na, nH, nL are 1-d, 2-d, or 3-d row or column vectors of birefringent refractive indices.

Next, we discuss some mirror design examples from [698] that illustrate some properties that are specific to birefringent media. The resulting optical effects in such mirror structures are referred to as *giant birefringent optics* (GBO) in [698,1853].

Example 8.13.1: We consider a GBO mirror consisting of 50-bilayers of high and low index quarter-wave layers with refractive indices $\mathbf{n}_H = [1.8, 1.8, 1.5]$, $\mathbf{n}_L = [1.5, 1.5, 1.5]$ (birefringent polyester and isotropic PMMA.) The surrounding media are air, $n_a = n_b = 1$.

The layers are quarter wavelength at the normalization wavelength $\lambda_0 = 700$ nm at normal incidence, so that for both polarizations we take $L_H = L_L = 1/4$.

Because the high/low index layers are matched along the z-direction, $n_{H3} = n_{L3}$, the TM reflection coefficient at the high/low interface will be constant, independent of the incident angle θ_a, as in Eq. (8.10.16). However, some dependence on θ_a is introduced through the cosine factors c_H, c_L of Eq. (8.13.2).

The left graph of Fig. 8.13.1 shows the reflectance $|\Gamma_T(\lambda)|^2$ as a function of λ for an angle of incidence $\theta_a = 60^\circ$. The TM and TE bandedge wavelengths were calculated from `omniband2` to be: $[\lambda_1, \lambda_2] = [540.24, 606.71]$ and $[\lambda_1, \lambda_2] = [548.55, 644.37]$ nm.

The typical MATLAB code used to generate the left graph and the bandedge wavelengths was as follows:

```
LH = 0.25; LL = 0.25;

na = [1; 1; 1];
nH = [1.8; 1.8; 1.5];
nL = [1.5; 1.5; 1.5];
nb = [1; 1; 1];

la0 = 700;
la = linspace(400,1000,601);

th = 60;                    % angle of incidence

N = 50;                     % number of bilayers
```
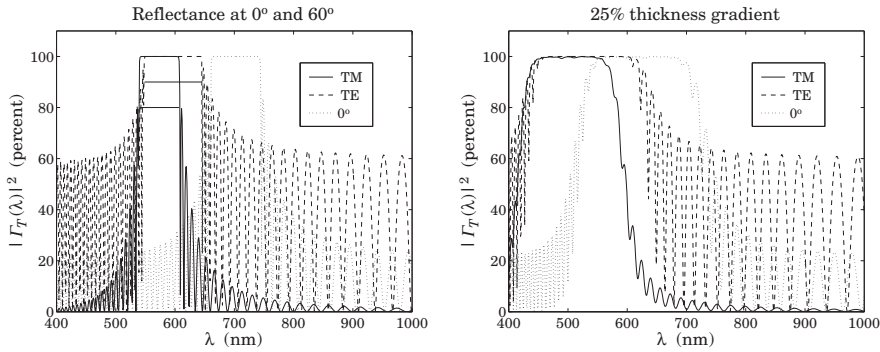
Fig. 8.13.1   Reflectance of birefringent mirror.

Fig. 8.13.2   Birefringent mirror with identical TM and TE reflection bands.

```
n = [na, repmat([nH,nL], 1, N), nb];            % 3×(2N + 2) matrix
L = [repmat([LH,LL], 1, N)];

Ge = 100*abs(multidiel(n, L, la/la0, th, 'te')).^2;
Gm = 100*abs(multidiel(n, L, la/la0, th, 'tm')).^2;
G0 = 100*abs(multidiel(n, L, la/la0)).^2;

plot(la,Gm,'-', la,Ge,'--', la,G0,':');

[F1,F2]=omniband2(na,nH,nL,LH,LL,th,'tm',3);
la1 = la0/F2;   la2 = la0/F1;                   % TM bandedge wavelengths
```

The right graph shows the reflectance with a 25% thickness gradient (the layer thicknesses L_H, L_L decrease linearly from quarter-wavelength to 25% less than that at the end.) This can be implemented in MATLAB by defining the thickness vector L by:

```
L = [repmat([LH,LL], 1, N)];
L = L .* (1 - linspace(0, 0.25, 2*N));    % 25% thickness gradient
```

The thickness gradient increases the effective bandwidth of the reflecting bands [696]. However, the bandwidth calculation can no longer be done with `omniband2`. The band centers can be shifted to higher wavelengths by choosing λ_0 higher. The reflecting bands can be made flatter by increasing the number of bilayers.   □

Example 8.13.2:   In this example, we design a 30-bilayer GBO mirror with $\mathbf{n}_H = [1.8, 1.8, 1.5]$ and $\mathbf{n}_L = [1.5, 1.5, 1.8]$, so that $n_{H1} = n_{H2} = n_{L3}$ and $n_{H3} = n_{L1} = n_{L2}$. As we discussed in Sec. 8.10, it follows from Eq. (8.10.14) that $\rho_{TM} = \rho_{TE}$ for all angles of incidence.

As in Ref. [698], the media a, b are taken to be isotropic with $n_a = n_b = 1.4$. The normalization wavelength at which the high and low index layers are quarter-wavelength is $\lambda_0 = 700$ nm.

The left graph of Fig. 8.13.2 shows the reflectance for a 45° angle of incidence. Because $\rho_{TM} = \rho_{TE}$, the reflection bands for the TM and TE cases are essentially the same.

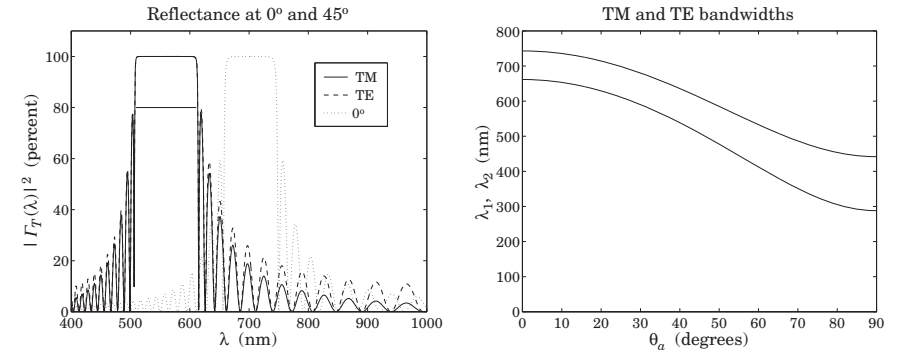The right graph depicts the asymptotic (for large number of bilayers) bandedges of the reflecting band versus incident angle. They were computed with `omniband2`. Unlike the isotropic case, the TM and TE bands are exactly identical. This is a consequence of the following relationships between the cosine factors in this example: $c_{H,TM} = c_{L,TE}$ and $c_{H,TE} = c_{L,TM}$. Then, because we assume quarter-wave layers in both the TE and TM cases, $L_H = L_L = 1/4$, we will have:

$$L_{+,TM} = L_{H,TM}c_{H,TM} + L_{L,TM}c_{L,TM} = = \frac{1}{4}\left(c_{H,TM} + c_{L,TM}\right) = \frac{1}{4}\left(c_{L,TE} + c_{H,TE}\right) = L_{+,TE}$$

$$L_{-,TM} = L_{H,TM}c_{H,TM} - L_{L,TM}c_{L,TM} = = \frac{1}{4}\left(c_{H,TM} - c_{L,TM}\right) = \frac{1}{4}\left(c_{L,TE} - c_{H,TE}\right) = -L_{+,TE}$$

Because the computational algorithm (8.8.17) for the bandwidth does not depend on the sign of L_-, it follows that Eq. (8.8.17) will have the same solution for the TM and TE cases. The typical MATLAB code for this example was:

```
LH = 0.25; LL = 0.25;

na = [1.4; 1.4; 1.4];
nb = [1.4; 1.4; 1.4];
nH = [1.8; 1.8; 1.5];
nL = [1.5; 1.5; 1.8];

la0 = 700;
la = linspace(400,1000,601);

tha = 45;

N = 30;
n = [na, repmat([nH,nL], 1, N), nb];
L = [repmat([LH,LL], 1, N)];

Ge = 100*abs(multidiel(n, L, la/la0, tha, 'te')).^2;
Gm = 100*abs(multidiel(n, L, la/la0, tha, 'tm')).^2;
G0 = 100*abs(multidiel(n, L, la/la0)).^2;

plot(la,Gm,'-', la,Ge,'--', la,G0,':');
```

In Fig. 8.13.3, the low-index material is changed slightly to $\mathbf{n}_L = [1.5, 1.5, 1.9]$. The main behavior of the structure remains the same, except now the TM and TE bands are slightly different.
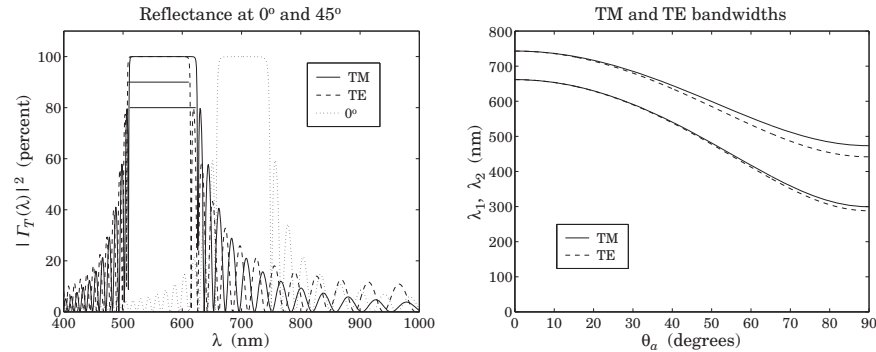

Fig. 8.13.3 Birefringent mirror with slightly different TM and TE reflection bands.

The MATLAB code used to compute the right graph was:

```
theta = linspace(0,90,361);                        % incident angles

F1e = []; F2e = [];
F1m = []; F2m = [];

Ni = 3;                                            % refinement iterations

for i=1:length(theta),
  [f1e,f2e] = omniband2(na,nH,nL,LH,LL,theta(i),'te',Ni);
  [f1m,f2m] = omniband2(na,nH,nL,LH,LL,theta(i),'tm',Ni);
  F1e = [F1e,f1e]; F2e = [F2e,f2e];
  F1m = [F1m,f1m]; F2m = [F2m,f2m];                % frequency bandedges
end

la1e = la0 ./ F2e; la2e = la0 ./ F1e;              % wavelength bandedges
la1m = la0 ./ F2m; la2m = la0 ./ F1m;

plot(theta,la1m,'-', theta,la2m,'-', theta,la1e,'--', theta,la2e,'--');
```

As the incident angle increases, not only does the TM band widen but it also becomes wider than the TE band—exactly the opposite behavior from the isotropic case.   □

Example 8.13.3: *GBO Reflective Polarizer.* By choosing biaxial high/low layers whose refractive indices are mismatched only in the x or the y direction, one can design a mirror structure that reflects only the TM or only the TE polarization.

Fig. 8.13.4 shows the reflectance of an 80-bilayer mirror with $\mathbf{n}_H = [1.86, 1.57, 1.57]$ for the left graph, and $\mathbf{n}_H = [1.57, 1.86, 1.57]$ for the right one. In both graphs, the low index material is the same, with $\mathbf{n}_L = [1.57, 1.57, 1.57]$.

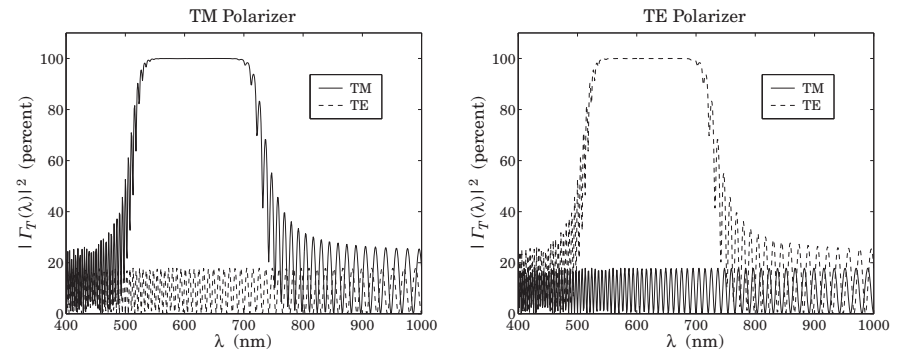The angle of incidence was $\theta_a = 0^o$. The typical MATLAB code was:

Fig. 8.13.4 TM and TE mirror polarizers.

```
LH = 0.25; LL = 0.25;

na = [1; 1; 1];
nb = [1; 1; 1];
nH = [1.86; 1.57; 1.57];
nL = [1.57; 1.57; 1.57];

la0 = 700;
la = linspace(400,1000,601);

N = 80;
n = [na, repmat([nH,nL], 1, N), nb];
L = [repmat([LH,LL], 1, N)];
L = L .* linspace(1,0.75,2*N);                     % 25% thickness gradient

Ge = 100*abs(multidiel(n, L, la/la0, 0, 'te')).^2;
Gm = 100*abs(multidiel(n, L, la/la0, 0, 'tm')).^2;

plot(la,Gm,'-', la,Ge,'--');
```

A 25% thickness gradient was assumed in both cases. In the first case, the x-direction indices are different and the structure will act as a mirror for the TM polarization. The TE polarization will be reflected only by the air-high interface.

In the second case, the materials are matched in their y-direction indices and therefore, the structure becomes a mirror for the TE polarization, assuming as always that the plane of incidence is still the xz plane.   □

   Giant birefringent optics is a new paradigm in the design of multilayer mirrors and polarizers [698], offering increased flexibility in the control of reflected light. The recently manufactured *multilayer optical film* by 3M Corp. [1853] consists of hundreds to thousands of birefringent polymer layers with individual thicknesses of the order of a wavelength and total thickness of a sheet of paper. The optical working range of such films are between 400–2500 nm.

   Applications include the design of efficient waveguides for transporting visible light over long distances and piping sunlight into interior rooms, reflective polarizers for

improving liquid crystal displays, and other products, such as various optoelectronic components, cosmetics, and "hot" and "cold" mirrors for architectural and automotive windows.

8.14 Problems

8.1 Prove the reflectance and transmittance formulas (8.4.6) in FTIR.

8.2 *Computer Experiment—FTIR.* Reproduce the results and graphs of Figures 8.4.3–8.4.5.

8.3 *Computer Experiment—Surface Plasmon Resonance.* Reproduce the results and graphs of Figures 8.5.3–8.5.7.

8.4 Working with the electric and magnetic fields across an negative-index slab given by Eqs. (8.6.1) and (8.6.2), derive the reflection and transmission responses of the slab given in (8.6.8).

8.5 *Computer Experiment—Perfect Lens.* Study the sensitivity of the perfect lens property to the deviations from the ideal values of $\epsilon = -\epsilon_0$ and $\mu = -\mu_0$, and to the presence of losses by reproducing the results and graphs of Figures 8.6.3 and 8.6.4. You will need to implement the computational algorithm listed on page 329.

8.6 *Computer Experiment—Antireflection Coatings.* Reproduce the results and graphs of Figures 8.7.1–8.7.3.

8.7 *Computer Experiment—Omnidirectional Dielectric Mirrors.* Reproduce the results and graphs of Figures 8.8.2–8.8.10.

8.8 Derive the generalized Snel's laws given in Eq. (8.10.10). Moreover, derive the Brewster angle expressions given in Eqs. (8.11.4) and (8.11.5).

8.9 *Computer Experiment—Brewster angles.* Study the variety of possible Brewster angles and reproduce the results and graphs of Example 8.11.1.

8.10 *Computer Experiment—Multilayer Birefringent Structures.* Reproduce the results and graphs of Figures 8.13.1–8.13.2.

9

Waveguides

Waveguides are used to transfer electromagnetic power efficiently from one point in space to another. Some common guiding structures are shown in the figure below. These include the typical coaxial cable, the two-wire and mictrostrip transmission lines, hollow conducting waveguides, and optical fibers.

In practice, the choice of structure is dictated by: (a) the desired operating frequency band, (b) the amount of power to be transferred, and (c) the amount of transmission losses that can be tolerated.
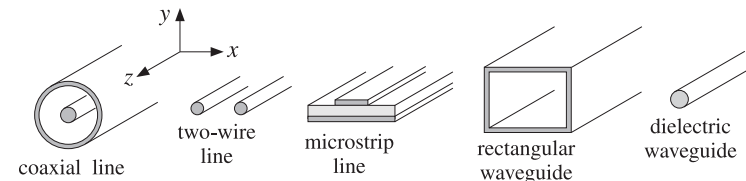

Fig. 9.0.1   Typical waveguiding structures.

Coaxial cables are widely used to connect RF components. Their operation is practical for frequencies below 3 GHz. Above that the losses are too excessive. For example, the attenuation might be 3 dB per 100 m at 100 MHz, but 10 dB/100 m at 1 GHz, and 50 dB/100 m at 10 GHz. Their power rating is typically of the order of one kilowatt at 100 MHz, but only 200 W at 2 GHz, being limited primarily because of the heating of the coaxial conductors and of the dielectric between the conductors (dielectric voltage breakdown is usually a secondary factor.) However, special short-length coaxial cables do exist that operate in the 40 GHz range.

Another issue is the single-mode operation of the line. At higher frequencies, in order to prevent higher modes from being launched, the diameters of the coaxial conductors must be reduced, diminishing the amount of power that can be transmitted.

Two-wire lines are not used at microwave frequencies because they are not shielded and can radiate. One typical use is for connecting indoor antennas to TV sets. Microstrip lines are used widely in microwave integrated circuits.