25. Lecture, 30 November 1999

25.1 Interference filters

Despite the wide variety of bulk absorption and reflection bands available in solids, it is usually necessary also to use interference filters in order to obtain filter transmission profiles as desired, with cut-on and cut-off edges sharp and precisely placed. As the name implies, this style of filter involves interference between light reflected partially at different places in the beam. Partial reflection at visible through mid-infrared wavelengths is conveniently provided by dielectric surfaces, or more commonly by stacks of thin ($<\lambda$) dielectric layers with different refractive index, as illustrated in Figure 25.1. Dielectric reflection can be independent of wavelength over wide bands, so the path-length differences in the stack, and resulting interference, determine such a filter's properties. It is straightforward to fabricate dielectric layers with this range of thickness, and the materials used in the films can usually be chosen so that negligible absorption is introduced; by contrast, a very thin, partially-transparent metal film can absorb substantially at these wavelengths and prevent good passband transmission. We will discuss this important case in detail in the following section.

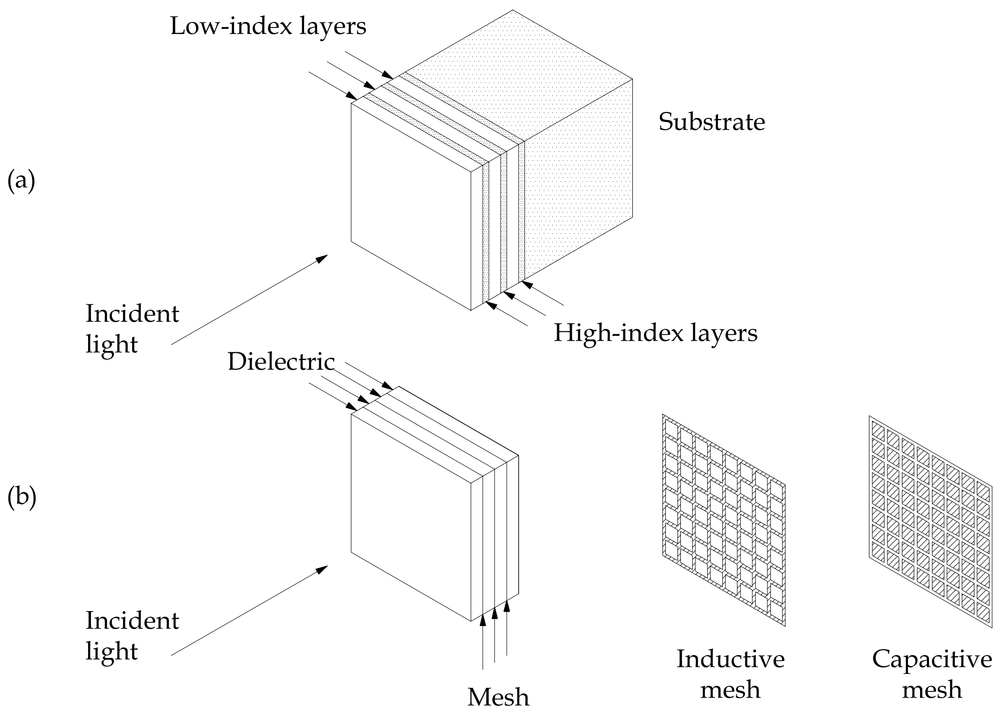

Figure 25.1: schematic diagram of multilayer interference filters (not drawn to scale): (a) alternating high- and low-index dielectric layers; (b) a series of metal meshes, spaced by low-index dielectric layers. Depending upon desired transmission characteristics, the mesh pattern would generally be either *inductive* or *capacitive*, illustrated at right.

At longer wavelengths the bulk absorption of light in metal films can be avoided by use of metal *mesh*, in which a thin film of metal (often free standing) is fabricated in the form of a grid pattern, and in which the smallest dimensions on which the grid is periodic is a small fraction of the wavelength of light. Such reflectors can be described accurately as regular arrays of waveguides, used near or below the cutoff wavelength of the lowest-order modes. The reflectance of such a mesh is of course not wavelength-independent, but like dielectric films, they can be arranged in a series as shown in Figure 25.1 to produce an interference filter. At the longest wavelengths used in astronomy, for which light is conducted to detectors by waveguides, the reflection provided by waveguide transitions can also be used in the same manner as the basis of interference filters. The theoretical details of metal-mesh reflectors, waveguides and associated interference filters are considerably more complicated in detail than those of dielectric layers, and will be considered to lie beyond the scope of our course.

25.2 Transmission through a plane-parallel dielectric slab

As a first step in the consideration of interference filters it is useful to consider the simplest one, consisting of two plane-parallel, partially reflecting surfaces separated by a dielectric medium with refractive index n and thickness d (Figure 25.2). It doesn't matter what the index of refraction outside the reflectors is, but we will assume here that it is unity on both sides, to be concrete. Let light be incident at angle θ_i. At every encounter with a reflecting surface the light is partially reflected and partially transmitted, leading to the distribution of displaced, parallel reflected and transmitted rays shown in Figure 25.2. The paths traversed by transmitted or reflected rays that have taken different numbers of internal reflections are of course different, and the waves they represent are generally out of phase. If the transmitted or reflected rays are focussed then the waves interfere.

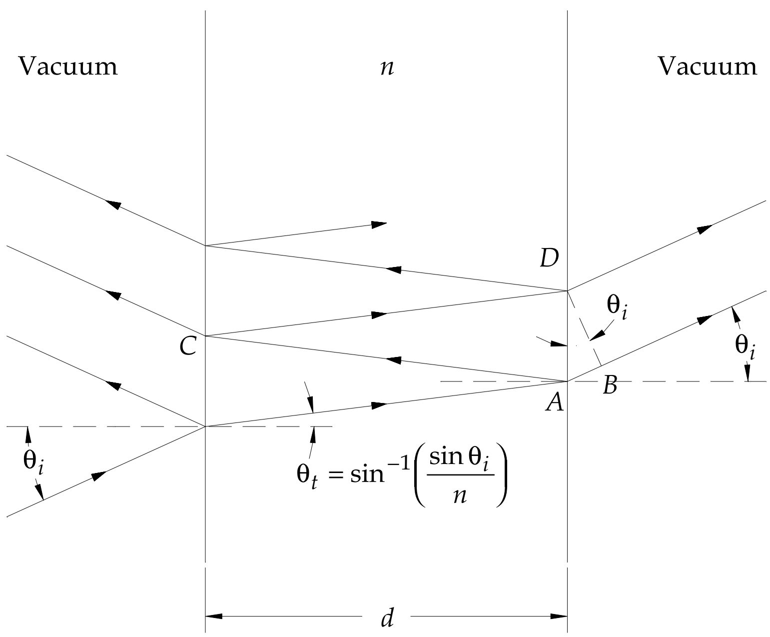

Figure 25.2: transmission through a plane-parallel dielectric slab.

It is easy to see that the path length difference between any two successive transmitted rays is the same; let us calculate it for the first two. The path these rays follow to the detector is the same before point A, and after line BD, in Figure 25.2, so the path length difference is that between AB and ACD. First note that

$$AD = 2d\tan\theta_t \quad , \tag{25.1}$$

where $n\sin\theta_t = \sin\theta_i$; then

$$AB = 2d\tan\theta_t \sin\theta_i = 2dn\frac{\sin^2\theta_t}{\cos\theta_t} \quad ,$$
$$ACD = \frac{2d}{\cos\theta_t} \quad . \tag{25.2}$$

The light has wavelength λ in vacuum and λ / n in the medium between the reflectors, so the phase lags from these paths are

$$\phi(AB) = 2\pi\frac{AB}{\lambda} = \frac{4\pi dn}{\lambda\cos\theta_t}\sin^2\theta_t \quad ,$$
$$\phi(ACD) = 2\pi\frac{nACD}{\lambda} = \frac{4\pi dn}{\lambda\cos\theta_t} \quad . \tag{25.3}$$

If the phase difference is an integer multiple of 2π, then the interference between the two wavefronts corresponding to these paths is completely constructive:

$$\Delta\phi = \phi(ACD) - \phi(AB) = \frac{4\pi dn}{\lambda\cos\theta_t}\left(1-\sin^2\theta_t\right) = \frac{4\pi dn\cos\theta_t}{\lambda} \quad ,$$
$$= 2\pi m \qquad (m = 0,1,2,\ldots), \tag{25.4}$$

and thus there are maxima in the spectrum of the transmission of the dielectric slab, at wavelengths given by

$$\lambda_m = \frac{2dn\cos\theta_t}{m} \qquad (m = 0,1,2,\ldots), \tag{25.5}$$

The wavelength, at which a peak, or *order*, that corresponds to a given value of *m* occurs, can be adjusted by changing *d*, *n*, or the angle of incidence. Near normal incidence, there is little "walk-off" or broadening of a beam from its multiple internal reflections, as in Figure 25.2, so it is usually the former of the two parameters that are varied. We will see below that if the reflecting surfaces are *highly* reflecting, the transmission is quite small except near the orders, at which it approaches unity. "Isolation" of a single order, through the rejection of all the other orders with additional filters, results in a useful, potentially very narrow-band (high-resolution) filter. In this application our simple device is usually known as a *Fabry-Perot interferometer.*

25.3 Reflection and transmission for dielectric multilayers

In E&M classes one usually learns to derive the coefficients of amplitude reflection and transmission from dielectrics – the Fresnel equations – by application of the boundary conditions for dielectric surfaces that arise from Maxwell's equations, to incident plane waves. It turns out that a simple rearrangement of these same boundary conditions, applied to one dielectric layer and its surfaces, gives rise to a matrix formulation for the reflection and transmission coefficients that can be extended to arbitrary numbers of plane-parallel dielectric layers, and is thus a useful tool in the design of interference filters.

Consider a plane electromagnetic wave incident obliquely from a dielectric medium with refractive index $n_0 = \sqrt{\varepsilon_0\mu_0}$, on a dielectric layer of thickness *d* with index $n_1 = \sqrt{\varepsilon_1\mu_1}$, as shown in Figure 25.3. Beyond this layer is a medium with index $n_2 = \sqrt{\varepsilon_2\mu_2}$. (ε and μ are as usual the dielectric constant and permeability). For simplicity we consider one linear polarization at a time, starting with the electric field amplitude perpendicular to the plane of incidence (the so-called *TE*-wave configuration). At each surface part of the incident wave will be reflected and part transmitted, so that each medium except the last contains waves propagating left and right. The electric fields of the incident and transmitted waves, which travel generally to the right in Figure 25.3, are of the form

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{i(n\boldsymbol{\kappa}\cdot\boldsymbol{r}-\omega t)} \quad , \tag{25.6}$$

while the leftward-bound reflected waves are described generically by

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{i(-n\boldsymbol{\kappa}\cdot\boldsymbol{r}-\omega t)} \quad , \tag{25.7}$$

where $\kappa = 1/\lambda$ is the wavenumber in vacuum. The components of $\boldsymbol{E}$ and $\boldsymbol{H} = \boldsymbol{B}/\mu$ parallel to each dielectric interface must be continuous across that interface. For interface 1 in Figure 25.3, in which lies the origin of the coordinate system and the phase reference, the tangential field components are

$$\begin{aligned} E_1 &= E_I + E_{R1} = E_{T1} + E_{R2} \\ H_1 &= \frac{1}{\mu_0}\left(B_I \cos\theta_I - B_{R1}\cos\theta_I\right) = \frac{1}{\mu_1}\left(B_{T1}\cos\theta_{T1} - B_{R2}\cos\theta_{T1}\right) \quad , \end{aligned} \tag{25.8}$$

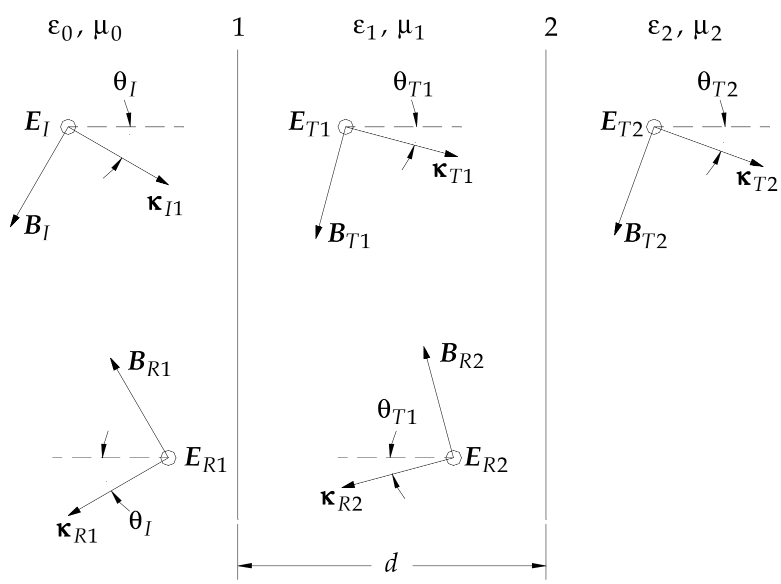

Figure 25.3: diagram for transmission and reflection of light incident upon a uniform dielectric slab between two uniform dielectric media, with the electric field polarized perpendicular to the plane of incidence (*TE* waves), pointing out of the page.

where the factors of $e^{-i\omega t}$ have been cancelled. Using $\boldsymbol{B} = \sqrt{\mu\varepsilon}\,\hat{\boldsymbol{\kappa}} \times \boldsymbol{E}$,

$$\begin{aligned} E_1 &= E_I + E_{R1} = E_{T1} + E_{R2} \\ H_1 &= \sqrt{\frac{\varepsilon_0}{\mu_0}}\cos\theta_I\left(E_I - E_{R1}\right) = \sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1}\left(E_{T1} - E_{R2}\right) \quad , \end{aligned} \tag{25.9}$$

As a wave crosses the slab it travels a distance $PQ = d/\cos\theta_{T1}$ (see Figure 25.4), and, with respect to the undisplaced wave that would have resulted if the slab were not there, undergoes a phase change of

$$\begin{aligned}\kappa\ell &= \frac{2\pi n_1}{\lambda}PQ - \frac{2\pi n_2}{\lambda}RS = \frac{2\pi n_1 d}{\lambda\cos\theta_{T1}} - \frac{2\pi n_2}{\lambda}d\tan\theta_{T1}\sin\theta_{T2} \\ &= \frac{2\pi n_1 d}{\lambda\cos\theta_{T1}}\left(1-\sin^2\theta_{T1}\right) = \frac{2\pi n_1 d}{\lambda}\cos\theta_{T1} \quad ,\end{aligned} \tag{25.10}$$

which is, sensibly, half the phase difference in Equation 25.4. Thus the tangential field components at interface 2 are

$$\begin{aligned}E_2 &= E_{T1}e^{i\kappa\ell} + E_{R2}e^{-i\kappa\ell} = E_{T2}e^{i\kappa\ell} \\ H_2 &= \sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1}\left(E_{T1}e^{i\kappa\ell} - E_{R2}e^{-i\kappa\ell}\right) = \sqrt{\frac{\varepsilon_2}{\mu_2}}\cos\theta_{T2}E_{T2}e^{i\kappa\ell} \quad .\end{aligned} \tag{25.11}$$

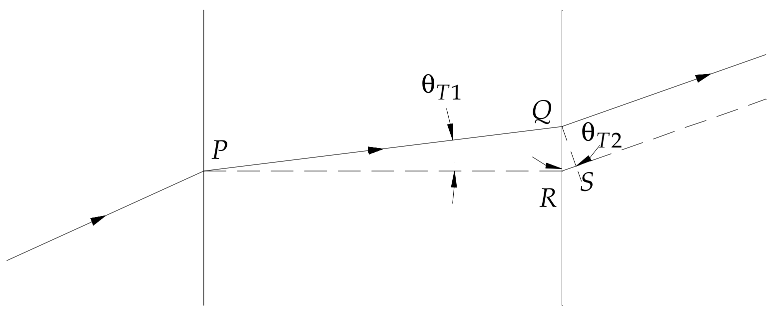

Figure 25.4: setup for calculation of phase lag from traversal of dielectric slab.

If we solve this equations for E_{T1} and E_{R2} in terms of E_2 and H_2, and then substitute the results into Equation 25.9, then we will obtain a relation among the fields at the two successive interfaces. Multiply the first expression by $\sqrt{\varepsilon_1/\mu_1}\cos\theta_{T1}$ and add, to obtain

$$\begin{aligned}&\sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1}E_2 + H_2 = 2\sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1}E_{T1}e^{i\kappa\ell} \quad \text{, or} \\ &E_{T1} = \frac{1}{2}e^{-i\kappa\ell}\left(E_2 + \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right) \quad ,\end{aligned} \tag{25.12}$$

and then replace in one of the original expressions to solve for E_{R2}:

$$\begin{aligned}&E_2 = \frac{1}{2}e^{-i\kappa\ell}\left(E_2 + \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right)e^{i\kappa\ell} + E_{R2}e^{-i\kappa\ell} \quad \text{, or} \\ &E_{R2} = \frac{1}{2}e^{i\kappa\ell}\left(E_2 - \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right) \quad .\end{aligned} \tag{25.13}$$

Now put these last two results into Equation 25.9:

$$\begin{aligned} E_1 &= \frac{1}{2}e^{-i\kappa\ell}\left(E_2 + \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right) + \frac{1}{2}e^{i\kappa\ell}\left(E_2 - \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right) \\ &= E_2\cos\kappa\ell - H_2\sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{i\sin\kappa\ell}{\cos\theta_{T1}} \quad , \end{aligned} \tag{25.14}$$

and

$$\begin{aligned} H_1 &= \sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1}\left[\frac{1}{2}e^{-i\kappa\ell}\left(E_2 + \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right) - \frac{1}{2}e^{i\kappa\ell}\left(E_2 - \sqrt{\frac{\mu_1}{\varepsilon_1}}\frac{H_2}{\cos\theta_{T1}}\right)\right] \\ &= -E_2\sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1} i\sin\kappa\ell + H_2\cos\kappa\ell \quad . \end{aligned} \tag{25.15}$$

We have set these equations up for *TE* waves, with ***E*** polarized perpendicular to the plane of incidence. However, Maxwell's equations without sources (i.e. zero charge and current density) are invariant if ***E*** and ***H***, and simultaneously ε and -μ, are switched. This transformation also turns our *TE* wave into a *TM* wave, for which ***E*** is polarized parallel to the plane of incidence. Thus the equations for *TM* waves, corresponding to Equations 25.14 and 25.15 above, are

$$H_1 = H_2\cos\kappa\ell - E_2\sqrt{\frac{\varepsilon_1}{\mu_1}}\frac{i\sin\kappa\ell}{\cos\theta_{T1}} \tag{25.16}$$

and

$$E_1 = -H_2\sqrt{\frac{\mu_1}{\varepsilon_1}}\cos\theta_{T1} i\sin\kappa\ell + E_2\cos\kappa\ell \quad . \tag{25.17}$$

Either way, these pairs of equations can each be written as a matrix equation:

$$\begin{aligned} \begin{bmatrix} E_1 \\ H_1 \end{bmatrix} &= \begin{bmatrix} \cos\kappa\ell & -i\sin\kappa\ell / Y_1 \\ -iY_1\sin\kappa\ell & \cos\kappa\ell \end{bmatrix}\begin{bmatrix} E_2 \\ H_2 \end{bmatrix} \\ &= M_1\begin{bmatrix} E_2 \\ H_2 \end{bmatrix} \quad , \end{aligned} \tag{25.18}$$

where for *TE* waves we have

$$Y_{1,TE} = \sqrt{\frac{\varepsilon_1}{\mu_1}}\cos\theta_{T1} \quad , \tag{25.19}$$

and, for *TM* waves,

$$Y_{1,TM} = \sqrt{\frac{\varepsilon_1}{\mu_1}}\frac{1}{\cos\theta_{T1}} \quad . \tag{25.20}$$

These are the same for normal incidence, as they must be. The matrix M_1 is called the *characteristic matrix* of the dielectric layer, and it relates the tangential components of the fields at its surfaces. If there were yet a third surface to the right in Figure 25.3, the fields there could therefore be determined from

$$\begin{bmatrix} E_2 \\ H_2 \end{bmatrix} = M_2 \begin{bmatrix} E_3 \\ H_3 \end{bmatrix} \quad , \tag{25.21}$$

which can be combined with Equation 25.18 to produce

$$\begin{bmatrix} E_1 \\ H_1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} E_3 \\ H_3 \end{bmatrix} \quad , \tag{25.22}$$

and so forth. Evidently, if there are p layers in the stack, then the fields at the first and last surfaces are related by

$$\begin{bmatrix} E_1 \\ H_1 \end{bmatrix} = M_1 M_2 \cdots M_p \begin{bmatrix} E_{p+1} \\ H_{p+1} \end{bmatrix} \quad , \tag{25.23}$$

and the whole stack can be said to have a characteristic matrix M given by

$$M = M_1 M_2 \cdots M_p = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \quad . \tag{25.24}$$

Now return to the original boundary conditions, Equations 25.9 and 25.11. Writing the fields at the surfaces in terms of incident, reflected and transmitted components, Equation 25.23 becomes

$$\begin{bmatrix} E_I + E_{R1} \\ Y_0 \left(E_I - E_{R1} \right) \end{bmatrix} = M \begin{bmatrix} E_{T,p+1} \\ Y_{p+1} E_{T,p+1} \end{bmatrix} \quad . \tag{25.25}$$

Multiply it out:

$$\begin{aligned} E_I + E_{R1} &= m_{11} E_{T,p+1} + m_{12} Y_{p+1} E_{T,p+1} \\ Y_0 \left(E_I - E_{R1} \right) &= m_{21} E_{T,p+1} + m_{22} Y_{p+1} E_{T,p+1} \quad . \end{aligned} \tag{25.26}$$

This pair of equations can be solved for the amplitude reflection and transmission coefficients, $r = E_{R1} / E_I$ and $t = E_{T,p+1} / E_I$:

$$\begin{aligned} 1 + r &= \left(m_{11} + m_{12} Y_{p+1} \right) t \\ Y_0 (1 - r) &= \left(m_{21} + m_{22} Y_{p+1} \right) t \quad . \end{aligned} \tag{25.27}$$

Multiply the first of these by Y_0 and add, to get

$$t = \frac{2Y_0}{m_{11} Y_0 + m_{12} Y_0 Y_{p+1} + m_{21} + m_{22} Y_{p+1}} \quad , \tag{25.28}$$

and substitute the result into either to get

$$r = \frac{m_{11}Y_0 + m_{12}Y_0Y_{p+1} - m_{21} - m_{22}Y_{p+1}}{m_{11}Y_0 + m_{12}Y_0Y_{p+1} + m_{21} + m_{22}Y_{p+1}} \quad . \tag{25.29}$$

Usually we will be most interested in the fraction of the *power* transmitted and reflected by the stack of layers. Power per unit area perpendicular to $\boldsymbol{\kappa}$ is expressed by the Poynting vector, which is

$$\boldsymbol{S} = \frac{c}{4\pi}\boldsymbol{E}\times\boldsymbol{H} = \frac{c}{4\pi}\sqrt{\frac{\varepsilon}{\mu}}\boldsymbol{E}\times\left(\hat{\kappa}\times\boldsymbol{E}\right) = \hat{\kappa}\frac{c}{4\pi}\sqrt{\frac{\varepsilon}{\mu}}|E|^2 \quad , \tag{25.30}$$

and the power per unit area crossing the interface is the time average of the component of $\boldsymbol{S}$ normal to the surface, or

$$\overline{S_n} = \frac{c}{8\pi}\sqrt{\frac{\varepsilon}{\mu}}\cos\theta|E|^2 = \frac{c}{8\pi}Y|E|^2 \quad , \tag{25.31}$$

where θ is the angle of the wavevector $\boldsymbol{\kappa}$ with respect to the interface normal. The fraction of the power per unit area transmitted through the interface, or *power transmission*, is therefore

$$\tau = \frac{\overline{S_{T,p+1,n}}}{\overline{S_{I,n}}} = \frac{Y_{p+1}}{Y_I}|t|^2 \quad , \tag{25.32}$$

and the fraction of the power per unit area reflected by the interface, or *power reflectance*, is

$$\rho = \frac{\overline{S_{R1,n}}}{S_{I,n}} = |r|^2 \quad . \tag{25.33}$$

In this week's homework you will show that $\tau + \rho = 1$, as energy conservation would imply. These relations, along with Equation 25.23, comprise a powerful method for the design of interference filters.

