

2

BOUNDARIES, NEAR-FIELD OPTICS, AND NEAR-FIELD IMAGING

When light travels from one medium into another, reflection and refraction usually take place at the boundary. A portion of the light is reflected back into the first medium, and the other portion is transmitted (refracted) into the second medium at the angle set by Snell's law. The laws of reflection and refraction are among the most basic principles of optics.

Optical systems are made up of boundaries. One cannot have a component without boundaries. A clear understanding of the boundary phenomena is most important, particularly in the field of integrated optics, where the boundaries are so close together that interactions become a serious problem. What will be treated here are simple configurations, but an in-depth understanding of these configurations will provide a solid foundation for solving more complex real-life problems, such as the photon tunneling microscope, the scanning near-field optical microscope (SNOM), and the high-density digital video disk (DVD).

2.1 BOUNDARY CONDITIONS

We will consider the simplest case when light is incident from medium 1 onto its border with medium 2, as indicated in Fig. 2.1. Let the indices of refraction of medium 1 and 2 be n_1 and n_2 , respectively. On the boundary, there are five important conditions:

1. The frequency of light does not change across the border, unless one of the media happens to be a nonlinear medium. In nonlinear media higher harmonics are generated.
2. The wavelength either expands or contracts, according to the ratio of the indices of refraction:

$$\frac{n_1}{n_2} = \frac{\lambda_2}{\lambda_1} \quad (2.1)$$

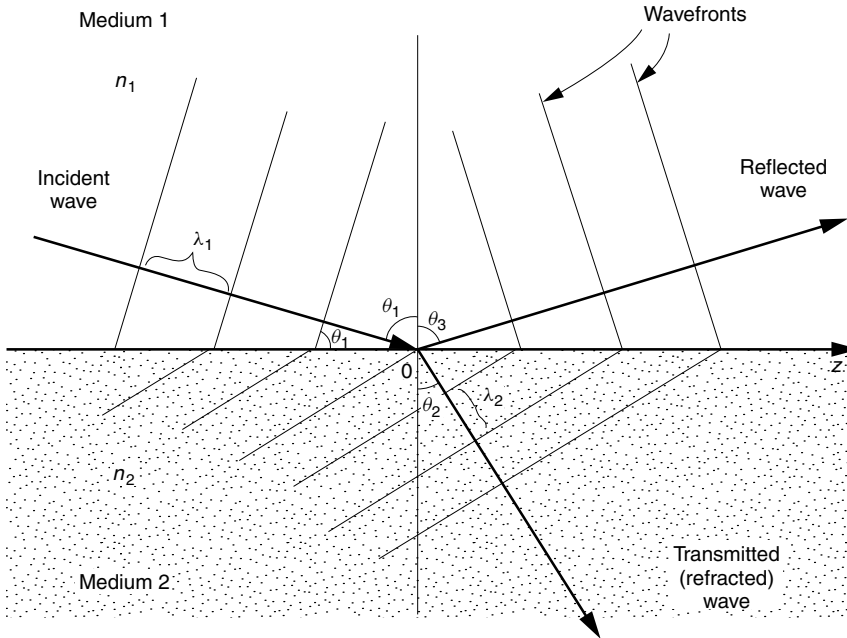


Figure 2.1 Wavefront and ray directions near the boundary.

The wavelength expands when light goes into a medium with a smaller index of refraction. A medium with a smaller index of refraction is often called an optically less dense medium.

3. Speaking in terms of electromagnetic theory, the ratio of light energy, contained in the form of electric energy, to that in the form of magnetic energy is changed at the boundary depending on the intrinsic impedance of the medium. The intrinsic impedance η_1 of medium 1 is

$$\frac{|\mathbf{E}_1|}{|\mathbf{H}_1|} = \frac{|\mathbf{E}_3|}{|\mathbf{H}_3|} = \eta_1 \quad (2.2)$$

$$\eta_1 = \sqrt{\frac{\mu_0 \mu_{r1}}{\epsilon_0 \epsilon_{r1}}} = \frac{\eta_0}{n_1}$$

where \mathbf{E}_1 and \mathbf{H}_1 are the electric and magnetic fields of the incident wave, and \mathbf{E}_3 and \mathbf{H}_3 are the electric and magnetic fields of the reflected wave. Both waves are in medium 1. The intrinsic impedance η_0 of the vacuum is

$$\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \quad (2.3)$$

and its value in MKS units is $120\pi\Omega$. In Eq. (2.2), μ_0 represents the magnetic permeability of free space, and μ_{r1} is the relative magnetic permeability of medium 1. With most of the materials treated here, the value of μ_r is unity, and we will treat it as such unless stated otherwise. In MKS units, $\mu_0 = 1.2566 \times 10^{-6}$ henry/meter. (MKS units will be used throughout the book.) In Eq. (2.3),

the quantities ϵ_0 and ϵ_r are the absolute and relative dielectric constants. The value of the former is $\epsilon_0 = 8.855 \times 10^{-12}$ farad/meter, and that of the latter varies from medium to medium. In optics the index of refraction n is used more often than ϵ_r .

The index of refraction n is defined as the ratio of the wavelength in the medium to that in vacuum. The wavelength λ in a nonmagnetic medium is

$$\lambda = v/f \quad \text{where } v = 1/\sqrt{\mu_0\epsilon_0\epsilon_r}$$

and that in the vacuum is

$$\lambda_0 = c/f \quad \text{where } c = 1/\sqrt{\mu_0\epsilon_0}$$

hence,

$$n = \lambda_0/\lambda = \sqrt{\epsilon_r}.$$

With media 1 and 2,

$$\begin{aligned} \sqrt{\epsilon_{r1}} &= n_1 \\ \sqrt{\epsilon_{r2}} &= n_2 \end{aligned} \quad (2.4)$$

Similarly, for medium 2, the intrinsic impedance η_2 is

$$\frac{|\mathbf{E}_2|}{|\mathbf{H}_2|} = \eta_2 \quad \eta_2 = \sqrt{\frac{\mu_0\mu_{r2}}{\epsilon_0\epsilon_{r2}}} = \frac{\eta_0}{n_2} \quad (2.5)$$

The ratio of the electric to magnetic field has to be changed as soon as the light crosses the boundary.

The manner in which the transition takes place is governed by Maxwell's boundary conditions. These boundary conditions represent the fourth and fifth items in our list of important conditions, and are stated below.

4. The tangential components of \mathbf{E} and \mathbf{H} are continuous across the boundary.
5. The normal components of \mathbf{D} and \mathbf{B} are continuous, where \mathbf{D} and \mathbf{B} are the electric and magnetic flux densities.

The \mathbf{E} field is often related to the voltage V and the \mathbf{H} field to the current I , and the ratio \mathbf{E}/\mathbf{H} is often related to the impedance V/I , even though no tangible medium is present. It is more appropriate to say that the intrinsic impedance is defined by the ratio of \mathbf{E} to \mathbf{H} of a plane wave in that medium.

The angles and amplitudes of the reflected and transmitted light are governed by the above five conditions; three by frequency, wavelength and impedance conditions, and two by Maxwell's continuity conditions. In the following, using these conditions, quantities associated with the boundary are found.

2.2 SNELL'S LAW

The relationship between the angles of incidence, reflection, and transmission (refraction) are found. If the incident light is sinusoidally varying in both time and space, then

both the reflected and the transmitted waves have to vary sinusoidally accordingly. To understand this fact, suppose that at a particular instance and at a particular location of the boundary, the oscillation of the incident wave is at its maximum; then both reflected and transmitted waves have to be at their maxima. Otherwise, even though the boundary conditions can be met at a specific instance or at a specific location, they cannot be met throughout time and the entire boundary surface. In other words, in order to meet Maxwell's continuity boundary condition, the wavelengths along the interface surface must have the same temporal and spatial variation,

$$\lambda_{z1} = \lambda_{z2} = \lambda_{z3} \quad (2.6)$$

where λ_{z1} , λ_{z2} , and λ_{z3} are the wavelengths measured in the z direction in the plane of the incident, transmitted, and reflected waves. Instead of Eq. (2.6) in wavelengths, the propagation constant $\beta_{1,2,3} = 2\pi/\lambda_{z1,2,3}$ is often used to express this condition and Eq. (2.6) is rewritten as

$$\beta_1 = \beta_2 = \beta_3 \quad (2.7)$$

Equation (2.7) is known as either β , k , phase, or momentum matching, but all mean the same thing: wavelength matching.

The free-space wavelengths λ_1 , λ_2 , and λ_3 are determined solely by the refractive indices of the media, and at the moment $n_1 \neq n_2$ and $\lambda_1 \neq \lambda_2$, according to Condition 2. How then is the condition of wavelength matching to be satisfied? The only way to meet the wavelength matching requirement is for the directions of the transmitted and reflected waves to be bent, as shown in Fig. 2.1. By decreasing the incident angle θ_1 from 90° to 0° , the wavelength λ_z along the z direction increases from λ_1 to infinity. Then, the condition for the wavelength matching is

$$\frac{\lambda_1}{\sin \theta_1} = \frac{\lambda_1}{\sin \theta_3} = \frac{\lambda_2}{\sin \theta_2} \quad (2.8)$$

One immediate result from Eq. (2.8) is that the angle of reflection θ_3 is identical to the angle of incidence θ_1 :

$$\theta_1 = \theta_3 \quad (2.9)$$

Applying Eq. (2.1) to (2.8) gives

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (2.10)$$

which is Snell's law. Snell's law is one of the most used laws in optics.

Next, the relative amplitudes of the transmitted and reflected waves are considered.

2.3 TRANSMISSION AND REFLECTION COEFFICIENTS

The electric field transmission coefficient t_E is defined as the ratio of the field \mathbf{E}_2 of the transmitted wave to the field \mathbf{E}_1 of the incident wave. The electric field reflection coefficient r_E is defined as the ratio of the field \mathbf{E}_3 of the reflected wave to the field \mathbf{E}_1 of the incident wave. Analogous definitions t_H and r_H hold for the magnetic field \mathbf{H} .

The reflection and transmission coefficients depend on the angle of incidence and the directions of polarization. Let us begin with the simplest case of normal incidence and then proceed to the general case of an arbitrary angle of incidence.

2.3.1 Transmission and Reflection Coefficients (at Normal Incidence)

When the direction of propagation is normal to the boundary (see Fig. 2.2), both the **E** and **H** fields are parallel to the interface. As shown in Fig. 2.2, the convention of positive **E** and **H** is chosen with regard to the Poynting vectors **s**₁, **s**₂, and **s**₃, which point down, down, and up, respectively. For a description [1] about the directions of the field, see Example 2.1. The Poynting vector **s** is the time rate of flow of electromagnetic energy per unit area (the power intensity), and **s** = **E** × **H**. Boundary Condition 4 is used, and

$$\mathbf{E}_1 + \mathbf{E}_3 = \mathbf{E}_2 \tag{2.11}$$

$$\mathbf{H}_1 - \mathbf{H}_3 = \mathbf{H}_2 \tag{2.12}$$

From Eqs. (2.2) and (2.5), the above equations are written as

$$\eta_1 \mathbf{H}_1 + \eta_1 \mathbf{H}_3 = \eta_2 \mathbf{H}_2 \tag{2.13}$$

$$\frac{\mathbf{E}_1}{\eta_1} - \frac{\mathbf{E}_3}{\eta_1} = \frac{\mathbf{E}_2}{\eta_2} \tag{2.14}$$

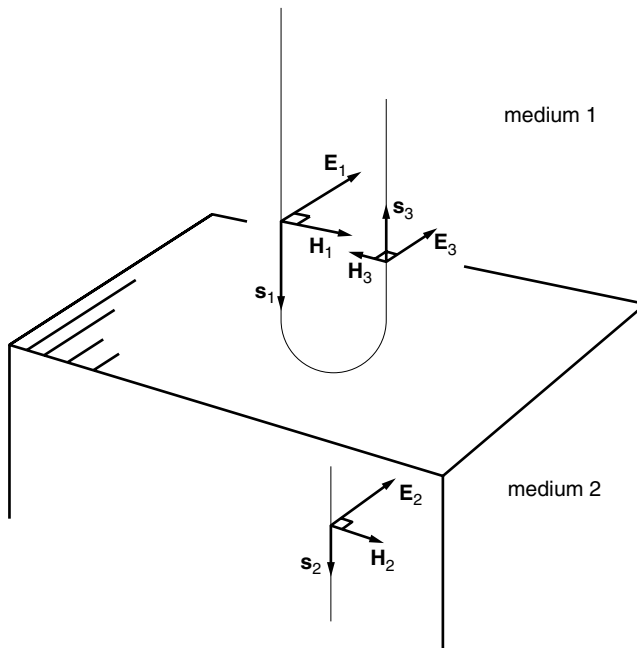


Figure 2.2 Transmission and reflection coefficients at an interface between two media at normal incidence.

From Eqs. (2.11) to (2.14), all of the following expressions for r_E , r_H , t_E , and t_H are derived:

$$r_E = r_H = \frac{E_3}{E_1} = \frac{H_3}{H_1} = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1} \quad (2.15)$$

$$t_E = \frac{E_2}{E_1} = \frac{2\eta_2}{\eta_2 + \eta_1} \quad (2.16)$$

$$t_H = \frac{H_2}{H_1} = \frac{2\eta_1}{\eta_2 + \eta_1} \quad (2.17)$$

Note that t_E and t_H are not equal.

In optics, the index of refraction n rather than the intrinsic impedance η is used. With the assumption that $\mu_{r1} = \mu_{r2}$, Eqs. (2.15) to (2.17) can be rewritten as

$$r_E = r_H = \frac{n_1 - n_2}{n_1 + n_2} \quad (2.18)$$

$$t_E = \frac{2n_1}{n_1 + n_2} \quad (2.19)$$

$$t_H = \frac{2n_2}{n_1 + n_2} \quad (2.20)$$

Example 2.1 Let medium 1 be air and medium 2 be glass. The index of refraction of glass is $n_2 = 1.5$ and of air is $n_1 = 1$. Calculate the coefficients of reflection and transmission associated with the air-glass interface for normal incidence.

Solution With $n_1 = 1.0$, $n_2 = 1.5$, Eqs. (2.18) to (2.20) become

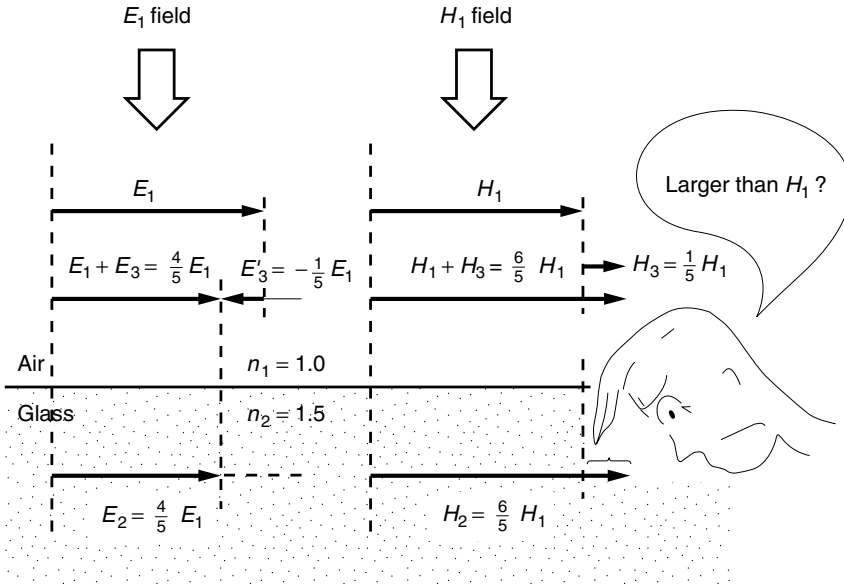
$$r_E = r_H = -\frac{1}{5} \quad (2.21)$$

$$t_E = \frac{4}{5} \quad (2.22)$$

$$t_H = \frac{6}{5} \quad (2.23)$$

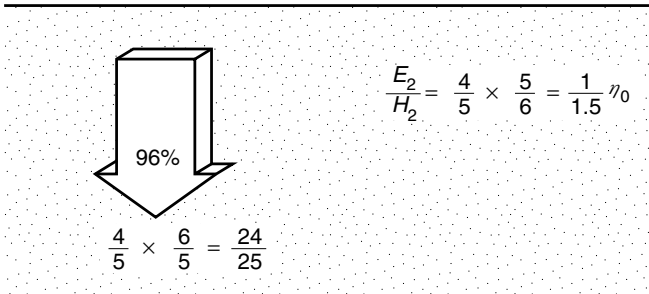
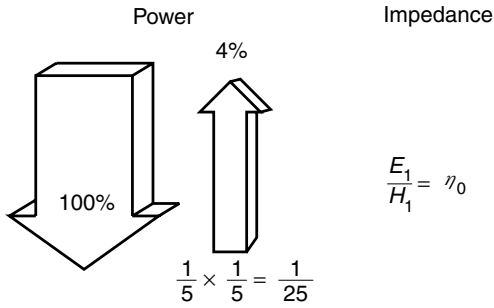
Everything looks all right except Eq. (2.23). Can the transmission coefficient be larger than 1? Let us examine this more closely [2].

Note from Eq. (2.18) when $n_2 > n_1$, the reflection coefficient becomes negative. What does a negative value for the reflection coefficient mean? Reexamine Fig. 2.2 closely. The positive directions of \mathbf{E} and \mathbf{H} were chosen such that the positive directions of E_1 and E_3 are the same but those of H_1 and H_3 are opposite; hence, when $r_H = -\frac{1}{5}$ it really means H_3 points in the same direction as H_1 and its magnitude is $\frac{1}{5}|\mathbf{H}_1|$, as illustrated in Fig. 2.3a. Likewise, since r_E is also negative according to Eq. (2.21), it really means that the direction of \mathbf{E}_3 is opposite to \mathbf{E}_1 and its magnitude is $\frac{1}{5}|\mathbf{E}_1|$, as illustrated in Fig. 2.3a. The resultant \mathbf{E} field just above the boundary is $|\mathbf{E}_1 + \mathbf{E}_3| = \frac{4}{5}|\mathbf{E}_1|$ and that just below the boundary is $|\mathbf{E}_2| = \frac{4}{5}|\mathbf{E}_1|$ so that the boundary condition for \mathbf{E} is satisfied. Similarly, the resultant \mathbf{H} field just above the interface is $\frac{6}{5}|\mathbf{H}_1|$ and that just below is $\frac{6}{5}|\mathbf{H}_1|$ and both happily match up with the results of Eqs. (2.22) and



$$\eta_2 = \frac{\eta_0}{n_2} = \frac{E_2}{H_2} = \frac{\eta_0}{1.5}$$

(a)



(b)

Figure 2.3 Field and power redistributions and impedance at an air–glass interface: $n_1 = 1, n_2 = 1.5$. (a) Field redistribution. (b) Power redistribution and impedance.

(2.23). The intrinsic impedance in the glass is

$$\eta_2 = \frac{\eta_0}{n_2} = \frac{E_2}{H_2} = \frac{\frac{4}{5}E_1}{\frac{6}{5}H_1} = \frac{\eta_0}{1.5}$$

From this example, the following generalization can be made. When light is incident from a medium of a higher intrinsic impedance (lower index of refraction) onto that of a lower intrinsic impedance (higher index of refraction), the transmitted \mathbf{E} field decreases, whereas the transmitted \mathbf{H} field increases as seen from Eqs. (2.16) and (2.17). Whenever $n_2 > n_1$, t_H in Eq. (2.20) is larger than unity! \square

The transmittance T is defined as the power ratio of transmitted light to incident light. The reflectance R is the power ratio of reflected light to incident light. The value of T and R will be calculated for Example 2.1. (Note that transmission and reflection coefficients are field ratios.)

Let P denote light power, \mathbf{s} denote the Poynting vector, and A denote the area on the interface intercepted by the light beams. For simplicity, A will be taken as a unit area. For the present case, $P = |\mathbf{s}|A = |\mathbf{s}|$.

The incident light power is therefore

$$|\mathbf{s}_1| = |\mathbf{E}_1 \times \mathbf{H}_1| = \frac{1}{\eta_1} |\mathbf{E}_1|^2 \quad (2.24)$$

The reflected power is

$$|\mathbf{s}_3| = |\mathbf{E}_3 \times \mathbf{H}_3| = \frac{1}{\eta_1} |\mathbf{E}_3|^2 = \frac{1}{\eta_1} |r_E \mathbf{E}_1|^2 \quad (2.25)$$

Hence, the reflectance is

$$R = \frac{|\mathbf{s}_3|}{|\mathbf{s}_1|} = |r_E|^2 = \frac{1}{25} \quad (2.26)$$

The transmitted power is

$$|\mathbf{s}_2| = |\mathbf{E}_2 \times \mathbf{H}_2| = \frac{1}{\eta_2} |\mathbf{E}_2|^2 = \frac{1}{\eta_2} |t_E \mathbf{E}_1|^2 \quad (2.27)$$

Hence, the transmittance T is

$$T = \frac{(1/\eta_2)|t_E \mathbf{E}_1|^2}{(1/\eta_1)|\mathbf{E}_1|^2} = \frac{\eta_1}{\eta_2} |t_E|^2 = \frac{1.5}{1} \left(\frac{4}{5}\right)^2 = \frac{24}{25} \quad (2.28)$$

In terms of the \mathbf{H} field, one has

$$|\mathbf{s}_1| = \eta_1 |\mathbf{H}_1|^2, \quad |\mathbf{s}_2| = \eta_2 |t_H \mathbf{H}_1|^2, \quad |\mathbf{s}_3| = \eta_1 |r_H \mathbf{H}_1|^2$$

$$R = \frac{|\mathbf{s}_3|}{|\mathbf{s}_1|} = |r_H|^2 = \frac{1}{25} \quad (2.29)$$

$$T = \frac{|\mathbf{s}_2|}{|\mathbf{s}_1|} \frac{\eta_2}{\eta_1} |t_H|^2 = \frac{1}{1.5} \left(\frac{6}{5}\right)^2 = \frac{24}{25} \quad (2.30)$$

which is consistent with the values calculated using the \mathbf{E} field.

2.3.2 Transmission and Reflection Coefficients (at an Arbitrary Incident Angle)

The values of the transmission and reflection coefficients depend not only on the angle of incidence, but also on the direction of polarization, which is the direction of the \mathbf{E} field (some books define polarization in terms of \mathbf{H}).

The propagation direction of the incident light and the normal to the interface define a plane, which is referred to as the plane of incidence. Two polarizations of particular interest are the \mathbf{E} field polarized in the plane of incidence, and \mathbf{E} polarized perpendicular to the plane of incidence.* An incident wave with an arbitrary direction of polarization can always be decomposed into these two polarizations.

The in-plane polarization, more commonly called parallel polarization, or p wave, is considered first. In Fig. 2.4a, the $y = 0$ plane is the plane of incidence. The convention of the positive directions of \mathbf{E} and \mathbf{H} are taken such that the Poynting vector $\mathbf{s} = \mathbf{E} \times \mathbf{H}$ coincides with the direction of light propagation (see boxed note). The continuity Conditions 4 and 5 are used. In medium 1, components of both incident and reflected waves must be considered. The tangential components of the \mathbf{E} and \mathbf{H} fields on both

There is more than one combination of \mathbf{E} and \mathbf{H} that satisfies this convention. In order to maintain consistency in the derivation of equations, henceforth in this text, we shall impose the following additional constraint on the choice of positive \mathbf{E} and \mathbf{H} . For both parallel and perpendicular polarizations, in the limit that the angle of incidence approaches 90° , all \mathbf{E} vectors shall point in the same direction. (Picture Figs. 2.4a and 2.4b with θ_1 approaching grazing incidence.) This is convenient when dealing with optical guides, where the angles of incidence at boundaries are often large. However, this constraint does lead to a peculiarity at normal incidence. For parallel polarization at normal incidence, incident and reflected \mathbf{E} vectors point in opposite directions, whereas for perpendicular polarization, incident and reflected \mathbf{E} vectors point in the same direction, giving $r_\perp = -r_\parallel$ for $\theta_1 = 0$. For example, if you examine Eqs. (2.35) and (2.40) for $\theta_1 = \theta_2 = 0$, you will find $r_\parallel = -r_\perp$. The difference in sign is simply due to the difference in the conventions for positive \mathbf{E} for the two cases. In Fig. 2.4a, the direction of positive E_{3r} is taken in the opposite direction to that of E_{1r} in Fig. 2.4a, while the positive directions of E_1 and E_3 are taken the same. Whichever convention one chooses, the final result for a given configuration comes out the same, as explained in Example 2.1.

* Sometimes these waves are called p waves and s waves, originating from the German terms of parallel wave and senkrecht wave.

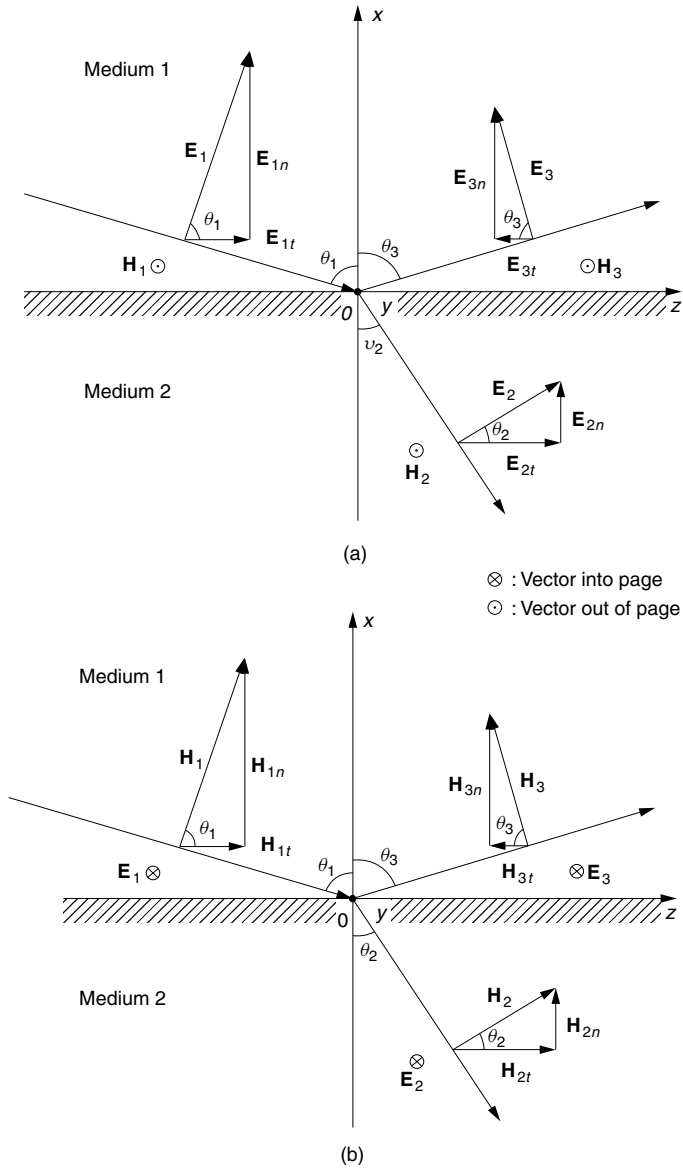


Figure 2.4 Boundary condition at the interface between two media. (a) Parallel polarization. (b) Perpendicular polarization.

sides of the boundary are equal; hence,

$$E_1 \cos \theta_1 - E_3 \cos \theta_3 = E_2 \cos \theta_2 \tag{2.31}$$

$$H_1 + H_3 = H_2 \tag{2.32}$$

Equation (2.32) can be rewritten further using Eqs. (2.2) and (2.5) as

$$\frac{1}{\eta_1}(E_1 + E_3) = \frac{1}{\eta_2}E_2 \tag{2.33}$$

From the continuity of the normal component of \mathbf{D} , and from the equality $\theta_1 = \theta_3$, one obtains

$$\epsilon_{r1}(E_1 + E_3) \sin \theta_1 = \epsilon_{r2}E_2 \sin \theta_2 \quad (2.34)$$

A combination of any two boundary conditions listed above can provide the electric field reflection (see boxed note) and transmission coefficients. The reflection coefficient r_{\parallel} for the \mathbf{E} field is

$$r_{\parallel} = \frac{E_3}{E_1} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_2 \cos \theta_1 + n_1 \cos \theta_2} \quad (2.35)$$

Similarly, the transmission coefficient t_{\parallel} for the \mathbf{E} field is

$$t_{\parallel} = \frac{E_2}{E_1} = \frac{2n_1 \cos \theta_1}{n_2 \cos \theta_1 + n_1 \cos \theta_2} \quad (2.36)$$

The subscript \parallel indicates the case when the \mathbf{E} field is in the plane of incidence, as shown in Fig. 2.4a. There is a difference between the values r_{\parallel} and t_{\parallel} for the \mathbf{E} field and the same quantities for the \mathbf{H} field. (See Problem 2.1.)

For the s wave case of light polarized perpendicular to the plane of incidence as shown in Fig. 2.4b, the reflection and transmission coefficients for the \mathbf{E} field are obtained in a similar manner. Boundary Condition 4 for the continuity of the tangential components of \mathbf{E} and \mathbf{H} is used.

$$\mathbf{E}_1 + \mathbf{E}_3 = \mathbf{E}_2 \quad (2.37)$$

$$H_1 \cos \theta_1 - H_3 \cos \theta_3 = H_2 \cos \theta_2 \quad (2.38)$$

Why does the wave reflect when a discontinuity in the index of refraction is encountered? Let's start with the case of reflection from a perfect conductor (mirror). As the electromagnetic wave is incident upon the conductor, a current is induced on the surface of the conductor, just as a current is induced on a metal wire receiving antenna. The induced current sets up a field. The reflected wave is just the induced field. The amount of induced current is such that the induced field is identical to the incident wave but exactly opposite in phase. Thus, the resultant field just above the surface becomes zero. On the other hand, the field just below the surface (inside the conductor) is zero. Thus, the continuity condition is satisfied.

When the second medium is a dielectric medium, the situation is slightly more complicated. In the optical spectrum, when an \mathbf{E} field impinges on a dielectric medium, the orbits of the electrons in the medium become slightly displaced due to the \mathbf{E} field, and a dipole moment is induced. The time-varying nature of the incident \mathbf{E} field causes these dipole moments to vibrate. This vibrating dipole moment establishes the reradiating field. This field radiates both in the direction of the original medium as well as into the second medium. The former is the reflected wave and the latter is the transmitted wave. The relative amplitudes of these two waves are determined such that Eqs. (2.35) and (2.36) are satisfied.

Using Eq. (2.2), and $\theta_1 = \theta_3$, Eq. (2.38) becomes

$$\frac{\cos \theta_1}{\eta_1}(E_1 - E_3) = \frac{E_2}{\eta_2} \cos \theta_2 \quad (2.39)$$

Equations (2.37) and (2.39) are used to derive the expressions for r_{\perp} and t_{\perp} :

$$r_{\perp} = \frac{E_3}{E_1} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \quad (2.40)$$

$$t_{\perp} = \frac{E_2}{E_1} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \quad (2.41)$$

The above four equations for r_{\parallel} , t_{\parallel} , r_{\perp} , and t_{\perp} are called Fresnel's equations. Note that for $\theta_1 = \theta_2 = 0$, Eqs. (2.40) and (2.41) match Eqs. (2.18) and (2.19), respectively.

In order to calculate Fresnel's equations, it is necessary to know θ_2 using Snell's law, Eq. (2.10). One can also make use of Snell's law to completely eliminate both indices of refraction, n_1 and n_2 . With Eq. (2.10), Eq. (2.40) can be reduced to

$$r_{\perp} = -\frac{\sin(\theta_1 - \theta_2)}{\sin(\theta_1 + \theta_2)} \quad (2.42)$$

Similarly, but with a slightly longer derivation (see Problem 2.2), Eqs. (2.35), (2.36), and (2.41) are reduced to

$$r_{\parallel} = \frac{\tan(\theta_1 - \theta_2)}{\tan(\theta_1 + \theta_2)} \quad (2.43)$$

$$t_{\perp} = \frac{2 \cos \theta_1 \sin \theta_2}{\sin(\theta_1 + \theta_2)} \quad (2.44)$$

$$t_{\parallel} = \frac{2 \cos \theta_1 \sin \theta_2}{\sin(\theta_1 + \theta_2) \cos(\theta_1 - \theta_2)} \quad (2.45)$$

If Eqs. (2.42) and (2.44) are examined closely, one can verify that

$$1 + r_{\perp} = t_{\perp} \quad (2.46)$$

This is a direct consequence of Maxwell's continuity condition, as can be seen by multiplying both sides of Eq. (2.46) by the incident field E_1 . The left-hand side of Eq. (2.46) represents the tangential component of the resultant field just above the interface and the right-hand side, the same just below the interface. It should be noted that the same is not true for the relationship between r_{\parallel} and t_{\parallel} (see Problem 2.3).

To gain some insight into the actual quantities, let us calculate the values of the transmission and reflection coefficients as a function of incident angle for a particular interface. Figure 2.5 shows curves for the glass ($n_2 = 1.5$) and air ($n_1 = 1$) interface. Figure 2.5 reveals some noteworthy features. All curves turn downward as θ_1 increases. The transmission coefficients t_{\parallel} and t_{\perp} both decrease as the incident angle θ_1 increases, and when θ_1 approaches 90° ; that is, when the incident beam becomes nearly parallel to the interface, the amount of transmitted light is almost zero. The decrease is especially significant for $\theta_1 > 70^\circ$.

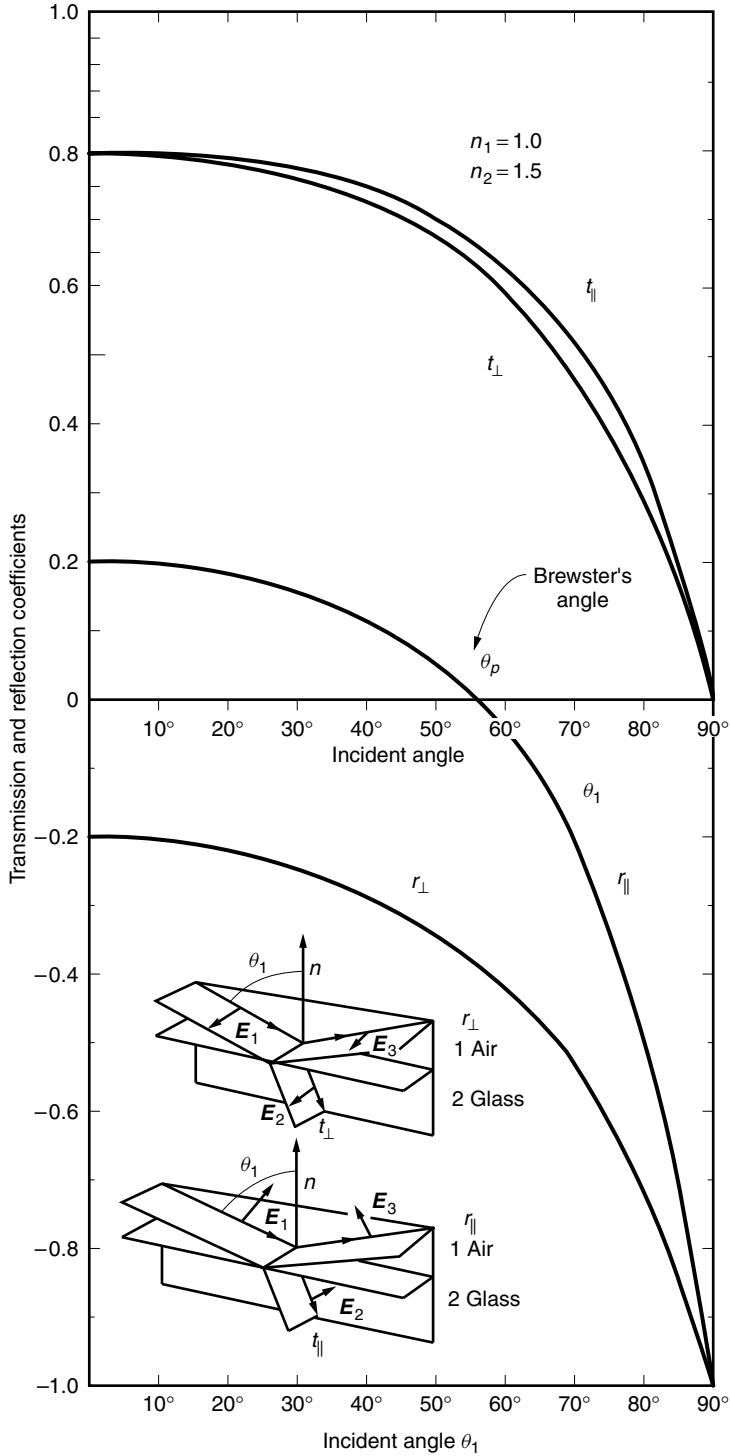


Figure 2.5 Reflection and transmission coefficients as a function of incident angle for air to glass interface ($n_1 = 1.0$, $n_2 = 1.5$).

This information is put to good use in the following practical example. Suppose that a glass lens is used to collect a diverging light source such as the emission from a light-emitting diode (LED). The edge of a double convex lens, as shown in Fig. 2.6a, should be avoided because the angle of incidence is large in this area and transmission is small. The result is an inefficient light collection.

The meniscus lens alleviates this problem, as shown in Fig. 2.6b. But if the meniscus lens is used incorrectly, the situation becomes even worse than with a double convex lens, as indicated in Fig. 2.6c.

Another interesting observation from Fig. 2.5 is that the curve for r_{\parallel} crosses the zero axis. This means that there is no reflected wave at this angle. This angle is called Brewster's angle and is discussed in detail in Section 2.5. At this angle, there is no reflected wave.

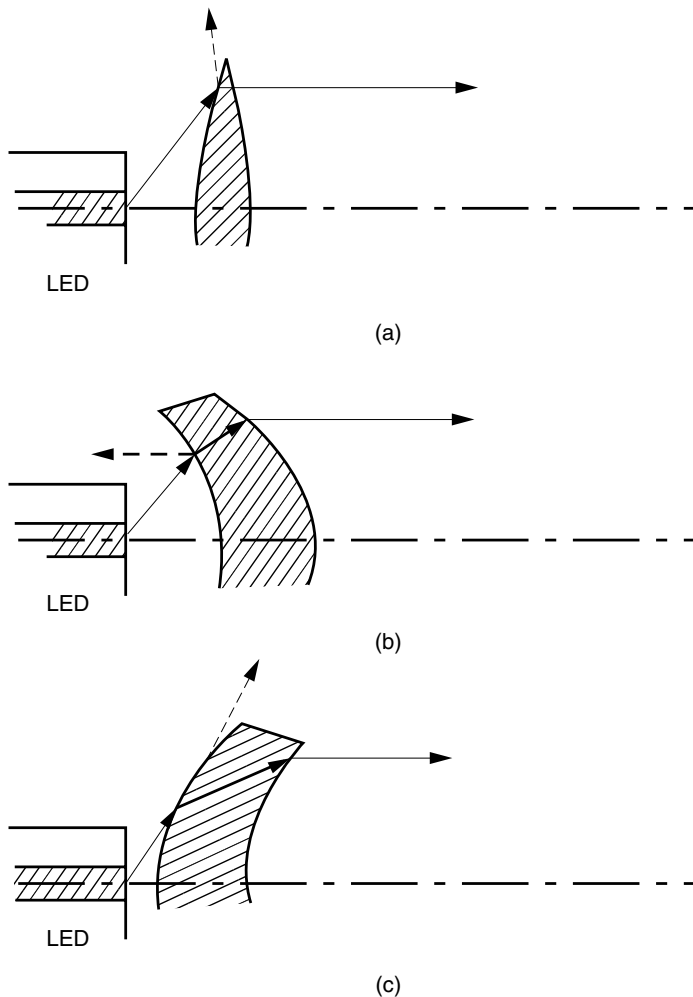


Figure 2.6 Selecting a lens to collect the emission from an LED. (a) Increase in reflection toward the edge of a double convex lens. (b) Minimal increase in reflection toward the edge of a meniscus lens. (c) Increase in reflection toward the edge with a misused meniscus lens.

2.3.3 Impedance Approach to Calculating Transmission and Reflection Coefficients

In the case of normal incidence the formulas for t and r are Eqs. (2.15) to (2.17), which are simpler and easier to remember than the equations for an arbitrary angle of incidence. The simplicity of Eqs. (2.15) to (2.17) is the basis for the impedance approach to calculating transmission and reflection coefficients. One defines an impedance out of certain components of \mathbf{E} and \mathbf{H} chosen such that their Poynting vectors point normal to the interface. From Fig. 2.4b, the impedances referring to the normal direction, which are sometimes called *characteristic wave impedances referred to the x direction* [2], are defined for incident and transmitted waves as

$$\eta_{1x} = \frac{E_1}{H_1 \cos \theta_1} = \frac{\eta_1}{\cos \theta_1} \quad (2.47)$$

$$\eta_{2x} = \frac{E_2}{H_2 \cos \theta_2} = \frac{\eta_2}{\cos \theta_2} \quad (2.48)$$

Replacing η_1 and η_2 in Eqs. (2.15) to (2.17) by η_{1x} and η_{2x} above, and then converting η to n , one can obtain the expressions for an arbitrary angle of incidence given by Eqs. (2.40) and (2.41). Equations (2.35) and (2.36) of the corresponding expressions for the other polarization are also obtained in a similar manner, referring to Fig. 2.4a and noting the sign convention mentioned in the boxed note. The impedance approach is quite powerful in dealing with surface acoustic wave devices, where the index of refraction is modulated in space and time, and the boundary conditions become more complex.

2.4 TRANSMITTANCE AND REFLECTANCE (AT AN ARBITRARY INCIDENT ANGLE)

The previous sections dealt primarily with the transmission and reflection coefficients t and r , which are field ratios. Transmittance and reflectance, which are power ratios, are studied in this section.

Referring to Fig. 2.7, consider the light power passing through a unit area on the interface. According to the law of conservation of energy, the power incident on this unit area has to be identical to the power emergent from the area. If this were not so, the energy would accumulate, resulting in an ever increasing build up of energy on the boundary. In applying conservation of energy, one has to realize two things:

1. Even if the amplitude of \mathbf{E} is the same, the energy of the wave is different if the medium in which the wave is propagating is different. The same is true for \mathbf{H} .
2. The power flowing through a unit area on the interface depends on the direction of incidence to the area.

The time rate of flow of the electromagnetic energy per unit area is given by the Poynting vector \mathbf{s} , where $\mathbf{s} = \mathbf{E} \times \mathbf{H}$. The magnitude of the Poynting vector, $s = |\mathbf{s}|$, represents the maximum instantaneous value of the electromagnetic power per unit area (s has units of W/m^2). Another quantity often used in optics is the irradiance I .

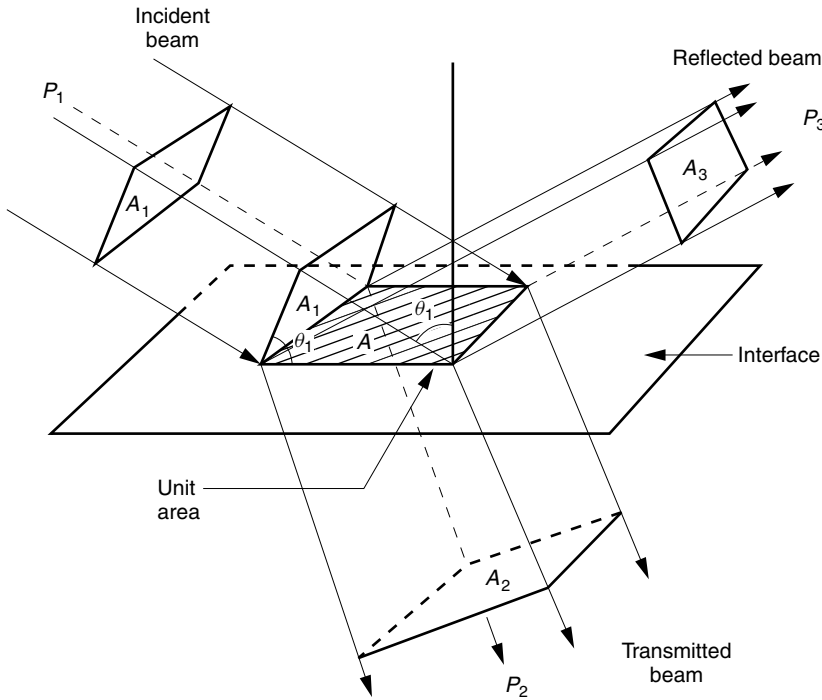


Figure 2.7 Reflectance and transmittance.

The irradiance I is the magnitude of the time average of the Poynting vector, $I = |\langle \mathbf{s} \rangle|$, and it also has units of W/m^2 . Since \mathbf{E} and \mathbf{H} are sinusoidally varying, the irradiance is $I = \frac{1}{2} |\mathbf{E} \times \mathbf{H}|$.

Using Eq. (2.2) or (2.5), the magnitude of the Poynting vector is given by

$$s = |\mathbf{s}| = \frac{1}{\eta} |\mathbf{E}|^2 \tag{2.49}$$

or

$$s = |\mathbf{s}| = \eta |\mathbf{H}|^2 \tag{2.50}$$

Suppose that the same energy flow, s_0 , is present in two different media such that $s_0 = (1/\eta_1) |\mathbf{E}_1|^2 = (1/\eta_2) |\mathbf{E}_2|^2$. What this means is that if $\eta_1 > \eta_2$, then $|\mathbf{E}_1|^2 > |\mathbf{E}_2|^2$. The medium with the greater impedance has the larger \mathbf{E} field.

As mentioned earlier, the power passing through the unit area in Fig. 2.7 depends on the angle of incidence. Let us use the symbol P to denote maximum instantaneous light power. For the geometry of Fig. 2.7, the incident, transmitted, and reflected light powers are

$$\begin{aligned} P_1 &= s_1 A_1 \\ P_2 &= s_2 A_2 \\ P_3 &= s_3 A_3 \end{aligned} \tag{2.51}$$

Also from Fig. 2.7, the areas A_1 , A_2 , and A_3 are related to the area A on the interface as follows:

$$\begin{aligned} A_1 &= A \cos \theta_1 \\ A_2 &= A \cos \theta_2 \\ A_3 &= A \cos \theta_1 \end{aligned} \quad (2.52)$$

As A is assumed to be a unit area, Eqs. (2.49), (2.51), and (2.52) are combined to give the incident, transmitted, and reflected powers:

$$\begin{aligned} P_1 &= \frac{1}{\eta_1} |\mathbf{E}_1|^2 \cos \theta_1 \\ P_2 &= \frac{1}{\eta_2} |\mathbf{E}_2|^2 \cos \theta_2 \\ P_3 &= \frac{1}{\eta_1} |\mathbf{E}_3|^2 \cos \theta_1 \end{aligned} \quad (2.53)$$

The transmittance T is defined as the ratio of transmitted power to incident power:

$$T = \frac{P_2}{P_1} = \frac{\eta_1 |\mathbf{E}_2|^2 \cos \theta_2}{\eta_2 |\mathbf{E}_1|^2 \cos \theta_1} \quad (2.54)$$

When $\mu_{r1} = \mu_{r2}$, $E_2 = E_1 t'_E$

$$T = \frac{n_2 \cos \theta_2}{n_1 \cos \theta_1} |t'_E|^2 \quad (2.55)$$

The value of T depends on the direction of polarization. For parallel (T_{\parallel}) and perpendicular (T_{\perp}) polarizations, $t'_E = t_{\parallel}$ and $t'_E = t_{\perp}$, respectively, have to be used in Eq. (2.55).

The reflectance R is the ratio of reflected power to incident power,

$$R = \frac{P_3}{P_1} = \frac{|\mathbf{E}_3|^2}{|\mathbf{E}_1|^2} = |r'_E|^2 \quad (2.56)$$

where R also depends on the polarization, and for parallel (R_{\parallel}) and for perpendicular (R_{\perp}) polarizations, $r'_E = r_{\parallel}$ and $r'_E = r_{\perp}$, respectively, have to be used in Eq. (2.56). (For general directions of polarization, see Ref. 1).

It is left as an exercise (see Problem 2.4) to show that

$$R + T = 1 \quad (2.57)$$

for both directions of polarization. In practice, Eq. (2.55) is seldom used to calculate T . Rather, the reflectance is first calculated via the simple relationship $R = |r'_E|^2$ and then the transmittance is found by making use of Eq. (2.57), $T = 1 - R$.

Figure 2.8 shows the plots of transmittance and reflectance at the interface between air and glass ($n_1 = 1.5$) as a function of the incident angle θ_1 . Note that light with perpendicular polarization has a narrower range of good transmission. To achieve 90% transmittance, the incident angle θ_1 has to stay smaller than 45° when perpendicular polarization is used, but the incident angle can be extended to 75° if parallel polarization is used.

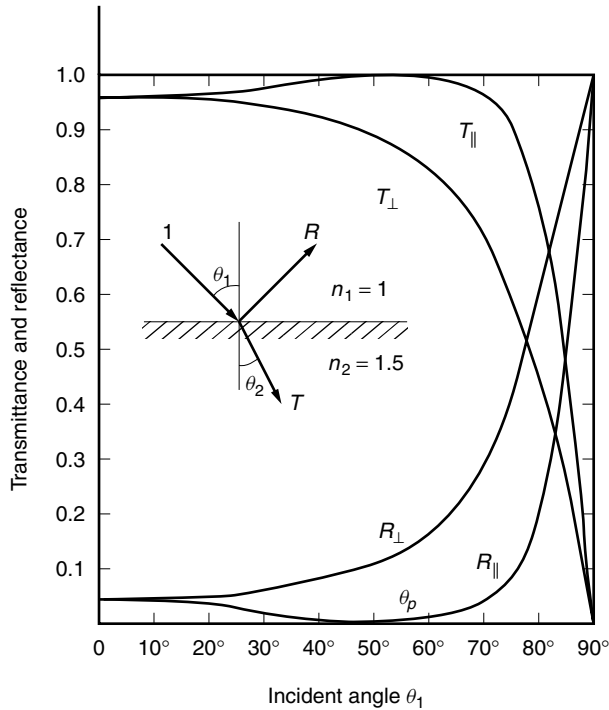


Figure 2.8 Transmittance and reflectance at the air-glass interface.

2.5 BREWSTER'S ANGLE

Of the four curves (for $n_1 = 1.0$, $n_2 = 1.5$) in Fig. 2.5, only one curve r crosses the horizontal axis. The reflection coefficient r becomes zero near $\theta_1 = 56^\circ$, or to be more exact, at $\theta_1 = 56^\circ 19'$; and there is no reflected wave.* This angle, labelled as θ_p in Fig. 2.5, is called Brewster's angle. Brewster's angle is used to avoid reflection from the surface. It is important to realize that Brewster's angle exists only when the direction of polarization is parallel to the plane of incidence.

The reflection coefficient r vanishes when the denominator of Eq. (2.43) becomes infinity, namely, when

$$\theta_1 + \theta_2 = 90^\circ \tag{2.58}$$

which is known as Brewster's condition. It is interesting to note that when Brewster's condition Eq. (2.58) is satisfied, the direction in which one would expect to find a reflection if it were to exist (dashed line in Fig. 2.9) coincides with the direction of polarization of the transmitted wave as shown in Fig. 2.9. From a microscopic viewpoint, the reflected wave is generated by the oscillation of electric dipoles in the transmission medium. The oscillating dipole does not radiate in the direction of oscillation.

* Substituting $\theta_p = 56^\circ 19'$ into Eq. (2.36), one can see that t is not unity at $\theta_1 = \theta_p$, even though r is zero. Why is this so? See Example 2.3.

Snell's law in impedances can be expressed as

$$\eta_1 \sin \theta_2 = \eta_2 \sin \theta_1 \quad (2.64)$$

With Eqs. (2.62)–(2.64)

$$\eta_{1x} = \eta_{2x} \quad (2.65)$$

Thus, Brewster's angle has led to the condition of impedance matching.

A good example of the use of Brewster's angle is the window of a gas laser tube of the external mirror type, as described in Example 2.2.

Example 2.2 Figure 2.10 shows a diagram of a gas laser tube with external mirrors. The end windows of the tube are tilted at Brewster's angle so that the reflection from the windows will be minimized. What is the Brewster's angle? Is the output of such a laser polarized? If so, in which direction is it polarized? Assume that the index of refraction of the glass window is $n_2 = 1.54$.

Solution From Eq. (2.59) with $n_1 = 1$, $n_2 = 1.54$, Brewster's angle is $\theta_p = 57^\circ$. Brewster's angle is applicable only to the wave with parallel polarization. The reflection coefficient for the perpendicular polarization r_\perp is calculated for comparison. From Eq. (2.40) or (2.42), $r_\perp = 0.4$. Compared to $r_\parallel = 0$, there is a significant difference. The external mirrors of the laser are usually highly reflective so that a large portion of the light makes thousands of passes back and forth between the two mirrors. For each pass through the laser medium, the light is amplified (this is a characteristic of laser media and represents a situation dealt with in Chapter 14 in Volume II). Because the perpendicular polarization suffers reflection losses at the windows, the losses for this polarization outweigh the gain through the amplifying medium. In comparison, the parallel polarization, which has no reflection losses at the windows, reaps the gain of the amplifying medium. Light exits the laser by transmission through the external mirrors. Even though the mirrors are highly reflective, a small amount is transmitted at each light pass. After the first few passes, the light that exits the laser is predominantly the parallel polarization: that is, the light that exits the laser of the external cavity type

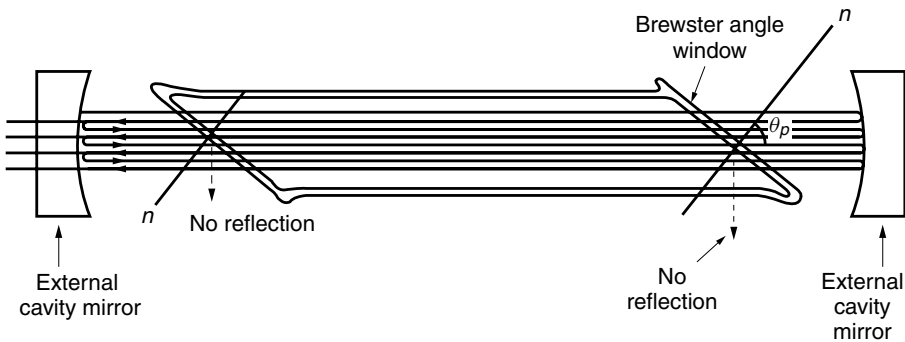


Figure 2.10 Laser cavity of the external mirror type.

is predominately polarized in the plane made by the normal to the window and the axis of the light beam. \square

Example 2.3 Referring to Fig. 2.5, $r_{\parallel} = 0$ at Brewster's angle but the corresponding transmission coefficient is 0.66. Since $r_{\parallel} = 0$, why isn't t_{\parallel} unity?

Solution In applying conservation of energy, the angle dependency of the power flow as indicated in Fig. 2.8 has to be considered, as well as the difference in \mathbf{E} due to the difference in media. These considerations were mentioned in Section 2.4. The law of conservation of energy does not say that $r_{\parallel} + t_{\parallel} = 1$; but the law of conservation of energy does say that $R_{\parallel} + T_{\parallel} = 1$. As a matter of fact, that $r_{\parallel} + t_{\parallel} = 1$ is not true is seen from Eqs. (2.35) and (2.36). Equations (2.55) and (2.56) give

$$R_{\parallel} = r_{\parallel}^2 = 0$$

$$T_{\parallel} = \frac{1.5 \cos(33^{\circ}41')}{\cos 56^{\circ}19'} \times 0.66^2 \approx 1.0 \quad \square$$

2.6 TOTAL INTERNAL REFLECTION

Consider the case when the light is incident from an optically dense medium (larger index of refraction n_1) onto that of an optically less dense medium (smaller index of refraction n_2) with the geometry shown in Fig. 2.11. If and only if the light is incident from an optically dense medium onto a less dense medium, is there a chance that the angle of the emergent light reaches 90° with an incident angle less than

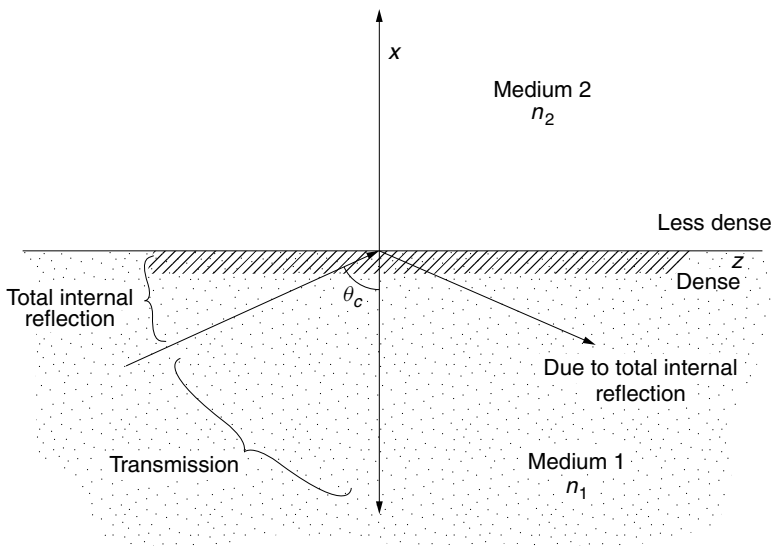
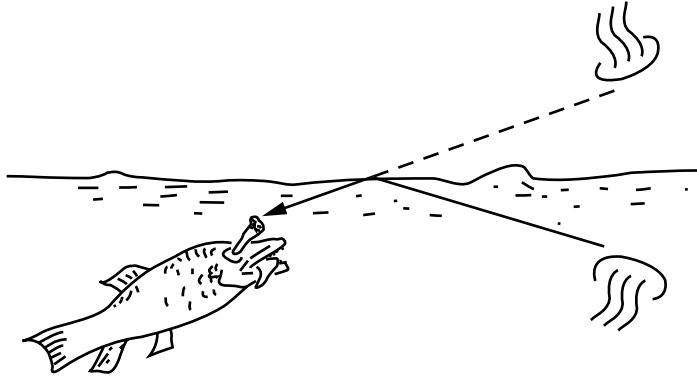


Figure 2.11 Total internal reflection, $n_1 > n_2$, $n = n_2/n_1 < 1$.



Total internal reflection.

90° . As soon as the emergent angle reaches 90° , *total internal reflection** takes place. The angle of total internal reflection follows the law of reflection, which means the angle of reflection is equal to the angle of incidence. The angle θ_c at which total internal reflection starts to take place is called the *critical angle* and is calculated from Eq. (2.10) as

$$\theta_c = \sin^{-1} \left(\frac{n_2}{n_1} \right) \quad (2.66)$$

In the field of fiber optics, total internal reflection at the boundary between the core and cladding layers supports the guided wave in the fiber. The light propagates as it is bounced back and forth.

Because of its durability, total internal reflection is sometimes used as a reflective device in place of a metal-deposited surface mirror, but the disadvantage is the restriction over the choice of incident angle.

2.6.1 Bends in a Guide

Changing the direction of an optical guide is often required. In many cases, a change of direction by a total internal reflection mirror is preferred to gradual bending of the guide. Gradual bending is limited to small bend angles. Typically, the radius of curvature of the bends cannot be less than 10–30 mm because, if any smaller, the incident angle to some portion of the wall of the guide becomes less than the critical angle, as indicated by the dotted line in Fig. 2.12a, and the light leaks.

A sharp bend, however, can be made if a flat mirror is placed at the bend [3,4]. The flat mirror preserves the angle of the zigzag path of the ray in the guide, as shown by the solid line, and the light does not leak. For instance, if the angle of incidence to the flat mirror is 65° , that of the circular bend in this particular geometry is 55° .

*The terms total reflection and total internal reflection are interchangeable. Johannes Kepler (1571–1630) experimentally discovered the existence of total internal reflection.

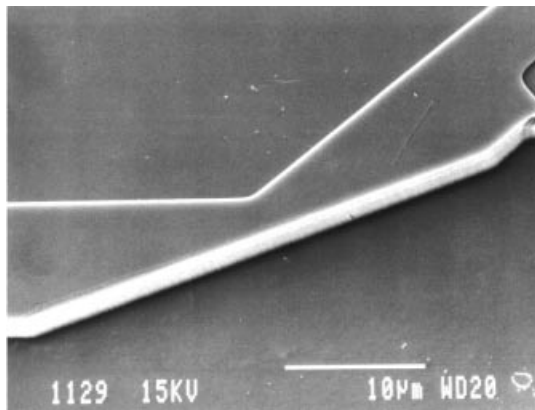
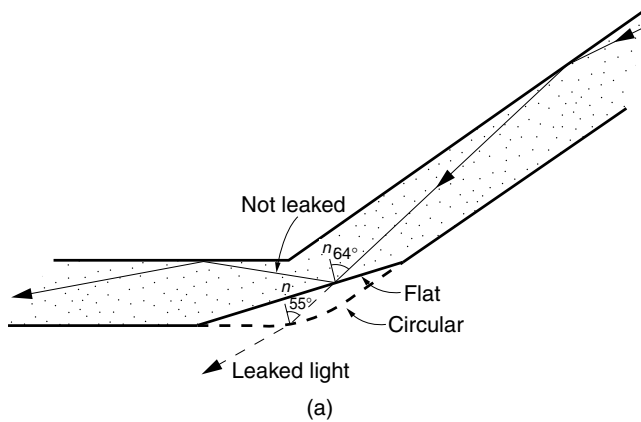


Figure 2.12 Low-loss corner mirrors with 45° deflection angle for integrated optics. (a) Bend by the total internal reflection flat mirror compared with that of a circular bend. (b) Scanning electron microscope picture of the total internal reflection corner mirror. (Photograph courtesy of E. Gini, G. Guekos, and H. Melchior [4].)

Figure 2.12b shows an electron microscope photograph of a GaAs rib guide mirror. The insertion loss[†] of such a bend was 0.3 dB [4].

2.7 WAVE EXPRESSIONS OF LIGHT

Up to this point, boundary phenomena, such as reflection, refraction, total internal reflection, and Brewster’s angle, have been described without explicit reference to the wave nature of light. In the next few sections, the wave nature will be emphasized. By using the wave expressions, finer details can be reworked.

[†] Let P_{in} be the power into the device and P_{out} be the power out of the device. The insertion loss of the device is defined as

$$10 \log \frac{P_{in}}{P_{out}}$$

2.7.1 Fields Near the Boundary

The expressions representing the fields near a discrete boundary between media 1 and 2 with the geometry shown in Fig. 2.11 will be described [5].

In medium 1, a plane wave propagating in the direction

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{k}} \quad (2.67)$$

is expressed from Eq. (1.22) as

$$\mathbf{E}_1 = |\mathbf{E}_1| e^{j(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (2.68)$$

where \mathbf{r} is the position vector

$$\mathbf{r} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + z \hat{\mathbf{k}}$$

The incident wave propagating in medium 1 is given by

$$\mathbf{E}_1 = |\mathbf{E}_1| e^{j(k_{1x}x + k_{1y}y + k_{1z}z - \omega t)} \quad (2.69)$$

Similarly, the transmitted wave in medium 2 is

$$\mathbf{E}_2 = |\mathbf{E}_2| e^{j(k_{2x}x + k_{2y}y + k_{2z}z - \omega t)} \quad (2.70)$$

Inserting \mathbf{E}_1 and \mathbf{E}_2 into the wave equation, $\nabla^2 \mathbf{E} + (nk)^2 \mathbf{E} = 0$, gives the condition that the sum of the squares of the k components has to equal the square of the propagation constant in that medium; that is,

$$k_{1x}^2 + k_{1y}^2 + k_{1z}^2 = (n_1 k)^2 \quad (2.71)$$

and

$$k_{2x}^2 + k_{2y}^2 + k_{2z}^2 = (n_2 k)^2 \quad (2.72)$$

Equations (2.71) and (2.72) will be manipulated in order to clarify the allowed and the prohibited regions of propagation.

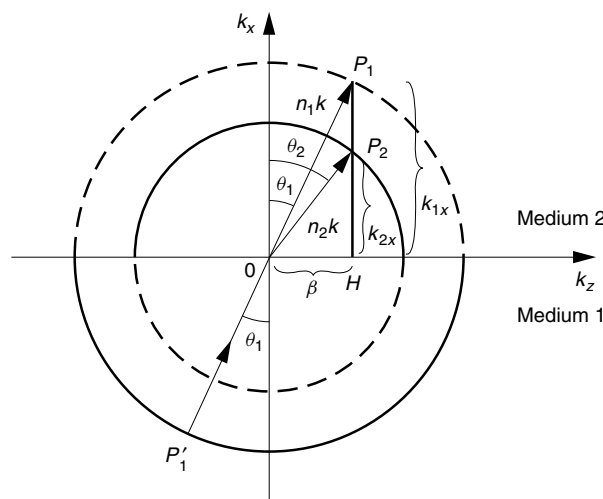


Figure 2.13 Wave vectors propagating in k space.

Equations (2.71) and (2.72) are equations of spheres in k space, with radii n_1k and n_2k . Figure 2.13 is a two-dimensional representation in k space, with $k_y = 0$. In this figure, propagation takes place only in the x and z directions. The angles that the radial vectors \vec{OP}_1 (translated from $\vec{P}_1'0$ for convenience) and \vec{OP}_2 make with the k_x axis are the incident and emergent angles, θ_1 and θ_2 , respectively. The wavelength matching condition (phase matching condition) on the boundary for the z direction is satisfied by ensuring that the projections of \vec{OP}_1 and \vec{OP}_2 onto the k_z axis are the same, namely,

$$k_{1z} = k_{2z} = \beta \quad (2.73)$$

Equation (2.73) fixes the value of θ_2 for a given θ_1 . With the phase matching condition, Eqs. (2.71) and (2.72) become

$$\begin{aligned} k_{1x}^2 + \beta^2 &= (n_1k)^2 \\ k_{2x}^2 + \beta^2 &= (n_2k)^2 \end{aligned} \quad (2.74)$$

Expressions for E_1 and E_2 will now be found. Referring to Fig. 2.13, one has $k_{2x} = n_2k \cos \theta_2$, so that the transmitted wave can be expressed as

$$E_2 = |E_2|e^{jn_2k \cos \theta_2 \cdot x + j\beta z - j\omega t} \quad (2.75)$$

Now, Eq. (2.75) will be rewritten in terms of the incident angle θ_1 and the transmission coefficients $t_{\parallel, \perp}$. Snell's law, Eq. (2.10), gives

$$\cos \theta_2 = \frac{1}{n} \sqrt{n^2 - \sin^2 \theta_1} \quad (2.76)$$

where

$$n = \frac{n_2}{n_1}$$

Henceforth, the symbol n without the suffix is used to denote the ratio of the indices of refraction.

The expressions for E_2 and E_1 are

$$E_2 = t_{\parallel, \perp} |E_1| e^{jn_1k[\sqrt{n^2 - \sin^2 \theta_1} x + (\sin \theta_1) z] - j\omega t} \quad (2.77)$$

and

$$E_1 = |E_1| e^{jn_1k[(\cos \theta_1)x + (\sin \theta_1)z] - j\omega t} \quad (2.78)$$

Note that Eqs. (2.77) and (2.78) are equivalent if n is unity, as one would expect.

2.8 THE EVANESCENT WAVE

When light is incident from an optically dense medium onto a less dense medium, total internal reflection takes place as soon as the emergent angle θ_2 reaches 90° . Let us take a look at total internal reflection from the viewpoint of the requirements of the boundary condition on the border. Since the wave is totally reflected back into the denser medium, no field is supposed to be present inside the less dense medium. Does the field abruptly

become zero on the boundary? Because of the boundary condition of continuity, the field in the less dense medium cannot abruptly become zero. The presence of the so-called evanescent wave (also called surface wave) solves this problem. The amplitude of the evanescent wave at the boundary is the same as the field on the boundary of the dense medium, but the amplitude of the evanescent wave decays very rapidly as it goes away from the boundary into the less dense medium. Moreover, the evanescent wave moves along the boundary in the z direction at the same speed as the incident wave in order to satisfy the phase matching condition.

Let us study this mysterious evanescent wave quantitatively. Equation (2.77) is the expression for the wave in the transmitted medium. The exponential term in this equation contains a square root. In the present case, $n < 1$, so that the quantity inside the square root becomes negative as θ_1 is increased. This means that, for large enough θ_1 , the square root term becomes a pure imaginary number $j\gamma$ and the expression for the transmitted wave becomes

$$E_2 = t_{\parallel,\perp} |E_1| e^{-\gamma x + j(\beta z - \omega t)} \quad (2.79)$$

where

$$\gamma = n_1 k \sqrt{\sin^2 \theta_1 - n^2} \quad (2.80)$$

The wave expressed by Eq. (2.79) is precisely the evanescent wave. Note from Eq. (2.79) that the amplitude of E_2 decays exponentially in the x direction, and there is no sinusoidal phase variation in the x direction. On the other hand, the phase variation in the z direction synchronizes with that of the incident wave. These are the major characteristics of the evanescent wave. The *effective depth* h of the penetration of the evanescent wave, which is defined as the depth where the amplitude decays to $1/e$ of that on the boundary, is $1/\gamma$ from Eq. (2.79). Equation (2.79) is graphically represented in Fig. 2.14. The equiamplitude lines are horizontal. The amplitude decays exponentially with x . The equiphase lines are vertical and they translate in the z direction with the incident wave.

In order to determine the amplitude of the evanescent wave at the boundary, let us first turn our attention to the fields inside the optically dense medium. When total internal reflection occurs, the optically dense medium contains both the incident wave, and the wave that is totally reflected. While their z components of propagation are the same, their x components of propagation are opposite. Because of the opposite directions of propagation, they form a standing wave in the x direction in the optically dense medium. In the less dense medium, only the evanescent wave exists. On the boundary, the amplitude of the standing wave and that of the evanescent wave have to be matched.

In the following sections, the reflection coefficient for the case of total internal reflection will be derived in order to calculate the reflected wave. Then, the reflected wave will be used to find the standing wave in the dense medium, and finally the amplitude of the evanescent wave at the boundary will be found.

2.8.1 Transmission and Reflection Coefficients for Total Internal Reflection

The reflection coefficients [5] for parallel and perpendicular polarizations will be extended to the case of total internal reflection. It will be assumed that Snell's law still holds true even in the case of total internal reflection because after all Snell's law is derived

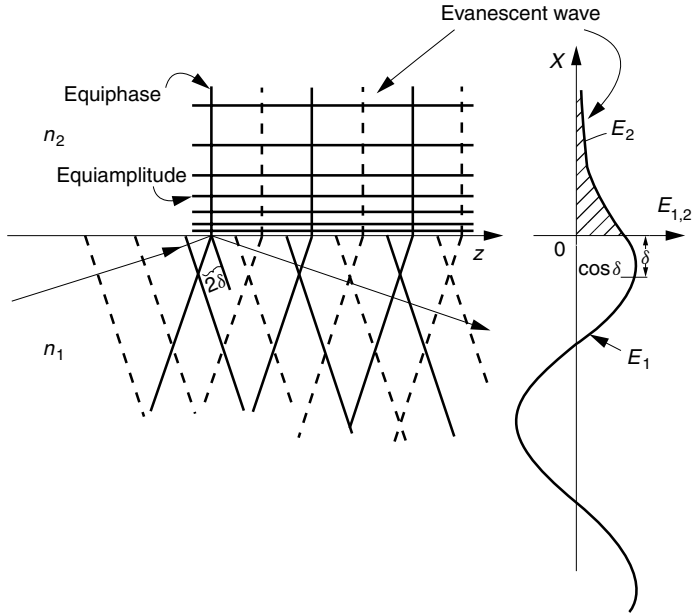


Figure 2.14 Evanescent wave ($n_1 > n_2$).

from the k -matching on the boundary as illustrated in Section 2.2 and has to be satisfied at all times. Writing Eqs. (2.35) and (2.40) in terms of the ratio $n = n_2/n_1$ gives

$$r_{\parallel} = \frac{\cos \theta_1 - (1/n) \cos \theta_2}{\cos \theta_1 + (1/n) \cos \theta_2} \tag{2.81}$$

$$r_{\perp} = \frac{\cos \theta_1 - n \cos \theta_2}{\cos \theta_1 + n \cos \theta_2} \tag{2.82}$$

For a given θ_1 , the only unknown quantity in Eqs. (2.81) and (2.82) is $\cos \theta_2$. Equation (2.76) gives an expression for $\cos \theta_2$. In the case of total internal reflection, $n < \sin \theta_1$ and Eq. (2.76) is rewritten as

$$\cos \theta_2 = \frac{j}{n} \sqrt{\sin^2 \theta_1 - n^2} \tag{2.83}$$

in order to explicitly show that $\cos \theta_2$ is a pure imaginary number. Insertion of Eq. (2.83) into Eqs. (2.81) and (2.82) gives

$$r_{\parallel} = \frac{\cos \theta_1 - j \sqrt{\sin^2 \theta_1 - n^2}/n^2}{\cos \theta_1 + j \sqrt{\sin^2 \theta_1 - n^2}/n^2} \tag{2.84}$$

$$r_{\perp} = \frac{\cos \theta_1 - j \sqrt{\sin^2 \theta_1 - n^2}}{\cos \theta_1 + j \sqrt{\sin^2 \theta_1 - n^2}} \tag{2.85}$$

Note that, in both the denominator and the numerator, the first term is real while the second is imaginary. Equations (2.84) and (2.85) can be written in simpler form

if the polar coordinate representation of a complex number is used. First, since the magnitudes of the denominator and numerator are identical, the absolute value of the reflection coefficient is unity. Next, if the phase angle of the numerator is $-\delta_{\parallel,\perp}$, and that of the denominator is $\delta_{\parallel,\perp}$, then

$$r_{\parallel} = e^{j\phi_{\parallel}} \quad \text{where} \quad \phi_{\parallel} = -2\delta_{\parallel} \quad (2.86)$$

$$r_{\perp} = e^{j\phi_{\perp}} \quad \text{where} \quad \phi_{\perp} = -2\delta_{\perp} \quad (2.87)$$

and

$$\tan \delta_{\parallel} = \frac{\sqrt{\sin^2 \theta_1 - n^2}}{n^2 \cos \theta_1} \quad (2.88)$$

$$\tan \delta_{\perp} = \frac{\sqrt{\sin^2 \theta_1 - n^2}}{\cos \theta_1} \quad (2.89)$$

Thus, the magnitude of the reflection coefficient for the total internal reflection is indeed unity but there is a phase lag on the boundary by $\phi_{\parallel,\perp}$ radians. The value of $\phi_{\parallel,\perp}$ is zero at the critical angle and approaches $-\pi$ radians as θ_1 increases. Using a trigonometric relationship with Eqs. (2.88) and (2.89), the difference becomes

$$\tan(\delta_{\parallel} - \delta_{\perp}) = \frac{\cos \theta_1 \sqrt{\sin^2 \theta_1 - n^2}}{\sin^2 \theta_1} \quad (2.90)$$

The phase shift difference $\phi_{\parallel} - \phi_{\perp} = -2(\delta_{\parallel} - \delta_{\perp})$ can be used to make waveplates (Section 6.3). A phase difference of $\pi/2$ radians between the waves of parallel and perpendicular polarizations results in a quarter-waveplate that can convert a linearly polarized wave into a circularly polarized wave. See Example 2.4 for more details. Such a quarter-waveplate is quite useful because it is independent of wavelength. Another good feature of this device is its slow dependence on θ_1 , as shown by the dashed line in Fig. 2.15.

Example 2.4 A Fresnel rhomb makes use of the phase shift difference produced by total internal reflection to convert a linearly polarized incident wave into a circularly polarized emergent wave. With the proper choice of incident angle, the Fresnel rhomb is designed to create a 45° phase difference at each of two total internal reflections, in order to achieve the 90° difference between the waves of parallel and perpendicular polarizations. The direction of the polarization of the incident wave is tilted 45° as shown in Fig. 2.16, so that the incident wave is decomposed into equal amplitudes of parallel and perpendicular polarizations.

These equal-amplitude component waves with 90° relative phase difference constitute a circularly polarized wave (mentioned in Chapter 6). Find the angle θ of incidence into the Fresnel rhomb shown in Fig. 2.16. Assume $n = 1/1.5$.

Solution Since each phase difference $2(\delta_{\parallel} - \delta_{\perp})$ is $\pi/4$, from Eq. (2.90), we have

$$\tan \frac{\pi}{8} = \frac{\cos \theta_1 \sqrt{\sin^2 \theta_1 - n^2}}{\sin^2 \theta_1}$$

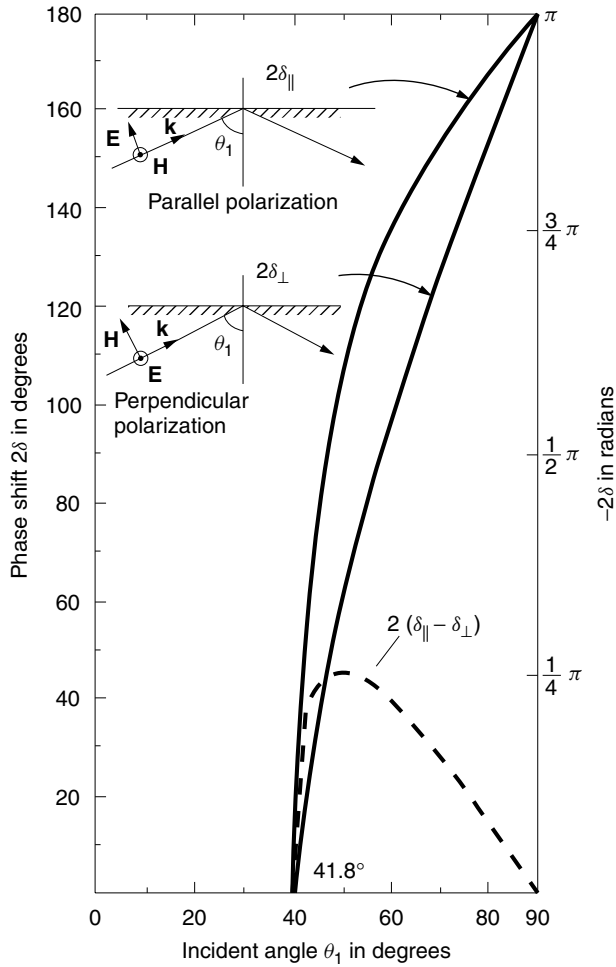


Figure 2.15 Phase shift associated with the total internal reflection at the glass ($n = 1.5$) and air interface.

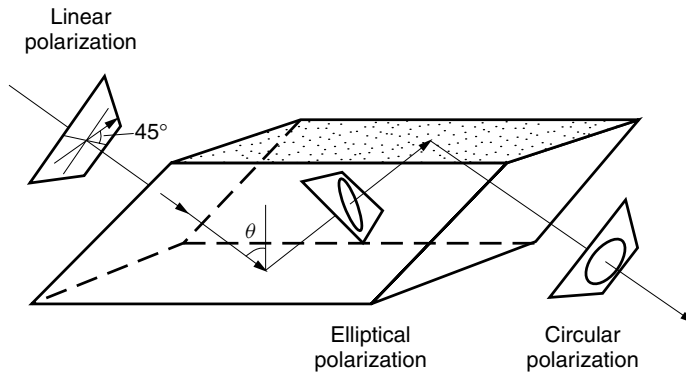


Figure 2.16 Fresnel rhomb.

Solving for $\sin^2 \theta_1$ gives

$$(1 - x)(x - n^2) = A^2 x^2$$

where

$$A = \tan \pi/8$$

$$x = \sin^2 \theta_1$$

This is a quadratic equation,

$$(A^2 + 1)x^2 - (n^2 + 1)x + n^2 = 0$$

with solution

$$x = \frac{(n^2 + 1) \pm \sqrt{(n^2 + 1)^2 - 4(A^2 + 1)n^2}}{2(A^2 + 1)}$$

$$\theta_1 = \sin^{-1}(\sqrt{x})$$

$$\theta_1 = 50.23^\circ \text{ and } 53.26^\circ \quad \square$$

Here, Eqs. (2.88) and (2.89) will be rewritten using the k -diagram mentioned in Section 2.8.3 so that they will be in a form suitable for describing the phase delay at the boundaries of the optical guide in the chapter to come. Let both the numerators and the denominators of Eqs. (2.88) and (2.89) be multiplied by $n_1 k$. Comparing with Eqs. (2.80), the numerators of both equations become γ . From Fig. 2.13, the denominators of Eqs. (2.88) and (2.89) become $n^2 k_{1x}$ and k_{1x} , respectively. Thus,

$$\tan \delta_{\parallel} = \frac{\gamma}{n^2 k_{1x}} \quad (2.91)$$

$$\tan \delta_{\perp} = \frac{\gamma}{k_{1x}} \quad (2.92)$$

In the particle theory of light, a propagation constant represents something proportional to the momentum of a photon, so that $\tan \delta_{\parallel}$ and $\tan \delta_{\perp}$ can then be interpreted as the ratio of momenta in medium 1 and medium 2.

Next, the transmission coefficient of total internal reflection will be considered. Keep in mind that the meaning of the transmission coefficient in the case of the total internal reflection is slightly different from the usual refraction. A simple model is that the evanescent wave is a result of light energy that goes into the less dense medium only for a short distance and then comes back again into the dense medium. The transmission coefficient describes the magnitude on the boundary but just a short distance inside the less dense medium.

An attempt will be made to find out how much light goes into the less dense medium by calculating the transmission coefficient of the total internal reflection at the boundary. Note from Eqs. (2.35) and (2.36) that

$$t_{\parallel} = \frac{1}{n}(1 + r_{\parallel}) \quad (2.93)$$

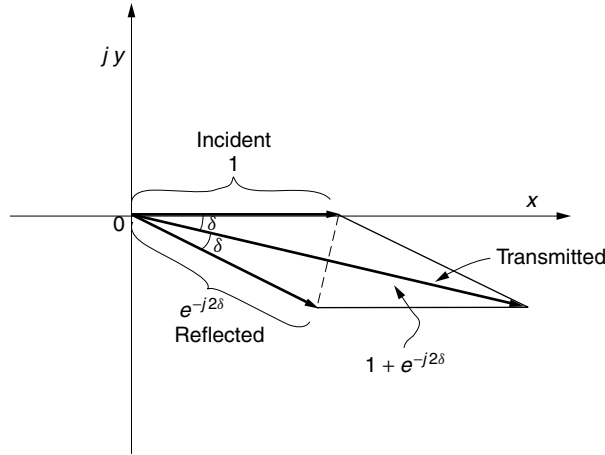


Figure 2.17 Polar coordinates of a complex-valued transmission coefficient.

and from Eqs. (2.40) and (2.41)

$$t_{\perp} = 1 + r_{\perp} \quad (2.94)$$

Use of Eqs. (2.86) and (2.87) in Eqs. (2.93) and (2.94) and with the aid of Fig. 2.17 gives

$$t_{\parallel} = \frac{2}{n} \cos \delta_{\parallel} e^{-j\delta_{\parallel}} \quad (2.95)$$

$$t_{\perp} = 2 \cos \delta_{\perp} e^{-j\delta_{\perp}} \quad (2.96)$$

Using Eqs. (2.88) and (2.89), $\cos \delta_{\parallel}$ and $\cos \delta_{\perp}$ in Eqs. (2.95) and (2.96) can be further rewritten as

$$t_{\parallel} = \frac{2n \cos \theta_1}{\sqrt{1-n^2} \sqrt{1-(n^2+1)\cos^2 \theta_1}} e^{-j\delta_{\parallel}} \quad (2.97)$$

$$t_{\perp} = \frac{2 \cos \theta_1}{\sqrt{1-n^2}} e^{-j\delta_{\perp}} \quad (2.98)$$

Recall that the momentum perpendicular to the boundary is $kn_1 \cos \theta_1$. With an increase in $\cos \theta_1$, this momentum increases, and also the magnitude of the transmission coefficient into the second medium increases. This is as if the photon in the optically dense medium is being pushed out into the less dense medium by this momentum. The extent to which the evanescent wave protrudes out into the less dense material reaches its maximum at the critical angle. The amount of phase shift, however, reaches zero at the critical angle from Eqs. (2.88) and (2.89). This fact rejects the simple explanation that the phase delay is the time needed for the evanescent wave to go out into the less dense medium and come back into the dense medium. The evanescent wave and phase delay indeed exist, but it cannot safely be said that the phase delay is due to the round-trip time. Another example where the phase delay at the boundary has nothing to do with penetration into the second medium is reflection from a perfect mirror. The

reflection from the perfect mirror creates π radians of phase shift but has nothing to do with the round-trip time in the mirror, since light does not go into the mirror at all. The phase shift is needed to match the boundary condition. Goos and Hänchen conducted interesting experiments to clarify these matters as described in the next section.

2.8.2 Goos-Hänchen Shift

Goos and Hänchen investigated the ray path associated with total internal reflection. They wanted to determine whether or not a spatial lateral shift accompanies the phase shift when light undergoes total internal reflection. One possibility is that the light ray is reflected back immediately into the dense medium without a spatial lateral shift but with $\phi_{\parallel, \perp}$ -radian phase delay as in the case of reflection from a surface mirror. The other possibility is that the reflected light ray first penetrates into the less dense medium and is laterally shifted along the boundary before going back into the dense medium as well as undergoing the phase delay.

In 1947, Goos and Hänchen [6] looked at the difference between reflection from a silver surface and total internal reflection at a glass-air interface. A clever experiment was performed using a glass prism onto which a narrow strip of silver was deposited along the center. As shown in Fig. 2.18, a wide light beam was incident on one end of the prism at such an incident angle that total internal reflections took place repeatedly inside the prism before the beam exited the prism and was projected onto the screen. These total internal reflections took place only with the portion excluding the silver deposit in the center. In the center, the light is reflected repeatedly by the silver strip as

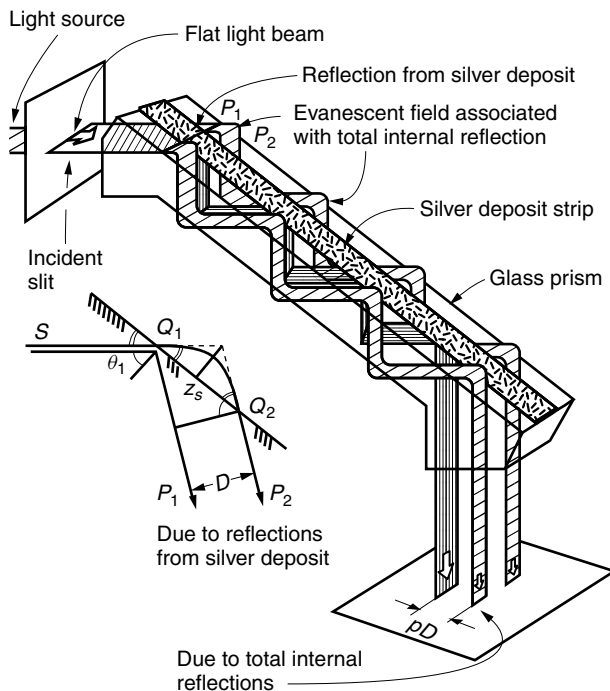


Figure 2.18 Goos and Hänchen experiment.

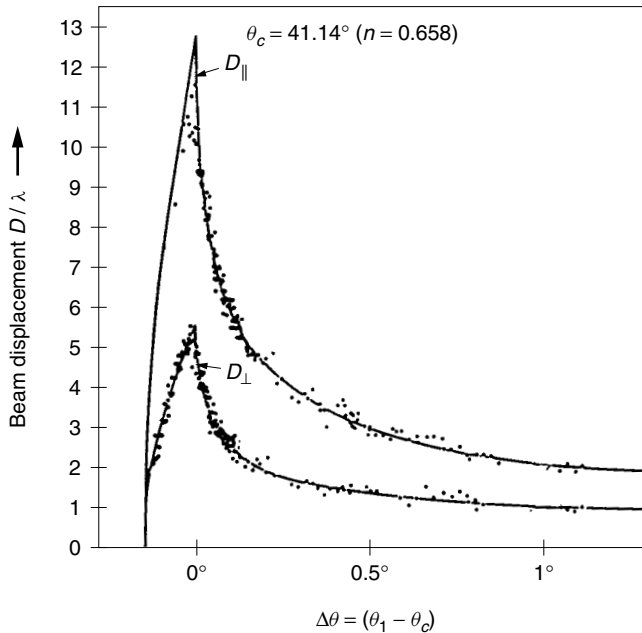


Figure 2.19 The beam displacement of the Goos-Hänchen shift versus the angle difference $(\theta_1 - \theta_c)$ in degrees. (After H. K. V. Lotsch [7] and H. Wolter [8].)

it travels through the prism. On the screen, the translation between the two light beams with different mechanisms of reflection can be compared, as shown in Fig. 2.18.

The explanation of the observed results is indicated by the diagram on the lower left-hand side in Fig. 2.18. While the reflected light from the silver surface takes the direct path SQ_1P_1 , the light that has undergone total internal reflection penetrates the optically less dense medium and then comes back into the optically dense medium taking the route $SQ_1Q_2P_2$, thus causing a beam displacement D between the two beams after reflection. If the number of reflections is p , the lateral shift of the output beams on the screen is pD . Goos and Hänchen experimentally verified the shift. Such a translation is known as the Goos-Hänchen shift. The Goos-Hänchen shift is an important quantity to consider when attempting to accurately estimate the penetration depth of the light field outside an optical guide or to fully account for the phase shift associated with an optical guide.

The original Goos-Hänchen experiments were refined to such an extent that the modified ray model agrees quite well with the experimental values, as shown in Fig. 2.19 [7,8]. For these experiments, glass with $n = 1.520$ was used as the optically dense medium, and air as the optically less dense medium. Near the critical angle, where the δ 's are at their minimum, the D 's are at their maximum. The value of the beam displacement D_{\parallel} with light polarized parallel to the plane of incidence is over 10 times the wavelength.

2.8.3 Evanescent Field and Its Adjacent Fields

Now we are ready to calculate the fields around the interface. In the dense medium, a standing wave E_{s1} is formed by the incident and reflected waves. Using Eq. (2.86),

the expression for the standing wave is given by

$$E_{s1} = E_1[e^{j(k_{1x}x+\beta z)} + e^{j(-2\delta-k_{1x}x+\beta z)}]e^{-j\omega t} \tag{2.99}$$

and finally

$$E_{s1} = 2E_1 \cos(k_{1x}x + \delta)e^{j(\beta z - \omega t - \delta)}, \quad x < 0 \tag{2.100}$$

where the subscripts \parallel and \perp have been suppressed for convenience. Equation (2.100) is the expression for a standing wave in the x direction and a propagating wave in the z direction. The amplitude of Eq. (2.100) is plotted on the right in Fig. 2.14. The position of maximum intensity is shifted downward by δ radians.

Because of the continuity condition, the amplitude $|E_2|$ of the evanescent wave has to be identical to that of the standing wave E_{s1} evaluated on the boundary; hence, from Eq. (2.79)

$$E_2 = 2E_1 e^{-j\delta} (\cos \delta) e^{-\gamma x + j(\beta z - \omega t)}, \quad x > 0 \tag{2.101}$$

Equations (2.100) and (2.101) together complete the expressions near the boundary.

With a decrease in the incident angle θ_1 from 90° approaching the critical angle θ_c , both δ and γ decrease and $h(= \gamma^{-1})$ increases according to Eqs. (2.80), (2.88), and (2.89). This means that with a decrease in θ_1 , the maxima shift toward the interface and the evanescent wave shifts into the optically less dense medium, as illustrated in Fig. 2.20. An analogy can be made with breaking a piece of wood. The sharper the bend, the more the broken ends protrude.

In practice, with the wavelengths $0.85\text{--}1.55 \mu\text{m}$ and $n_1 \simeq 1.55$, $n_2 \simeq 1.54$, which are normally used for the components of fiber optical communication systems, the extent of the evanescent wave is of the order of $1\text{--}10 \mu\text{m}$. The minimum required diameter of the cladding layer can be estimated (see Chapter 11).

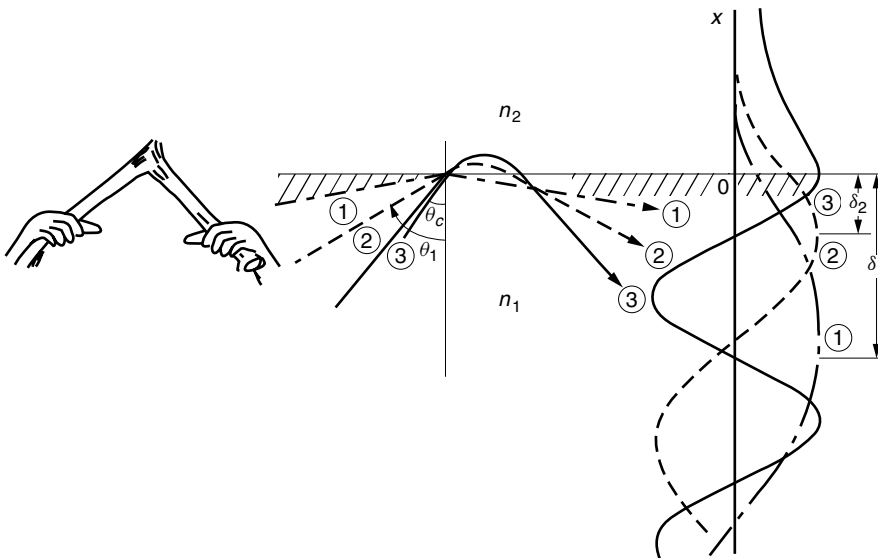


Figure 2.20 Relationship between incident angle θ_1 and the shape of the evanescent waves, $\theta_1 > \theta_c$.

In integrated optics, where the dimensions of the elements are quite small, a thorough understanding of the evanescent wave is very important. A typical example of the use of the evanescent wave is an optical guide coupler, such as shown in Fig. 2.21. An optical coupler is used either for monitoring the signal of the main guide or dividing the power into desired proportions.

When an auxiliary optical guide is put in the evanescent field of the main optical guide, the auxiliary guide is excited by the evanescent field. The amount of excitation can be controlled by either adjusting the spacing between the guides, the indices of refraction of the media surrounding the guides, or the length of the region of interaction. In fact, by adjusting these parameters properly, almost the entire energy of guide 1 can be transferred into guide 2.

Another interesting phenomenon is that if the interaction region is made long enough, the energy that was transferred to the auxiliary guide will itself set up an evanescent field, which in turn excites the main guide and transfers energy back into the main guide. If the guide is made out of an electrooptic material like lithium niobate, LiNbO_3 , the index of refraction of the guide can be changed by an external electric field (see Chapter 5). By careful design of the parameters of the coupler, an electronically controlled optical switch can be fabricated. The optical switch based on coupling of the evanescent field can achieve nearly 100% switching with a low control voltage, typically less than 10 volts.

The phasefront of the evanescent wave moves at the same speed as that of the bulk wave (wave inside the optically dense medium). It does not leave the surface, and so no loss of energy takes place. However, as

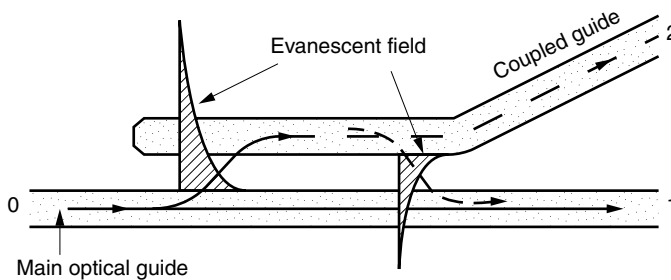
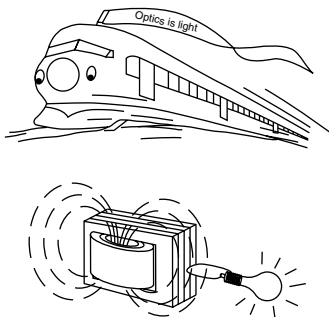


Figure 2.21 Optical guide coupling by the evanescent wave.



in the example of the optical guide coupler, if there is another guide that can carry the energy away, the evanescent wave provides the means of conveying energy from the main guide to the second guide.

The evanescent wave can be compared to a banner attached to a train. The banner moves with the train but has very little to do with the loss of energy of the train. Should the banner be caught by an external obstacle, it will convey the train's energy into the external obstacle.

Another analogy might be the magnetic flux leaking out of a transformer. The leaked magnetic flux is there all the time but contributes very little to the loss of energy of the main source. But once a loop circuit is placed in the leakage field, the energy of the main source can be transferred into the loop circuit.

2.8.4 k Diagrams for the Graphical Solution of the Evanescent Wave

Many of the expressions derived thus far have been expressed in terms of angles. Here the same quantities will be expressed in terms of the k_x vectors. From Fig. 2.13, the expressions for the k_x vectors are $k_{1x} = n_1 k \cos \theta_1$ and $k_{2x} = n_2 k \cos \theta_2$. There is a one-to-one correspondence between k_{1x} and θ_1 , and likewise between k_{2x} and θ_2 inside a given medium. It is sometimes easier to use k_{1x} and k_{2x} rather than angles. For example, when analyzing the fields inside an optical guide having a rectangular shape, the rectangular quantities k_{1x} and k_{2x} are much more convenient variables to use than θ_1 or θ_2 .

From Eq. (2.74), which was obtained by phase matching, β can be eliminated to give

$$k_{1x}^2 - k_{2x}^2 = (n_1^2 - n_2^2)k^2 \quad (2.102)$$

Equation (2.102) is Snell's law in k -coordinates (see Problem 2.8). Note that when $n_1 > n_2$, the right-hand side of Eq. (2.102) is positive and there is a region where k_{1x} becomes too small (i.e., θ_1 too large) to maintain the left-hand side of Eq. (2.102) positive and there is no real number k_{2x} . Consequently, there is no transmitted light. This is the region of total internal reflection discussed in Section 2.6. The only possible way to keep the left-hand side positive is to make k_{2x} a pure imaginary number $\pm j\gamma$; and in the case of the total internal reflection k_{2x} is imaginary.

Thus, when total internal reflection exists,

$$k_{2x} = \pm j\gamma \quad k_{1x} = K \quad (2.103)$$

where k_{1x} is more conveniently written without the subscript. Inserting Eq. (2.103) into Eqs. (2.74) and (2.102) gives

$$K^2 + \beta^2 = (n_1 k)^2 \quad (2.104)$$

$$K^2 + \gamma^2 = (n_1^2 - n_2^2)k^2 \quad (2.105)$$

Both of these curves are circles. The graph of Eq. (2.104) is the K - β diagram and the radius is $n_1 k$. The graph of Eq. (2.105) is the K - γ diagram and the radius is $k\sqrt{n_1^2 - n_2^2}$, as shown in Fig. 2.22. Note that the radius of the K - γ diagram is smaller than that of the K - β diagram. For a given θ_1 , draw in a straight line at θ_1 with respect

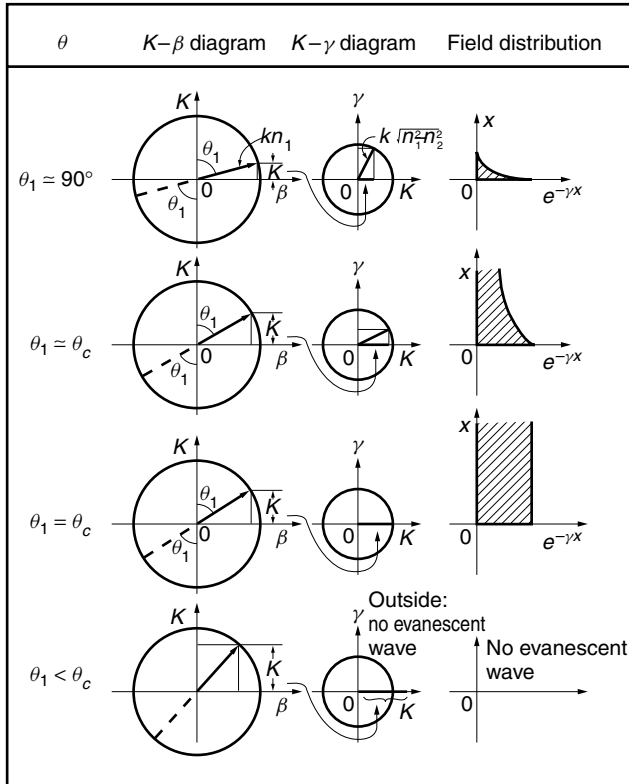


Figure 2.22 $K-\beta$ and $K-\gamma$ diagrams for the graphical solution of the evanescent wave.

to the K axis. The $K-\beta$ diagram provides the values of K and β directly. From the value of K thus obtained, the $K-\gamma$ diagram gives the value of γ and also determines whether or not total internal reflection takes place. Thus, the two circles combined provide γ for a given θ_1 . Figure 2.22 demonstrates how to use the diagram based on what has already been calculated in Section 2.8.3. The diagrams are arranged in descending order of the incident angle θ_1 in the optically dense medium.

In the top row of the diagram, θ_1 is large and is close to 90° . From the $K-\beta$ diagram, K is small. For a small value of K , the $K-\gamma$ diagram provides a large value of γ . A large value of γ means an evanescent wave, which decays very fast with distance away from the boundary. The corresponding situation is labeled ① in Fig. 2.20.

In the diagrams in the second row of Fig. 2.22, θ_1 is reduced to a value just slightly above the critical angle. Compared to the top row diagrams, K increases and hence γ decreases, and the evanescent wave extends further into the optically less dense medium. This corresponds to cases ② and ③ in Fig. 2.20.

The diagrams in the third row explain what happens when $\theta_1 = \theta_c$. The value of K becomes exactly the same as $k\sqrt{n_1^2 - n_2^2}$ and $\gamma = 0$. The field does not decay with x and the energy is lost into the less dense medium.

In the bottom row diagrams, θ_1 is reduced to a value below the critical angle, $\theta_1 < \theta_c$. The value of K now becomes large and falls outside the $K-\gamma$ diagram. This means that no evanescent wave exists.

2.9 WHAT GENERATES THE EVANESCENT WAVES?

So far, only the evanescent wave that was generated by total internal reflection has been discussed. Is this the only way to generate an evanescent wave? In this section, the evanescent wave will be treated in a more general manner.

2.9.1 Structures for Generating Evanescent Waves

Let us consider a medium with index of refraction n_i . As with Eqs. (2.71) and (2.72), solving the wave equation in this medium requires that

$$k_x^2 + k_y^2 + k_z^2 = (n_i k)^2 \tag{2.106}$$

If $k_x^2 + k_y^2 + k_z^2$ were bigger than $(n_i k)^2$, in order to satisfy the boundary condition, at least one of k_x^2 , k_y^2 , or k_z^2 has to become negative, and the corresponding k_x , k_y , or k_z becomes an imaginary number. There is an evanescent wave in that direction. For the geometry so far discussed, k_{2z} was made so large (in order to meet the phase match boundary condition in medium 1) that k_{2z} had to be an imaginary number.

A corrugated metal surface is another example of a geometry that supports an evanescent wave [9]. The corrugation is oriented in the z direction, and the teeth in the x direction, as shown in Fig. 2.23a. Assume there is no variation in the y direction, meaning $\lambda_y = \infty$ and $k_y = 0$. A further assumption is made that the electric field is zero on the contour of the corrugation.

In order to match the boundary conditions, a number of component waves have to be summed. We realize immediately that the field distribution of the component

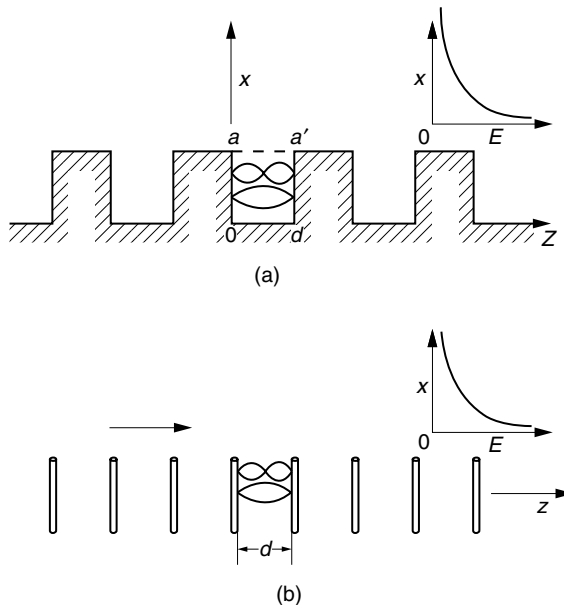


Figure 2.23 Structures that support evanescent waves. (a) Corrugated metal surface. (b) Array of metal pins.

waves inside the slot space is proportional to a trigonometric function so as to satisfy the condition of vanishing fields on both sides of the wall. To satisfy the boundary condition, an integral multiple of a half-wavelength $\lambda_{zn}/2$ should fit the wall dimension d :

$$\frac{\lambda_{zn}}{2}n = d \quad \text{or} \quad k_{zn} = \frac{\pi}{d}n \quad (2.107)$$

where n is an integer and the order number of the component waves. Note that k_{zn} , which describes the propagation constant of the n th order component wave in the vicinity of $a-d'$, can be increased by reducing d . From Eq. (2.106), we have

$$k_{xn}^2 = \left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{\pi}{d}n\right)^2 \quad (2.108)$$

where k_{xn} is the propagation constant of the n th order component wave in the x direction.

The lower order component waves can propagate in both the x and z directions because all values in Eq. (2.106) are real. For the higher order components whose n is larger than a certain value, k_x becomes an imaginary number. In this case, the wave is an evanescent wave in the x direction and a propagating wave in the z direction. As a matter of fact, when d is smaller than one-half of the free-space wavelength, all component waves become evanescent waves.

Another structure that supports evanescent waves is shown in Fig. 2.23b. It is an array of conductors. Again, on the surface of the conductors the field has to be zero. By the superposition of many component waves with various propagation constants, this boundary condition is met. Some of the component waves have to satisfy the condition of Eq. (2.107) and become evanescent waves. These structures are used as evanescent waveguides for both microwaves and light waves.

Example 2.5 A lossy glass is bordered by air in the $x = 0$ plane. The propagation constant in the z direction on the border is $\beta + j\alpha$. Find the expression for the evanescent wave in the air near the boundary. Assume $k_y = 0$.

Solution Let the lossy glass be medium 1, and the air be medium 2. The propagation constants in air have to satisfy

$$k_{2x}^2 + k_{2z}^2 = k^2 \quad (2.109)$$

while the propagation constant along the boundary in the z direction in the glass is $\beta + j\alpha$. Phase matching along the z direction requires that

$$k_{2z} = \beta + j\alpha \quad (2.110)$$

Inserting Eq. (2.110) into (2.109) gives

$$k_{2x}^2 = k^2 - \beta^2 + \alpha^2 - j2\alpha\beta \quad (2.111)$$

The real part K and imaginary part γ of k_{2x} will be found. Putting

$$\begin{aligned} A &= k^2 - \beta^2 + \alpha^2 \\ B &= 2\alpha\beta \end{aligned} \quad (2.112)$$

gives

$$k_{2x}^2 = \sqrt{A^2 + B^2} e^{j(-2\phi + 2n\pi)}$$

where $\tan 2\phi = (B/A)$ and n is an integer. Depending on n being even or odd, the value of k_{2x} is

$$k_{2x} = \begin{cases} K - j\gamma, & n \text{ is even} \\ -K + j\gamma, & n \text{ is odd} \end{cases} \quad (2.113)$$

where

$$\begin{aligned} K &= \sqrt[4]{A^2 + B^2} \cos \phi \\ \gamma &= \sqrt[4]{A^2 + B^2} \sin \phi \end{aligned}$$

Finally, the expressions for the evanescent field become

$$E_2 = E_{20} e^{\gamma x - \alpha z + j(Kx + \beta z - \omega t)} \quad (2.114)$$

or

$$E_2 = E_{20} e^{-\gamma x - \alpha z + j(-Kx + \beta z - \omega t)} \quad (2.115)$$

Equation (2.114) is not acceptable as a solution because E_2 increases indefinitely with an increase in x . Figure 2.24 illustrates the field distribution expressed by Eq. (2.115). Comparing Fig. 2.24 to the distribution for lossless glass shown in Fig. 2.14, the entire

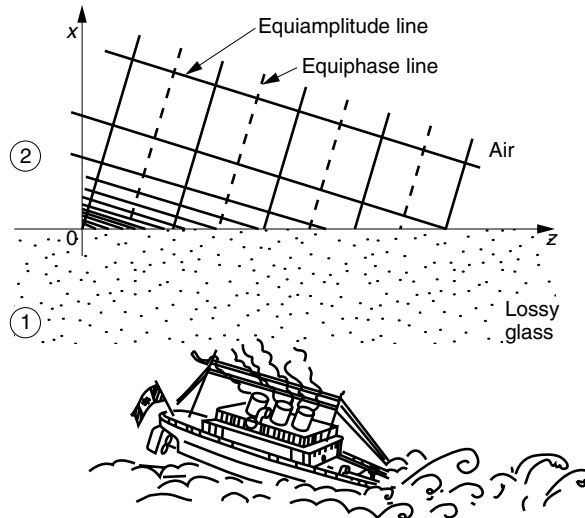


Figure 2.24 Field distribution of a leaky wave.

distribution is tilted and looks as if the wave is sinking into the glass to supply the energy lost in the glass. The angle of the tilt depends on $-K$, and K vanishes when $\alpha = 0$. This kind of wave associated with the surface is called a leaky wave, meaning that it loses energy as it propagates. \square

2.10 DIFFRACTION-UNLIMITED IMAGES OUT OF THE EVANESCENT WAVE

In Chapter 1, the observation regions of the diffraction pattern were classified into two areas: far field and near field. The near field in this classification, however, means that the distance between the aperture and the observation region is still large compared to the light wavelength. In this section, the region of concern is within a few wavelengths of the aperture. It would be more appropriate to call this region the very-near-field region. To avoid confusion, we will use the term “very near field” whenever necessary, but commonly used proper names like SNOM (scanning near-field optical microscope) will be left unchanged.

A lens-type microscope forms a magnified image out of the radiating light wave by means of lenses. The resolution of the microscope, however, is limited to the order of one wavelength of the light.

On the other hand, the evanescent-field-type microscope forms the image out of the evanescent wave generated by the object. An evanescent field to radiating field converter is used to form the image. The resolution of the microscope is 10–20 times better than lens-type microscopes. The resolution of the lens-type microscope is first reviewed in the next section.

2.10.1 Resolution of a Lens-Type Microscope

The minimum detectable variation of a microscope objective lens will be calculated. For the geometry shown in Fig. 2.25, a grating is used as the object. Assume that the grating is illuminated by a parallel laser beam and the light distribution on the surface of the grating is expressed by

$$E(x) = E_0(1 + \cos 2\pi f_g x) \quad (2.116)$$

where f_g is the spatial frequency of the grating. The spatial frequency is the inverse of the period t ; that is,

$$f_g = \frac{1}{t} \quad (2.117)$$

For the simplicity at the expense of accuracy, the far field approximation will be used. Inserting Eq. (2.116) into (1.36) gives

$$E(\theta) = K \left[\delta \left(\frac{\sin \theta}{\lambda} \right) + \delta \left(\frac{\sin \theta}{\lambda} - f_g \right) + \delta \left(\frac{\sin \theta}{\lambda} + f_g \right) \right] \quad (2.118)$$

where K absorbs the necessary amplitude and phase factors. Besides the center radiation lobe, Eq. (2.118) gives two side lobes in the directions of

$$\theta = \pm \sin^{-1} \lambda f_g \quad (2.119)$$

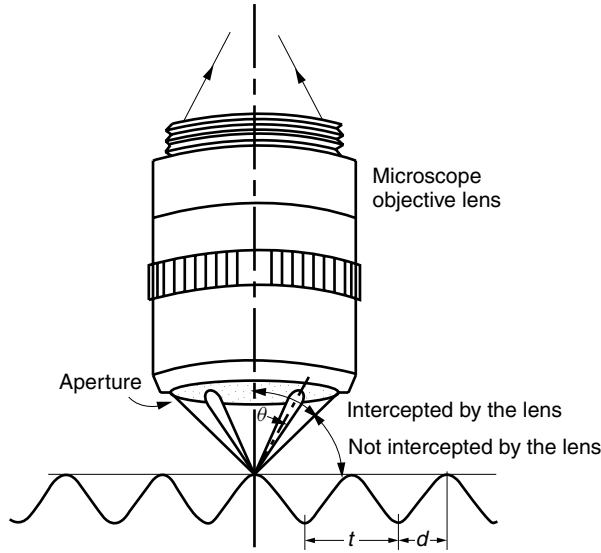


Figure 2.25 Radiating field from the sinusoidal grating into the microscope objective lens.

Note that as the spatial frequency f_g of the grating increases from zero, the direction of radiation of the two side lobes moves from normal to parallel to the grating.

As shown in Fig. 2.25, only the portion of the scattered light that is intercepted by the objective lens contributes to the formation of the image. Then, the maximum spatial frequency f_{gM} of the grating whose radiation lobe can still be intercepted by the aperture of the lens is, from Eq. (2.119),

$$\sin \theta_M = \lambda f_{gM} \tag{2.120}$$

where θ_M is the angle that the aperture subtends with respect to the normal of the grating. The quantity $\sin \theta_M$ is called the numerical aperture (NA) of the objective lens and Eq. (2.120) is rewritten as

$$f_{gM} = \frac{\text{NA}}{\lambda} \tag{2.121}$$

The surface variation d is the distance between the peak and valley of the sinusoidal variation of the grating. The surface variation is expressed as

$$d = \frac{t}{2} \tag{2.122}$$

Then, the minimum detectable variation d_m is, from Eq. (2.117), (2.121), and (2.122),

$$d_m = \frac{\lambda}{2\text{NA}} \tag{2.123}$$

Equation (2.123) is known as Rayleigh’s resolution criteria of a diffraction-limited lens (a lens with a perfect shape whose resolution is limited only by the finite size of the lens). The conclusion is that $\sin \theta_M$ cannot be larger than unity and the best lateral (in

the x direction) resolution d_m that can be achieved is one half-wavelength. This is the case when the radiating field alone is used for forming the image. Is there any other means of forming a light image and obtaining a resolution better than a half-wavelength of light? The answer is yes, as will be shown next.

Let us turn our attention from the microscope objective to the field scattered out of the grating. Inserting Eq. (2.117) into (2.119) gives the direction of the radiation lobe in terms of the period of the grating:

$$\theta = \sin^{-1} \left(\frac{\lambda}{t} \right) \quad (2.124)$$

What happens if the period of the grating becomes shorter than one wavelength? There is no θ that satisfies Eq. (2.124), meaning there is no radiating field. Still, the boundary conditions have to be satisfied along the contour of teeth of the grating. The evanescent wave coasts the contour of the teeth and appears only in the region within a few wavelengths of light from the surface and decays away exponentially with the distance from the surface. Roughly speaking, information about variations longer than one half-wavelength of light is carried by the radiating portion of the field, while that of variations shorter than a half-wavelength, by the evanescent portion of the field.

The mathematical formulation of the evanescent wave has already been given by Eq. (1.197) in connection with Eq. (1.201) or by Eq. (2.101). Optical microscopes have been devised to collect the evanescent field and convert it into an image. These microscopes have achieved resolutions much shorter than a half-wavelength of light.

2.10.2 Near-Field Optical Microscopes

The microscopes whose operation is based on the evanescent wave may broadly be divided into two categories: the photon tunneling microscope (PTM) and the scanning near-field optical microscope (SNOM) [10,11].

2.10.2.1 Photon Tunneling Microscope

Figure 2.26 shows the geometry of the photon tunneling microscope [12]. Input light is incident from the convex side of the plano-convex objective lens, so that the condition of total internal reflection is satisfied on the flat bottom surface of the flexible transducer placed beneath the flat side of the plano-convex objective lens. The evanescent wave is excited in the region of the tunneling gap. The reflected light is fed to the imaging system of a video camera through the other side of the plano-convex objective lens. The photons lost into the sample due to the short tunneling gap are responsible for the reduction in the detected signal, which leads to information about the depth of the tunneling gap.

The depth pattern of the sample is displayed as the brightness distribution in the reconstructed image. The lateral resolution of the photon tunneling microscope is diffraction limited in the same manner as a normal imaging system, but as far as the depth information is concerned, the exponential decay of the evanescent wave is used and the resolution is limited only by the gray scale of the video camera.

Figure 2.27 shows the depth image (gray-scale image) of a single polyethylene crystal and the processed 3D image of the same crystal when measured by a photon tunneling microscope [13].

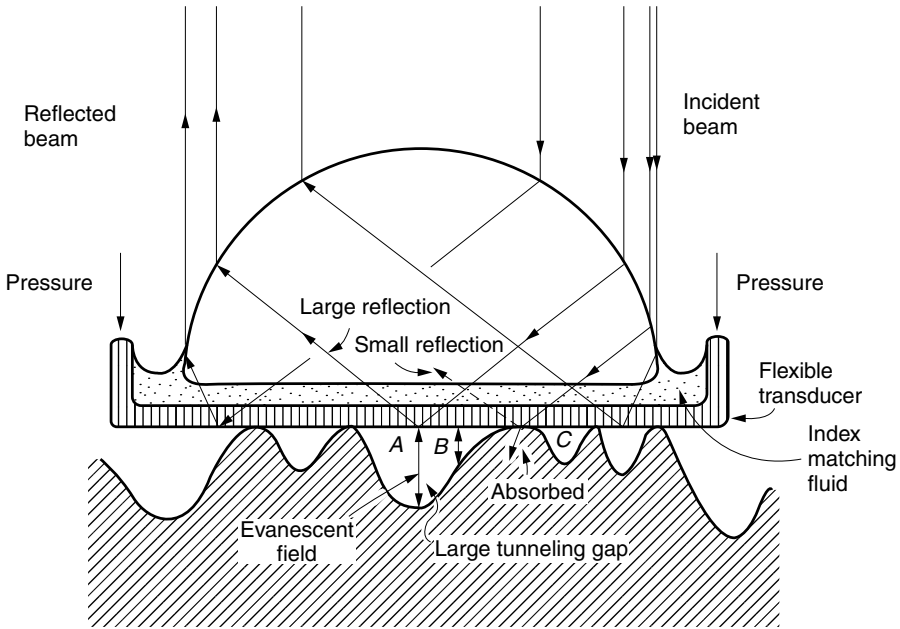


Figure 2.26 Geometry of the photon tunneling microscope.

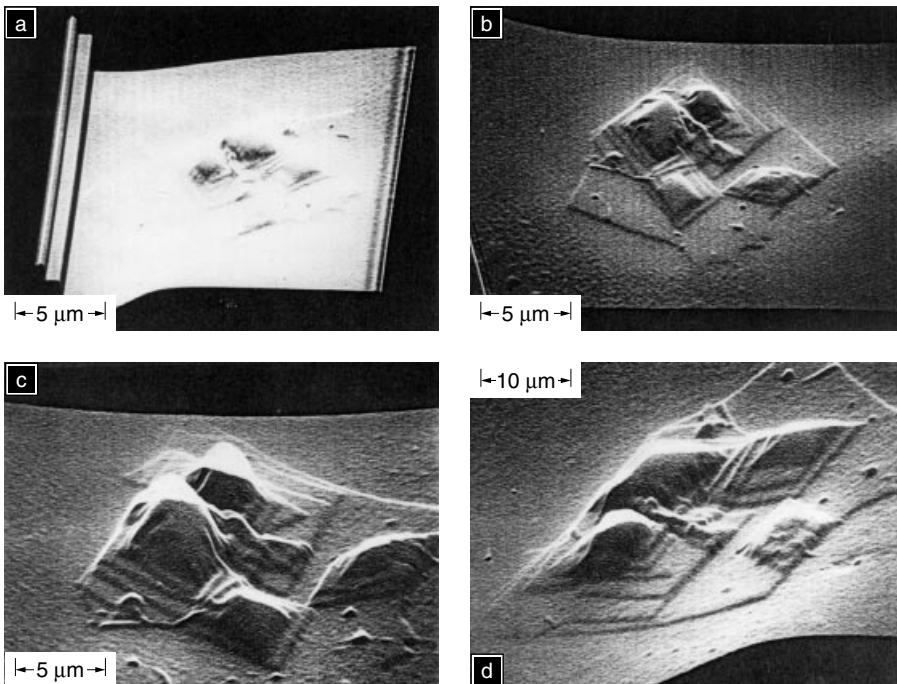


Figure 2.27 (a) Gray-scale PTM image of a single polyethylene crystal. (b,c) Three-dimensional (3D) images showing the topography from different viewpoints. (d) A single large crystal can be imaged with the PTM in contact interference mode. (Courtesy of J. M. Guerra, M. Srinivasarao, and R. S. Stein [13].)

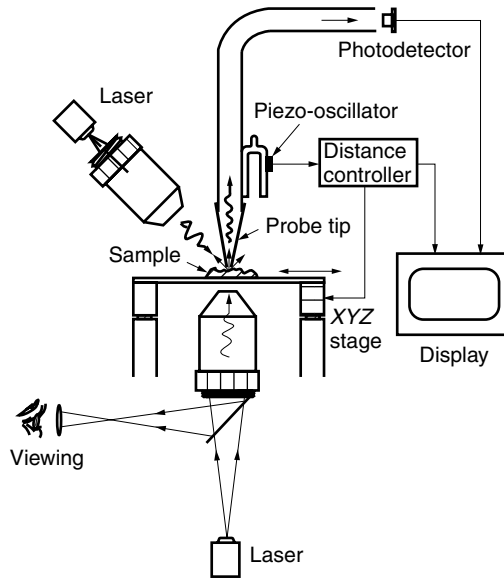


Figure 2.28 Schematics of the scanning near-field optical microscope (SNOM).

2.10.2.2 Scanning Near-Field Optical Microscope (SNOM)

Figure 2.28 shows the geometry of the SNOM [14–16]. Images of subwavelength resolution are obtainable. The surface of the sample can be illuminated from above, underneath, or both by ordinary radiating laser light. The amplitude of the resultant radiating and evanescent fields is collected with a 50–500-nm diameter probe at a distance of 10–60 nm from the sample surface.

The evanescent field decays exponentially with distance away from the surface, and information about the surface variation is lost as soon as the probe is outside the very near field. Placing the tip in the very near field without hitting and possibly damaging the sample is a challenge. There are more than a few techniques for preventing the probe from crashing into the sample. When a voltage is applied between the metalized tip and the sample, an electron tunnel current starts to flow as soon as the two are brought close together. From the tunneling current the clearance is monitored. Another technique is to monitor the change in the frequency of the mechanical vibration that takes place when the probe is placed in proximity to the sample surface. Another method is by the change of the capacitance between the tip and the sample, or by the change of the shear force between the tip and the sample. As a matter of fact, the shear force itself can be used to image the surface [17].

Figure 2.29 is the image of deoxyribonucleic acid, commonly known as DNA, obtained by shear force imaging [18].

2.10.3 Probes to Detect the Evanescent Field

We will now explain how the probe converts the evanescent field into a radiating field [19,20]. The evanescent field that is established on the boundary between the optically dense and less dense media is used as an example. The direction of the polarization of the wave is assumed to be perpendicular to the page or an s wave.

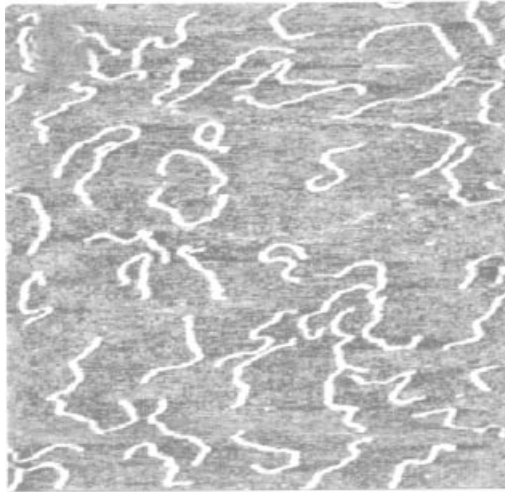


Figure 2.29 Shear-force image of DNA. (Courtesy of M. F. Garcia-Parajo et al. [18].)

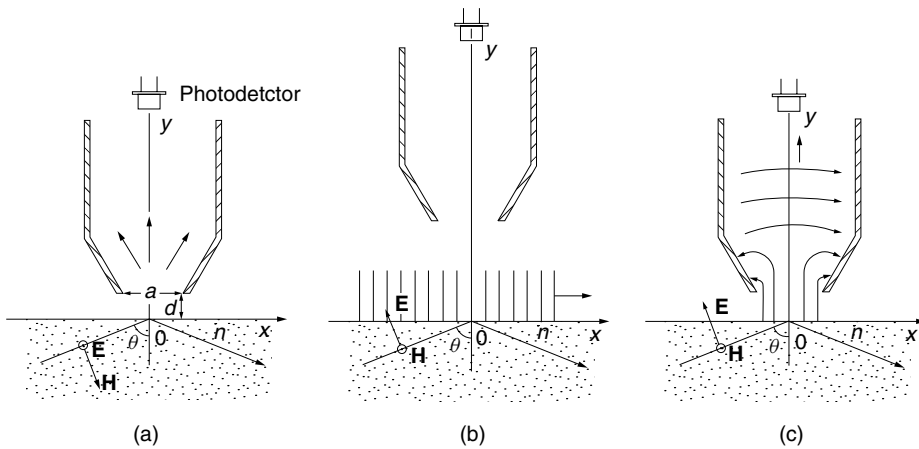


Figure 2.30 Probe aperture excited by evanescent waves. (a) Probe excited by s wave. (b) Probe above the p wave. (c) Probe excited by the p wave.

As shown in Fig. 2.30a, the two-dimensional model is considered. The x axis is taken in the direction of the boundary and the y axis, vertically perpendicular to it. The probe is placed with its aperture parallel to the boundary at a distance d away from the boundary. The evanescent field in which the aperture of the probe is immersed is, from Eq. (1.197),

$$E(d, z) = E_0 \exp[-2\pi f_s \sqrt{n^2 \sin^2 \theta - 1} d + j2\pi n f_s (\sin \theta) x] \tag{2.125}$$

where n is the refractive index of the more dense medium, and the less dense medium is air. The angle θ is the angle of incidence of the light inside the more dense medium

to the boundary. The field across the aperture is

$$E_a = \Pi\left(\frac{x}{a}\right) E(x, d) \quad (2.126)$$

where a one-dimensional aperture with width a is assumed. The field reradiated toward the photodetector due to E_a will be found using Eq. (1.177). For this purpose, the Fourier transform of Eq. (2.126) is calculated:

$$\epsilon(f_x, 0) = aA \operatorname{sinc}[a(f_x - f_0)] \quad (2.127)$$

where

$$A = E_0 \exp(-2\pi f_s \sqrt{n^2 \sin^2 \theta - 1} d)$$

and

$$f_0 = n f_s \sin \theta > f_s$$

where the inequality sign is due to the evanescent wave.

The radiating field component of Eq. (2.126) is distinguished from that of the evanescent field component before the equation is put into Eq. (1.177). The spectrum of Eq. (2.127) is plotted in Fig. 2.31. As explained earlier in Section 2.10.1, whether the component is radiative or evanescent is determined by whether the spatial frequency is smaller or larger, respectively, than $f_s (= 1/\lambda)$. The hatched region in Fig. 2.31 is radiative, and the unhatched region is evanescent. When a photodetector is placed on the y axis at (x, y) , the radiating field received by the detector is obtained by inserting Eq. (2.127) into (1.177):

$$E(x, y) = aA \int_{-f_s}^{f_s} \operatorname{sinc}[a(f_x - f_0)] e^{j2\pi \sqrt{f_s^2 - f_x^2} (y-d)} e^{j2\pi f_x x} df_x \quad (2.128)$$

The propagation medium between the aperture and the detector is assumed to be free space rather than an optical fiber.

The proper choice of the aperture size of the probe is crucial. Let's first consider what happens if the aperture is too narrow, as shown in Fig. 2.31a. The factor a in front of the integral in Eq. (2.128) becomes small. In addition, the sinc function is spread so wide that the relative portion of the radiative component, which is inside the $-f_s$ to f_s region, is small compared to the entire region. Consequently, the radiating field reaching the detector is weak, and the signal-to-noise ratio of the image deteriorates even though a narrow aperture provides a higher resolution as a microscope. On the other hand, if the aperture is chosen too wide, such as shown in Fig. 2.31b, the sinc function shrinks around $f_x = f_0$. The radiative portion, which is inside $-f_s$ to f_s , is reduced, and the radiating field reaching the detector is again weak. In this case, the spatial resolution of the microscope deteriorates as well.

Next, let's consider the case when the direction of polarization is parallel to the page, or a p wave as shown in Figs. 2.30b and 2.30c. When the aperture stays high above the boundary as shown in Fig. 2.30b, the E field of the evanescent wave is undisturbed. The direction of the E field is parallel to the y axis, and certainly there is no propagation in the y direction because the electromagnetic wave does not propagate in the same

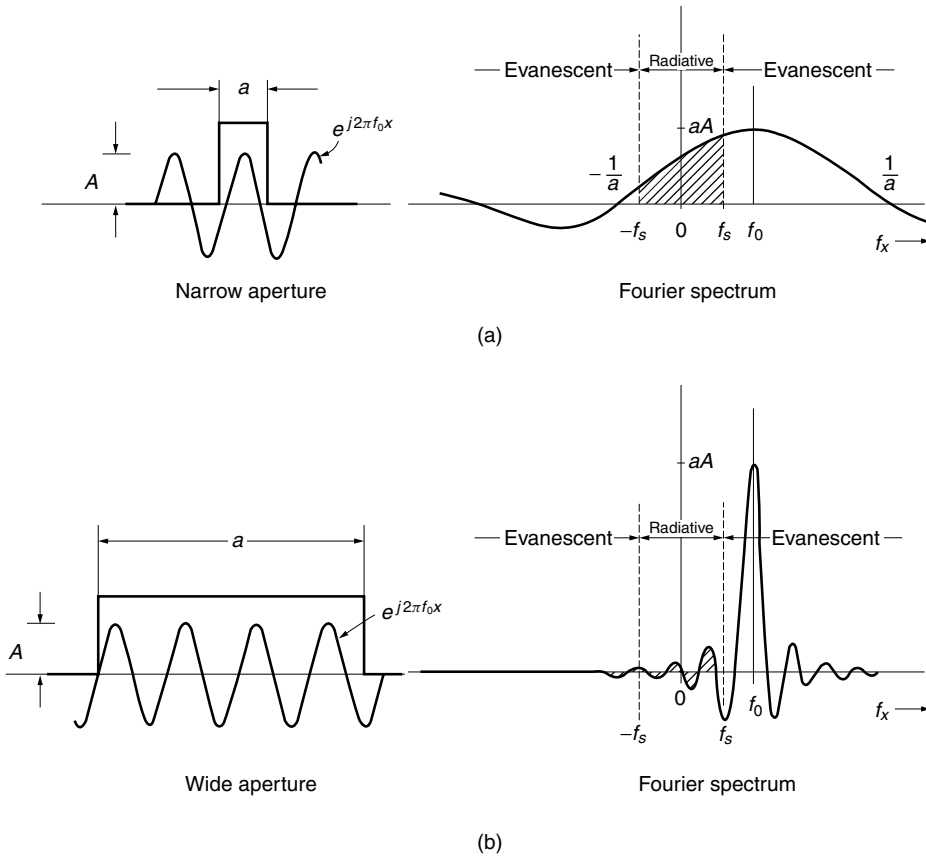


Figure 2.31 Aperture size of the probe and Fourier transform spectrum. (a) A case of narrow aperture. (b) A case of wide aperture.

direction as either the E or H field, but propagates in the direction perpendicular to either the E or H field.

When the aperture is lowered to the boundary, however, the electric lines of force start to curl so that they are terminated perpendicular to the metal surface of the probe in order to satisfy the boundary condition of the E field, as shown in Fig. 2.30c. Now, this curled E field has a component in the x direction and provides the possibility of creating a radiating wave toward the y direction. Note, however, that if the shape of the probe is symmetric, then half of the lines of force are bent toward the positive x direction while the other half are bent toward the negative x direction. One would think that the contributions of the lines of force curled in opposite directions would cancel each other, thereby failing to excite a radiating wave in the y direction. This thinking is not quite correct. Remember that there is a phase lag between the left and right lines of force, and they do not quite cancel each other. The difference contributes to a weak radiating wave toward the y direction. It is safe to say that the probe in this configuration is more sensitive to the s wave than to the p wave.

In short, the probe aperture redistributes the evanescent wave energy into radiative and evanescent field components, and the energy that has been converted into the radiative component reaches the photodetector.

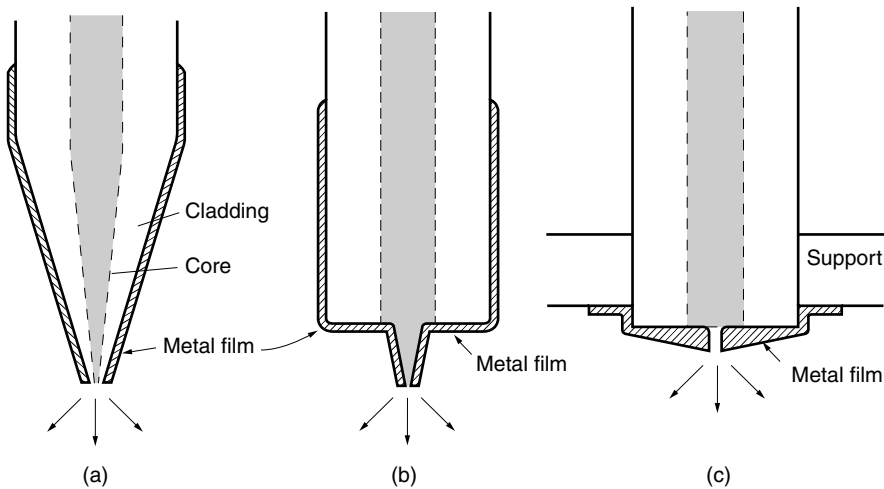


Figure 2.32 Various apertures of SNOM probes at the tip of the optical fiber. (a) Heating and pulling. (b) Selectively etched. (c) Flat tip.

2.10.4 Apertures of the SNOM Probes

The SNOM probes were first fabricated by applying a heating-and-pulling method to an optical fiber. The tip was either bare or coated with such metals as silver (Ag), aluminum (Al), or chromium (Cr). Probes of this type, however, have a long tapered section, such as shown in Fig. 2.32a. The transmission loss is usually high and is in the vicinity of 50 dB for an aperture diameter of 100 nm.

A shorter tip section, such as shown in Fig. 2.32b, is made by chemically etching the tip of the optical fiber using an aqueous mixture of hydrofluoric acid (HF) and ammonium fluoride (NH_4F). The strength of the etchant and the immersion time control the finish of the probe. A smoother etched surface is obtainable by dipping the fiber without removing the acrylate jacket [21]. As the final step, the etched surface is coated with a metal film.

The flat tip shown in Fig. 2.32c almost completely eliminates the tapered section. The metal film is coated directly onto the probe (Fig. 2.32c), the flat end of the optical fiber, and then the film is drilled by a focused ion beam (FIB) [22].

The diameter of the drilled aperture is approximately 50–100 nm. The transmission loss is around 30 dB for an aperture diameter of 100 nm. This is an improvement of 20 dB over the tapered probe.

2.10.5 Modes of Excitation of the SNOM Probes

Five different modes of operation of the SNOM probe are listed in Fig. 2.33. [23]. The collection mode shown in Fig. 2.33(a) is the mode that has been dealt with so far. The probe is used to pick up the field scattered from the target.

If the light path of the collection mode is reversed, the result is the illumination mode shown in Fig. 2.33b. In the illumination mode, the optical fiber is driven by a light source and the target is illuminated locally as the probe is scanned, analogous to using a flashlight to illuminate a large object. The advantage of this mode is the minimization of the background light.

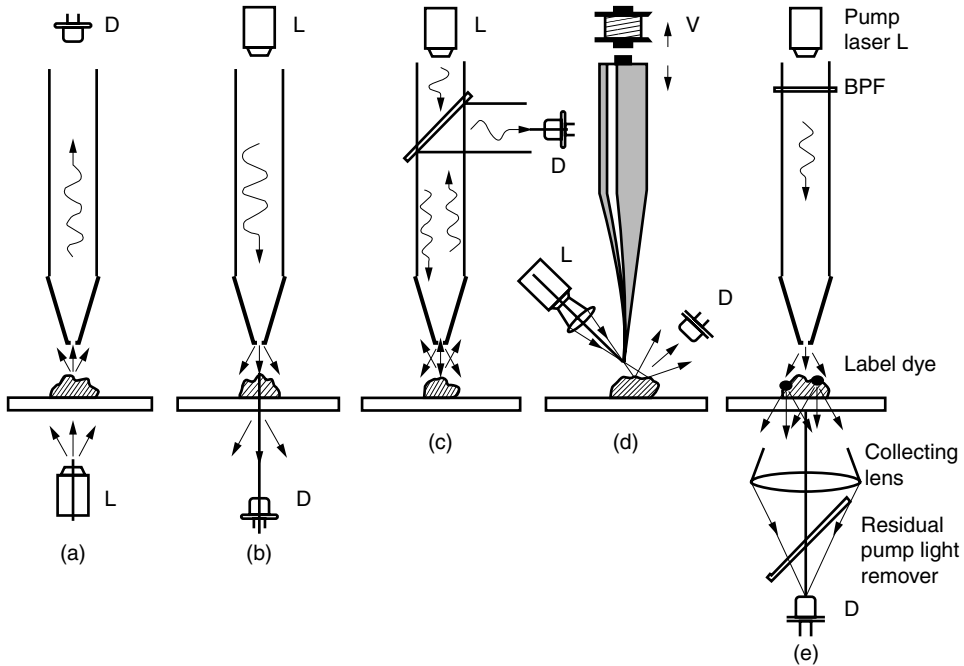


Figure 2.33 Various modes of SNOM operation. (a) Collection. (b) Illumination. (c) Reflection. (d) Apertureless. (e) Fluorescence. L, laser; D, photodetector; V, vibrator; BPF, band-pass filter.

The reflective mode shown in Fig. 2.33c is the combination of the above two modes using just one probe. The probe driven by a laser illuminates the object. The field scattered from the object is collected by the same probe. A half-mirror separates the transmitted and received light. The resultant directivity is the product of the directivities of the collection and illumination modes, and the resultant directivity is narrower than the individual directivity.

The apertureless mode is shown in Fig. 2.33d. The probe is made out of a solid metal wire with a sharp tip. The tip of the probe is mechanically dithered along its axis. The tip is illuminated with laser light, and because of the probe's dithering, the wave scattered from the probe is modulated. The photodetector exclusively detects the modulated signal. The unmodulated background light is eliminated. Because the probe does not contain an aperture, the tip can be made small to minimize the unwanted disturbance of the field due to the presence of the probe.

With all modes of operation, the illuminating light can be either an evanescent or a radiating wave, or even a combination of the two. The illuminating light does not have to be an evanescent wave. However, for the microscope to obtain a subwavelength resolution, the scattered field has to contain the evanescent field.

When a fluorescent substance is illuminated by light of a specific wavelength (pump light), light with wavelengths longer than the pump light is reemitted due to interactions occurring between light and matter. Figure 2.33e is an example of a SNOM probe in the illumination mode that utilizes this fluorescent spectroscopy.

By chemically treating the sample, it is possible to label only one fluorophore per molecule [18]. The area of illumination of the pump light through the SNOM probe is

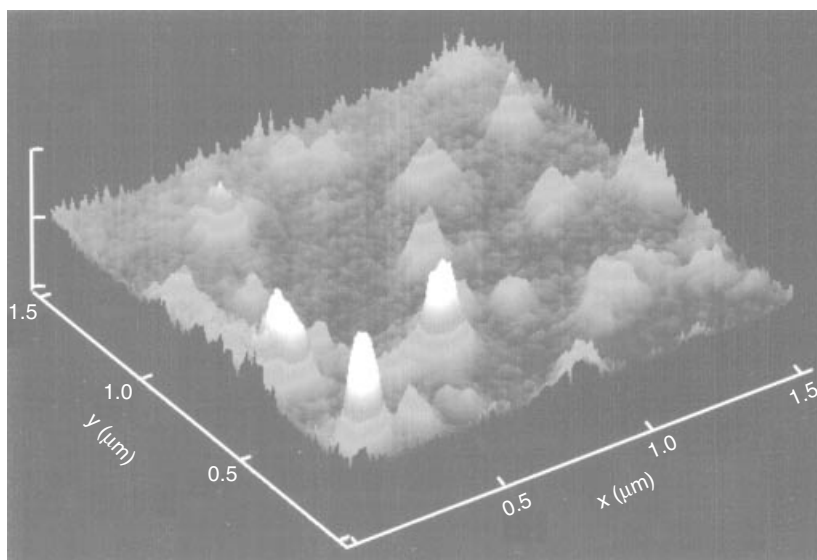


Figure 2.34 Near-field fluorescence image of DNA fragments labelled with rhodamine dye. (Courtesy of M. F. Garcia-Parajo et al. [18].)

restricted to a diameter of approximately 100 nm. Figure 2.34 shows the fluorescent image of a double-stranded DNA fragment labeled at one end with rhodamine dye. Thus, the spectroscopic use of the SNOM makes it possible to monitor the structural changes of a single biomacromolecule.

2.10.6 SNOM Combined with AFM

Figure 2.35 shows a SNOM combined with an atomic force microscope (AFM). In the AFM, a sharp probe tip is attached to the bottom side of the cantilever, and a small mirror is attached on the top side. The orientation of the cantilever is monitored by the change in the direction of the laser beam reflected from the small mirror.

When the tip of the probe is brought closer than several nanometers to the sample surface, an atomic force is experienced between the probe tip and the sample surface in accordance with the distance between them. The cantilever is excited to vibrate by a piezoelectric crystal. The phase and amplitude of the vibration are monitored by the laser beam reflected from the small mirror. The phase and amplitude of vibration vary with the distance of the tip to the sample surface because the atomic force varies. The phase and amplitude are maintained constant as the cantilever is scanned laterally over the sample surface by means of a feedback loop current to another piezoelectric crystal that raises and lowers the sample stage. At the same time, this feedback loop current provides information about the movement of the tip as it coasts the sample surface, and this information is used to form the AFM image.

In Fig. 2.35a, a collection mode SNOM is combined with an AFM. The photodiode is incorporated into the tip of the probe, and the SNOM signal is directly detected by the photodiode. Images obtained by mechanical means, such as by the shear force or atomic force microscopes, are often called topographic images as compared to the images obtained by optical means.

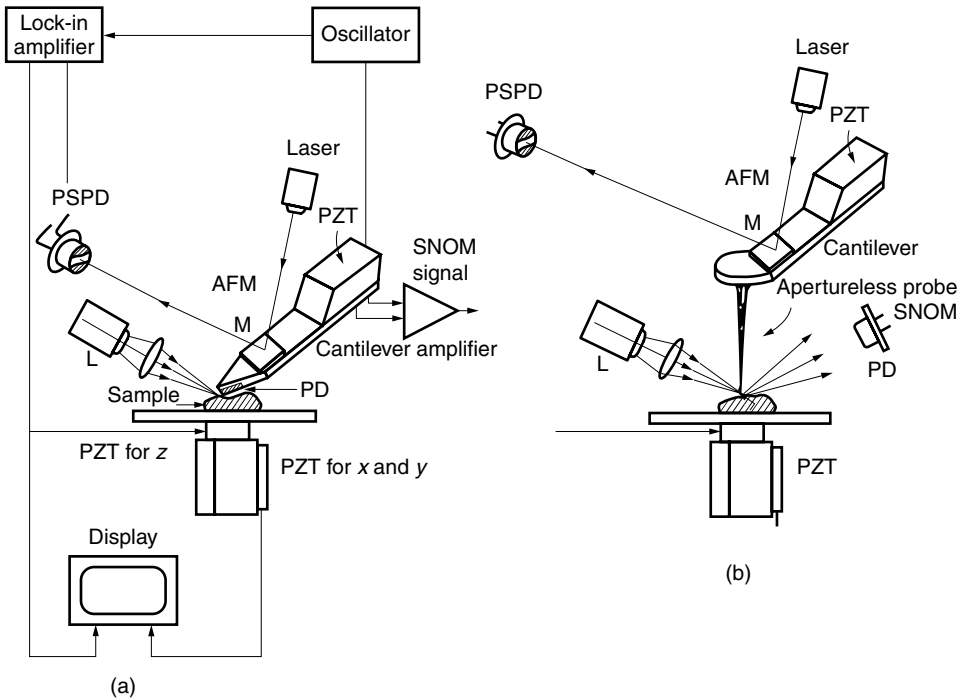


Figure 2.35 Combination of SNOM with the AFM. (a) Collection mode SNOM with AFM. (b) Apertureless SNOM with AFM. PSPD, position-sensitive photodetector; M, mirror; PZT, piezoelectric transducer; PD, photodetector.

In Fig. 2.35b, an apertureless mode SNOM is combined with the AFM. The AFM tip can be used as the apertureless optical probe with minimum modification [24].

Figure 2.36 compares images of the same test sample taken by the AFM and the SNOM in the apertureless mode.

2.10.7 Concluding Remarks

The resolving power of the SNOM is of the order of tens of nanometers, while that of an electron microscope is of the order of a nanometer. Even though the ultimate resolution of the near-field optical microscopes is poorer than that of the scanning electron microscope (SEM), the advantages of the photon tunneling microscope are that it neither requires metallizing the sample nor causes the intrusive effects of electrons to the sample. The very-near-field optical microscope is a valuable tool for revealing faults, especially during the fabrication of optical devices, because the detected signal is directly related to the severity of the faulty function.

The usefulness of the SNOM goes beyond that of being a mere microscope. When combined with optical spectroscopy, the SNOM becomes a unique tool for studying the local interaction of light with matter, even to the extent of allowing the *in vivo* observation of the movement of a single biomacromolecule.

Another area of importance is the application of the SNOM to the development of a high-density data storage. In the digital video disk (DVD) recorder, a focused beam of laser light is used to read or write the marks and spaces (or “1” and “0” bits) in

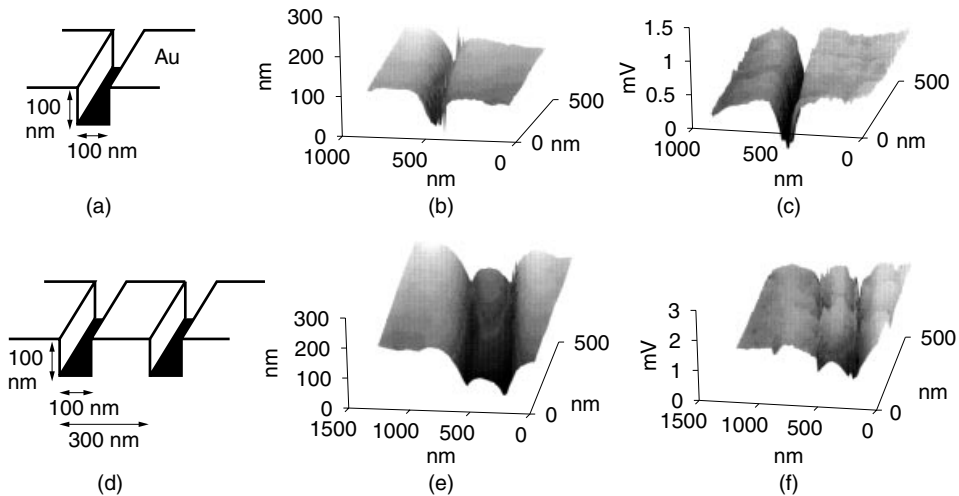


Figure 2.36 Comparison of images taken by AFM and SNOM. (a) Single-groove test sample engraved by X-ray lithography. (b) Topographic image of the single groove determined by AFM. (c) Optical image of the single groove determined by SNOM (amplitude). (d) Double-groove test sample engraved by X-ray lithography. (e) Topographic image of the double groove determined by AFM. (f) Optical image of the double groove determined by SNOM (amplitude). (After R. Bachelot, P. Gleyzes, and A. C. Boccarda [24].)

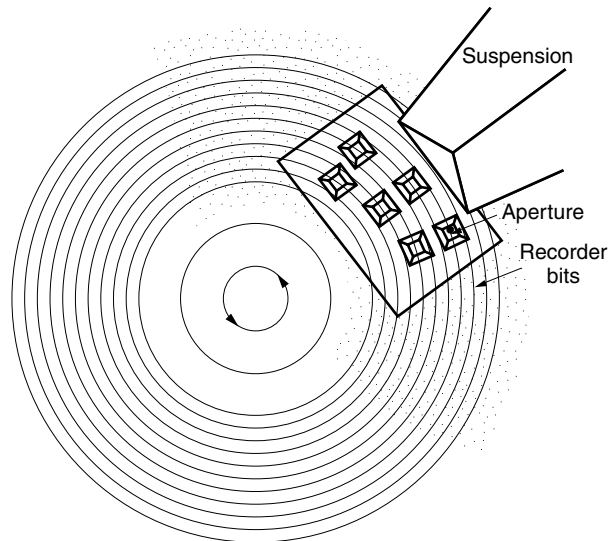


Figure 2.37 Multichannel flat chip used as a scanner over a high-density memory disk. (After M. Kourogi et al. [25].)

memory. Being a focused light beam, the narrowest spacing between the marks and spaces cannot be less than a wavelength of the light. If, however, the light from the SNOM probe is used for writing or reading the memory, the required spacings between codes can be reduced by a factor of 10. The memory area density can be increased by 100 times. The difficulty, however, is how to maintain a 10-nm separation between

the probe and the memory disk, which is spinning at a rate equivalent to 1 m/s in linear distance speed. This problem can be alleviated by the multichannel flat chip approach, such as shown in Fig. 2.37. On a flat silicon chip ($30 \mu\text{m} \times 150 \mu\text{m}$) an array of 80-nm apertures are micromachined, and this chip is slid over the oiled memory disk surface. The flat bottom surface of the chip prevents the crash as well as vibration. The multichannel approach reduces the required speed of spinning the memory disk [25].

PROBLEMS

- 2.1 Derive the reflection coefficient r_{\parallel} and transmission coefficient t_{\parallel} for the **H** field.
- 2.2 Derive the following equation:

$$r_{\parallel} = \frac{\tan(\theta_1 - \theta_2)}{\tan(\theta_1 + \theta_2)}$$

- 2.3 Explain why $1 + r_{\parallel} \neq t_{\parallel}$ even though $1 + r_{\perp} = t_{\perp}$ is true.
- 2.4 Prove that $R + T = 1$ for both directions of polarization.
- 2.5 From Eq. (2.42),

$$r_{\perp} = -\frac{\sin(\theta_1 - \theta_2)}{\sin(\theta_1 + \theta_2)}$$

Prove that a wave with perpendicular polarization cannot have Brewster's angle.

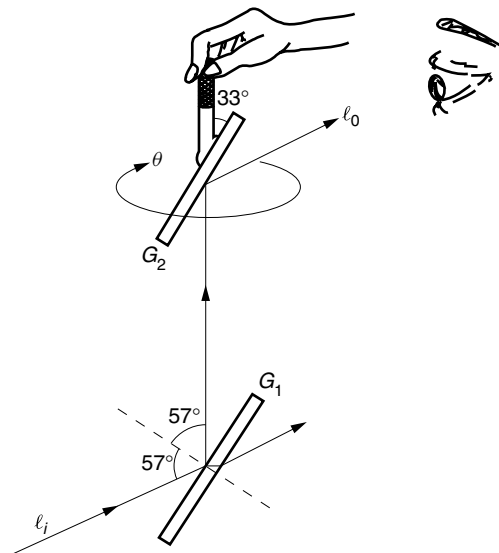


Figure P2.6 When G_1 and G_2 are parallel, light ℓ_0 is reflected; however, when G_2 is rotated 90° from this position, the light ℓ_0 disappears.

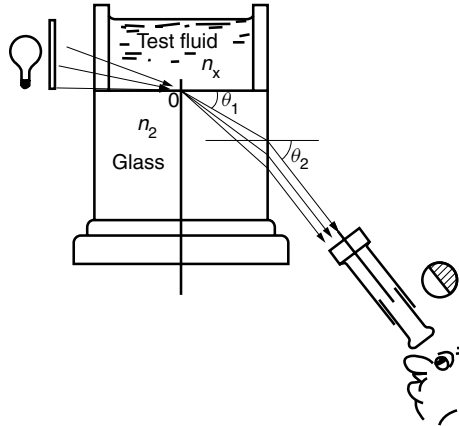


Figure P2.7 Pulfrich's refractometer.

- 2.6 As shown in the Fig. P2.6, glass plates G_1 and G_2 are initially set parallel to each other. While G_1 remains fixed, G_2 is able to rotate fully. The normals of both glass plates are tilted by 57° from the vertical direction. The light beam l_i is incident at 57° to the normal of the lower glass plate. The reflected light from G_1 is directed vertically upward, reflecting the light l_0 . When G_2 is rotated by 90° , however, the exit light disappears. Explain why this exit light disappears.
- 2.7 Pulfrich's refractometer, which measures the index of refraction of fluid, is constructed as shown in Fig. P2.7. As the viewing angle θ_2 of the telescope changes, an angle of θ_2 is created such that half of the field is bright and the other half dark. The surface of the interface between the liquid under test and the base glass is illuminated by diffuse light (light with all angles of incidence). The index of refraction (n_2) of the glass is larger than n_x of the fluid. Find the index of refraction n_x of the fluid when the emergent angle of this half-bright and half-dark condition is θ_2 .
- 2.8 The expression for Snell's law in k coordinates is

$$k_{1x}^2 - k_{2x}^2 = (n_1^2 - n_2^2)k^2$$

By referring to Fig. 2.13, rewrite the k -coordinate expression of Snell's law into the angular expression of Snell's law.

REFERENCES

1. E. Hecht, *Schaum's Outline of Theory and Problems of Optics*, Schaum's Outline Series. McGraw-Hill, New York, 1975.
2. H. H. Skilling, *Fundamentals of Electric Waves*, Wiley, New York, 1948.
3. N. Mizutani and T. Numai, "Analysis of reflectivity for a beveled corner mirror in semiconductor ring lasers," *J. Lightwave Technol.* **19**(2), 222–229 (2001).
4. E. Gini, G. Guekos, and H. Melchior, "Low loss corner mirrors with 45° deflection angle for integrated optics," *Electron. Lett.* **28**(5), 499–501 (1992).

5. G. R. Fowles, *Introduction to Modern Optics*, 2nd ed., Dover Publications, New York, 1989.
6. V. F. Goos and H. Hänchen, "Ein neuer and fundamentaler Versuch zur Totalreflexion," *Ann. Phys.* 6. Folge, Band 1, Heft 7/8, 333–346 (1947).
7. H. K. V. Lotsch, "Beam displacement at total reflection: The Goos–Hänchen effect, II, III, IV" *Optik* **32**, 189–204, 299–319, 553–569 (1971).
8. H. Wolter, "Untersuchungen zur Strahlversetzung bei Totalreflexion des Lichtes mit der Methode der Minimumstrahlkennzeichnung," *Z. Naturforsch.* **5a**, 143–153 (1950).
9. M. Tsuji, S. Matsumoto, H. Shigesawa, and K. Takiyama, "Guided-wave experiments with dielectric waveguide having finite periodic corrugation," *IEEE Trans. Microwave Theory Tech.* **MTT-31**(4), 337–344 (1983).
10. A. Lewis, "Diffraction unlimited optics," in *Current Trends in Optics*, J. C. Dainty, Ed., Academic Press, London, 1994, Chap. 17.
11. V. P. Tychinsky and C. H. F. Velzel, "Super-resolution in microscopy," in *Current Trends in Optics*, J. C. Dainty, Ed., Academic Press, London, 1994, Chap. 18.
12. J. M. Guerra, "Photon tunneling microscopy applications," *Mat. Res. Soc. Symp. Proc.* **332**, 449–460 (1994).
13. J. M. Guerra, M. Srinivasarao, and R. S. Stein, "Photon tunneling microscopy of polymeric surfaces," *Science* **262**, 1395–1400 (26 Nov. 1993).
14. M. A. Paesler and P. J. Moyer, *Near-Field Optics, Theory, Instrumentation and Applications*, Wiley, New York, 1996.
15. D. W. Pohl, "Some thoughts about scanning probe microscopy, micromechanics and storage," *IBM J. Res. Dev. USA* **39**(6), 701–711 (Nov. 1995).
16. D. W. Pohl and D. Courjon (Eds.), *Near Field Optics*, Kluwer Academic Publishers, Boston, MA, 1993.
17. A. G. Ruiter, J. A. Veerman, K. O. van der Werf, and N. F. van Hulst, "Dynamic behavior of tuning fork shear-force feedback," *Appl. Phys. Lett.* **71**(1), 28–30 (1997).
18. M. F. Garcia-Parajo, J.-A. Veerman, S. J. T. van Noort, B. G. de Grooth, J. Greve, and N. F. van Hulst, "Near field optical microscopy for DNA studies at the single molecular level," *Bioimaging* **6**, 43–53 (1998).
19. J. M. Vigoureux, F. Depasse, and C. Girard, "Superresolution of near-field optical microscopy defined from properties of confined electromagnetic waves," *Appl. Opt.* **31**(16), 3036–3045 (1992).
20. J. M. Vigoureux and D. Courjon, "Detection of nonradiative fields in light of the Heisenberg uncertainty principle and the Rayleigh criterion," *Appl. Opt.* **31**(16), 3170–3177 (1992).
21. P. Lambelet, A. Sayah, M. Pfeffer, C. Philipona, and F. Marquis-Weible, "Chemical etching of fiber tips through the jacket; new process for smoother tips," *5th International Conference on Near Field Optics*, Shirahama, Japan, pp. 218–219, C3, Dec. 6–10, 1998.
22. K. Ito, A. Kikukawa, and S. Hosaka, "A flat probe for a near field optical head on a slider," *5th International Conference on Near Field Optics*, Shirahama, Japan, pp. 480–481, I16, Dec. 6–10, 1998.
23. J.-J. Greffet and R. Carminati, "Image formation in near-field optics," *Prog. Surf. Sci.* **56**(3), 133–237 (1997).
24. R. Bachelot, P. Gleyzes, and A. C. Boccarda, "Reflection-mode scanning near-field optical microscopy using an apertureless metallic tip," *Appl. Opt.* **36**(10), 2160–2170 (1997).
25. M. Kourogi, T. Yatsui, S. Ishimura, M. B. Lee, N. Atoda, K. Tsutsui, and M. Ohtsu, "A near-field planar apertured probe array for optical near-field memory," *SPIE* **3467**, 258–267 (1998).