5.1 THE BEHAVIOUR OF LIGHT AT A DIELECTRIC INTERFACE

The previous Chapters have been concerned with the propagation of waves in empty space or in uniform, homogeneous media. It is now time to introduce more complicated geometries, building towards those appropriate to optical waveguides. We begin with a qualitative discussion of the phenomena that occur when light strikes a dielectric interface. Figure 5.1-1 shows a plane wave incident at an angle \( \theta_1 \) on two semi-infinite media, with refractive indices \( n_1 \) and \( n_2 \), respectively. These are chosen such that \( n_1 > n_2 \).

\[ n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \]  
Equation 5.1-1

Experience suggests that for near-normal incidence (small \( \theta_1 \)) the incident beam will give rise to a plane reflected beam in Medium 1, together with a transmitted beam in Medium 2. The reflected beam travels at an angle equal and opposite to that made by the incident wave with the interface normal. The angle of the transmitted beam is defined by Snell’s law (named after Willebrord Snell, 1591-1626). This was originally an experimental observation, and requires that:

However, we note that for \( \theta_1 = \sin^{-1}(n_2/n_1) \), we get \( \theta_2 = \pi/2 \). In this case, the transmitted wave travels parallel to the interface, as shown in Figure 5.1-2.

This angle of incidence is called the critical angle, \( \theta_c \). For an interface between glass and air (when \( n_1 = 1.5 \) and \( n_2 = 1.0 \)), the critical angle is \( \theta_c = 41.8^\circ \). However, \( \theta_1 \) can never reach \( \pi/2 \) if we exchange the values of \( n_1 \) and \( n_2 \), so there is no critical angle when the wave is
incident from the low-index side. For $\theta_1 > \theta_c$, however, there is no real solution for $\theta_2$, which implies that a propagating transmitted wave cannot arise at all. This is known as total internal reflection, and is illustrated in Figure 5.1-3.

![Figure 5.1-3 Total internal reflection.](image)

It is easy to show that when total internal reflection occurs, no power crosses the interface. Total internal reflection therefore presents a mechanism for confining a field in one region of space. We shall now proceed to a more rigorous analysis of the problem. The first point to note is that the presence of two waves (one reflected, one transmitted) in Medium 1 implies that there must be a more complicated field distribution in this region. We will therefore start by examining the way plane wave solutions to Maxwell’s equations can be combined. Since the equations are linear, this is very simple.

### 5.2 Superposition of Fields

Perfectly monochromatic fields can produce a range of strikingly beautiful effects, which fall under the general heading of interference phenomena. The simplest of these occurs when two plane waves of the same frequency cross at an angle, as shown in Figure 5.2-1.

![Figure 5.2-1 Interference between two plane waves.](image)

We assume that both waves are polarised in the y-direction, travel in free space, and have equal amplitude $E_y$, so their time-independent electric fields are given by:

$$E_1 = j E_y \exp[-jk_0(z \cos(\theta) + x \sin(\theta))]$$  \hspace{1cm} 5.2-1

for the upward-travelling wave, and:

$$E_2 = j E_y \exp[-jk_0(z \cos(\theta) - x \sin(\theta))]$$  \hspace{1cm} 5.2-2

for the downward one. Because Maxwell’s equations are themselves linear, the total field can be written as a linear superposition of these two fields, in the form:

$$E = j E_y \{ \exp[-jk_0(z \cos(\theta) + x \sin(\theta))] + \exp[-jk_0(z \cos(\theta) - x \sin(\theta))] \}$$
Here, we have combined the x-dependent parts of the two fields to yield a real, cosinusoidal amplitude distribution. We can write Equation 5.2-3 in the alternative form:

\[
\mathbf{E} = j \mathbf{E}(x) \exp(-j\beta z)
\]

where:

\[
\mathbf{E}(x) = 2\mathbf{E}_y \cos[(k_0 x \sin(\theta))]
\]

and

\[
\beta = k_0 \cos(\theta)
\]

The result can therefore be regarded as a field with non-uniform amplitude \(\mathbf{E}(x)\), travelling in the z-direction. However, its propagation constant is reduced from the normal value by the factor \(\cos(\theta)\).

**DESIGN EXAMPLE**

As an exercise, we shall now calculate the time average of the power carried by the combined field above, per unit area in the z-direction. This is representative of the way the field would be detected by the eye, or by photographic film oriented in a plane normal to the z-axis. Once again, we need to evaluate the quantity \(P_z = \frac{1}{2} \text{Re} \{\mathbf{E} \times \mathbf{H}^*\} \cdot \mathbf{k}\). Now, the time independent electric field is given by Equation 5.2-3. The time independent magnetic field, on the other hand, is best calculated from the curl relationship between \(\mathbf{E}\) and \(\mathbf{H}\):

\[
\nabla \times \mathbf{E} = -j\omega\mu_0 \mathbf{H}
\]

Since \(\mathbf{E}\) has just a y-component, it is easy to show that \(\mathbf{H}\) has only x- and z-components, given by:

\[
H_x = 2\mathbf{E}_y \left( -\frac{\beta}{\mu_0 \omega} \right) \exp(-j\beta z) \cos[k_0 x \sin(\theta)]
\]

\[
H_z = 2\mathbf{E}_y \left( -\frac{j k_0 \sin \theta}{\mu_0 \omega} \right) \exp(-j\beta z) \sin[k_0 x \sin(\theta)]
\]

Substituting into the expression for \(P_z\), we get the following non-uniform distribution:

\[
P_z = -\mathbf{E}_y \mathbf{H}_x^* = 4\mathbf{E}_y^2 \left( \frac{\beta}{2\mu_0 \omega} \right) \cos^2[k_0 x \sin(\theta)]
\]

This is known as an **interference pattern**. Notice that it has no dependence on z, but varies periodically with x. It would therefore be visible as a pattern of straight fringes, oriented parallel to the z-axis, as shown in Figure 5.2-2.

The spatial frequency of the pattern is \(K = k_0 \sin(\theta)\), so the separation between fringes is \(\Lambda = 2\pi/K\). Notice that \(\Lambda\) decreases as the angle between the two beams increases, and is of the same order as the wavelength of the light. Interference patterns of this type are particularly important in many branches of optics. However, they do not arise if the two beams are orthogonally polarized, or if the light used is not coherent.
5.3 BOUNDARY MATCHING

There is one further complication to the interface problem. When electromagnetic waves meet changes in dielectric constant, boundary conditions must be satisfied. We can find the necessary conditions by applying Maxwell's equations at a boundary discontinuity. Figure 5.3-1 shows a suitable junction between two different semi-infinite media, Medium 1 (where the time-dependent electric field is $E_1$), and Medium 2 (where it is $E_2$).

First, recall from Faraday's law that:

$$\int_L E \cdot dL = - \int_A \frac{\partial B}{\partial t} \cdot da$$  \hspace{1cm} 5.3-1

We shall now evaluate the left-hand side of Equation 5.3-1 by performing a line integration round the rectangular dashed loop shown in Figure 5.3-1. This crosses the boundary, and has sides of length $L_1$ and $L_2$. Here $L_2 \ll L_1$, and $L_1$ is small enough for the fields to be approximately uniform along each of the long edges of the loop. The right hand side is of course a surface integral, over the enclosed area.

We now let $L_2$ tend to zero. As this happens, the enclosed area tends to zero, so the right-hand side of Equation 5.3-1 must tend to zero as well. In this case, we must have:
∫_L E \cdot dL = 0 \quad \text{5.3-2}

Now at this point, the integral in Equation 5.3-2 is well approximated by:

∫_L E \cdot dL = (E_{t1} - E_{t2}) L_i \quad \text{5.3-3}

Where \(E_{t1}\) and \(E_{t2}\) are the components of \(E_1\) and \(E_2\) tangential to the boundary. This implies that:

\(E_{t1} - E_{t2} = 0\) \quad \text{5.3-4}

In other words, the tangential components of \(E\) must match across the boundary. This is the first of our boundary conditions. We can also write this without explicitly mentioning \(E_{t1}\) and \(E_{t2}\), as:

\(n \times (E_2 - E_1) = 0\) \quad \text{5.3-5}

Where the vector \(n\) is a new vector, the local normal to the boundary.

We can do the same sort of thing for \(H\), \(D\) and \(B\). The complete set of boundary conditions is then:

\(n \times (E_2 - E_1) = 0\)
\(n \times (H_2 - H_1) = K\)
\(n \cdot (D_2 - D_1) = \rho_s\)
\(n \cdot (B_2 - B_1) = 0\) \quad \text{5.3-6}

Here \(K\) is the surface current (which can be ignored in dielectrics) and \(\rho_s\) is the surface charge density (again, zero in dielectrics). Generally, not all of Equations 5.3-6 are needed to solve any particular problem; in fact, picking the most useful subset of conditions is something of an art in itself.

Similar equations can be obtained for the single-frequency, time-independent representation, simply by replacing \(E\) by \(E_0\) and so on. For dielectric media, we therefore obtain:

\(n \times (E_2 - E_1) = 0\)
\(n \times (H_2 - H_1) = 0\)
\(n \cdot (D_2 - D_1) = 0\)
\(n \cdot (B_2 - B_1) = 0\) \quad \text{5.3-7}

5.4 ELECTROMAGNETIC TREATMENT OF THE INTERFACE PROBLEM

We must now calculate what happens at the interface more exactly. We assume that the governing equation is the time-independent vector wave equation. For a uniform dielectric, this takes the form:

\(\nabla^2 \overline{E} + n^2 k_0^2 \overline{E} = 0\) \quad \text{5.4-1}
We can use Equation 5.4-1 to solve the interface problem, by first finding solutions for each region as if it were infinite, and then matching them at the boundary. The geometry is shown in Figure 5.4-1. The junction is assumed to lie between two infinite sheets of dielectric, with indices $n_1$ and $n_2$. These are lying in the $x-z$ plane, with the interface at $x = 0$, and the aim is to calculate the field distribution with a plane wave incident from Medium 1 at an angle $\theta_1$.

![Figure 5.4-1](image)

**Figure 5.4-1** Geometry for treatment of the interface problem.

TE INCIDENCE

If we assume as before that the incident wave is polarized in the $y$-direction, the electric field is entirely parallel to the interface and only has a $y$-component. This is known as transverse electric or TE incidence. Without justification (apart from common sense) we assume that the transmitted and reflected waves are similarly polarized, so we can work with $E_y$ throughout. The vector wave equation therefore reduces to a single scalar equation.

Furthermore, because everything is uniform in the $y$-direction, $\partial/\partial y$ must be zero for all field quantities. The relevant equation is therefore:

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} + n_2^2 k_0^2 E_y = 0$$  \hspace{1cm} 5.4-2

**1. REFLECTION AND REFRACTION**

We start by assuming that total internal reflection does not occur, so we need only consider the case of reflection and refraction. The equations we must solve in each medium are:

$$\frac{\partial^2 E_{y1}}{\partial x^2} + \frac{\partial^2 E_{y1}}{\partial z^2} + n_1^2 k_0^2 E_{y1} = 0 \quad \text{in Medium 1}$$

$$\frac{\partial^2 E_{y2}}{\partial x^2} + \frac{\partial^2 E_{y2}}{\partial z^2} + n_2^2 k_0^2 E_{y2} = 0 \quad \text{in Medium 2}$$  \hspace{1cm} 5.4-3

We now assume that the field in each region is a combination of plane waves, the solutions we would find if the media were infinite and separate. A suitable guess is to choose incident and reflected waves in Medium 1, and just a transmitted wave in Medium 2, so that:

$$E_{y1} = E_i \exp[-jk_0n_1(z \sin(\theta_1) - x \cos(\theta_1))] + E_R \exp[-jk_0n_1(z \sin(\theta_1') + x \cos(\theta_1'))]$$

$$E_{y2} = E_T \exp[-jk_0n_2(z \sin(\theta_2) - x \cos(\theta_2))]$$  \hspace{1cm} 5.4-4

Here we have assumed for generality that the reflected wave travels upwards at an angle $\theta_1'$, which may be different from $\theta_1$. Now, there are a number of unknowns in Equation 5.4-4. Given that we know $\theta_1$ and $E_i$, we must find $\theta_1'$, $\theta_2$, $E_R$ and $E_T$. We can find them from the boundary conditions in Equation 5.3-7. In principle, we could use any of these conditions, but here we choose to start with:
This requires that the tangential components of $E$ must match on the boundary. Since $E$ is wholly tangential anyway, we must have $E_{y1} = E_{y2}$ on $x = 0$, or:

$$E_I e^{-jk_0 n_1 z \sin(\theta_1)} + E_R e^{-jk_0 n_1 z \sin(\theta_1')} = E_T e^{-jk_0 n_2 z \sin(\theta_2)}$$

Because it contains exponential terms that vary with $z$, this equation can only be satisfied for all $z$ if:

$$n_1 \sin(\theta_1) = n_1 \sin(\theta_1')$$

$$= n_2 \sin(\theta_2)$$

If this is the case, the $z$-variation is removed, and Equation 5.4-6 can be satisfied by a correct choice of the constants $E_R$ and $E_T$. Equation 5.4-7 requires firstly that $\theta_1' = \theta_1$, so the angle of the reflected beam is indeed equal to that of the incident beam (verifying our initial intuitive guess). However, it also shows that Snell's law appears automatically in the electromagnetic analysis, which is highly satisfactory. Equation 5.4-6 then reduces to:

$$E_I + E_R = E_T$$

Performing the boundary matching for the electric field, we therefore find that the wave amplitudes must be continuous at the interface, which is physically reasonable.

The second boundary condition we choose is that:

$$n \times (H_2 - H_1) = 0 \quad \text{on} \quad x = 0$$

This implies that the tangential components of $H$ must also match at $x = 0$. We can find the necessary components of $H$ by using the curl relation between $E$ and $H$ given previously in Equation 5.2-6. Since $E_x = E_z = 0$, we have:

$$H_x = (-j/\omega\mu_0) \frac{\partial E_y}{\partial z} ; \quad H_y = 0 ; \quad H_z = (+j/\omega\mu_0) \frac{\partial E_y}{\partial x}$$

The only tangential component of $H$ is $H_z$, so the second boundary condition will be satisfied if:

$$\frac{\partial E_{y1}}{\partial x} = \frac{\partial E_{y2}}{\partial x} \quad \text{on} \quad x = 0.$$ 

Using this condition, we find that:

$$j\kappa n_1 \cos(\theta_1) E_I e^{-jk_0 n_1 z \sin(\theta_1)} - j\kappa n_1 \cos(\theta_1) E_R e^{-jk_0 n_1 z \sin(\theta_1)}$$

$$= j\kappa n_2 \cos(\theta_2) E_T e^{-jk_0 n_2 z \sin(\theta_2)}$$

Following the same argument as before, we can remove all the exponentials to get:

$$n_1 \cos(\theta_1) E_I - n_1 \cos(\theta_1) E_R = n_2 \cos(\theta_2) E_T$$

Equations 5.4-8 and 5.4-13 now represent two simultaneous equations, with two unknowns. By a fairly simple rearrangement, we can therefore find $E_R$ in terms of $E_I$. Normally the result is expressed in terms of a reflection coefficient $\Gamma_E$, which can be shown to reduce to:
\[ \Gamma_E = \frac{E_R}{E_I} = \left[ n_1 \cos(\theta_1) - n_2 \cos(\theta_2) \right] / \left[ n_1 \cos(\theta_1) + n_2 \cos(\theta_2) \right] \]

We can also find the transmission coefficient \( T_E \), which is given by:

\[ T_E = \frac{E_T}{E_I} = 2n_1 \cos(\theta_1) / \left[ n_1 \cos(\theta_1) + n_2 \cos(\theta_2) \right] \]

These expressions satisfy physical intuition. For example, if the two media have the same refractive index, the reflection coefficient is zero, and the transmission coefficient is unity. Surprisingly, the reflection coefficient can be large, even at normal incidence. For a glass/air interface (where \( n_1 = 1.5 \) and \( n_2 = 1 \)), we find that \( \Gamma_E = (1.5 - 1) / (1.5 + 1) = 0.2 \).

**THE OTHER BOUNDARY CONDITIONS**

At this point we pause, and consider the fact that there were two more boundary conditions we could have used. What about them? The third condition is that:

\[ \mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = 0, \quad \text{or} \quad \mathbf{n} \cdot (\sqrt{n_2} \mathbf{E}_2 - \sqrt{n_1} \mathbf{E}_1) = 0 \quad \text{on} \quad x = 0 \]

Since \( \mathbf{E} \) has no component normal to the boundary, all this equation says is \( 0 = 0 \). The third condition is therefore not too helpful, and we can safely ignore it. The fourth is that:

\[ \mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0, \quad \text{or} \quad \mathbf{n} \cdot (\mathbf{H}_2 - \mathbf{H}_1) = 0 \quad \text{on} \quad x = 0 \]

Since the only normal component of \( \mathbf{H} \) is \( H_z \), this implies that:

\[ H_{z2} = H_{z1} \quad \text{on} \quad x = 0 \]

Bearing in mind the definition of \( \mathbf{H} \) given in Equation 5.4-10, we find that:

\[ \partial E_{y2} / \partial z = \partial E_{y1} / \partial z \quad \text{on} \quad x = 0. \]

Doing the necessary differentiation, we get:

\[ -j\kappa n_1 \sin(\theta_1) \mathbf{E}_i \exp[-j\kappa n_1 z \sin(\theta_1)] - j\kappa n_1 \sin(\theta_1) \mathbf{E}_r \exp[-j\kappa n_1 z \sin(\theta_1)] \]

\[ = -j\kappa n_2 \sin(\theta_2) \mathbf{E}_t \exp[-j\kappa n_2 z \sin(\theta_2)] \]

If we now use Snell's law again to remove the exponentials, this reduces to Equation 5.4-8. The two other boundary conditions therefore introduce no new features, so our original choice was a good one.

**TM INCIDENCE**

The analysis can easily be repeated for the case when the magnetic field is oriented in the y-direction. This is known as transverse magnetic or TM incidence. We shall not work through all the mathematics, which is somewhat repetitive, but merely quote the results. These are slightly different; for the reflection coefficient, we get:

\[ \Gamma_H = \frac{E_R}{E_I} = \left[ n_2 \cos(\theta_1) - n_1 \cos(\theta_2) \right] / \left[ n_2 \cos(\theta_1) + n_1 \cos(\theta_2) \right] \]

and for the transmission coefficient, we find that:

\[ T_H = \frac{E_T}{E_I} = 2n_1 \cos(\theta_1) / \left[ n_2 \cos(\theta_1) + n_1 \cos(\theta_2) \right] \]
Collected together, $\Gamma_E$, $T_E$, $\Gamma_H$ and $T_H$ are often known as the Fresnel coefficients. Most interestingly, $\Gamma_H$ becomes zero at a particular angle of incidence $\theta_1$. This occurs when:

$$n_2 \cos(\theta_1) = n_1 \cos(\theta_2)$$

We can find a suitable angle $\theta_1$ for Equation 5.4-23 to hold, as follows. Squaring both sides of the equation, and using Snell's Law, we obtain:

$$n_2^2 \cos^2(\theta_1) = n_1^2 [1 - (n_1^2/n_2^2) \sin^2(\theta_1)]$$

Or:

$$n_2^2 \cos^2(\theta_1) + (n_1^4/n_2^4) \sin^2(\theta_1) = n_1^2$$

The $\cos^2$ term above can now be eliminated by combining it with the $\sin^2$ term. If this is done, we get:

$$n_2^2 [\cos^2(\theta_1) + \sin^2(\theta_1)] + [(n_1^4 - n_2^4)/n_2^4] \sin^2(\theta_1) = n_1^2$$

so that:

$$[(n_1^4 - n_2^4)/n_2^4] \sin^2(\theta_1) = n_1^2 - n_2^2$$

We therefore find that:

$$\sin^2(\theta_1) = n_2^2 / (n_1^2 + n_2^2)$$

So the relevant angle is given by:

$$\theta_1 = \sin^{-1}[n_2 / \sqrt{(n_1^2 + n_2^2)}]$$

This is often written in the alternative form:

$$\theta_1 = \tan^{-1}(n_2/n_1)$$

The reduction of the reflection coefficient to zero for TM incidence is called the Brewster effect, after David Brewster (1781 - 1868), and is often exploited in the design of polarizing components. Similarly, the angle defined in Equation 5.4-30 is known as the the Brewster angle $\theta_B$. This can be calculated for an glass/air interface (for which $n_1 = 1.5$ and $n_2 = 1.0$) as $\theta_B \approx 33.7^\circ$. Note that the Brewster effect does not occur for TE incidence.

Figure 5.4-2 shows one way in which the Brewster effect may be used to make a polarizer. A number of identical plates are piled up, and oriented so that an incident, unpolarized beam strikes the stack at the Brewster angle. Any TE polarization component in the beam will be diminished by reflection at each interface, while the TM component will pass through unaffected. After many such reflections, the output beam will therefore be largely TM polarized. This is known as the pile-of-plates polarizer. Though simple, it has been entirely superseded by dichroic polarizers.
Figure 5.4-2 The pile of plates polarizer.

Figure 5.4-3 shows a much more effective use of the Brewster effect, inside the resonant cavity of a He-Ne laser. We have previously described the required optical layout in Chapter 4, and concentrate here on the provision of optical gain by an electrically excited gas plasma, which is contained inside a glass tube. In lasers that emit unpolarized light, the plasma tube is simply capped with end plates that are orthogonal to the bore axis. In polarized lasers, the tube is sealed with windows slanted at the Brewster angle (known, unsurprisingly, as Brewster windows). Now, TE polarized light must suffer some reflection loss at these windows. However, because of the Brewster effect, TM polarized light suffers no loss. Consequently, the lasing threshold for TM light is much lower than for TE, and the device will lase preferentially on a TM-polarized mode.

2. TOTAL INTERNAL REFLECTION

We now consider the case when $\theta_1 > \theta_c$, i.e. when the input beam is incident at an angle greater than the critical angle. In this regime, total internal reflection occurs, so that a propagating wave cannot arise in Medium 2. How do we find the solution now? Well, there is nothing wrong with the analysis we have used so far, but we cannot satisfy the Snell's law equation, $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$. Or rather, we cannot satisfy it with a real value of $\theta_2$. It turns out that we can satisfy it with a complex value instead. This might seem a bizarre idea, but there is no need to dwell on the physical implications involved; it is simplest just to consider this as a mathematical trick. The important thing is that we can now calculate $\cos(\theta_2)$ for $\theta_1 > \theta_c$ as:

$$\cos(\theta_2) = \pm j \sqrt{(n_1/n_2)^2 \sin^2(\theta_1) - 1}$$

5.4-31
This identity will allow us to repeat the previous calculation under the new conditions.

**TE INCIDENCE**

Once again, we will begin with the case when the electric field vector is parallel to the interface. Returning to the solution in Medium 2, \( E_{y2} = E_T \exp[-jk_0n_2(z \sin(\theta_2) - x \cos(\theta_2))] \), and inserting the value for the cosine given in Equation 5.4-31, we find:

\[
E_{y2} = E_T \exp \left[ \pm k_0 n_2 x \sqrt{(n_1/n_2)^2 \sin^2(\theta_1) - 1} \right] \exp[-jk_0 n_1 z \sin(\theta_1)].
\]

Equation 5.4-32

The solution no longer has the form of a plane wave propagating at an angle \( \theta_2 \), because its \( x \)-dependence is not a complex exponential. Instead, it is a wave whose amplitude distribution varies exponentially in the \( x \)-direction, and that propagates along the interface in the \( z \)-direction. This is called an **evanescent wave**. Clearly, the amplitude must decay to zero for large negative \( x \), so we will keep only the positive sign in Equation 5.4-32, which corresponds to the negative sign in Equation 5.4-31.

What does the complete field distribution look like now? Well, we now know that:

\[
E_{y1} = E_I \left\{ \exp(+j\kappa x) + \Gamma E \exp(-j\kappa x) \right\} \exp(-j\beta z)
\]

\[
E_{y2} = E_I T E \exp(\gamma x) \exp(-j\beta z)
\]

Equation 5.4-33

Where the transmission and reflection coefficients are given by:

\[
\Gamma E = \frac{[n_1 \cos(\theta_1) + jn_2 \sqrt{(n_1/n_2)^2 \sin^2(\theta_1) - 1}]}{[n_1 \cos(\theta_1) - jn_2 \sqrt{(n_1/n_2)^2 \sin^2(\theta_1) - 1}]}
\]

Equation 5.4-34

and:

\[
T E = 2n_1 \cos(\theta_1) / [n_1 \cos(\theta_1) - jn_2 \sqrt{(n_1/n_2)^2 \sin^2(\theta_1) - 1}]\]

Equation 5.4-35

These expressions are rather long-winded. However, we can simplify them by defining three new parameters \( \beta, \kappa \) and \( \gamma \) as:

\[
\beta = k_0 n_1 \sin(\theta_1)
\]

\[
\kappa = \sqrt{(n_1/n_2)^2 - \beta^2}
\]

\[
\gamma = \sqrt{(\beta^2 - n_2^2 k_0^2)}
\]

Equation 5.4-36

in which case Equation 5.4-33 becomes:

\[
E_{y1} = E_I \left\{ \exp(+j\kappa x) + \Gamma E \exp(-j\kappa x) \right\} \exp(-j\beta z)
\]

\[
E_{y2} = E_I T E \exp(\gamma x) \exp(-j\beta z)
\]

Equation 5.4-37

And Equations 5.4-34 and 5.4-35 simplify to:

\[
\Gamma E = (\kappa + j\gamma) / (\kappa - j\gamma)
\]

\[
T E = 2\kappa / (\kappa - j\gamma)
\]

Equation 5.4-38

The reflection coefficient \( \Gamma E \) is therefore now a complex number, in the form \( \Gamma E = z/z^* \), where \( z \) is also a complex number and * denotes the operation of complex conjugation. It is easy to show that the **power reflection coefficient** \( |\Gamma E|^2 \) is unity, since:

\[
|\Gamma E|^2 = \Gamma E \Gamma E^* = (z/z^*)(z^*/z) = 1
\]
This implies that all of the incident power is reflected, which is consistent with our earlier qualitative discussion of total internal reflection. It might be argued that a non-zero transmission coefficient is at odds with a unity power reflection coefficient. The evanescent wave will then have a finite amplitude, and must therefore carry some power. However, it is easy to show that it does not carry any power normal to the interface, but only parallel to it. Power conservation is therefore not violated after all.

Since $\Gamma_E$ is a complex number with modulus unity, it can be written in complex exponential form, as:

$$\Gamma_E = \exp(j2\phi)$$

where the phase term $\phi$ is given by:

$$\tan(\phi) = \gamma/\kappa$$

Similarly, the transmission coefficient $T_E$ can be written as:

$$T_E = 2 \cos(\phi) \exp(j\phi)$$

We can obtain entirely analogous results for the TM transmission and reflection coefficients when total internal reflection occurs. The phase shift is different, however; this time we get $\tan(\phi) = (n_1^2/n_2^2) \gamma/\kappa$.

**DESIGN EXAMPLE**

We can combine the solutions obtained so far, and plot the power reflectivity $|\Gamma|^2$ versus angle over the complete range of incidence. Figure 5.4-4 shows the curves for TE and TM incidence, and the particular example of a glass/air interface (where $n_1 = 1.5$ and $n_2 = 1$).

![Figure 5.4-4](image)

Figure 5.4-4  Power reflection coefficients, for a glass-air interface.
Naturally, the reflectivities are the same for normal incidence, when $|\Gamma|^2 = 0.2^2 = 0.04$. In each case, the reflectivity is unity after the critical angle, $\theta_1 = 41.8^\circ$. Notice, however, that it falls to zero, for TM incidence alone, at the Brewster angle ($33.7^\circ$).

THE TRANSVERSE FIELD DISTRIBUTION NEAR THE INTERFACE

Knowing the transmission and reflection coefficients, we can now rewrite the field distributions in the two media as:

\[
E_{y1} = 2E_i \exp(j\phi) \cos(\kappa x - \phi) \exp(-j\beta z) = E \cos(\kappa x - \phi) \exp(-j\beta z)
\]

And:

\[
E_{y2} = 2E_i \exp(j\phi) \cos(\phi) \exp(\gamma x) \exp(-j\beta z) = E \cos(\phi) \exp(\gamma x) \exp(-j\beta z)
\]

5.4-43

Both fields again have the form $E_y = E(x) \exp(-j\beta z)$, so the pattern corresponds to a non-uniform wave travelling in the z-direction, parallel to the interface. More specifically, it is a standing wave pattern in Medium 1, and a decaying or evanescent wave in Medium 2. Near the interface, the transverse field function $E(x)$ is therefore as shown in Figure 5.4-5. Notice that the transverse field is continuous at the boundary, as is its first derivative. This follows from application of the boundary matching conditions. It is also worth noting that as $\beta$ tends to $n_2 k_0$, $\gamma$ tends to zero, implying that, as the angle of incidence approaches the critical angle, the decay of the evanescent field gets slower. The confinement of the field is therefore less good, and it extends further and further into Medium 2. If we wish to avoid this, the incidence angle should be chosen to lie suitably far away from the critical angle.

The results are qualitatively similar for TM incidence; we still get standing waves in Medium 1 and an evanescent wave in Medium 2. However, the details at the boundary are slightly different. The main change is that the transverse electric field distribution is no longer continuous.
THE REFLECTIVITY OF METALS

We conclude with a brief discussion of the reflectivity of metals, which we assumed without much justification to be high in Chapter 4. The analysis developed above will still be valid, provided merely that we correct the expression for refractive index. In Chapter 3, we found that (neglecting collision damping losses) a metal may be described at optical frequencies as a medium with a large, negative value of \( \varepsilon' \). Consequently, the refractive index of a metal will be almost purely imaginary, given by \( n \approx jn' \), where \( n' = \sqrt{\varepsilon'} \).

At normal incidence, for example, the reflection coefficient at the interface between a dielectric (of refractive index \( n_1 \)) and a metal (of index \( jn_2' \)) may then be found from Equation 5.4-14 as:

\[
\Gamma_E \approx \frac{(n_1 - jn_2')}{(n_1 + jn_2')}
\]

Once again, \( \Gamma_E \) is a complex number, of the form \( \Gamma_E = z / z^* \). We therefore conclude without further ado that the power reflectivity will be 100%, justifying our previous assumptions. In practice, collision damping losses reduce this figure somewhat, but (as we know from our bathrooms) even the most basic metal layer will act as a good reflector. Furthermore, if \( n_2 \gg n_1 \) (as is the case for good conductors), the reflection coefficient \( \Gamma_E \) is \( \approx -1 \). This means the total electric field at the interface (i.e. the sum of the incident and reflected waves) must be close to zero, providing a simple justification for the heuristic boundary condition adopted in Chapter 4.

5.5 MODAL TREATMENT OF THE DIELECTRIC INTERFACE PROBLEM

We now consider an easier route to the solution to the interface problem, which will prove a useful tool for the analysis of waveguides in Chapter 6. Since the solution has the form \( E_0(x, z) = E(x) \exp(-j\beta z) \) in both media, the thing to do is to assume this at the outset. Substituting this solution directly into the wave equation 5.4-2, we get:

\[
\frac{d^2E}{dx^2} + (n_2^2k_0^2 - \beta^2) E = 0
\]

This type of equation is known as a waveguide equation, and it links the transverse field \( E(x) \) with the propagation constant \( \beta \). In this case, it is a standard second-order differential equation, of the form:

\[
\frac{d^2E}{dx^2} + C^2 E = 0
\]

Note that its solutions are sines and cosines if \( C^2 > 0 \), and exponentials if \( C^2 < 0 \).

For our particular geometry, the waveguide equations we must solve are:

\[
\frac{d^2E_1}{dx^2} + (n_1^2k_0^2 - \beta^2) E_1 = 0 \quad \text{in Medium 1}
\]
\[
\frac{d^2E_2}{dx^2} + (n_2^2k_0^2 - \beta^2) E_2 = 0 \quad \text{in Medium 2}
\]

Here \( E_1 \) and \( E_2 \) are the transverse field distributions in the two media. Knowing (as we do) the exact solution, the sensible thing is to choose a trial solution in the same form. We therefore take:

\[
E_1 = E \cos(kx - \phi)
E_2 = E' \exp(\gamma x)
\]

R.R.A.Syms and J.R.Cozens  Optical Guided Waves and Devices 14
Where κ and γ are as in Equation 5.4-36 and E, E' and \( \phi \) are unknown constants. Once again, we can determine their values by boundary matching. Now, the first of our original boundary conditions were that \( E_1 = E_2 \) at the interface; for the transverse fields, this implies that \( E_1 = E_2 \) on \( x = 0 \), so:

\[
E' = E \cos(\phi)
\]

Similarly, the second condition was that \( \partial E_1 / \partial x = \partial E_2 / \partial x \) at the interface; for the transverse fields, this implies that \( \frac{dE_1}{dx} = \frac{dE_2}{dx} \) on \( x = 0 \). Performing the necessary differentiation, we get:

\[
-\kappa E \sin(-\phi) = \gamma E'
\]

Dividing Equation 5.5-6 by Equation 5.5-5, we then obtain:

\[
\tan(\phi) = \frac{\gamma}{\kappa}
\]

Since this is identical to Equation 5.4-41, we have obtained exactly the same results as before, without invoking the Fresnel coefficients at all. A solution of this type is called a modal solution.

### 5.6 SURFACE PLASMA WAVES

We have become accustomed to solutions to interface problems that involve either propagating waves on both sides of the interface, or propagating waves on one side and evanescent ones on the other. It is now reasonable to ask whether other possibilities exist. Can we, for example, have a solution that is evanescent on both sides of an interface? It turns out that this is indeed possible, and is of engineering significance.

The geometry involved is the interface between a metal (most often, silver or gold) and a dielectric, as shown in Figure 5.6-1. Because the wave is evanescent on both sides, it is effectively confined to the interface, and is therefore a surface wave. Its correct title is a surface plasma wave or surface plasmon, since propagation in the metal layer involves plasma oscillations of the free electrons inside the metal, but this is not important to our electromagnetic analysis. What is important, however, is that surface plasma waves can be shown to exist only for TM polarization.

![Figure 5.6-1](image-url)  
**Figure 5.6-1** Geometry for surface plasmon waves.
We shall tackle the problem using the modal approach above. First, however, we recall once again that at optical frequencies a metal can be adequately modelled as a dielectric, but with a relative permittivity that is large and negative. The methods we have used so far will therefore still prove suitable. Now, it is sensible to perform the analysis in terms of the magnetic field, since this is always oriented in the same direction. In this case, the relevant time-independent wave equation is:

$$\nabla^2 \mathbf{H} + \varepsilon_r k_0^2 \mathbf{H} = 0$$  \hspace{1cm} 5.6-1

Assuming that the magnetic field is parallel to the y-axis, a suitable modal solution has the form:

$$\mathbf{H} = H(x) \exp(-j\beta z) \hat{\mathbf{j}}$$  \hspace{1cm} 5.6-2

Where $H(x)$ is the transverse variation of the magnetic field, and $\beta$ is the propagation constant. Substituting this solution into Equation 5.6-1, we obtain the following waveguide equation:

$$\frac{d^2H}{dx^2} + (\varepsilon_r k_0^2 - \beta^2) H = 0$$  \hspace{1cm} 5.6-3

To solve the plasmon problem, we simply assume a relative dielectric constant of $\varepsilon_r^1$ in Medium 1 (the metal layer), and $\varepsilon_r^2$ in Medium 2 (the dielectric). The equations we must solve are therefore:

$$\frac{d^2H_1}{dx^2} + (\varepsilon_r^1 k_0^2 - \beta^2) H_1 = 0 \quad \text{in Medium 1}$$
$$\frac{d^2H_2}{dx^2} + (\varepsilon_r^2 k_0^2 - \beta^2) H_2 = 0 \quad \text{in Medium 2}$$  \hspace{1cm} 5.6-4

Assuming exponential decay of the transverse magnetic field on either side of the interface, the solutions in each region can then be specified as:

$$H_1(x) = H \exp(-\gamma_1 x)$$
$$H_2(x) = H' \exp(\gamma_2 x)$$  \hspace{1cm} 5.6-5

Where $H$ and $H'$ are arbitrary constants, and $\gamma_1$ and $\gamma_2$ are given by:

$$\gamma_1 = \sqrt{(\beta^2 - \varepsilon_r^1 k_0^2)}$$
$$\gamma_2 = \sqrt{(\beta^2 - \varepsilon_r^2 k_0^2)}$$  \hspace{1cm} 5.6-6

As before, the boundary conditions are that the tangential components of the magnetic and electric fields must match at the interface, which lies at $x = 0$. Matching the magnetic field (which is wholly tangential) we see simply that $H' = H$. To match the electric fields, we must first evaluate $E$. This is best done using the curl relation between $\mathbf{H}$ and $\mathbf{E}$:

$$\nabla \times \mathbf{H} = j\omega\varepsilon \mathbf{E}$$  \hspace{1cm} 5.6-7

Given that $H_y = H_z = 0$, we can then find the individual components of $\mathbf{E}$ as:

$$E_x = (+j/\omega\varepsilon) \frac{\partial H_y}{\partial z} \quad ; \quad E_y = 0 \quad ; \quad E_z = (-j/\omega\varepsilon) \frac{\partial H_y}{\partial x}$$  \hspace{1cm} 5.6-8

Since the only tangential component of the electric field is $E_x$, the second condition will be satisfied if:
For the modal solution we have assumed, this condition requires:

\[
\frac{1}{\varepsilon_r} \frac{\partial H_y}{\partial x} = \frac{1}{\varepsilon_r} \frac{\partial H_y}{\partial x} \quad \text{on} \quad x = 0
\]

5.6-9

Performing the necessary differentiation, we then find that:

\[
-\gamma_1 / \varepsilon_r = \gamma_2 / \varepsilon_r
\]

5.6-11

Although Equation 5.6-11 looks insoluble (given that \(\gamma_1\) and \(\gamma_2\) are positive quantities), it can be satisfied in this particular case, because \(\varepsilon_r\) is negative. Squaring both sides, we obtain:

\[
(\beta^2 - \varepsilon_r k_0^2 \varepsilon_r^2) / (\varepsilon_r^2) = (\beta^2 - \varepsilon_r k_0^2 \varepsilon_r^2) / (\varepsilon_r^2)
\]

After some rearrangement, we may then extract the propagation constant as:

\[
\beta = k_0 \sqrt{\varepsilon_r^2 / (\varepsilon_r + \varepsilon_r^2)}
\]

5.6-13

Knowing \(\varepsilon_r\) and \(\varepsilon_r^2\), we may therefore calculate \(\beta\). The transverse magnetic field distribution is then much as shown in Figure 5.6-1. The field decays exponentially on both sides of the interface, but the decay constant is different in each medium. Because of the sign of \(\varepsilon_r\), Equation 5.6-6 shows that the rate of decay in the metal is much larger than that in the dielectric. As a result, the light hardly penetrates the metal, and the major part of the field extends into the dielectric.

The analysis above is an extremely simple demonstration that a single interface may actually 'guide' an optical wave. However, to be more realistic, one major modification must be made to our model. As we saw in Chapter 3, the dielectric constant of a metal is actually complex (because of collision damping), so \(\varepsilon_r\) really has the form \(\varepsilon_r = \varepsilon_r' - j \varepsilon_r''\). If this expression is used in Equation 5.6-13, it is found that the propagation constant is also complex, so the plasmon mode decays as it propagates. In real materials, this is very significant, and plasmon modes have an extremely short range.

**DESIGN EXAMPLE**

We shall now calculate the range of a surface plasmon in a typical material, silver. In this case, the data in Figure 3.4-1 shows that \(\varepsilon_r' = -16.9\) and \(\varepsilon_r'' = 0.55\), at \(\lambda_0 = 0.633\) \(\mu\)m.

Assuming that the surrounding dielectric is air, for which \(\varepsilon_r = 1\), \(\beta\) is given by:

\[
\beta = (2\pi / 0.633 \times 10^6) \times \sqrt{\left(\frac{-16.9 - j 0.55}{-15.9 - j 0.55}\right)}
\]

\[
= 10.2 \times 10^6 - j 1.05 \times 10^4
\]

5.6-14

Now, the actual form of the plasmon wave is given by Equation 5.6-2. Splitting the propagation constant into real and imaginary parts, we can write:

\[
\beta = \beta_r - j \beta_i
\]

5.6-15

With this notation, Equation 5.6-2 then becomes:

\[
H = H(x) \exp(-j\beta_i z) \exp(-\beta_r z) j
\]
This shows that the mode amplitude will decay to 1/e of its initial value in a distance $z_e$, given by:

$$ z_e = \frac{1}{\beta_i} $$

We may take $z_e$ to be representative of the range of the plasmon mode. Inserting the relevant value from Equation 5.6-14, we then obtain a range of $z_e \approx 100 \mu m$, a very small figure. In practise, other effects reduce the range still further, making a surface plasmon wave useless for communication. Nonetheless, one highly important application has been found for optical plasmons: sensing.

**SENSING WITH SURFACE PLASMON WAVES**

We have already mentioned that because of the imbalance between $\gamma_1$ and $\gamma_2$, the vast proportion of the field extends into the dielectric. As a result, it is possible to show that the value of $\beta_r$ is highly sensitive to the exact nature of the dielectric. If this medium is made in the form of a sensing layer, with a dielectric constant that can be altered by some parameter of interest, a variation in that parameter will then result in a detectable change in the plasmon propagation constant.

The experimental apparatus needed to measure this change is extremely simple. It is known as the **prism coupler**, and will be described in detail in Chapter 7. For the present, we merely note that many sensor materials have already been tested for suitability. The most promising application area is biosensing, where it is hoped that surface plasmon devices will provide a range of extremely cheap, disposable medical sensors (for example, for measuring enzyme levels).
CHAPTER FIVE

PROBLEMS

5.1. (a) Find the critical angle for an interface between a high-index glass \((n_1 = 1.7)\) and air \((n_2 = 1)\). (b) What is the corresponding Brewster angle, if the incident beam travels in the glass?
\[(a) 36.03^\circ; (b) 30.46^\circ\]

5.2. The critical angle for an interface between two media is 30°. Find the transmission coefficient, for normal incidence from the high-index side of the interface.
\[1.333\]

5.3. Find the power reflectivity \(\Gamma_2\) and transmissivity \(T_2\), for TE incidence at \(\theta_1 = 30^\circ\) on an interface between air \((n_1 = 1)\) and glass \((n_2 = 1.5)\).
\[0.0577; 0.577\]

5.4. The figures below show the definitions of the reflection and transmission coefficients at a dielectric interface. In the left-hand figure, incidence is from Medium 1, and the coefficients are \(\Gamma\) and \(T\), respectively. In the right-hand figure, incidence is from Medium 2, and the corresponding values are \(\Gamma'\) and \(T'\). Using only the fact that light rays are reversible, show that (a) \(\Gamma = -\Gamma'\), and (b) \(\Gamma^2 + TT' = 1\).

5.5. Verify the results of Question 5.4 analytically, using the Fresnel coefficients.

5.6. The figure below shows a three-layer dielectric stack. Media 1 and 2 are semi-infinite, with refractive indices \(n_1\) and \(n_2\), respectively; Medium 3 has thickness \(d\), and index \(n_3\). The effect of the stack on a plane optical wave may be analysed by summing the contributions from all possible paths between the input and the output, using the transmission and reflection coefficients to calculate the path amplitudes and taking account of any optical phase changes.

For example, the reflectivity may be found by summing all paths starting and ending in Medium 1. Path 1 simply involves reflection at the first interface. The amplitude for this is \(\Gamma_{13}\), where \(\Gamma_{13}\) is the reflection coefficient at the boundary between Media 1 and 3. Path 2 involves transmission through the first interface, reflection at the second, and transmission through the first interface again; the amplitude for this is \(T_{13}\Gamma_{32}T_{31}\exp(-j2k_0n_3d)\), where \(T_{13}\) and \(T_{31}\) are transmission coefficients at the first interface, and \(\Gamma_{32}\) is the reflection coefficient at the second.

Write down the amplitude for Path 3. Generalise your results to form an infinite series for the reflectivity of the stack, and sum the series. Hence show that there will be no reflection at all, if the refractive index of Medium 3 is \(n_3 = \sqrt{n_1n_2}\) and if its thickness \(d\) represents a quarter of a wavelength in the medium. What application might this have?
5.7. Derive the Fresnel coefficients for TM incidence on a dielectric interface.

5.8. The figure below shows a plane wave incident from air on a glass slab of finite thickness. Show that there is no reflection from either interface, for TM incidence at the Brewster angle $\theta_B$.

5.9. Show (a) analytically, and (b) numerically, using data from Question 5.3, that the Fresnel coefficients for TE incidence satisfy the relation:

$$\Gamma^2 + T^2 \left( \frac{n_2}{n_1} \right) \left[ \cos(\theta_2) / \cos(\theta_1) \right] = 1.$$ 

5.10. Interpret the result of Question 5.9 in terms of power conservation.
SUGGESTIONS FOR FURTHER READING

Kretschmann E. "The determination of the optical constants of metals by the excitation of surface plasmons" Z. Phys. 214, 313-324 (1971)