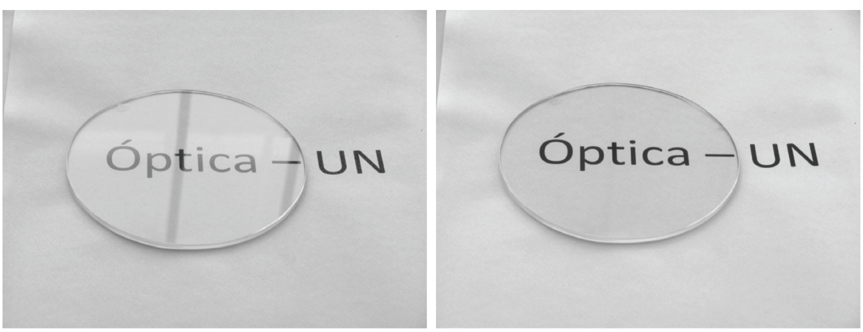
Chapter 2
Polarization

According to classical physics, light is an electromagnetic wave and its properties are obtained from Maxwell's equations. One of these properties is that light is a transverse wave; i.e., the electric and magnetic vectors (optical field) vibrate orthogonally to the direction of wave propagation. If we assume that the light source is composed of oscillators that emit electromagnetic energy, then in general, the directions of the electric and magnetic vectors are random. However, it is possible to maintain the vibration of the resulting electric (magnetic) vector in a fixed plane or following an elliptical or circular curve. In such a case, the wave is said to be polarized. This chapter defines polarization and shows some of its applications (Fig. 2.1).

Taking into account the linearity of Maxwell's equations, one can limit the study of polarization to plane harmonic waves. Although the emitted or reflected optical field can have any form, Fourier analysis shows that the complex form of the optical field wavefront can be expressed by the sum of



(a) Image with nonpolarized light. (b) Image with polarized light.

Figure 2.1 Polarization by reflection. The light that enters through a window in a room is not polarized. When reflecting off a glass plate (smooth surface), as in (a), part of the window is visible along with the text below the glass plate. If the reflection is viewed at an angle close to Brewster's angle, the light will be linearly polarized, which is verified by placing a linear polarizer between the glass plate and the photographic camera taking the image. This eliminates reflected light, and the text below the page is seen clearly, as shown in (b).

harmonic plane waves. Thus, the results for plane waves can be extended to more complex forms of the optical field.

This chapter begins by developing the algebra to describe linear, elliptical, and circular polarization. Among the polarization mechanisms, dichroism, polarization by total internal and external reflection, and birefringence are discussed in detail, with the latter limited to the case in which the principal directions of the refractive indices coincide with the axes of the crystal glass. The refractive media considered here are dielectrics without absorption. Finally, the Jones formalism to describe polarization states and polarizing elements is presented.

2.1 Plane Waves and Polarized Light

In a vacuum, for a vector point $\mathbf{r} = (x, y, z)$ and time t, the optical field is described by the electric vector $\mathbf{E}$ and the magnetic vector $\mathbf{H}$, which are related to each other according to Maxwell's equations, given by

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t}, \tag{2.1}$$

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \tag{2.2}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{2.3}$$

$$\nabla \cdot \mathbf{H} = 0. \tag{2.4}$$

From Eqs. (2.1), (2.2), and (2.3), the wave equation for the electric field is*

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}, \tag{2.5}$$

with $c^2 = 1/\mu_0 \epsilon_0$. For the magnetic field, an equation analogous to Eq. (2.5) is obtained.

Because $\mathbf{E}(x, y, z) = \{E_x(x, y, z), E_y(x, y, z), E_z(x, y, z)\}$ for a time t, Eq. (2.5) represents a set of three equations, one for each component of the electric field $\mathbf{E}$. If any of these components is represented by $V = V(x, y, z)$, we have a scalar equation of the form

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 V}{\partial t^2}. \tag{2.6}$$

Let $\hat{\mathbf{s}} = (s_x, s_y, s_z)$ be a unit vector in a fixed direction in space. A solution of Eq. (2.6) of the form $V(\mathbf{r}, t) = g(\mathbf{r} \cdot \hat{\mathbf{s}}, t)$ represents a homogeneous plane

*The wave equation is obtained using the identity vector $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2(\mathbf{E})$.

wave propagating in the $\hat{\mathbf{s}}$ direction, since at a given time g is constant for the planes $\mathbf{r} \cdot \hat{\mathbf{s}} = a$ (with a constant).

In particular, a harmonic plane wave can be written as

$$V(\mathbf{r}, t) = A_0 \cos(k\mathbf{r} \cdot \hat{\mathbf{s}} - \omega t + \delta), \tag{2.7}$$

where A_0 is the amplitude of the wave. In addition, $k = 2\pi/\lambda$, where λ is the wavelength in vacuum, called the wavenumber; $\omega = 2\pi\nu$, where ν is the wave temporal frequency, called the angular frequency; and δ is the initial phase shift. Then, the phase of the wave is composed of three terms: a spatial phase given by the surface $\varphi(x, y, z) = k\mathbf{r} \cdot \hat{\mathbf{s}}$, a temporal phase ωt, and a constant phase δ that allows the value to be adjusted at the origin (spatial and/or temporal). Surfaces of constant spatial phase are called *wavefronts* and correspond to the geometrical wavefronts derived from Fermat's principle in Chapter 1.

2.1.1 Maxwell's equations with plane waves

If in Maxwell's equations the fields $\mathbf{E}$ and $\mathbf{H}$ describe plane waves, these differential equations simplify to algebraic equations. To see this, let us write $\mathbf{E}$ and $\mathbf{H}$ in complex form, i.e.,

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \tag{2.8}$$

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \tag{2.9}$$

where $\mathbf{k}$ is the wave vector whose modulus is the angular wavenumber and whose direction is that of the unit vector $\hat{\mathbf{s}}$; i.e., $\mathbf{k} = k\hat{\mathbf{s}}$. Although the term of the initial phase has been omitted to simplify the treatment, it will be included again when required. Then, the result of the operators ∇ and $\partial/\partial t$ on plane waves is

$$\nabla \times \mathbf{E} = i(\mathbf{k} \times \mathbf{E}), \tag{2.10}$$

$$\nabla \cdot \mathbf{E} = i(\mathbf{k} \cdot \mathbf{E}), \tag{2.11}$$

$$\frac{\partial \mathbf{E}}{\partial t} = -i\omega \mathbf{E} \tag{2.12}$$

for the field $\mathbf{E}$. For the field $\mathbf{H}$, similar relationships are obtained. Therefore, Maxwell's equations can be reduced to

$$\mathbf{k} \times \mathbf{E} = \mu_0 \omega \mathbf{H}, \tag{2.13}$$

$$\mathbf{k} \times \mathbf{H} = -\epsilon_0 \omega \mathbf{E}, \tag{2.14}$$

$$\mathbf{k} \cdot \mathbf{E} = 0, \tag{2.15}$$

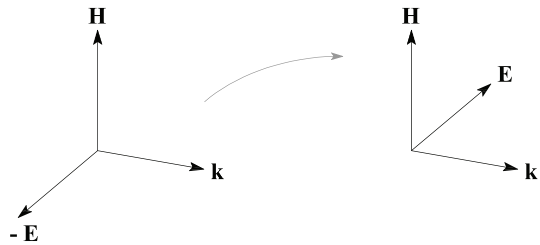


Figure 2.2 Orientation of the fields **E** and **H** and the wave vector **k**.

$$\mathbf{k} \cdot \mathbf{H} = 0. \tag{2.16}$$

From these equations,

$$\mathbf{E} = -\frac{1}{\epsilon_0 \omega} \mathbf{k} \times \mathbf{H}, \tag{2.17}$$

which with Eq. (2.16) implies that **E**, **H**, and **k** form an orthogonal system of vectors, as illustrated by Fig. 2.2.

Let $H = |\mathbf{H}|$ (the modulus of **H**) and $E = |\mathbf{E}|$ (the modulus of **E**). Given the mutual orthogonality between **E**, and **H**, and **k** from Eq. (2.14),

$$H = \epsilon_0 c E, \tag{2.18}$$

where $c = \omega / k$.

Thus the fields **E** and **H** vibrate in phase [Eqs. (2.8) and (2.9)] in a plane orthogonal to **k** and propagate as illustrated by Fig. 2.3.

2.1.2 Irradiance

Experimentally, in the visible range, the field **E** is not measured due to the lack of detectors that can respond as fast as the **E** vibrations. Instead of measuring the amplitude of the field, a time average of the square of the field can be measured, i.e., the average energy per unit time per unit area. The time

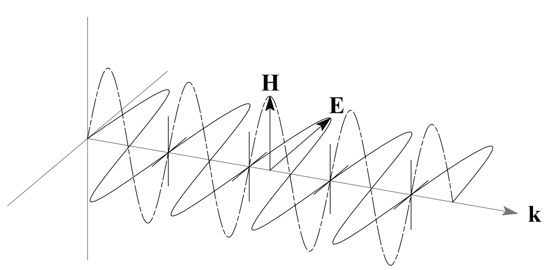


Figure 2.3 Illustration of the propagation of harmonic fields **E** and **H**.

required to do the averaging is determined by the response of the detector. Detector response times are several orders of magnitude of the time period.

Formally, the mean value of the energy is calculated from the Poynting vector

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}. \tag{2.19}$$

From the shared orthogonality between **E**, **H**, and **k**, taking into account Eq. (2.13), the Poynting vector can be written as

$$\mathbf{S} = \frac{1}{\mu_0 \omega}(\mathbf{E} \cdot \mathbf{E})\mathbf{k}; \tag{2.20}$$

and because $\omega/k = c$,

$$\mathbf{S} = \frac{1}{\mu_0 c}(\mathbf{E} \cdot \mathbf{E})\hat{\mathbf{s}} = \epsilon_0 c(\mathbf{E} \cdot \mathbf{E})\hat{\mathbf{s}}. \tag{2.21}$$

This formula indicates that the direction in which the energy flows is normal to the wavefront, since $\hat{\mathbf{s}}$ is the unit vector that defines the normal of the wavefront. This result is also valid in dielectric (isotropic) media.

The irradiance is then defined as

$$I = \langle S \rangle_{\check{T}}, \tag{2.22}$$

where $S = |\mathbf{S}|$. Here, $\langle\, \rangle$ denotes the mean value of the function* and $\check{T}$ denotes the integration (detection) time. By inserting Eq. (2.21) into Eq. (2.22), the irradiance is

$$I = \frac{\epsilon_0 c}{\check{T}} \int_{-\check{T}/2}^{\check{T}/2} (\mathbf{E} \cdot \mathbf{E}) dt. \tag{2.23}$$

In particular, for a periodic function, the mean value is taken with respect to the period of the signal. So for a harmonic plane wave, the irradiance will be

$$I = \frac{\epsilon_0 c}{T} \int_{-T/2}^{T/2} (\mathbf{E} \cdot \mathbf{E}) dt, \tag{2.24}$$

where $T = 1/\nu$ is the period of the wave.

*The mean value of the function f over time period τ is defined as $\langle f \rangle = \frac{1}{\tau} \int_{-\tau/2}^{\tau/2} f dt$.

If the harmonic plane wave is given by

$$\mathbf{E} = \mathbf{E_0} \cos(k\mathbf{r} \cdot \hat{\mathbf{s}} - \omega t), \tag{2.25}$$

the irradiance would be

$$I = \epsilon_0 c(\mathbf{E_0} \cdot \mathbf{E_0}) \frac{1}{T} \int_{-T/2}^{T/2} \cos^2(k\mathbf{r} \cdot \hat{\mathbf{s}} - \omega t) dt. \tag{2.26}$$

The mean value of the cosine-squared function in a period is equal to 1/2; therefore;

$$I = \frac{\epsilon_0 c}{2} (\mathbf{E_0} \cdot \mathbf{E_0}) = \frac{\epsilon_0 c}{2} E_0{}^2. \tag{2.27}$$

Another common way of representing the harmonic plane wave, also used in this book, is through the complex exponential function. Suppose the wave is given by

$$\mathbf{E} = \mathbf{E_0} e^{i(k\mathbf{r} \cdot \hat{\mathbf{s}} - \omega t)}, \tag{2.28}$$

where the amplitude $\mathbf{E_0}$ is also complex and $\|\mathbf{E}_0\| = E_0$. In this case, the irradiance should be defined as

$$I = \frac{\epsilon_0 c}{2T} \int_{-T/2}^{T/2} (\mathbf{E} \cdot \mathbf{E}^*) dt, \tag{2.29}$$

where $\mathbf{E}^*$ is the conjugate of $\mathbf{E}$. In this way, it is guaranteed that the irradiance value is equal to the one obtained if the wave is represented as a cosine (or sine) function. By inserting Eq. (2.28) into Eq. (2.29), the irradiance is

$$I = \frac{\epsilon_0 c}{2} (\mathbf{E_0} \cdot \mathbf{E_0^*}). \tag{2.30}$$

2.1.3 Natural light and polarized light

Because $\mathbf{E}$ and $\mathbf{H}$ are in phase and related according to Eq. (2.17), we will only consider the vector $\mathbf{E}$ to refer to the propagation of the optical field.

Let us assume that the field $\mathbf{E}$ is the result of monochromatic plane waves emitted by the oscillators that make up a light source. Simplifying the model, let us also assume that the waves propagate in the z direction, so that at a given instant of time in an (x, y) plane at a distance z, the field $\mathbf{E}$ can be written as

$$\mathbf{E}(x, y, z; t) = \left\{ |E_{ox}| e^{i\delta_x} e^{i(kz-\omega t)}, \; |E_{oy}| e^{i\delta_y} e^{i(kz-\omega t)} \right\}. \tag{2.31}$$

The amplitude of each field component has a phase term (δ_x and δ_y), so their difference allows us to measure the delay of one component with respect to the other. Choosing the plane $z = z_0$, from the phase difference $\Delta\delta = \delta_y - \delta_x$, we can observe how the vector **E** evolves in time. This depends on the nature of the source. In general, the oscillations in the sources are such that $\Delta\delta$ is a random variable. Consequently, we cannot predict how the vector **E** evolves in time (what amplitude and direction it has at a given moment). It is said that these types of sources, very common in nature, emit natural polarized or nonpolarized light. But if $\Delta\delta$ remains stable in time, i.e., $\Delta\delta = \text{constant}$, then it is possible to determine how the vector **E** evolves. In this case, it is said that the light is polarized.

2.1.4 Elliptical, circular, and linear polarization

Polarized light implies that, given two components for the field **E**, the phase difference of the amplitude components is a constant. Depending on the value of this difference, the electric vector evolves confined to a plane or following an ellipse (circle). In the first case, there is linear polarization; in the second, there is elliptical (circular) polarization to the left or to the right. To see this, let us consider an example. Suppose

$$\mathbf{E}(x, y, z; t) = \left\{ E_{ox}, E_{oy} \right\} e^{i(kz-\omega t)} \tag{2.32}$$

with

$$E_{ox} = |E_{ox}| e^{i0} \quad \text{and} \quad E_{oy} = |E_{oy}| e^{i\pi/2}; \tag{2.33}$$

i.e., $\delta_x = 0$ and $\delta_y = \pi/2$. Therefore, $\Delta\delta = \pi/2$. First, let us see what the endpoint of the electric vector projected onto a plane, say $z = 0$, looks like. The components of **E** take the form

$$E_x = |E_{ox}| e^{-i2\pi t/T}, \tag{2.34}$$

$$E_y = |E_{oy}| e^{-i(2\pi t/T - \pi/2)}. \tag{2.35}$$

Considering each component as a phasor, i.e., a rotating vector of radius $|E_{ox}|$ ($|E_{oy}|$) and a phase δ_x (δ_y), graphically each component will look as illustrated in Fig. 2.4(a). The y component (top left), at $t = 0$, starts with a lag of $\pi/2$, i.e., at point 1. On the other hand, the x component (bottom), at $t = 0$, starts with a phase shift of 0, which is also indicated by the dot 1. As time increases, particularly for $t = T/4$, $T/2$, $3T/4$, and $2T$, the positions for the phasors for x and y will be indicated by points 2, 3, 4, and 5, respectively. The

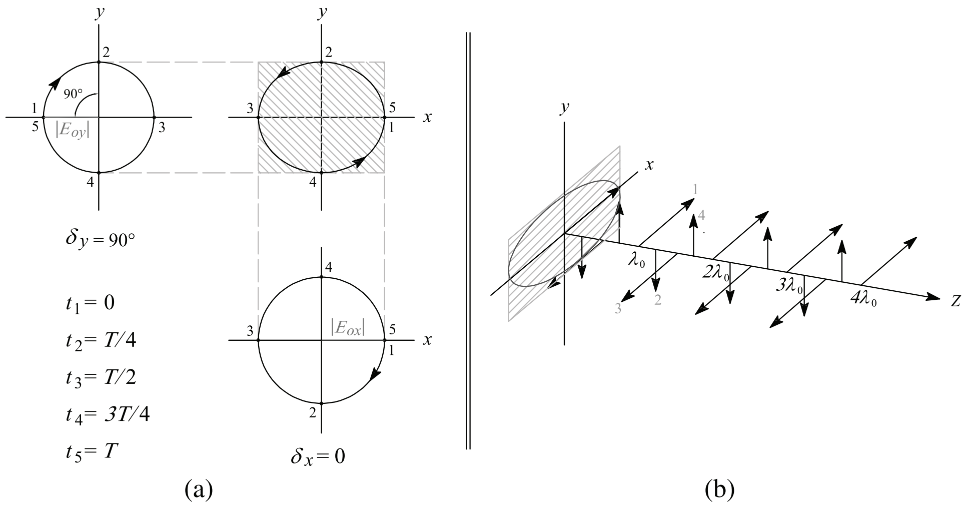


Figure 2.4 Left elliptical polarization. (a) Endpoint of the electric vector projected onto a fixed plane ($z = 0$). The end of the vector rotates counterclockwise. (b) Spatial evolution ($t = 0$) of the electric vector. The end of the vector describes a helix that advances in the z direction in a clockwise rotation.

projections of points 1, 2, 3, 4, and 5 in the vertical direction for the x phasor and in the horizontal direction for the y phasor intersect in the shaded rectangular region with sides $2|E_{ox}|$ and $2|E_{oy}|$. The locus of the intersections describes the evolution of the electric vector in time, as shown at the top right of Fig. 2.4(a). In this example, the path is an ellipse, since $|E_{ox}| > |E_{oy}|$, and it builds counterclockwise. In this case, the electric vector is said to have *left elliptical polarization*. The opposite direction of the trajectory (clockwise) is obtained if $\Delta\delta = -\pi/2$; then the electric vector is said to have *right elliptical polarization*. In particular, if $|E_{ox}| = |E_{oy}|$, the trajectory is a circle, corresponding to left or right circular polarization.

Now if we look at how the electric vector evolves in space, we see something different. Fixing the time at $t = 0$, the **E** components take the form

$$E_x = |E_{ox}| e^{i2\pi z/\lambda}, \tag{2.36}$$

$$E_y = |E_{oy}| e^{i(2\pi z/\lambda + \pi/2)}. \tag{2.37}$$

For the real part, in $z = 0$, $E = \{|E_{ox}|, 0\}$; in $z = \lambda/4$, $E = \{0, -|E_{oy}|\}$; in $z = \lambda/2$, $E = \{-|E_{ox}|, 0\}$; in $z = 3\lambda/4$, $E = \{0, |E_{oy}|\}$; and finally, in $z = \lambda$, $E = \{|E_{ox}|, 0\}$. This is illustrated by arrows in Fig. 2.4(b). In this case, the path followed by the end of the electric vector is a helix moving clockwise. Of course, the two images are equivalent. Indeed, if in (b) we observe how the ends of the electric vector reach the plane $z = 0$, we will have the trajectory given in (a).

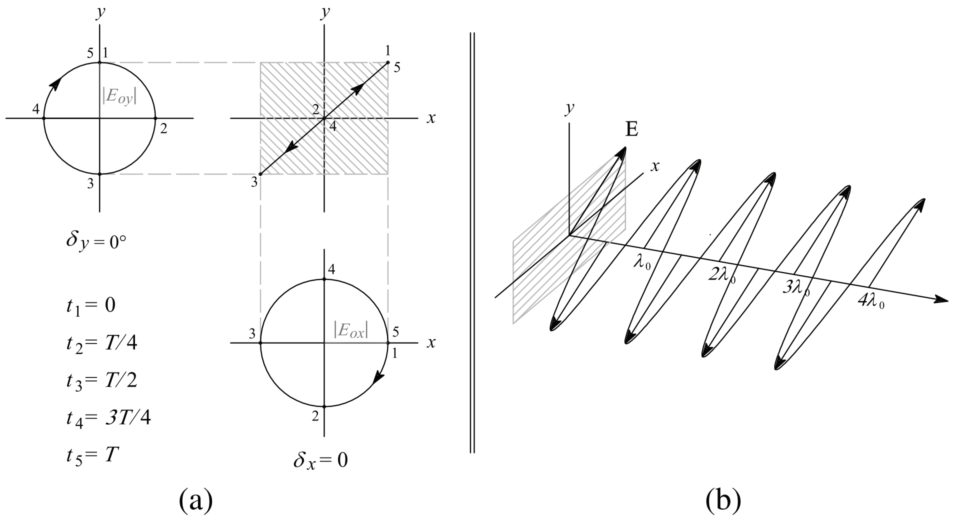


Figure 2.5 Linear polarization when the two components are in phase. (a) The trajectory of the end of the electric vector is a straight line bounded by the rectangular region of sides $2|E_{ox}|$ and $2|E_{oy}|$, tilted by an angle $\tan^{-1}(|E_{oy}|/|E_{ox}|)$. (b) The field **E** is confined to a plane inclined by an angle $\tan^{-1}(|E_{oy}|/|E_{ox}|)$ and moves harmonically.

When $\Delta\delta = 0$, the observed path of the endpoint of the electric vector in a fixed plane ($z = 0$) is a straight line inclined at an angle $\arctan(|E_{oy}|/|E_{ox}|)$, as illustrated in Fig. 2.5(a). The extension of the line is limited by the rectangle with sides $2|E_{ox}|$ and $2|E_{oy}|$. This is the case of linear polarization.

If the spatial evolution of the electric vector is observed, the field **E** will vibrate harmonically in an inclined plane with the angle $\arctan(|E_{oy}|/|E_{ox}|)$, as illustrated in Fig. 2.5(b).

2.1.5 Polarization: general case

In the previous section, left elliptical and linear polarization are shown, which result from the phase differences $\Delta\delta = \pi/2$ and $\Delta\delta = 0$. However, the value of $\Delta\delta$ can be any constant. In general, the polarization state of an electromagnetic wave is given by a rotated ellipse,

$$\frac{E_x^2}{|E_{ox}|^2} - 2\frac{E_x}{|E_{ox}|}\frac{E_y}{|E_{oy}|}\cos(\Delta\delta) + \frac{E_y^2}{|E_{oy}|^2} = \sin^2(\Delta\delta), \tag{2.38}$$

as shown in Fig. 2.6. The ellipse remains confined to the rectangle of sides $2|E_{ox}|$ and $2|E_{oy}|$ and the angle of rotation ψ of the ellipse is determined from

$$\tan(2\psi) = \tan(2\alpha)\cos(\Delta\delta), \tag{2.39}$$

where

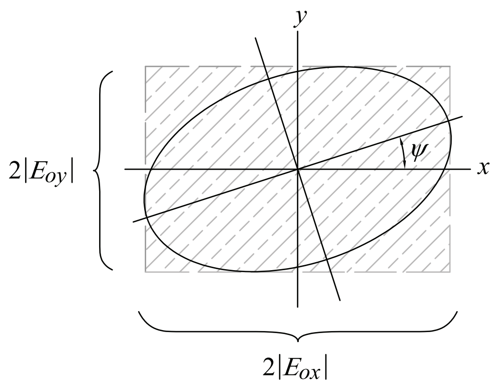


Figure 2.6 Polarization ellipse in the general case.

$$\tan\alpha = \frac{|E_{oy}|}{|E_{ox}|}. \tag{2.40}$$

The sign of sin(Δδ) indicates the direction of the polarization. If sin(Δδ) > 0, the polarization is to the left; if sin(Δδ) < 0, the polarization is to the right. Appendix F shows the complete derivation to obtain Eqs. (2.38), (2.39), and (2.40).

Case 1. $\Delta\delta = m\pi$; $m = 0, \ \pm 1, \ \pm 2, \ \ldots,$
Equation (2.38) becomes

$$\frac{E_x^2}{|E_{ox}|^2} - 2(-1)^m \frac{E_x}{|E_{ox}|}\frac{E_y}{|E_{oy}|} + \frac{E_y^2}{|E_{oy}|^2} = 0, \tag{2.41}$$

which is equal to

$$\left(\frac{E_x}{|E_{ox}|} - (-1)^m \frac{E_y}{|E_{oy}|}\right)^2 = 0, \tag{2.42}$$

which represents the straight line

$$E_y = (-1)^m(\tan\alpha)E_x, \tag{2.43}$$

i.e., linear polarization with the plane of vibration inclined by the angle ±α.

Case 2. $\Delta\delta = (2m-1)\pi/2;\ m = 0,\ \pm 1,\ \pm 2,\ \ldots,$
Equation (2.38) becomes

$$\frac{E_x^2}{|E_{ox}|^2} + \frac{E_y^2}{|E_{oy}|^2} = 1, \tag{2.44}$$

which represents an ellipse with its major and minor axes oriented with the axes x (E_x) and y (E_y), respectively. Thus, there is elliptical polarization to the left or to the right. In particular, if $|E_{ox}| = |E_{oy}|$, there is left or right circular polarization.

Case 3. $\Delta\delta \neq m\pi,\ (2m-1)\pi/2$.
For phase shifts different from those treated in cases 1 and 2, there is elliptical polarization, in which the ellipse is rotated according to Eqs. (2.39) and (2.40).

Example: ellipse from line to circle and vice versa
To illustrate these three cases, suppose we have an optical field with $|E_{ox}| = |E_{oy}|$ and $\Delta\delta = 0$, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, $7\pi/4$, and 2π. In this example, because $\tan 2\alpha = \infty$, the angle of rotation of the general ellipse will be given by $\tan 2\psi = \pm\infty$, where the sign is given by the sign of $\cos(\Delta\delta)$. In other words, the angle of rotation of the ellipse would be $\psi = 45°$, if $(0 \leq \Delta\delta < \pi/2) \cup (3\pi/2 < \Delta\delta \leq 2\pi)$ and $\psi = -45°$, if $\pi/2 < \Delta\delta < 3\pi/2$.

In Fig. 2.7, the polarization states for the phase changes mentioned above are shown: $\mathbf{E}_{+45}$ denotes linear polarization with the plane of vibration at 45°, $\mathbf{E}_{-45}$ denotes linear polarization with the plane of vibration at −45°, $\mathbf{E}_L$ indicates left circular or elliptical polarization, and $\mathbf{E}_R$ indicates right circular or elliptical polarization.

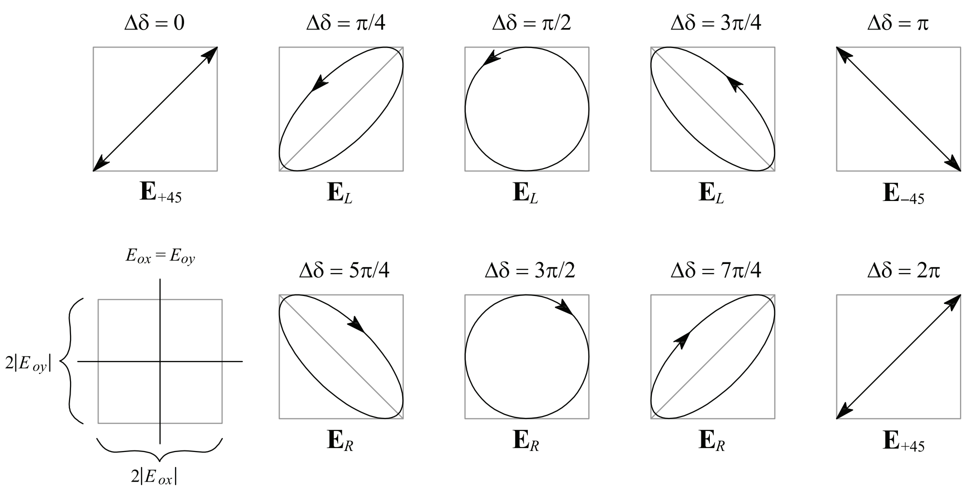


Figure 2.7 Polarization states for various phase shifts with $|E_{ox}| = |E_{oy}|$.

On the other hand, $\alpha = 0$ defines a horizontal linear polarization state ($\mathbf{E}_H$) and $\alpha = 90°$ defines a vertical linear polarization state ($\mathbf{E}_V$).

2.2 Dichroism Polarization

One way to remove one of the **E**-field components is by absorbing that component. This can be achieved by designing a device that performs this task or by using a natural material that has this property [1]. In both cases, the selective absorption of one of the **E**-field components is called *dichroism*. The final effect on the field will be linearly polarized light.

2.2.1 Linear polarizer

To see how a dichroism-based linear polarizer works, suppose that a grid of parallel conducting wires is constructed, as shown in Fig. 2.8, and an unpolarized field **E** (natural light) is incident in a direction orthogonal to the grid plane. Because the electric charges have the possibility of greater displacement in the horizontal direction (along the wires) compared with the vertical direction (cross section of the wires), there will be a greater absorption of electric energy in the direction of the wires; thus, the net component E_x experiences a greater attenuation than the net component E_y. If, ideally, the component E_x is completely attenuated, we will have a linear polarizer and the field will have a vertical linear polarization state, $\mathbf{E}_V$. The direction in which the field is not attenuated is called the *transmission axis of the linear polarizer*.

Using lithographic methods, polarizers based on a grid of conductive wires for the visible spectrum are manufactured, achieving arrangements with a separation of 100 nm between wires. Aluminum microwires are deposited on glass substrates.

The most common dichroic linear polarizers are made of sheets of a special transparent plastic (polyvinyl alcohol). These sheets have been stretched in one direction to align their long molecules, which are then

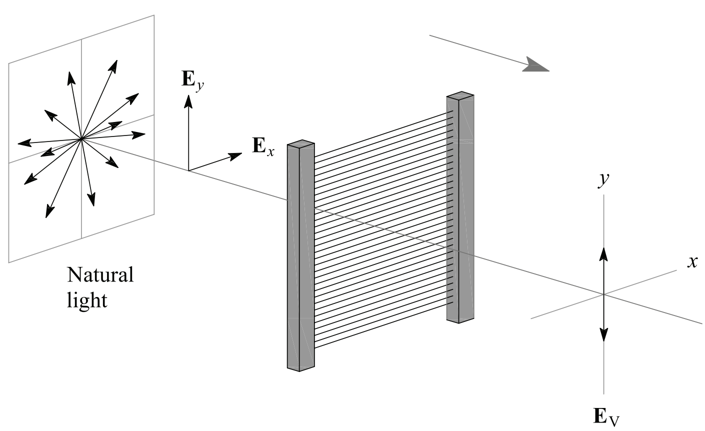


Figure 2.8 Linear polarizer made with a grid of conducting wires.

coated with iodine. In this way, something similar to the arrangement of the threads shown in Fig. 2.8 is obtained, but at a microscopic level (type H polarizers).

Extinction coefficient and degree of polarization

In practice, it is not possible to completely attenuate the component orthogonal to the transmission axis of the polarizer, and the component parallel to the transmission axis of the polarizer is not completely transmitted. If we represent the linear polarizer, as shown in Fig. 2.9, and we decompose the resulting incident field on the polarizer into a component parallel to the transmission axis, $\mathbf{E}_{\parallel}$, and into a component orthogonal to the transmission axis, $\mathbf{E}_{\perp}$, then the incident field would be

$$\mathbf{E} = \{E_{\parallel}, E_{\perp}\}e^{i(kz-\omega t)}. \tag{2.45}$$

To characterize the linear polarizer taking into account the absorption of the components $\mathbf{E}_{\parallel}$ and $\mathbf{E}_{\perp}$, two quantities are defined: the *extinction coefficient*,

$$\rho_P = \frac{t_{\parallel}}{t_{\perp}}, \tag{2.46}$$

and the *degree of polarization*,

$$P_P = \frac{t_{\parallel} - t_{\perp}}{t_{\parallel} + t_{\perp}}, \tag{2.47}$$

where $t_{\parallel} = |E'_{\parallel}|/|E_{\parallel}|$ is the fraction that transmits the component parallel to the transmission axis and $t_{\perp} = |E'_{\perp}|/|E_{\perp}|$ is the transmission fraction of the

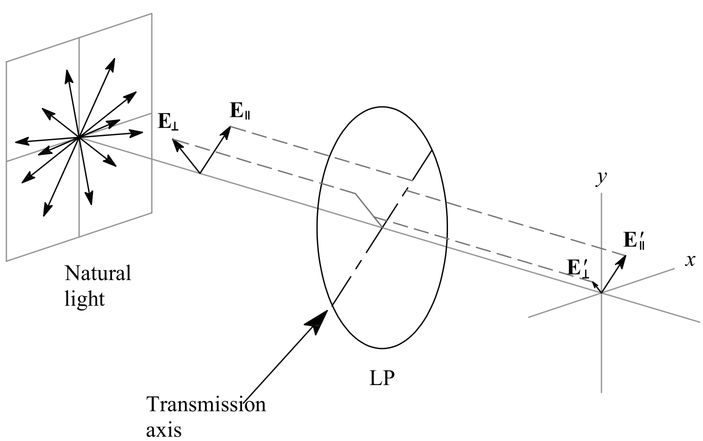


Figure 2.9 In a real polarizer, 100% of the component parallel to the transmission axis is not transmitted, and the component orthogonal to the transmission axis is not completely canceled.

Table 2.1 Technical specifications for a linear dichroic polarizer.

Bandwidth (nm)	100–1000
Extinction	10^2–10^6
Transmission (%)	>50, <90
Optical quality	λ/5–λ/2

component orthogonal to the transmission axis. The extinction ratio is usually given as ρ_P:1. The transmittance of the polarizer is also used, defined as the fraction of the intensity of linearly polarized light parallel to the transmission axis that is transmitted by the polarizer, $T = t_{\parallel}^2$. For example, it is typical to find technical specifications for linear dichroic polarizers such as those shown in Table 2.1.

The bandwidth tells us the spectral range for which the extinction coefficient and the transmission are maintained according to the nominal value. A 100:1 extinction corresponds to a low quality polarizer, whereas a 1,000,000:1 extinction is a high-quality polarizer. On the other hand, near 100% transmission requires a high-quality polarizer. Ultimately, the optical quality of the polarizing foil comes down to the maximum distortion the foil generates in a flat wavefront once it passes through the polarizer.

2.2.2 Malus' law

Combining several linear polarizers not only allows control of the resulting polarization state, but also the irradiance. Suppose we have several ideal linear polarizers, i.e., with $\rho_P = \infty$ and $P_P = 1$. In Fig. 2.10, a configuration with two polarizers is shown where their transmission axes form an angle θ. The first polarizer receives natural light. The second receives linearly polarized light vibrating in planes parallel to the transmission axis of the

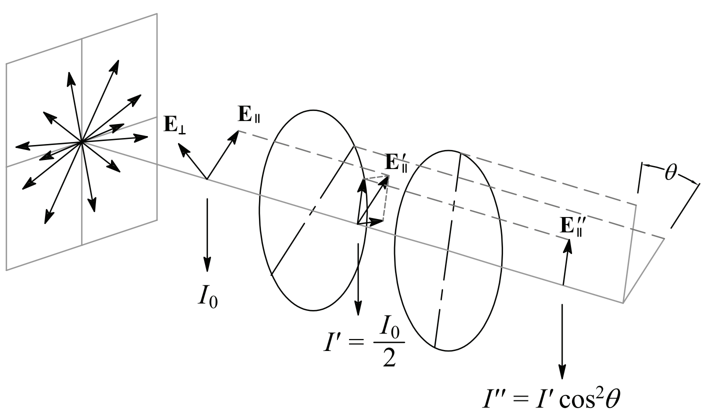


Figure 2.10 Malus' law.

first polarizer and transmits linearly polarized light, but changes the orientation of the vibrating planes parallel to the transmission axis of the second polarizer.

To determine the irradiance obtained at the end of the second polarizer, let us assume that the irradiance of natural light is I_0. This irradiance is the mean value of the energy flow per unit area, emitted by the oscillators that make up the light source. The oscillators are randomly oriented and emit short-duration (10^{-8} s) electromagnetic wave trains. When looking at plane waves at a great distance from the oscillator, each wave train will be linearly polarized. Therefore, there will be a superposition of wave trains with different initial phases (origin of time) and with different planes of occasional vibration. Because the integration time is much longer than the duration of the wave trains, it is not possible to predict the state of polarization. At a given moment, we will have a resultant vector for the field **E** and then, an instant later, it will have randomly changed its orientation and amplitude. However, we can define an average vector in a given direction. Because the process is random for any other direction, we will have a mean vector with the same amplitude as the first. Consequently, if we decompose the vector **E** incident on the first polarizer into a component parallel to the transmission axis and another one orthogonal to the transmission axis [Eq. (2.45)], then

$$I_0 = \frac{\epsilon_0 c}{2}\{E_{\parallel}, E_{\perp}\} \cdot \{E_{\parallel}, E_{\perp}\}^*, \tag{2.48}$$

which is equal to

$$I_0 = \frac{\epsilon_0 c}{2}(|E_{\parallel}|^2 + |E_{\perp}|^2). \tag{2.49}$$

And, since natural light is randomly polarized, $|E_{\parallel}| = |E_{\perp}|$; thus,

$$I_0 = \frac{\epsilon_0 c}{2}(2|E_{\parallel}|^2). \tag{2.50}$$

After the first polarizer, $\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}$; thus, the transmitted irradiance would be

$$I' = \frac{\epsilon_0 c}{2}|E'_{\parallel}|^2 = \frac{I_0}{2}. \tag{2.51}$$

Finally, after the second polarizer $E''_{\parallel} = E'_{\parallel}\cos\theta$; therefore, the transmitted irradiance will be

$$I'' = \frac{\epsilon_0 c}{2}|E''_{\parallel}|^2 = \frac{\epsilon_0 c}{2}|E'_{\parallel}|^2\cos^2\theta, \tag{2.52}$$

i.e.,

$$I'' = I'\cos^2\theta. \tag{2.53}$$

This expression, known as *Malus' law*, states that when two linear polarizers are placed one behind the other with their transmission axes parallel to each other, the second polarizer transmits 100% of the linearly polarized light emerging from the first polarizer ($I'' = I_0/2$); and when two linear polarizers are placed one behind the other with their transmission axes orthogonal to each other, the second polarizer transmits 0% of the linearly polarized light emerging from the first polarizer ($I'' = 0$). For other orientations of the transmission axes, the transmission will have a value in the range $0 < I'' < I'$ according to Eq. (2.53). In other words, the two-polarizer system shown in Fig. 2.10 is an irradiance-attenuating device (which also rotates the plane of vibration of the transmitted **E** field).

2.3 Polarization by Reflection

Another way to generate linearly polarized light from natural light is by reflecting it off a smooth surface (mirror) at an appropriate angle (Brewster's angle). To determine this angle, one observes how the amplitude of the reflected wave changes as a function of the angle of incidence. This depends on the boundary conditions for the incident, reflected and transmitted waves at the surface. The boundary conditions must be met for both the phases and the amplitudes of the waves. From the boundary conditions of the phases, the laws of reflection and refraction (Snell's law) are derived; from the boundary conditions of the wave amplitudes, the Fresnel equations are derived. From these equations, the condition to obtain linearly polarized light is established.

2.3.1 Laws of reflection and refraction

Let us first consider the vibration of frequency ν of the field **E** at a point in space, given by the expression

$$\mathbf{E}(t) = \mathbf{E}_0 e^{-i2\pi\nu t}. \tag{2.54}$$

Let us assume that the observation point is immersed in a homogeneous dielectric medium of refractive index n. Then the (phase) speed with which the progressive wave propagates would be $\upsilon = c/n$, where c is the speed of light in a vacuum. The expression for the progressive harmonic plane wave in the direction of the unit vector $\hat{\mathbf{s}}$ becomes

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp\left[-i2\pi\nu\left(t - \frac{\hat{\mathbf{s}}\cdot\mathbf{r}}{\upsilon}\right)\right], \tag{2.55}$$

which can be rewritten as

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp\left[i2\pi\nu\left(n\frac{\hat{\mathbf{s}}\cdot\mathbf{r}}{c} - t\right)\right] \tag{2.56}$$

or

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp\left[i\left(2\pi n\frac{\hat{\mathbf{s}}\cdot\mathbf{r}}{\lambda} - 2\pi\nu t\right)\right]. \tag{2.57}$$

Defining the wave vector in a medium of refractive index n as

$$\mathbf{k} = \frac{2\pi}{\lambda}n\hat{\mathbf{s}}, \tag{2.58}$$

the plane wave is

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.59}$$

with $\mathbf{k}$ given by Eq. (2.58).

In a homogeneous dielectric medium, Maxwell's equations are written as in Section 2.1.1, but changing ϵ_0 to ϵ, μ_0 to μ, and the wave vector $\mathbf{k}$ according to Eq. (2.58). With these changes, the wave equation becomes

$$\nabla^2\mathbf{E} = \frac{1}{\upsilon^2}\frac{\partial^2\mathbf{E}}{\partial t^2}, \tag{2.60}$$

with $\upsilon = 1/\sqrt{\epsilon\mu}$ (see Appendix B).

Assuming that we have an interface (plane surface) separating two media of refractive indices n_i and n_t, and a plane wave traveling in the medium of refractive index n_i and incident on the interface, then we will have a reflected plane wave and a transmitted (or refracted) plane wave. Let $\hat{\mathbf{s}}_i$, $\hat{\mathbf{s}}_r$, and $\hat{\mathbf{s}}_t$ be the unit vectors that indicate the direction of propagation of the incident, reflected, and transmitted plane waves, respectively, at a point of the interface located with the vector $\mathbf{r}$. The incident, reflected, and transmitted plane waves would then be given by

$$\mathbf{E}_i(\mathbf{r}, t) = \mathbf{E}_{0i} e^{i(\mathbf{k}_i\cdot\mathbf{r}-\omega t)}, \tag{2.61}$$

$$\mathbf{E}_r(\mathbf{r}, t) = \mathbf{E}_{0r} e^{i(\mathbf{k}_r\cdot\mathbf{r}-\omega t)}, \tag{2.62}$$

$$\mathbf{E}_t(\mathbf{r}, t) = \mathbf{E}_{0t} e^{i(\mathbf{k}_t\cdot\mathbf{r}-\omega t)}, \tag{2.63}$$

with $\mathbf{k}_i = (2\pi/\lambda)n_i\hat{\mathbf{s}}_i$, $\mathbf{k}_r = (2\pi/\lambda)n_r\hat{\mathbf{s}}_r$, and $\mathbf{k}_t = (2\pi/\lambda)n_t\hat{\mathbf{s}}_t$. Normally, the reflected and incident waves are in the same medium as the incident wave, so $n_r = n_i$. Because for any point on the interface the phases of all three waves must be equal, then

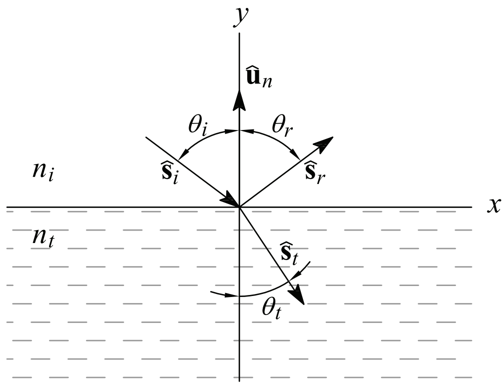


Figure 2.11 Unit vectors corresponding to the directions of propagation of the incident, reflected, and transmitted plane waves.

$$\mathbf{k}_i \cdot \mathbf{r} = \mathbf{k}_r \cdot \mathbf{r} = \mathbf{k}_t \cdot \mathbf{r}. \tag{2.64}$$

This double equality implies that the vectors $\mathbf{k}_i$, $\mathbf{k}_r$, and $\mathbf{k}_t$ are coplanar. In Fig. 2.11, the unit vectors $\hat{\mathbf{s}}_i$, $\hat{\mathbf{s}}_r$, and $\hat{\mathbf{s}}_t$ are shown with the normal unit vector $\hat{\mathbf{u}}_n$ of the interface at the point of incidence. The interface corresponds to the *xz* plane, and the plane containing the vectors $\hat{\mathbf{s}}_i$, $\hat{\mathbf{s}}_r$, and $\hat{\mathbf{s}}_t$ corresponds to the *xy* plane and is called the *plane of incidence*. The angles of incidence θ_i, reflection θ_r, and transmission θ_t are measured with respect to the normal interface at the point of incidence.

Law of reflection

From Eq. (2.64), for the reflected wave

$$(\mathbf{k}_i - \mathbf{k}_r) \cdot \mathbf{r} = 0; \tag{2.65}$$

i.e., $(\mathbf{k}_r - \mathbf{k}_i)$ is parallel to the normal unit vector of the interface. Thus,

$$(\mathbf{k}_i - \mathbf{k}_r) \| \hat{\mathbf{u}}_n. \tag{2.66}$$

Therefore,

$$\frac{2\pi}{\lambda} n_i (\hat{\mathbf{s}}_i - \hat{\mathbf{s}}_r) \times \hat{\mathbf{u}}_n = 0 \tag{2.67}$$

and

$$\hat{\mathbf{s}}_r \times \hat{\mathbf{u}}_n = \hat{\mathbf{s}}_i \times \hat{\mathbf{u}}_n, \tag{2.68}$$

from which

$$\sin\theta_r = -\sin\theta_i, \tag{2.69}$$

and the law of reflection, $\theta_r = -\theta_i$, is obtained.

Law of refraction

From Eq. (2.64), for the transmitted wave,

$$(\mathbf{k}_i - \mathbf{k}_t)\cdot\mathbf{r} = 0, \tag{2.70}$$

so the vector $(\mathbf{k}_t - \mathbf{k}_i)$ is also parallel to the normal unit vector of the interface. Thus,

$$(\mathbf{k}_i - \mathbf{k}_t)\|\hat{\mathbf{u}}_n. \tag{2.71}$$

Hence,

$$\frac{2\pi}{\lambda}(n_i\hat{\mathbf{s}}_i - n_t\hat{\mathbf{s}}_t)\times\hat{\mathbf{u}}_n = 0, \tag{2.72}$$

and

$$n_t\hat{\mathbf{s}}_t\times\hat{\mathbf{u}}_n = n_i\hat{\mathbf{s}}_i\times\hat{\mathbf{u}}_n \tag{2.73}$$

from which

$$n_t\sin\theta_t = n_i\sin\theta_i, \tag{2.74}$$

which is the law of refraction or Snell's law.

Note that Eq. (2.71) implies that

$$n_i\hat{\mathbf{s}}_i - n_t\hat{\mathbf{s}}_t = \Gamma\hat{\mathbf{u}}_n, \tag{2.75}$$

which turns out to be the vector form of the geometrical construction shown in Fig. 1.9(b), where Γ is the length of side $\overline{BC}$.

2.3.2 Fresnel equations

Let us now consider the boundary conditions for the amplitudes of the incident, reflected, and transmitted waves. Because we are dealing with dielectric media (without absorption), the tangential components of the fields **E** and **H** are continuous at the interface. It is convenient, then, to decompose vectors **E** and **H** into components orthogonal to the plane of incidence and components parallel to the plane of incidence, as shown in Fig. 2.12.

Whereas the orthogonal component of the electric field $\mathbf{E}^{\perp}$ and its corresponding magnetic field $\mathbf{H}^{\parallel}$ are shown in Fig. 2.12(a), the parallel component of the electric field $\mathbf{E}^{\parallel}$ and its corresponding magnetic field $\mathbf{H}^{\perp}$ are shown in Fig. 2.12(b). Taking into account the right-handed coordinate

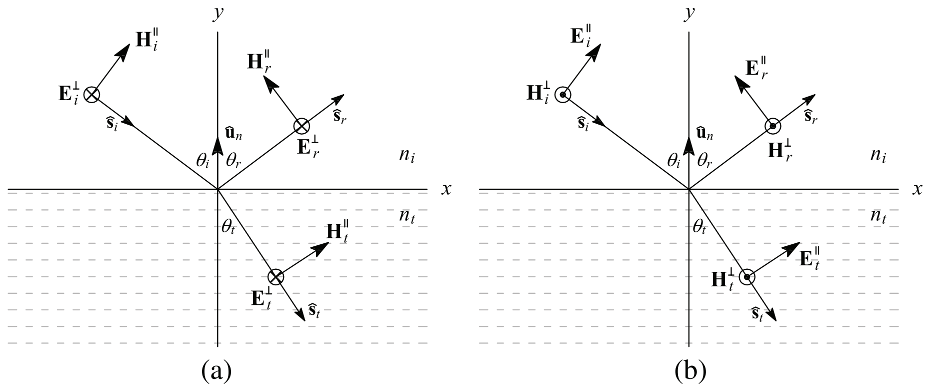


Figure 2.12 Parallel and orthogonal components of the **E** and **H** fields with respect to the plane of incidence.

system for the fields **E** and **H**, as shown in Fig. 2.12, the normal components in the positive direction are represented by a circle with a cross to indicate that the vector enters the plane of the paper (away from the reader). Similarly, the normal components in the negative direction are represented by a circle with a center point to indicate that the vector exits the plane of the paper (toward the reader). For parallel components, we also have components in the x and y directions. The vectors must have their origin at the point of incidence, but for the sake of clarity they have been drawn at some distance from the point of incidence. The orientations of the reflected and transmitted vectors can change according to the phase changes that these components experience at the interface (this will be shown later). The field $\mathbf{E}^{\perp}$ is usually called the "TE" (transverse electric field) polarization or "s" polarization. The field $\mathbf{E}^{\parallel}$ is usually called the "TM" (transverse magnetic field) polarization or "p" polarization. With these components, the amplitudes of the electric fields are written as

$$\mathbf{E}_{0i} = \{E_i^{\perp}, E_i^{\parallel}\}, \tag{2.76}$$

$$\mathbf{E}_{0r} = \{E_r^{\perp}, E_r^{\parallel}\}, \tag{2.77}$$

$$\mathbf{E}_{0t} = \{E_t^{\perp}, E_t^{\parallel}\}; \tag{2.78}$$

the magnetic fields are written as

$$\mathbf{H}_{0i} = \{H_i^{\perp}, H_i^{\parallel}\}, \tag{2.79}$$

$$\mathbf{H}_{0r} = \{H_r^{\perp}, H_r^{\parallel}\}, \tag{2.80}$$

$$\mathbf{H}_{0t} = \{H_t^{\perp}, H_t^{\parallel}\}. \tag{2.81}$$

Taking Eq. (2.18) into account for a homogeneous dielectric medium, in each case

$$H^{\parallel} = \epsilon \upsilon E^{\perp} \tag{2.82}$$

and

$$H^{\perp} = \epsilon \upsilon E^{\parallel}. \tag{2.83}$$

Applying the boundary conditions for the components of the fields, for the TE polarization state [Fig. 2.12(a)]

$$E_i^{\perp} + E_r^{\perp} = E_t^{\perp}, \tag{2.84}$$

$$H_i^{\parallel} \cos\theta_i - H_r^{\parallel} \cos\theta_r = H_t^{\parallel} \cos\theta_t. \tag{2.85}$$

Thus, Eq. (2.85) can be rewritten as

$$\epsilon_i \upsilon_i E_i^{\perp} \cos\theta_i - \epsilon_i \upsilon_i E_r^{\perp} \cos\theta_i = \epsilon_t \upsilon_t E_t^{\perp} \cos\theta_t \tag{2.86}$$

using Eq. (2.82) and the law of reflection. In a dielectric (nonmagnetic) material, the magnetic permeability can be approximated to that of a vacuum; thus, multiplying Eq. (2.86) by $\mu_0 c$, and then exchanging μ_0 for μ_i on the left-hand side of the equality and μ_0 for μ_t on the right-hand side leads to

$$n_i E_i^{\perp} \cos\theta_i - n_i E_r^{\perp} \cos\theta_i = n_t E_t^{\perp} \cos\theta_t. \tag{2.87}$$

Given the incident field, Eqs. (2.84) and (2.87) constitute a system of two equations with two unknowns, namely $E_r^{\perp}$ and $E_t^{\perp}$. Solving this system of equations for each of the unknowns leads to

$$E_r^{\perp} = \frac{n_i \cos\theta_i - n_t \cos\theta_t}{n_i \cos\theta_i + n_t \cos\theta_t} E_i^{\perp} \tag{2.88}$$

and

$$E_t^{\perp} = \frac{2n_i \cos\theta_i}{n_i \cos\theta_i + n_t \cos\theta_t} E_i^{\perp}. \tag{2.89}$$

If the boundary conditions for the components in the TM polarization state [Fig. 2.12(b)] are applied, then

$$-H_i^{\perp} - H_r^{\perp} = -H_t^{\perp}, \tag{2.90}$$

$$E_i^{\parallel}\cos\theta_i - E_r^{\parallel}\cos\theta_r = E_t^{\parallel}\cos\theta_t. \tag{2.91}$$

Using Eq. (2.83) to write Eq. (2.90) in terms of the parallel components of **E**, a procedure analogous to that used in the case of TE polarization to solve the system of equations leads to

$$E_r^{\parallel} = \frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_t\cos\theta_i + n_i\cos\theta_t}E_i^{\parallel} \tag{2.92}$$

and

$$E_t^{\parallel} = \frac{2n_i\cos\theta_i}{n_t\cos\theta_i + n_i\cos\theta_t}E_i^{\parallel}. \tag{2.93}$$

Equations (2.88), (2.89), (2.92), and (2.93) are the Fresnel equations. The coefficients multiplying $E_i^{\perp}$ and $E_i^{\parallel}$ on the right-hand side of these equations determine the fraction of the amplitude of the electric field components that is reflected and transmitted at the interface. These coefficients are

$$r_{\perp}(=r_s) = \left(\frac{E_r}{E_i}\right)_{\perp}, \quad t_{\perp}(=t_s) = \left(\frac{E_t}{E_i}\right)_{\perp} \tag{2.94}$$

for the TE polarization state and

$$r_{\parallel}(=r_p) = \left(\frac{E_r}{E_i}\right)_{\parallel}, \quad t_{\parallel}(=t_p) = \left(\frac{E_t}{E_i}\right)_{\parallel} \tag{2.95}$$

for the TM polarization state.

Then, for the reflected wave at the interface, the amplitudes of each component are $E_r^{\perp} = r_{\perp}E_i^{\perp}$ and $E_r^{\parallel} = r_{\parallel}E_i^{\parallel}$, where

$$r_{\perp} = \frac{n_i\cos\theta_i - n_t\cos\theta_t}{n_i\cos\theta_i + n_t\cos\theta_t}, \tag{2.96}$$

$$r_{\parallel} = \frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_t\cos\theta_i + n_i\cos\theta_t}. \tag{2.97}$$

For the wave transmitted at the interface, the amplitudes of each component are $E_t^{\perp} = t_{\perp}E_i^{\perp}$ and $E_t^{\parallel} = t_{\parallel}E_i^{\parallel}$, where

$$t_{\perp} = \frac{2n_i\cos\theta_i}{n_i\cos\theta_i + n_t\cos\theta_t}, \tag{2.98}$$

$$t_{\parallel} = \frac{2n_i\cos\theta_i}{n_t\cos\theta_i + n_i\cos\theta_t}. \tag{2.99}$$

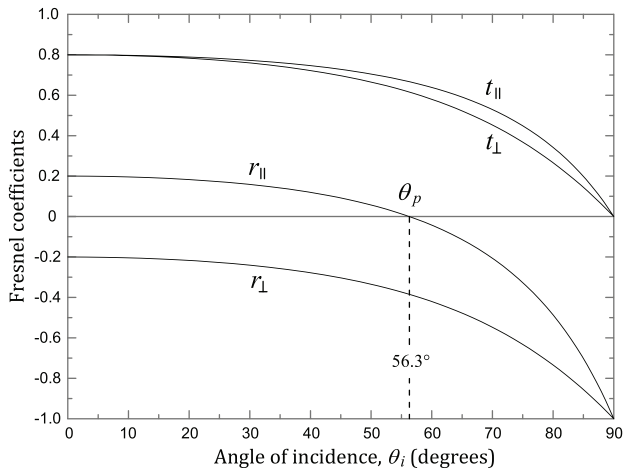


Figure 2.13 Fresnel coefficients for reflection and transmission at an air ($n_i = 1.0$) – glass ($n_t = 1.5$) interface. The polarization angle is θ_p.

The behavior of the reflection coefficients $r_{\perp}$ and $r_{\parallel}$ and the transmission coefficients $t_{\perp}$ and $t_{\parallel}$, as a function of the angle of incidence θ_i, is shown in Fig. 2.13, when the refractive indices are $n_i = 1.0$ (air) and $n_t = 1.5$ (glass).

From Fig. 2.13, the following facts can be established:

- At normal incidence, $\theta_i = 0$, the highest transmission is obtained, with $t_{\perp} = 0.8$ and $t_{\parallel} = 0.8$. For reflection, $r_{\parallel} = 0.2$ and $r_{\perp} = -0.2$. The negative sign for $r_{\perp}$ indicates that at the interface, while the incident vector points toward the paper, the reflected vector points away from the paper. Then it is said that the orthogonal component undergoes a phase change of $\pm\pi$. Therefore, $E_r^{\perp} = 0.2e^{\pm i\pi} E_i^{\perp}$ at $\theta_i = 0$.
- At grazing incidence, $\theta_i \to 90°$, the transmission tends to zero for both $t_{\perp}$ and $t_{\parallel}$. For reflection, both $r_{\parallel}$ and $r_{\perp}$ tend to –1; i.e., the incident field is completely reflected and each of the components suffers a phase shift of $\pm\pi$.
- In reflection, the parallel component $r_{\parallel}$ is in phase at $0 < \theta_i < \theta_p$ and experiences a phase shift of $\pm\pi$ in $\theta_p < \theta_i < \pi/2$. For the angle $\theta_i = \theta_p$, the parallel component vanishes ($r_{\parallel} = 0$). This means that if an unpolarized electromagnetic plane wave hits a reflecting surface with an angle of incidence equal to θ_p, a linearly polarized wave is obtained from the reflection in the TE polarization state.

To explicitly find the polarization angle θ_p, the numerator of Eq. (2.97) is set equal to zero, with $\theta_i = \theta_p$. Thus,

$$n_t \cos\theta_p = n_i \cos\theta_t. \tag{2.100}$$

Changing the cosine functions to sine functions and using Snell's law,

$$n_t\sqrt{1-\sin^2\theta_p} = n_i\sqrt{1-n_i^2\sin^2\theta_p/n_t^2}. \tag{2.101}$$

Squaring and multiplying by n_t^2 leads to

$$n_t^4 - n_t^4\sin^2\theta_p = n_i^2 n_t^2 - n_i^4\sin^2\theta_p, \tag{2.102}$$

and then rearranging and factoring leads to

$$\sin^2\theta_p(n_i^4 - n_t^4) = n_t^2(n_i^2 - n_t^2), \tag{2.103}$$

from where

$$\sin\theta_p = \frac{n_t}{\sqrt{n_t^2 + n_i^2}}, \tag{2.104}$$

which is equivalent to

$$\theta_p = \arctan\left(\frac{n_t}{n_i}\right). \tag{2.105}$$

This angle is also known as Brewster's angle.*

Figure 2.14 illustrates how to obtain linearly polarized light from natural light when it is reflected from a smooth interface. When the angle of incidence of natural light is equal to the Brewster angle given by Eq. (2.105), the component of the incident field parallel to the plane of incidence vanishes on reflection. Consequently, only the component orthogonal to the plane of incidence is reflected (with a phase shift of π); i.e., linearly polarized light is obtained in the TE polarization state. For the example shown in Fig. 2.13, $r_\perp = -0.3846$. The transmitted light remains natural. The two components of the incident field are transmitted with a similar amplitude (with $t_\parallel = 0.6667$ and $t_\perp = 0.6145$ for the example shown in Fig. 2.13).

Of course, the Brewster angle depends on the wavelength. For example, suppose the interface is the face of a BK7 glass plate (Appendix C). The refractive indices of the wavelengths used to characterize optical glasses are $n_F(\lambda = 486.13\,\text{nm}) = 1.5223$, $n_d(\lambda = 587.56\,\text{nm}) = 1.5168$, and $n_C(\lambda = 656.27\,\text{nm}) = 1.5143$. With these values, the angles of incidence to have linear polarization in each case are $\theta_p^F = 56.6991°$, $\theta_p^d = 56.6038°$, and $\theta_p^C = 56.5604°$. These values are close to each other, so in practice, with white light, if we are close to the polarization angle corresponding to the central

*David Brewster (1781–1868) was a Scottish scientist who investigated the polarization of light. In 1815, he formulated this law that is used to find the angle of polarization.

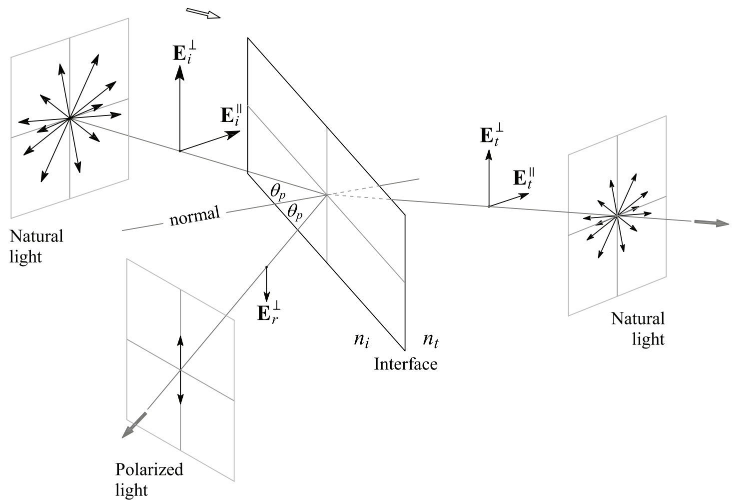


Figure 2.14 Polarization by reflection. The component of the incident field that is parallel to the plane of incidence vanishes on reflection when the angle of incidence is equal to the Brewster angle θ_p.

wavelength of the visible spectrum, the effect of polarization by reflection can be observed very well. Note that even the effect is well observed if the angle of incidence is a few degrees away from Brewster's angle, since the value of $r_{\parallel}$ remains close to zero, as shown in Fig. 2.1.

2.3.3 Reflectance and transmittance

The reflection and transmission coefficients measure the change in the amplitude of the incident wave. To measure the amount of reflected and transmitted energy, the flux of energy [Eq. (1.64)] is used. On a flat surface, the flux energy can be calculated as

$$dF = I(Area). \tag{2.106}$$

Suppose that the region where there is an incident light beam (plane wave) has a circular shape with area σ. The incident, reflected, and transmitted beams will be contained in tubes of cross section σ cos θ_i, σ cos θ_r, and σ cos θ_t, as illustrated in Fig. 2.15.

Reflectance is the ratio of reflected flux to incident flux, and *transmittance* is the ratio of transmitted flux to incident flux. For reflectance,

$$R = \frac{dF_r}{dF_i} = \frac{I_r \sigma \cos\theta_r}{I_i \sigma \cos\theta_i} = \frac{(\epsilon_r \upsilon_r/2)|E_r|^2 \sigma \cos\theta_r}{(\epsilon_i \upsilon_i/2)|E_i|^2 \sigma \cos\theta_i}. \tag{2.107}$$

Because the reflection occurs in the same medium as the incident beam, and taking into account the law of reflection,

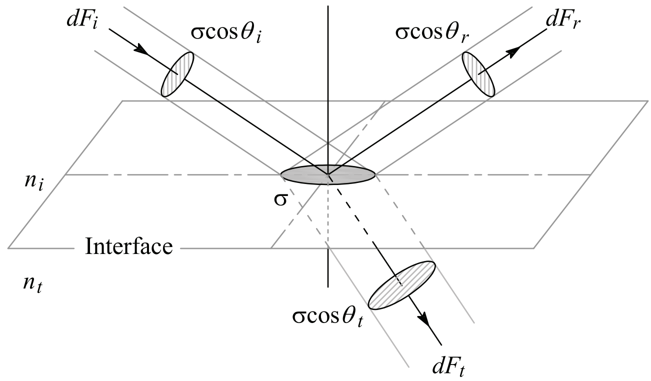


Figure 2.15 Reflected and transmitted intensities at an interface.

$$R = r^2, \tag{2.108}$$

where $r = E_r/E_i$ is the reflection coefficient.

For transmission,

$$T = \frac{dF_t}{dF_i} = \frac{I_t\sigma\cos\theta_t}{I_i\sigma\cos\theta_i} = \frac{(\epsilon_t\upsilon_t/2)|E_t|^2\sigma\cos\theta_t}{(\epsilon_i\upsilon_i/2)|E_i|^2\sigma\cos\theta_i}, \tag{2.109}$$

and assuming that in dielectric materials $\mu_i = \mu_0$, $\mu_t = \mu_0$, multiplying and dividing Eq. (2.109) by $\mu_0 c$ leads to

$$T = \frac{n_t\cos\theta_t}{n_i\cos\theta_i}t^2, \tag{2.110}$$

where $t = E_t/E_i$ is the transmission coefficient.

Conservation of energy means that

$$R + T = 1 \tag{2.111}$$

if the dielectric media do not absorb energy. In the case of energy absorption, $R + T < 1$ and an absorption term must be added to make the energy balance.

For the TE and TM polarization states, Eqs. (2.107) and (2.110) both apply for each component, orthogonal and parallel. Therefore, the energy balance (assuming no absorption) in each case would be

$$R_\perp + T_\perp = 1 \tag{2.112}$$

for the TE polarization state and

$$R_\parallel + T_\parallel = 1 \tag{2.113}$$

for the TM polarization state.

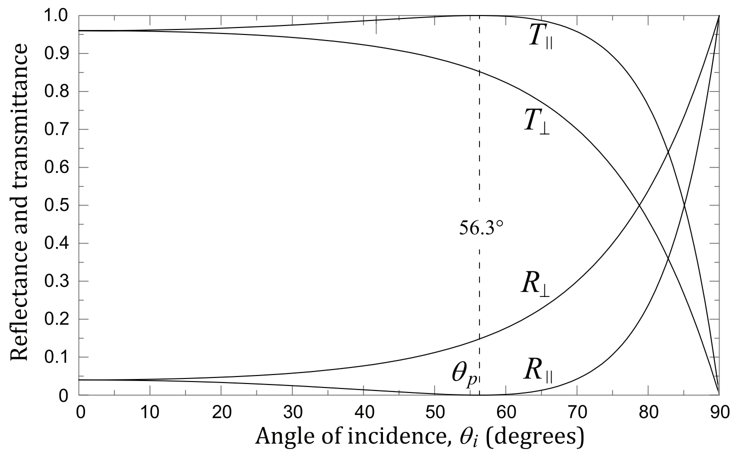


Figure 2.16 Parallel and orthogonal components of reflectance and transmittance.

In Fig. 2.16, $R_{\perp}$, $R_{\parallel}$, $T_{\perp}$, and $T_{\parallel}$ are shown as functions of the angle of incidence when the refractive indices are $n_i = 1.0$ (air) and $n_t = 1.5$ (glass). For any angle, it can be verified that the sum of the reflectance and the transmittance is equal to 1. For example, if the incidence is given at the Brewster angle, then $\theta_i = 56.3°$ and $\theta_t = 33.7°$; $r_{\parallel} = 0$, $r_{\perp} = -0.3846$, $t_{\parallel} = 0.6667$, and $t_{\perp} = 0.6145$. Note that the sum $r_{\parallel} + t_{\parallel} \neq 1$; however, $R_{\parallel} + T_{\parallel} = T_{\parallel} = [1.5\cos(33.7°)/\cos(56.3°)](0.6667^2) = 1$. As another example, when the incidence occurs with an angle $\theta_i = 0$ (incidence normal to the interface), the coefficients are $r_{\parallel} = 0.2$, $r_{\perp} = -0.2$, $t_{\parallel} = 0.8$, and $t_{\perp} = 0.8$. Therefore, the reflectance and transmittance are $R_{\parallel} = R_{\perp} = 0.04$ and $T_{\parallel} = T_{\perp} = 1.5(0.8)^2 = 0.96$, and again Eqs. (2.112) and (2.113) hold. It is common to give the reflectance and transmittance as percentages, multiplying by 100%; e.g., for $\theta_i = 0$, the reflectance is 4% and the transmittance is 96%.

For a glass optical surface ($n_i = 1.0$ and $n_t = 1.5$) the reflectance is small. However, in a multi-lens system, the transmitted light can be greatly reduced. For example, in a system with three lenses, the light must pass through six interfaces, so in a first approximation, the total transmittance would be $T^6 = 78,3\%$.* For this reason, antireflective thin-film dielectric coatings are very common on lenses.

Finally, it is worth noting how the reflectance approaches 1 as the angle of incidence approaches 90°. Any smooth surface behaves like a mirror at

*Note that in this example, there are three air–glass interfaces and three glass–air interfaces. With n_a denoting the refractive index of air and n_v denoting the refractive index of glass, for the first type of interface the transmittance is $T = (n_v/n_a)[2n_a/(n_a + n_v)]^2$ and for the second type of interface the transmittance is $T' = (n_a/n_v)[2n_v/(n_v + n_a)]^2$. Because $T = T' = 4n_a n_v/(n_a + n_v)^2$, the total transmittance $T^3 T'^3$ is equal to T^6. This result is approximate, since successive reflections and transmissions on the faces are being omitted. However, these are minor contributions, and the final result is very close to the approximate one.

grazing incidence (at the interface). This even happens with flat opaque surfaces; e.g., if we look at a sheet of paper from a grazing angle, we can see the specular reflection of light very well.

2.4 Polarization by Total Internal Reflection

In the previous section, we show how the orthogonal and parallel components of the reflected field can undergo a phase change with respect to the components of the incident field. The parallel component is reflected in phase, when $0 < \theta_i < \theta_p$, and with a phase shift of $\pm\pi$, when $\theta_p < \theta_i < \pi/2$. On the other hand, the orthogonal component is always reflected with a phase shift of $\pm\pi$. This, together with the change in amplitude of the reflected components, implies that a linearly polarized wave when reflected remains linearly polarized but with a rotation in the plane of vibration. When total internal reflection occurs, the components of the reflected field have phase shifts that vary between 0 and $\pm\pi$, and the phase difference between the components is no longer limited to 0 or $\pm\pi$. Therefore, the reflected wave can have an elliptical polarization state.

2.4.1 Total internal reflection

If the incident wave goes from a medium with a higher refractive index to one with a lower refractive index, from a certain angle, called the *critical angle*, the reflection and transmission coefficients obtain the values $|r_{\parallel}| = |r_{\perp}| = 1$ and $t_{\parallel} = t_{\perp} = 0$. In other words, the energy of the reflected wave is equal to the energy of the incident wave. In Section 1.1, the condition in which the transmission ray is tangential to the interface is illustrated in Fig. 1.11. The angle of incidence for which this occurs is [Eq. (1.5)]

$$\theta_c = \arcsin\left(\frac{n_t}{n_i}\right), \tag{2.114}$$

with $n_i > n_t$. From this angle the phenomenon of *total internal reflection* occurs.

The Fresnel equations for $n_i > n_t$ and $\theta_i < \theta_c$ apply in the same way as in the external reflection case ($n_i > n_t$), and the only phase changes of the reflected components with respect to the incident components are 0 or $\pm\pi$, as shown in Fig. 2.17 for $n_i = 1.5$ and $n_t = 1.0$. Unlike external reflection (Fig. 2.13), the orthogonal component does not undergo a phase change. In contrast, the parallel component has a phase shift of $\pm\pi$ for $0 < \theta_i < \theta'_p$ and is in phase for $\theta'_p < \theta_i < \theta_c$. The angle θ'_p is the angle at which the polarization of the reflection occurs and is given by $\tan\theta'_p = (n_t/n_i)$. This angle together with the external polarization angle satisfies the relation

$$\theta_p + \theta'_p = \pi/2. \tag{2.115}$$

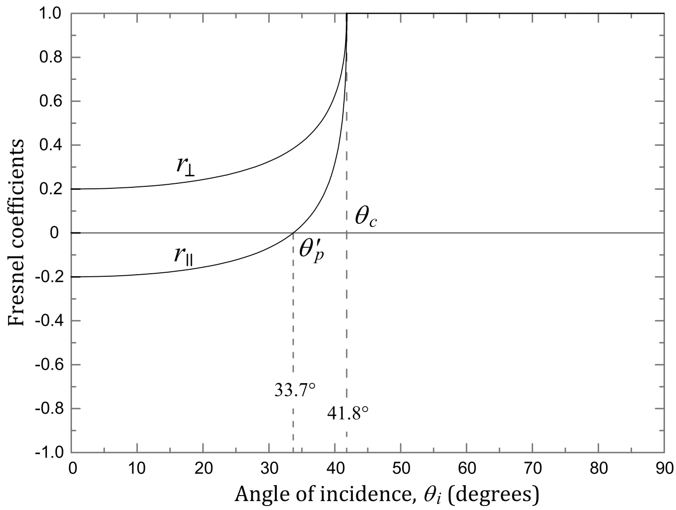


Figure 2.17 Reflection coefficients for the parallel and orthogonal components in internal reflection ($0 < \theta_i < \theta_c$) and total internal reflection ($\theta_c < \theta_i < \pi/2$).

When $n_i > n_t$ and $\theta_i > \theta_c$, the reflection coefficients are complex variable quantities. To see this, note that from Snell's law $\cos\theta_t = \sqrt{1-(n_i/n_t)^2\sin^2\theta_i} = \sqrt{1-(\sin\theta_i/\sin\theta_c)^2}$, and for $\theta_i > \theta_c$, the term inside the square root is negative. Then it is convenient to write

$$\cos\theta_t = \frac{i}{n}\sqrt{\sin^2\theta_i - n^2}, \tag{2.116}$$

with $i = \sqrt{-1}$ and $n = n_t/n_i < 1$. Thus, the reflection coefficients can be rewritten as

$$r_\parallel = \frac{n^2\cos\theta_i - i\sqrt{\sin^2\theta_i - n^2}}{n^2\cos\theta_i + i\sqrt{\sin^2\theta_i - n^2}} \tag{2.117}$$

and

$$r_\perp = \frac{\cos\theta_i - i\sqrt{\sin^2\theta_i - n^2}}{\cos\theta_i + i\sqrt{\sin^2\theta_i - n^2}}. \tag{2.118}$$

In the two coefficients, the numerator is the conjugate of the denominator; therefore,

$$|r_\parallel| = |r_\perp| = 1. \tag{2.119}$$

This is shown in Fig. 2.17 for $\theta_c < \theta_i < \pi/2$. Then the reflection coefficients can be written as

$$r_{\parallel} = |r_{\parallel}|e^{i\delta_{\parallel}} = e^{i\delta_{\parallel}} = e^{-i2\phi_{\parallel}}, \tag{2.120}$$

$$r_{\perp} = |r_{\perp}|e^{i\delta_{\perp}} = e^{i\delta_{\perp}} = e^{-i2\phi_{\perp}}, \tag{2.121}$$

where

$$\tan\phi_{\parallel} = \frac{\sqrt{\sin^2\theta_i - n^2}}{n^2\cos\theta_i}, \tag{2.122}$$

$$\tan\phi_{\perp} = \frac{\sqrt{\sin^2\theta_i - n^2}}{\cos\theta_i}. \tag{2.123}$$

The behavior of the phase shifts ($\delta_{\parallel} = -2\phi_{\parallel}$ and $\delta_{\perp} = -2\phi_{\perp}$) of the parallel and orthogonal reflection coefficients in the internal reflection $0 < \theta_i < \theta_c$ and the total internal reflection $\theta_c < \theta_i < \pi/2$, when $n_i = 1.5$ and $n_t = 1.0$, is shown in Fig. 2.18.

The polarization state of the reflected wave will be determined by $\Delta\delta = \delta_{\parallel} - \delta_{\perp}$. For the range $0 < \theta_i < \theta_c$, a linearly polarized incident wave is reflected linearly polarized (except for a change in the orientation of the plane of vibration). For the range $\theta_c < \theta_i < \pi/2$, the phase difference can be determined from

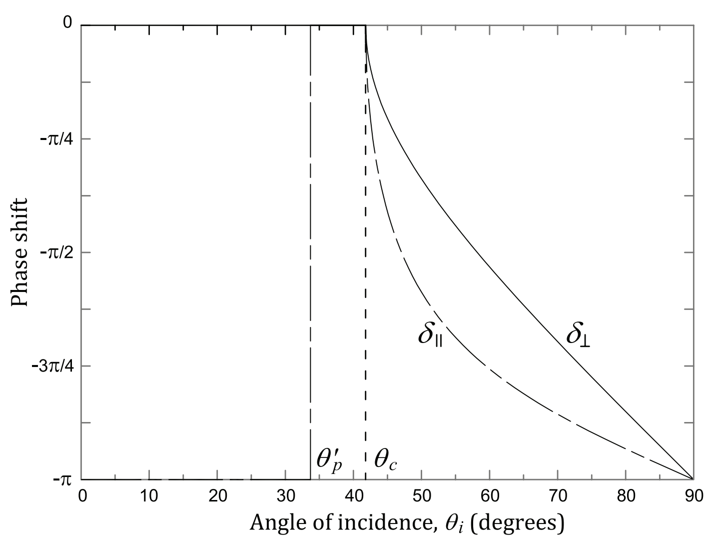


Figure 2.18 Phase shifts of the parallel and orthogonal components in internal reflection ($0 < \theta_i < \theta_c$) and total internal reflection ($\theta_c < \theta_i < \pi/2$).

$$\tan\left(\frac{\delta_{\parallel}}{2}-\frac{\delta_{\perp}}{2}\right)=\tan(\phi_{\perp}-\phi_{\parallel}), \tag{2.124}$$

i.e.,

$$\tan\left(\frac{\delta_{\parallel}}{2}-\frac{\delta_{\perp}}{2}\right)=\frac{\sqrt{\sin^2\theta_i-n^2}/\cos\theta_i-\sqrt{\sin^2\theta_i-n^2}/(n^2\cos\theta_i)}{1+(\sin^2\theta_i-n^2)/(n^2\cos^2\theta_i)}. \tag{2.125}$$

Simplifying, this can be rewritten as,

$$\tan\left(\frac{\delta_{\parallel}}{2}-\frac{\delta_{\perp}}{2}\right)=-\frac{\cos\theta_i\sqrt{\sin^2\theta_i-n^2}}{\sin^2\theta_i}. \tag{2.126}$$

Therefore, the phase differences will be

$$\Delta\delta=\delta_{\parallel}-\delta_{\perp}=-2\arctan\left(\frac{\cos\theta_i\sqrt{\sin^2\theta_i-n^2}}{\sin^2\theta_i}\right). \tag{2.127}$$

The difference of phase shifts by internal reflection $0 < \theta_i < \theta_c$ and total internal reflection $\theta_c < \theta_i < \pi/2$, when $n_i = 1.5$ and $n_t = 1.0$, are shown in Fig. 2.19. In this figure it can be seen that for total internal reflection the phase difference varies between 0 and a value close to $-\pi/4$. Consequently, a wave reflected in the domain of total internal reflection will have an elliptical polarization state.

To determine the minimum value of the phase difference in total internal reflection, $d(\Delta\delta)/d\theta_i = 0$ can be solved for θ_i. The result obtained for the angle of incidence is

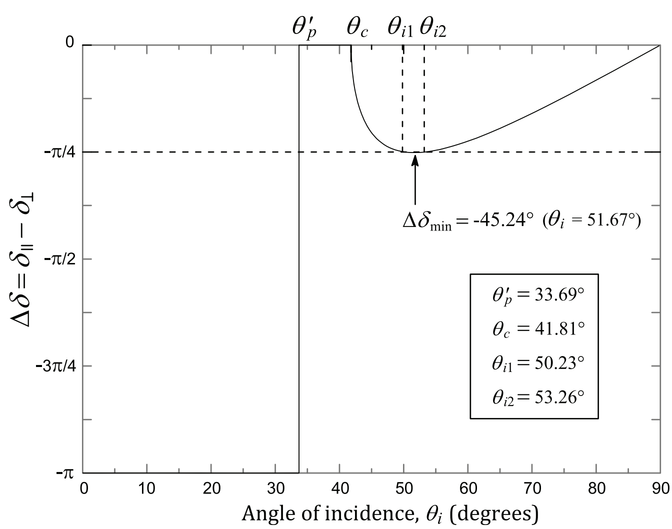


Figure 2.19 Difference of the phase shifts of the parallel and orthogonal reflected components, when $n_i = 1.5$ and $n_t = 1.0$.

$$\sin^2\theta_i = \frac{2n^2}{1+n^2}, \tag{2.128}$$

and the minimum value of the phase difference turns out to be

$$\Delta\delta_{\min} = 2\arctan\left(\frac{n^2-1}{2n}\right). \tag{2.129}$$

As shown in Fig. 2.19, the minimum of the phase difference is $\Delta\delta_{\min} = -45.24°$, corresponding to the angle of incidence $\theta_i = 51.67°$. This tells us that with a reflection, in the domain of total internal reflection, when $n_i = 1.5$ and $n_t = 1.0$, it is not possible to obtain a wave with a state of circular polarization. One option to achieve $\Delta\delta_{\min} = -\pi/2$ is to have a material whose refractive index n_i is such that $n^2 - 1 = -2n$. The positive solution of this equation is $n = -1 + \sqrt{2} = 0.4142$. Assuming that $n_t = 1$, then $n_i = 2.4142$. This is a very high index for an optical glass. Another option is to keep a common optical glass and achieve two successive reflections, as long as the sum of the two phase differences is $-\pi/2$. The most common option is the one in which each reflection has a phase shift of $-\pi/4$. The values of the angles of incidence, for which the phase differences of $-\pi/4$ are obtained, are illustrated in Fig. 2.19 with segmented lines. These are $\theta_{i1} = 50.23°$ and $\theta_{i2} = 53.26°$.

An optical element that changes from linear to circular polarization, based on these ideas, is the Fresnel rhomb, such as the one shown in Fig. 2.20, with glass of refractive index 1.5 for the angle $\theta_{i2} = 53.26°$. Suppose that an incident linearly polarized light beam (traveling from left to right) with the plane of vibration at $-45°$ to the plane of incidence (45° to the positive direction of the orthogonal component) is perpendicular to the first side of the rhomb. Omitting the spatial and temporal phase terms, the electric field shown in Fig. 2.20 at position (1) can be written as

$$\mathbf{E}(1) = \{e^{i0}, e^{i0}\}E_0.$$

Position (2) would be

$$\mathbf{E}(2) = \{t_\perp^{(2)} e^{i0}, t_\parallel^{(2)} e^{i0}\}E_0 = \{e^{i0}, e^{i0}\}0.8E_0;$$

i.e., it is still linearly polarized, but the amplitude of the components is 0.8 times the initial one.

Position (3) would be

$$\mathbf{E}(3) = \left\{r_\perp^{(3)} t_\perp^{(2)} e^{i0}, r_\parallel^{(3)} t_\parallel^{(2)} e^{i0}\right\}E_0 = \{e^{-i(1.2785)}, e^{-i(2.0639)}\}0.8E_0;$$

i.e., $\{e^{i0}, e^{-i\pi/4}\}0.8e^{-i(1.2785)}E_0$, which represents a state of right elliptical polarization.

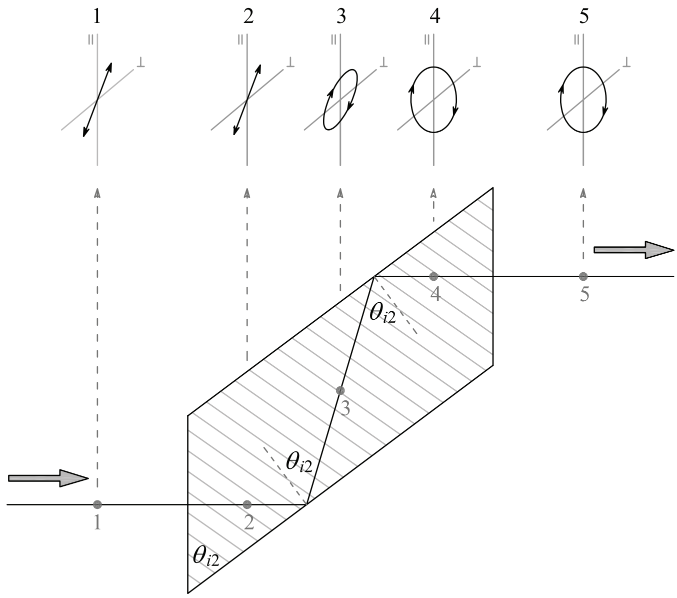


Figure 2.20 Fresnel rhomb in air ($n_t = 1.0$) for glass of refractive index $n_i = 1.5$. The angle of incidence on the diagonal faces is $\theta_i = 53.26°$. Incident light linearly polarized at $-45°$ (with respect to the plane of incidence) emerges as right circularly polarized light.

Position (4) would be

$$\mathbf{E}(4) = \left\{ r_{\perp}^{(4)} r_{\perp}^{(3)} t_{\perp}^{(2)} e^{i0},\ r_{\parallel}^{(4)} r_{\parallel}^{(3)} t_{\parallel}^{(2)} e^{i0} \right\} E_0 = \{e^{-i(2.5570)},\ e^{-i(4.1277)}\} 0.8 \times E_0;$$

i.e., $\{e^{i0}, e^{-i\pi/2}\}0.8e^{-i(2.5570)}E_0$, which represents a state of right circular polarization. The radius of the circle is $0.8E_0$.

Position (5) will be

$$\begin{aligned} \mathbf{E}(5) &= \left\{ t_{\perp}^{(5)} r_{\perp}^{(4)} r_{\perp}^{(3)} t_{\perp}^{(2)} e^{i0},\ t_{\parallel}^{(5)} r_{\parallel}^{(4)} r_{\parallel}^{(3)} t_{\parallel}^{(2)} e^{i0} \right\} E_0 \\ &= \{e^{-i(2.5570)},\ e^{-i(4.1277)}\} 0.96 E_0; \end{aligned}$$

i.e., $\{e^{i0}, e^{-i\pi/2}\}0.96e^{-i(2.5570)}E_0$, which represents a state of right circular polarization. The radius of the circle is $0.96E_0$.

If the incident beam is linearly polarized but with a plane of vibration other than +45°, the result will be a polarization state other than circular. Specifically, if the angle of the vibration plane is 0 or 90°, the state of linear polarization does not change; it remains in a plane perpendicular or parallel to the plane of incidence. Note that the plane of incidence is defined with the normal of the second surface, since with the first surface the plane of incidence is not defined (the incident wave vector is collinear with the normal of the interface).

2.4.2 Reflectance and transmittance

Reflectance and transmittance curves are shown in Fig. 2.21. The reflectance is the square of the curves shown in Fig. 2.17 [Eq. (2.108)]. For transmittance,

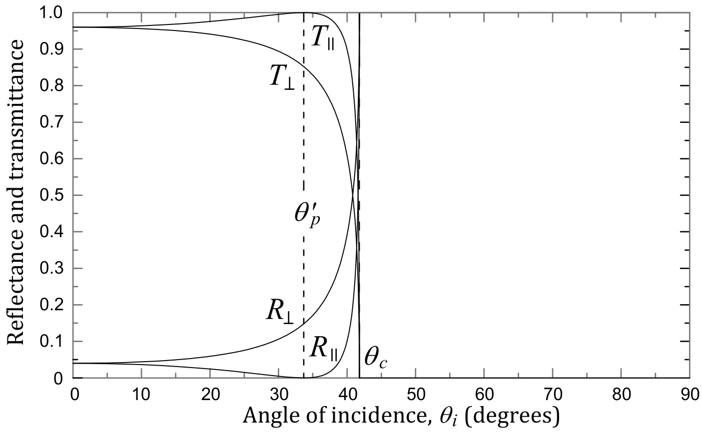


Figure 2.21 Reflectance and transmittance for internal reflection ($0 < \theta_i < \theta_c$) and total internal reflection ($\theta_c < \theta_i < \pi/2$).

we also have two intervals. In the first, $0 < \theta_i < \theta_c$, the transmission coefficients are calculated according to Eqs. (2.98) and (2.99). In the second interval, $\theta_c < \theta_i < \pi/2$, the transmission angle is $\theta_t = 90°$; therefore, according to Eq. (2.110), the transmittance for both the parallel and the orthogonal components becomes equal to 0.* Of course, it is again verified that $R + T = 1$ for the two components of the electric (and magnetic) fields.

2.5 Polarization with Birefringent Materials

The electric polarization vector in dielectric materials is related to the electric field by the electric susceptibility according to Eq. (B.9):

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E}.$$

When the material is isotropic, the susceptibility quantity is described by a scalar and the wave equation [Eq. (B.14)] is reduced to $\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 (1 + \chi) \times \partial^2 \mathbf{E} / \partial t^2$, where $\upsilon = c / \sqrt{1 + \chi}$ is the speed of light in the material. When the material is anisotropic, the susceptibility is described by a tensor (3 × 3 matrix) and the wave equation is a bit more complex.

In general, electric susceptibility can be described as

$$\chi = \begin{pmatrix} \chi_{11} & \chi_{12} & \chi_{13} \\ \chi_{21} & \chi_{22} & \chi_{23} \\ \chi_{31} & \chi_{32} & \chi_{33} \end{pmatrix}, \tag{2.130}$$

*In total internal reflection, the incident energy is completely reflected. However, there is still an electromagnetic wave beyond the interface that is rapidly fading. This wave is known as an evanescent wave.

and the electric polarization vector can be described as

$$\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \epsilon_0 \begin{pmatrix} \chi_{11} & \chi_{12} & \chi_{13} \\ \chi_{21} & \chi_{22} & \chi_{23} \\ \chi_{31} & \chi_{32} & \chi_{33} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}, \tag{2.131}$$

which is not necessarily parallel to the electric field vector.

In the case of nonabsorbing anisotropic dielectric materials (of particular interest in this book), χ is a symmetric matrix and can be reduced on a system of principal axes to [2]

$$\chi = \begin{pmatrix} \chi_1 & 0 & 0 \\ 0 & \chi_2 & 0 \\ 0 & 0 & \chi_3 \end{pmatrix}. \tag{2.132}$$

The quantities χ_1, χ_2, and χ_3 are called *principal susceptibilities*.

In general, crystals are anisotropic materials whose elements (atoms, molecules, ions, etc.) are located in regular geometrical arrangements: cubic, trigonal, tetragonal, hexagonal, triclinic, monoclinic, and orthorhombic. For nonabsorbing dielectric crystals, the principal susceptibilities are related as follows: for cubic, $\chi_1 = \chi_2 = \chi_3$; i.e., it behaves like an isotropic material; for trigonal, tetragonal, and hexagonal, $\chi_1 = \chi_2 \neq \chi_3$; and for triclinic, monoclinic, and orthorhombic, $\chi_1 \neq \chi_2 \neq \chi_3$.

By writing the wave equation [Eq. (B.14)] as

$$\nabla \times (\nabla \times \mathbf{E}) + \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\frac{1}{c^2}\chi\frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{2.133}$$

and proposing a plane wave solution of the form $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, a set of algebraic equations are obtained from which the behavior of the components of the electric field inside the anisotropic material can be analyzed. According to Eqs. (2.10) and (2.12), for harmonic plane waves, the following changes can be made: $\nabla \rightarrow i\mathbf{k}$ and $\partial/\partial t \rightarrow -i\omega$. With this in mind, Eq. (2.133) becomes

$$\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) + \frac{\omega^2}{c^2}\mathbf{E} = -\frac{\omega^2}{c^2}\chi\mathbf{E}. \tag{2.134}$$

In terms of components of $\mathbf{k}$ and $\mathbf{E}$,* it leads to

$$\left(-k_y^2 - k_z^2 + \frac{\omega^2}{c^2}\right)E_x + k_x k_y E_y + k_x k_z E_z = -\frac{\omega^2}{c^2}\chi_1 E_x, \tag{2.135}$$

*To obtain each of the components, the identity vector $[\mathbf{x} \times (\mathbf{y} \times \mathbf{z})]_i = y_i x_j z_j - z_i x_j y_j$ can be used, where $x_j y_j$ ($x_j z_j$) denotes the scalar product $x \cdot y$ ($x \cdot z$). With this equation, we obtain the ith component of the double cross product.

$$k_xk_yE_x + \left(-k_x^2 - k_z^2 + \frac{\omega^2}{c^2}\right)E_y + k_yk_zE_z = -\frac{\omega^2}{c^2}\chi_2E_y, \tag{2.136}$$

$$k_xk_zE_x + k_yk_zE_y + \left(-k_x^2 - k_y^2 + \frac{\omega^2}{c^2}\right)E_z = -\frac{\omega^2}{c^2}\chi_3E_z. \tag{2.137}$$

2.5.1 Phase retarder plates

A first result of Eqs. (2.135–2.137) is obtained directly if we assume that a wave propagates within the material in one of the directions, e.g., $\mathbf{k} = (k_x, 0, 0)$, where the magnitude of $\mathbf{k}$ is $k = k_x$. In this case, the wave equations reduce to

$$\frac{\omega^2}{c^2}E_x = -\frac{\omega^2}{c^2}\chi_1E_x, \tag{2.138}$$

$$\left(-k_x^2 + \frac{\omega^2}{c^2}\right)E_y = -\frac{\omega^2}{c^2}\chi_2E_y, \tag{2.139}$$

$$\left(-k_x^2 + \frac{\omega^2}{c^2}\right)E_z = -\frac{\omega^2}{c^2}\chi_3E_z. \tag{2.140}$$

Because $\epsilon \neq \epsilon_0$ in the dielectric material, then χ_1 (and χ_2, χ_3) $\neq 0$. Therefore, in Eq. (2.138), $E_x = 0$; i.e., $\mathbf{E}$ is transverse to the direction of propagation, or $\mathbf{E} = \{0, E_y, E_z\}$.

In Eq. (2.139), if $E_y \neq 0$, then

$$k^2 = \frac{\omega^2}{c^2}(1 + \chi_2). \tag{2.141}$$

Because the index of refraction is defined as $n = \sqrt{1 + \chi}$ [Eq. (B.18)], the wavenumber of the component of the electric field that vibrates in the yx plane and propagates in the x direction is given by

$$k = \frac{2\pi}{\lambda}n_2. \tag{2.142}$$

In other words, the E_y component propagates in a material of refractive index $n_2 = \sqrt{1 + \chi_2}$.

Similarly, in Eq. (2.140), if $E_z \neq 0$, then

$$k^2 = \frac{\omega^2}{c^2}(1 + \chi_3); \tag{2.143}$$

therefore, the wavenumber of the component of the electric field that vibrates in the zx plane and propagates in the x direction is given by

$$k = \frac{2\pi}{\lambda} n_3. \qquad (2.144)$$

Now, the E_z component of the field propagates in a material of refractive index $n_3 = \sqrt{1+\chi_3}$. So for the wave propagating in the x direction, the material has two refractive indices, one for each component of the electric field. These types of materials are called *birefringent materials* (double index).

In short, the electromagnetic wave propagating in the x direction has the form

$$\mathbf{E} = \{0,\ E_{0y}e^{i(2\pi x n_2/\lambda - \omega t)},\ E_{0z}e^{i(2\pi x n_3/\lambda - \omega t)}\}. \qquad (2.145)$$

Thus, E_y and E_z components are continuously out of phase. For the distance x, the phase shift between the components would be

$$\Delta\delta = \frac{2\pi(n_3 - n_2)}{\lambda} x. \qquad (2.146)$$

Similar results are obtained whether the propagation direction is y or z. Thus, the principal refractive indices n_1, n_2, and n_3 are associated with the x, y, and z, respectively.

Based on these results, birefringent plates are fabricated to generate phase delays between components by controlling the thickness of the plate. Suppose we have a birefringent plate of thickness d, as shown in Fig. 2.22, whose edges coincide with the principal directions and $n_3 > n_2$. A wave that vibrates in the zx plane propagates with speed $\upsilon_3 = c/n_3$, and a wave that vibrates in the yx plane propagates with speed $\upsilon_2 = c/n_2$. Because $n_3 > n_2$, then $\upsilon_3 < \upsilon_2$. In other words, for Eq. (2.145), the $\mathbf{E}_z$ component travels slower than the $\mathbf{E}_y$ component. So it is said that there is a *slow axis* in z and a *fast axis* in y. Thus,

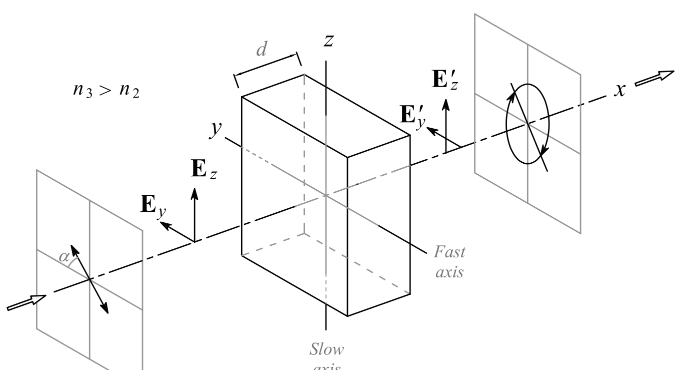


Figure 2.22 Phase retarder plate.

it is clear that the $\mathbf{E}_y$ component is ahead of $\mathbf{E}_z$ when they emerge from the plate, and the phase difference between the two would be $2\pi(n_3 - n_2)d/\lambda$.

Quarter-wave plate

Suppose that we want to generate a phase difference equal to $\pi/2$, i.e., $2\pi(n_3 - n_2)d/\lambda = \pi/2$. The thickness of the plate should be

$$d = \frac{\lambda}{4}\left(\frac{1}{\Delta n}\right),$$

with $\Delta n = |n_3 - n_2|$. These types of plates are called λ/4 *plates* and are often used to generate circularly polarized light. For example, if a linearly polarized plane wave with the plane of vibration at an angle $\alpha = 45°$ is incident on the front face of the plate with $\theta_i = 0$ (normal incidence), the light exiting the plate will have a left circular polarization state, as shown in Fig. 2.22 Note that if the plate is rotated 90° around the x axis (now the fast axis will be vertical and the slow axis will be horizontal) maintaining the polarization of the incident beam, the result will be light with a right circular polarization state.

Half-wave plate

To generate a phase difference of π, the thickness of the plate must be doubled, i.e,

$$d = \frac{\lambda}{2}\left(\frac{1}{\Delta n}\right).$$

These types of plates are called λ/2 *plates* and are often used to rotate the plane of vibration of a linearly polarized wave. Specifically, if the incident wave is as in the previous case, with $\alpha = 45°$, the wave leaving the λ/2 plate will be linearly polarized with $\alpha = -45°$.

A notable advantage of using a λ/2 plate to rotate the plane of vibration of a linearly polarized wave is that the rotation is performed without attenuating the amplitude of the wave, which occurs when dichroic polarizers are used, according to Malus' law [Eq. (2.53)].

To estimate the thickness of phase-retarding plates, suppose we want to build a λ/4 plate of calcite (trigonal crystal structure). The principal refractive indices of calcite are $n_1 = n_2 = 1.658$ and $n_3 = 1.486$,* i.e., $\Delta n = 0.172$. Thus, the thickness for the line "d" of helium ($\lambda = 587.56$ nm) would be $d = 0.8540$ μm. This is a very small thickness for a practical device. Thus, at a commercial level, the plates are made with birefringent crystals whose thickness is an odd multiple of λ/4 (or an even multiple of λ/2).

*These are the ordinary and extraordinary indices. These definitions appear in Section 2.5.2.

2.5.2 Birefringent crystals

The previous section shows how a birefringent crystal behaves assuming that the wave propagates along one of the principal directions. A more general case assumes any direction of propagation. In this case, Eqs. (2.135–2.137) must be considered. These equations constitute a system of homogeneous linear equations. The trivial solution is $E_x = E_y = E_z = 0$. The nontrivial solution assumes that the determinant of the coefficients is equal to 0, i.e.,

$$\begin{vmatrix} (n_1\omega/c)^2 - k_y^2 - k_z^2 & k_x k_y & k_x k_z \\ k_x k_y & (n_2\omega/c)^2 - k_x^2 - k_z^2 & k_y k_z \\ k_x k_z & k_y k_z & (n_3\omega/c)^2 - k_x^2 - k_y^2 \end{vmatrix} = 0. \tag{2.147}$$

In a k_x, k_y, and k_z coordinate system, this equation represents a double-layered surface. To see these surfaces in a simple way, let us examine the cuts of the surfaces with the planes $k_x k_y$ ($k_z = 0$), $k_x k_z$ ($k_y = 0$), and $k_y k_z$ ($k_x = 0$), assuming that $n_1 < n_2 < n_3$. Starting with the plane $k_z = 0$, Eq. (2.147) is reduced to

$$\begin{vmatrix} (n_1\omega/c)^2 - k_y^2 & k_x k_y & 0 \\ k_x k_y & (n_2\omega/c)^2 - k_x^2 & 0 \\ 0 & 0 & (n_3\omega/c)^2 - k_x^2 - k_y^2 \end{vmatrix} = 0, \tag{2.148}$$

from where

$$\{(n_3\omega/c)^2 - k_x^2 - k_y^2\}\{[(n_1\omega/c)^2 - k_y^2][(n_2\omega/c)^2 - k_x^2] - k_x^2 k_y^2\} = 0. \tag{2.149}$$

This is the product of two factors (those in curly brackets), and at least one of them must be equal to 0. From the first factor,

$$k_x^2 + k_y^2 = (n_3\omega/c)^2, \tag{2.150}$$

i.e., a circle of radius $n_3\omega/c$. From the second factor,

$$\frac{k_x^2}{(n_2\omega/c)^2} + \frac{k_y^2}{(n_1\omega/c)^2} = 1, \tag{2.151}$$

i.e., an ellipse with semiaxes $n_2\omega/c$ and $n_1\omega/c$ along k_x and k_y, respectively.

The two curves resulting from the intersection of the surface of two shells with the plane $k_z = 0$ are shown in Fig. 2.23: the circle with radius $n_3\omega/c$ and the ellipse with semiaxes $n_1\omega/c$ and $n_2\omega/c$.

Performing an analogous analysis for the planes $k_x = 0$ and $k_z = 0$, the intersections of the two surfaces in each plane are a circle and an ellipse, as shown in Fig. 2.24. So, in general, for any direction **k**, there will be two values

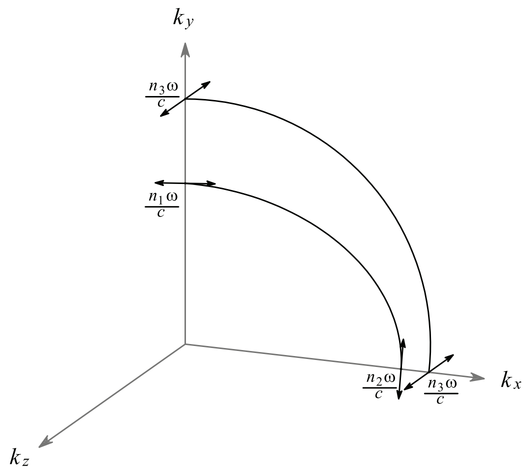


Figure 2.23 Wave vector curves in a birefringent crystal in the plane $k_z = 0$.

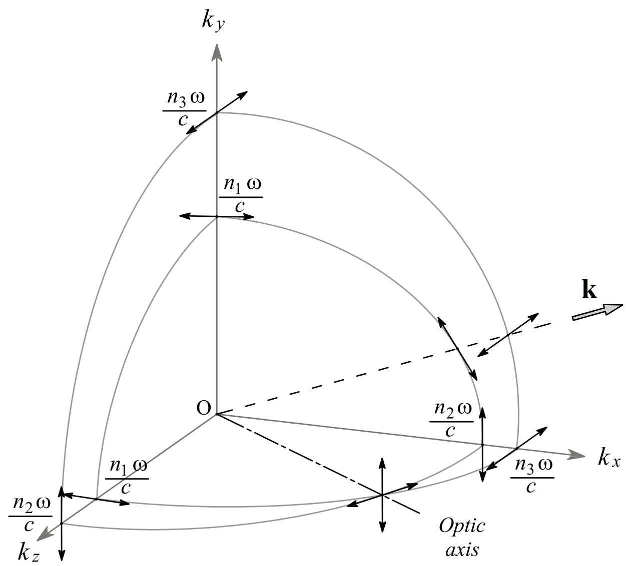


Figure 2.24 Wave vector surfaces in a birefringent crystal with $n_1 < n_2 < n_3$. In the direction of the optic axis, the phase velocities associated with each of the polarization directions are equal.

for the wavenumber k, i.e., two refractive indices. In particular, for propagation in the x direction, there will be two phase velocities, $\upsilon_2 = c/n_2$ for the E_y component and $\upsilon_3 = c/n_3$ for the E_z component; for propagation in the y direction, we will also have two phase velocities, $\upsilon_1 = c/n_1$ for the E_x component and $\upsilon_3 = c/n_3$ for the E_z component; and similarly for propagation in the z direction, i.e., $\upsilon_1 = c/n_1$ for the E_x component and $\upsilon_2 = c/n_2$ for the E_y component. This means that for each direction of propagation there are two phase velocities corresponding to two mutually orthogonal components of the electric field. For any other direction of propagation the same thing happens: there are two phase velocities that always correspond to two polarization directions orthogonal to each other. This is illustrated in Fig. 2.24, with the wave vector **k** in the plane $k_z = 0$. The values of the refractive indices for each direction of polarization are the distances from the origin of the coordinate system to the two intersections of the vector **k**, divided by ω/c.

A particular situation for phase velocities occurs when the two surfaces of the wave vector intersect at a point, as shown in Fig. 2.24 for the $k_y = 0$ plane. There, the wavenumber k is the same for the two polarization directions; therefore, the refractive indices are the same and, consequently, the phase velocities will also be the same. The direction of **k** for which this occurs is called the *optic axis of the crystal*. Thus, a wave that propagates in an anisotropic crystal along the optic axis of the crystal does so in the same way as in an isotropic material. The two orthogonal components of the field are not out of phase with each other. In birefringent crystals in which the three principal refractive indices are different from each other, there are two optic axes; these crystals are called *biaxial crystals*. In birefringent crystals in which two of the three principal refractive indices are equal to each other, there is only one optic axis; these crystals are called *uniaxial crystals*. The two surfaces of the wave vector in uniaxial crystals consist of a sphere and an ellipsoid of revolution. Depending on which surface contains the other, uniaxial crystals are classified as positive or negative. They are positive if the ellipsoid circumscribes the sphere; they are negative if the sphere circumscribes the ellipsoid. In the case of biaxial crystals, these surfaces are intersecting ellipsoids of revolution. The crystal classification of the sphere and ellipsoid sections in the plane containing the optic axes is shown in Fig. 2.25.

In uniaxial crystals $\chi_1 = \chi_2$, the corresponding index of refraction is called the *ordinary refractive index*, n_o. The index corresponding to χ_3 is called the *extraordinary refractive index*, n_e. Thus, in a positive uniaxial crystal, $n_o < n_e$,

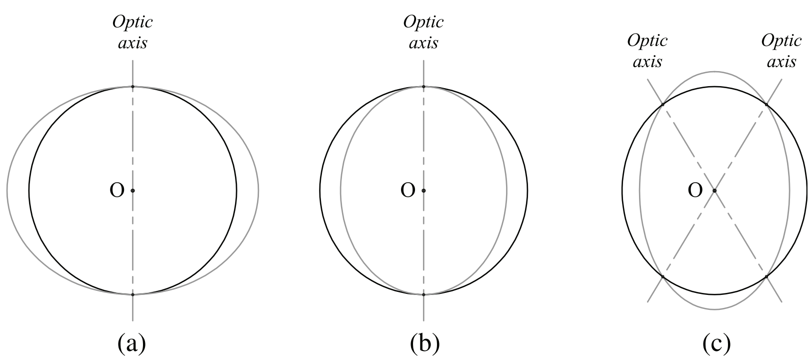


Figure 2.25 Classification of birefringent crystals according to the optic axis of the crystal. (a) Positive uniaxial, (b) negative uniaxial, and (c) biaxial.

Table 2.2 Some crystals. In uniaxial crystals, the indices that are the same are called ordinary, and the index that is different is called extraordinary [2].

Structure	Susceptibility	Crystal	n_1	n_2	n_3
Isotropic *cubic*	$\chi = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}$	*Sodium chloride*	1.544	1.544	1.544
		Diamond	2.417	2.417	2.417
Uniaxial					
trigonal *tetragonal* *hexagonal*	$\chi = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$	+ *Quartz*	1.544	1.544	1.553
		+ *Ice*	1.309	1.309	1.310
		– *Calcite*	1.658	1.658	1.486
		– *Sodium nitrate*	1.587	1.587	1.336
Biaxial					
triclinic *monoclinic* *orthorhombic*	$\chi = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$	*Topaz*	1.619	1.620	1.627
		Mica	1.552	1.582	1.588

and in a negative uniaxial crystal, $n_o > n_e$. On the other hand, the component of the field that vibrates with the wavenumber corresponding to n_o is called an *ordinary wave* (*ordinary ray*), and the component of the field that vibrates with the wavenumber corresponding to n_e is called an *extraordinary wave* (*extraordinary ray*).

Some examples of positive (+) and negative (–) uniaxial and biaxial crystals are given in Table 2.2 [2]. In particular, calcite has a relatively large Δn, which explains why it is a widely used material for the manufacture of optical elements (phase retarders and prisms).

2.5.3 Refraction in crystals

In general, a birefringent crystal has two refractive indices for one direction of light propagation. Therefore, in the refraction of light at an interface separating two media, one isotropic and the other anisotropic, an incident light beam with $\theta_i \neq 0$ (coming from the isotropic medium) will separate into two refracted light beams. In particular, for refraction we will consider interfaces that are parallel or orthogonal to the optic axis of the crystal.

In Fig. 2.24, in the $k_y = 0$ plane, it can be seen that the ordinary wave vibrates orthogonally to the optic axis of the crystal. This is the case in all crystals and allows us to determine the direction that ordinary and extraordinary waves follow in refraction. To see the refraction, let us consider a negative uniaxial crystal immersed in air. Assume that the interface is a flat face of the crystal that contains the optic axis and that the plane of incidence is orthogonal to the optic axis. From Fig. 2.25(b), we can see that the incidence and refraction of rays occur as shown in Fig. 2.26. The refractive index curves (wavenumber divided by ω/c) are circles. Because the optic axis indicates one

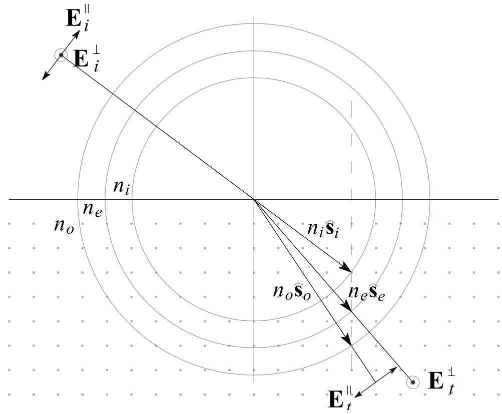


Figure 2.26 Refraction in a negative uniaxial crystal ($n_o > n_e$). In refraction, there are two separate beams: the ordinary wave $E_t^{\parallel}$ and the extraordinary wave $E_t^{\perp}$.

direction, the optic axis is shown as points (lines orthogonal to the plane of the paper) in Fig. 2.26. If we decompose the incident beam into a vector parallel to the plane of incidence (paper plane), $\mathbf{E}_i^{\parallel}$, and into a vector orthogonal to the plane of incidence, $\mathbf{E}_i^{\perp}$, then the parallel component will be orthogonal to the optic axis; this will be the ordinary wave.

To obtain the directions of propagation of ordinary and extraordinary waves in refraction, we can make use of the graphical ray tracing of Fig. 1.9, as shown in Fig. 2.26. Of course, both rays obey the vector form of Snell's law (Eq. [2.75)]. For the ordinary ray, $n_t = n_o$, and for the extraordinary ray, $n_t = n_e$. Therefore, for the ordinary ray,

$$n_o \sin\theta_{to} = n_i \sin\theta_i, \tag{2.152}$$

and for the extraordinary ray,

$$n_e \sin\theta_{te} = n_i \sin\theta_i. \tag{2.153}$$

The final result is the separation of the two components of the incident wave: $E_t^{\parallel}$ vibrates propagating in the direction of the ordinary ray, and $\mathbf{E}_t^{\perp}$ vibrates propagating in the direction of the extraordinary ray. The separation of the beams by double refraction explains the double image that can be observed with calcite crystals, as shown in Fig. 2.27.

Changing the orientation of the optic axis with respect to the interface also changes the refractive index curves and thus the way ordinary and extraordinary rays are refracted. Some basic configurations of the index curves for positive and negative uniaxial crystals are shown in Fig. 2.28. The optic axis is represented by parallel lines in gray or by dots (lines orthogonal to the plane of the paper). In addition to the separation of the ordinary (o) and extraordinary (e) rays, in negative uniaxial crystals the ordinary ray lags behind the extraordinary ray; in positive uniaxial crystals the opposite occurs.

Figure 2.27 Double image generated with a calcite crystal.

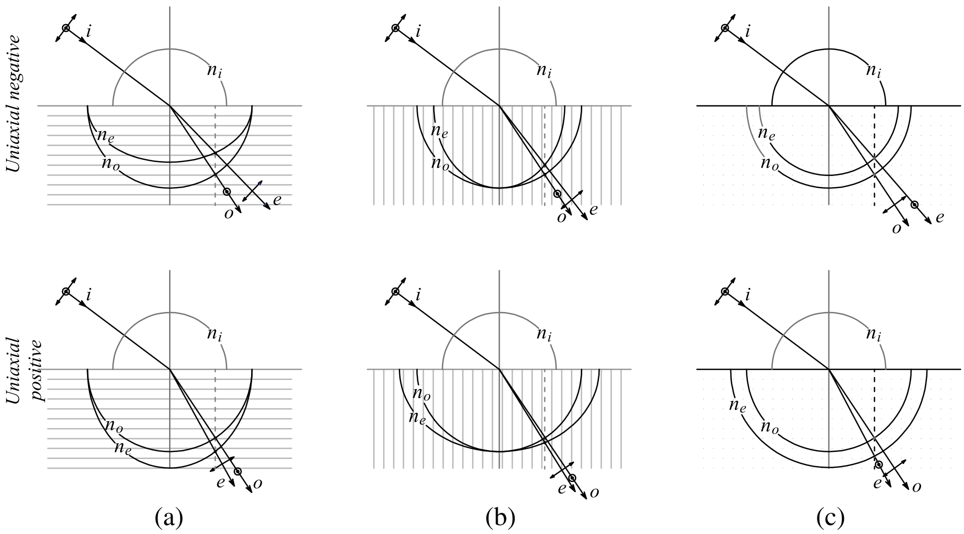


Figure 2.28 Some configurations of the interface and the plane of incidence in uniaxial crystals whose optic axis is parallel or orthogonal to the interface.

2.5.4 Polarizing prisms

From the configurations shown in Fig. 2.28, optical elements can be built to obtain linearly polarized light from natural light. Because the two components of the refracted field have mutually orthogonal linear polarization states, one of them can be blocked to obtain linearly polarized light. However, the obtained beam will have a different direction of propagation than the incident beam.

To obtain linearly polarized light in the same direction as the incident natural light beam, a pair of birefringent prisms can be used with their diagonal faces joined by an optical medium or by a film of air. In the first prism, the rays with mutually orthogonal polarizations maintain the direction

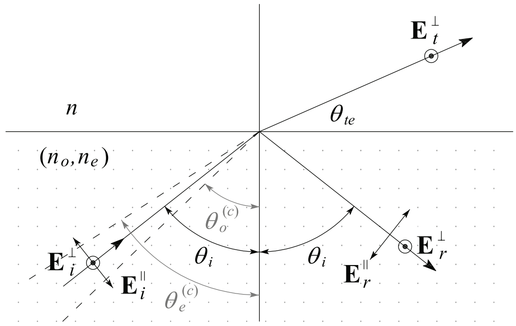


Figure 2.29 Total internal reflection in calcite for ordinary rays, $n < n_o$ and $n < n_e$.

of the incident ray (orthogonal to the first face). In the diagonal face, one of the rays is completely reflected while the other is reflected and transmitted to the second prism to maintain the direction of the incident beam. What is remarkable here is the total internal reflection of one of the beams.

To see the working principle, let us consider the internal reflection of a beam in calcite, as shown in Fig. 2.29. Because there are two refractive indices in calcite, $n_o = 1.658$ for the wave that vibrates parallel to the plane of incidence (ordinary ray) and $n_e = 1.486$ for the wave that vibrates orthogonal to the plane of incidence (extraordinary ray), there will also be two critical angles for total internal reflection. If $n = 1$ (air), for ordinary rays,

$$\theta_o^{(c)} = \arcsin\left(\frac{n}{n_o}\right) = 37.09^\circ,$$

and for the extraordinary,

$$\theta_e^{(c)} = \arcsin\left(\frac{n}{n_e}\right) = 42.29^\circ.$$

Based on these two angles, if the incident ray arrives at the interface with an angle $\theta_o^{(c)} < \theta_i < \theta_e^{(c)}$, then the ordinary ray is reflected with total internal reflection, while the extraordinary ray is reflected and transmits, so that in the second medium (air) there is a linearly polarized beam that vibrates orthogonally to the plane of incidence.

Figure 2.30 shows a possible system with a pair of calcite prisms in which the output beam is linearly polarized in the "s" state and in the same direction as the incident beam (natural light). The normal of the diagonal face (of the first prism) and the incident beam determine the plane of incidence. On the other hand, the inclination of the diagonal is such that the rays transmitted in the first prism arrive at an angle $\theta_i = 40°$ with the diagonal. A total internal reflection of the ordinary ray and an internal reflection of the extraordinary ray occur in the reflection of this diagonal. In transmission there is only the

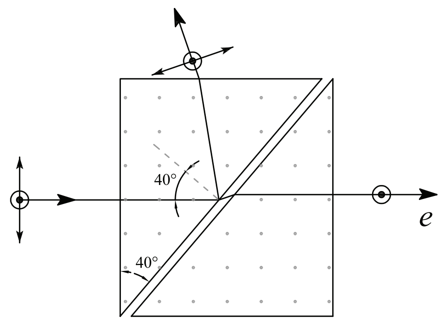


Figure 2.30 System of two calcite prisms with the optic axis orthogonal to the plane of incidence to generate linearly polarized light in the "s" state. The two diagonals are parallel, and the medium between them is air ($n = 1.000$).

extraordinary ray that deviates (according to Snell's law), and when it reaches the second diagonal it returns to take the direction it had in the first prism because the two diagonals are faced parallel to each other.

Commercially, there is a wide variety of polarizing prisms. The diagonal faces are usually bonded with an optical medium other than air, e.g., *Canada balsam*,* which has a refractive index in the range of 1.54 to 1.55. This index is in the middle of the two refractive indices of calcite, n_o and n_e, so there will only be the critical angle for the ordinary ray, i.e., $\theta_o^{(c)} = 68.25^\circ$. A Glan–Thompson prism made of calcite, with the optic axis orthogonal to the plane of incidence, is shown in Fig. 2.31. The geometry of the prisms is such that the angle of incidence on the first diagonal face is greater than $\theta_o^{(c)}$.

Other types of prisms made of quartz are shown in Fig. 2.32. In these prisms, the optic axes change direction and the diagonals are inclined by 45°. In Wollaston prisms, $\mathbf{E}_i^{\perp}$ in the first prism is the ordinary ray; it changes to an extraordinary ray when passing to the second prism, so that on the diagonal it

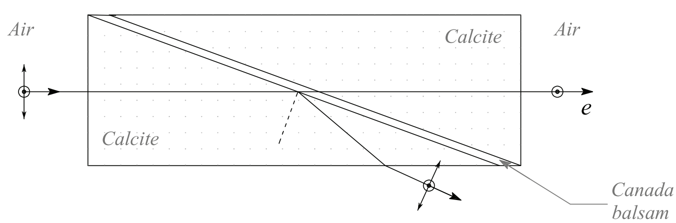


Figure 2.31 Glan–Thompson prism made with two calcite prisms bonded with Canada balsam.

*Canada balsam is a tree resin that, due to its transparency and its refractive index close to glass (after being subjected to an oil evaporation process), is used to glue lenses in optical instruments.

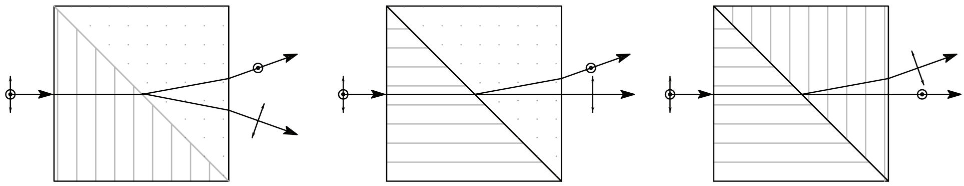


Figure 2.32 Some types of polarizing prisms that separate the components of the optical field into the "s" and "p" states. In all cases, the prisms are quartz [2].

is refracted approaching the normal. On the other hand, $\mathbf{E}_i^{\parallel}$ is the extraordinary ray in the first prism; it changes to an ordinary ray when passing to the second prism, so that on the diagonal it refracts away from the normal. In Rochon prisms, $\mathbf{E}_i^{\perp}$ and $\mathbf{E}_i^{\parallel}$ enter the first prism along the optic axis, so both rays will see the ordinary index. In the second prism, $\mathbf{E}_i^{\perp}$ changes to an extraordinary ray and refracts closer to the normal. In contrast, $\mathbf{E}_i^{\parallel}$ continues as an ordinary ray, so its propagation direction does not change. Finally, in Sénarmont prisms, $\mathbf{E}_i^{\perp}$ and $\mathbf{E}_i^{\parallel}$ also enter the first prism along the optic axis, so both rays will see the ordinary index. In the second prism, $\mathbf{E}_i^{\perp}$ continues as an ordinary ray, so its direction of propagation does not change, and $\mathbf{E}_i^{\parallel}$ changes to an extraordinary ray and refracts closer than normal.

There are also polarizing prisms made of isotropic optical glass. These are beamsplitter cubes with a dielectric film between the diagonals that joins the prisms, allowing the transmission of the p-polarization state and the reflection in the diagonal of the s-polarization state. These are mentioned briefly in Appendix E.

2.6 Vectors and Jones Matrices

To describe the polarization state of a plane wave in Section 2.1, we use the vector [Eq. (2.32)]

$$\mathbf{E}(x, y, z; t) = \{\hat{\imath}E_{ox} + \hat{\jmath}E_{oy}\}e^{i(kz-\omega t)},$$

where $E_{ox} = |E_{ox}|e^{i\delta_x}$ and $E_{oy} = |E_{oy}|e^{i\delta_y}$. Because the temporal and spatial phase terms are common to the complex amplitudes of the two wave components, it is convenient to represent the state of polarization as a column vector in which its elements determine the relationship between the components of the wave. This representation is known as a *Jones vector*:

$$\begin{bmatrix} E_{ox} \\ E_{oy} \end{bmatrix} = \begin{bmatrix} |E_{ox}|e^{i\delta_x} \\ |E_{oy}|e^{i\delta_y} \end{bmatrix}. \tag{2.154}$$

For some of the more common polarization states, the explicit representations are as follows.

Linear polarization

The phase shift between the components must be $\Delta\delta = m\pi$ ($m = 0, \pm 1, \pm 2, \ldots$). Therefore,

$$\begin{bmatrix} |E_{ox}| \\ \pm|E_{oy}| \end{bmatrix}. \tag{2.155}$$

In the most common cases of linear polarization, the Jones vector can be reduced depending on the orientation of the plane of vibration. For horizontal polarization ($\mathbf{E}_H$), $|E_{oy}| = 0$,

$$|E_{ox}| \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \tag{2.156}$$

for vertical polarization ($\mathbf{E}_V$), $|E_{ox}| = 0$,

$$\pm|E_{oy}| \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \tag{2.157}$$

and for diagonal polarization at ±45° ($\mathbf{E}_{\pm 45}$), $|E_{ox}| = |E_{oy}|$,

$$|E_{ox}| \begin{bmatrix} 1 \\ \pm 1 \end{bmatrix}. \tag{2.158}$$

Circular polarization

The phase shift between the components must be $\Delta\delta = (2m - 1)\pi/2$; ($m = 0, \pm 1, \pm 2, \ldots$) and $|E_{ox}| = |E_{oy}|$. Therefore, the Jones vector becomes

$$|E_{ox}| \begin{bmatrix} 1 \\ \pm i \end{bmatrix}. \tag{2.159}$$

The (+) sign is for the left circular polarization ($\mathbf{E}_L$), and the (–) sign is for the right circular polarization ($\mathbf{E}_R$).

Elliptical polarization

When the phase shift between the components is $\Delta\delta = (2m - 1)\pi/2$; ($m = 0, \pm 1, \pm 2, \ldots$) and $|E_{ox}| \neq |E_{oy}|$, which corresponds to an unrotated ellipse, the Jones vector would be

$$|E_{ox}| \begin{bmatrix} 1 \\ \pm ib \end{bmatrix}, \tag{2.160}$$

Table 2.3 Some polarization states in Jones vector notation.

Jones vector	Polarization state
$\begin{bmatrix}1\\0\end{bmatrix}$	$\mathbf{E}_H$, (Horizontal linear)
$\begin{bmatrix}0\\1\end{bmatrix}$	$\mathbf{E}_V$, (Vertical linear)
$\begin{bmatrix}1\\\pm1\end{bmatrix}$	$\mathbf{E}_{\pm45}$, (Linear diagonal ±45°)
$\begin{bmatrix}1\\i\end{bmatrix}$	$\mathbf{E}_L$, (Left circular)
$\begin{bmatrix}1\\-i\end{bmatrix}$	$\mathbf{E}_R$, (Right circular)

with $b = |E_{oy}|/|E_{ox}|$. The (+) sign is for the left elliptical ($\mathbf{E}_L$) polarization, and the (–) sign is for the right elliptical polarization ($\mathbf{E}_R$).

Any other state of polarization will be described by Eq. (2.154). The polarization states most commonly used are shown in Table 2.3. Note that the factor that multiplies the column vector is omitted since it is common to both elements. Thus, to represent a state of polarization, only the simplest form of the Jones vector is used.

Operations between polarization states can be performed with Jones vectors. For example, the sum of two mutually orthogonal linear polarization states

$$\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}$$

results in a state of diagonal polarization; the sum of a linear and circular polarization state

$$\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}1\\i\end{bmatrix}=2\begin{bmatrix}1\\0.5i\end{bmatrix}$$

results in a state of elliptical polarization; and the sum of two circular polarization states, one to the left and one to the right,

$$\begin{bmatrix}1\\i\end{bmatrix}+\begin{bmatrix}1\\-i\end{bmatrix}=2\begin{bmatrix}1\\0\end{bmatrix}$$

results in a state of linear polarization.

On the other hand, polarizing elements can also be conveniently represented as 2×2 matrices, such that the effect of a polarizing element on a polarization state is described by a linear transformation. Thus, if

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{2.161}$$

describes the polarizing element, the result on a state of polarization $\begin{bmatrix} A \\ B \end{bmatrix}$ is

$$\begin{bmatrix} A' \\ B' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}. \tag{2.162}$$

For example, consider a linear polarizer with its transmission axis horizontal. The result over $\mathbf{E}_H$ is again $\mathbf{E}_H$, and the result over $\mathbf{E}_V$ is 0; i.e.,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.163}$$

and

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.164}$$

It is immediately clear that from Eq. (2.163) $a=1$ and $c=0$, and from Eq. (2.164) $b=0$ and $d=0$. Therefore, the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{2.165}$$

represents a linear polarizer with its transmission axis horizontal.

Table 2.4 shows some matrices that represent polarizing elements. Some of the matrices are accompanied by the factors 1/2 and $1/\sqrt{2}$, which are necessary when energy balance is required but can be omitted for the analysis of polarization changes. Multiple representations can exist for the same element; e.g., a left circular polarizer element is represented in Table 2.4 by

Table 2.4 Some polarizers represented as Jones matrices.

Element	Jones matrix		
Linear polarizer	horizontal $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	vertical $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$	diagonal $\frac{1}{2}\begin{bmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{bmatrix}$
Plate λ/4	fast axis horizontal $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$	fast axis vertical $\begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix}$	fast axis ±45° $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & \pm i \\ \pm i & 0 \end{bmatrix}$
Plate λ/2	fast axis horizontal $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	fast axis vertical $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	
Circular polarizer	right $\frac{1}{2}\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$	left $\frac{1}{2}\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$	

$$\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}.$$

But the same polarizer can also be represented by the matrix that results from combining a diagonal linear polarizer with a λ/4 plate, i.e.,

$$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ i & i \end{bmatrix}.$$

With either arrangement, a linear polarization state ($\mathbf{E}_H$, $\mathbf{E}_V$, $\mathbf{E}_{\pm 45}$) can be transformed into a circular $\mathbf{E}_L$ polarization state, which can be easily verified as follows:

$$\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}\begin{bmatrix} A \\ B \end{bmatrix} = (A - iB)\begin{bmatrix} 1 \\ i \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 1 \\ i & i \end{bmatrix}\begin{bmatrix} A \\ B \end{bmatrix} = (A + B)\begin{bmatrix} 1 \\ i \end{bmatrix}.$$

As an example, let us consider the system shown in Fig. 2.33. There are three optical elements aligned with the z axis: a linear polarizer LP with its transmission axis rotated 45°, a λ/4 plate with its fast (f) axis vertical, and a glass plate reflecting $r_{\parallel} = 0.2$ and $r_{\perp} = -0.2$.

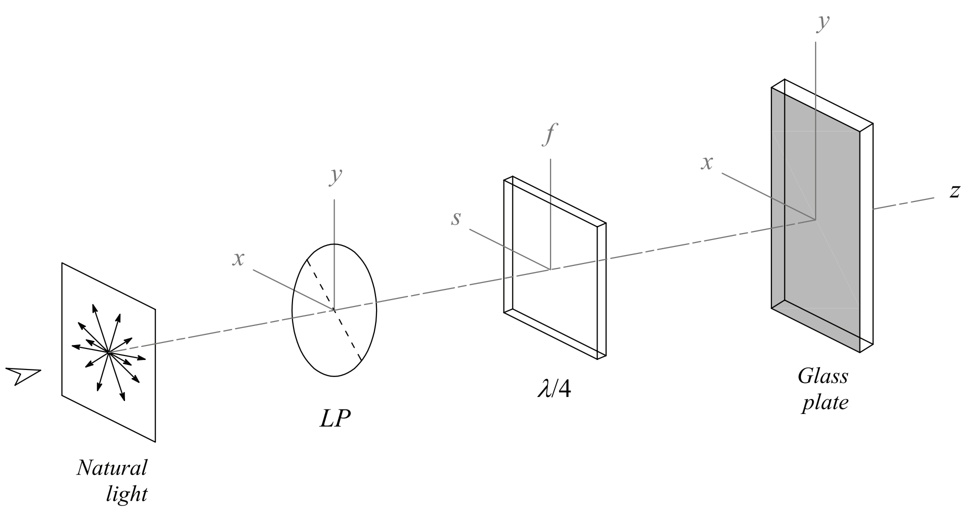


Figure 2.33 Optical system that cancels the reflected light after passing through the λ/4 plate.

If (from the left) natural light hits the linear polarizer, the reflected light will vanish when it hits the linear polarizer again. Qualitatively (observing from the positive side of the z axis), it follows that after the linear polarizer the light becomes linearly polarized at 45°; passing through the $\lambda/4$ plate it becomes right-circularly polarized; reflecting off the glass plate it changes to left-circularly polarized; and when it passes again through the $\lambda/4$ plate during its return, it becomes linearly polarized at $-45°$ and no light is transmitted back through the polarizer LP. Mathematically, with vectors and Jones matrices, it will be:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}.$$

Note that the arrays are written in the reverse order of the optical elements (from left to right). The product of the matrices is indeed 0. The middle matrix represents the reflecting surface at normal incidence, which is written as

$$\begin{bmatrix} r_\perp & 0 \\ 0 & r_\parallel \end{bmatrix} = 0.2 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

On the other hand, when light is reflected, the positive direction of the x axis changes. Although this has no effect on the $\lambda/4$ plate, since its fast (f) and slow (s) axes are still oriented vertically and horizontally, it impacts the linear polarizer, since its transmission axis will now be at $-45°$. This explains the sign change in the matrix representing the linear polarizer (first matrix, after the = sign) on the return of the light.

In addition to Jones vectors and matrices, there are other possible mathematical representations, such as Stokes parameters, the Poincaré sphere, and Mueller matrices.

References

[1] E. H. Land, "Some aspects of the development of sheet polarizers," *JOSA* **41**(12), 957–963 (1951).
[2] G. R. Fowles, *Introduction to Modern Optics*, 2nd ed., Dover Publications, Mineola, New York (1989).