Chapter 2
Optical Components and Methods

Abstract In this chapter, after presenting Maxwell equations inside matter and describing reflection and refraction processes, the interaction of optical waves with simple optical components such as mirrors, prisms, and lenses is examined. The different methods for measuring the power spectrum of an optical beam are described and compared in Sect. 2.3. In Sect. 2.4, after treating wave propagation in anisotropic materials, the methods used for fixing or modifying the polarization state of an optical beam are presented. Finally, Sect. 2.5 introduces the important subject of optical waveguides.

2.1 Electromagnetic Waves in Matter

Inside a medium Maxwell's equations take the following form:

$$\nabla \cdot \mathbf{D} = \rho \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.2}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.4}$$

where $\mathbf{D}$ is the electric displacement vector (measured in units of C/m^2). The quantities ρ and $\mathbf{J}$ are, respectively, the electric charge density (C/m^3) and the current density (A/m^2). Here the treatment is limited to dielectric media, in which there are no free electric charges or currents, so that one can put $\mathbf{J} = \rho = 0$.

The relations connecting $\mathbf{D}$ to $\mathbf{E}$ and $\mathbf{B}$ to $\mathbf{H}$ inside the medium become:

$$\mathbf{D} = \varepsilon_o \mathbf{E} + \mathbf{P} \tag{2.5}$$

$$\mathbf{B} = \mu_o (\mathbf{H} + \mathbf{M}). \tag{2.6}$$

V. Degiorgio and I. Cristiani, *Photonics*, Undergraduate Lecture Notes in Physics,
DOI 10.1007/978-3-319-20627-1_2

The new vectors $\mathbf{P}$ and $\mathbf{M}$ are, respectively, the unit-volume electric polarization and the unit-volume magnetization inside the medium. In this book non-ferromagnetic media are only considered, and so it is possible to simplify the treatment by putting $\mathbf{M} = 0$. The wave equation inside the medium is the following:

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu_o \frac{\partial^2 \mathbf{P}}{\partial t^2}. \tag{2.7}$$

In order to solve wave propagation problems, it is necessary to associate an equation relating $\mathbf{P}$ to $\mathbf{E}$ to (2.7). It is useful to recall what is the physical meaning of $\mathbf{P}$. The medium is a collection of positively charged nuclei and negatively-charged bound electrons. Under the action of an electric field the electron clouds are slightly displaced from their equilibrium position, so that a macroscopic electric dipole is created. Limiting the discussion to the time-dependence, it is clear that the medium response to the application of an electric field cannot be instantaneous. Therefore, a convolution relation should be used:

$$\mathbf{P}(\mathbf{r}, t) = \varepsilon_o \int_{-\infty}^{t} R(t - t')\mathbf{E}(\mathbf{r}, t')dt', \tag{2.8}$$

where $R(t - t')$ is the response function of the medium.

The electric field $E(\mathbf{r}, t)$ associated with a generic light beam can always be described as a weighted superposition of sinusoidal functions oscillating at different frequencies, as expressed by the integral:

$$E(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega E(\mathbf{r}, \omega)\exp(-i\omega t). \tag{2.9}$$

The complex quantity $E(\mathbf{r}, \omega)$ represents the weight of the component at angular frequency ω. From a mathematical point of view, $E(\mathbf{r}, \omega)$ is the Fourier transform of $E(\mathbf{r}, t)$.

The inversion of (2.9) gives:

$$E(\mathbf{r}, \omega) = \int_{-\infty}^{\infty} dt E(\mathbf{r}, t)\exp(i\omega t). \tag{2.10}$$

Recalling that the Fourier transform of the convolution of two functions is the product of their Fourier transforms, the Fourier transform of $\mathbf{P}(\mathbf{r}, t)$ is expressed by:

$$\mathbf{P}(\mathbf{r}, \omega) = \varepsilon_o \chi(\omega)\mathbf{E}(\mathbf{r}, \omega), \tag{2.11}$$

where $\chi(\omega)$, called electric susceptibility, is given by:

$$\chi(\omega) = \int_{-\infty}^{\infty} R(t)\exp(i\omega t)dt. \tag{2.12}$$

By substituting (2.11) inside (2.5), it is found that the electric displacement $\mathbf{D}(\mathbf{r}, \omega)$, generated by a monochromatic electric field oscillating at the frequency ω, is proportional to $\mathbf{E}(\mathbf{r}, \omega)$:

$$\mathbf{D}(\mathbf{r}, \omega) = \varepsilon_o (1 + \chi) \mathbf{E}(\mathbf{r}, \omega) = \varepsilon_o \varepsilon_r \mathbf{E}(\mathbf{r}, \omega). \tag{2.13}$$

The dimensionless constant ε_r is called relative dielectric constant or relative electric permittivity. The wave propagation velocity inside the dielectric medium is:

$$u = \frac{1}{\sqrt{\varepsilon_o \varepsilon_r \mu_o}} = \frac{c}{n} \tag{2.14}$$

The quantity n, called index of refraction of the medium, is given by:

$$n = \sqrt{\varepsilon_r} = \sqrt{1 + \chi}. \tag{2.15}$$

In the more general case, in which the relative magnetic permeability of the material, μ_r, may be different from 1, it is found that $n = \sqrt{\varepsilon_r \mu_r}$.

Since the frequency of oscillation cannot change when the wave goes from one medium to the other, the change of velocity is associated with a change of wavelength. Specifically, if λ is the vacuum wavelength, then the wavelength inside a medium of index of refraction n is: $\lambda' = \lambda / n$.

Since $\chi(\omega)$ is frequency-dependent, also $n(\omega) = \sqrt{1 + \chi(\omega)}$ is frequency-dependent. This phenomenon is called optical dispersion. As an example, over its transparency window going from 0.35 to 2.3 μm, the index of refraction of the borosilicate glass called BK7 decreases monotonically from 1.54 at 0.35 μm to 1.49 at 2.3 μm. This behavior is called normal dispersion. The dispersion is said to be anomalous when the index of refraction increases as a function of wavelength. As it will be seen in Chap. 6, the propagation of ultrashort light pulses is strongly influenced by optical dispersion.

The electric field of the monochromatic plane wave inside the medium is still given by (1.9). The only difference is that the modulus of the propagation vector is now:

$$k = \frac{\omega}{u} = \frac{\omega}{c} \sqrt{1 + \chi} = \frac{\omega}{c} n, \tag{2.16}$$

and (1.12) becomes:

$$I = c \frac{\varepsilon_o n E_o^2}{2}. \tag{2.17}$$

The propagation inside a medium with a real χ proceeds similarly to vacuum propagation, the only difference being in the propagation velocity. However, if χ is a complex quantity, then n is also complex, and can be written as: $n = n' + i n''$. Assuming that a monochromatic plane wave, traveling along z, enters into the medium at $z = 0$, the electric field at the generic coordinate z is:

$$\mathbf{E} = \mathbf{E_o}\exp\{-i([\omega t - (\omega/c)(n' + in'')z]\}. \tag{2.18}$$

The intensity $I(z)$, which is proportional to the square of the field modulus, is given by:

$$I(z) = I_o\exp[-2(\omega/c)n''z] = I_o\exp(-\alpha z), \tag{2.19}$$

where I_o is the intensity at $z = 0$. Equation (2.19) shows that, if $n'' > 0$, the wave intensity decays exponentially during propagation with an attenuation coefficient, $\alpha = 2(\omega/c)n''$, proportional to the imaginary part of the refractive index. The attenuation may be due to absorption or to scattering.

2.2 Reflection and Refraction

In this section the laws of reflection and refraction are introduced and applied to describe the behavior of different optical components, such as mirrors and lenses.

2.2.1 Dielectric Interface

A monochromatic plane wave is incident at the planar boundary between two dielectric media, characterized by refractive indices n_1 and n_2. According to the geometry shown in Fig. 2.1, the wave is coming from medium 1, and its wave-vector $\mathbf{k_i}$ forms an angle θ_i with the normal to the boundary. The planar boundary is taken to coincide with the x-y plane, and $\mathbf{k_i}$ lies on the y-z plane. The plane containing $\mathbf{k_i}$ and the normal to the planar boundary is known as the incidence plane. In the case of Fig. 2.1, y-z is the incidence plane.

Part of the wave will be reflected and part will be transmitted. Let θ_r and θ_t be the reflection and transmission angles, respectively. By imposing the continuity conditions at the boundary for the tangential components of the electric and magnetic

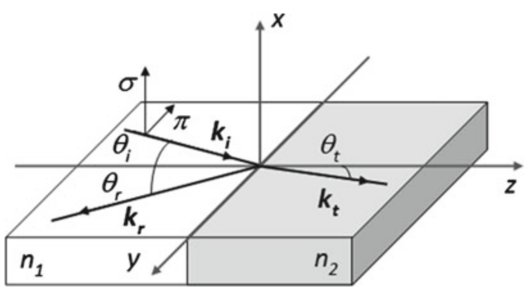


Fig. 2.1 Reflection and refraction at a dielectric boundary

fields, it is possible to calculate the direction and the complex amplitude of the reflected and transmitted (also called refracted) waves.

It is useful to consider separately two different polarization states of the incident wave: (a) a wave linearly polarized in a direction perpendicular to the incidence plane, this is usually called σ case; (b) a wave linearly polarized in the incidence plane, this is usually called π case. Any arbitrary polarization state of the incident wave can always be described as a linear superposition of these two states.

- σ case.

Putting $|\mathbf{k_i}| = k_o n_1$, where $k_o = 2\pi/\lambda$, the electric field of the incident wave is:

$$\mathbf{E_i} = E_i \exp(-i\omega t + ik_o y n_1 \sin\theta_i + ik_o z n_1 \cos\theta_i)\mathbf{x}, \tag{2.20}$$

where $\mathbf{x}$ is the unit vector directed along x. The electric fields $\mathbf{E_r}$ and $\mathbf{E_t}$ of the reflected and transmitted waves are written as:

$$\mathbf{E_r} = E_r \exp(-i\omega t + ik_o y n_1 \sin\theta_r - ik_o z n_1 \cos\theta_r)\mathbf{x}, \tag{2.21}$$

and

$$\mathbf{E_t} = E_t \exp(-i\omega t + ik_o y n_2 \sin\theta_t + ik_o z n_2 \cos\theta_t)\mathbf{x}. \tag{2.22}$$

The continuity condition for the tangential component of the electric field at $z = 0$ is:

$$E_i \exp(ik_o y n_1 \sin\theta_i)\mathbf{x} + E_r \exp(ik_o y n_1 \sin\theta_r)\mathbf{x} = E_t \exp(ik_o y n_2 \sin\theta_t)\mathbf{x}. \tag{2.23}$$

In order to satisfy (2.23) for any y, the reflection and transmission angles should satisfy the relations:

$$\theta_i = \theta_r, \tag{2.24}$$

and

$$n_1 \sin\theta_i = n_2 \sin\theta_t. \tag{2.25}$$

Equation (2.25) is known as Snell's law.

By using (2.24) and (2.25), (2.23) simply becomes

$$E_i + E_r = E_t \tag{2.26}$$

In order to derive E_r and E_t, a second equation is needed besides that of (2.26). The continuity of the tangential component of the magnetic field is expressed as:

$$H_i \cos\theta_i - H_r \cos\theta_r = H_t \cos\theta_t \tag{2.27}$$

Recalling that the electric field amplitude is related to the magnetic field amplitude by the expression $H = \sqrt{\varepsilon_r \varepsilon_o/\mu_o} E$, (2.27) can be written as:

$$E_i \cos\theta_i - E_r \cos\theta_r = \frac{n_2}{n_1} E_t \cos\theta_t. \tag{2.28}$$

The solution of the system of (2.26) and (2.28) can be presented in terms of the transmission and reflection coefficients, τ_σ and ρ_σ, respectively:

$$\tau_\sigma = \frac{E_t}{E_i} = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t} = \frac{2 \sin\theta_t \cos\theta_i}{\sin(\theta_i + \theta_t)}, \tag{2.29}$$

$$\rho_\sigma = \frac{E_r}{E_i} = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} = -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)}. \tag{2.30}$$

If the incident beam is coming from medium 2, one can repeat the treatment finding two new coefficients, τ'_σ and ρ'_σ, related to the previous ones as follows:

$$\rho'_\sigma = -\rho_\sigma; \quad \tau_\sigma \tau'_\sigma = 1 - \rho_\sigma^2 \tag{2.31}$$

- π case.

Following a treatment similar to that of the σ case, it is found that:

$$\tau_\pi = \frac{E_t}{E_i} = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_t + n_2 \cos\theta_i} = \frac{2 \sin\theta_t \cos\theta_i}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)}, \tag{2.32}$$

$$\rho_\pi = \frac{E_r}{E_i} = \frac{n_1 \cos\theta_t - n_2 \cos\theta_i}{n_1 \cos\theta_t + n_2 \cos\theta_i} = -\frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)}. \tag{2.33}$$

If the beam arrives at the boundary from medium 2, it is found that the relations between the coefficients τ'_π, ρ'_π and the coefficients τ_π and ρ_π are identical to those of (2.31).

In the case of normal incidence, $\theta_i = \theta_r = \theta_t = 0$, and ρ_σ coincides with ρ_π:

$$\rho_\sigma(\theta_i = 0) = \rho_\pi(\theta_i = 0) = -\frac{n_2 - n_1}{n_2 + n_1}. \tag{2.34}$$

If $n_2 > n_1$, the field reflection coefficient is negative. This indicates that the phase of E_r is changed by π with respect to E_i. If $n_2 < n_1$, E_r has the same phase as E_i.

In order to derive how the incident power splits between transmission and reflection, it should be recalled that the power arriving on a generic surface Σ, is determined as the time average of the flux of the Poynting vector through the surface. Projecting the Poynting vector on the direction perpendicular to the surface, allows the following expression for P_i to be given:

$$P_i = (1/2)c\varepsilon_o n_1 \Sigma |E_i|^2 \cos\theta_i. \tag{2.35}$$

Similarly, for the transmitted and reflected power:

$$P_r = (1/2)c\varepsilon_o n_1 \Sigma |E_r|^2 \cos\theta_i \qquad P_t = (1/2)c\varepsilon_o n_2 \Sigma |E_t|^2 \cos\theta_t. \tag{2.36}$$

Therefore the reflectance R and the transmittance T are given by:

$$R = \frac{P_r}{P_i} = \frac{|E_r|^2}{|E_i|^2} = |\rho|^2 \tag{2.37}$$

$$T = \frac{P_t}{P_i} = |\tau|^2 \frac{n_2 \cos\theta_t}{n_1 \cos\theta_i}. \tag{2.38}$$

Note that $R + T = 1$, whereas the sum $|\rho|^2 + |\tau|^2$ is, in general, $\neq 1$.

The behavior of R as a function of θ_i depends on whether n_1 is smaller or larger than n_2. Therefore the two cases are discussed separately.

- $n_1 < n_2$

Considering, as an example, an air/glass boundary with a ratio $n_2/n_1 = 1.5$, the behavior of R_π e R_σ for a wave coming from air is illustrated in Fig. 2.2. At normal incidence ($\theta_i = 0$), using (2.34), it is found that $R_\pi(\theta_i = 0) = R_\sigma(\theta_i = 0) = 0.04$. The reflectance R_σ is monotonically increasing with θ_i reaching a value of 1 at $\theta_i = \pi/2 = 90°$. On the other hand, R_π at first decreases till it vanishes at $\theta_i = \theta_B$, and then increases, becoming equal to 1 at $\theta_i = \pi/2$.

The angle θ_B is derived by noting that the reflection coefficient ρ_π vanishes if the denominator of the fraction at right-hand side of (2.33) becomes infinite. This happens if: $\theta_i + \theta_t = \pi/2$. From this condition, by using Snell's law, the following expression of θ_B, called Brewster's angle, is derived:

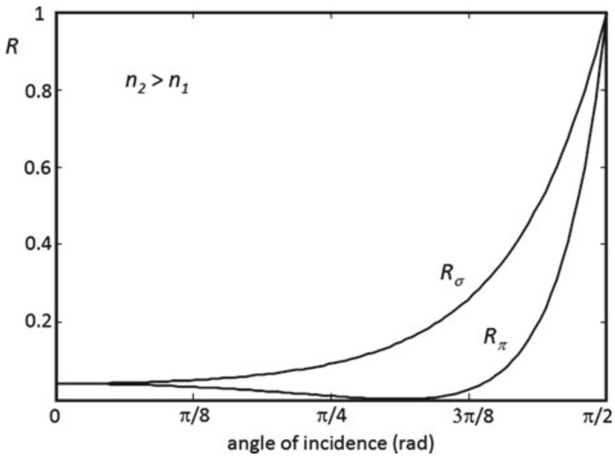


Fig. 2.2 Reflectance versus incidence angle for $n_2/n_1 = 1.5$

$$\tan\theta_B = \frac{n_2}{n_1} \tag{2.39}$$

As an example, if $n_2/n_1 = 1.5$, $\theta_B = 56°$, and $R_\sigma(\theta_B) = 0.147$.

In general, ρ is a complex quantity that can be written as $\rho = \sqrt{R}\exp(i\phi)$, where ϕ is the phase shift of the reflected field with respect to the incident field. In the σ case, (2.30) shows that $\phi_\sigma = \pi$ over the whole interval $0 \le \theta_i \le \pi/2$. In the π case, (2.33) indicates that ϕ_π is equal to π within the interval $0 \le \theta_i < \theta_B$, and vanishes for $\theta_B \le \theta_i \le \pi/2$.

- $n_1 > n_2$

In this case, Snell's law indicates that θ_t is larger than θ_i. Therefore, as θ_i grows, there will be an incidence angle, θ_c, smaller than $\pi/2$, at which θ_t becomes $\pi/2$. If $\theta_i \ge \theta_c$ there is no transmitted beam, and so the power reflection coefficient is equal to 1. The total reflection angle or limit angle, θ_c, is defined by the relation:

$$\sin\theta_c = n_2/n_1 \tag{2.40}$$

Considering the glass/air interface, $n_1/n_2 = 1.5$, for which $\theta_c = 41.8°$, the behavior of R_π and R_σ is illustrated in Fig. 2.3. As in the previous case, R_π vanishes at the Brewster angle, $\theta_B = 33.7°$.

The field reflection coefficient is a real quantity for $\theta_i \le \theta_c$. If $\theta_i > \theta_c$, $\cos\theta_t$ becomes imaginary:

$$\cos\theta_t = \sqrt{1-\sin^2\theta_t} = \sqrt{1-\frac{n_1^2}{n_2^2}\sin^2\theta_i} = i\sqrt{\frac{n_1^2}{n_2^2}\sin^2\theta_i - 1}. \tag{2.41}$$

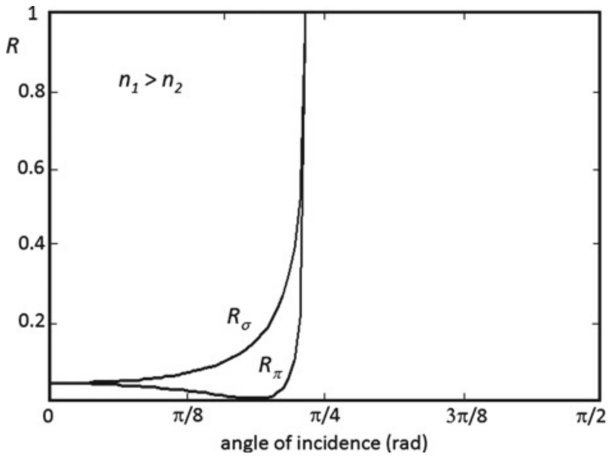


Fig. 2.3 Reflectance versus incidence angle for $n_1/n_2 = 1.5$

By substituting (2.41) into the expressions for ρ, it is seen that the field reflection coefficient becomes complex. In the σ case, the phase ϕ_σ of ρ_σ is 0 for $\theta_i \leq \theta_c$, and grows from 0 to π in the interval $\theta_c \leq \theta_i \leq \pi/2$ according to the expression:

$$\tan\left(\frac{\phi_\sigma}{2}\right) = \frac{\sqrt{\sin^2\theta_i - \sin^2\theta_c}}{\cos\theta_i}. \tag{2.42}$$

In the π case, ϕ_π vanishes in the interval $0 \leq \theta_i \leq \theta_B$, and is equal to π in the interval $\theta_B < \theta_i \leq \theta_c$, decreasing from π to 0 in the interval $\theta_c \leq \theta_i \leq \pi/2$ according to the expression:

$$\tan\left(\frac{\phi_\pi}{2}\right) = \frac{\cos\theta_i \sin^2\theta_c}{\sqrt{\sin^2\theta_i - \sin^2\theta_c}}. \tag{2.43}$$

2.2.2 Reflection from a Metallic Surface

The most commonly used mirrors are made by depositing a thin metallic layer on a glass substrate. It is therefore interesting to discuss the case in which medium 2 in Fig. 2.1 is a metal. Metals contain electrons that are free to move inside the crystal. In a simplified approach, the relative dielectric constant of the metal is written as the sum of two terms, one due to bound electrons, ε', and the other related to free electrons:

$$\varepsilon_m = \varepsilon' + i\frac{\sigma}{\varepsilon_o\omega}, \tag{2.44}$$

where σ is the frequency-dependent electric conductivity of the metal:

$$\sigma(\omega) = \sigma(0)\frac{1 + i\omega\tau}{1 + \omega^2\tau^2}. \tag{2.45}$$

The quantity τ is a relaxation time that depends on the nature of the metal. The zero-frequency conductivity is given by:

$$\sigma(0) = \frac{Ne^2\tau}{m_e}, \tag{2.46}$$

where N is the density of free electrons, m_e is the electron mass, and e is the electron charge.

The free electron gas inside the metal has a characteristic frequency of oscillation, called the plasma frequency, related to $\sigma(0)$ by the expression:

$$\omega_p = \sqrt{\frac{\sigma(0)}{\varepsilon' \varepsilon_o \tau}} = \sqrt{\frac{Ne^2}{\varepsilon' \varepsilon_o m_e}} \tag{2.47}$$

Electromagnetic radiation having frequency close to the plasma frequency is strongly absorbed by the metal, and so the refractive index must be written as a complex quantity, $n_m = \sqrt{\varepsilon_m} = n' + in''$. Limiting the discussion to normal incidence ($\theta_i = 0$), the reflectance of the air-metal interface, derived by using (2.34), is given by:

$$R = \frac{|n_m - 1|^2}{|n_m + 1|^2} = \frac{n'^2 + n''^2 + 1 - 2n'}{n'^2 + n''^2 + 1 + 2n'}. \tag{2.48}$$

Approximate expressions for n' and n'' can be derived from (2.44) by assuming that the contribution of bound electrons is negligible in comparison to that of free electrons. Taking into account that, for a typical metal, $\tau \approx 10^{-13}$ s and $\omega_p \approx 10^{15}$ s^{-1}, three different frequency intervals can be distinguished:

- infrared: $\omega \ll 1/\tau \ll \omega_p$. The result is: $R = 1$.
- near infrared and visible: $1/\tau \ll \omega \ll \omega_p$. It is found that: $R = 1 - 2/(\omega_p \tau)$, the reflectance is large, but lower than 100 %.
- ultraviolet: $\omega_p \ll \omega$. It is found: $R = \omega_p^4/(16\omega^6\tau^2) \ll 1$. In this case the reflectance is very low.

Experimental values of n', n'' and R for three metals at three different wavelengths are reported in Table 2.1. The data refer to thin layers deposited by evaporation on a glass substrate.

Table 2.1 Complex index of refraction and normal-incidence reflectance for some metal surfaces at three distinct wavelengths

Metal	λ (nm)	n′	n″	R
Silver	400	0.08	1.9	0.94
	550	0.06	3.3	0.98
	700	0.08	4.6	0.99
Aluminum	400	0.4	4.5	0.93
	550	0.8	6	0.92
	700	1.5	7	0.89
Gold	450	1.4	1.9	0.40
	550	0.33	2.3	0.82
	700	0.13	3.8	0.97

One problem with metallic mirrors is that the fraction of incident power that is not reflected is absorbed. When reflecting high-intensity beams, such as those coming from laser sources, the temperature rise due to this absorption may cause damage to the metallic layer.

2.2.3 Anti-reflection Coating

The reflectance of a boundary separating two transparent media can be strongly modified by inserting thin dielectric layers. The structure investigated in this section, shown in Fig. 2.4, consists of a single layer of refractive index n_2 and thickness d separating media 1 and 3 which have refractive indices n_1 and n_3, respectively. It is assumed that $n_1 < n_2 < n_3$.

Considering a monochromatic plane wave perpendicularly incident on the boundary 1–2, and calling E_i the incoming electric field, the first contribution to the reflected field is: $E_{r1} = \rho_{12} E_i$. The transmitted field $E_{t1} = \tau_{12} E_i$ is partially back-reflected by the interface 2–3, producing a second contribution to the reflected field given by: $E_{r2} = \exp(i\Delta)\rho_{23}\tau_{21}E_{t1}$. This contribution is phase-shifted with respect to the first one because the wave has traveled back and forth inside layer 2. The shift is: $\Delta = 4\pi n_2 d/\lambda$, where λ is the vacuum wavelength of the incident wave. Note that, in this case, the coefficients ρ_{12} and ρ_{23} have the same sign, so that no contribution to the phase shift between E_{r1} and E_{r2} is coming from the reflections at the interfaces.

Since the wave back-reflected from the surface 2–3 is partially reflected again when it reaches the surface 2–1, the total reflected field should be calculated by adding an infinite number of multiple reflections. Assuming that the reflectance of the interfaces 1–2 and 2–3 is small, which also means that $\tau_{21}\tau_{12} \approx 1$, the summation can be truncated by taking only the first two terms:

$$E_r = E_{r1} + E_{r2} = [\rho_{12} + \exp(i\Delta)\rho_{23}]E_i \tag{2.49}$$

It is clear from (2.49) that E_r vanishes through destructive interference if E_{r2} and E_{r1} have the same amplitude, but opposite sign. The amplitudes of the two

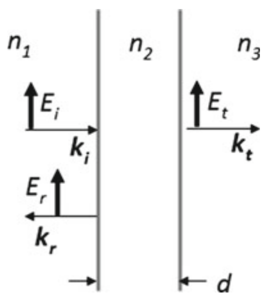


Fig. 2.4 Anti-reflection coating

fields are the same if $\rho_{12} = \rho_{23}$. According to (2.34), this is equivalent to saying that $(n_2/n_1) = (n_3/n_2)$, or: $n_2 = \sqrt{n_3 n_1}$. This means that the intermediate layer should have an index of refraction that is the geometric mean of n_1 and n_3. The two contributions have opposite sign if $\Delta = \pi$, which implies $d = \lambda/(4n_2)$. This means that the layer thickness must be a quarter of the wavelength inside the medium.

The fact that the condition imposed on the layer thickness d is wavelength-dependent indicates that there is a price to pay for eliminating reflection: the method works exactly at only one specific wavelength. In other words this anti-reflection coating is chromatic.

Instead of using the approximate expression given by (2.49), E_r can be exactly calculated by summing up an infinite number of multiple reflections or by writing the boundary conditions at the two interfaces. The latter approach is the more convenient, also because it can be more easily generalized to an arbitrary number of layers.

Consider first the interface 1–2; at left there are two waves, the incident wave with field E_i and the reflected wave with field E_r, at right there is a wave leaving the interface, with field E_t' and one arriving at the interface, with field E_r'. The boundary conditions are:

$$\begin{aligned} E_i + E_r &= E_t' + E_r' \\ E_i - E_r &= (n_2/n_1)(E_t' + E_r'). \end{aligned} \tag{2.50}$$

Taking into account that the wave crossing layer 2 acquires a phase shift of $\Delta/2$, the following boundary conditions are written for the interface 2–3:

$$\begin{aligned} \exp(i\Delta/2)E_t' + \exp(-i\Delta/2)E_r' &= E_t \\ \exp(i\Delta/2)E_t' + \exp(-i\Delta/2)E_r' &= (n_3/n_2)E_t. \end{aligned} \tag{2.51}$$

Equations (2.50) and (2.51) constitute a linear system of four equations in four variables that can be exactly solved. The expression obtained for E_r is:

$$E_r = \rho_{12}E_i + \rho_{23}E_i \frac{\tau_{12}\tau_{21}\exp(i\Delta)}{1 - \rho_{23}\rho_{21}\exp(i\Delta)} \tag{2.52}$$

Recalling that $\tau_{12}\tau_{21} = 1 - \rho_{12}^2$, it is immediately apparent that the conditions for a vanishing E_r derived from (2.52) coincide with those imposed by the approximate expression (2.49).

Table 2.2 gives the index of refraction and the transparency window of several dielectric materials. The index of refraction of the thin film depends not only on the wavelength, but also on the preparation method, and so the values of n listed in Table 2.2 should be only considered as indicative. The deposition of layers with precisely controlled thickness is usually performed by vacuum evaporation methods. In order to have an idea of the required thickness, consider the anti-reflection coating of an air-glass surface. Taking $\lambda = 550$ nm, $n_1 = 1$ (air), and $n_3 = 1.52$ (glass), the

Table 2.2 Properties of materials used for deposition of thin films

Material	n	Transparency window (μm)
Criolite Na_3AlF_6	1.35	0.15–14
Magnesium fluoride MgF_2	1.38	0.12–8
Silicon oxide SiO_2	1.46	0.17–8
Aluminum oxide Al_2O_3	1.62	0.15–6
Silicon monoxide SiO	1.9	0.5–8
Zirconium oxide ZrO_2	2.00	0.3–7
Cerium oxide CeO_2	2.2	0.4–16
Titanium oxide TiO_2	2.3	0.4–12
Zinc sulfide ZnS	2.32	0.4–14
Cadmium telluride $CdTe$	2.69	0.9–16
Silicon Si	3.5	1.1–10
Germanium Ge	4.05	1.5–16
Lead telluride $PbTe$	5.1	3.9–20

coating should be made with a material having $n_2 = \sqrt{1.52} = 1.23$ and thickness $d = 550/(4n_2) = 112$ nm. In practice, given the list of available materials in Table 2.2, the best choice would be magnesium fluoride. A quarter-wave layer of MgF_2 reduces the reflectance from 4 to 1.3 %.

By using more than one layer it is possible to cancel completely the reflectance of the air-glass surface. For instance, the sequence of two quarter-wave layers with refractive indices n_2 and n_2', produces a vanishing reflectance if $n_2'/n_2 = \sqrt{n_3/n_1}$. Putting $n_1 = 1$ e $n_3 = 1.52$, one finds $n_2'/n_2 = 1.23$. Zero-reflectance is obtained by choosing from Table 2.2 ZrO_2 ($n_2' = 2.00$) and Al_2O_3 ($n_2 = 1.62$).

Up to now normal incidence has been assumed. What happens if θ_i is different from 0, as shown in Fig. 2.5? The phase difference between the first and the second reflection becomes:

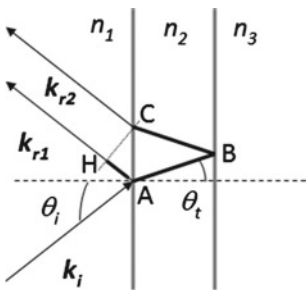


Fig. 2.5 Anti-reflection coating: non-normal incidence. The optical path difference between $\mathbf{k}_{r2}$ and $\mathbf{k}_{r1}$ is: $2n_2\overline{AB} - n_1\overline{AH}$

$$\Delta' = 2\,\frac{2\pi n_2}{\lambda}\,\frac{d}{\cos\theta_t} - \frac{2\pi n_1 d}{\lambda}\,tg\theta_t\,\sin\theta_i = \frac{4\pi n_2 d}{\lambda}\cos\theta_t, \tag{2.53}$$

where θ_t, the propagation angle inside medium 2, is related to θ_i by Snell's law. It is important to note that, once d and n_2 are fixed, the wavelength at which the reflectance vanishes, corresponding to $\Delta' = \pi$, depends on θ_i.

2.2.4 Multilayer Dielectric Mirror

By using a sequence of dielectric layers it is possible to obtain high reflectance at any desired wavelength. An important advantage of multilayer dielectric mirrors, as compared to metallic mirrors, is that there is no absorption. This means they can stand high intensities of visible light without being damaged. An important difference is that metallic mirrors have a broad-band reflectance, whereas the reflectance of dielectric mirrors is narrow-banded.

The structure of a multilayer dielectric mirror is sketched in Fig. 2.6: between medium 1 (air) and medium 2 (glass substrate) there is a sequence of alternating layers ABAB..A with refractive indices $n_A > n_B$, and $n_A > n_2$.

Clearly, the maximum reflectance will be obtained when the fields reflected by all the interfaces add up in phase at the first interface, so that constructive interference is exploited. The field reflected at the interface 1-A is phase-shifted by π with respect to the incident field because $n_A > n_1$. The field reflected at the first A–B interface does not suffer any phase-shift because $n_A > n_B$, but, upon arriving at the interface 1-A, it will acquire the phase shift $4\pi n_A d_A/\lambda$ because of the propagation back and forth within the layer thickness d_A. This second reflection will be in phase with the first one if the layer thickness corresponds to a quarter wavelength: $d_A = \lambda/(4n_A)$. By analyzing all subsequent reflections in a similar way, it is immediately clear that the maximum reflectance, $R_{\max}$, is obtained if all the layers A and B have a quarter wavelength thickness. The number $2N + 1$ of layers is always an odd number, in order to ensure that the last reflection (occurring at the A-2 interface) is also in phase with the previous ones.

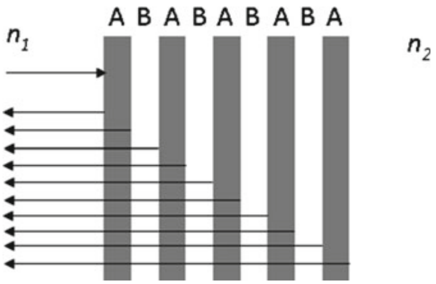


Fig. 2.6 Multilayer dielectric mirror

A detailed calculation can be developed by writing the boundary conditions at each interface, and noting that each pair of boundary conditions corresponding to a specific interface, such as (2.50) and (2.51), can be expressed in terms of a matrix equation. Therefore each interface introduces a transmission matrix, and the solution is always expressible as a cascade of matrices, independently from the number of layers.

The maximum reflectance at normal incidence is found to be:

$$R_{\max} = \left[\frac{1-(n_A/n_1)(n_A/n_2)(n_A/n_B)^{2N}}{1+(n_A/n_1)(n_A/n_2)(n_A/n_B)^{2N}}\right]^2. \tag{2.54}$$

Note that an expression of the type

$$\left(\frac{1-a}{1+a}\right),$$

can be transformed, by introducing the new variable $b = \ln a$, into

$$\left(\frac{1-a}{1+a}\right) = \left(\frac{e^{-b/2}-e^{b/2}}{e^{-b/2}+e^{b/2}}\right) = \tanh\left(\frac{b}{2}\right).$$

Therefore $R_{\max}$ can be expressed as:

$$R_{\max} = \tanh^2(N\xi), \tag{2.55}$$

where:

$$\xi = \ln\left(\frac{n_A}{n_B}\right) + \frac{1}{2N}\left[\ln\left(\frac{n_A}{n_1}\right) + \ln\left(\frac{n_A}{n_2}\right)\right] \tag{2.56}$$

By increasing N, $R_{\max}$ is increased and, at the same time, the high-reflectance band becomes narrower and narrower. In principle, as N tends towards infinity, $R_{\max}$ tends to 1. The greater is the ratio of n_A/n_B, the quicker $R_{\max}$ converges to its maximum value. For example, alternating zinc sulfide ($n_A = 2.32$) and magnesium fluoride ($n_B = 1.38$) layers, a reflectance of $R = 0.99$ can be reached using only 13 layers.

Similarly to the case of the anti-reflection coating, analogous calculation of the reflectance of the multilayer mirror can be performed for any incidence angle, provided the propagation path inside each layer is appropriately expressed. The dielectric multilayer mirror offers a great flexibility because it can be designed for any value of reflectance at any wavelength and any incidence angle. In particular, beam splitters having any arbitrary power partition between transmitted and reflected beam can be made.

The idea that the optical properties of a periodic structure can be considerably varied not by changing materials, but by simply acting on the geometry has gained more and more importance in the last decade. Two-dimensional and three-dimensional periodic structures made of dielectric or semiconductor materials, called photonic crystals, are increasingly utilized in photonic devices. It is interesting to note that this kind of phenomenon occurs also in Nature. For example, the brilliant colors of some butterflies are due not to absorption, but to the presence of periodic dielectric layers on their wings. A phenomenon typical of periodic structures is iridescence, due to the fact that the wavelength of maximum reflectance depends on the incidence angle. This pleasing visual effect, also called opalescence, can be observed with opals, semi-precious stones consisting of a three-dimensional periodical arrangement of sub-micrometric silica spheres.

2.2.5 Beam Splitter

As mentioned in the previous section, beam splitters are semi-reflecting mirrors that divide an optical beam into two parts propagating in different directions. In order to understand the behavior of some optical systems utilizing beam splitters, it is useful to know the phase relation between transmitted and reflected fields.

Let $\tau = \tau_o \exp(i\phi_\tau)$ and $\rho = \rho_o \exp(i\phi_\rho)$ be, respectively, the transmission and reflection coefficients for the electric field. The actual value taken by these coefficients depends on many parameters, such as the beam-splitter structure, the incidence angle, and the wavelength of the incident beam. On the other hand, as will be shown, the phase difference $\phi_\tau - \phi_\rho$ is expressed by a general law.

The scheme of Fig. 2.7 shows a beam-splitter symmetrically illuminated on both sides by two identical light beams having the same phase, polarization, incidence angle, and intensity I_o. Calling I_1 and I_2 the intensities of the two output beams and assuming no losses, it should be expected, for symmetry reason, that $I_1 = I_2 = I_o$. Since the field E_1 at output 1 is the sum of two contributions, one coming from reflection and the other from transmission, $E_1 = \rho E_o + \tau E_o$, the intensity at output 1 is given by: $I_1 = |\rho + \tau|^2 I_o$. The condition $I_1 = I_o$ is satisfied if:

$$|\rho + \tau|^2 = [\tau\tau^* + \rho\rho^* + 2Re(\tau\rho^*)] = 1. \tag{2.57}$$

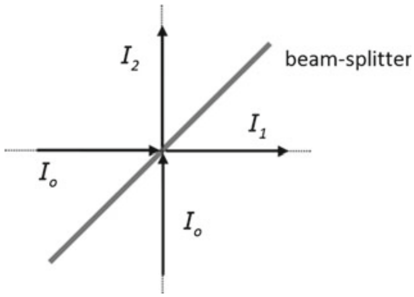


Fig. 2.7 Two identical light beams impinging on the two sides of a beam splitter

Since $\tau\tau^* + \rho\rho^* = 1$, the quantity $\text{Re}(\tau\rho^*) = \tau_o\rho_o\cos(\phi_\tau - \phi_\rho)$ must be nil, therefore:

$$\phi_\tau - \phi_\rho = \frac{\pi}{2}. \tag{2.58}$$

Equation (2.58) is a general relation, valid for any lossless beam-splitter, regardless of the structure of the mirror or the incidence angle. The only limitation is that the beam-splitter should be symmetrical. It is easy to generalize to the non-symmetrical case: calling ϕ_τ and ϕ_ρ the phase shifts for a beam incident to the left of the mirror, and ϕ'_τ and ϕ'_ρ the phase shifts for a beam incident to the right of the mirror, instead of (2.58), the following relation is found:

$$\phi_\tau + \phi'_\tau - \phi_\rho - \phi'_\rho = \pi. \tag{2.59}$$

2.2.6 Total-Internal-Reflection Prism

The total-internal-reflection glass prism, shown in Fig. 2.8, behaves as a 100 %-reflectance mirror if the incident beam undergoes total reflections at the two glass-air interfaces. This happens if the incidence angles at the glass-air boundaries are larger than the limit angle. In the case of the figure, $\theta_i = 45°$, which is larger than θ_c for a normal optical glass. The spurious reflection of about 4 % at the air-glass input interface can be eliminated by an anti-reflection coating.

Whereas multilayer dielectric mirrors present high reflectance in a narrow wavelength band, the total-internal-reflection mirror is a broadband component. The only limitations could come from the decrease of the refractive index of the glass as the wavelength is increased, which could move θ_c to a value larger than 45°.

An interesting property of the total-internal-reflection prism is that the reflected beam is always propagating in the opposite direction with respect to the incident beam, whatever is the angle of incidence, provided that the propagation vector of the beam lies in a plane perpendicular to the prism axis (this is the plane of Fig. 2.8). Such a property is not shared by normal mirrors. Instead of a prism, it is possible to use a trihedral (see Fig. 2.9), known as the corner cube. With this, the beam is retro-reflected after three internal reflections, and it is possible to obtain a back-propagating reflected beam for any direction of incidence. In practical applications,

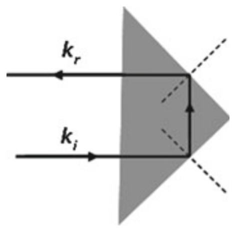


Fig. 2.8 Total internal reflection prism

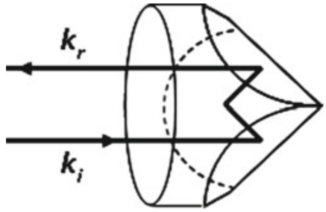


Fig. 2.9 Corner cube

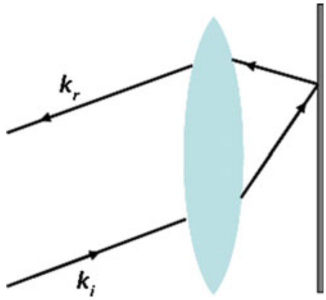


Fig. 2.10 Cat's eye

the back-reflectors placed on vehicles or on wayside posts have a structure imitating an array of corner cubes.

This same property of the corner cube is also shared by the combination of a lens plus a mirror placed in the focal plane of the lens, known as a cat's eye (see Fig. 2.10).

2.2.7 Evanescent Wave

A linearly polarized beam impinges on the glass-air planar boundary with an incidence angle $\theta_i \geq \theta_c$. Assume that the boundary coincides with the x-y plane, that the propagation vector $\mathbf{k_1}$ of the incident beam lies in the y-z plane, and that the beam is polarized along x as shown in Fig. 2.11a. Using the subscript 1 for glass and 2 for air, the electric fields of the incident and transmitted waves can be written as:

$$\begin{aligned}\mathbf{E_i} &= E_i \exp[-i(\omega t - k_{1y}y - k_{1z}z)]\mathbf{x}\\ \mathbf{E_t} &= E_t \exp[-i(\omega t - k_{2y}y - k_{2z}z)]\mathbf{x},\end{aligned} \tag{2.60}$$

where $k_{1y} = -k_o n_1 \sin\theta_i$ and $k_{1z} = k_o n_1 \cos\theta_i$.

The continuity condition at the boundary $z = 0$ imposes $k_{1y} = k_{2y}$. Using the relation:

$$k_2 = k_o n_2 = \sqrt{{k_{2y}}^2 + {k_{2z}}^2}, \tag{2.61}$$

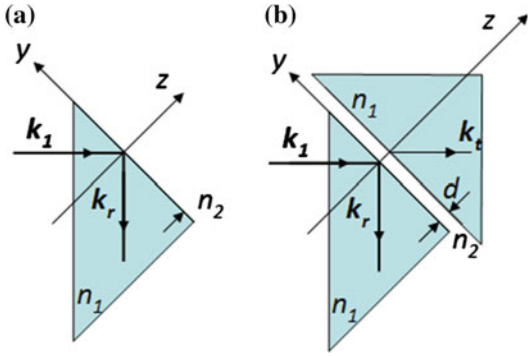


Fig. 2.11 **a** Geometry for total internal reflection. **b** Transmission of the evanescent wave by a pair of prisms

it is found: $k_{2z} = k_o\sqrt{{n_2}^2 - {n_1}^2\sin^2\theta_i}$. If $\theta_i > \theta_c$, the term under square root is negative, so that k_{2z} becomes imaginary. Putting $\sqrt{{n_2}^2 - {n_1}^2\sin^2\theta_i} = i\Gamma n_2$, where Γ is a real quantity, it is found:

$$\mathbf{E_t} = E_t\exp[-i(\omega t + k_o n_1\sin\theta_i y)]\exp(-\Gamma k_o n_2 z)\mathbf{x}. \tag{2.62}$$

The condition of total internal reflection at the boundary between glass and air does not imply that there is no electromagnetic field in the air beyond the glass. Indeed Eq. (2.62) describes a wave, known as evanescent wave, running parallel to the glass-air interface, with an amplitude that is an exponentially decreasing function of the penetration distance z. The electromagnetic energy density inside medium 2, being proportional to $|E_t|^2$, is also exponentially decreasing with z, becoming smaller by a factor e^{-1} at the distance:

$$\delta = \frac{1}{2\Gamma k_o n_2} = \frac{\lambda}{4\pi\sqrt{{n_1}^2\sin^2\theta_i - {n_2}^2}}. \tag{2.63}$$

For instance, if $n_1 = 1.5$, $n_2 = 1$, $\theta_i = \pi/4$, it follows that $\Gamma = 0.354$, therefore $\delta = \lambda/4.44$. Taking the central wavelength of the visible range, $\lambda = 0.55$ μm, it is found that $\delta = 120$ nm.

In order to derive the amplitude of the transmitted and reflected fields, the following boundary conditions are written:

$$E_i + E_r = E_t \tag{2.64}$$

$$E_i - E_r = i\gamma E_t, \tag{2.65}$$

where $\gamma = \Gamma n_2/(n_1\cos\theta_i)$ is a real quantity. The solutions of (2.64) and (2.65) are:

$$E_t = \frac{2}{1+i\gamma}E_i\ ; \quad E_r = \frac{1-i\gamma}{1+i\gamma}E_i. \tag{2.66}$$

Note that $|E_i| = |E_r|$, thus confirming total reflectance, even if $|E_t|$ is different from 0.

By placing a second prism in the vicinity of the totally reflecting one, it is possible to exploit a tunnel effect, by "catching" some part of the evanescent wave. Such a technique, called frustrated total internal reflection, allows for the realization of a partially transmitting mirror, with a transmittance controlled by the thickness of the air-gap between the two prisms. Consider the configuration in Fig. 2.11b, where there is an air-gap of thickness d between the two prisms. In order to write down the boundary conditions at the two interfaces, it is necessary to assume that inside the gap there are two evanescent waves decaying in opposite directions. The continuity conditions at the first interface are:

$$E_i + E_r = E_t' + E_r', \tag{2.67}$$

and

$$E_i - E_r = i\gamma(E_t' - E_r'), \tag{2.68}$$

where E_t' and E_r' are the fields of the two evanescent waves, the first decaying as a function of z and the second growing as a function of z. Similarly, the continuity conditions at the second interface are:

$$\exp[-d/(2\delta)]E_t' + \exp[d/(2\delta)]E_r' = E_t, \tag{2.69}$$

and

$$\exp[-d/(2\delta)]E_t' - \exp[d/(2\delta)]E_r' = -(i/\gamma)E_t. \tag{2.70}$$

By eliminating E_t' and E_r', E_t and E_r can be calculated. In particular, the transmittance T is found to be:

$$T = \frac{|E_t|^2}{|E_i|^2} = \frac{\exp(-d/\delta)}{\left[\Big(1 + \exp(-d/\delta)\Big)/2\right]^2 + A\Big(1 - \exp(-d/\delta)\Big)^2}, \tag{2.71}$$

where $A = (1 - \gamma^2)^2/(16\gamma^2)$. It is seen from (2.71) that, for d tending to 0, T tends to 1, as expected. If $d \gg \delta$, T is approximated by:

$$T \approx \frac{4\exp(-d/\delta)}{1 + 4A}, \tag{2.72}$$

which shows an exponential decay of T as a function of d. The set of two prisms is therefore equivalent to a variable-transmittance attenuator.

2.2.8 Thin Lens and Spherical Mirror

So far in this chapter only situations involving plane waves and plane boundaries have been considered. In this section spherical boundaries are discussed, concentrating the attention on the important optical component of the lens.

The lens, as shown in Fig. 2.12, is formed by joining two spherical caps made of glass. The radii of curvature of the two surfaces, ρ_1 and ρ_2, are conventionally taken to be positive if the surfaces are convex. The thin-lens approximation consists in assuming that the lens thickness is much smaller than the radii of curvature of the two spherical surfaces. Neglecting reflections at the two surfaces, the effect of the lens on the incident field E_1 is described by a transmission function $\tau_f(x, y)$:

$$E_2(x, y, z, t) = \tau_f(x, y) E_1(x, y, z, t). \tag{2.73}$$

This transmission function has a unitary modulus and a phase that is locally dependent on the glass thickness:

$$\tau_f(x, y) = \exp\{ik_o[n\Delta(x, y) + \Delta_o - \Delta(x, y)]\}, \tag{2.74}$$

where n is the glass refractive index, $\Delta(x, y)$ is the lens thickness at the transversal position (x, y), and $\Delta_o = \Delta(0, 0)$ is the thickness on the optical axis.

The phase shift is the sum of two contributions, one due to the path inside the glass, $k_o n\Delta(x, y)$, and the other due to the path in air, $k_o[\Delta_o - \Delta(x, y)]$. The lens thickness at the generic point (x, y) is:

$$\Delta(x, y) = \Delta_o + \sqrt{\rho_1{}^2 - (x^2 + y^2)} - \rho_1 + \sqrt{\rho_2{}^2 - (x^2 + y^2)} - \rho_2 \tag{2.75}$$

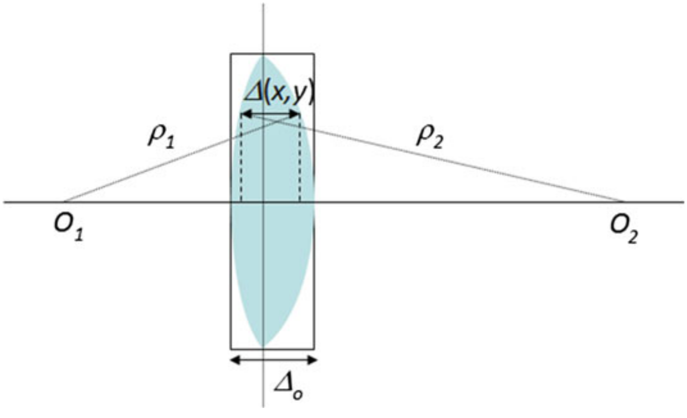


Fig. 2.12 Thin lens made of two glass caps with radii of curvature ρ_1 and ρ_2

Equation (2.75) can be simplified by assuming that the lens diameter is much smaller than the radii of curvature of the two surfaces, so that:

$$\sqrt{{\rho_1}^2-(x^2+y^2)}=\rho_1\sqrt{1-\frac{x^2+y^2}{{\rho_1}^2}}\approx\rho_1\left[1-\frac{x^2+y^2}{2{\rho_1}^2}\right], \tag{2.76}$$

and

$$\sqrt{{\rho_2}^2-(x^2+y^2)}\approx\rho_2\left[1-\frac{x^2+y^2}{2{\rho_2}^2}\right]. \tag{2.77}$$

In the power expansion of the square-root terms, only the first-order terms in (x^2+y^2) are kept. This is equivalent to approximate the spherical surface with a paraboloidal surface, similarly to the paraxial approach used in Chap. 1 for the description of spherical waves.

Inserting (2.76) and (2.77) into (2.75):

$$\varDelta(x,y)=\varDelta_o-\frac{x^2+y^2}{2}\left(\frac{1}{\rho_1}+\frac{1}{\rho_2}\right). \tag{2.78}$$

By substituting inside (2.74):

$$\tau_f(x,y)=\exp(ik_o n\varDelta_o)\exp\left[-ik_o\frac{x^2+y^2}{2f}\right], \tag{2.79}$$

where f, called focal length, is defined by the relation:

$$\frac{1}{f}=(n-1)\left(\frac{1}{\rho_1}+\frac{1}{\rho_2}\right) \tag{2.80}$$

Note that f has a sign: $f>0$ means converging lens, $f<0$ diverging lens. Since f depends on n, and, in turn, n changes with wavelength, the lens focuses different wavelengths to different distances, that is, it suffers from a chromatic aberration.

The spherical mirror is described similarly to the lens, the main difference being that it works in reflection instead of transmission. The mirror shown in Fig. 2.13 has center in O and radius of curvature ρ. Consider a beam parallel to the optical axis, incident at point P with an incidence angle α. The reflected beam intersects the optical axis at point F. The distance of F from O is: $\overline{OF}=\overline{OP}/(2\cos\alpha)$. Since $\overline{OF}$ depends on α, parallel beams impinging on the mirror in different points produce reflected beams crossing the optical axis at different positions. However, if only beams traveling close to the optical axis are considered (i.e. where $\cos\alpha\approx 1$), it is found that $\overline{OF}\approx\overline{OP}/2=\rho/2$. Under this approximation, which is equivalent to treating the surface as paraboloidal instead of spherical, it can be said that the spherical mirror brings all parallel beams to converge at a single point a distance $\rho/2$ from the mirror surface. In other words, the mirror behaves in reflection just as a lens

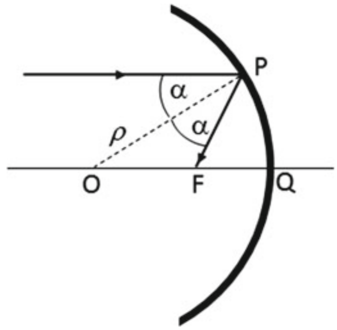


Fig. 2.13 Spherical mirror with radius of curvature ρ

having a focal length $f = \rho/2$. Note that the focal length of the spherical mirror is independent of the wavelength. All large astronomical telescopes use parabolic mirrors instead of lenses.

2.2.9 Focusing Spherical Waves

A spherical wave traveling along the z axis impinges on a thin lens located at the plane $z = 0$. If R_1 is the radius of curvature of the wavefront, then, under the paraxial approximation, the electric field E_1 of the wave at $z = 0$ is given by:

$$E_1(x, y, 0) = \frac{A_o}{R_1} \exp\left(ik_o \frac{x^2 + y^2}{2R_1}\right). \tag{2.81}$$

The field distribution transmitted through the lens, $E_2 = \tau_f E_1$, calculated by using (2.79) and (2.81), is again that of a spherical wave, with the modified radius of curvature R_2, which is related to R_1 by:

$$\frac{1}{R_2} = \frac{1}{R_1} - \frac{1}{f}. \tag{2.82}$$

If $R_1 = f$, R_2 becomes infinite, which means that a spherical wave originating from the left focus is converted into a parallel beam (plane wave). If $R_1 > f$, $R_2 < 0$, indicating that the diverging wave is converted into a converging wave. If $R_1 \gg f$, $R_2 \approx f$, that is, a plane wave is converted into a spherical wave converging at the right focal plane. In practice, the wave will not converge to a point, but the beam spot at focus will be described by an Airy function (see Sect. 1.6.1) because of the diffraction effect introduced by the finite dimension of the lens-diameter.

The ratio between the lens radius and the focal length is called numerical aperture (NA) of the lens. It describes the light-gathering capacity of the lens.

Relation (2.82) can also be interpreted in the following way: given a point-like object P at distance R_1 on the left, the lens creates a real image Q of P at distance R_2 on the right. This is the standard interpretation given by geometrical optics. If P is on the optical axis, Q is also on the same axis. What happens if the point-like object P′ is off-axis? Assigning to P′ the coordinates $(x_o, 0, R_1)$, the electric field of the spherical wave originated at P′, evaluated at the lens plane $z = 0$, is now:

$$E_1(x, y, 0) = \frac{A_o}{R_1}\exp\left\{\frac{ik_o}{2R_1}[(x - x_o)^2 + y^2]\right\}. \tag{2.83}$$

Using again the relation $E_2 = \tau_f E_1$, after some algebra the following expression is found for E_2:

$$E_2(x, y, 0) = \frac{A_o\exp(i\phi)}{R_1}\exp\left\{\frac{ik_o}{2R_2}\left[(x + \frac{f}{R_1 - f}x_o)^2 + y^2\right]\right\}, \tag{2.84}$$

where ϕ depends on x_o and on the lens parameters, but not on x and y. Equation (2.84) shows that the image of P′ is at point Q′ having coordinates $[-f x_o/(R_1 - f), 0, R_2]$. It can be said that the lens creates an image QQ′ of the segment PP′. If $R_1 > f$, the real image is upside-down, with a magnification $f/(R_1 - f)$. This approach is therefore equivalent to the geometrical-optics treatment. The topic of image formation has important applications but will not be further discussed in this text.

Another case to be considered is that of a plane wave having a propagation direction slightly tilted with respect to the optical axis. To simplify the discussion it is assumed that the wave vector **k** is in the plane x-z, forming an angle α with x. At $z = 0$ the incident field is expressed as $E_1 = A_o\exp(ik_o\alpha x)$, where $\sin\alpha$ has been approximated as α. After using $E_2 = \tau_f E_1$ again, and applying some algebra, it is found that:

$$E_2(x, y, 0) = A_o\exp(i\phi)\exp\left\{-\frac{ik_o}{2f}\left[(x - \alpha f)^2 + y^2\right]\right\} \tag{2.85}$$

where $\phi = k\alpha^2 f/2$. Equation (2.85) describes a spherical wave that converges at the point $(\alpha f, 0, f)$ located on the right focal plane, as shown in Fig. 2.14.

Generalizing the treatment, it can be stated that a biunique correspondence is established between the propagation direction of the incident wave and points on the right focal plane.

2.2.10 *Focusing Gaussian Spherical Waves*

If the wave impinging on the lens is a Gaussian spherical wave with a complex radius of curvature of the wavefront q_1, then it can be immediately shown, by using (2.73) and (2.79), that the transmitted wave is still a Gaussian spherical wave. The new

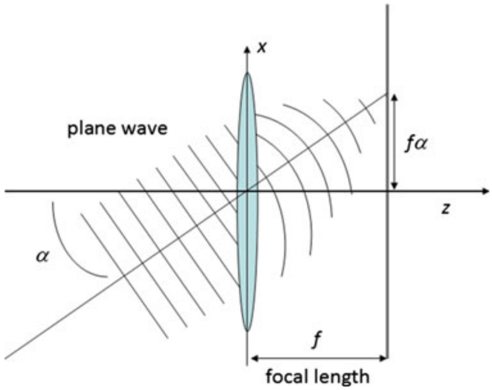


Fig. 2.14 A plane wave corresponds to a point at the focal plane of a lens

complex radius of curvature q_2 satisfies the relation:

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f}. \tag{2.86}$$

Since the beam radius is unchanged by crossing the lens, $w_1 = w_2$, one finds that (2.86) coincides with (2.82). As an example, if $f > 0$ and E_1 describes a diverging Gaussian wave with $R_1 > f$, the transmitted Gaussian wave has a negative radius of curvature of the wavefront, $R_2 < 0$, and is, therefore, converging. There is a basic difference between this case and the one treated in the previous section: the theory of Gaussian waves accounts intrinsically for diffraction effects, such that the wave converges to a finite-size beam spot.

Once q_2 at $z = 0$ is assigned, it is possible to obtain the distance z_f at which the focused beam takes the minimum radius, and also the minimum radius w_f, by using relations 1.35 and 1.36. The results of the calculation are:

$$z_f = \frac{|R_2|}{1 + [\lambda |R_2|/(\pi w_1^2)]^2}, \quad w_f = \frac{w_1}{\sqrt{1 + [\pi w_1^2/(\lambda |R_2|)]^2}}. \tag{2.87}$$

Assuming that $R_1 \gg f$, which means $|R_2| \approx f$, and $\pi w_1^2 \gg \lambda f$, the relations of (2.87) simplify as follows:

$$z_f \approx f, \quad w_f \approx \frac{\lambda f}{\pi w_1} \tag{2.88}$$

It is useful to calculate the Rayleigh length at focus:

$$z_{fR} = \frac{\pi w_f^2}{\lambda} \approx \frac{\lambda}{\pi} \left(\frac{f}{w_1} \right)^2. \tag{2.89}$$

The length $2z_{fR}$ defines the field depth, that is, the interval of distance over which the beam remains focused. As an example, putting $f = 20$ cm, $w_1 = 2$ mm, $\lambda = 0.5\,\mu$m (note that $\pi w_1^2/(\lambda f)$ is $\gg 1$), one finds that $w_f = 16\,\mu$m and $z_{fR} = 1.6$ mm.

From a mathematical point of view, the Gaussian spherical wave has an infinite extension on the transversal plane. In reality, practical optical components have a limited size, and so the tails of the Gaussian distribution are lost. The empirical criterion for applying the theory of Gaussian waves is if the lost power is sufficiently small. Considering a lens (or a mirror) with diameter d, the power fraction intercepted by the lens is given by:

$$T_f = \frac{\int_0^{d/2} 2\pi r \, \exp(-\frac{2r^2}{w^2})dr}{\int_0^{\infty} 2\pi r \, \exp(-\frac{2r^2}{w^2})dr} = 1 - \exp\left(-\frac{d^2}{2w^2}\right) \tag{2.90}$$

As an example, if $d = 2w$, the intercepted power fraction is 86 %, if $d = 3w$, $T_f = 0.99$. Therefore, if $d > 3w$, it is reasonable to assume the applicability of the Gaussian wave theory.

2.2.11 Matrix Optics

Within the geometrical optics approximation, a useful approach is that of associating a 2×2 matrix, called $ABCD$ matrix, with each optical component.

The ray present at the coordinate z is characterized by two parameters, the distance $r(z)$ from the optical axis, and the angle its direction forms with the optical axis, $\phi(z)$. Within the paraxial approximation, $\phi(z)$ is simply the derivative of r with respect to z. The ray is represented by a one-column matrix:

$$\mathbf{R} = \begin{pmatrix} r \\ \phi \end{pmatrix}. \tag{2.91}$$

Every optical component is represented by a 2×2 matrix of the type:

$$\mathbf{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{2.92}$$

connecting the output vector $\mathbf{R_u}$ to the input vector $\mathbf{R_i}$ according to:

$$\mathbf{R_u} = \mathbf{M R_i}. \tag{2.93}$$

Using refraction laws and the properties of lenses and mirrors discussed in the preceding sections, the $ABCD$ matrices of the components shown in Fig. 2.15 can be written as follows.

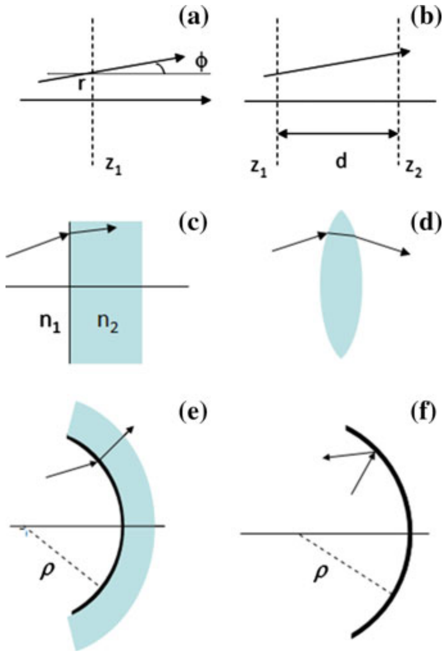


Fig. 2.15 Different optical components that are described by ABCD matrices: **a** generic ray, **b** free propagation, **c** planar boundary between two different media, **d** lens, **e** spherical boundary between two media, **f** spherical mirror

- Free propagation

$$\mathbf{M} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \tag{2.94}$$

- Lens

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ \frac{-1}{f} & 1 \end{pmatrix} \tag{2.95}$$

- Planar boundary between two media

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix} \tag{2.96}$$

- Spherical boundary between two media

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2} \frac{1}{\rho} & \frac{n_1}{n_2} \end{pmatrix} \tag{2.97}$$

- Spherical mirror

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ \frac{-2}{\rho} & 1 \end{pmatrix} \tag{2.98}$$

Note that the determinant of a generic $ABCD$ matrix is given by:

$$AD - BC = n_1/n_2. \tag{2.99}$$

If the two media are the same, the determinant is equal to 1, as expected if the optical element does not introduce losses.

An optical system made by a cascade of various elements is represented by a matrix that is the ordered product of all the individual matrices describing each element. The order is important because the matrix product is generally non-commutative.

2.3 Fourier Optics

As discussed in Sect. 1.6, the Fraunhofer approximation of the diffraction integral is only valid in the far field, at distances from the screen that are much larger than the typical size of optical instruments. Therefore, it is of great practical importance to find a method of transporting the far-field behavior to short distances. In this section it will be shown that the insertion of a lens can solve the problem.

Consider a thin lens having a focal distance f and positioned at $z = 0$, that is illuminated from the left by a wave propagating along the z axis. By assuming that the field distribution in the left focal plane of the lens is known, $E_i(x_o, y_o, -f)$, the aim of the calculation is to derive the field distribution at the right focal plane, $E_f(x_f, y_f, f)$. The calculation is performed according to the following scheme: (a) the wave is propagated from the plane $z = -f$ to the plane $z = 0$ by using Fresnel diffraction, as expressed by (1.50); (b) at $z = 0$ the field distribution is multiplied by the lens transmission function $\tau_f(x, y)$; (c) the obtained field distribution is propagated to the plane $z = f$ by using again Fresnel diffraction.

Step (a) of the calculation gives:

$$E(x, y, 0) = \frac{i \exp(ikf)}{\lambda f} \int_{-\infty}^{\infty} dx_o$$
$$\int_{-\infty}^{\infty} E_i(x_o, y_o, -f) \exp\left\{\frac{ik}{2f}[(x - x_o)^2 + (y - y_o)^2]\right\} dy_o. \tag{2.100}$$

Steps (b) and (c) require calculating the integral:

$$E_f(x_f, y_f, f) = \frac{i \exp(ikf)}{\lambda f} \int_{-\infty}^{\infty} dx$$
$$\int_{-\infty}^{\infty} E(x, y, 0)\tau_f(x, y) \exp\left\{\frac{ik}{2f}[(x - x_f)^2 + (y - y_f)^2]\right\} dy. \tag{2.101}$$

By inserting (2.100), and the expression of τ_f given by (2.79), into (2.101), one obtains:

$$E_f(x_f, y_f, f) = -\frac{\exp[2ikf + ik(n-1)\Delta_o]}{\lambda^2 f^2} \int_{-\infty}^{\infty} dx_o \int_{-\infty}^{\infty} dy_o E_i(x_o, y_o, -f) \int_{-\infty}^{\infty} dx$$
$$\int_{-\infty}^{\infty} dy \exp\left\{\frac{ik}{2f}[(x-x_f)^2 + (y-y_f)^2 + (x-x_o)^2 + (y-y_o)^2 - x^2 - y^2]\right\}. \quad (2.102)$$

The $dxdy$ integral is solved by writing the integrand function as the exponential of a square term. Noting that: $(x-x_f)^2+(y-y_f)^2+(x-x_o)^2+(y-y_o)^2-x^2-y^2 = (x-x_o-x_f)^2 + (y-y_o-y_f)^2 - 2(x_o x_f + y_o y_f)$, it is finally found:

$$E_f(x_f, y_f, f) = \frac{\exp[2ikf + ik(n-1)\Delta_o]}{\lambda f}$$
$$\int_{-\infty}^{\infty} dx_o \int_{-\infty}^{\infty} dy_o E_i(x_o, y_o, -f) \exp\left[\frac{ik}{f}(x_o x_f + y_o y_f)\right]. \quad (2.103)$$

Equation (2.103) represents a very important result: it states that the field distribution at the right-hand focal plane of the lens is the two-dimensional Fourier transform of the distribution at the left-hand plane. Such a result is consistent with the consideration, developed at the end of Sect. 2.2.9, that a biunique correspondence exists between the propagation direction of the incident wave and points on the right focal plane. In fact to perform the spatial Fourier transform of the field distribution is the same as describing the field by a superposition of plane waves.

Equation (2.103) has a fundamental role in many optical processing techniques. For instance, it suggests that spatial filtering of optical signals can be performed by placing an appropriate filter at the right-hand focal plane, and reconstructing the filtered signal with a second confocal lens. An example is shown in Fig. 2.16, where a pinhole is used as a low-pass filter, i.e. a filter transmitting only low spatial frequencies.

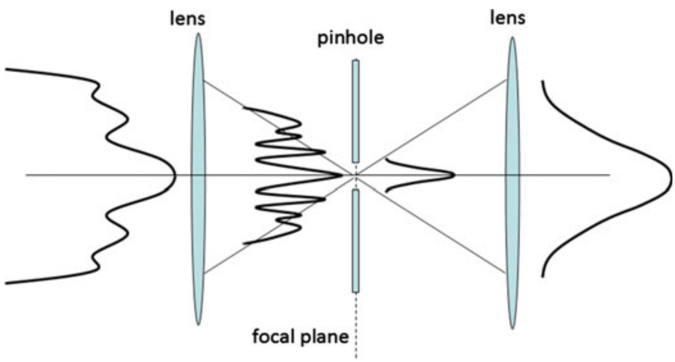


Fig. 2.16 Scheme of spatial filtering of an optical beam

2.4 Spectral Analysis

The electric field $E(t)$ associated with a generic light beam can always be described, as discussed in Sect. 2.1, as a weighted superposition of sinusoidal functions oscillating at different frequencies. The Fourier transform of $E(t)$ is the complex quantity $E(\omega)$ (given by (2.10)) that represents the amplitude and the phase of the field component at angular frequency ω.

The power spectrum of the light beam, $S(\omega)$, is defined as:

$$S(\omega) = c\frac{\varepsilon_o |E(\omega)|^2}{2} \tag{2.104}$$

It should be noted that $S(\omega)$ does not contain any information on the phase of $E(\omega)$. This means that the knowledge of $S(\omega)$ is not sufficient, in general, for reconstructing the temporal behavior of $E(t)$.

The usual method to measure the power spectrum of an optical signal is that of collimating a sample beam and sending it through a dispersive system in order to angularly separate the different frequency components. This method was introduced by Newton, who first observed the spectrum of solar light through a glass prism. Instead of using the dispersion properties of glass, angular separation can be generated by diffraction gratings, i.e. structures with periodic transmission functions. Almost all modern commercial spectrometers are based on gratings. As will be seen, there is another method, consisting of the use of a narrow-band tunable filter. This method is conceptually similar to that used to measure the power spectrum of electrical signals. In the optical case, the narrow-band filter is a Fabry-Perot interferometer.

An important parameter for assessing the performance of a spectrometer is the frequency resolution $\delta\nu$, defined as the minimum frequency separation that can be resolved by the spectrometer. Calling ν_o the central frequency of the signal to be analyzed, the resolving power of the instrument at that frequency is calculated as:

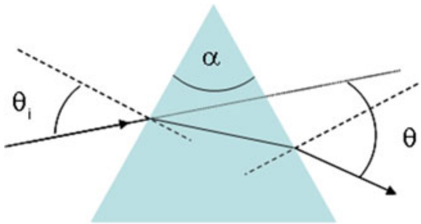
$$P_r = \frac{\nu_o}{\delta\nu} = \frac{\lambda_o}{\delta\lambda}, \tag{2.105}$$

where $\lambda_o = c/\nu_o$ and $\delta\lambda$ is the minimum wavelength separation that can be resolved. When writing the right-hand member of (2.105), the assumption $\delta\nu \ll \nu_o$ has been made. In fact, by differentiating $\nu\lambda = c$ around ν_o, it is obtained: $\lambda_o d\nu - \nu_o d\lambda = 0$, hence (2.105).

2.4.1 Dispersive Prism

Optical components fabricated from dispersive materials refract light of different wavelengths by different angles. Consider a collimated light beam that impinges on a glass prism at an incidence angle θ_i, as shown in Fig. 2.17. Since the angle

Fig. 2.17 Dispersive prism



of refraction depends on the refractive index, which is wavelength dependent, the direction of the emergent beam is also wavelength dependent.

The deflection angle θ, calculated from the refraction law, is given by:

$$\theta = \theta_i - \alpha + \arcsin\left[\sin\alpha\sqrt{n^2 - \sin^2\theta_i} - \sin\theta_i\cos\alpha\right], \tag{2.106}$$

where α is the prism angle and n the refractive index of the glass.

What's the minimum wavelength separation $\delta\lambda$ that can be resolved? Clearly, this depends on the minimum angular resolution. Even assuming that the incident beam is monochromatic and perfectly collimated, the output beam will present an angular spread due to diffraction. If L is the minimum size between prism size and incident beam diameter, and λ_o is the wavelength of the incident beam, the angular spread will be $\delta\theta \approx \lambda_o/L$. In order to resolve two different wavelengths, the difference in their deflection angles should be larger than $\delta\theta$. Therefore, the wavelength resolution is given by:

$$\delta\lambda = \left(\frac{d\theta}{d\lambda}\right)^{-1}\delta\theta, \tag{2.107}$$

where $d\theta/d\lambda$ is calculated by assigning the dependence of n on λ. Taking typical values for the parameters appearing inside (2.106), $\alpha = \theta_i = \pi/4$, $n = 1.5$, it is found that $d\theta/d\lambda \approx dn/d\lambda$. Therefore, the resolving power of the dispersive prism can be written as:

$$P_r = \frac{\lambda_o}{\delta\lambda} \approx L\frac{dn}{d\lambda} \tag{2.108}$$

If $dn/d\lambda = (10\ \mu\text{m})^{-1}$, and $L = 1$ cm, then $P_r = 10^3$.

2.4.2 *Transmission Grating*

The diffraction grating is an optical component that periodically modulates (in the transversal direction) the amplitude or phase of an incident wave. The treatment here is limited to linear gratings in which the diffracting elements consist of parallel lines ruled on glass or metal. These gratings may be used either in transmission or in reflection. The most common transmission grating is made of a periodic array of narrow slits on an opaque screen. The slit width is smaller than the wavelength, so that the diffracted field has a wide angular spread.

Assume that the grating plane coincides with the x-y plane, and that the slits are parallel to the x axis and have spacing d. The wave vector of the incident plane wave lies in the y-z plane and forms an angle θ_i with the z axis, as shown in Fig. 2.18. The spatial distribution of the field in the half-space beyond the grating is determined by the interference among the diffracted waves coming from all the slits. Note that in the approximation of a far-field projection the interfering beams can be considered as parallel. Diffractive interference will occur only in those directions satisfying the condition:

$$q\lambda = d(\sin\theta_q - \sin\theta_i), \tag{2.109}$$

where $q = 0, \pm1, \pm2, \ldots$. Equation (2.109) defines a discrete set of directions in which a diffracted beam can be observed.

In particular, the beam corresponding to the order of $q = 0$ is simply the transmitted beam: $\sin\theta_0 = \sin\theta_i$.

Equation (2.109) says that the directions of the diffraction orders (except for the order 0) are wavelength-dependent. Therefore the grating is a dispersive structure. As such it can be used, like the dispersive prism, to analyze the power spectrum of an optical signal.

Similarly to the case of the prism, the limit to the wavelength resolution comes from the diffraction effect due to the finite size of the grating. At $\theta_i = 0$ the first-order diffracted field, as derived from the theory of Fraunhofer diffraction, is expressed by (1.69). This equation shows that, in the far-field plane at distance z from the grating plane, the peak corresponding to the diffracted field is transversally displaced by

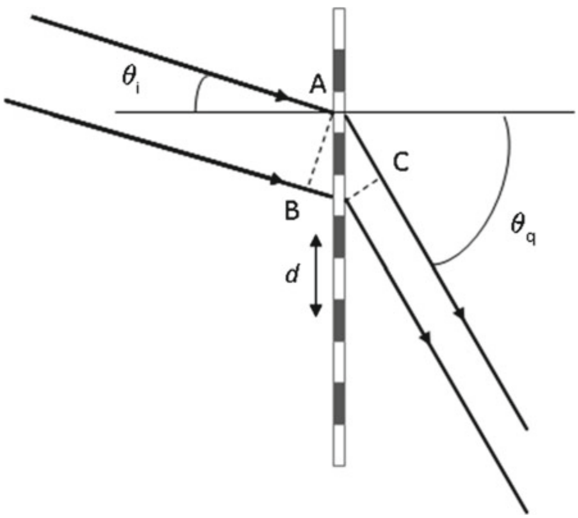


Fig. 2.18 Transmission grating. The path difference between the two diffracted beams is: $\overline{AC} - \overline{BH} = d\sin\theta_q - d\sin\theta_i$

$\lambda z/d$ and has a width $\lambda z/L$. This means that the resolving power of the grating is:

$$P_r = \frac{\lambda z/d}{\lambda z/L} = \frac{L}{d} = N, \tag{2.110}$$

where L is the size of the grating, and N the number of illuminated slits. Note that N coincides with L/d only if the collimated beam fully covers the grating surface. By introducing the number of slits per unit length, $f_o = 1/d$, $N = f_o L$. Taking, e.g., $f_o = 10^3\ \text{mm}^{-1}$ and $L = 1$ cm, the resolving power is 10^4, one order of magnitude larger than that of the dispersive prism. As shown by (2.109), the angular separation between two wavelengths grows by increasing the diffraction order: for the order q, the resolving power becomes $P_r = qN$. However, the fraction of diffracted optical power in a given order decreases with the value of q.

A problem arising from the use of a grating in spectral analysis is an ambiguity that can appear if the optical signal has a very broad wavelength spectrum. Indeed, considering (2.109), it can be seen that the first-order diffraction of the wavelength λ_1 appears in the same direction as the second-order diffraction of the wavelength $\lambda_1/2$. This means that there is an upper limit to the spectral width that can be unambiguously analyzed by the grating. For very broad spectra it is more convenient to use a dispersing prism.

Instead of an amplitude grating, a phase grating can be used: this is discussed in Sect. 1.6.2. Most of the properties of the amplitude grating are also shared by the phase grating.

The description of the diffraction properties of gratings assumes that the diffracted field is observed in the far field. Since, in order to have a good resolving power, the grating size must be of the order of centimeters, far field means a distance of kilometers, as discussed in Sect. 1.6. However, by exploiting the properties of lenses (or spherical mirrors) discussed in Sect. 2.3, the far field can be brought to the focal plane of a lens placed beyond the grating: this is the approach adopted by all practical spectrometers.

2.4.3 *Reflection Grating*

Reflection gratings are made by creating a linear alternation of reflecting and opaque zones on a surface, as shown in Fig. 2.19.

In the case of the reflection grating, the diffractive effects are observed in the backward direction. The directions of constructive interference are given by:

$$q\lambda = d(\sin\theta_i + \sin\theta_q), \tag{2.111}$$

where q is $0, \pm 1, \pm 2, \ldots$ and θ_i is the incidence angle. Inside (2.111) diffraction angles are positive if the corresponding directions stay in the upper half-plane with

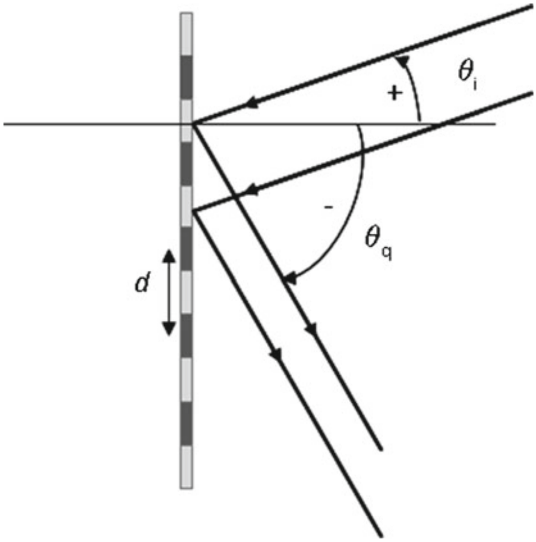


Fig. 2.19 Reflection grating. The *black* zones are opaque, and the *grey* zones are reflecting

respect to the grating plane normal. The zeroth-order beam propagates at the angle $\theta_0 = -\theta_i$, corresponding to specular reflection, that is independent from λ.

In a transmission grating the diffraction lobe due to a single slit presents maximum intensity in correspondence with the direction of incidence. As a consequence, the largest fraction of diffracted power goes to order 0. This is useless from the point of view of spectral analysis because the zeroth order is not dispersive. Similarly, in the case of a reflection grating the largest fraction of diffracted power would appear in specular reflection. Diffraction gratings can be optimized such that most of the power goes into a certain diffraction order, $q_0 \neq 0$, leading to a high diffraction efficiency for that order. This is indeed possible by using a so-called blazed grating (see Fig. 2.20), in which the reflecting elements are tilted by an angle θ_g with respect to the grating plane, obtaining a sawtooth-like profile.

The specular reflection from the tilted elements appears in a direction forming an angle $\theta_r = -(\theta_i - \theta_g) + \theta_g = (2\theta_g - \theta_i)$ with the normal to the grating plane. It is intuitive to expect that the maximum of the diffraction lobe associated to a single reflecting element occurs in the direction of specular reflection relative to that segment. By imposing that such direction coincides with that of the diffraction order q_0, $\theta_{q_o} = \theta_r$, the following relation is found:

$$\theta_g = \frac{\theta_i + \theta_{q_o}}{2}. \tag{2.112}$$

By substituting (2.112) inside (2.111):

$$q_o\lambda = d[\sin\theta_i + \sin(2\theta_g - \theta_i)]. \tag{2.113}$$

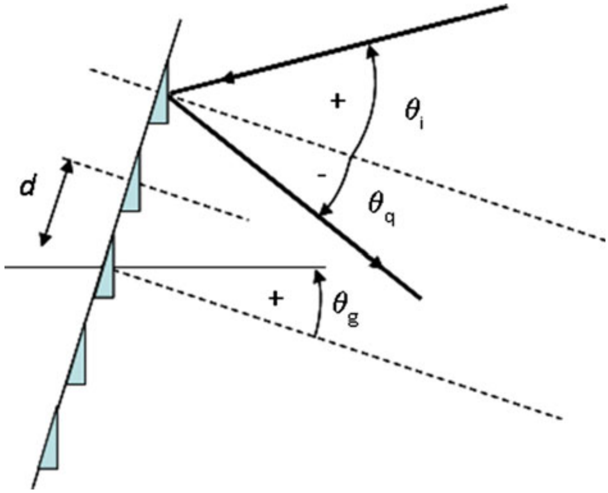


Fig. 2.20 Blazed reflection grating

Once λ and θ_i are fixed, (2.113) prescribes which is the value of θ_g to be chosen in order to maximize the power diffracted at order q_o.

There are two particularly interesting cases. The first corresponds to the Littrow configuration, in which the incident beam arrives normally to the tilted elements, $\theta_i = \theta_g$, so that (2.113) becomes:

$$q_o \lambda = 2d \sin \theta_g. \tag{2.114}$$

In this case $\theta_{q_o} = \theta_i$, which means that the order q_o propagates in the opposite direction with respect to the incident beam. The grating behaves as a wavelength-selective mirror, because it reflects only the wavelength that satisfies (2.114). As an example, consider a grating with 1200 lines per millimeter used in the Littrow configuration at the first order ($q_o = 1$) for $\lambda = 600$ nm. Equation (2.114) gives $\theta_g \approx 21°$. The back-reflected wavelength can be tuned by slightly rotating the grating, in order to change the incidence angle.

The second interesting case is that of normal incidence. If $\theta_i = 0$, $\theta_g = -\theta_{q_o}/2$. The condition of constructive interference is: $q_o \lambda = d \sin(2\theta_g)$. This is the configuration utilized in standard spectrometers.

2.4.4 Fabry-Perot Interferometer

The Fabry-Perot interferometer is a frequency filter, made by two parallel mirrors, as shown in Fig. 2.21.

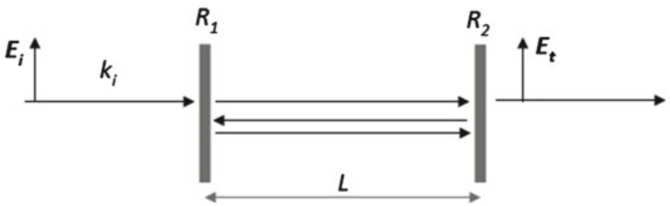


Fig. 2.21 Fabry-Perot interferometer

Considering a monochromatic plane wave at normal incidence, the aim is to calculate the transmittance of the interferometer, T_{FP}, as a function of the frequency of the wave. If R_1 and R_2 are the reflectances of the two mirrors, one might naively guess that T_{FP} is given by the product $(1 - R_1)(1 - R_2)$. This is wrong, because the field of the transmitted wave needs to be calculated by taking into account the interference process arising by the superposition of the multiple reflections of the waves bouncing back and forth between the two mirrors.

The treatment is analogous to the one performed for the anti-reflection coating. By introducing the field reflection and transmission coefficients of the two mirrors, ρ_1, τ_1, ρ_2, τ_2, and calling L the distance between the two mirrors, the electric field of the transmitted wave, E_t, is given by a converging geometric series:

$$E_t = E_i \tau_1 \tau_2 e^{i\Delta/2}[1 + \rho_1\rho_2 e^{i\Delta} + (\rho_1\rho_2 e^{i\Delta})^2 + \cdots] = E_i \frac{\tau_1\tau_2 e^{i\Delta/2}}{1 - \rho_1\rho_2 e^{i\Delta}}, \quad (2.115)$$

where E_i is the electric field of the incident wave, and Δ is the phase shift introduced by the round-trip propagation between the two mirrors. If λ is the wavelength of the incident wave and n the refractive index of the medium that fills the cavity, $\Delta = 4\pi nL/\lambda$. Putting $\rho_1 = \sqrt{R_1}\mathrm{e}^{\mathrm{i}\phi_1}$, $\rho_2 = \sqrt{R_2}\mathrm{e}^{\mathrm{i}\phi_2}$, $|\tau_1|^2 = 1 - R_1$, and $|\tau_2|^2 = 1 - R_2$, it is found:

$$T_{FP} = \frac{|E_t|^2}{|E_i|^2} = \frac{1 + R_1R_2 - R_1 - R_2}{1 + R_1R_2 - 2\sqrt{R_1R_2}\cos\delta}, \quad (2.116)$$

where $\delta = \Delta + \phi_1 + \phi_2$. The Fabry-Perot interferometer is usually made of two identical mirrors. Putting $R_1 = R_2 = R$, (2.116) becomes:

$$T_{FP} = \frac{(1 - R)^2}{1 + R^2 - 2R\cos\delta}. \quad (2.117)$$

As shown in Fig. 2.22, T_{FP} is a periodic function of the phase delay δ, and takes the maximum value, $T_{\max} = 1$, when $\cos\delta = 1$, that is $\delta = 2q\pi$, q being a positive integer. T_{FP} is minimum when $\cos\delta = -1$, that is, $\delta = (2q + 1)\pi$. The minimum value is: $T_{\min} = [(1 - R)/(1 + R)]^2$. For instance, if $R = 0.99$, $T_{\min} = 0.25 \times 10^{-4}$. The ratio between $T_{\max}$ and $T_{\min}$ is the contrast factor F:

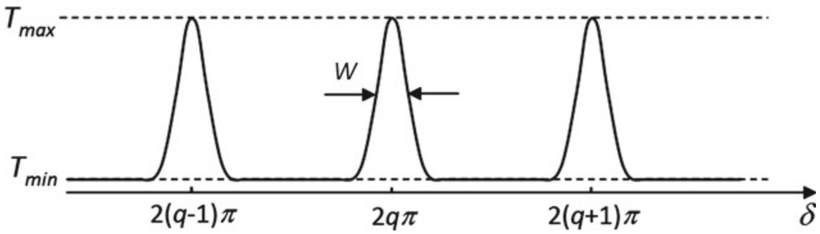


Fig. 2.22 Transmittance of the Fabry-Perot interferometer as a function of the round-trip phase shift δ, which is linearly related to the frequency of the light wave

$$F = \frac{T_{\max}}{T_{\min}} = 1 + \frac{4R}{(1-R)^2}. \tag{2.118}$$

Since δ is linearly dependent on the frequency ν of the incident wave, the abscissa scale in Fig. 2.22 can be considered as a frequency scale. As such the figure shows that the interferometer is a periodic filter in frequency-space. The frequencies corresponding to maximum transmission (resonance frequencies) are given by:

$$\nu_q = \frac{c}{2nL}(q - \frac{\phi_1 + \phi_2}{2\pi}). \tag{2.119}$$

The frequency distance between two consecutive transmission maxima, called free spectral range, is:

$$\Delta\nu = \nu_{q+1} - \nu_q = \frac{c}{2nL}. \tag{2.120}$$

An important parameter characterizing the performance of the interferometer is the width at half-height, W, of the transmission peak, expressed in units of radians. If R is close to 1, the approximate expression of W derived from (2.117) is: $W \simeq 4/\sqrt{F} \simeq 2(1-R)$. In frequency units, the peak width, $\delta\nu$, is derived by noting that $\delta\nu/\Delta\nu = W/(2\pi)$. Therefore:

$$\delta\nu = \frac{1-R}{\pi}\Delta\nu \tag{2.121}$$

As an example, if $R = 0.99$, $\delta\nu \approx \Delta\nu/300$. Taking $L = 1$ cm and $n = 1$, one finds $\Delta\nu = 15$ GHz and $\delta\nu = 50$ MHz.

Equation (2.121) seems to suggest that, once fixed $\Delta\nu$, the resolving power of the interferometer could be made arbitrarily large by choosing a value of R very close to 1. However, when $1-R$ is made very small, there are effects due to diffraction or misalignment losses that become important. In practice, it is difficult to obtain values of the ratio $\Delta\nu/\delta\nu$ (usually called finesse of the Fabry-Perot interferometer) larger than 300. The minimum achievable frequency resolution is of the order of 1 MHz.

The finding that a sequence of two highly reflecting mirrors could give rise to a complete transmission of the beam power, when $\cos\delta = 1$, may appear rather counter-intuitive. It is important to take into account that T_{FP} is the result of a steady-state calculation. If it is assumed that the interferometer is illuminated by an intensity step-function, starting at $t = 0$, it is clear that the initial transmittance is very low. If the incident frequency corresponds to a resonance, what happens is that optical power is progressively accumulated inside the interferometer, and the output intensity progressively grows, by approaching asymptotically the value of the input intensity. As an example, if $R = 0.99$ and the incident power is 1 W, the output power of 1 W will be reached only when the power bouncing back and forth inside the cavity attains the value of 100 W. It can be shown that the time constant controlling such a transient is the reciprocal of the frequency resolution $\delta\nu$. In other terms, the larger is the resolving power of the instrument, the longer is its time-response.

In order to use the Fabry-Perot interferometer for spectral analysis, one should be able to continuously shift the transmission peaks by changing the optical path nL, so that a scan of the frequency interval between two consecutive resonant frequencies can be performed. What is the change in L, δL, required to scan a given free spectral range? Assuming that the interferometer is initially tuned to ν_q, the condition to satisfy is that the new resonance ν'_{q-1} corresponding to the modified length $L + \delta L$ should coincide with ν_q:

$$\nu_q = q\frac{c}{2L} = \nu'_{q-1} = (q-1)\frac{c}{2(L-\delta L)} \tag{2.122}$$

From (2.122) it is found:

$$\delta L = \frac{L}{q} = \frac{c}{2\nu_q} = \frac{\lambda_q}{2}. \tag{2.123}$$

Therefore in order to scan a free spectral range it is sufficient to change the mirror distance by half wavelength.

In practical cases, the scan is performed by attaching one of the mirrors to a piezoceramic, and varying the ceramic thickness by the application of a slow voltage ramp. The mirror displacement is proportional to the applied voltage, and therefore changes linearly with time. In such a measurement, the time scale is transformed into a frequency scale, and so the observed temporal dependence of the transmitted power coincides with the spectral distribution of the incoming optical signal.

As an alternative approach, scanning of the resonant frequency can also be performed by keeping L fixed and changing the refractive index of the medium inside the Fabry-Perot cavity. This can be done by putting the interferometer inside a sealed gas cell and sweeping the pressure.

When L is larger than a few centimeters, spherical mirrors are used instead of plane mirrors. The advantage of this approach is that the alignment tolerance is less critical, and diffraction losses are lower.

To avoid ambiguity in spectral analysis, the spectral width of the incoming beam should be narrower than the free spectral range. This means that the Fabry-Perot interferometer is useful only for narrow-band signals.

2.5 Waves in Anisotropic Media

Manipulation of light polarization is a key issue in several applications. Main goals include controlled modification of the polarization state of a beam or selective transmission of that portion of beam polarized in a particular linear direction. To this end, components based on materials that exhibit an anisotropy in transmission or refraction are usually exploited. All media constituted by randomly positioned atoms (or molecules), like gases, liquids, and amorphous solids (glasses) are isotropic, that is, their macroscopic properties are independent of the orientation with respect to an applied field. In contrast, ordered media (crystals) or partially ordered media (liquid crystals) may be anisotropic. A consequence of this anisotropy is that the velocity of an optical wave inside the medium will depend on its propagation direction and polarization state. As will be seen, anisotropic materials can be used to change the polarization state of an optical beam.

Inside an anisotropic medium the relation between the electric field and the electric displacement is described by a tensor:

$$D_i(\mathbf{r}, \omega) = \varepsilon_o \sum \varepsilon_{ij} E_j(\mathbf{r}, \omega). \tag{2.124}$$

The indices i, j can take values 1, 2, 3, corresponding to the axes x, y, z. Equation (2.124) describes a system of three linear equations connecting the Cartesian components of $\mathbf{E}$ to those of $\mathbf{D}$. The electric permittivity tensor, ε_{ij}, is expressed by a 3×3 matrix. It can be demonstrated that such a matrix is symmetric, which means that $\varepsilon_{ij} = \varepsilon_{ji}$. Symmetric matrices can be made diagonal by an appropriate choice of the coordinate system. With reference to the diagonal axes, usually called principal axes, the off-diagonal elements of the electric permittivity matrix vanish:

$$\begin{pmatrix} \varepsilon_{11} & 0 & 0 \\ 0 & \varepsilon_{22} & 0 \\ 0 & 0 & \varepsilon_{33} \end{pmatrix} \tag{2.125}$$

From now on, unless otherwise specified, it will be assumed that the reference system is the system of principal axes.

In the general case in which $\mathbf{E}$ has an arbitrary direction inside the anisotropic medium, the vector $\mathbf{D}$ will not be parallel to $\mathbf{E}$. Considering a monochromatic plane wave with wave vector $\mathbf{k}$, propagating inside an anisotropic medium, Maxwell's equations indicate that $\mathbf{D}$ and $\mathbf{H}$ are both perpendicular to $\mathbf{k}$, whereas $\mathbf{E}$ is perpendicular to $\mathbf{H}$, but not, in general, to $\mathbf{k}$. However, if $\mathbf{E}$ is directed along one of the principal axes, $\mathbf{E}$ and $\mathbf{D}$ are parallel, but the relative permittivity will take different values depending on the specific principal axis. Note that the polarization direction inside an anisotropic medium is the direction of $\mathbf{D}$, not of $\mathbf{E}$.

It is useful to distinguish three situations:

- isotropic medium, $\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{33}$, the refractive index $n = \sqrt{\varepsilon_{11}}$ is a scalar quantity. The wave propagation velocity is the same for all directions and all polarizations.
- uniaxial medium, $\varepsilon_{11} = \varepsilon_{22} \neq \varepsilon_{33}$, it is possible to define an ordinary refractive index $n_o = \sqrt{\varepsilon_{11}} = \sqrt{\varepsilon_{22}}$ and an extraordinary refractive index $n_e = \sqrt{\varepsilon_{33}}$. The z axis is called optical axis.
- biaxial medium, $\varepsilon_{11} \neq \varepsilon_{22} \neq \varepsilon_{33}$, the medium is characterized by three distinct refractive indices.

In order to describe wave propagation in a uniaxial medium, it is better to start with a simple situation, namely that of a wave propagating in a direction perpendicular to the optical axis. Assuming that **k** is parallel to the x axis, **D** lies in the y-z plane. If the wave is linearly polarized along y, then its velocity will be c/n_o, where $n_o = \sqrt{\varepsilon_{22}}$. Analogously, if the wave is polarized along z, then the propagation velocity will be c/n_e, where $n_e = \sqrt{\varepsilon_{33}}$. In both cases, the wave polarization does not change during propagation. What happens now if the linear polarization direction is not coincident with a principal axis? By exploiting the linearity of Maxwell equations, the plane wave can be described as a linear combination of two waves, one polarized along y and the other polarized along z. The two waves travel with different velocities, and so a phase delay arises. When recombined, the resulting wave is elliptically polarized.

In the general case, in which **k** is not directed along a principal axis, the situation is more complicated. It would be useful to discover if it is also the case that there are two "special" linear polarization directions possessing the property that the polarization remains unchanged during propagation. Luckily, straightforward theory shows not only that those special directions exist, but also provides the means of determining them. Once the directions are found, the propagation is treated by describing **D** as the sum of two principal polarization modes.

The principal directions are found by using a geometric construction based on the index ellipsoid that is described by the equation:

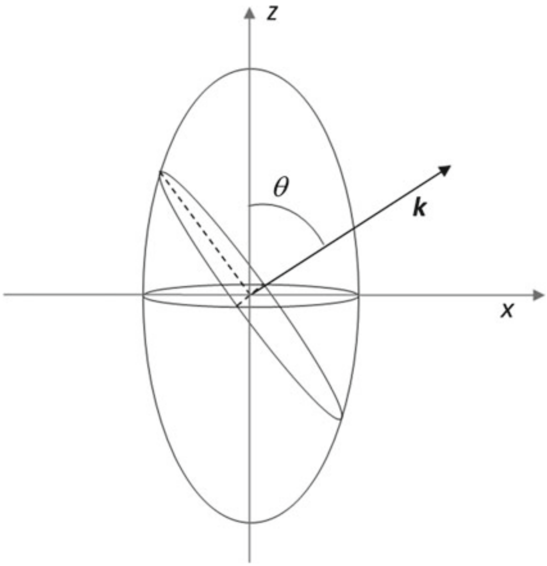
$$\frac{x^2}{{n_1}^2} + \frac{y^2}{{n_2}^2} + \frac{z^2}{{n_3}^2} = 1. \tag{2.126}$$

Consider the ellipse generated by intersecting the ellipsoid with the plane perpendicular to **k** and containing the origin of the coordinate system, as shown in Fig. 2.23. The two principal modes have polarization directions given by the ellipse axes. In addition, if n_a and n_b are the lengths of the ellipse semi-axes, then the propagation velocities of the two modes are c/n_a and c/n_b, respectively.

In the case of a uniaxial medium, by taking the symmetry axis parallel to z, and putting $n_1 = n_2 = n_o$, and $n_3 = n_e$, (2.126) becomes:

$$\frac{x^2 + y^2}{{n_o}^2} + \frac{z^2}{{n_e}^2} = 1. \tag{2.127}$$

Fig. 2.23 Index ellipsoid



From now on, the discussion is restricted to uniaxial media. It can be assumed without loss of generality that **k** is in the plane x-z, forming an angle θ with the z axis. The plane that is perpendicular to **k** and passes through the origin has equation: $x\sin\theta + z\cos\theta = 0$. The intersection of this plane with the index ellipsoid gives an ellipse that has one axis coincident with y, and the other one, placed in the x-z plane, forming an angle θ with z. The length of the first semi-axis is n_o. The length of the second semi-axis, called n_θ, satisfies the equation:

$$\frac{1}{n_\theta^2} = \frac{\cos^2\theta}{{n_o}^2} + \frac{\sin^2\theta}{{n_e}^2} \tag{2.128}$$

A generic plane wave with wave vector **k** can be described as the superposition of an ordinary wave and an extraordinary wave. The former, polarized perpendicularly to the plane defined by **k** and z, travels at velocity c/n_o. The latter, polarized in the plane defined by **k** and z, travels at velocity c/n_θ. If $\theta = 0$, then $n_\theta = n_o$, that is, for a wave propagating along the optical axis the medium behaves as though it were isotropic. If $\theta = 90°$, the extraordinary wave is polarized along z, and has a propagation velocity c/n_e.

For the ordinary wave, **D** has no components along z, therefore **D** is parallel to **E**, and the Poynting vector is parallel to **k**. For the extraordinary wave, if $0 < \theta < 90°$, **D** has nonzero components along both z and x. Therefore **D** is not parallel to **E**. The consequence is that the Poynting vector forms with **k** a non-zero angle, ψ, called walk-off angle. ψ is given by the relation:

$$\tan\psi = \frac{\sin(2\theta)n_\theta^2}{2}\left(\frac{1}{{n_e}^2} - \frac{1}{{n_o}^2}\right) \tag{2.129}$$

If the propagating wave is a collimated beam, the effect of walk-off is to generate a spatial separation between ordinary and extraordinary beams.

Up to now the treatment has concerned the propagation inside the crystal. It is also important to see what happens when a wave coming from an isotropic medium, impinges on the crystal surface. The discussion is here limited to the case of a uniaxial crystal having the optical axis in the incidence plane. An incident wave that is linearly polarized perpendicularly to the incidence plane (σ case) propagates in the crystal as an ordinary wave, whereas an incident π wave will behave as an extraordinary wave. Since the refractive index is different in the two cases, the refraction angles calculated by Snell's law will also be different. In the case of arbitrary polarization, an object seen through the uniaxial crystal will present a double image, hence the name birefringent crystal.

The refractive indices of a few birefringent crystals are given in Table 2.3. Birefringence is said to be positive if $n_e > n_o$, as in the case of quartz crystal, and negative if $n_e < n_o$, as in the case of calcite crystal.

2.5.1 Polarizers and Birefringent Plates

Many applications require a well-defined polarization state. Conventional optical sources produce light having a polarization state that changes randomly with time (unpolarized light). Some lasers emit polarized light. Birefringent crystals are used to generate polarized light or to modify the state of polarization of a light beam.

The simplest polarizer is made from a dichroic material. This is an anisotropic material having, for instance, a large absorption coefficient for light polarized along the optical axis, and a small absorption for light polarized perpendicularly to the

Table 2.3 Refractive indices of some birefringent crystals

Crystal	n_o	n_e	λ (nm)
Calcite, $CaCO_3$	1.658	1.486	589
Quartz, SiO_2	1.543	1.552	589
Rutile, TiO_2	2.616	2.903	589
Corundum, Al_2O_3	1.768	1.760	589
Lithium niobate, $LiNbO_3$	2.304	2.212	589
Barium borate, BaB_2O_4	1.678	1.553	589

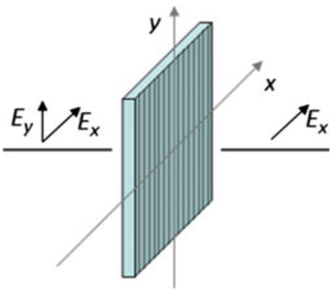


Fig. 2.24 Polaroid polarizer. The polymer chains are parallel to y, and so only the linear polarization parallel to x is transmitted

optical axis. The most common dichroic material, Polaroid, is composed of a parallel arrangement of linear chains of conjugated polymers (see Fig. 2.24).

The optical axis coincides with the chain direction. The polymer presents high electric conductivity along the chain and low conductivity across the chain. In a grossly simplified description it behaves as a conductor if the applied electric field is parallel to the optical axis, and as an insulator if the electric field is perpendicular to the optical axis. In practice the component polarized perpendicularly to the chains also suffers a non-negligible attenuation. Polarizers made of Polaroid are cheap, but the selectivity is not very large (a typical value of the extinction ratio, defined as the transmittance ratio between unwanted and wanted polarization state, could be $\approx 1/500$). Since absorption heats the material, high-intensity beams damage the polarizer.

High-selectivity polarizers, having an extinction ratio about 10^5, consist of a pair of prisms made of a birefringent crystal, typically calcite. One of the most utilized polarizers is the Glan–Thompson polarizer, shown in Fig. 2.25, which is made of two calcite prisms cemented together by their long faces. The cement has a refractive index intermediate between the ordinary and the extraordinary indices of calcite. The optical axes of the calcite crystals are parallel and aligned perpendicular to the incidence plane, that is, the plane of the figure. Light of arbitrary polarization is split into two rays, experiencing different refractive indices; the π-polarized ordinary ray is totally internally reflected from the calcite-cement interface, leaving the σ-polarized extraordinary ray to be transmitted. The prism can therefore be used as a polarizing beam splitter. Traditionally a resin called Canada balsam was used as the cement in assembling these prisms, but this has largely been replaced by synthetic polymers.

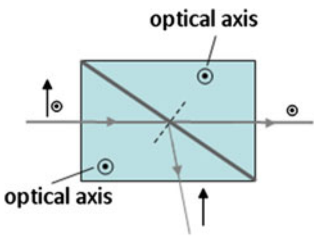
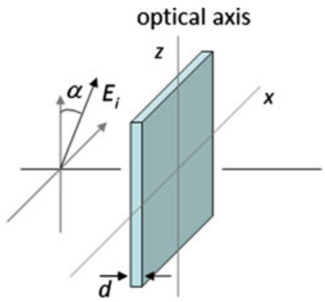


Fig. 2.25 Glan-Thompson prism

Fig. 2.26 Birefringent plate



The Glan-Thompson polarizer operates over a very wide wavelength range, from 350 nm to 2.3 µm. This type of polarizer can handle much higher optical powers than the dichroic polarizers.

Thin plates made of a birefringent crystal are used to change the polarization state of a light beam. The plate faces are parallel to the x-z plane, and the light beam travels in the y direction, perpendicular to the optical axis z, as shown in Fig. 2.26. If the incident beam is linearly polarized in the x direction (ordinary wave), then the propagation across the plate does not modify the polarization state, but simply introduces a phase shift $\Delta\phi_o = 2\pi n_o d/\lambda$, where d is the plate thickness. Similarly, no modification of the polarization state occurs for the beam polarized along z, which is phase-shifted by: $\Delta\phi_e = 2\pi n_e d/\lambda$. An input wave linearly polarized in an arbitrary direction can be described as the sum of an ordinary and an extraordinary wave. The two waves are initially in phase, but acquire, on crossing the plate, a phase difference $\Delta\phi = 2\pi(n_e - n_o)d/\lambda$. In general, the output polarization is elliptical. An important particular case is a birefringent plate that gives a phase difference of $|\Delta\phi| = \pi$. This is known as a half-wave plate. When linearly-polarized light is incident upon a half-wave plate, the emergent light is also linearly-polarized. If the polarization direction of the incident light makes an angle α with the optical axis, then the polarization direction of the emergent beam makes an angle $-\alpha$ with the same axis, therefore the direction of polarization has been rotated through an angle 2α, as shown in Fig. 2.27. A plate for which $|\Delta\phi| = \pi/2$ is called a quarter-wave plate. When linearly-polarized light is incident upon a quarter-wave plate, the emergent light is in general elliptically polarized. If, however, $\alpha = \pi/4$, the emergent light is circularly polarized.

2.5.2 *Jones Matrices*

The propagation of polarized light beams trough anisotropic optical elements is conveniently described by using the Jones formalism, in which the electric field of the light wave is represented by a two-component vector and the optical element by a 2×2 transfer matrix.

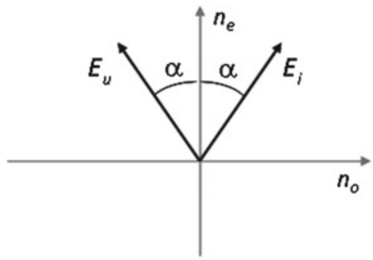


Fig. 2.27 Rotation of the polarization direction in a half-wave plate

In general, the electric field of a monochromatic plane wave propagating along z can be described as the sum of two components:

$$\mathbf{E}(z,t) = E_x \exp[-i(\omega t - kz + \phi_x)]\mathbf{x} + E_y \exp[-i(\omega t - kz + \phi_y)]\mathbf{y}, \quad (2.130)$$

where **x** and **y** are unit vectors, directed along x and y, respectively. Equation (2.130) can be written as:

$$\mathbf{E}(z,t) = E_o \exp[-i(\omega t - kz + \phi_x)]\mathbf{V}, \quad (2.131)$$

where $E_o = \sqrt{{E_x}^2 + {E_y}^2}$, and **V** is the Jones vector represented by the column matrix:

$$\begin{pmatrix} \cos\alpha \\ e^{i\phi}\sin\alpha \end{pmatrix}, \quad (2.132)$$

where $\alpha = \arctan(E_y/E_x)$ and $\phi = \phi_y - \phi_x$.

The adopted sign convention considers the angles measured counter-clockwise as being positive.

The vector (2.132) corresponds to a generic elliptical polarization. If the wave is linearly polarized, $\phi = 0$. The following Jones vectors describe a linearly polarized field with a polarization direction forming the angle α with the x axis, or directed along the x axis, or along the y axis, respectively:

$$\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}; \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.133)$$

For a circularly polarized wave, $\alpha = \pi/4$ and $\phi = \pm\pi/2$. Therefore the Jones vectors associated with clockwise and counter-clockwise circular polarization are respectively given by:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}; \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (2.134)$$

Two waves represented by the vectors $\mathbf{V_1}$ and $\mathbf{V_2}$ have orthogonal polarization states if the scalar product of the two vectors is zero. As an example, the linear polarization states along x and y are mutually orthogonal. An arbitrary polarization state can always be described as the sum of two orthogonal polarization states.

An optical component modifying the polarization state of an incident beam is described by a transfer matrix $\mathbf{T}$. If $\mathbf{V_i}$ represents the incident wave, then the Jones vector of the emergent wave is given by: $\mathbf{V_u} = \mathbf{TV_i}$. If the beam goes through a sequence of components with matrices $\mathbf{T_1}, \ldots, \mathbf{T_n}$, then the emergent wave is given by: $\mathbf{V_u} = \mathbf{T_n} \ldots \mathbf{T_1 V_i}$. Note that, in general, the $\mathbf{T}$ matrices do not commute.

The polarizers having optical axis along x and y are represented, respectively, by the matrices $\mathbf{T_{px}}$ e $\mathbf{T_{py}}$:

$$\mathbf{T_{px}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad \mathbf{T_{py}} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.135}$$

The birefringent plate having optical axis directed along x introduces a phase shift $\Delta\phi$ between the x and y components of the electric field, without modifying the amplitude ratio E_x/E_y. It is represented by:

$$\mathbf{T_r} = \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\Delta\phi) \end{pmatrix}. \tag{2.136}$$

The phase shift is given by $\Delta\phi = 2\pi(n_e - n_o)d/\lambda$, where d is the plate thickness. The half-wave plate corresponds to $\Delta\phi = \pi$, and the quarter-wave plate to $\Delta\phi = \pi/2$.

The polarization rotation matrix,

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{2.137}$$

represents an optical component that rotates the polarization direction. If the incident wave is linearly polarized at an angle θ_1, the emergent wave is linearly polarized at the angle $\theta_2 = \theta_1 + \theta$.

It is important to take into account that Jones vectors and matrices change when the coordinate system is changed. If a field is described by the vector $\mathbf{V}$, once the coordinate system is rotated by the angle θ, the same field is described by the new vector $\mathbf{V}'$:

$$\mathbf{V}' = \mathbf{R}(-\theta)\mathbf{V}. \tag{2.138}$$

A polarizer having its optical axis that forms an angle θ with the x axis is represented by the matrix:

$$\mathbf{T}_{\mathbf{p}\theta} = \mathbf{R}(\theta)\mathbf{T_{px}}\mathbf{R}(-\theta) = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}. \tag{2.139}$$

Similarly, the matrix describing an optical retarder (birefringent plate) having the optical axis rotated by θ with respect to x is written as:

$$\mathbf{T}_{\mathbf{r}\theta} = \mathbf{R}(\theta)\mathbf{T}_{\mathbf{r}}\mathbf{R}(-\theta)$$
$$= \begin{pmatrix} 1 + [\exp(i\Delta\phi) - 1]\sin^2\theta & [1 - \exp(i\Delta\phi)]\sin\theta\cos\theta \\ [1 - \exp(i\Delta\phi)]\sin\theta\cos\theta & \exp(i\Delta\phi) - [\exp(i\Delta\phi) - 1]\sin^2\theta \end{pmatrix}. \quad (2.140)$$

2.5.3 Rotatory Power

Certain materials present a particular type of birefringence: the two fundamental modes of propagation are circularly rather than linearly polarized waves, with right- and left-circular polarizations traveling at different velocities, c/n^+ and c/n^-. This property, known as optical activity, is typical of materials having an intrinsically helical structure at the molecular or crystalline level.

As mentioned in Sect. 1.3, a linearly polarized wave can be decomposed into the sum of two circularly polarized waves having the same amplitude. By using Jones matrices, one can write:

$$\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \frac{1}{2}\exp(-i\theta)\begin{pmatrix} 1 \\ i \end{pmatrix} + \frac{1}{2}\exp(i\theta)\begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (2.141)$$

If L is the thickness of the optically active medium, the phase delay between the two circularly polarized waves is given by: $\phi = 2\pi L(n^+ - n^-)/\lambda$. The recombination of the two emergent waves produces a linearly polarized beam (see Fig. 2.28) with a polarization direction rotated with respect to that of the incident beam:

$$\frac{1}{2}\exp(-i\theta)\begin{pmatrix} 1 \\ i \end{pmatrix} + \frac{1}{2}\exp[i(\theta - \phi)]\begin{pmatrix} 1 \\ -i \end{pmatrix}$$
$$= \exp(-i\phi/2)\begin{pmatrix} \cos(\theta - \phi/2) \\ \sin(\theta - \phi/2) \end{pmatrix}. \quad (2.142)$$

The rotation per unit length, called rotatory power ρ, is:

$$\rho = \frac{\phi}{2L} = \frac{\pi(n^+ - n^-)}{\lambda}. \quad (2.143)$$

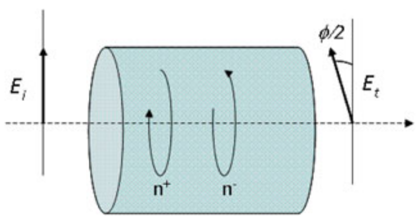


Fig. 2.28 Rotatory power

A typical optically active crystal is quartz. For propagation along the optical axis, the rotatory power is 4 rad/cm at $\lambda = 589$ nm. Also disordered media containing chiral molecules, as, for instance, sugar molecules, can show optical activity. For an aqueous solution containing 0.1 $\mathrm{g/cm^3}$ of sugar, $\rho = 10^{-2}$ rad/cm. The sugar concentration can therefore be determined by measuring the rotatory power.

From a formal point of view, the optical activity can be explained only assuming the presence of anti-symmetric off-diagonal terms inside the ε_{ij} matrix. Such a presence would contradict the assertion that the matrix ε_{ij} is symmetric. However, it can be shown that such a symmetry property applies only for spatially homogeneous fields, that is, infinite wavelength fields. If λ is finite, the local value of **D** depends, in general, on **E** and its spatial derivatives. At first order:

$$D_i = \varepsilon_o \sum_{jm} \left(\varepsilon_{ij} E_j + i\gamma_{ijm} \frac{\partial E_j}{\partial x_m} \right). \tag{2.144}$$

whereas ε_{ij} is symmetrical, the third-order tensor γ_{ijm} is anti-symmetrical with respect to the indices ij: $\gamma_{ijm} = -\gamma_{jim}$. If there is no absorption, then γ_{ijm} is real. In the case of an incident plane wave with wave vector **k**, it is found that $\partial E_j/\partial x_m = iE_j k_m$. Hence:

$$D_i = \varepsilon_o \sum_{jm} \left(\varepsilon_{ij} E_j + i\gamma_{ijm} E_j k_m \right). \tag{2.145}$$

D is related to **E** by the new electric permittivity tensor:

$$\varepsilon'_{ij} = \varepsilon_{ij} + i \sum_{m} \gamma_{ijm} k_m, \tag{2.146}$$

which is the sum of a symmetric and an anti-symmetric term, the latter vanishing if the modulus of **k** vanishes, that is, if the wavelength becomes infinite.

For an isotropic medium, (2.145) becomes:

$$D_i = \varepsilon_o [\varepsilon_r E_i + i\gamma (\mathbf{E} \times \mathbf{k})_i], \tag{2.147}$$

where $(\mathbf{E} \times \mathbf{k})_i$ denotes the i-th component of the cross product, and γ is a scalar quantity. Assuming **k** parallel to z, the **D** components along x and y, as derived from (2.147), are:

$$D_x = \varepsilon_o [\varepsilon_r E_x + i\gamma k E_y] \tag{2.148}$$

$$D_y = \varepsilon_o [\varepsilon_r E_y - i\gamma k E_x]. \tag{2.149}$$

These two equations show that a linear polarization cannot propagate unchanged inside an optically active medium, due to the fact that E_x generates D_y, and viceversa. Considering that the electric field of a right- (left-) circularly polarized beam is written as: $E^{\pm} = E_x \pm iE_y$, by using (2.148) and (2.149) the quantities $D^{\pm} = D_x \pm iD_y$ may be written as:

$$D^+ = \varepsilon_o[\varepsilon_r + \gamma k]E^+ \tag{2.150}$$

$$D^- = \varepsilon_o[\varepsilon_r - \gamma k]E^- \tag{2.151}$$

These two equations show that circular polarization indeed represents a propagation mode inside the optically active medium. The refractive indices seen by the two circular polarizations are: $n^{\pm} = \sqrt{\varepsilon_r \pm \gamma k}$. Assuming that the second term under square root is much smaller than the first one, putting $k = n\omega/c$ and $\varepsilon_r = n^2$, one obtains the approximate relation: $n^{\pm} = n[1 \pm \gamma k/(2)]$, and, consequently, $\rho = \pi\gamma k/(n\lambda)$.

2.5.4 Faraday Effect

There are various means to change the optical properties of a medium by applying an external field. In this section the magneto-optic effect known as Faraday effect is described. It consists in the induction of rotatory power through the application of a static magnetic field. The effect is longitudinal, since the applied magnetic field $\mathbf{B}$ is parallel to the propagation direction of the linearly-polarized optical wave. The polarization direction is rotated by an angle $\theta = C_V BL$, where L is the propagation length and C_V is the Verdet constant. The sign of θ changes with the sign of $\mathbf{B}$, but it does not change by reversing the propagation direction. Therefore the rotation effect is doubled if the beam travels back and forth inside the medium.

In the standard configuration of the Faraday rotator the medium has a cylindrical shape and is placed inside a solenoid. The longitudinal magnetic field is created by an electric current flowing into the solenoid.

In the presence of the magnetic field, the relation between $\mathbf{D}$ and $\mathbf{E}$ can be written, similarly to (2.144), as:

$$\mathbf{D} = \varepsilon_o[\varepsilon_r\mathbf{E} + i\gamma_B(\mathbf{B} \times \mathbf{E})], \tag{2.152}$$

where γ_B is called magneto-gyration coefficient. The rotatory power is given by:

$$\rho = C_V B = -\frac{\pi\gamma_B B}{n\lambda}. \tag{2.153}$$

The big difference between a passive rotator, such as a quartz plate, and the Faraday rotator is made clear by considering the situation in which the wave transmitted in one direction is reflected back and retransmitted in the opposite direction. In the case of the passive rotator there is no net polarization rotation because the rotation generated in the forward path is canceled during the backward path, while in the case of the Faraday rotator the effect is doubled. The Faraday rotator is said to be a nonreciprocal device.

The Faraday effect can be found in gases, liquids, and solids. Since its microscopic origin is connected to the Zeeman effect, the Verdet constant becomes larger when the wavelength of the optical wave gets close to absorption lines. The most commonly utilized materials are doped glasses and crystals, like YIG (yttrium iron garnet), TGG (terbium gallium garnet) and TbAlG (terbium aluminum garnet). As an example, the Verdet constant of TGG is: $C_V = -134$ rad/(T $\cdot$ m) at $\lambda = 633$ nm, and becomes -40 rad/(T $\cdot$ m) at $\lambda = 1064$ nm.

There are some ferromagnetic materials (containing Fe, Ni, Co) that exhibit large values of C_V, but they also have a very large absorption coefficient, so the Faraday rotation can only be observed in reflection. They have an important application in optical memories, as will be mentioned in Chap. 8.

2.5.5 *Optical Isolator*

An optical isolator is a device that transmits light in only one direction, thereby preventing reflected light from returning back to the source. Isolators are utilized in many laser applications, since back-reflected beams often have deleterious effects on the operating conditions of optical devices.

A Faraday isolator may be constructed by placing a Faraday rotator between two polarizers, as shown in Fig. 2.29. The optical axis of polarizer 1 is along x, the Faraday cell rotates clockwise the polarization direction by $\pi/4$, the optical axis of polarizer 2 makes a $\pi/4$ angle with the x axis, so that the forward traveling wave is fully transmitted by the device. If a portion of the transmitted wave is back-reflected by a

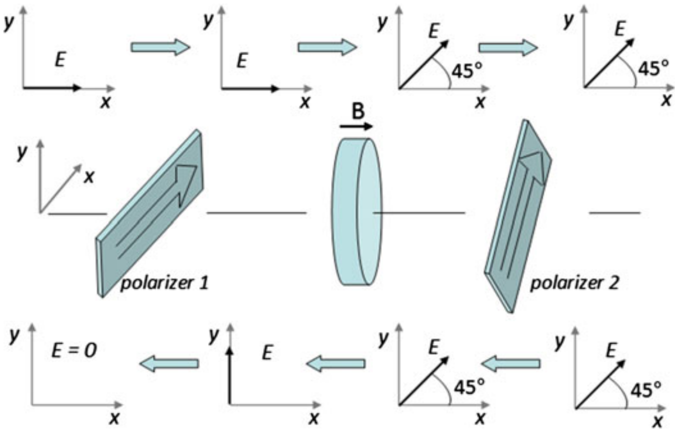


Fig. 2.29 Faraday isolator

subsequent optical component, it travels inside the isolator in the backward direction. Assuming that reflection does not alter the polarization state, the reflected beam is fully transmitted by polarizer 2, and its polarization direction suffers a clockwise rotation of $\pi/4$ by the Faraday cell. The backward beam becomes y-polarized, and therefore is blocked by polarizer 1.

The behavior of the isolator can be mathematically described by Jones matrices. Let $\mathbf{V}_\mathrm{i}$ and $\mathbf{V}_\mathrm{u}$ be the Jones vectors of the incident and emergent waves, respectively, and call $\mathbf{T_1}$, $\mathbf{T_2}$, $\mathbf{T_3}$, the matrices representing polarizer 1, Faraday rotator, and polarizer 2, respectively. It is important to recall that all vectors and matrices must be expressed with reference to the same x-y coordinate system. The output field is related to the input field by the following expression:

$$\begin{aligned}\mathbf{V}_\mathrm{u} &= \mathbf{T_3T_2T_1V_i}\\ &= \begin{pmatrix}1/2 & 1/2\\ 1/2 & 1/2\end{pmatrix}\begin{pmatrix}\sqrt{2}/2 & -\sqrt{2}/2\\ \sqrt{2}/2 & \sqrt{2}/2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{pmatrix}.\end{aligned} \tag{2.154}$$

As expected, the emergent wave is linearly polarized at $\pi/4$ with respect to the x axis. The back-reflected beam crosses the sequence of devices in opposite order, and is completely blocked by the isolator:

$$\mathbf{V'_u} = \mathbf{T_1T_2T_3V_u} = \begin{pmatrix}0\\ 0\end{pmatrix}. \tag{2.155}$$

Note that $\mathbf{T_3}$ seems to have no role in the proposed scheme. In practice, however, it may happen that back-reflection modifies the polarization state. In such a case the role of polarizer 2 becomes important.

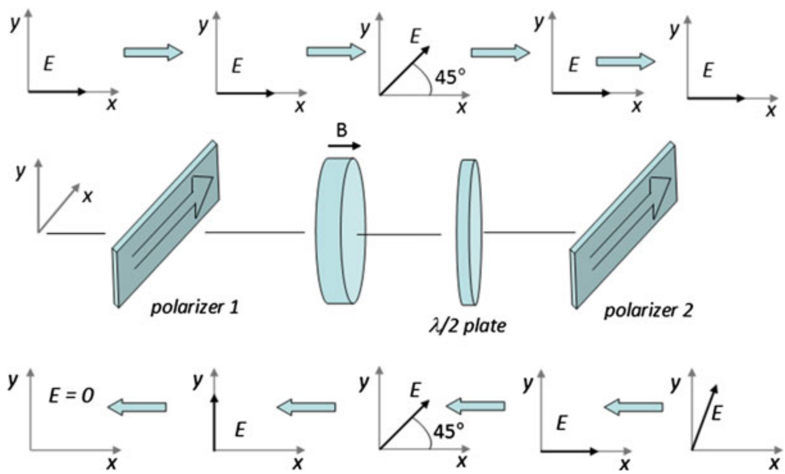


Fig. 2.30 Faraday isolator including a half-wave plate

The sequence of a Faraday rotator plus a half-wave plate constitutes another nonreciprocal device, as schematically depicted in Fig. 2.30. If the former rotates the polarization direction by $\pi/4$, and the latter by $-\pi/4$, a linearly polarized beam does not change its polarization state by crossing the device. On the contrary, for a beam coming from the opposite direction, the two rotations have the same sign, so that its polarization direction is rotated by $\pi/2$. Such a device can be used to make uni-directional the wave propagation in a circular path.

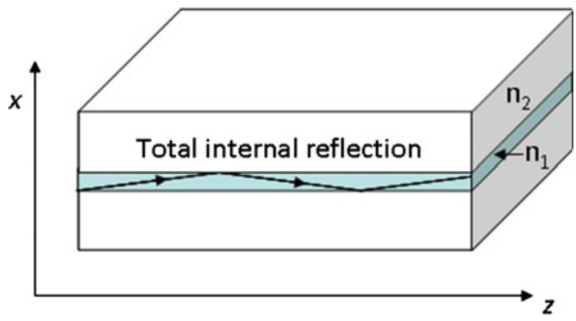
2.6 Waveguides

Waveguides are strips of dielectric material embedded in another dielectric material of lower index of refraction. Starting from the advent of optical communications, the development of integrated optical systems has become more and more important. In such systems the light propagates in a confined regime through optical waveguides, instead of free space, without suffering from the effect of diffraction.

A light beam traveling in a direction that is only slightly tilted with respect to the guide axis is totally reflected at the boundary between the two media, provided the incidence angle is larger than the limit angle. So that the guide can completely trap a light beam in a diffraction-less propagation: the optical power should be mainly confined inside the guide, with only an evanescent tail extending outside.

Many aspects of guided propagation can be understood by studying the simple case of a planar dielectric waveguide, consisting of a slab of thickness d and refractive index n_1, usually called core, surrounded by a cladding of lower refractive index n_2, as shown in Fig. 2.31. In a planar waveguide, confinement is guaranteed only in one dimension. Light rays making angles θ with the z-axis, in the x-z plane, undergo total internal reflection if $\cos\theta \geq n_2/n_1$. The electromagnetic theory predicts the existence of waveguide modes, that is, transverse field distributions that are invariant in the propagation. Here, the discussion is limited to waves linearly polarized along x, that is, to transverse electric-field modes (TE modes). The propagation properties are derived by writing the Helmholtz equation inside the two media,

Fig. 2.31 Scheme of a planar waveguide



$$(\nabla^2 + (n_1 k_o)^2)\mathbf{E}(\mathbf{r}) = 0 \tag{2.156}$$

$$(\nabla^2 + (n_2 k_o)^2)\mathbf{E}(\mathbf{r}) = 0, \tag{2.157}$$

and imposing the boundary conditions.

By assuming that the wave propagates along z with a propagation constant β, the electric field can be generally written as:

$$E(x, z, t) = A(x)\exp[i(\beta z - \omega t)] \tag{2.158}$$

By substitution of (2.158) into (2.156), the equation relative to medium 1 (the core) becomes:

$$\frac{d^2 A}{dx^2} + [(n_1 k_o)^2 - \beta^2]A(x) = 0. \tag{2.159}$$

Similarly, inside the cladding:

$$\frac{d^2 A}{dx^2} + [(n_2 k_o)^2 - \beta^2]A(x) = 0 \tag{2.160}$$

Since the propagation constant of the guided wave must be intermediate between $n_2 k_o$ and $n_1 k_o$, the solution of (2.159) is a periodic function, $A_1 \cos(\gamma_1 x)$, and that of (2.160) is an exponential function, $A_2 \exp(-\gamma_2 x)$, where the real quantities γ_1 and γ_2 are given, respectively, by:

$$\gamma_1 = \sqrt{{n_1}^2 {k_o}^2 - \beta^2}, \tag{2.161}$$

and

$$\gamma_2 = \sqrt{\beta^2 - {n_2}^2 {k_o}^2} \tag{2.162}$$

By imposing the continuity of the field amplitude and its derivative at $x = d/2$, it is found:

$$A_1 \cos(\gamma_1 d/2) = A_2 \exp(-\gamma_2 d/2), \tag{2.163}$$

and

$$A_1 \gamma_1 \sin(\gamma_1 d/2) = A_2 \gamma_2 \exp(-\gamma_2 d/2). \tag{2.164}$$

Non-zero field amplitudes A_1 and A_2 exist only if the following condition is satisfied:

$$\tan(\gamma_1 d/2) = \frac{\gamma_2}{\gamma_1}. \tag{2.165}$$

This is a transcendental equation in the variable β. A graphic solution can be found by plotting the right- and left-hand sides as functions of β.

What has been calculated is the propagation constant of the fundamental mode. It can be shown that the waveguide can support a discrete sequence of modes, whose propagation constants are derived by generalizing (2.165) as follows:

$$\tan(\gamma_1 d/2 - m\pi/2) = \frac{\gamma_2}{\gamma_1}, \tag{2.166}$$

where $m = 0, 1, \ldots$.

Whereas a solution always exists for the fundamental mode, a value of the propagation constant β_1 of the $m = 1$ mode satisfying the condition $n_2 k_o \leq \beta_1 \leq n_1 k_o$ is found only if the mode frequency ν is larger than the cut-off frequency ν_c given by:

$$\nu_c = \frac{c}{2d\sqrt{{n_1}^2 - {n_2}^2}}, \tag{2.167}$$

By increasing the mode order m, the cut-off frequency increases. Therefore, by increasing the frequency, the guide accepts more and more propagation modes. According to (2.167), to be single-mode the waveguide must be thin, and the refractive index difference $n_1 - n_2$ needs to be small. As it will be seen in the following chapters, there are various applications requiring single-mode propagation.

Concerning the field distribution of the waveguide modes, it should be noted that, by increasing m, the field penetrates deeper into the cladding. Correspondingly, the propagation constant β_m decreases with m.

Most of the considerations developed for the propagation in slab waveguides also apply for the case of channel waveguides, in which the confinement of the light beam is guaranteed on two dimensions. A very important example of channel waveguide is the optical fiber (Fig. 2.32a), which is a waveguide with a cylindrical symmetry that will be discussed in greater details in Chap. 6. Figure 2.32b–d illustrate the typical geometries of different channel waveguides fabricated on planar substrates, which constitute the basic structures of several integrated optical devices, like optical modulators (Chap. 4) and semiconductor lasers or amplifiers (Chap. 5). In the recent years, the evolution of communication systems has driven the development of optical circuits mainly based on silicon deposited on silicon oxide (silicon on insulator, SOI) in which several electronic and photonic functions are integrated on the same

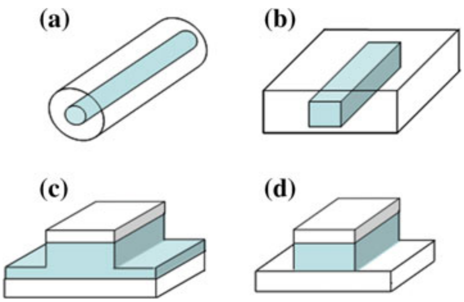


Fig. 2.32 Scheme of the most commonly used optical waveguide geometries: **a** optical fiber, **b** embedded waveguide, **c** rib waveguide, **d** ridge waveguide. The *darker regions* represent the cores in which the light is confined

chip. Miniaturization is a key issue for this application, as it will be discussed in Sect. 8.1.3, and waveguides with cross-sections of hundreds of nanometers and very high refractive contrast are needed. This provides high confinement of the light that can be bent in curved waveguides, thus making feasible the fabrication of complex networks of waveguides on a very small area.

Problems

2.1 Given that the wavelength of a light signal in vacuum is 600 nm, what will it be in a glass block having a refractive index $n = 1.50$?

2.2 A glass block having an index of 1.55 is coated with a layer of magnesium fluoride of index 1.32. For light traveling in the glass, what is the total reflection angle at the interface?

2.3 Yellow light from a sodium lamp ($\lambda = 589$ nm) traverses a tank of benzene (of index 1.50), which is 30-m long, in a time τ_1. If it takes time τ_2 for the light to pass through the same tank when filled with carbon disulfide (of index 1.63), determine the value of $\tau_2 - \tau_1$.

2.4 Light having a vacuum wavelength of 600 nm, traveling in a glass ($n = 1.48$) block, is incident at 45° on a glass-air interface. It is then totally internally reflected. Determine the distance into the air at which the amplitude of the evanescent wave has dropped to a value of 1/e of its maximum value at the interface.

2.5 A beam of light in air strikes the surface of a transparent material having an index of refraction of 1.5 at an angle with the normal of 40°. The incident light has component E-field amplitudes parallel and perpendicular to the plane-of-incidence of 10 and 20 V/m, respectively. Determine the corresponding reflected field amplitudes.

2.6 The complex index of refraction of silver at $\lambda = 532$ nm is $n' + in'' = 0.14 - 3.05i$. Calculate: (i) the reflectance of an air-silver surface at normal incidence; (ii) the transmittance of a 100 μm silver plate.

2.7 A Gaussian beam at wavelength $\lambda = 0.63$ μm having a spot size at the beam waist of $w_o = 0.5$ mm is to be focused to a spot size of 30 μm by a lens positioned at a distance of 0.5 m from the beam waist. What focal length should the lens have?

2.8 A Gaussian beam with $\lambda = 800$ nm and $w_o = 0.1$ mm at $z = 0$ is propagating along the z axis. A plano-convex lens, made of a spherical glass cap having a radius of curvature of 4 cm, is placed at $z_1 = 20$ cm. The refractive index of the glass is $n = 1.8$. Determine: (i) at which coordinate z_2 the beam is focused; (ii) the beam spot size at z_2; (iii) the lens diameter to be chosen in order to collect 99 % of the beam power.

2.9 A blazed reflection grating with pitch $d = 1200\,\mathrm{nm}$ and blazing angle $\theta_g = 20°$ is used in the Littrow configuration. Determine which incident wavelength is diffracted at second order in the backward direction.

2.10 A Fabry-Perot interferometer made by two identical mirrors at distance $d = 1\,\mathrm{mm}$ is illuminated by a light beam containing two wavelengths $\lambda_1 = 1064\,\mathrm{nm}$ and $\lambda_2 = \lambda_1 + \Delta\lambda$, where $\Delta\lambda = 0.05\,\mathrm{nm}$. Determine which is the minimum reflectance of the two mirrors required to obtain a resolving power $P_r \geq \lambda_1/\Delta\lambda$.

2.11 Consider a light beam at $\lambda = 589\,\mathrm{nm}$ traveling along the z-axis and linearly polarized along the x-axis. The light beam crosses, in succession, a quarter-wave plate and a polarizer. The optical axis of the quarter-wave plate is in the x-y plane, forming an angle $\pi/4$ with the x-axis. The optical axis of the polarizer is parallel to the y-axis. Write the Jones matrix of the quarter-wave plate in the x-y reference frame, and calculate the fraction of the incident power that is transmitted by the polarizer.

2.12 Consider the following sequence of components: a polarizer P_1, a Faraday rotator that rotates counterclockwise the linear polarization direction of the incoming beam by 45°, a calcite half-wave plate, and a second polarizer P_2. Assume that the incoming beam travels along y and is linearly polarized along x, that the two polarizers have optical axes parallel to x, and that the optical axis of the calcite plate lies in the x-z plane. (i) Calculate which angle the optical axis of the plate should form with the x axis in order to ensure a unitary transmission. (ii) Write the Jones matrix of the half-wave plate in the x-z reference frame.

http://www.springer.com/978-3-319-20626-4

Photonics
A Short Course
Degiorgio, V.; Cristiani, I.
2016, XII, 253 p. 124 illus., Softcover
ISBN: 978-3-319-20626-4