Chapter 3

Thin Films

Contents

3.1   Introduction

This chapter deals with the propagation of optical waves in and through thin films. Thin films are used for a very wide range of optical applications ranging from a thin metal film for use as a mirror to stacks of quarter-wave layers for highly reflective coatings in laser cavities. Thin films also serve as material systems for integrated photonic circuits such as polymer films for polymer waveguide circuits or SOI (silicon-on-insulator) circuits. In this chapter we will gain insight in the optical properties of thin films and provide tools to deal with thin film problems in a quantitative way.

The problem considered is schematically represented in Figure 3.1: a light source generates electromagnetic waves and illuminates a stack of layers of different media. It is clear that the propagation of electromagnetic waves generated by the light source will be heavily influenced by the layered stack of media. Along their path the travelling light waves will be reflected and transmitted multiple times by interfaces between different media. As a result, different waves will contribute to the electromagnetic field in each layer of the stack giving rise to interference phenomena. As a general result, part of the light incident on the stack of layers will be transmitted, part of it will be absorbed and part of it will be reflected. In this chapter we will introduce a systematic approach that allows to calculate the reflection and transmission of light at such a stack of layers. Moreover, this approach can be used to calculate the electromagnetic field in all layers of the stack, even for a very high number of layers. This method uses the transfer matrices for wave propagation through layers and at interfaces and will provide an efficient toolbox for a wide variety of thin film optical problems. Throughout this chapter, we consider linear, homogeneous and isotropic media. How-
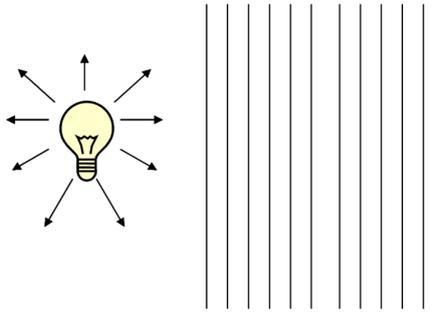
Figure 3.1: The problem of light propagation through layered media.

ever, we do not restrict ourselves to lossless media. As we will see, the transfer matrix formalism of wave propagation perfectly allows to deal with complex refractive indices.

Consider a general light source producing light with a discrete or continuous spectrum of frequencies. Given that the materials involved are linear, each frequency component of the light interacts independently with the media. Therefore it is legitimate to investigate the properties of layered media using monochromatic light, that is light with a unique frequency of oscillation, and do so for every frequency component in the spectrum. Moreover, we learned in the chapter on Fourier optics that in any plane in space each vector component of the electric (or magnetic) field $U(x, y, z)$ of a monochromatic wave can be expanded into plane waves. It follows that we can limit our investigation to the propagation of monochromatic plane waves.

The phenomenon of interference is fundamental for the understanding of the optical behaviour of thin films. This follows from the insight that multiple reflections and refractions by the different interfaces in the stack give rise to numerous field contributions that interfere with each other. Therefore, we will first treat interference before we go on with the transfer matrix formalism. We will end by investigating some examples of thin film applications.

3.2 Basics of interference

The term interference generally indicates that two (or more) phenomena are interacting with each other. Within the context of electromagnetism, this term may be a bit misleading. The reason is this: wave propagation of electromagnetic fields is described by Maxwell's equations. In linear media, these equations are linear differential equations and thus, waves will not interact with each other: the total field vector is always the sum of the individual field vectors. In other words, for a wave problem with two (or more) sources, the propagating fields do not influence each other, as the principle of superposition holds.

However, propagating waves often *seem* to interact in a non-linear way. This is because the total field depends strongly on the phase difference between the waves. In the end, this means that the total intensity is (often) not equal to the sum (or superposition) of the individual intensities. For example, if two waves with the same frequency and the same intensity come together, the total intensity can take any value between zero (destructive interference) and four times (constructive interference) the value of the intensity of the individual wave, depending on their relative phase

difference. The reason for the apparent interaction is that the quantity observed is not the amplitude of the waves but the energy density, which is a non-linear function of the electromagnetic field. Note that the concept of *phase difference* only makes sense for sources with (nearly) equal wavelengths.

3.2.1 Intensity

Any detector - be it the human eye, a photographic plate or an optical power meter - will detect the energy density $U(\mathbf{r}, t)$ associated with the total electric field $\mathbf{E}(\mathbf{r}, t)$. The response time of a detector is of course finite and therefore, a time-averaged energy density $\langle U(\mathbf{r}, t) \rangle$ will be detected, defined as the intensity of the wave:

$$I = I(\mathbf{r}, t) = \langle U(\mathbf{r}, t) \rangle = \epsilon \langle \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{E}(\mathbf{r}, t) \rangle, \tag{3.1}$$

where the time-average of the function $U(\mathbf{r}, t)$ is defined as

$$\langle U(\mathbf{r}, t) \rangle = \frac{1}{T} \int_0^T U(\mathbf{r}, t) dt. \tag{3.2}$$

The averaging is performed over a time interval T.

If two fields $\mathbf{E}_1(\mathbf{r}, t)$ and $\mathbf{E}_2(\mathbf{r}, t)$ are present, the time-averaged energy density is the result of the total field $\mathbf{E}(\mathbf{r}, t)$ which is the superposition of $\mathbf{E}_1(\mathbf{r}, t)$ and $\mathbf{E}_2(\mathbf{r}, t)$:

$$
\begin{aligned}
I = I(\mathbf{r}, t) \quad &= \langle U(\mathbf{r}, t) \rangle \\
&= \epsilon \langle (\mathbf{E}_1(\mathbf{r}, t) + \mathbf{E}_2(\mathbf{r}, t)) \cdot (\mathbf{E}_1(\mathbf{r}, t) + \mathbf{E}_2(\mathbf{r}, t)) \rangle \\
&= \epsilon \langle \mathbf{E}_1(\mathbf{r}, t) \cdot \mathbf{E}_1(\mathbf{r}, t) + \mathbf{E}_2(\mathbf{r}, t) \cdot \mathbf{E}_2(\mathbf{r}, t) + 2\mathbf{E}_1(\mathbf{r}, t) \cdot \mathbf{E}_2(\mathbf{r}, t) \rangle. \\
&= I_1 + I_2 + 2I_{12}
\end{aligned}
\tag{3.3}
$$

The first two terms in 3.3 are the time-averaged energy densities or intensities associated with each wave separately. The third term is a cross-term that gives rise to interference phenomena. The same holds when multiple waves interfere. In that case, multiple cross terms will add to the sum of the individual intensities. Before we start a more detailed discussion of the cross-term, we first investigate the aspect of polarization within the context of interference.

3.2.2 Polarization

The vectorial nature of the cross-term is of great importance. The cross-term is a dot-product of two vectors and is zero only when the two vectors have orthogonal directions. For the case of plane waves, this implies that interference occurs always except when the waves have orthogonal polarization states. This property will help us to simplify the mathematical treatment of interference problems. Consider for example a plane wave, propagating in the z-direction:

$$\mathbf{E}(z, t) = \Re[\mathbf{A}e^{+j(\omega t - kz)}]. \tag{3.4}$$

$\mathbf{A}$ is a complex vector that lies in the xy-plane and is defined as

$$\mathbf{A} = A_x e^{j\phi_x(t)}\mathbf{e}_x + A_y e^{j\phi_y(t)}\mathbf{e}_y, \tag{3.5}$$

or equivalently, the x and y-components of the field vector $\mathbf{E}$ are described by:

$$\begin{aligned}
E_x &= A_x \cos(\omega t - kz + \phi_x(t)) \\
E_y &= A_y \cos(\omega t - kz + \phi_y(t)),
\end{aligned} \tag{3.6}$$

with A_x and A_y positive numbers.

As we know from chapter 2, the phase difference $\delta(t) = \phi_x(t) - \phi_y(t)$ determines the polarization state (linear, circular, elliptic, partial, unpolarized) of the plane wave. Each of the components E_x and E_y can be regarded as linearly polarized plane waves with orthogonal directions of oscillation. And given that those waves will never interfere, they can be treated separately. In conclusion, any interference problem of waves with arbitrary polarization states can be treated by first decomposing the waves into orthogonally polarized wave components and then investigate interference between the wave components of colinear polarization (thus having parallel field vectors). This is what we will do in the next section.

3.2.3   Interference of two plane waves

Consider two monochromatic plane waves with frequencies ω_1 and ω_2, wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ and with colinear polarisation.

$$\begin{aligned}
\mathbf{E}_1(\mathbf{r}, t) &= \mathbf{A}_1 \cos(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r} + \phi_1(t)) \\
\mathbf{E}_2(\mathbf{r}, t) &= \mathbf{A}_2 \cos(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r} + \phi_2(t)).
\end{aligned} \tag{3.7}$$

Since both fields have colinear polarization, we may discard the vectorial nature of the fields and write for each of the fields ($n = 1, 2$):

$$E_n(\mathbf{r}, t) = A_n \cos(\omega_n t - \mathbf{k}_n \cdot \mathbf{r} + \phi_n(t)). \tag{3.8}$$

The intensity associated with this plane wave can be easily calculated and yields:

$$\begin{aligned}
I &= \epsilon \left\langle |A_n|^2 \cos^2\left(\omega_n t - \mathbf{k}_n \cdot \mathbf{r} + \phi_n(t)\right) \right\rangle \\
&= \tfrac{\epsilon}{2} |A_n|^2 \left\langle \{1 + \cos\left(2\left[\omega_n t - \mathbf{k}_n \cdot \mathbf{r} + \phi(t)\right]\right)\}\right\rangle \\
&= \tfrac{\epsilon}{2} |A_n|^2
\end{aligned} \tag{3.9}$$

One will get the same result when working with the complex vector notation: the intensity is directly related to the modulus squared of the complex amplitude.

We will now examine the effect of the cross term in detail. As we will be calculating energy-densities involving products of field vectors, it is safe practice to use the real field vector notation.

The intensity or time-averaged energy density thus becomes:

$$
\begin{aligned}
I \;&=\; \epsilon \left\langle E_1^2 \right\rangle + \epsilon \left\langle E_2^2 \right\rangle + 2\epsilon \left\langle E_1 E_2 \right\rangle \\
&=\; \epsilon \frac{A_1^2}{2} + \epsilon \frac{A_2^2}{2} + \epsilon A_1 A_2 \left[\langle \cos \left[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r} + \phi_1(t) - \phi_2(t) \right] \rangle \right] + \\
&\quad \epsilon A_1 A_2 \left[\langle \cos \left[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r} + \phi_1(t) + \phi_2(t) \right] \rangle \right] \\
&=\; I_1 + I_2 + 2\sqrt{I_1 I_2} \left[\langle \cos \left[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r} + \phi_1(t) - \phi_2(t) \right] \rangle \right] + \\
&\quad 2\sqrt{I_1 I_2} \left[\langle \cos \left[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r} + \phi_1(t) + \phi_2(t) \right] \rangle \right],
\end{aligned}
\tag{3.10}
$$

where I_1 and I_2 are the intensities of the individual waves.

The sum frequency term $\langle \cos \left[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r} + \phi_1(t) + \phi_2(t) \right] \rangle$ will be averaged out by every detector as $T >> \frac{1}{\omega_1 + \omega_2}$, so we find:

$$
I = I_1 + I_2 + 2\sqrt{I_1 I_2} \left\langle \cos \left[(\omega_1 - \omega_2)t - (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r} + \phi_1(t) - \phi_2(t) \right] \right\rangle.
\tag{3.11}
$$

If $\omega_1 \neq \omega_2$, the cross term will be non-zero only when $\frac{1}{\omega_1 - \omega_2} >> T$, i.e. if the two frequencies are only slightly different. The result is a time-dependent beating of $\langle U \rangle$. However, when $\frac{1}{\omega_1 - \omega_2} << T$ the last term in 3.11 will be zero.

If $\omega_1 = \omega_2$, then the phase difference between the two fields is $\mathbf{K} \cdot \mathbf{r} + \delta$ where $\mathbf{K} = \mathbf{k}_2 - \mathbf{k}_1$ and $\delta(t) = \phi_1(t) - \phi_2(t)$. If $\delta(t)$ is constant over a long enough period of time $(>> T)$, then the two waves are mutually coherent and a stationary interference pattern will be observed in space. For $A_1 = A_2 = A_0$ and thus $I_1 = I_2 = I_0$ the interference pattern is described by:

$$
\begin{aligned}
I \;&=\; 2\epsilon \frac{A_0^2}{2} \left[1 + \langle \cos(\mathbf{K} \cdot \mathbf{r} + \delta(t)) \rangle \right] \\
&=\; 4 I_0 \cos^2 \left(\frac{\mathbf{K} \cdot \mathbf{r} + \delta(t)}{2} \right) \\
&=\; 4 I_0 \cos^2 \left(\frac{\Phi}{2} \right),
\end{aligned}
\tag{3.12}
$$

where $\Phi = \mathbf{K} \cdot \mathbf{r} + \delta(t)$ and I_0 is the intensity of the individual waves. The intensity thus varies periodically in space in the direction of $\mathbf{K}$ and varies between 0 and 4 times the intensity of a single wave. The spatial period of the interference pattern is

$$
\Delta = \frac{2\pi}{|\mathbf{K}|} = \frac{\lambda_0}{2 sin(\theta/2)},
\tag{3.13}
$$

where θ is the angle between the two wave vectors and $\lambda_0 = 2\pi c/\omega$ is the vacuum wavelength of light.

Equation (3.12) indicates a strong dependence of the total intensity on the phase difference. This dependence is plotted in Figure 3.3. Also when the two waves originate from the same source, but propagate along different paths with different optical path lengths before they come together, as illustrated in figure 3.2, the detected intensity will depend on the value of the phase $\Phi = \mathbf{K} \cdot \mathbf{r} + \delta(t)$.

Assuming a path length difference (also called delay) of d, the phase equals

$$
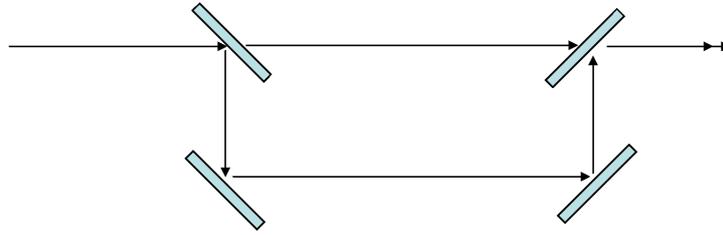\Phi = \frac{2\pi d}{\lambda} = \frac{2\pi n d}{\lambda_0}
\tag{3.14}
$$

Figure 3.2: Interference of two waves originating from the same source, after having travelled a different path length

(a) $I_1 = I_2$
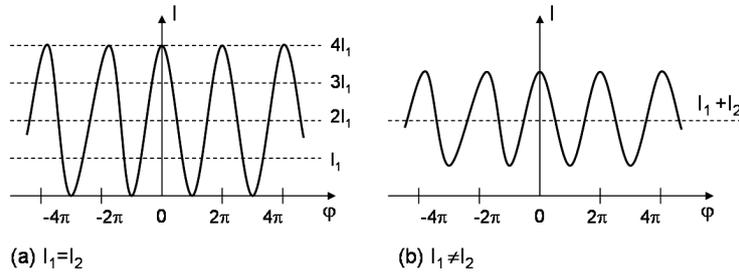
(b) $I_1 \neq I_2$

Figure 3.3: Interference of two waves, (a) $I_1 = I_2 = I_0$, (b) $I_1 \neq I_2$.

An interferometer uses the strong phase-dependence of the intensity to measure small variations of distance d, index n or wavelength λ_0 (or frequency ν). If $d/\lambda = 10^4$, then an index variation of $\delta n = 10^{-4}$ realizes a phase difference $\delta\phi = 2\pi$. Analogously, the phase changes over 2π, if d increases with a wavelength $\delta d = \lambda$. An increase of the frequency $\delta\nu = c/d$ has the same effect. A tiny change in optical path length causes a change in Φ and thus a change in detected intensity.

3.2.4 Interference of multiple plane waves

The optical properties of thin films are mostly due to interference effects of more than two waves. When multiple waves come together, the total field is given by

$$\mathbf{E}(\mathbf{r}, t) = \sum_{n=1}^{N} \mathbf{E}_n(\mathbf{r}, t) \tag{3.15}$$

and the intensity of the total field can be calculated by definition (3.1)

Considering fields of colinear polarization,

$$\mathbf{E}_n(\mathbf{r}, t) = \mathbf{A}_n \cos(\omega_n t - \mathbf{k}_n \cdot \mathbf{r} + \phi_n(t)), \tag{3.16}$$

we may discard the vectorial nature of the fields to calculate the intensity of the total field:

$$I = \sum_{n=1}^{N} I_n + \sum_{n=1}^{N} \sum_{m=1,m\neq n}^{N} \sqrt{(I_n)}\sqrt{(I_m)} \left\langle \cos\left[(\omega_n - \omega_m)t - (\mathbf{k}_n - \mathbf{k}_m)\cdot\mathbf{r} + \phi_n(t) - \phi_m(t)\right]\right\rangle, \quad (3.17)$$

where $I_n = \epsilon\frac{|\mathbf{A}_n|^2}{2}$. The first term is the sum of the individual intensities and the second term determines the interference pattern.

In this section, we will briefly examine two special cases that are regularly encountered in thin film applications. To simplify the notations and calculations, we will from now on work with the complex notation of the field amplitudes.

Interference of M plane waves with equal amplitudes and constant phase difference

Consider M plane waves $E_m = A_m e^{j\omega t}$ with complex amplitudes defined by

$$A_m = \sqrt{I_0}e^{[j(m-1)\delta]}, \quad m = 1, 2, \ldots, M. \quad (3.18)$$

The waves have equal intensity I_0 and a constant phase difference δ. We define $h = e^{j\delta}$, so that $A_m = I_0^{1/2}h^{m-1}$. The complex amplitude of the total wave becomes

$$
\begin{aligned}
A &= \sqrt{I_0}\left(1 + h + h^2 + \ldots + h^{M-1}\right) \\
&= \sqrt{I_0}\frac{1 - h^M}{1 - h} \\
&= \sqrt{I_0}\frac{1 - e^{jM\delta}}{1 - e^{j\delta}}
\end{aligned}
\quad (3.19)
$$

and the intensity is

$$I = |A|^2 = I_0\left|\frac{e^{-jM\delta/2} - e^{jM\delta/2}}{e^{-j\delta/2} - e^{j\delta/2}}\right|^2 \quad (3.20)$$

so that

$$I(\delta) = I_0\frac{sin^2\left(M\delta/2\right)}{sin^2\left(\delta/2.\right)} \quad (3.21)$$

This function is plotted in Figure 3.4 for different values of M. For M=2, the same interference pattern is found as described above. For higher values of M, the interference pattern exhibits peaks. Indeed, for $\delta \to 0$, we find that

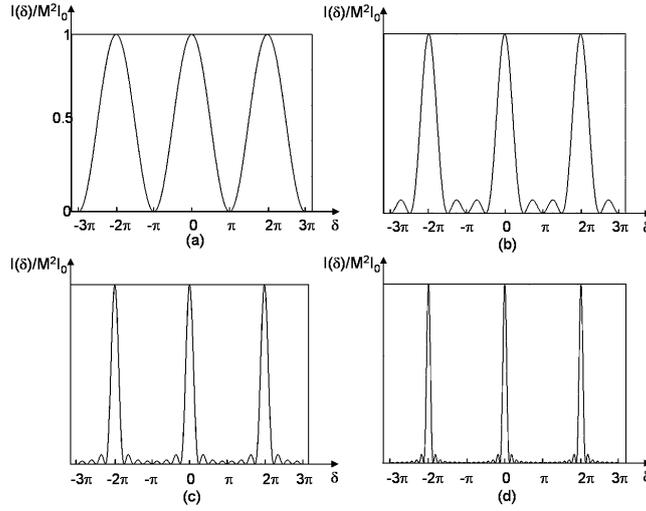$$\lim_{\delta/2\to 0} I(\delta) = I_0 M^2, \quad (3.22)$$

Figure 3.4: Interference of a finite number of plane waves with equal amplitudes and constant phase difference: intensity $\frac{I(\delta)}{M^2 I_0}$ in function of phase difference δ for (a) $M = 2$, (b) $M = 4$, (c) $M = 8$ and (d) $M = 16$.

indicating constructive interference of all the fields. This happens each time when $\delta = 2m\pi$. However, for certain values of δ, namely $\frac{2\pi}{M}, \frac{4\pi}{M}, ..., \frac{2(M-1)\pi}{M}$, $I(\delta)$ becomes 0. That is why the interference patterns exhibit $M - 1$ minima and $M - 2$ secondary maxima.

This example of interference between M waves is common in practice. Probably the most well-known case is the illumination of a screen through M slits by a plane wave. The diffracted field exhibits the behavior described above, in function of the angle.

Interference of an infinite number of waves with progressively declining amplitude and equal phase difference

Let us now consider an infinite number of plane waves with exponentially decreasing amplitude coming together. The complex amplitudes are given by:

$$A_1 = \sqrt{I_0}, \quad A_2 = hA_1, \quad A_3 = hA_2 = h^2 A_1, \quad \ldots \tag{3.23}$$

with now $h = |h| \, e^{j\delta}$ and $|h| < 1$. I_0 is again the intensity of the initial wave. The superposition of all these waves has complex amplitude

$$
\begin{aligned}
A &= A_1 + A_2 + A_3 + \ldots \\
&= \sqrt{I_0}(1 + h + h^2 + \ldots) \\
&= \frac{\sqrt{I_0}}{1 - h} \\
&= \frac{\sqrt{I_0}}{1 - |h| \, e^{j\delta}},
\end{aligned}
\tag{3.24}
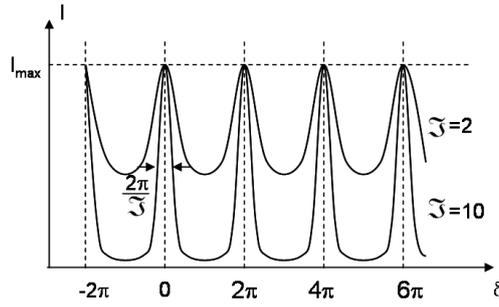$$

which can be rewritten as

Figure 3.5: Interference of an infinite number of waves with progressively declining amplitude and equal phase difference: intensity I in function of phase difference δ.

$$I = \frac{I_0}{(1 - |h|)^2 + 4\,|h|\,sin^2(\delta/2)}. \tag{3.25}$$

or

$$I = \frac{I_{max}}{1 + \left(\frac{2\mathfrak{F}}{\pi}\right)^2 sin^2\left(\frac{\delta}{2}\right)} \tag{3.26}$$

with

$$I_{max} = \frac{I_0}{(1 - |h|)^2} \tag{3.27}$$

and

$$\mathfrak{F} = \frac{\pi\,|h|^{\frac{1}{2}}}{1 - |h|} \tag{3.28}$$

a parameter called finesse.

As illustrated in figure 3.5 the intensity is a periodic function of δ with period 2π. It reaches the maximum I_{max} for $\delta = 2\pi q$, with q an integer. When the finesse $\mathfrak{F}$ is large (so r is close to one), the function I is sharply peaked. As the finesse $\mathfrak{F}$ decreases the peaks become less sharp and they disappear when $r = 0$. An important parameter associated with this interference pattern, is the so-called *Full Width at Half Maximum* (FWHM), equal to

$$\Delta\delta = \frac{2\pi}{\mathfrak{F}}. \tag{3.29}$$

The finesse $\mathfrak{F}$ is the ratio between the period 2π of the peaks and the FWHM of the transmission peaks.

This example is especially relevant in practice, in particular for the Fabry-Perot interferometer. We will come back to this structure in section 3.4.1.
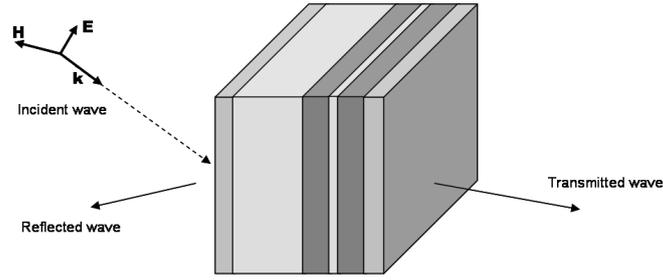
Figure 3.6: Reflection and transmission at a multilayer stack.

3.3   Transfer Matrix Formulation for Multilayer Systems

We will now investigate the general problem of wave propagation in and through a multilayer structure. Consider the dielectric structure depicted in Figure 3.6. The multilayer structure can be regarded as an optical system with one input and one output port. Thus, a transfer matrix **T** can be assigned to the system which relates the incident and reflected waves at the input port with the incident and reflected waves at the output port. If light is incident from the left, the reflected and transmitted field can thus be calculated once the transfer matrix **T** is known. In this section we will build the matrix **T** by regarding the multilayer stack as a cascaded system of interfaces and layers, each with their own transfer matrix. We get the complete transfer matrix **T** by multiplying the individual transfer matrices.

The layered medium consists of a stack of layers with thickness d_i and refractive indices n_i, separated by interface planes. A plane wave with wave vector **k** is incident on the stack, as depicted in Figure 3.6. It is convenient to choose the following coordinate system. Let the normal to the interface planes be the x-axis, then the interface planes are parallel to the (y,z)-plane and the multilayer stack (n_i, d_i) is defined as follows:

$$n(x) = \begin{array}{ll} n_0, & x < x_0, \\ n_1, & x_0 < x < x_1, \\ n_2, & x_1 < x < x_2, \\ \vdots & \\ n_N, & x_{N-1} < x, \end{array} \tag{3.30}$$

where n_i is the complex refractive index of layer i and where x_i marks the position of the planar interface between the layers i and $(i+1)$. The layers have a thickness d_i equal to $x_{i+1} - x_i$.

We now define the z-axis. The normal to the interface planes and the wave vector of the incident plane wave define a plane, which is called the incident plane. Let the z-axis be the coordinate axis in this plane orthogonal to the x-axis. As a result, the y-axis is defined and the wave vector lies entirely in the (x,z)-plane: its y-component is zero, $\mathbf{k}_y = 0$. The multilayer structure in this coordinate system is depicted in Figure 3.7 (a).

As we noted before, it is sufficient to treat the problem for two orthogonal polarizations. The easiest way is to solve the problem once for the TE-polarization (s-wave) with the electric field parallel to the interfaces of the stack and once for the TM-polarization (p-wave) with the magnetic
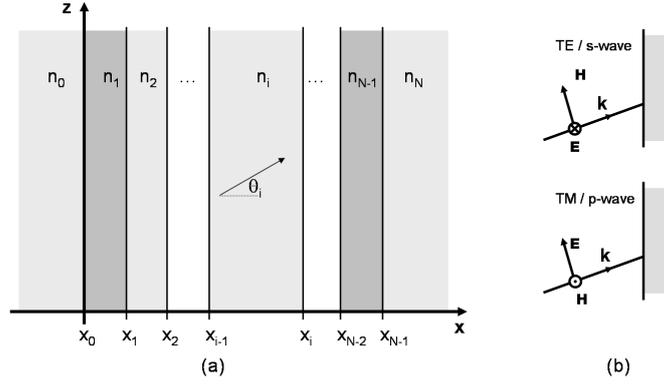
Figure 3.7: A multilayer stack of dielectric media.

field parallel to the interfaces of the stack. The orientation of both polarizations is illustrated in Figure 3.7 (b). In the following calculations, the two problems - one for each polarization - are discriminated by the polarization dependent reflection and transmission coefficients. When we consider only one polarization (TE or TM), the electric field is fully described by its amplitude $E(x, y, z)$. In each layer, all contributions to the forward propagating wave (towards increasing x) have the same direction and thus form one plane wave. The same holds for the backward propagating wave (towards decreasing x). The total field in layer i can thus be described as

$$
\begin{aligned}
E(x, y, z) &= A_F e^{-j(k_{x,i}x + k_z z)} + A_B e^{-j(-k_{x,i}x + k_z z)} \\
&= A_F e^{-jk_{x,i}x} e^{-jk_z z} + A_B e^{+jk_{x,i}x} e^{-jk_z z} \\
&= E_F(x) e^{-jk_z z} + E_B(x) e^{-jk_z z}
\end{aligned}
\tag{3.31}
$$

where $k_{x,i}$ is the x-component of the wave vector $\mathbf{k}_i$ in layer i:

$$
k_{x,i} = \left[(n_i k_0)^2 - k_z^2 \right]^{1/2}, i = 0, 1, 2, ..., N,
\tag{3.32}
$$

where $k_0 = (\omega/c)$. Due to boundary conditions that relate the field amplitudes of the incident, reflected and transmitted wave at the interfaces, k_z remains constant. When working with lossless media and in the absence of total internal reflection, $k_{x,i}$ can be related to the ray angle θ_i (see Figure 3.8) in the following way:

$$
k_{x,i} = n_i k_0 \cos \theta_i.
\tag{3.33}
$$

However, it should be noted that $k_{x,i}$ is a complex number and that its imaginary part can be nonzero. This will be the case when $\Im(n_i)$ is nonzero or when $k_z > n_i k_0$. The former corresponds to wave propagation in a lossy medium, the latter corresponds to incident angles that are bigger than the critical angle for total internal reflection. When total internal reflection occurs, the electric field vector decreases exponentially in the x-direction and the attenuation occurs within a distance of q^{-1} where (if we choose $+j$ as the solution for $\sqrt{-1}$):

$$
q_i = \left[k_z^2 - n_i k_0^2 \right]^{1/2}, i = 0, 1, 2, ..., N,
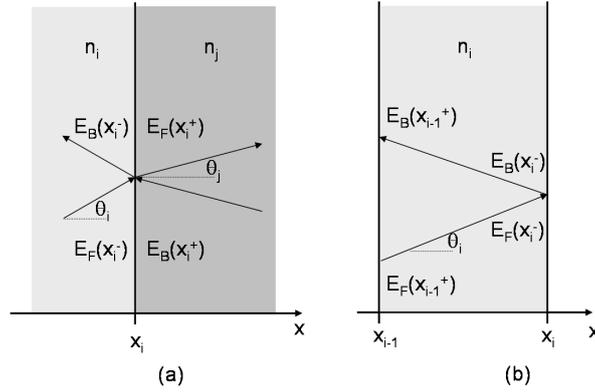\tag{3.34}
$$

Figure 3.8: (a) Reflection and transmission at an interface. (b) Translation through a layer with thickness d_i and refractive index n_i.

In that case, the plane wave 3.31 is an evanescent wave propagating parallel to the interface surfaces in the z-direction.

For a given z, the complete transfer matrix **T** for wave propagation relates the complex amplitudes E_F and E_B just before and just behind the first and the last interface of the multilayer stack.

$$\begin{bmatrix} E_F(x_0^-) \\ E_B(x_0^-) \end{bmatrix} = \mathbf{T}_{0N} \begin{bmatrix} E_F(x_{N-1}^+) \\ E_B(x_{N-1}^+) \end{bmatrix} = \begin{bmatrix} T_{11}^{0N} & T_{12}^{0N} \\ T_{21}^{0N} & T_{22}^{0N} \end{bmatrix} \begin{bmatrix} E_F(x_{N-1}^+) \\ E_B(x_{N-1}^+) \end{bmatrix}. \tag{3.35}$$

We first build the individual transfer matrices for wave propagation through an interface and for wave propagation through a layer and will then calculate the complete transfer matrix by multiplying all individual ones. We start with the transfer matrix for wave propagation through an interface. Consider the interface between layer i and layer j (Figure 3.8 (a)). For a given z, the amplitudes of the forward and backward propagating waves just before and just behind the interface are related in the following way:

$$\begin{bmatrix} E_F(x_i^+) \\ E_B(x_i^-) \end{bmatrix} = \begin{bmatrix} t_{ij} & r_{ji} \\ r_{ij} & t_{ji} \end{bmatrix} \begin{bmatrix} E_F(x_i^-) \\ E_B(x_i^+) \end{bmatrix} \tag{3.36}$$

where we use the complex reflection and transmission coefficients (Fresnel coefficients) for the situation where only one wave is incident on the interface. This is actually the scattering matrix description of the interface between two layers.

In the case of TE-polarization, the Fresnel coefficients are given by:

$$r_{ij} = \frac{E_B(x_i^-)}{E_F(x_i^-)} = \frac{k_x^i - k_x^j}{k_x^i + k_x^j} \tag{3.37}$$

$$t_{ij} = \frac{E_F(x_i^+)}{E_F(x_i^-)} = 1 + r_{ij} = \frac{2k_x^i}{k_x^i + k_x^j} \tag{3.38}$$

and in the case of TM-polarization:

$$r_{ij} = \frac{E_B(x_i^-)}{E_F(x_i^-)} = \frac{n_i^2 k_x^j - n_j^2 k_x^i}{n_i^2 k_x^j + n_j^2 k_x^i} \tag{3.39}$$

$$t_{ij} = \frac{E_F(x_i^+)}{E_F(x_i^-)} = \frac{n_i}{n_j}(1 + r_{ij}) \tag{3.40}$$

or their equivalent as a function of the refractive indices n_i, n_j and ray angles θ_i, θ_j, for TE-polarization:

$$r_{ij} = \frac{E_B(x_i^-)}{E_F(x_i^-)} = \frac{n_i \cos\theta_i - n_j \cos\theta_j}{n_i \cos\theta_i + n_j \cos\theta_j} \tag{3.41}$$

$$t_{ij} = \frac{E_F(x_i^+)}{E_F(x_i^-)} = 1 + r_{ij} = \frac{2n_i \cos\theta_i}{n_i \cos\theta_i + n_j \cos\theta_j} \tag{3.42}$$

and for TM-polarization:

$$r_{ij} = \frac{E_B(x_i^-)}{E_F(x_i^-)} = \frac{n_j \cos\theta_i - n_i \cos\theta_j}{n_j \cos\theta_i + n_i \cos\theta_j} \tag{3.43}$$

$$t_{ij} = \frac{E_F(x_i^+)}{E_F(x_i^-)} = \frac{n_i}{n_j}(1 + r_{ij}) = \frac{2n_i \cos\theta_i}{n_j \cos\theta_j + n_i \cos\theta_j} \tag{3.44}$$

For perpendicular incidence, there is no difference between the TE and TM case. Equation (3.37) differs however from equation (3.39) in this case. This is caused by the different definition of the direction of the unit vectors for the E-field, in the TE and TM case. Figure 3.9 depicts the definition of the unit vectors for the E and H-field for the incident, reflected and refracted wave.

Equations (3.36) can be rewritten in such a way that a relationship is described between the forward and backward propagating wave in layer i and the forward and backward propagating wave in layer j:

$$\begin{bmatrix} E_F(x_i^-) \\ E_B(x_i^-) \end{bmatrix} = \begin{bmatrix} \frac{1}{t_{ij}} & -\frac{r_{ji}}{t_{ij}} \\ \frac{r_{ij}}{t_{ij}} & t_{ji} - \frac{r_{ij}}{t_{ij}}r_{ji} \end{bmatrix} \begin{bmatrix} E_F(x_i^+) \\ E_B(x_i^+) \end{bmatrix}. \tag{3.45}$$

We make use of the symmetry relations of the Fresnel coefficients

$$r_{ij} = -r_{ji} \tag{3.46}$$

$$t_{ij}t_{ji} - r_{ij}r_{ji} = 1 \tag{3.47}$$

to simplify this expression and get:

$$\begin{bmatrix} E_F(x_i^-) \\ E_B(x_i^-) \end{bmatrix} = \frac{1}{t_{ij}} \begin{bmatrix} 1 & r_{ij} \\ r_{ij} & 1 \end{bmatrix} \begin{bmatrix} E_F(x_i^+) \\ E_B(x_i^+) \end{bmatrix}. \tag{3.48}$$
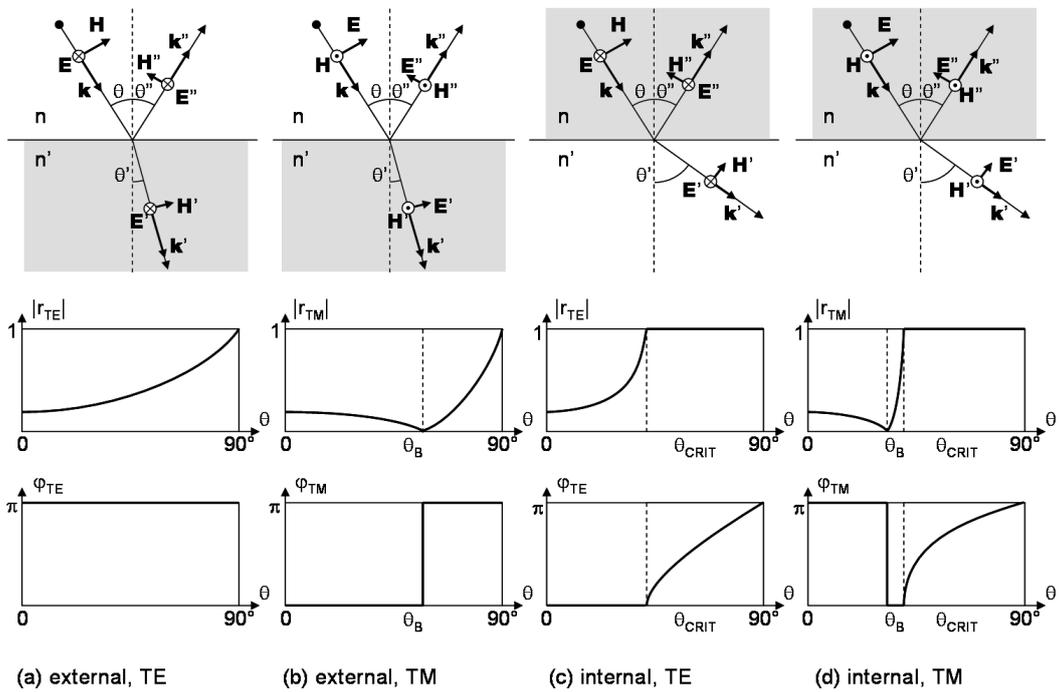
Figure 3.9: Magnitude and phase of the reflection coefficient in function of incidence angle for (a) external reflection ($n_j/n_i = 1.5$) and TE polarization, (b) external reflection ($n_j/n_i = 1.5$) and TM polarization, (c) internal reflection ($n_i/n_j = 1.5$) and TE polarization and (d) internal reflection ($n_i/n_j = 1.5$) and TM polarization.
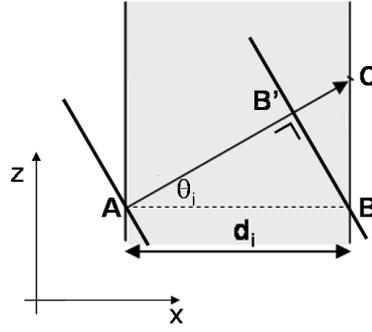
Figure 3.10: Inspection of Equation (3.52). The phase change Φ corresponds to the path length $|AB'|$ and not to the path length $|AC|$.

Thus, the transfer matrix $\mathbf{T}_{ij}$ for wave propagation through the interface between layers i and j reads:

$$\mathbf{T}_{ij} = \frac{1}{t_{ij}} \begin{bmatrix} 1 & r_{ij} \\ r_{ij} & 1 \end{bmatrix}.$$ (3.49)

The next step is to build the transfer matrix for wave propagation through a layer. Consider the layer i. (Figure 3.8 (b)). For a given z, the amplitudes of the forward and backward propagating waves just before and just behind the interfaces of the adjacent layers are related in the following way:

$$\begin{aligned} E_F(x_i^-) &= E_F(x_{i-1}^+)e^{-jk_{x,i}d_i} \\ E_B(x_{i-1}^+) &= E_B(x_i^-)e^{-jk_{x,i}d_i}, \end{aligned}$$ (3.50)

with d_i the thickness of layer i. This follows straightforwardly from Equations (3.31). Thus, the transfer matrix $\mathbf{T}_i$ for wave propagation through layer i reads:

$$\mathbf{T}_i = \begin{bmatrix} e^{j\Phi_i} & 0 \\ 0 & e^{-j\Phi_i} \end{bmatrix}$$ (3.51)

in which we defined $\Phi_i = k_{x,i}d_i$, which in general is a complex quantity. In a lossless medium and in the absence of total internal reflection, Φ_i is a real quantity:

$$\Phi_i = k_{x,i}d_i = \frac{2\pi}{\lambda_0}n_i d_i \cos\theta_i,$$ (3.52)

where $k_{x,i}$ was substituted using (3.33). Equation (3.52) might look surprising at first sight. Clearly, Φ_i represents a phase change but the phase change does not correspond to the path length $|AC|$. Careful inspection of Equation (3.52) shows that the phase change Φ corresponds to the path length $|AB'|$. The reason for this is that in the transfer matrix formalism, the transfer matrices always describe relationships between E_F and E_B at a constant z-level. Thus Φ corresponds to the distance between the two phase fronts through points A and B, equal to $|AB'|$.

The complete transfer matrix for the wave propagation through the layered medium can now be found by multiplication of all individual transfer matrices.

$$\begin{bmatrix} E_F(x_0^-) \\ E_B(x_0^-) \end{bmatrix} = \mathbf{T}_{0N} \begin{bmatrix} E_F(x_{N-1}^+) \\ E_B(x_{N-1}^+) \end{bmatrix} = \begin{bmatrix} T_{11}^{0N} & T_{12}^{0N} \\ T_{21}^{0N} & T_{22}^{0N} \end{bmatrix} \begin{bmatrix} E_F(x_{N-1}^+) \\ E_B(x_{N-1}^+) \end{bmatrix} \tag{3.53}$$

where

$$\mathbf{T}_{0N} = \mathbf{T}_{01}\mathbf{T}_1\mathbf{T}_{12}\mathbf{T}_2...\mathbf{T}_{(N-1)}\mathbf{T}_{(N-1)N}. \tag{3.54}$$

Equation (3.53) is known as the matrix formulation for wave propagation through multilayer systems. Now, this expression can be used for solving a variety of wave propagation problems. Let us first apply the transfer matrix method to our original problem: a plane wave is incident on a stack of layers. First calculate the individual transfer matrices according to the described method (equations (3.49) and (3.51)). Next, calculate the complete transfer matrix by multiplication using Equation (3.53). As there is no incident field from the right side, we have $E_B(x_{N-1}^+) = 0$. So, we are left with the following matrix formulation for wave propagation in the multilayer stack:

$$\begin{bmatrix} E_F(x_0^-) \\ E_B(x_0^-) \end{bmatrix} = \begin{bmatrix} T_{11}^{0N} & T_{12}^{0N} \\ T_{21}^{0N} & T_{22}^{0N} \end{bmatrix} \begin{bmatrix} E_F(x_{N-1}^+) \\ 0 \end{bmatrix}. \tag{3.55}$$

This leads to two equations and two unknowns, namely $E_B(x_0^-)$ and $E_F(x_{N-1}^+)$, which can now easily be solved for. The reflection and transmission coefficients follow from the ratios $E_B(x_0^-)/E_F(x_0^-)$ and $E_F(x_{N-1}^+)/E_F(x_0^-)$ respectively.

$$\begin{aligned} r &= \frac{E_B(x_0^-)}{E_F(x_0^-)} = \frac{T_{21}^{0N}}{T_{11}^{0N}} \\ t &= \frac{E_F(x_{N-1}^+)}{E_F(x_0^-)} = \frac{1}{T_{11}^{0N}} \end{aligned} \tag{3.56}$$

From the amplitude reflection and transmission coefficients, the power reflection and transmission coefficients can be easily calculated. The power density of a (forward or backward) plane wave in layer i is given by the real part of the Poynting vector:

$$\begin{aligned} \langle P(x) \rangle &= Re(\mathbf{S}(x)) \\ &= Re(\mathbf{E}(x) \times \mathbf{H}(x)) \\ &= Re(n)\frac{|\mathbf{E}(x)|^2}{2Z_0} \end{aligned} \tag{3.57}$$

with Z_0 the impedance of vacuum (= $\sqrt{\mu_0/\epsilon_0}$= 377Ω). This power expresses the power per unit area perpendicular to the propagation direction (which differs from layer to layer). In order to calculate the power reflection and transmission of the multilayer thin film, it is better to express the power density per unit of area parallel to the layers. This modified power density is given by:

$$\begin{aligned} \langle P(x_i)_{||} \rangle &= \cos\theta_i \langle P(x_i) \rangle \\ &= Re(n)\cos\theta_i \frac{|\mathbf{E}(x_i)|^2}{2Z_0} \end{aligned} \tag{3.58}$$

This allows to calculate the power reflection and transmission coefficients of the thin film:

$$
\begin{aligned}
R &= \frac{Re(n_0) \ \cos\theta_0 \ |\mathbf{E}_B(x_0)|^2}{Re(n_0) \ \cos\theta_0 \ |\mathbf{E}_F(x_0)|^2} = |r|^2 \\
T &= \frac{Re(n_N) \ \cos\theta_N \ |\mathbf{E}_F(x_N)|^2}{Re(n_0) \ \cos\theta_0 \ |\mathbf{E}_F(x_0)|^2} = \frac{Re(n_N) \ \cos\theta_N}{Re(n_0) \ \cos\theta_0} \ |t|^2 \ .
\end{aligned}
\tag{3.59}
$$

To conclude our discussion of the transfer matrix formalism and to show its broad applicability, let us consider a different problem. What if no light is incident, nor from the left nor from the right? This means: $E_F(x_0^-) = 0$ and $E_B(x_{N-1}^+) = 0$. In this case, we are left with the following matrix formulation for wave propagation in the multilayer stack:

$$
\left[\begin{array}{c} 0 \\ E_B(x_0^-) \end{array} \right] = \left[\begin{array}{cc} T_{11}^{0N} & T_{12}^{0N} \\ T_{21}^{0N} & T_{22}^{0N} \end{array} \right] \left[\begin{array}{c} E_F(x_{N-1}^+) \\ 0 \end{array} \right].
\tag{3.60}
$$

So, even if no light is incident, this system has nonzero solutions provided that the following condition is fulfilled:

$$
T_{11}^{0N} = 0.
\tag{3.61}
$$

Let us briefly examine this condition. T_{11}^{0N} is a function of k_z, ω, indices n_i and thicknesses t_i of the layers. Given a multilayer structure (n_i, d_i) and a frequency ω, Equation (3.61) can thus be used to solve for the propagation constant k_z of all the confined modes supported by the multilayer structure. This example is an illustration of the application of the transfer matrix formalism to waveguide problems. We will come back to the transfer matrix method in a later chapter where it will be used to analyze slab waveguides.

3.4 Applications of thin films

In this section we discuss two important applications of thin film structures: the Fabry-Perot etalon and optical coatings.

3.4.1 Fabry-Perot etalon

The Fabry-Perot interferometer, or etalon, can be considered as the simplest type of optical resonator. Normally, such an instrument consists of two parallel dielectric mirrors separated at a distance l. Here, we consider a simple structure that consists of a plane-parallel plate of thickness l and refractive index n immersed in a medium of index n' (Figure 3.11). This is a simple multilayer stack and the transfer matrix formalism can be applied in a straightforward way. 3 transfer matrices need to be calculated and multiplied in order to get the complete transfer matrix of the 3-layer system. First, we calculate the transfer matrices associated with the two interfaces:

$$
\mathbf{T}_{12} = \frac{1}{t_{12}} \left[\begin{array}{cc} 1 & r_{12} \\ r_{12} & 1 \end{array} \right] = \frac{1}{t_{12}} \left[\begin{array}{cc} 1 & r \\ r & 1 \end{array} \right].
\tag{3.62}
$$

and

$$\mathbf{T}_{23} = \frac{1}{t_{23}} \begin{bmatrix} 1 & r_{23} \\ r_{23} & 1 \end{bmatrix} = \frac{1}{t_{23}} \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix}. \tag{3.63}$$

where $r = r_{12} = -r_{23}$. The transfer matrix associated to the layer between the two interfaces is the following:

$$\mathbf{T}_2 = \begin{bmatrix} e^{j\Phi} & 0 \\ 0 & e^{-j\Phi} \end{bmatrix} = e^{j\Phi} \begin{bmatrix} 1 & 0 \\ 0 & e^{-2j\Phi} \end{bmatrix} \tag{3.64}$$

where $\Phi = \frac{2\pi}{\lambda_0} nl \cos\theta$. Multiplying yields the complete transfer matrix:

$$\mathbf{T}_{13} = \mathbf{T} = \mathbf{T}_{12}\mathbf{T}_2\mathbf{T}_{23} = \frac{1}{t_{12}t_{23}} e^{j\Phi} \begin{bmatrix} 1 - r^2 e^{-2j\Phi} & -r(1 - e^{-2j\Phi}) \\ r(1 - e^{-2j\Phi}) & -r^2 + e^{-2j\Phi} \end{bmatrix}. \tag{3.65}$$

When such a Fabry-Perot etalon is illuminated by a light beam, the situation corresponds to one incident wave from the left of the structure. Thus, the reflection and the transmission of the structure are described by the formulas in Equation (3.55). Substitution and simplification yields for the reflection and transmission coefficients r_{FP} and t_{FP} according to 3.56:

$$\begin{aligned} r_{FP} &= \frac{T_{21}}{T_{11}} = \frac{r(1 - e^{-2j\Phi})}{1 - r^2 e^{-2j\Phi}} \\ t_{FP} &= \frac{1}{T_{11}} = \frac{t_{12}t_{23}e^{-j\Phi}}{1 - r^2 e^{-2j\Phi}}, \end{aligned} \tag{3.66}$$

and for the power reflection and transmission coefficients R_{FP} and T_{FP} according to 3.59:

$$\begin{aligned} R_{FP} &= \left| \frac{T_{21}}{T_{11}} \right|^2 = \frac{4r^2 \sin^2 \Phi}{(1 - r^2)^2 + 4r^2 \sin^2 \Phi} \\ T_{FP} &= \left| \frac{1}{T_{11}} \right|^2 = \frac{|t_{12}t_{23}|^2}{(1 - r^2)^2 + 4r^2 \sin^2 \Phi}, \end{aligned} \tag{3.67}$$

describing the reflection and transmission of a Fabry-Perot etalon. Note that our basic model of the Fabry-Perot etalon does not contain any loss mechanisms, so conservation of energy requires $R_{FP} + T_{FP} = 1$, which is indeed the case. In what follows, we will work with the fraction of the intensity reflected R and the fraction of the intensity transmitted T at one interface and refer to it as the mirror's reflection and transmittance:

$$\begin{aligned} R &= r_{12}^2 = r_{23}^2 = r^2 \\ T &= t_{12}t_{23}. \end{aligned} \tag{3.68}$$

For the moment, we consider no losses, so $R + T = 1$. Let us now discuss the results of our calculations given by 3.67. First of all, the formula for the transmitted intensity (3.67) looks very similar to the formula (3.25) we found in section 3.2.4 describing the interference pattern of an infinite number of plane waves with progressively declining amplitude and equal phase difference. In fact, they are the same. This is not surprising at all. It is clear from Figure 3.11 that the transmission of a Fabry-Perot etalon is resulting from an infinite number of waves interfering with each other. If the incident intensity is taken as unity, the first transmitted contribution has intensity $|t_{12}t_{23}|$. The second contribution is decreased by a factor $|r_{12}r_{23}|$, the third contribution again, and so on. The phase difference between each contribution remains constant and equal to $\delta = 2\Phi = 2\frac{2\pi}{\lambda_0} nl \cos\theta$.
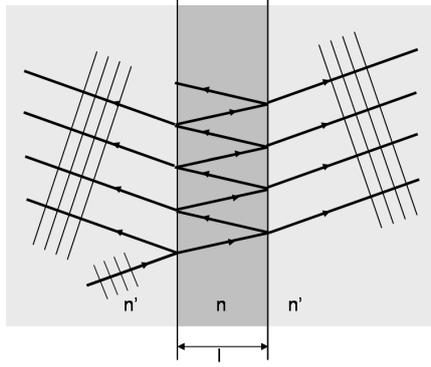
Figure 3.11: Reflection and transmission at a parallel plate structure such as the Fabry-Perot etalon.

Thus, from the viewpoint of interference, we end up with exactly the formula 3.25 for the tranmission through this structure. The transfer matrix formalism for wave propagation gives the same result, as it should be.

Let us now take a closer look at the transmission characteristics of a Fabry-Perot etalon. According to Equation (3.67), transmission is maximal and equal to unity when

$$\Phi = \frac{2\pi n}{\lambda} l \cos\theta = m\pi, m = integer, \tag{3.69}$$

where θ is the ray angle in the medium. Φ is called the resonator round trip phase. On the other hand, the transmission is minimal and equal to

$$T_{FP,min} = \frac{(1-R)^2}{(1-R)^2 + 4R} = \frac{1}{1 + \frac{4R}{(1-R)^2}} \tag{3.70}$$

when the resonator round trip phase equals

$$\Phi = \frac{2\pi n}{\lambda} l \cos\theta = (2m-1)\frac{\pi}{2}, m = integer. \tag{3.71}$$

By using $\lambda = c/\nu$, the condition for maximal transmission can also be written as

$$\nu_m = m\frac{c}{2nl\cos\theta}, m = integer, \tag{3.72}$$

where ν is the optical frequency. For a given l and n, Equation (3.72) defines the resonance frequencies of the Fabry-Perot etalon. At resonance, all transmitted contributions interfere constructively. Off resonance, the contributions do no longer interfere constructively. Two neighboring resonance frequencies are separated by the so-called free spectral range:

$$\Delta\nu = \nu_{m+1} - \nu_m = \frac{c}{2nl\cos\theta}, m = integer. \tag{3.73}$$

At this point, it is interesting to look back at Figure 3.5 depicting the intensity of an infinite number of interfering waves with declining amplitude and equal phase difference. It should be clear now that this figure also depicts the transmission of a Fabry-Perot cavity as a function of phase, see

Fig. 3.12. However, note that the Fabry-Perot peaks are at the positions $\Phi = m\pi$, which correspond to $\delta = 2m\pi$ in Fig. 3.5.

The finesse $\mathfrak{F}$ is directly related to the mirror's reflection R at each interface by

$$\mathfrak{F} = \frac{\pi\sqrt{R}}{1-R}. \tag{3.74}$$

Thus, the higher the mirror's reflection R at each interface, the higher is the finesse and the sharper are the transmission peaks:

$$
\begin{array}{lll}
R = 50\% & \to & \mathfrak{F} \cong 4 \\
R = 90\% & \to & \mathfrak{F} = 30 \\
R = 98\% & \to & \mathfrak{F} = 156
\end{array}
\tag{3.75}
$$

Another important dimensionless quantity in connection with resonances is the *quality factor* or Q-factor. It is defined as

$$Q = \frac{\omega_r}{\Delta\omega} \tag{3.76}$$

with ω_r the (angular) resonant frequency, and $\Delta\omega$ the FWHM-bandwidth (Full Width at Half Maximum). It is a measure of the sharpness of the resonance in the spectrum. We see in Fig. 3.12 e.g. that a Fabry-Perot with strongly reflecting mirrors gives sharper peaks with a larger Q, intuitively indicating that the resonance is 'better defined' or 'purer'.

Now we investigate the phase characteristics of the Fabry-Perot etalon. If we write the amplitude transmission coefficient 3.66 as

$$t = |t|\, e^{-j\angle t}, \tag{3.77}$$

and examine the phase shift $\angle t$ as a function of the cavity length l (or equivalently Φ), we note that a strong dispersion of the phases versus Φ exists at resonance. This is shown in Figure 3.12. Note that there is a link to the concept of *group delay*, defined as

$$\frac{\mathrm{d}\angle t}{\mathrm{d}\omega} \tag{3.78}$$

which measures the propagation time of a signal through the structure. We clearly note an increase of the group delay at resonance, which corresponds to the intuition that light then spends a longer time in the cavity. Remark that without the cavity structure, the phase would be a straight line ($\angle t = k_0 nd$) in Figure 3.12.

When designing a Fabry-Perot resonator with high finesse, one needs to ensure that $T_{FP,\min}$ given by Equation (3.70) is as small as possible. Therefore R needs to be close to 1. In practice, this is difficult because of the available materials. Indeed, the refractive index of optical materials is limited to about $n \approx 4$, which means for a Fabry-Perot in air a maximal reflection of only $R \approx 36\%$.

One solution to this problem is offered by applying metal coatings on the interfaces. Due to the presence of such a metal coating, an additional phase difference will occur upon reflection. So assume:

$$r = \sqrt{R}e^{+j\alpha}. \tag{3.79}$$

With this r one can calculate for the transmission

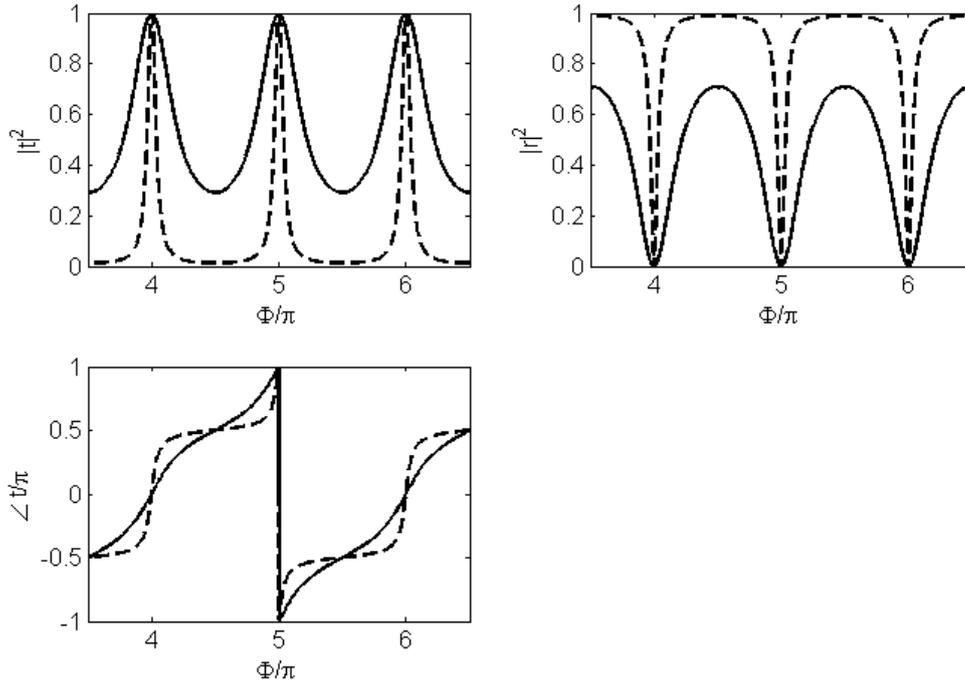$$T_{FP} = \left| \frac{T}{1 - Re^{-j2(\phi-\alpha)}} \right|^2 \tag{3.80}$$

Figure 3.12: Power transmission/reflection and phase of transmission for a Fabry-Perot etalon as a function of cavity length l (or equivalently Φ). Solid line: $R = 30\%$, dashed line: $R = 80\%$.

where T is no longer equal to $1 - R$ but to $1 - R - A$ with A the intensity absorption in the metal film. As a result, the maximal transmission drops from unity to

$$T_{FP,\max} = \frac{T^2}{(1-R)^2} = \left[1 - \frac{A}{1-R}\right]^2 . \tag{3.81}$$

The relation between the ratio $\frac{T_{FP,\max}}{T_{FP,\min}}$ and the mirror's reflection R is depicted in Figure 3.13.

To provide more insight into the operation of the Fabry-Perot cavities we plot the energy in and around the cavities in various situations, see Figure 3.14. One notices that at resonance the energy in the cavities is enhanced, and the transmission is unity. Off resonance (middle graphs) we find a standing wave in reflection, and a small field in the cavity and also a small transmission.

Most Fabry-Perot etalons are made of two identical mirrors. In fact, the symmetry of the mirror reflection is important to obtain high finesse. Any asymmetry in the reflection will lead to a decrease in either transmission or finesse. (It is left as an exercise to investigate the transmission properties of an asymmetric Fabry-Perot etalon in detail.)

3.4.2 Coatings

Multilayer stacks can be used as coatings that alter the optical properties of a substrate, for example increasing or decreasing reflection. Anti-reflective coatings and highly-reflective coatings are two examples that are widely used.
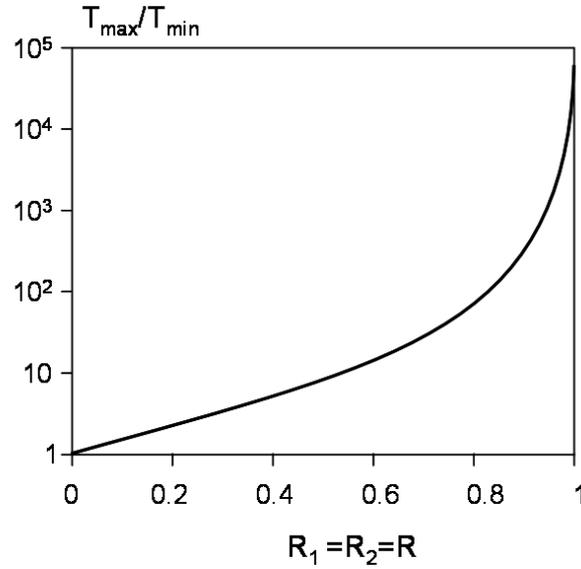
Figure 3.13: $\frac{T_{FP,\max}}{T_{FP,\min}}$ as a function of R.

AR-coatings: quarter-wave layer

Minimal reflection or equivalently maximal transmission is desirable in many applications. One example is efficient coupling of light from a fibre into a waveguide. In designing an anti-reflective-coating one needs to ensure that the reflection at the front of the coating interferes destructively with the reflection at the back of the coating. Assume a film thickness d. If $n_1 < n_2 < n_3$ with n_2 the refractive index of the coating, the two waves will interfere destructively if

$$d = \frac{1}{4}\frac{\lambda_0}{n_2}, \tag{3.82}$$

hence the name quarter-wave layer. The question remains which refractive index is needed for the coating. From the transfer matrix formalism for this 3-layer structure, one easily obtains

$$r = \frac{r_{12} + r_{23}e^{-2j\Phi}}{1 + r_{12}r_{23}e^{-2j\Phi}} = \frac{r_{12} - r_{23}}{1 - r_{12}r_{23}}, \tag{3.83}$$

where we used $\Phi = \frac{2\pi}{\lambda_0}n_2 d = \frac{\pi}{2}$. Given that $r_{ij} = \frac{n_i - n_j}{n_i + n_j}$, the reflection will be zero when

$$n_2 = \sqrt{n_1 n_3}. \tag{3.84}$$

In practice however, materials with such a refractive index may not exist. Nevertheless, using available materials with an index of refraction close to that given by Equation (3.84), a great reduction in reflection is obtained. We also notice that by using materials with an index of refraction between n_1 and n_3, a single-layer coating will always reduce the reflection, regardless of the layer thickness. Thus, the reflection of a coated surface is always lower than that of an uncoated one,
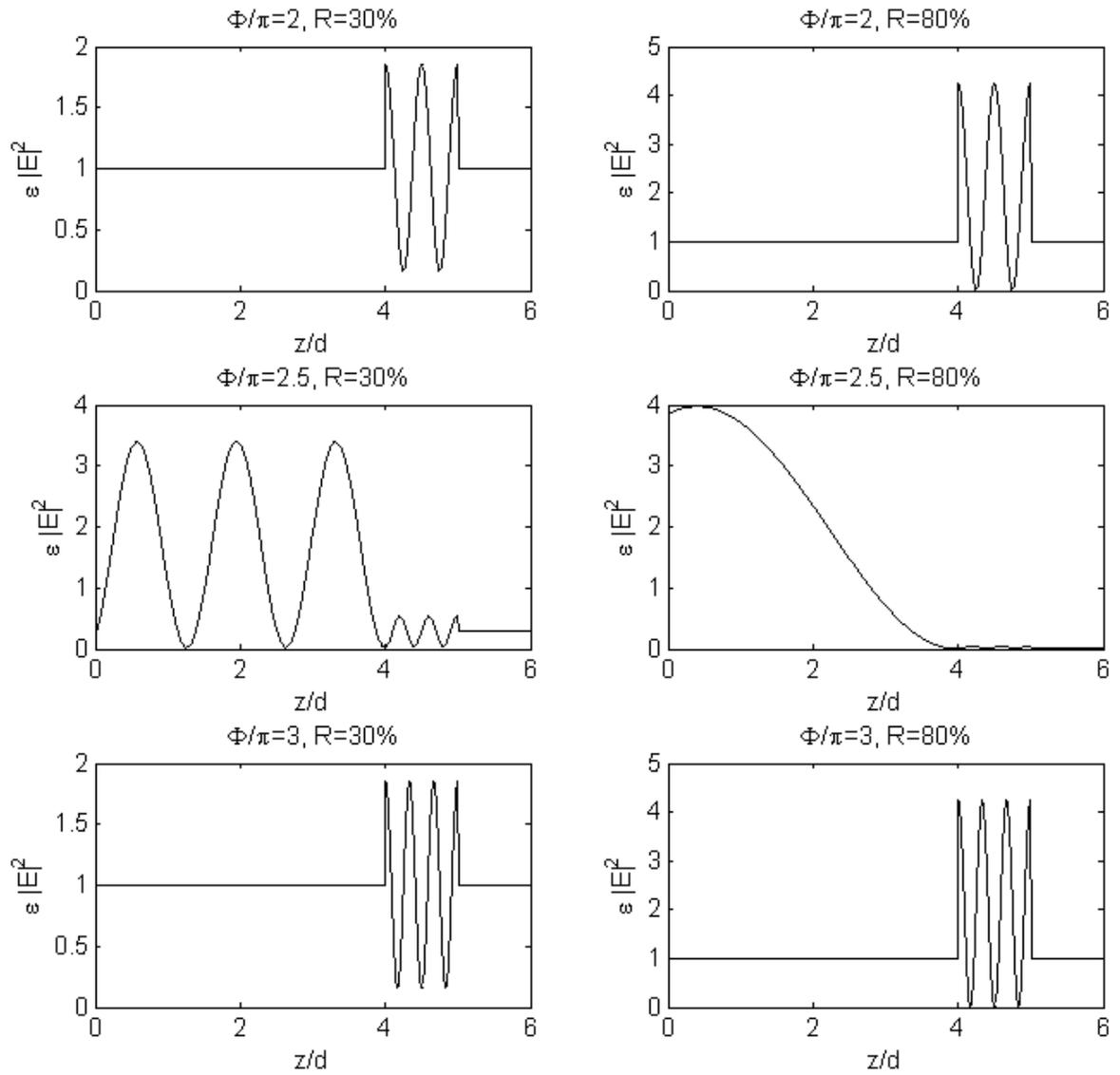
3–22

Figure 3.14: Energy ($\epsilon|E|^2$) in the Fabry-Perot cavities for $R = 30\%$ (left) and $R = 80\%$ (right). The situation is shown for $\Phi/\pi = 2$ (top), $\Phi/\pi = 2.5$ (middle) and $\Phi/\pi = 3$ (bottom). The cavity is located between $z = 4d$ and $z = 5d$. Light is incident from the left with $|E|^2 = 1$.
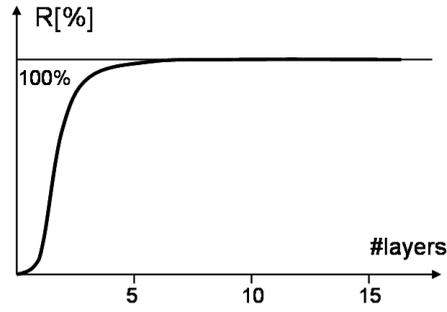
Figure 3.15: Reflection of a HR-coating for a He-Ne laser at wavelength λ_0 =633nm. n_1 =2.32 (ZnS), n_2=1.38 (MgF$_2$). A reflection of 98.9 % is achieved already after 13 layers.

provided that the index of refraction of the coating is between that of the two media. In some sense, the coating smooths dielectric discontinuity.

High-reflective coatings

High-reflective mirrors are desirable in many applications. These include high-finesse Fabry-Perot interferometers and low-loss laser resonators. Mirrors made of metallic films such as silver, aluminum or gold are generally of high reflection. For example, a silver mirror can achieve reflection approaching 99 % in the visible spectrum. Approximately 1 % of light energy penetrates the surface of the metal and gets absorbed in the bulk of the metal. These metallic mirrors cannot be used with high-power lasers because even a small fraction of absorption can cause severe heating problems. Thus there is a need to design high-reflection mirrors by using materials that have (almost) no absorption.

The dielectric layered structure that consists of alternating quarter-wave layers of two different materials is the simplest way to obtain high reflection. This is the so-called Bragg reflector. If for a certain wavelength λ_0, the thicknesses d_1, d_2 and the refractive indices n_1 and n_2 of the consecutive layers can be controlled so that:

$$n_1 d_1 = n_2 d_2 = \frac{\lambda_0}{4}, \tag{3.85}$$

then the reflected beams from the different interfaces will all interfere constructively, leading to a peak in the reflection spectrum for this wavelength.

Using the matrix method, the peak reflection $R_{HR,\mathrm{max}}$ can be calculated and is given by

$$R_{HR,\mathrm{max}} = \left(\frac{1 - \left(\frac{n_s}{n_a} \right) \left(\frac{n_2}{n_1} \right)^{2N}}{1 + \left(\frac{n_s}{n_a} \right) \left(\frac{n_2}{n_1} \right)^{2N}} \right)^2, \tag{3.86}$$

where n_s is the refractive index of the substrate, n_a that of air and N is the number of periods. $R_{HR,\mathrm{max}}$ converges to 1 as N increases. The convergence improves as the ratio $\frac{n_1}{n_2}$ becomes larger. This is illustrated in Figure 3.15 for the case of a high reflective coating for a He-Ne laser at wavelength $\lambda_0 = 633nm$ and using a quarter-waves stack of ZnS (n_1 =2.32) and MgF (n_2=1.38). However, Bragg reflectors can provide high reflection over any desired spectral regime of interest by

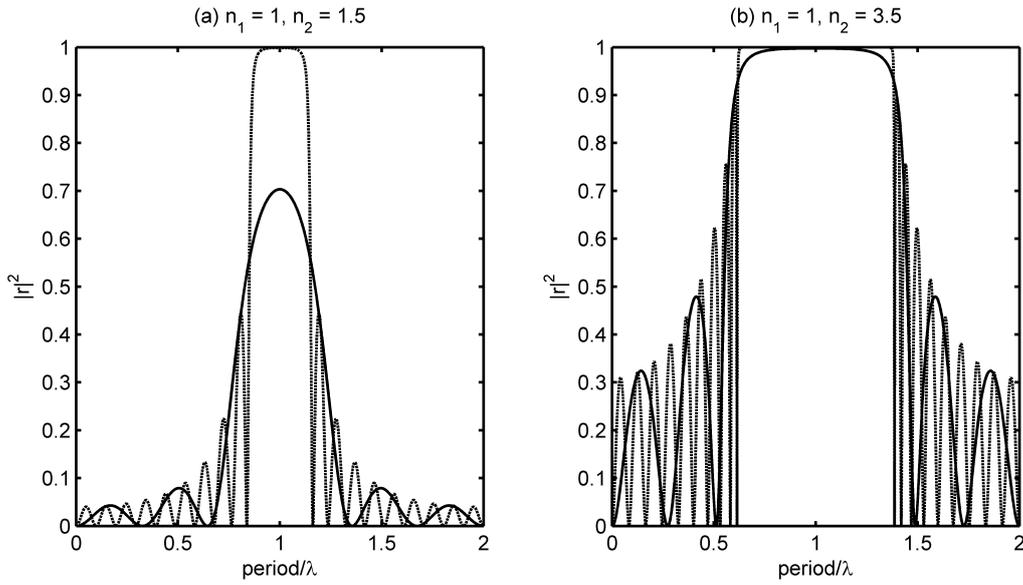**Figure 3.16:** Reflection spectrum of Bragg coatings with different index contrasts Δn and number of layers N: (a) $\Delta n = 0.5$, (b) $\Delta n = 2.5$. Solid (dashed) curves correspond with $N = 3$ ($N = 10$), respectively.

properly tailoring the layer thicknesses. It can be proven that the bandwidth of these reflectors is given by

$$\frac{\Delta\lambda}{\lambda} = \frac{4}{\pi} \sin^{-1} \frac{|n_2 - n_1|}{n_2 + n_1}, \qquad (3.87)$$

which can be approximately written as

$$\frac{\Delta\lambda}{\lambda} = \frac{2}{\pi} \frac{\Delta n}{n}, \qquad (3.88)$$

provided that $n_1 \cong n_2 \cong n$. Thus, a high bandwidth is obtained by increasing the difference between the refractive indices of the layers. These trends are illustrated in Figure 3.16.

To obtain more insight in the Bragg reflector we plot a few field profiles in Figure 3.17. In the bandgap we expect a strong reflection, leading to an exponentially decaying field in the reflector, which is confirmed in Fig. 3.17(d). On the sides of the bandgap there are frequencies with reflection equal to zero (and thus total transmission), these correspond to Fabry-Perot like resonances, built up in between the ends of the Bragg reflector. Indeed, the field profile in Fig. 3.17(c) shows a bump in the profile in the reflector, and total transmission. In addition, the next reflection minimum is shown in Fig. 3.17(a), which indicates a higher order resonance, as it has two maxima in the field profile. In Fig. 3.17(b) the situation in between the previous resonances is plotted, at the point where the reflection has a (relative) maximum. We notice an intermediate situation, with one and a half lobes inside the cavity, leading to a significant reflection (indicated by the standing wave pattern).

Bragg reflectors can be made to reflect broad bands of light by stacking up several periodic layered media with different periods. In this case, each reflector acts as a band rejection filter for each wavelength. If the bandwidths are wide enough to have substantial overlap, the whole structure can reject a broad band of light.
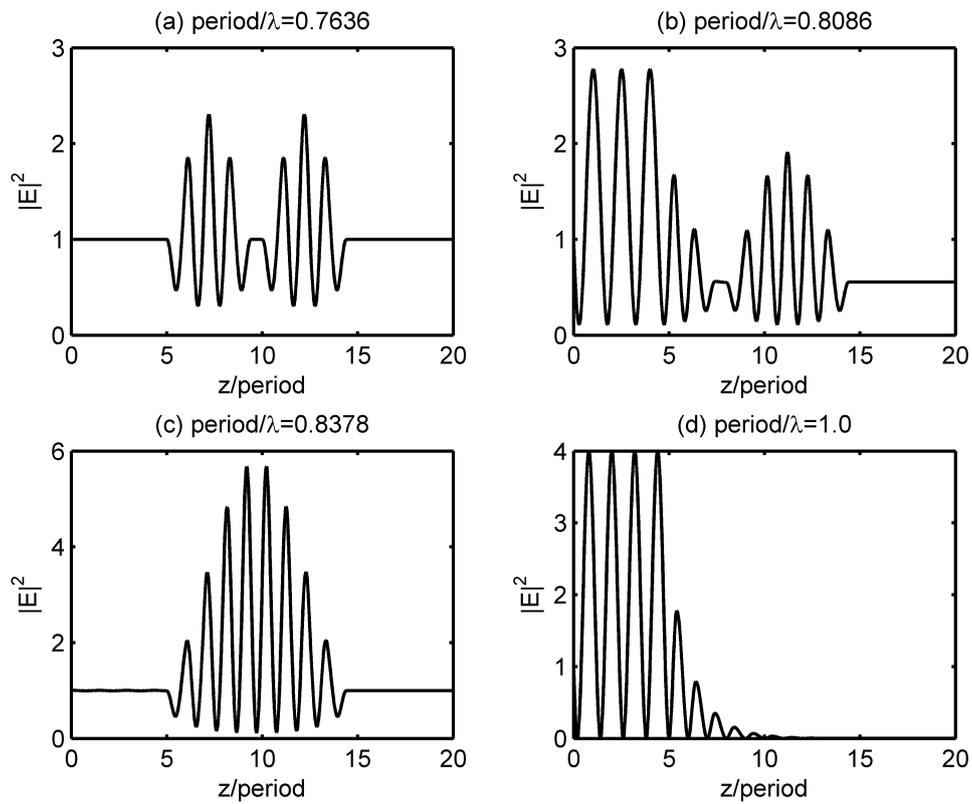
Figure 3.17: Field plots in and around a Bragg reflector. Light with $|E| = 1$ is incident from the left side. The reflector, located between $z/\text{period} = 5$ and 10, has 10 periods with $n_1 = 1, n_2 = 1.5$.
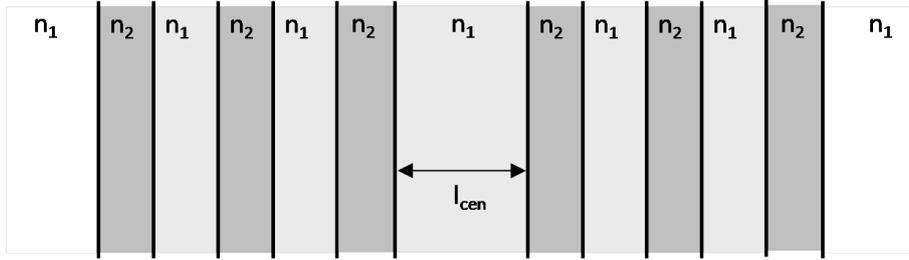
Figure 3.18: Schematic of a Fabry-Perot cavity with Bragg reflector mirrors.
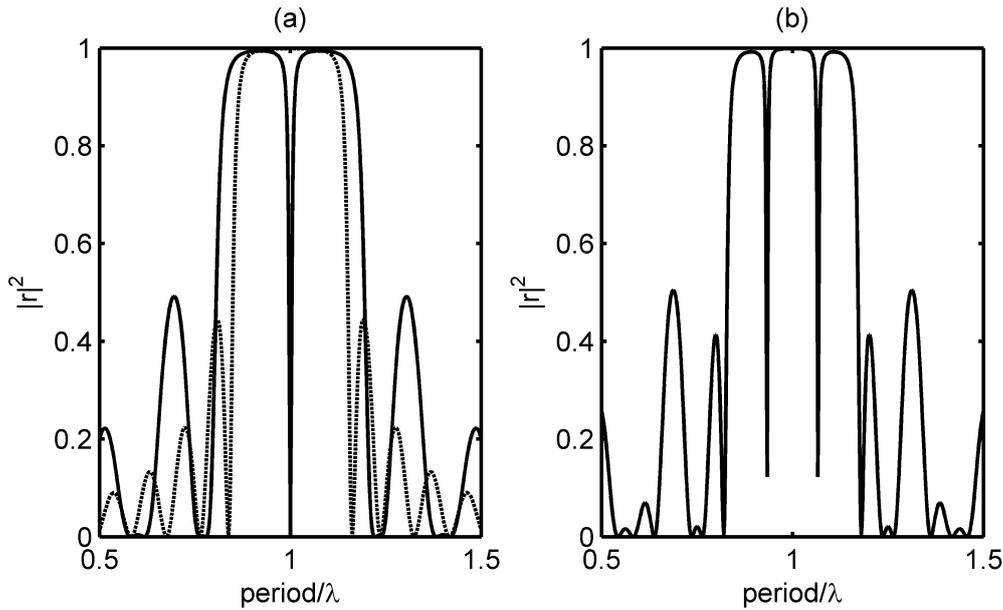

Figure 3.19: Reflection spectra of Bragg structures with $n_1 = 1, n_2 = 1.5$. (a) Solid line: n_1-defect in center with $l_{cen} = \lambda_0/2n_1$, 5 periods on left and right. Dashed line: Bragg without defect and 10 periods. (b) Thicker n_1-defect in center with $l_{cen} = 5.5\lambda_0/2n_1$, 5 periods on left and right.

Fabry-Perot with Bragg reflector

One can combine the concepts of the Fabry-Perot cavity with the Bragg reflector to create a very high quality resonance. This can be done by enlarging one of the layers in the middle of a Bragg reflector. In this way the sides of the Bragg reflector act as good mirrors (in the bandgap), and the 'defect' layer accommodates a Fabry-Perot type cavity mode. Thus, ideally the Bragg wavelength should correspond with the Fabry-Perot resonance wavelength.

A sketch of such a structure is shown in Fig. 3.18, with l_{cen} the thickness of the defect layer. Reflection spectra are depicted in Fig. 3.19. We see a sharp dip in reflection in the center of the bandgap in Fig. 3.19(a)(solid). The spectrum of the Bragg reflector without the defect is indicated with a dashed line. For this particular defect mode we chose $l_{cen} = \lambda_0/2n_1$, thus giving a fundamental Fabry-Perot resonance at λ_0, which is also the center of our Bragg bandgap (for $l_1 = \lambda_0/4n_1, l_2 = \lambda_0/4n_2$). A thicker defect layer can give rise to multiple Fabry-Perot type modes in one bandgap, this is illustrated in Fig 3.19(b).

Bibliography

[Bae06] R. Baets. *Photonics (Syllabus)*. Universiteit Gent, 2006.

[Yeh88] P. Yeh. *Optical Waves in Layered Media*. John Wiley and Sons, 1988.