

Chapter 2

Guided Wave Propagation

MEINT SMIT, XAVEER LEIJTENS

The theory of wave propagation is based on four elementary equations, the Maxwell equations. The solutions to these equations are rather complicated, unless the waveguide structure is very simple. A rigorous analysis requires numerical mode solvers, but it appears to be possible for many structures to obtain approximate solutions, which are often sufficiently accurate for design purposes, with a limited amount of computational labor. Where possible, emphasis will be on such an approximate approach and on the physical understanding which is required in order to apply it successfully.

In the following sections a description of the most important wave propagation phenomena will be given, starting with the most simple structure: free space (section 2.2), and ending with the most complicated one, which is relevant for the development of planar opto-electronic circuits: a 3-dimensional waveguide structure with varying cross-section in the propagation direction.

2.1 Introduction

In a source free medium the Maxwell equations for a field with time dependence $\exp(j\omega t)$ take the following form: *Maxwell equations*

$$\begin{aligned}\vec{E} &= \frac{1}{j\omega\bar{\epsilon}}\nabla\times\vec{H} \\ \vec{H} &= \frac{-1}{j\omega\bar{\mu}}\nabla\times\vec{E} \\ \nabla\cdot\vec{D} &= 0 & (\vec{D}=\bar{\epsilon}\vec{E}) \\ \nabla\cdot\vec{B} &= 0 & (\vec{B}=\bar{\mu}\vec{H})\end{aligned}\tag{2.1}$$

in which \vec{E} and \vec{H} represent the electric and the magnetic field vectors, \vec{D} and \vec{B} are the dielectric displacement and the magnetic flux density, $\bar{\epsilon}$ and $\bar{\mu}$ are the dielectric permittivity and the magnetic permeability of the medium and $\omega = 2\pi f$ is the angular

Support from Alexandr Zozulia, David de Vocht, Limeng Zhang, Lukas Puts, Yi Wang, Zhaowei Chen and Bart Bas for checking the text and the calculations, and generating some of the graphs, is gratefully acknowledged.

Maxwell: $E = \frac{1}{j\omega\epsilon} \nabla \times H$ $H = \frac{-1}{j\omega\mu} \nabla \times E$

Helmholtz: $\nabla^2 U + k^2 U = 0$ $k^2 = \omega^2 \epsilon \mu$

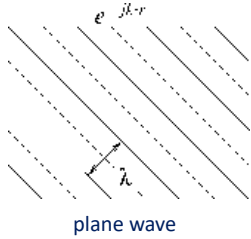
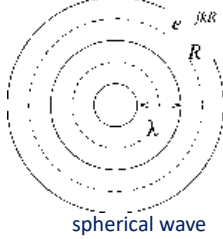
Solutions:  

Figure 2.1: The wave equation.

angular frequency frequency of the wave. In anisotropic media $\bar{\epsilon}$ and $\bar{\mu}$ are tensors, but for isotropic media they are simple constants. Substituting these equations into each other and assuming that the medium is homogeneous, so that $\nabla \bar{\epsilon}$ and $\nabla \bar{\mu}$ are zero, we find the same wave equation for the E -field and the H -field:

$$\nabla^2 \vec{U} + k^2 \vec{U} = 0, \quad k^2 = \omega^2 \epsilon \mu \quad (2.2)$$

Helmholtz equation in which \vec{U} represents either the \vec{E} -field or the \vec{H} -field. This result can be obtained by applying the identities $\nabla \times \nabla \times \vec{U} = \nabla(\nabla \cdot \vec{U}) - \nabla^2 \vec{U}$, $\nabla \cdot \vec{E} = 0$, and $\nabla \cdot \vec{H} = 0$. It is known as the homogeneous Helmholtz equation. Although it is restricted to homogeneous media, it can also be employed for media with a stepwise uniform refractive index profile. For such media, the total solution can be found by combining the solutions for the separate homogeneous regions, with the constraint that the continuity conditions for the field be satisfied at the interfaces. In the following we will focus on this approach. It can be employed for gradually varying media too, by approximating their index profile by a stepwise uniform one with many small steps (staircase approximation).

The homogeneous Helmholtz equation allows for a number of solutions, the most elementary one being the plane wave. More complicated wave forms can be described as combinations of plane waves, so that the plane wave forms the basis for understanding more complicated wave forms. In the following sections we will take the plane wave as the basis for the description of waveguide phenomena in planar waveguides. For a more extensive treatment the reader is referred to Unger [96].

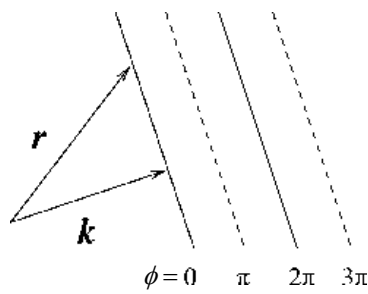
2.2 Wave propagation in free space

2.2.1 The plane wave

plane wave A plane wave can be described as:

$$U(\vec{r}) = A \cdot e^{j(\omega t - \vec{k} \cdot \vec{r})} \quad (2.3)$$

wave vector in which U can be either the E -field or the H -field, A is the amplitude of the field, $\exp(j\omega t)$ is its time dependence, \vec{r} is the position vector and \vec{k} is the wave vector. See



$$U(\mathbf{r}) = A e^{-j\mathbf{k}\cdot\mathbf{r}}$$

$$\phi = \mathbf{k} \cdot \mathbf{r}$$

$$2\pi = k \cdot \lambda$$

$$k = \frac{2\pi}{\lambda}$$

$$\lambda = \frac{2\pi}{k}$$

$$k = \omega\sqrt{\epsilon\mu_0} = \sqrt{\epsilon_r} \cdot \omega\sqrt{\epsilon_0\mu_0}$$

$$n = \sqrt{\epsilon_r}$$

$$c = 1/\sqrt{\epsilon_0\mu_0}$$

$$\omega = 2\pi f$$

$$k = nk_0$$

$$\lambda = \lambda_0/n$$

Figure 2.2: Wave number and wavelength of a plane wave.

Problem 2.1: A plane wave in the wave equation.

Problem a: Show that a field $U = A \exp(-j\vec{k} \cdot \vec{r})$ (plane wave) satisfies the wave equation.

Solution: If we write $U = A e^{-j\phi}$, with $\phi = \vec{k} \cdot \vec{r}$, then:

$$\begin{aligned} \nabla^2 U &= A \nabla \cdot \nabla e^{-j\phi} \\ &= -j A \nabla \cdot \{e^{-j\phi} \nabla \phi\} \\ &= -j A \{ \nabla e^{-j\phi} \cdot \nabla \phi + e^{-j\phi} \nabla^2 \phi \} \\ &= -A e^{-j\phi} \nabla \phi \cdot \nabla \phi - j A e^{-j\phi} \nabla^2 \phi \end{aligned}$$

With $\nabla \phi = \nabla \vec{k} \cdot \vec{r} = \vec{k}$ and $\nabla^2 \vec{k} \cdot \vec{r} = 0$ we find $\nabla^2 U = -k^2 A e^{-j\vec{k} \cdot \vec{r}}$, from which it follows that the Helmholtz equation is satisfied.

Problem b: What is the signal frequency of a light wave with a vacuum wavelength of $1.5 \mu\text{m}$?

Solution: $f = c/\lambda = 3 \cdot 10^8 / 1.5 \cdot 10^{-6} \text{ Hz} = 200 \text{ THz}$.

Figure 2.2. The term $\omega t - \vec{k} \cdot \vec{r}$ represents the phase ϕ of the wave. It is easily verified that this wave satisfies the Helmholtz equation. See Problem 2.1. The time dependence $\exp(j\omega t)$, which is the same for all field components at any position, will be omitted in the following. The field expressions without time dependence may be considered as the field “frozen” at $t = 0$.

time dependence

The direction of the wave vector \vec{k} is the propagation direction of the wave. Its modulus k is called the wave number. The wavelength λ is the distance, measured in the propagation direction, over which the phase of the wave changes by 2π : $\Delta\phi = k\lambda = 2\pi$,

wave number
wavelength

from which it follows that k and λ are related to each other according to:

$$k = 2\pi/\lambda \text{ or } \lambda = 2\pi/k \quad (2.4)$$

The wave number thus inversely depends on the wavelength and has the dimension rad/m or rad/ μm , the latter being most customary in integrated photonics.

dielectric permittivity From equation 2.2 we see that $k = \omega\sqrt{\epsilon\mu}$, in which $\omega = 2\pi f$ is the angular frequency and f is the frequency of the wave. For a medium with dielectric permittivity $\epsilon = \epsilon_r\epsilon_0$ ($\epsilon_0 = 8.854 \cdot 10^{-12}$ F/m) and a magnetic permeability $\mu = \mu_0$ ($= 4\pi \cdot 10^{-7}$ H/m) the wave number can be written as

$$\begin{aligned} k &= nk_0 & k_0 &= 2\pi f \sqrt{\epsilon_0\mu_0} = 2\pi f/c \\ \lambda &= \lambda_0/n & \lambda_0 &= c/f \end{aligned} \quad (2.5)$$

speed of light in which k_0 is the wave number in vacuum, ϵ_0 and μ_0 are the dielectric permittivity and the magnetic permeability of vacuum and $c = 1/\sqrt{\epsilon_0\mu_0} \approx 3 \cdot 10^8$ m/s is the speed of light¹. The refractive index n is defined as:

$$n \equiv \sqrt{\epsilon_r} \quad (2.6)$$

Most optical materials are nonmagnetic. Their magnetic permeability equals μ_0 and the wave number in such a medium is fully defined by the frequency f and the refractive index n .

longitudinal In the literature on optical waveguides it is customary to use the z -coordinate for the propagation direction and the x -coordinate for the direction perpendicular to the film interfaces. The y -coordinate is used for the third direction parallel to the film interfaces. We will refer to these three directions as the longitudinal (z), the transverse (x) and the lateral (y) direction. For the components of a vector we use the same terminology. By the transverse field components (plural) we mean both x and y .

In the first part of this chapter, we will restrict ourselves to plane waves propagating into the x - z -direction in vertically layered structures, so that both the fields and the media are y -invariant. The y -coordinate will, therefore, be omitted from the formulae. A plane wave propagating in an arbitrary direction θ can thus be written as:

$$U(x, z) = e^{-jk_x x} \cdot e^{-jk_z z} \quad (2.7)$$

directional wave number in which k_x and k_z are the directional wave numbers in the x - and z -direction, respectively:

$$\begin{aligned} k_x &= nk_0 \sin \theta \\ k_z &= nk_0 \cos \theta \end{aligned} \quad (2.8)$$

propagation angle If we know the transverse propagation constant k_x of the wave and the refractive index n of the medium, the propagation angle θ follows from

$$\theta = \arcsin(k_x/nk_0) \quad (2.9)$$

From expression 2.7 it is seen that the terms describing the x - and the z -dependence have the same form as the expressions describing plane waves propagating in the x - and the z -direction, respectively. Obviously, a plane wave propagating in an arbitrary

¹The speed of light is presently considered as a natural constant which is defined as 299 792 458 m/s. In this book we will approximate it as $3 \cdot 10^8$ m/s.

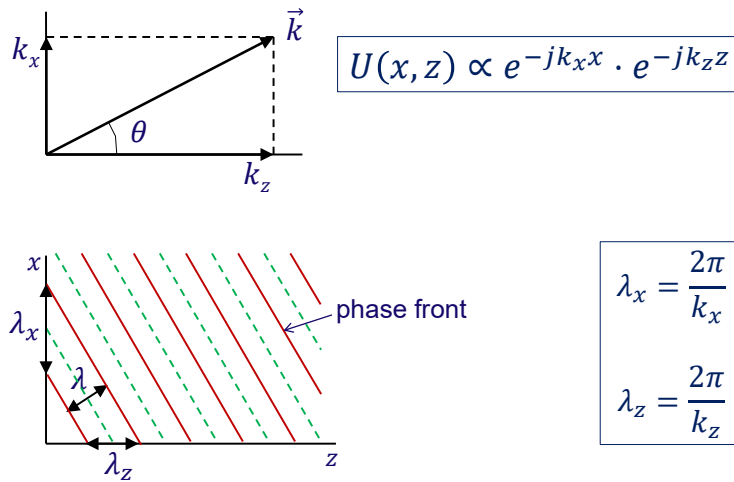


Figure 2.3: Wave decomposition.

direction can be conceived of as a composition of two plane waves with wave numbers k_x and k_z , propagating in the x - and the z -direction, respectively.

The directional wave numbers correspond to directional wavelengths λ_x and λ_z , the spatial period of the wave in the x - and z -direction. They are related to the directional wave numbers in the same way as for a “normal” wave:

$$\begin{aligned} \lambda_x &= 2\pi/k_x, \quad k_x = 2\pi/\lambda_x \\ \lambda_z &= 2\pi/k_z, \quad k_z = 2\pi/\lambda_z \end{aligned} \tag{2.10}$$

The concept of wave decomposition, i.e. considering a plane wave in an arbitrary direction as consisting of a longitudinal and a transverse wave (see figure 2.3), will prove useful for understanding propagation properties of waveguide structures. *wave decomposition*

2.2.2 Interference

Two plane waves with equal amplitudes A and opposite propagation angles $\pm\theta$ will produce a field pattern of the following form:

$$\begin{aligned} U(x, z) &= A(e^{-jk^+\cdot\vec{r}} + e^{-jk^-\cdot\vec{r}}) \\ &= A(e^{-jk_x x} + e^{jk_x x})e^{-jk_z z} \\ &= 2A\cos(k_x x)e^{-jk_z z} \end{aligned} \tag{2.11}$$

The corresponding wave pattern is depicted in figure 2.4. In the x -direction the two waves produce a standing wave $\cos(k_x x)$ with period $\lambda_x = 2\pi/k_x$, in the z -direction they behave as a traveling wave $\exp(-jk_z z)$ with wavelength $\lambda_z = 2\pi/k_z$. The whole field can be conceived of as a standing wave pattern $U(x) = \cos(k_x x)$ which propagates into the z -direction with longitudinal wave number k_z . *standing wave traveling wave*

If the field at a plane of constant z has the shape of a standing wave pattern with period λ_x , then the longitudinal wave number of the field and the propagation angles $\pm\theta$ of the corresponding plane waves at that location can be found from:

Problem 2.2: Standing wave pattern.

Problem a: Periodic optical gratings can be produced by illuminating a thin photo sensitive film, deposited onto a planar substrate, with two interfering plane waves (broad beams), so that the film is exposed to the standing wave pattern of the two beams. If the exposed film material is removed with a developer, a periodic line pattern will remain, which can be used as a mask for etching the grating pattern into the substrate. Calculate the required angle between the two beams if the grating period Λ has to be 200 nm and the vacuum wavelength λ_0 of the two beams is 364 nm.

Solution: The interference pattern has two maxima per transverse wavelength λ_x . For a grating period of 200 nm we thus need a transverse wavelength $\lambda_x = 2\Lambda = 400$ nm. This corresponds to an angle $\theta = \arcsin(k_x/k_0) = \arcsin(\lambda_0/\lambda_x) = 65.5^\circ$. The required angle between the beams is twice this angle, i.e. 131.0° .

Problem b: Calculate the wavelength λ_z of the wave pattern between the two plates in figure 2.4 for the cases that it has one and two maxima, respectively, between the plates, if the distance between the plates is 5 cm and the wavelength λ_0 in vacuum is 3 cm?

Solution: For one maximum: $k_z = \sqrt{k_0^2 - k_x^2} = \sqrt{(2\pi/3)^2 - (\pi/5)^2} = 2.0 \text{ cm}^{-1}$, $\lambda_z = 2\pi/k_z = 3.14$ cm. For two maxima $k_z = \sqrt{(2\pi/3)^2 - (2\pi/5)^2} = 1.68 \text{ cm}^{-1}$, $\lambda_z = 2\pi/k_z = 3.75$ cm.

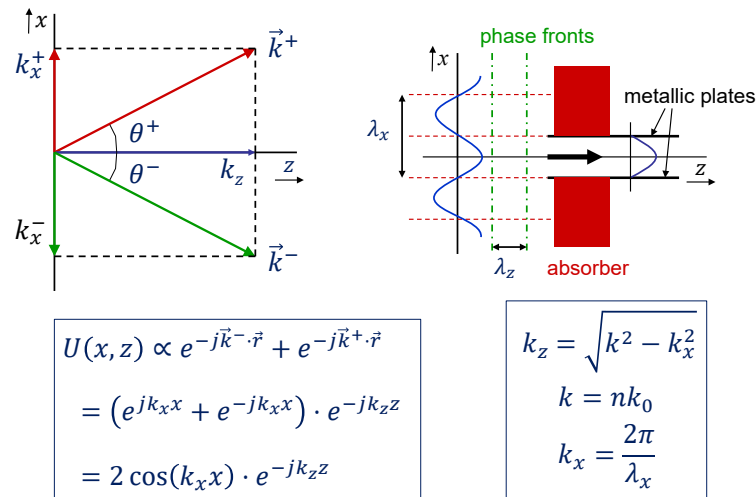


Figure 2.4: A metallic waveguide mode conceived as interference pattern of two plane waves.

$$\begin{aligned}
 k_z &= \sqrt{n^2 k_0^2 - k_x^2}, \quad k_x = \frac{2\pi}{\lambda_x} \\
 \theta &= \arcsin\left(\frac{k_x}{nk_0}\right) = \arctan\left(\frac{k_x}{k_z}\right)
 \end{aligned}
 \tag{2.12}$$

The interference is observed only when the two interfering waves are coherent and have the same polarisation state, i.e. if they have the same phase or a constant phase difference. If the waves are not coherent, which will usually be the case if they originate from different sources, the interference pattern will fluctuate so rapidly (dependent on the momentaneous frequency difference) that only the average intensity pattern is observed. *coherence*

The mechanism of interference is useful for understanding waveguide phenomena. From figure 2.4 it can be seen that at the planes $x = (\pi/2 \pm m\pi)/k_x$, $m = 0, 1, 2, \dots$ the field intensity is zero. For a field which is electrically polarised in the y -direction (perpendicular to the x - z -plane), we can introduce metallic plates at these zero field positions without affecting the field. If we introduce absorbers at both sides of the plates, the field outside the plates will vanish, but the field inside the plates will continue propagating. The plates act as a waveguide; the guided wave can be conceived of as a plane wave which is reflected back and forth between the plates. It also shows the relation with ray theory where propagation in a waveguide is described as a ray which is reflected between the edges of a waveguide, a ray being an asymptotic approximation of a plane wave. In section 2.4 it will be shown how the interference concept can also be used for understanding wave propagation in optical waveguides.

2.2.3 Fourier optics

In the previous section we saw that a sinusoidal wave pattern $U(x)$ can be conceived of as the sum of two waves with x -dependence $\exp(\pm jk_x x)$. More general an arbitrary field distribution $U(x)$ can be considered as being composed of a spectrum of plane waves $\exp(-jk_x x)$ with spectral density $U_k(k_x)$:

$$U(x) = \int_{-\infty}^{\infty} U_k(k_x) \cdot e^{-jk_x x} dk_x \tag{2.13}$$

Because the exponent represents the x -dependence of a plane wave with transverse wave number k_x and propagation angle $\theta = \arcsin[k_x/(nk_0)]$, the function $U(x)$ can be considered as resulting from the interference of a spectrum of plane waves $\exp(-j\vec{k}_\theta \cdot \vec{r}) = \exp[-j(k_x x + k_z z)]$ with a spectral distribution $U_k(k_x)$. *spectral distribution*

The spectral distribution $U_k(k_x)$ is closely related to the angular amplitude spectrum U_θ (see problem 2.3a). The formulas remain slightly more elegant if the spectral distribution $U_k(k_x)$ is used instead of $U_\theta(\theta)$. Because we recognize the above integral as a Fourier transformation, the spectral distribution is computed from the inverse transformation: *angular amplitude spectrum*

$$U_k(k_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(x) \cdot e^{+jk_x x} dx \tag{2.14}$$

An arbitrarily shaped field $U(x)$ can thus be decomposed into a spectrum of plane waves $U_k(k_x)$. This is an important result. It implies that the propagation properties of arbitrary fields can be reduced to the propagation properties of plane waves.

Problem 2.3: Angular spectrum

Problem a: What is the relation between the spectral distribution $U_k(k_x)$ and the angular spectrum $U_\theta(\theta)$, in which θ and k_x are related through $k_x = nk_0 \sin \theta$?

Solution: $U_\theta(\theta)d\theta = U_k(k_x)dk_x$, so that $U_\theta(\theta) = U_k(k_x)(dk_x/d\theta)$. The only difference is thus the factor $dk_x/d\theta = nk_0 \cos \theta$ which is almost constant for small values of θ .

Problem b: Compute the angular distribution (the far field) corresponding to a uniform amplitude distribution $U(x) = 1$, $|x| < a/2$ and a distribution $U(x) = \cos(\pi x/a)$, also for $|x| < a/2$.

Solution:

1. $U(x) = 1$, $|x| < a/2$
 $\Rightarrow U_k(k_x) = \frac{a}{2\pi} \text{sinc}(\frac{1}{2}k_x a) \Rightarrow U_\theta(\theta) \propto \text{sinc}(\frac{1}{2}nk_0 a \sin \theta) \cdot \cos \theta$
2. $U(x) = \cos(\pi x/a)$, $|x| < a/2$
 $\Rightarrow U_k(k_x) = a \cos(\frac{1}{2}k_x a) / (\pi^2 - k_x^2 a^2)$
 $\Rightarrow U_\theta(\theta) \propto \cos(\frac{1}{2}nk_0 a \sin \theta) \cdot \cos \theta / [\pi^2 - (nk_0 a \sin \theta)^2]$

Problem c: What is the width of the main lobe (between zeroes) of the far field of the metallic plate waveguide as shown in figure 2.4 if the distance between the plates is 5 cm and the vacuum wavelength is 3 cm?

Solution: From problem 2.3b-2 we see that the (first) zeroes of the angular sinc spectrum occur at $\frac{1}{2}k_x a = \pm \frac{3}{2}\pi$. So we find $k_x = \pm 3\pi/a = nk_0 \sin \theta_0$. With $a = 5$ cm and $n = 1$ this gives us $\theta_0 = \pm 64^\circ$.

diffraction The formula pair 2.13 and 2.14 offers a means for analyzing diffraction phenomena in a homogeneous medium. The field $U(x, z)$ through the whole x - z plane follows directly from the source field $U(x, z = 0)$ by integrating over all the plane waves contained in the spectrum $U_k(k_x)$. If we want to know the field at a point $P(x, z)$ we have to integrate the local amplitudes $U_k(k_x) \exp[-j(k_x x + k_z z)]$ at the position (x, z) over the whole spectrum of plane waves (see figure 2.5):

$$\begin{aligned}
 U(x, z) &= \int_{-\infty}^{\infty} U_k(k_x) e^{-j(k_x x + k_z z)} dk_x & (2.15) \\
 &= \int_{-\infty}^{\infty} [U_k(k_x) e^{-jk_z z}] e^{-jk_x x} dk_x \\
 &= \mathcal{F} \{ U_k(k_x) e^{-jk_z z} \}
 \end{aligned}$$

with

$$k_z = \sqrt{n^2 k_0^2 - k_x^2} \quad (2.16)$$

The computation thus comes down to computing the spectral distribution $U_k(k_x)$ corresponding to the distribution $U_x(x)$ by means of a Fourier transform, multiplying it with the z -propagation factor $\exp(-jk_z z)$ and then transforming it back. Note that k_z

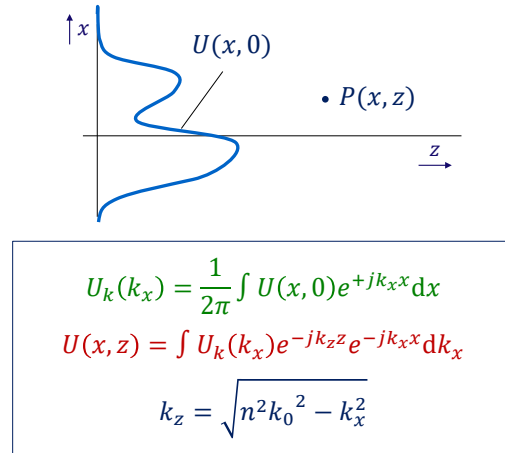


Figure 2.5: Diffraction analysis with Fourier optics.

is dependent on k_x (equation 2.16). The analysis of diffraction and imaging phenomena based on these formulas is known as Fourier Optics. Fourier Optics is particularly suited for analyzing the imaging properties of lenses, for example the field distribution in the focal plane of a lens.

Another application is the analysis of the shape of a beam coming out of a waveguide. For this application the amplitude distribution at the end face of the waveguide is substituted for the field $U_x(x)$ at $z = 0$. This field is usually called the near field. Note that the term near field is sometimes also used with respect to the intensity distribution $I(x) = U_x^2(x)$. What we are often interested in is the angular intensity distribution $I_\theta(\theta)$ of the field at a large distance from the aperture, the so called far field (also used with respect to the amplitude distribution). It can be derived from the spectral distribution by remembering that the power $I(\theta)d\theta$ contained in the angular interval $d\theta$ equals the power $U_k^2(k_x)dk_x$ contained in the corresponding interval dk_x :

$$I(\theta) = U^2(k_x) \frac{dk_x}{d\theta} \propto U_k^2(k_0 \sin \theta) \cos \theta \quad (2.17)$$

$$\approx U_k^2(k_0 \theta), \theta \ll \frac{\pi}{2}$$

The far field thus follows directly from the Fourier Transform of the near field.

2.2.4 Beam Propagation Method (BPM)

The main application of Fourier Optics is analysis of beam propagation in homogeneous media. It can also be used, however, for an approximate analysis of beam propagation in media with small perturbations of homogeneity, such as waveguides with a slightly higher refractive index than the background medium. An example is the directional coupler depicted in Figure 2.6. It consists of two parallel waveguides (the shaded regions) which are brought so close together that the power in the input waveguide couples to the output waveguides. The coupling ratio can be controlled with the length over which the waveguides couple. Such a structure can be analyzed by propagating the field through the structure in small steps as if it were homogeneous, and correcting after each step for the non-homogeneous perturbation.

Problem 2.4: Fourier optics.

Problem a: Calculate the angular width of the main lobe (between zeroes) of the far field of a uniform plane wave ($\lambda_0 = 0.8 \mu\text{m}$) which is passing through a slit with a width of $2 \mu\text{m}$.

Solution: $U_k(k_x) = 1/2\pi \int_{-1}^{+1} \exp(jk_x x) dx \propto \text{sinc}(k_x)$. The zeroes are at $k_x = \pm\pi$, i.e. $\theta = \pm \arcsin(k_x/k_0) = \pm \arcsin[\pi/(2\pi/\lambda_0)] = \pm \arcsin(0.4) = \pm 23.6^\circ$. So the full width of the beam is 47.2° .

Problem b: What is the resolution of a lens with numerical aperture NA if the spectral distribution of the converging beam behind the lens is uniform within the NA of the lens? The NA is the sine of the maximum angle within the converging beam behind the lens, i.e.: $NA = \sin[\arctan(\frac{1}{2}D_\ell/f)]$ with D_ℓ being the lens diameter and f the distance of the focal image to the lens.

Solution: For reasons of reciprocity (the relation between a near field and its far field holds for both propagation directions) the focal field follows as the Fourier transform of the spectral distribution behind the lens: $U_k(k_x) = U_k(k_0 \sin\theta) = 1$ for $|k_x| < k_0 NA$ and zero elsewhere. The focal field then follows as $U(x) = \int U_k(k_x) \exp(-jk_x x) dk_x = \int \exp(-jk_x x) dk_x$, in which the latter integration is over the interval $-k_0 NA < k_x < k_0 NA$. We thus find $U_x(x) = C \sin(k_0 NA x)/(k_0 NA x) = C \text{sinc}(k_0 NA x)$. The resolution Δx is the minimum distance for which two points can be distinguished separately, i.e. the distance between the maximum of the sinc function and its first zero, which occurs at $k_0 NA x = \pi$, so $\Delta x = \lambda/(2 NA)$.

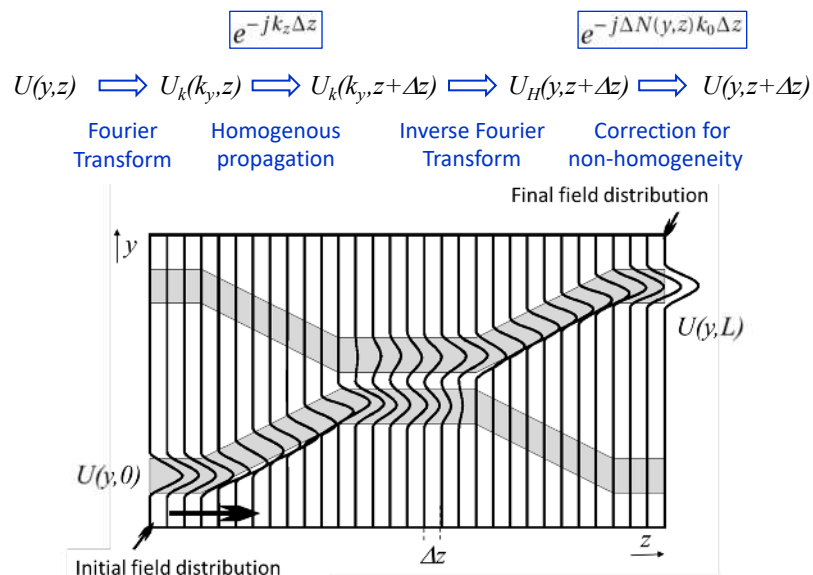


Figure 2.6: Example of the field calculated with a BPM for a directional coupler.

Beam Propagation Methods can be applied in two- as well as in three dimensions. Three-dimensional BPMs are very time consuming. BPMs are, therefore, mostly used for two-dimensional analysis. Although real waveguides in Photonic ICs are three-dimensional they can be analysed with a two-dimensional BPM using the Effective Index Method (EIM), as explained in section 2.5.1. In this approach the three-dimensional refractive index profile $n(x, y, z)$ is reduced to a two-dimensional effective index profile $N(y, z)$ which takes into account the effect of the vertical index profile $n(x)$ on the field propagation. In such a two-dimensional waveguide structure a Beam-Propagation Method works as follows.

*Effective Index
Method
effective index*

In the first step the field $U(y, 0)$ at the beginning of the section is propagated over a short distance Δz with equation 2.15 as if the structure were homogeneous with an x - and y -invariant index n_0 , for which we usually take the refractive index in the background region. In the second step we correct the phase of the field with a term $e^{-j\Delta n(y)k_0\Delta z}$, in which $\Delta n(y) = n(y) - n_0$ is the index contrast with the background index n_0 , which describes the perturbation from homogeneity. The thus corrected field is taken as starting point for the next propagation step, and in this way we can propagate the field in steps Δz through the whole structure.

In formulas a 2d-BPM can be described as follows (using the effective index profile $N(y, z)$ with background index N_0):

1. Transform the field $U(y, z)$ into its plane wave spectrum

$$U_k(k_y, z) = \mathcal{F}\{U(y, z)\}$$

2. Propagate the spectrum of plane waves to $z + \Delta z$

$$U_k(k_y, z + \Delta z) = U_k(k_y, z)e^{-jk_z\Delta z}$$

in which $k_z = \sqrt{N_0^2 k_0^2 - k_y^2}$

3. and transform it back:

$$U_H(y, z + \Delta z) = \mathcal{F}^{-1}\{U_k(k_y, z + \Delta z)\}$$

in which the field $U_H(y, z + \Delta z)$ is the field at $z + \Delta z$ after propagation through a homogeneous region.

4. Correct the field for the index difference $\Delta N(y, z) = N(y, z) - N_0$:

$$U(y, z + \Delta z) = U_H(y, z + \Delta z) \cdot e^{-j\Delta N(y, z)k_0\Delta z}$$

Figure 2.6 summarizes the four steps, of which the first three steps are homogeneous propagation using Fourier Optics (diffraction) and step 4 is a correction for non-homogeneity. The figure shows an example of the field calculated with a BPM for a directional coupler.

For the accuracy of the method it is important that changes in the field as well as in the index profile are small within a propagation step. A proper step size can be determined by reducing the step size until the calculated field is no longer affected by a further reduction. Practical step sizes are in the order of a few 10s of wavelengths for structures with a low index contrast to a few wavelengths for higher index contrasts.

As said, BPMs can also be used for analysing beam propagation in three dimensions, these methods require significantly more computation time, however. These and other methods for calculating propagation in 2d- and 3d-waveguide structures are described in more detail in Chapter 6.

2.2.5 Gaussian beam

The near field of a gas laser can be described in good approximation as a Gaussian beam:

$$U(r) = A e^{-\frac{r^2}{w_0^2}} \quad (2.18)$$

Gaussian beam waist in which w_0 is called the Gaussian beam waist, i.e. the beam radius at which the amplitude is reduced by a factor of $1/e$ (see figure 2.7). Also the field coming out of an optical waveguide can be approximated for many applications as a Gaussian beam.

Because a Gaussian function is invariant under the Fourier transformation, a beam with a Gaussian near field also has a Gaussian spectral distribution and a Gaussian far field :

$$U_k(k_x) = A \frac{w_0}{2\sqrt{\pi}} e^{-\frac{1}{4} k_x^2 w_0^2} \quad (2.19)$$

or with $k_x = nk_0 \sin\theta$:

$$U_\theta(\theta) = C \exp\left(-\frac{\sin^2\theta}{\sin^2\theta_0}\right) \approx C \exp\left(-\frac{\theta^2}{\theta_0^2}\right) \quad (2.20)$$

in which $\sin\theta_0 = \frac{\lambda}{\pi w_0}$ and $\lambda = \frac{\lambda_0}{n}$. So both the near field and the far field of a Gaussian beam have a Gaussian profile, and their beam waists w_0 and $\sin\theta_0$ (or θ_0) can be calculated from each other according to:

$$\sin\theta_0 = \frac{\lambda}{\pi w_0}, \quad w_0 = \frac{\lambda}{\pi \sin\theta_0} \quad (2.21)$$

or, if $\theta_0 \ll 1$:

$$\theta_0 = \frac{\lambda}{\pi w_0}, \quad w_0 = \frac{\lambda}{\pi \theta_0}.$$

For the Numerical Aperture of a Gaussian beam we find:

$$\text{NA} = n \sin\theta_0 \approx \frac{\lambda_0}{\pi w_0} \quad (2.22)$$

So the beam divergence θ_0 of a Gaussian beam is dependent on the refractive index n , whilst the NA is not. In air there is no difference, but in a slab waveguide or for immersion lithography both the divergence angle and the diffraction limited spot size will be smaller than in air.

angular width beam divergence The $1/e^2$ half width ($1/e^2$ intensity) θ_0 is called the angular (half) width of the beam. The full width $2\theta_0$ is often referred to as beam divergence. Note that half and full width definitions are used interchangeably in the literature and also in this text.

An advantage of Gaussian beams is that there exists an analytical expression, not only for the far field, but for the field throughout the whole beam (see Appendix 2A on page 2-69):

$$U(x, z) = A \frac{w_0}{w_c} \cdot e^{-\frac{x^2}{w_c^2}} \cdot e^{-jn k_0 z} \quad (2.23)$$

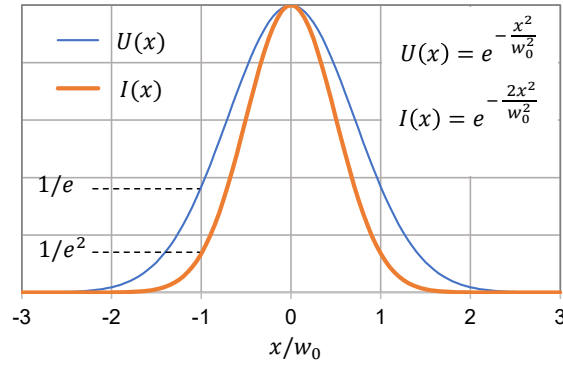


Figure 2.7: Gaussian beam

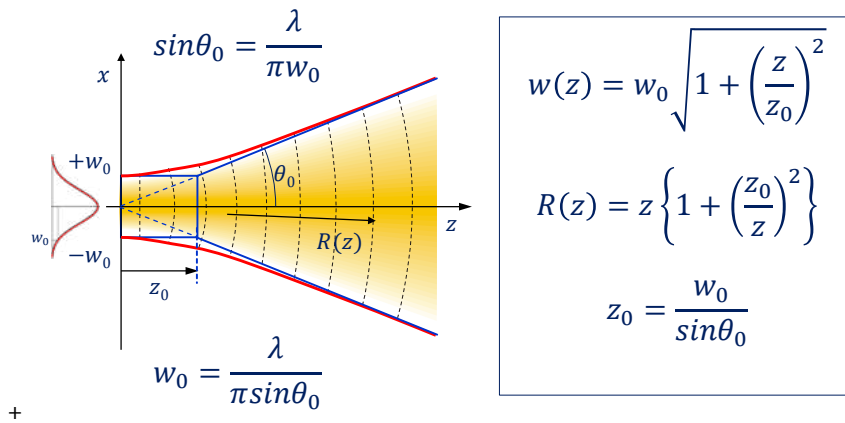


Figure 2.8: Diffraction of Gaussian beam.

in which w_c is a complex beam waist:

$$w_c = w_0 \sqrt{1 - j \frac{2\pi}{nk_0 w_0^2}} \tag{2.24}$$

The beam is illustrated in figure 2.8.

The x -dependence can be written as (see Appendix 2A):

$$U_x(x, z) = A e^{-\frac{x^2}{w_c^2}} = A e^{-\frac{x^2}{w^2} - \frac{jnk_0 x^2}{2R}} \tag{2.25}$$

in which w is the (real) width of the beam:

$$w = w_0 \sqrt{1 + (z/z_0)^2} \tag{2.26}$$

and the term $-\frac{jnk_0 x^2}{2R}$ describes the curvature of the phase fronts with

$$R = z[1 + (z_0/z)^2] \tag{2.27}$$

in which

$$z_0 = \frac{w_0}{\sin\theta_0} \tag{2.28}$$

Problem 2.5: Divergence and focal field of a Gaussian beam.

Problem a: A He-Ne gas laser ($\lambda = 632.8$ nm) produces a Gaussian beam with a $1/e^2$ (intensity) diameter of 0.5 mm, as specified by the supplier. What is the far field divergence of this beam? Estimate the beam diameter at a distance of 2 m. Do the same for a beam with 1 mm beam diameter.

Solution:

1. Far field divergence ($w_0 = 0.25$ mm): $\sin\theta_0 = \lambda/(\pi w_0) = 8.06 \cdot 10^{-4} \Rightarrow \theta_0 = 0.046^\circ$, so the far field divergence is $2\theta_0 \approx 0.1^\circ$.
2. Beam diameter at $z = 2$ m, for $w_0 = 0.25$ mm. From equation 2.26: $2w(z) = 2 \cdot w_0 \cdot 6.52 = 3.26$ mm.
3. Beam diameter at $z = 2$ m, for $w_0 = 0.5$ mm. From equation 2.26: $2w(z) = 2 \cdot w_0 \cdot 1.8965 = 1.90$ mm.

At distances larger than the focal depth the broader beam has become the narrower one!

Problem b: A Gaussian beam ($\lambda_0 = 1.3$ μm) with 1-mm $1/e^2$ -diameter is focused by a lens with 1 mm focal length. Calculate the spot diameter in the focal plane.

Solution: The numerical aperture of the beam behind the lens is $\text{NA} = \sin\theta_0 = \sin[\arctan(w/f)] = 0.45 \Rightarrow \text{spot diameter } 2w_0 = \lambda_0/(\pi\text{NA}) = 1.8$ μm .

focal depth is the focal (half) depth of the beam, i.e. the distance over which the beam width increases by a factor $\sqrt{2}$, as can be seen from equation 2.26. Graphically it can be interpreted as the distance z at which the parallel beam edges $x = \pm w_0$ intersect the asymptotes $\theta = \theta_0$ of the diverging part of the beam. From equation 2.23 we see that the beam propagates in the z -direction with wave number nk_0 , as might be expected. The beam profile is Gaussian not only at the near and the far field, but everywhere throughout the beam. Phase front curvature is zero for $z = 0$, for large values of z the radius R approaches z , which means that the diverging beam virtually comes from the origin.

In applications where the divergence angle θ_0 is small we apply the approximation $\sin\theta_0 \approx \theta_0$.

2.2.6 Gaussian approximation of beams and modes

Gaussian beam approximation Gaussian beams are useful for a quick approximate analysis of diffraction and coupling phenomena for beams with a Gaussian like beam shape. The actual beam is then considered as if it were a Gaussian beam. The translation of a beam into a Gaussian beam can be done in different ways. A practical one is by means of the effective width defined as the full width of a uniform profile with amplitude $U_{\max} = \max[U(x)]$, which has the same power contents (i.e. $\int U^2 dx$) as the function $U(x)$:

$$w_e = \frac{1}{U_{\max}^2} \int U^2(x) dx \quad (2.29)$$

Be aware that different definitions of the effective width exist! The concept is illustrated

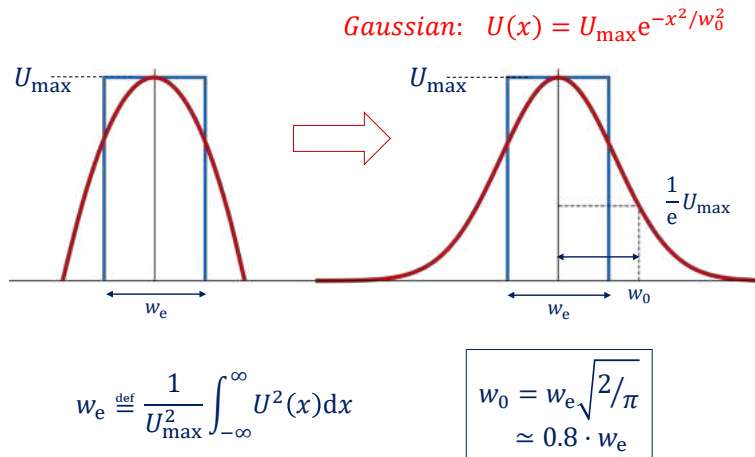


Figure 2.9: Approximating an arbitrary beam as a Gaussian beam.

in figure 2.9. For a Gaussian function $\exp(-x^2/w_0^2)$ evaluation of the above integral yields $w_e = w_0 \sqrt{\pi/2}$ so that the width w_0 of an equivalent Gaussian beam is found from the effective width w_e (as computed from equation 2.29) of the original function according to:

$$w_0 = w_e \sqrt{\frac{2}{\pi}} \quad (2.30)$$

The same relation is valid for the far field of a Gaussian beam, i.e.:

$$\theta_0 = \theta_e \sqrt{\frac{2}{\pi}} \quad (2.31)$$

in which

$$\theta_e = \frac{1}{I_{\max}} \int I(\theta) d\theta \quad (2.32)$$

and $I = U^2$. The above relations are only valid if $\sin \theta \approx \theta$, i.e. $\theta \ll \pi/2$. Note that the effective widths w_e and θ_e are full widths, whereas the Gaussian widths w_0 and θ_0 are half widths!

The calculation of an equivalent Gaussian beam comes down to calculating the effective width, according to equation 2.29, after which the waist of the equivalent Gaussian beam follows from eq.2.30.

Another beam width definition, which is often applied because of the ease in determining it from a measured curve, is the so called Full Width Half Maximum (FWHM), *FWHM* i.e. the full width at the 50% level. The FWHM is commonly defined with respect to the 50% intensity level (i.e. the $\frac{1}{2}\sqrt{2}$ -amplitude level). For many functions the FWHM and the effective width as defined above come close to each other (see problem 2.6).

2.2.7 Diffraction limited imaging

Figure 2.10 shows the most customary configurations for coupling light into a planar waveguide from a broad collimated beam or a fiber. In order to obtain a good coupling efficiency the spot size in the focal plane has to match that of the guided wave. Both configurations will be shortly discussed. More information about imaging is provided in chapter 7: Measurement Methods.

Problem 2.6: Full width half maximum.

Problem a: Calculate the effective width and the FWHM if the amplitude distribution is a Gaussian function $\exp(-x^2/w_0^2)$, a cosine function $\cos(x)$, $|x| < \pi/2$ or a square cosine function $\cos^2(x)$, $|x| < \pi/2$.

Solution:

1. Gaussian: $w_e = w_0\sqrt{\pi/2} \approx 1.25w_0$. The FWHM is found from: $\exp(-w_{0.5}^2/w_0^2) = \sqrt{1/2}$, so $\text{FWHM} = 2w_{0.5} = 2\sqrt{-\ln\sqrt{1/2}} = 1.18w_0$; so $w_e = \text{FWHM} \cdot 1.25/1.18 = 1.06\text{FWHM}$
2. Cosine: $w_e = \text{FWHM} = \frac{\pi}{2}w_0$;
3. Square cosine: $w_e = w_0 \int \cos^4(x)dx = \frac{3}{4}w_0 \int \cos^2(x)dx = 0.375\pi w_0$.
FWHM: $w_0 \cos^2(\frac{1}{2}\text{FWHM}) = w_0\sqrt{1/2} \Rightarrow \text{FWHM} = 0.364\pi w_0 \Rightarrow w_e = \frac{0.364}{0.375}\text{FWHM} = 0.97\text{FWHM}$.

Problem b: Estimate the effective vertical width of the far field of a semiconductor laser with a vertical near field FWHM of $1.2\mu\text{m}$ and an emission wavelength of $1.3\mu\text{m}$.

Solution: We start with the assumption that the effective width equals the FWHM. Next we approximate the beam as a Gaussian beam with $w_0 = w_e\sqrt{2/\pi} = 0.96\mu\text{m}$. The NA of the far field is then found from $\text{NA} = \lambda/(\pi w_0) = 0.43 \Rightarrow \theta_0 = \arcsin(\text{NA}) \approx 26^\circ$ and $\theta_e = \theta_0\sqrt{\pi/2} = 32^\circ$. Note that θ_e is the (effective) full width, whereas θ_0 is the (Gaussian) half width.

Focusing a broad beam

If a broad beam (e.g. from a gas laser or a fibre collimator) with Gaussian beam waist w_L has to be coupled into a waveguide we need, in principle, only one lens (usually a microscope objective). The spot size w_2 in the focal plane is determined by the numerical aperture of the beam after passing the lens:

$$\text{NA} = n \sin\theta_0 = n \frac{w_L}{\sqrt{w_L^2 + f^2}} \approx n \frac{w_L}{f}, \quad \text{if } \frac{w_L}{f} \ll 1 \quad (2.33)$$

in which f is the focal length of the lens or objective. For a Gaussian beam it follows directly as

$$w_2 = \frac{\lambda_0}{\pi \text{NA}} = \frac{\lambda_0}{\pi n \theta_0} \quad (2.34)$$

in which the refractive index n is usually 1. This spot size is called the diffraction limited spot size; it is only obtained with aberration free lenses. Microscope objectives are designed for imaging in the focal plane with minimal aberration. Therefore, they are indispensable for efficient coupling into narrow waveguides. The spot size in the focal plane can thus be chosen by selecting a lens with a proper numerical aperture. The diameter of the beam should match the aperture of the objective: if it is too large part of the beam will be truncated, if it is too small the numerical aperture of the objective will not be completely filled and the diffraction limited size of the spot will increase.

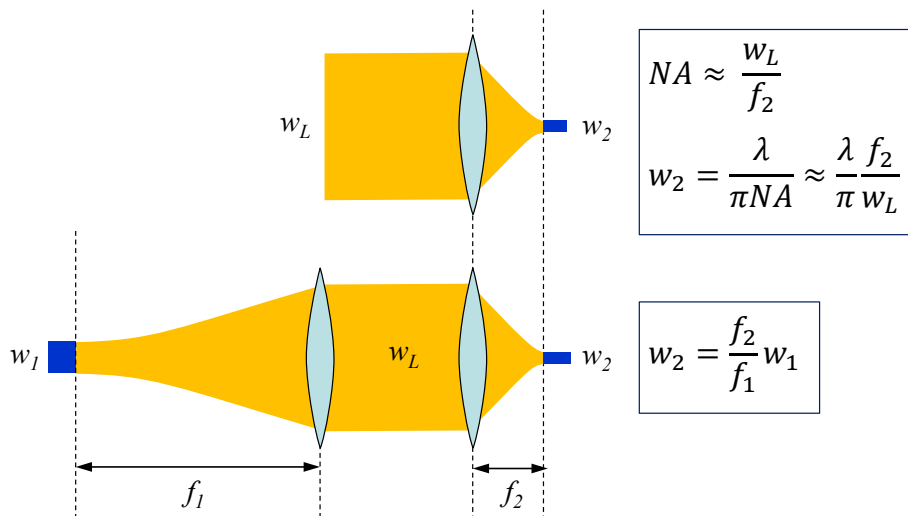


Figure 2.10: Coupling with one or two lenses.

Imaging a spot

If the *spot size* has to be changed, for example in coupling light from a fibre to a semiconductor waveguide, a two-lens configuration as depicted in figure 2.10 is often used. Such a configuration is also convenient if we want to insert additional optical elements into the parallel beam section, e.g. optical isolators, beamsplitters for power or wavelength monitors, waveplates, etc. If the spot to be imaged has a (Gaussian) spot radius w_1 , the beam radius w_L behind the lens is $w_L = f_1 \tan \theta_1 \approx f_1 \lambda / (\pi w_1)$. If $w_L \gg \lambda$, which is usually the case, the beam may be considered parallel in the region between the two lenses. The spot size w_2 behind the second lens is $w_2 = \lambda / (\pi \sin \theta_2) \approx \lambda / (\pi w_L / f_2) = (f_2 / f_1) w_1$. It is seen that the magnification formula of geometrical optics for the two lens imaging system ($M = f_1 / f_2$) is also valid for the magnification of diffraction limited spots.

2.3 Wave propagation in two-layer media

2.3.1 Polarization in layered structures

In section 2.2.2 we saw that waves can propagate in a 2-dimensional metal waveguide, formed by two parallel metal plates. The guided wave can be conceived of as the interference pattern of two plane waves which are reflected back and forth between the metal plates. More generally, waves can propagate between any two reflecting interfaces, which can be metal plates, but also reflecting dielectric interfaces, e.g. the interface between media with different refractive indices. As metal waveguides are very lossy in the optical domain, the most frequently used waveguide in Photonic Integrated Circuits is the dielectric waveguide, which consists of an optically transparent layer embedded between two layers with a lower refractive index. The guided wave in the central layer can propagate without loss if it is totally reflected at both dielectric interfaces, which will happen if the propagation angle of the plane waves, of which it consists, is below the critical angle. In this section we will study the reflection and trans-

*dielectric
waveguide*

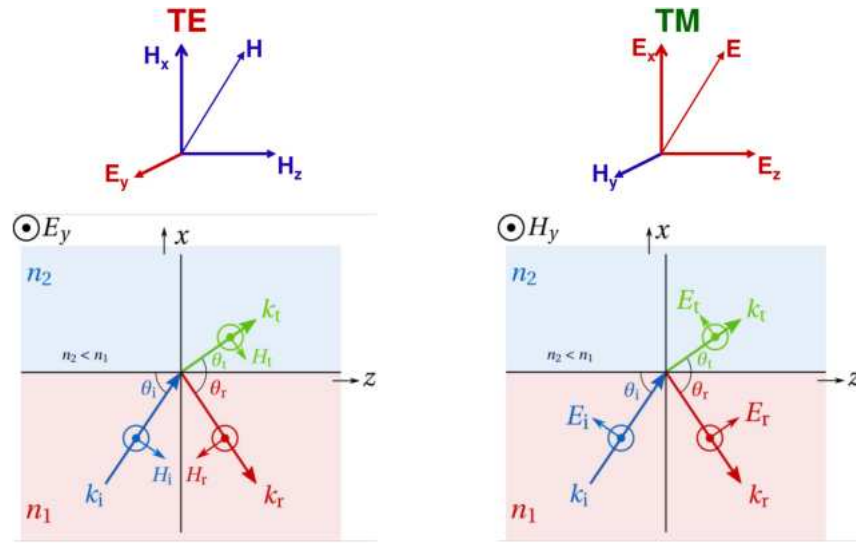


Figure 2.11: TE and TM-polarization .

mission properties of dielectric interfaces in some detail, and apply them to modes in three-layer waveguides in the next section.

For a plane wave which is incident on a plane interface between two different media the reflective and refractive properties of the interface will be dependent on the polarization of the wave. If the E -field is polarized in the y -direction (transverse to the incidence plane, i.e. the plane through the wave vector and its projection on the interface) the wave is called TE-polarized (Transverse Electric), if the H -field is polarized in the y -direction it is called TM-polarized (Transverse Magnetic).

From figure 2.11 it is seen that for a TE polarized wave the reflected and the refracted wave are also TE polarized. The direction of the H -field is changed (it remains perpendicular to the propagation direction) but the direction of the E -field is not affected. The same holds for the H -field of a TM polarized wave. Thus there will be no conversion from TE to TM or vice versa in layered media as long as their refractive index profile is invariant in the y - and z -direction.

The independence of both polarizations can also be shown using the Maxwell equations. If we substitute $\partial/\partial y = 0$ they can be written after some rearrangement:

$$\begin{array}{ll}
 \text{TE} & \text{TM} \\
 E_y = \frac{1}{j\omega\epsilon} \left[\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \right] & H_y = \frac{-1}{j\omega\mu} \left[\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \right] \\
 H_z = \frac{-1}{j\omega\mu} \frac{\partial E_y}{\partial x} & E_z = \frac{1}{j\omega\epsilon} \frac{\partial H_y}{\partial x} \\
 H_x = \frac{1}{j\omega\mu} \frac{\partial E_y}{\partial z} & E_x = \frac{-1}{j\omega\epsilon} \frac{\partial H_y}{\partial z}
 \end{array} \tag{2.35}$$

Inspection shows that the left set comprises only E_y , H_x and H_z , and the right one only H_y , E_x and E_z . There are thus two types of field:

$$\begin{array}{ll}
 \text{TE} & \text{TM} \\
 H_y = E_x = E_z = 0 & E_y = H_x = H_z = 0
 \end{array}$$

which do not couple and can thus propagate independently. Comparison with figure 2.11 reveals that these are the TE- and TM-polarized fields as shown in the picture. From the Maxwell equations as arranged in equation 2.35 it is seen that for both polarizations the x - and the z -components of the field are easily derived from the y -component. The whole field can therefore be described with one single component, the y -component. This is an important conclusion. It means that fields in layered media can be analyzed using the scalar Helmholtz equation, i.e. the Helmholtz equation applied to a single field component. Once a solution for the y -component has been found, the other field components follow from equation 2.35.

scalar Helmholtz equation

We will now discuss what happens at an interface between two dielectric layers. In order to satisfy the Maxwell equations, the field at a discontinuous interface has to satisfy the continuity conditions: the tangential field components (E_y and H_z for a TE polarized field, H_y and E_z for a TM polarized field) have to be continuous through the interface. If the field is described in terms of its y -component, it is useful to translate the continuity condition for the z -components into a condition for the y -component using equation 2.35, so that we arrive at the following continuity conditions:

continuity conditions

$$E_y \text{ and } \frac{\partial E_y}{\partial x} \text{ continuous} \quad \text{TE} \quad H_y \text{ and } \frac{1}{n^2} \cdot \frac{\partial H_y}{\partial x} \text{ continuous} \quad \text{TM} \quad (2.36)$$

For TE-polarization the field and its first derivative thus have to be continuous through the interface, so the field distribution has a smooth transition through the interface. For TM-polarization there will be a step in the first derivative because the term $(1/n^2)$ has a different value at both sides of the interface. The field distribution of a TM-polarized wave thus shows a kink at the transition through the interface. And also the electric field component E_x of a TM-mode, which is directed perpendicular to the interfaces, shows a step at the dielectric interface. This can also be seen from the requirement that the divergence of the dielectric displacement $\nabla \cdot D = \nabla \cdot \epsilon E = \nabla \cdot n^2 E = 0$, from which it follows that $n^2 E_x$ has to be equal at both sides of the interface. As a result the field at the low-index side will peak.

2.3.2 Reflection and refraction

Optical waveguides differ from metallic waveguides in that the wave reflects at a dielectric interface instead of a metallic one. If a plane wave is incident on such a dielectric interface, as depicted in figure 2.12, part of the beam will be reflected and part will be transmitted. The total field at both sides of the interface can be written as:

$$\begin{aligned} U_1 &= U_i + U_r \\ U_2 &= U_t \end{aligned} \quad (2.37)$$

with

$$\begin{aligned} U_i &= e^{-j\vec{k}_i \cdot \vec{r}} \\ U_r &= r \cdot e^{-j\vec{k}_r \cdot \vec{r}} \\ U_t &= t \cdot e^{-j\vec{k}_t \cdot \vec{r}} \end{aligned} \quad (2.38)$$

in which r and t are the (complex) reflection and transmission coefficients. Note that these coefficients refer to the field amplitude: r^2 is the power reflection coefficient, but t^2 is *not* the power transmission coefficient because the refractive index at both sides of the interface is different (analogous to voltage and power measured over different impedances).

reflection coefficient
transmission coefficient
power reflection coefficient

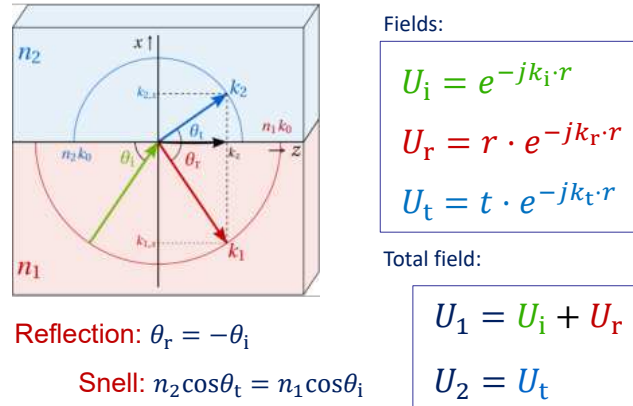


Figure 2.12: Reflection and refraction.

Because of the requirement that the total field be continuous from region 1 to region 2, the wave vectors of the different waves and their amplitudes are coupled. From the requirement that the continuity condition should be satisfied along the whole interface it follows that the z -dependence of all three waves should be equal, which implies that the z -components of their wave vectors have to be equal: $k_{1z} = k_{rz} = k_{tz} = \beta$, in which β can be interpreted as the propagation constant of the field in the z -direction. The whole field can be written as

$$U(x, z) = U(x)e^{-j\beta z} \quad (2.39)$$

in which $U(x) = U_i(x) + U_r(x)$ in medium 1 and $U(x) = U_t(x)$ in medium 2. The wave can thus be described as a field profile $U(x)$ which is independent of z and moves with propagation constant β in the z -direction.

The incident and the reflected field have identical wave numbers, because they propagate in the same medium, so that their transverse wave numbers k_{ix} and k_{rx} must also have equal magnitudes (but opposite sign), which gives us the well known *reflection law*:

reflection law

$$\theta_r = -\theta_i \quad (2.40)$$

From the requirement that k_{tz} and k_{iz} should be equal (remember that $k_{tz} = n_2 k_0 \cos \theta_t$ and $k_{iz} = n_1 k_0 \cos \theta_i$, see figure 2.12) we find Snell's refraction law:

Snell's law

$$n_2 \cos \theta_t = n_1 \cos \theta_i \quad (2.41)$$

The formula differs from the usual form of Snell's law ($n_2 \sin \theta_t = n_1 \sin \theta_i$) because we define the angles θ with respect to the wave propagation direction z instead of to the normal at the surface (the x -direction). From this equation it is clear that no transmission into medium 2 is possible if $n_1 \cos \theta_i > n_2$. For smaller values of θ_i the wave will be totally reflected. The largest angle for which the beam is totally reflected is called the *critical angle*:

critical angle

$$\theta_c = \arccos\left(\frac{n_2}{n_1}\right) \quad (2.42)$$

reflection coefficient
transmission coefficient

The reflection and transmission coefficients r and t can be found from the continuity conditions for the field at both sides of the interface:

Problem 2.7: Critical angle.

Problem: What is the critical angle of (a) an air-glass interface if $n_1 = n_{\text{glass}} = 1.5$ and for (b) an air-semiconductor interface if $n_1 = n_{\text{semi}} = 3.17$?

Solution: (a) $\theta_c = \arccos(1/1.5) = 48^\circ$. (b) $\theta_c = \arccos(1/3.17) = 72^\circ$.

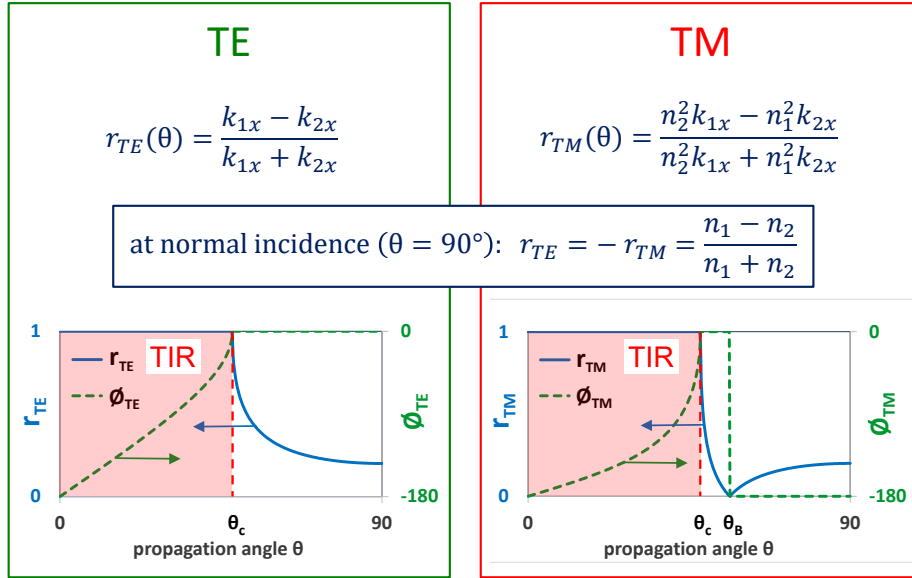


Figure 2.13: Reflection and transmission at a glass-air interface ($n_1 = 1.5$).

$$\begin{array}{ccc}
 \text{TE} & & \text{TM} \\
 E_{iy} + E_{ry} = E_{ty} & & H_{iy} + H_{ry} = H_{ty} \\
 \frac{\partial}{\partial x} [E_{iy} + E_{ry}] = \frac{\partial E_{ty}}{\partial x} & & \frac{1}{n_1^2} \frac{\partial}{\partial x} [H_{iy} + H_{ry}] = \frac{1}{n_2^2} \frac{\partial H_{ty}}{\partial x}
 \end{array} \quad (2.43)$$

After substitution of the field expressions (equation 2.38) these conditions reduce to:

$$\begin{array}{ccc}
 \text{TE} & & \text{TM} \\
 1 + r_{\text{TE}} = t_{\text{TE}} & & 1 + r_{\text{TM}} = t_{\text{TM}} \\
 k_{1x}(1 - r_{\text{TE}}) = k_{2x}t_{\text{TE}} & & n_2^2 k_{1x}(1 - r_{\text{TM}}) = n_1^2 k_{2x}t_{\text{TM}}
 \end{array} \quad (2.44)$$

from which the reflection and transmission coefficients follow as

$$\begin{array}{ccc}
 \text{TE} & & \text{TM} \\
 r_{\text{TE}} = \frac{k_{1x} - k_{2x}}{k_{1x} + k_{2x}} & & r_{\text{TM}} = \frac{n_2^2 k_{1x} - n_1^2 k_{2x}}{n_2^2 k_{1x} + n_1^2 k_{2x}} \\
 t_{\text{TE}} = \frac{2k_{1x}}{k_{1x} + k_{2x}} & & t_{\text{TM}} = \frac{2n_2^2 k_{1x}}{n_2^2 k_{1x} + n_1^2 k_{2x}}
 \end{array} \quad (2.45)$$

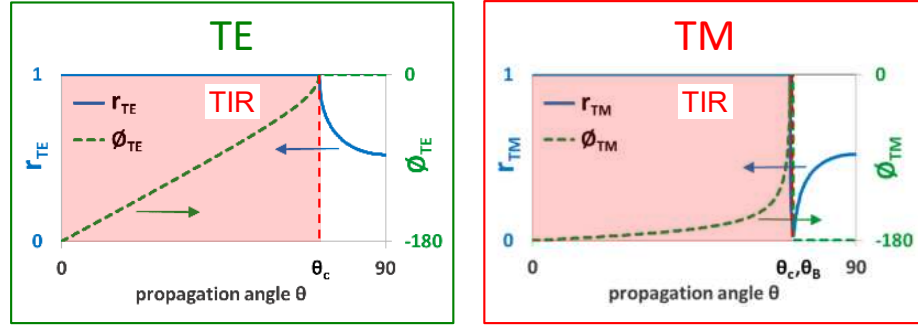


Figure 2.14: Reflection and transmission at a semiconductor-air interface ($n_1 = 3.17$).

If the propagation angle increases to 90° we arrive at the case of normal incidence. For this case the transverse propagation constants $k_{1,x}$ and $k_{2,x}$ equal $n_1 k_0$ and $n_2 k_0$, respectively, and the reflection coefficients reduce to:

$$\begin{array}{cc} \text{TE} & \text{TM} \\ r_{\text{TE}} = \frac{n_1 - n_2}{n_1 + n_2} & r_{\text{TM}} = -\frac{n_1 - n_2}{n_1 + n_2} \end{array} \quad (2.46)$$

We see that at normal incidence the reflection coefficients get the same magnitude, as should be expected because at normal incidence the difference between TE and TM-polarized waves disappears. The opposite signs are due to the fact that r_{TE} refers to the E -field, whereas r_{TM} refers to the H -field. The power reflection coefficient $R = r_{\text{TE}}^2 = r_{\text{TM}}^2$ is known as the Fresnel reflection coefficient:

$$R = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2} \quad (2.47)$$

From equation 2.45 we see that for a TM-polarized wave the reflection becomes zero when $n_2^2 k_{1,x} = n_1^2 k_{2,x}$. This occurs if the propagation angle θ takes the value

$$\theta_B = \arctan\left(\frac{n_1}{n_2}\right) \quad (2.48)$$

Brewster angle in which θ_B is the so-called Brewster angle. It occurs only for TM-polarized light.

Figure 2.13 shows the angular behaviour of the reflection at an air glass interface. We see that up to the critical angle the reflection is 100% both for TE and TM-polarized waves, which is caused by the fact that beyond the critical angle $k_{2,x}$ in equation 2.35 is imaginary. We call this Total Internal Reflection (TIR). In the TIR-region the phase of the reflection gradually increases from -180° for very small propagation angles to 0° at the critical angle. Beyond the critical angle the phase of the reflected signal remains 0° and the reflection coefficient decreases, but for TM-polarisation it goes through zero at the Brewster angle and is negative for large angles.

Figure 2.14 shows the transmission and reflection coefficients for an interface between InP ($n = 3.17$) and air. We see a significant difference with the curves for glass, as shown in Figure 2.13. The critical angle is significantly larger and is close to the Brewster angle. And the difference between the TE and TM polarized curves is much larger, which means that in semiconductor waveguides the difference between TE and TM-polarized modes will be much larger than in glass waveguides.

Problem 2.8: Transmission and reflection.**Problem:**

1. Calculate the power transmission coefficient of an air glass ($n_{\text{glass}} = 1.5$) interface for perpendicular incidence. Do the same for an air semiconductor interface ($n_{\text{InP}} = 3.17$).
2. Calculate the penetration depth for a TE-polarized wave with a wavelength of $1.55\mu\text{m}$, which is incident at an angle $\theta = 13^\circ$ at an interface between two semiconductor layers with $n_1 = 3.36$ (InGaAsP) and $n_2 = 3.17$ (InP).

Solution:

1. $T = 1 - R = 1 - (0.5/2.5)^2 = 0.96$
 $T = 1 - R = 1 - (2.17/4.17)^2 = 0.73$
2. The effective penetration depth follows from $\alpha = \sqrt{\beta^2 - n_2^2 k_0^2} = \sqrt{(n_1 k_0 \cos \theta)^2 - n_2^2 k_0^2} = 3.32\mu\text{m}^{-1}$ as $\Delta x_e = 1/\alpha = 0.30\mu\text{m}$.

Total reflection does not mean that there is no field in medium 2; it means that there is no net power flow into that region, but the field can extend over some distance. To study the field in medium 2, we consider the transverse propagation constant in that medium:

$$k_{1x} = \sqrt{n_2^2 k_0^2 - \beta^2} \quad (2.49)$$

in which $\beta = k_{iz} = k_{tz} = k_{tz} = k_0 \cos \theta$. If the incidence angle is smaller than the critical angle then β becomes larger than $n_2 k_0$ and the transverse propagation constant k_{2x} will become imaginary. If we define an extinction coefficient α according to:

$$\alpha = j k_{2x} = \sqrt{\beta^2 - n_2^2 k_0^2} \quad (2.50)$$

then the field in medium 2 can be written as:

$$U(x, z) = t \cdot e^{-\alpha x} e^{-j\beta z} \quad (2.51)$$

The field propagates in the z -direction with propagation constant β and decays exponentially in the transverse direction, which means that there is no power flow into medium 2. The effective penetration depth (i.e. the depth at which the field is reduced by a factor $1/e$) follows as:

$$\Delta x_e = 1/\alpha \quad (2.52)$$

A field as described above, with plane phase fronts but an exponentially decaying amplitude in a direction transverse to the propagation direction, is called a non-uniform plane wave.

For the case of total internal reflection, the reflection coefficients (equation 2.45) can be rewritten as:

*effective
penetration depth*

*non-uniform plane
wave*

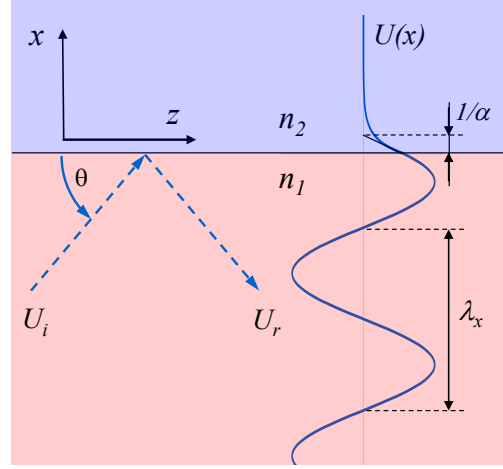


Figure 2.15: Total reflection at a semiconductor-air interface ($n_1 = 3.17, \theta = 50^\circ, \lambda = 1.55 \mu m$)

$$\begin{array}{cc}
 \text{TE} & \text{TM} \\
 r_{\text{TE}} = \frac{k_{1x} + j\alpha}{k_{1x} - j\alpha} & r_{\text{TM}} = \frac{n_2^2 k_{1x} + jn_1^2 \alpha}{n_2^2 k_{1x} - jn_1^2 \alpha}
 \end{array} \quad (2.53)$$

It is seen that $|r_{\text{TE}}| = |r_{\text{TM}}| = 1$, as might be expected. The reflection phase follows from equation 2.53 as

$$\begin{array}{cc}
 \text{TE} & \text{TM} \\
 \Phi_{\text{TE}} = 2 \arctan\left(\frac{\alpha}{k_{1x}}\right) & \Phi_{\text{TM}} = 2 \arctan\left(\frac{n_1^2 \alpha}{n_2^2 k_{1x}}\right)
 \end{array} \quad (2.54)$$

where the factor 2 appears because the numerator and the denominator in equation 2.53 have equal amplitude but opposite arguments.

Figure 2.15 shows the field profile of a wave which is totally reflected at a semiconductor-air interface.

2.4 Wave propagation in three-layer waveguides

2.4.1 Basic concepts of waveguiding

In section 2.2.2 it was shown that between the two plates of a (2-dimensional) metallic waveguide a field of the following form can propagate:

$$U(x, z) = A \cos(k_x x) e^{-jk_z z} \quad (2.55)$$

The field is illustrated in figure 2.16. It can be conceived of as the interference pattern of two plane waves, which produce a standing-wave pattern $U_x(x) = \cos(k_x x)$ with period $\lambda_x = 2d$ into the x -direction, and behave like a traveling wave with wavelength $\lambda_z = 2\pi/k_z$ into the z -direction. So the whole pattern propagates in the z -direction with a z -invariant transverse amplitude profile $U_x(x) = \cos(k_x x)$. The directional wave

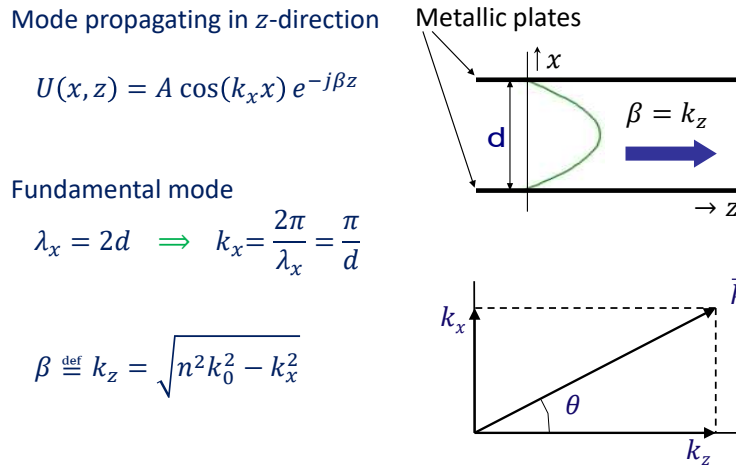


Figure 2.16: Metallic waveguide.

number into the propagation direction of such a wave pattern is usually denoted as the propagation constant β :

$$\beta \equiv k_z = \sqrt{n^2 k_0^2 - k_x^2} \tag{2.56}$$

propagation constant

The directional wave number k_x is called the transverse propagation constant. The interference pattern inside the metallic waveguide may be considered as being caused by a single wave which is reflected between the two plates. This corresponds to the ray concept of a guided wave as a ray which is reflected between two reflecting interfaces. A ray thus represents a (locally) plane wave. Although usually only a single zig-zagging ray is drawn, it should be remembered that both the direct and the reflected wave are present at each point of the waveguide. For understanding waveguide phenomena it will prove useful to apply both the field and the ray approach.

transverse propagation constant ray concept

In a ray approach there are no restrictions on the propagation angles of the rays which propagate between the plates. From the field concept it is clear, however, that the interference pattern of the corresponding plane waves should have zeroes at the positions of the plates, i.e. the transverse standing wave pattern should contain an integer number of (half) periods. Obviously propagation is only possible for a discrete number of wave patterns. These wave patterns are called the modes of a waveguide. The transverse amplitude distribution $U_m(x)$ of each wave pattern is called the mode profile.

mode mode profile

For metallic waveguides the modes are numbered according to the number m of half periods within the waveguides, for dielectric waveguides according to the number of zeroes. The number m is called the order of a mode or mode number. The mode with the lowest order is called the fundamental mode of the waveguide. In metallic waveguides the fundamental mode has order $m = 1$, in dielectric waveguides $m = 0$.

mode order mode number fundamental mode

The first three modes in a metallic waveguide are illustrated in figure 2.17. Because in the following we will restrict ourselves to dielectric waveguide modes we have used the numbering for dielectric waveguides, so that for the transverse propagation constant we find

$$k_{xm} = (m + 1)\pi / d, m = 0, 1, 2, \dots \tag{2.57}$$

Each mode corresponds to a ray angle in the ray theoretic approach; for the m^{th} mode

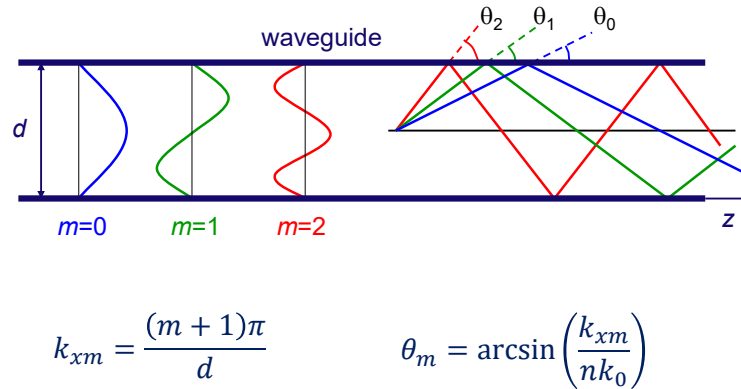


Figure 2.17: Modes and rays.

the angle follows as:

$$\theta_m = \arcsin\left(\frac{k_{xm}}{nk_0}\right) \quad (2.58)$$

Mode excitation

If a broad beam is incident perpendicular to the input of the waveguide, it may be considered locally as a uniform plane wave. At the input of the waveguide the illumination will thus have a rectangular shape. From Fourier theory we know that such a uniform illumination function can be decomposed into a number of cosine functions $\cos((m+1)k_x x)$, $m = 0, 2, 4, \dots$. Note that with a symmetric excitation only symmetric (even) modes can be excited. The anti-symmetric (odd) sine functions with $m = 1, 3, 5, \dots$ are not excited.

The uniform illumination will excite a number of modes, which will propagate independently through the waveguide. The phenomenon of mode excitation can thus be analyzed as the decomposition of the exciting field into the modes of the waveguide. The *excitation coefficients* follow as the Fourier coefficients (or the square of the latter, if we are talking about the power excitation coefficients). Because the different modes propagate with different propagation constants the phase difference between the modes will vary in the z -direction, so that the the transverse field profile becomes z -dependent.

Figure 2.18 illustrates propagation in waveguides where several modes are excited. The left figure illustrates the propagation for the case that a small part of the light is carried by the asymmetric first-order mode. This will happen if the exciting field is not centered. Dependent whether the phase difference between the fundamental and the first-order mode is positive or negative, the field maximum will shift towards one or the other side of the waveguide. As the different modes have different propagation constants, this will cause the propagating field wobbling through the waveguide as depicted in the figure.

The right picture shows the propagation for the case that part of the light is carried by a symmetric second-order mode. This will happen when the exciting field is wider or narrower than the modal field. Depending on whether the phase difference between the fundamental and the second-order mode is positive or negative the total field will show a higher and sharper peak or a lower and broader peak. We say that the mode is

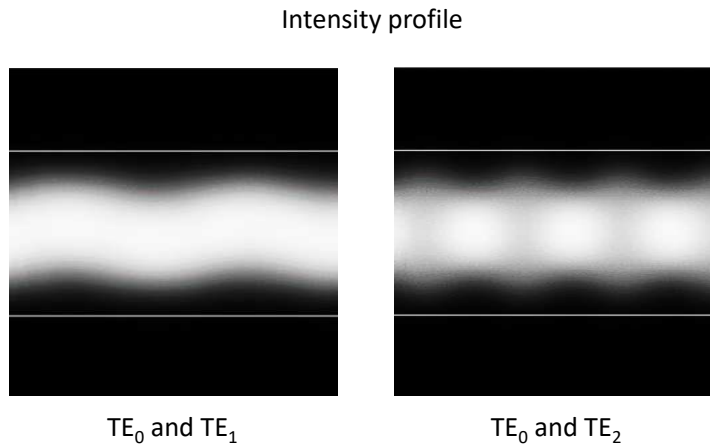


Figure 2.18: Multi-mode propagation in a metallic waveguide

'breathing'. In general a z -dependent field pattern indicates the existence of multiple modes. The more modes are involved the more complex the field pattern will look. The transverse field distribution of the modes themselves, however, will *not* change when they propagate in the z -direction.

Cutoff

From figure 2.17 it is seen that the maximum value of k_x for a guided mode is nk_0 , corresponding to a propagation angle of 90° , which means that physically there is no longer propagation and the wave is resonating vertically between the waveguide walls. The transverse period of a guided mode can, therefore, not become shorter than the plane-wave wavelength λ_0/n . For the uniform excitation mentioned above it is clear, however, that also modes with a shorter period will be excited. From equation 2.56 it is seen that these modes have imaginary propagation constants, i.e. $\beta = \pm j\alpha$, which means that they decay exponentially in the z -direction² according to $\exp(-\alpha z)$ with an effective penetration length (i.e. the length at which the amplitude is reduced by a factor $1/e$) $z_e = 1/\alpha$. These modes are called cut-off, because they cannot propagate over a long distance in the waveguide. However, they are able to penetrate through a short waveguide section with a power transmission coefficient $\exp(-2\alpha L)$ where L is the length of the waveguide section. The latter process is analogous to the tunneling process which is known from quantum mechanics.

2.4.2 Slab guide

The simplest optical waveguide structure is the slab guide: an optically transparent film enclosed between two transparent media with lower refractive indices. The lower medium is called the substrate, the upper one is called the superstrate or upper cladding layer. In such a waveguide a wave can propagate in the same way as in a metallic guide: if the propagation angle is below the critical angle of both waveguide interfaces the wave will be totally reflected in the film.

²A mode which increases exponentially in the propagation direction is excluded in passive media on physical grounds.

Problem 2.9: Guided and cutoff modes.

Problem a: Show that a mode which is cut-off also satisfies the wave equation.

Solution: If we substitute the field $U(x, z) = \cos(k_x x) \exp(-\alpha z)$ in the Helmholtz equation, we find: $\nabla^2 U + k^2 U = \frac{\partial^2}{\partial x^2} U + \frac{\partial^2}{\partial z^2} U + k^2 U = \{-k_x^2 U + \alpha^2 U + k_0^2 U\} = 0$. The field $U(x, z) = \cos(k_x x) \exp(-\alpha z)$ satisfies the Helmholtz equation if $\alpha = \sqrt{k_x^2 - k_0^2}$.

Problem b: How many modes can propagate in the waveguide of figure 2.17 if the wavelength λ_0 is 3 cm and the distance d between the plates is 2 cm?

Solution: For the m^{th} mode: $\lambda_x = 2d/(m+1)$. A mode can propagate if $\lambda_x > \lambda_0$, i.e. $2d/(m+1) = 4/(m+1) > \lambda_0 = 3$. So only the first order mode can propagate, the waveguide is single-mode.

Problem c: What are the propagation constants for the first two modes in this waveguide?

Solution: For the fundamental mode $k_{x,0} = 2\pi/2d \Rightarrow \beta = \sqrt{k_0^2 - k_{x,0}^2} = \sqrt{(2\pi/3)^2 - (\pi/2)^2}$, from which we find $\beta = 1.39$ rad/cm. For the first order mode β is imaginary, so the mode decays exponentially with extinction coefficient $\alpha = \sqrt{k_{x,1}^2 - k_0^2} = \sqrt{(2\pi/2)^2 - (2\pi/3)^2} = 2.34$ cm⁻¹, so the penetration length is $1/\alpha \approx 0.4$ cm.

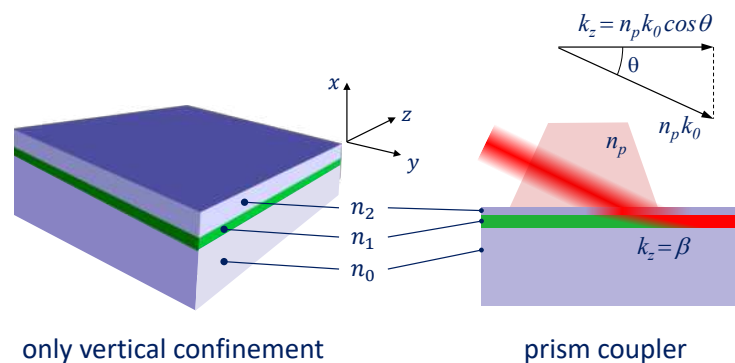


Figure 2.19:

Problem 2.10: Prism coupling.

Problem: A 1 mW laser beam with 1 mm diameter is coupled into a $1\ \mu\text{m}$ thick film with a coupling prism. Estimate the power density in the film (approximate the beam-shape as rectangular). Do the same if we couple the whole beam into a rectangular waveguide with dimensions $2 \times 1\ \mu\text{m}^2$.

Solution: In a coarse approximation the power density in the beam is $100\ \text{mW}/\text{cm}^2$. Behind the prism the beam size is reduced by a factor 1000 in one (the vertical) direction, so the power density is $100\ \text{W}/\text{cm}^2$. If the beam size in the other (the horizontal) direction is reduced from 1 mm to $2\ \mu\text{m}$ the power density is increased to $50\ \text{kW}/\text{cm}^2$.

A slab waveguide can be fabricated by depositing a transparent film (for example a compound glass) onto a substrate with a lower refractive index (for example another compound glass), with a low-index cover layer or air forming the superstrate or upper cladding layer.

In the early days of photonic integration, which was called Integrated Optics at the time, the most practical way to couple light into a planar waveguide was by using a prism with a higher refractive index than the waveguide film, as shown in figure 2.19. Most glass waveguides have refractive indices well below 1.7 and prisms are available with indices up to 2. With such a prism a laser beam can be coupled into a waveguide by adjusting the angle θ of the laser beam such that the z -component of its wave number in the prism ($n_p k_0$) matches the propagation constant β of the waveguide mode, as indicated in figure 2.19. If the beam is phase matched to the mode and the coupling strength is properly controlled by adjusting the gap between the prism and the waveguide, a significant part of the power in the laser beam is coupled to the waveguide [97].

prism

If the film quality is good the beam is seen to propagate several centimeters through the film. Good quality films have propagation losses in the range of 0.1-1 dB/cm. The losses are caused by absorption or scattering at small irregularities of the film medium or the film interfaces. The latter mechanism is responsible for the fact that the guided beam is visible with the eye or with an infrared camera: if there was no scattering, nothing could be seen outside the film. More information about the fabrication process is given in chapter 4.

For semiconductor PICs the waveguide layers are grown on a semiconductor substrate using epitaxial crystal growth. For InP-based PICs the substrate is a monocrystalline InP-wafer, on which composite semiconductor layers (e.g. InGaAsP, which has a higher refractive index) are grown.

If we couple a broad parallel laser beam into the slab guide, we will see that the beam coupled in the slab guide has the same width. If we would have coupled a narrow beam into the film, we would have observed broadening of the beam, subject to the laws of diffraction as discussed in section 2.2. The beam thus behaves as a guided wave in the x - z -plane (perpendicular to the plane of the film) and as a free propagating wave in the y - z -plane (the plane of the film). Obviously, the three-dimensional propagation problem can be separated into two two-dimensional ones. In section 2.5.1 it will be shown that for slab waveguides such a separation yields accurate solutions. As two-dimensional calculations are much faster than three-dimensional ones such a separation leads to a significant reduction of computation times. For three-dimensional waveguides it yields approximate solutions, which will prove adequate for many appli-

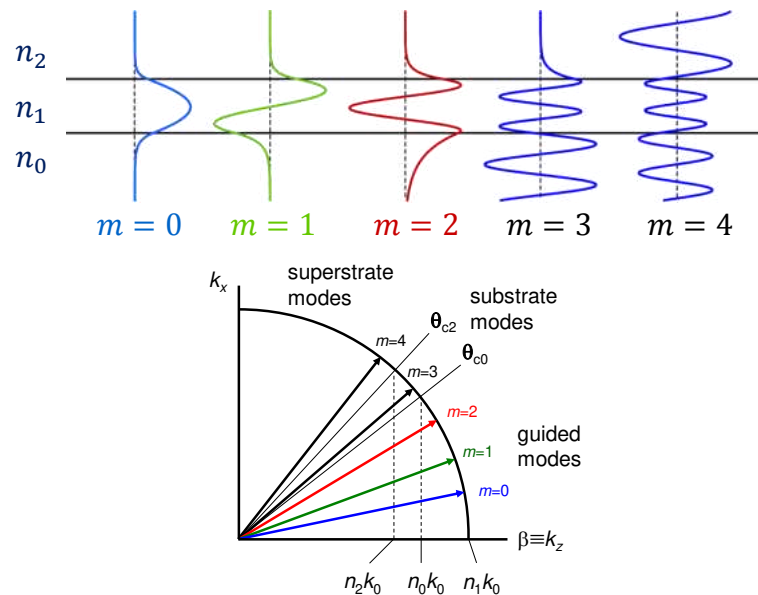


Figure 2.20: Optical waveguide modes in dielectric waveguide.

cations, however.

two-dimensional waveguide In the following subsection we will start with the analysis of the two-dimensional waveguide, i.e. the case of an infinitely broad beam (broad in the y -direction) propagating in a slab waveguide. The results of this two-dimensional analysis may be applied to practical three-dimensional beams which are sufficiently broad for diffraction effects to be neglected ($w \gg \lambda$), and to three-dimensional problems which can be separated into two two-dimensional problems.

2.4.3 Slab modes (two-dimensional optical waveguide modes)

dielectric waveguide Modes in optical waveguides (also called dielectric or open waveguides) show the same cosine shaped standing wave pattern as metallic waveguide modes. The main difference is that the field penetrates into the cladding layers as described in section 2.3.2, figure 2.15. Instead of becoming zero at the edge of the waveguide, the mode has exponential tails. The penetration depth increases when the propagation angle θ_m approaches the critical angle. Similar to metallic waveguide modes the standing wave pattern can have one or more maxima. Figure 2.20 shows the mode profiles corresponding to the first five modes of an asymmetrical waveguide.

mode order Optical modes are labeled with respect to the number of zeroes in the mode profile, this number is called the order of the mode or the mode number. The mode with order zero is called the fundamental mode of the waveguide, different from metallic waveguides where the fundamental mode has order one. Note that a waveguide has two fundamental modes, a TE-polarized and a TM-polarized mode. Modes with the same order but a different polarization generally have different propagation angles and propagation constants due to the fact that the reflection coefficients at the film interfaces are polarization dependent: a difference in the reflection coefficient causes the standing wave pattern in the film and, consequently, its transverse and longitudinal propagation constant to be different. In the most customary nomenclature a mode is labeled

with respect to its polarization and its mode number: the TE_m- and TM_m-mode are the TE- and TM-polarized modes, respectively, of order m .

Modes for which $\theta_m < \min(\theta_{c0}, \theta_{c2})$, in which θ_{c0} and θ_{c2} are the critical angles at the lower and the upper interface, respectively, are called guided modes. These modes experience total reflection at both waveguide interfaces and will propagate without being attenuated if the waveguide media are lossless. For guided modes the propagation angles are thus restricted to the interval $0 < \theta < \min(\theta_{c0}, \theta_{c2})$. From figure 2.20 it is seen that this requirement corresponds to (remember that $\beta = n_1 k_0 \cos \theta$):

$$\max(n_0 k_0, n_2 k_0) < \beta < n_1 k_0 \quad (2.59)$$

In the example of figure 2.20 it is satisfied by the modes with order 0, 1 and 2. If the propagation angle of a mode exceeds the critical angle, the mode is said to be cut-off. The concept of cut-off in an optical waveguide differs from the same concept in a metallic waveguide. In a metallic waveguide, a mode which is cut-off cannot propagate; power can only be transmitted through the waveguide by tunneling (i.e. exponential decay, no phase change). In an optical waveguide a mode which is cut-off can propagate in the same way as a guided mode, the difference being that part of the guided power is refracted into the cladding region, so that the mode will experience attenuation. Such a mode is called a radiating mode. If the propagation angle of a radiating mode comes close to the critical angle, the reflection coefficient approaches 100% and the propagation direction of the refracted beam is almost parallel to the waveguide so that the attenuation of the mode will be small. It may take a long distance for such modes to vanish. From figure 2.20 it is seen that the propagation constants of radiating modes satisfy the following relation: $\beta < \max(n_0 k_0, n_2 k_2)$. For most practical waveguide configurations the refractive index of the superstrate is smaller than or equal to the substrate index: $n_2 \leq n_0$. With this convention, the restrictions for the propagation constant can be written as follows:

$$\begin{aligned} \text{guided modes:} & \quad n_0 k_0 < \beta < n_1 k_0 \\ \text{radiating modes:} & \quad \beta < n_0 k_0 \\ \text{substrate modes:} & \quad n_2 k_0 < \beta < n_0 k_0 \end{aligned} \quad (2.60)$$

A substrate mode is a special type of radiating mode, for which the radiation only occurs into the substrate. Note that the above relations only hold for lossless media. For a film with high loss or high gain (e.g. in a laser), the propagation constants of a guided mode may become lower than the substrate index.

2.4.4 Dispersion relation

In the previous sections the optical waveguide modes were qualitatively discussed. We will proceed with a quantitative analysis which allows us to calculate propagation constants, field distributions and other properties of the modes.

In section 2.3 we saw that for the metallic waveguide only a discrete number of transverse propagation constants k_x , with their corresponding values of β , are possible. By substituting these values of k_x in equation 2.56 we find the corresponding values for β . This relation is known as the dispersion relation. For metallic waveguides the possible values of k_x follow directly from the waveguide thickness because the field strength at the plates should be zero, which corresponds to 180° phase change on reflection. For optical waveguides the dispersion relation is more complicated because the reflection phase is more complicated.

Problem 2.11: Single mode waveguide.

Problem a: What is the maximum thickness for which a silicon-nitride film ($n = 2.0$) embedded between two silicon-dioxide films ($n = 1.44$) is single-mode for $\lambda = 1.55 \mu\text{m}$? Consider only the TE-polarized case.

Solution: The waveguide is single-mode if the first order mode is cut-off. If the first-order mode of a symmetrical waveguide approaches cut-off, its two maxima coincide with the film interface (the penetration depth approaches infinity, so the field derivative in the cladding is zero and, due to the continuity condition, the derivative at the other side of the interface should also be zero). At cut-off the first-order mode profile thus contains exactly one half period, i.e. $\lambda_x = 2d$, and the propagation angle is $\theta = \arcsin(k_x/nk_0) = \arcsin(\lambda_0/n\lambda_x) = \arcsin(\lambda_0/2nd)$. This angle should equal the critical angle $\theta_c = \arccos(1.44/2.0) = 44^\circ$, so that $\lambda_0/4d = \sin 44^\circ$, i.e. $d = \lambda_0/4 \sin 44^\circ \approx 0.56 \mu\text{m}$.

Problem b: The same question for an InP-based waveguide with an InGaAsP film index of 3.36 and InP cladding layers with $n=3.17$.

Solution: The propagation angle is $\theta = \arcsin(\lambda_0/2nd)$. This angle should equal the critical angle $\theta_c = \arccos(3.17/3.36) = 19^\circ$, so that $\lambda_0/2 \cdot 3.36d = \sin 19^\circ$, i.e. $d \approx \lambda_0/6.72 \sin 19^\circ \approx 0.71 \mu\text{m}$. Note that this value is valid if the upper cladding layer has infinite thickness. For a superstrate with finite thickness and air on top the cut-off value will be lower.

The solution of the dispersion equation is fundamental to the solution of most waveguide problems. Once the possible value(s) of k_x and the corresponding value of β have been found, the other mode properties, such as mode profile, effective width, confinement factor, etc., are easily calculated. In this section we will derive the dispersion relation for a lossless three layer waveguide structure. To clarify the relation between the different approaches, the derivation will be done in three ways: with the transverse resonance (wave-decomposition) approach, with the ray-theoretic approach and with the field-theoretic (Maxwell) approach. For a detailed description of the different methods we refer the reader to Unger [96].

transverse resonance **Transverse resonance approach.** The dispersion equation for the metallic waveguide can be interpreted as a transverse resonance condition as depicted in figure 2.21: after being reflected twice, the wave should interfere constructively with itself, i.e. the phase change $-2k_x d$, which the wave experiences on traversing the waveguide twice, increased with the phase shifts Φ_0 and Φ_2 introduced by the reflection, should equal an integer multiple of 2π :

$$2k_x d - \Phi_0 - \Phi_2 = m \cdot 2\pi \quad (2.61)$$

In the example of the metallic waveguide, both Φ_0 and Φ_2 equal π , so that the relation reduces to the simpler form of equation 2.57. According to equation 2.22 the phase shift for an optical waveguide as depicted in figure 2.20 is:

$$\Phi_i = 2 \arctan \frac{\alpha_i}{k_x}, \quad \alpha_i = \sqrt{\beta^2 - n_i^2 k_0^2} \quad i = 0, 2 \quad (2.62)$$

in which

$$k_x = n_1 k_0 \sin \theta, \quad \beta = n_1 k_0 \cos \theta \quad (2.63)$$

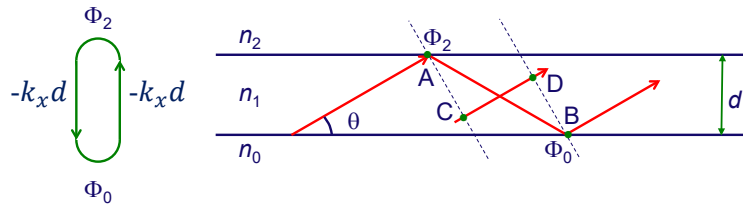


Figure 2.21: Transverse resonance and ray-theoretic approach of a guided wave in a planar waveguide.

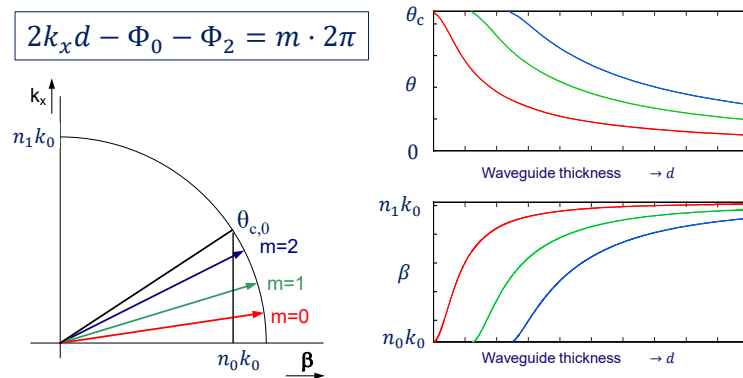


Figure 2.22: Dependence of the propagation angle and the propagation constant on the waveguide thickness.

With the above relations the dispersion equation can be written in terms of k_x , β or θ . It is a transcendental equation due to the dependence of Φ_0 and Φ_2 on θ . Its solutions are easily found, however, with a numerical root finding procedure. Figure 2.22 shows how the propagation angle θ and the propagation constant β depend on the film thickness d for different values of m . Each curve corresponds to a mode of the waveguide. From the figure it is seen that in such a waveguide the fundamental mode is guided for all values of d . For small values of d the propagation constant β is close to $n_0 k_0$, with increasing thickness d it increases to $n_1 k_0$. The higher order modes become guided only after d exceeds a certain value, below that value the mode is cut-off and it will be attenuated through radiation. Close to the cutoff thickness the propagation angle θ is close to the critical angle. The thicker the waveguide becomes, the closer the propagation angle goes to zero, as can be seen from the upper figure.

Ray theoretic approach. In a ray-theoretic approach the dispersion relation is derived by considering the phase difference along different ray trajectories as depicted in figure 2.21. Consider a number of parallel rays that are being reflected and upon reflection experience a phase shift Φ_0 and Φ_2 at the lower and upper interface. The rays are normal to the plane wave phase fronts (dashed lines in the figure). In order to maintain the phase fronts, the phase change over the trajectories AB and CD (the first including two reflections) should be an integer multiple of 2π . From this we find:

$$-n_1 k_0 \cdot (AB - CD) + \Phi_0 + \Phi_2 = -m \cdot 2\pi \tag{2.64}$$

From figure 2.21 we see that $CD = AB \cos 2\theta = AB(1 - 2 \sin^2 \theta)$ and $AB = d / \sin \theta$, from which it follows that $(AB - CD) = 2d \sin \theta$. This brings the equation to the same form as

equation 2.61.

field theoretic approach **Field theoretic approach.** In the field-theoretic approach the possible values of the propagation constant β follow from application of the boundary conditions. If the field in the three layers is written as:

$$\begin{aligned} \text{superstrate: } U_2(x) &= C e^{-\alpha_2(x-d)} \\ \text{film: } U_1(x) &= A \cos(k_{1x}x) + B \sin(k_{1x}x) \\ \text{substrate: } U_0(x) &= A e^{\alpha_0 x} \end{aligned} \quad (2.65)$$

in which we chose $C = U_1(d) = (A \cos(k_{1x}d) + B \sin(k_{1x}d))$, so that the continuity conditions for the field are always satisfied. The continuity conditions (relation 2.36) for the field derivative give us the following identities for TE-polarized waves:

$$\begin{aligned} x = d: \quad \left[\frac{\partial}{\partial x} U_1(x) \right]_{x=d} &= \left[\frac{\partial}{\partial x} U_2(x) \right]_{x=d} \\ &\Rightarrow -A k_{1x} \sin(k_{1x}d) + B k_{1x} \cos(k_{1x}d) = -\alpha_2 (A \cos(k_{1x}d) + B \sin(k_{1x}d)) \\ x = 0: \quad \left[\frac{\partial}{\partial x} U_1(x) \right]_{x=0} &= \left[\frac{\partial}{\partial x} U_0(x) \right]_{x=0} \\ &\Rightarrow B k_{1x} = \alpha_0 \end{aligned} \quad (2.66)$$

With the following substitutions:

$$k_{1x}d = u \quad \alpha_0 d = v \quad \alpha_2 d = w \quad (2.67)$$

these equations can be brought onto the form:

$$\begin{aligned} Av + Bu &= 0 \\ (Au + Bw) \tan u - Bu + Aw &= 0 \end{aligned} \quad (2.68)$$

Combination of both equations gives us the following dispersion relation:

$$\tan u = \frac{v/u + w/u}{1 + vw/u^2} = \tan(\arctan \frac{v}{u} + \arctan \frac{w}{u}) = \tan\left(\frac{\Phi_0}{2} + \frac{\Phi_2}{2}\right) \quad (2.69)$$

with

$$\Phi_0 = 2 \arctan \frac{v}{u} \quad \text{and} \quad \Phi_2 = 2 \arctan \frac{w}{u} \quad (2.70)$$

as defined in equation 2.54. The dispersion relation can thus be rewritten as:

$$u = k_x d = \frac{\Phi_0}{2} + \frac{\Phi_2}{2} + m\pi \quad (2.71)$$

or

$$2k_x d - \Phi_0 - \Phi_2 = m \cdot 2\pi \quad (2.72)$$

which is recognized as the well known dispersion relation (equation 2.61) for TE-polarized waves. For TM-polarized waves the dispersion relation becomes slightly more complicated because the refractive indices appear also in the boundary conditions, whereas for TE-polarized waves they only appear in the transverse propagation constants.

Problem 2.12: Mode properties of a standard waveguide.

Problem: Using a numerical mode solver we find a propagation constant $\beta = 13.25 \mu\text{m}^{-1}$ at a wavelength $\lambda_0 = 1.55 \mu\text{m}$ for the TE₀-mode of a standard InP-based semiconductor waveguide with a film thickness $d = 0.5 \mu\text{m}$, a film index $n_1 = 3.36$ and an index $n_0 = n_2 = 3.17$ for the substrate and the upper cladding layer. Calculate the relative mode intensity at the boundaries of the waveguide and the penetration depth Δx_e into the cladding.

Solution: The transverse wave number k_{1x} follows from $k_{1x} = \sqrt{n_1^2 k_0^2 - \beta^2} = 3.154 \mu\text{m}^{-1}$. In a symmetric waveguide we can write the mode profile as defined in equation 2.65 in the form $U_1(x) = A \cos(k_{1x}(x - \frac{1}{2}d))$, from which we see that the field strength at the film boundaries is $U_1(0) = U_1(d) = A \cos(\frac{1}{2}k_{1x}d)$. So the relative field strength is $\cos(\frac{1}{2}k_{1x}d) = 0.70$ and the relative intensity is 0.50. The effective penetration depth follows analogously from $\alpha_0 = \sqrt{\beta^2 - n_0^2 k_0^2} = 3.23 \mu\text{m}^{-1}$ as $\Delta x_e = 1/\alpha_0 = 0.31 \mu\text{m}$.

Reflection at a cleaved or etched waveguide facet. Because of the polarization dependent reflection behaviour at dielectric interfaces, as shown in figure 2.13, the reflection from a waveguide which is terminated at a cleaved or etched facet is also polarisation dependent. At first glance we might think that the reflection will be polarisation independent if the waveguide is positioned normal to the endface because for normal incidence TE and TM polarized light have the same reflection. However, as we have seen, a mode can be conceived of as the interference pattern of two plane waves with opposite propagation angles. For these angles the TE and TM reflection will be different and, consequently, waveguide modes will experience a polarisation-dependent reflection coefficient at cleaved end facets. The difference will increase if the propagation angle comes closer to the Brewster angle. As we can see from figure 2.13, TE-modes will always show a larger reflection coefficient than TM-modes. An example is elaborated in Problem 2.13.

reflection at a waveguide facet

Mode solvers. There are no analytical solutions for the dispersion relation of a dielectric waveguide. A variety of numerical mode solvers is available for calculating solutions for a given waveguide structure. Each solution represents a mode of the waveguide. The most frequently used methods are described in chapter 6.

mode solvers

2.4.5 Effective index

The wave number $k = nk_0$ of a plane wave is dependent on the vacuum wave number k_0 and the refractive index of the medium in which it propagates. The refractive index is a property not only of the medium, but also of the wave which propagates in the medium. For a plane wave it could be defined as the ratio between the physical values of k and k_0 : $n = k/k_0$. In the same way we can assign an effective index N to a mode which is propagating in the y - z -plane with wave number β :

effective index

$$N = \frac{\beta}{k_0} = n_1 \cos \theta \quad (2.73)$$

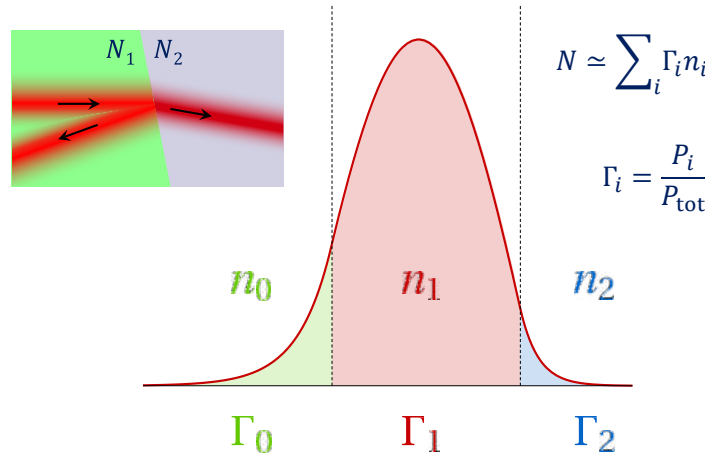


Figure 2.23: (inset at the left) Reflection and refraction of a broad beam in a slab waveguide. (right) Approximating the effective index as a weighted average of the material indices.

Although this definition may seem a little bit artificial at first glance, it has a clear physical interpretation. The wavelength of the mode in the film is λ_0/N , just as for a normal plane wave. Further, if a broad beam, which is (vertically) guided in a planar film, is impinging on an interface between two regions with different effective index, part of the beam will be reflected and part of it will be refracted as sketched in figure 2.23. The discontinuity may be obtained by etching a small step in the waveguiding film, thereby changing its thickness and, consequently, the propagation constant of the guided mode and its effective index. The reflection and refraction behaviour of the vertically guided beam appears to obey the laws of reflection and refraction if we substitute the effective indices N_1 and N_2 into the reflection and refraction formulae instead of the usual refractive index values. The physical interpretation of the effective index may be elucidated in still another way. Figure 2.23 shows the intensity distribution of a mode in a planar film. The power is distributed over all three layers. A fraction Γ_0 propagates through the substrate, a fraction Γ_1 through the waveguide film and a fraction Γ_2 through the superstrate. The mode thus "feels" all three media and its effective index can be considered as a weighted average of the refractive indices of the three media.

Estimating the effective index as a weighted average is not very useful because for calculating the confinement factors we need to know the mode profile, for which we need to know the effective index. However, for estimating absorption or gain of a waveguide with an active layer it is useful. Its absorption or gain will be a weighted average of the absorption or gain of the different media:

$$\alpha \approx \Gamma_0 \alpha_0 + \Gamma_1 \alpha_1 + \Gamma_2 \alpha_2 = \sum_i \Gamma_i \alpha_i \quad (2.74)$$

with

$$\Gamma_i = \frac{P_i}{P_{\text{total}}} \quad (2.75)$$

confinement factor The weight factors Γ_i are called the confinement factors of the films or media, they describe which fraction of the guided power is confined in the film.

The confinement factor is useful for estimating the propagation loss or gain of a mode if the properties of the bulk material are known. If the bulk loss of a doped cladding

layer is 20 dB/cm, and 5% of the modal power is confined in that layer, it will contribute 1 dB/cm (5% of 20 dB/cm) to the modal propagation loss. Also for estimating the gain of a laser it can be used. The loss or gain properties of a mode can thus be estimated by calculating the confinement factors from the real refractive index profile, without using a complex mode solver. The assumption is that the confinement factor (and, consequently, the mode profile from which it is computed) is not dependent on the bulk gain or absorption of the films. For small gains and absorptions (up to several hundred dB/cm) this assumption appears to be reasonable.

In terms of the effective index the restrictions on the propagation constant, as listed in equation 2.60, can be written as:

$$\begin{aligned} \text{guided modes:} & \quad n_0 < N < n_1 \\ \text{radiating modes:} & \quad N < n_0 \\ \text{substrate modes:} & \quad n_2 < N < n_0 \end{aligned} \quad (2.76)$$

Although the propagation constant β , the propagation angle θ and the effective index N are equivalent in describing the properties of a mode, the effective index gives an easy physical indication how well or how weakly guided a mode is: strongly guided modes have an effective index close to the film index, weakly guided modes close to the substrate index.

Effective indices in a standard waveguide stack. Figure 2.24 shows the effective indices of eight modes in a Q1.25 waveguide layer on an InP substrate, as a function of the Q1.25-layer thickness. This graph is representative for the standard waveguide stack of SMART Photonics. The upper two curves show the effective indices of the fundamental TE and TM modes in a symmetric slab waveguide, with an InP top cladding layer. We see that for these modes the effective index starts at the substrate value 3.17, which is to be expected because for very thin waveguide films only a small part of the mode travels through the film and most of the flux is propagating through the InP cladding layers. If the film thickness increases, the effective index increases. For large thicknesses the mode is almost completely confined in the waveguide layer and its effective index goes to 3.36, the bulk-value for Q1.25 material. We see that the TE-mode has a slightly higher index than the TM-mode. For the standard film thickness of 500 nm in the SMART-Photonics platform the effective index is 3.271 for the TE₀-mode and 3.264 for the TM₀-mode.

In a symmetric waveguide the fundamental modes can always propagate, in principle, even for very thin waveguide films. If the waveguide becomes asymmetric, e.g. by removing the top cladding layer, this is no longer true. The two following curves show the fundamental modes in an asymmetric waveguide without an InP top cladding layer in which the waveguide layer is covered with air. Because of the asymmetry, both modes are cutoff for small film thicknesses. Below the so-called cutoff film thickness the mode index becomes smaller than the substrate index, so that the mode radiates into the substrate.

The first order modes in a waveguide can only propagate if the waveguide is thicker than the cutoff thickness. They start with an effective index equal to the substrate index, below which they radiate into the substrate, and increase gradually towards the film index for very large film thickness. Below the cutoff thickness of the first order modes we call the waveguide single mode, even though it carries two modes: the fundamental TE and TM-mode.

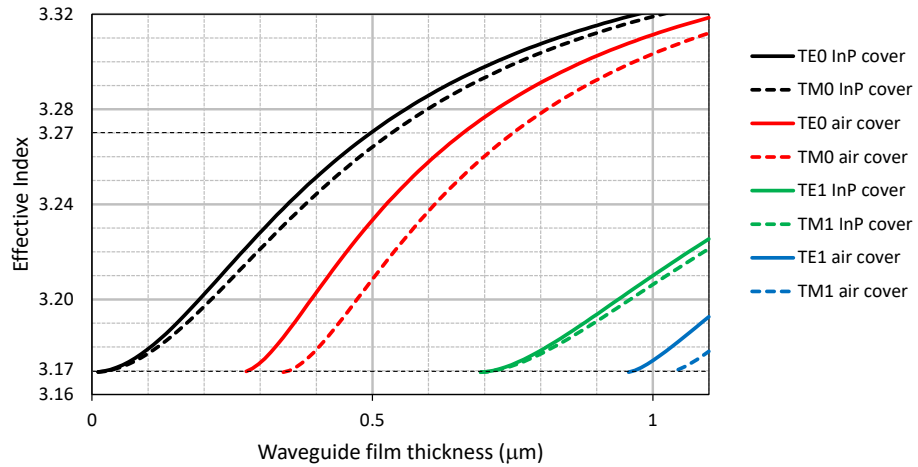


Figure 2.24: Effective indices of the first four modes in a standard slab waveguide with or without InP cover layer, as a function of the Q1.25 layer thickness

Problem 2.13: Effective index.

Problem a: Estimate, using Figure 2.24, the plane-wave propagation angle θ for the TE_0 and the TM_0 -mode in the standard waveguide stack, with $n_0 = 3.17$, $n_1 = 3.36$, a waveguide film thickness $d = 0.5 \mu\text{m}$ and $\lambda = 1.55 \mu\text{m}$.

Solution: From the figure we read that the effective index $N_{TE} = 0.271$ and $N_{TM} = 0.264$. The propagation angles follow from equation 2.73: $\theta_{TE} = \arccos N/n_1 = 13.2^\circ$ and $\theta_{TM} = 13.7^\circ$. e is required.

Problem b: Estimate the plane-wave reflection coefficients at a cleaved facet for these propagation angles using formula 2.45.

Solution: To use formula 2.45 for calculating the reflection at a cleaved end facet we have to replace the angles θ_{TE} and θ_{TM} by their complementary values, i.e. $\theta_{TE} = 76.8^\circ$ and $\theta_{TM} = 76.3^\circ$. For the TE-mode we find $k_{1x} = n_1 k_0 \sin(76.8^\circ) = 13.26$ and $k_{2x} = \sqrt{k_0^2 - n_1^2 k_0^2 \sin^2 13.2^\circ} = 2.60$. With these values we find $r_{TE} = (k_{1x} - k_{2x}) / (k_{1x} + k_{2x}) = 0.67$ and for the power reflection $r_{TE}^2 = 45\%$. For the TM-mode we find $r_{TM} = -0.34$ and $r_{TM}^2 = 12\%$. So there is a large difference between the two polarizations. For 2d waveguide modes the reflection coefficients differ from these values because this simple calculation is valid for plane waves in a medium with index 3.36, it does not account for the effect of the cladding. For practical waveguides the differences are usually smaller, but still significant. For accurate calculations dedicated software is required.

2.4.6 Normalized approach of mode properties

In the early days of photonic integration a popular way to calculate the propagation constant of a waveguide was by applying a normalization to the waveguide parameters, and reading the (normalized) propagation constant from a graph which showed the normalized propagation constant of a number of modes as a function of the normalized waveguide parameters. Although this approach is no longer used since the advent of a variety of numerical mode solvers, it is very useful for providing physical insight in the behaviour of the modes. We recommend to study it and do some exercises to calculate waveguide properties without a numerical mode solver.

In section 2.4.4 we have described how we can find the propagation constants β of the modes of a waveguide as solutions of the dispersion equation (equation 2.61). All other properties of the modes, such as the effective index, the confinement factors and the modal field $U(x, y, z)$, including all 6 field components ($E_x, E_y, E_z, H_x, H_y, H_z$), can be derived from β . For a three-layer waveguide structure β is determined by the following properties of the waveguide: the film thickness d and the refractive indices n_0, n_1, n_2 of the substrate, the waveguide film and the superstrate, respectively.

Normalization for TE-modes. For TE-polarized modes the structure of the dispersion relation (equation 2.61) allows for normalized solutions from which the solutions for a specific waveguide are easily derived. For these modes we can bring the dispersion relation of equation 2.61 to the following normalized form:

$$V\sqrt{1-b} - \arctan\sqrt{\frac{b}{1-b}} - \arctan\sqrt{\frac{b+a}{1-b}} = m\pi \quad (2.77)$$

in which the V-parameter or normalized frequency describes the film thickness normalized with respect to the index contrast: *V-parameter*

$$V = k_0 d \sqrt{n_1^2 - n_0^2} = k_0 d \sqrt{2n\Delta n} \quad (2.78)$$

with $n = (n_1 + n_0)/2$ and $\Delta n = n_1 - n_0$, and the parameter a describes the asymmetry of the waveguide: *asymmetry parameter*

$$a_{TE} = \frac{n_0^2 - n_2^2}{n_1^2 - n_0^2} \quad (2.79)$$

It is assumed that $n_2 \leq n_0 < n_1$. If the waveguide is symmetric, i.e. the substrate and superstrate have the same index ($n_2 = n_0$) then $a = 0$.

The parameter b is a normalized propagation constant:

$$b = \frac{\beta^2 - n_0^2 k_0^2}{n_1^2 k_0^2 - n_0^2 k_0^2} = \frac{N^2 - n_0^2}{n_1^2 - n_0^2} \quad (2.80)$$

normalized propagation constant

in which β is the propagation constant and N is the effective index. The normalized propagation constant b ranges from 0 when the mode is close to cutoff ($\beta \approx n_0 k_0$ or $N \approx n_0$) to 1 when the mode is strongly guided ($\beta \approx n_1 k_0$ or $N \approx n_1$).

A more detailed description of the normalization is given in Appendix 2B on page 2-69.

For TM polarized modes an exact normalization is not possible, but an approximate one can be applied to waveguides with small index contrasts. See Appendix 2B.

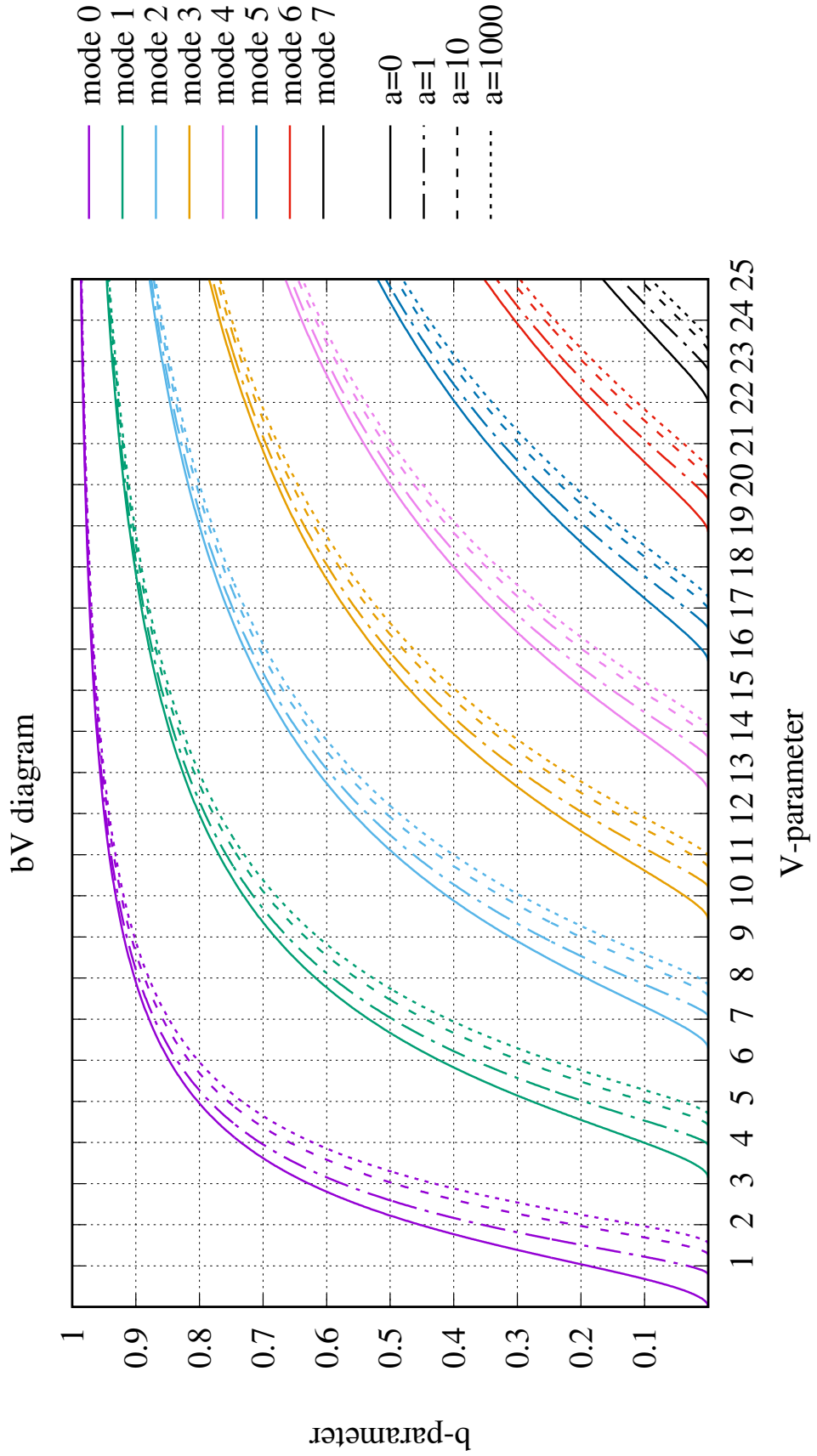


Figure 2.25: The b-V diagram.

Using these normalizations the lower graph in figure 2.21, which shows the dependence of β on the film thickness d , transforms into the normalized graph of figure 2.25, which is called the b-V-diagram. It shows the normalized propagation constant b as a function of the normalized film thickness, the V-parameter, for the guided modes of a waveguide. And for each mode it shows the dispersion curve for different values of the asymmetry. From this graph we can find the propagation constant β of a mode in a given waveguide by first calculating the V-parameter and the a-parameter of that waveguide and reading the corresponding value of b from the graph. The effective index and the propagation constant are then found from:

$$N = \beta/k_0 = \sqrt{n_0^2 + b(n_1^2 - n_0^2)} \quad (2.81)$$

$$\approx n_0 + b(n_1 - n_0) \quad \text{if } \Delta \ll 1 \quad (2.82)$$

in which the relative index contrast is defined as:

$$\Delta = \frac{\Delta n}{n} \approx \frac{n_1 - n_0}{n_1} \quad (2.83)$$

*relative index
contrast*

From the b-V-diagram we see that a symmetric waveguide is single-mode up to a V-parameter value $V = \pi$. For higher values the first-order mode is also guided, and for $V > 2\pi$ also the second-order mode. The V-parameter thus indicates how many (TE) modes can propagate:

$$m = \text{int}(V/\pi) + 1 \quad (2.84)$$

in which the *int*-function takes the largest integer value $<$ its argument. So, for $V < \pi$, $m = 1$ and the waveguide is single mode. From the diagram we also see that asymmetry reduces the effective index and increases the cut-off thickness of a waveguide. Note that most waveguides can carry TE as well as TM-modes, so that the total number of modes roughly doubles to $2m$, with the precaution that the cutoff-wavelengths of TE and TM-modes will be different, especially in the case of large index contrasts.

number of modes

2.4.7 Mode properties

Mode profile. Once the propagation constant β or the effective index N of a mode is known we can calculate the transverse propagation constants $k_{1,x}$, α_0 and α_2 . In a symmetric waveguide as shown in figure 2.26 the field has the shape $U_1(x) = A \cos(k_{1,x}(x - \frac{1}{2}d))$, in the cladding layers it has the form $C_0 \exp(\alpha_0 x)$ and $C_2 \exp(-\alpha_2[x - d])$, in which the coefficients C_0 and C_2 equal $U_1(0)$ and $U_1(d)$, respectively, because of the continuity conditions. The same formulas apply to even modes with a higher order. For odd modes the field in the waveguide is antisymmetric, it is described by a sine function instead of a cosine function, but otherwise it is the same. For asymmetrical waveguides the maximum of the mode profile in the waveguide is shifted. The formulas are summarized in equations 2.85 and 2.86.

*even modes
odd modes*

$$\text{Even mode: } U(x) = \begin{cases} U_1(x) = A \cos(k_{1,x}(x - \frac{1}{2}d) + \Phi_a) & 0 \leq x \leq d \\ U_0(x) = U_1(0) \cdot e^{\alpha_0 x} & x \leq 0 \\ U_2(x) = U_1(d) \cdot e^{-\alpha_2(x-d)} & x \geq d \end{cases} \quad (2.85)$$

Problem 2.14: Properties of a standard waveguide stack.

Problem a: Calculate the effective index at a wavelength $\lambda = 1.55 \mu\text{m}$ for the TE_0 -mode in a $0.5 \mu\text{m}$ thick InGaAsP waveguide with refractive index $n = 3.36$, between an InP substrate and a thick InP top layer with $n = 3.17$.

Solution: The V-parameter of this waveguide is $V = k_0 d \sqrt{n_1^2 - n_0^2} = \frac{2\pi \cdot 0.5}{1.55} \sqrt{3.36^2 - 3.17^2} = 2.258$ while the asymmetry $a = 0$. From the b-V-diagram (figure 2.25) we read $b \approx 0.51$, so that $N = \sqrt{n_0^2 + b(n_1^2 - n_0^2)} = 3.2682 \approx 3.27$. With a numerical mode solver we find $N = 3.2677$ for the TE_0 -mode and $N = 3.2622$ for the TM_0 -mode. So the difference is small.

Problem b: As in 1, but for an air-covered waveguide.

Solution: In this case the asymmetry parameter is $a_{TE} = (n_0^2 - n_2^2)/(n_1^2 - n_0^2) = 7.3$. From the b-V-diagram we find $b \approx 0.3$, and $N \approx 3.23$. With a numerical mode solver we find $N = 3.231$ for the TE_0 -mode and $N = 3.206$ for the TM_0 -mode. Due to the large index contrast between the InGaAsP waveguide layer and the air cover the difference between the TE and the TM mode is significant.

Problem c: How many modes can propagate in the waveguides described in 1 and 2?

Solution: The maximum order follows from $m = \text{int}(V/\pi) + 1 = 1$ for both waveguides. Because this waveguide also guides the TM_0 -mode, the total number of guided modes is 2.

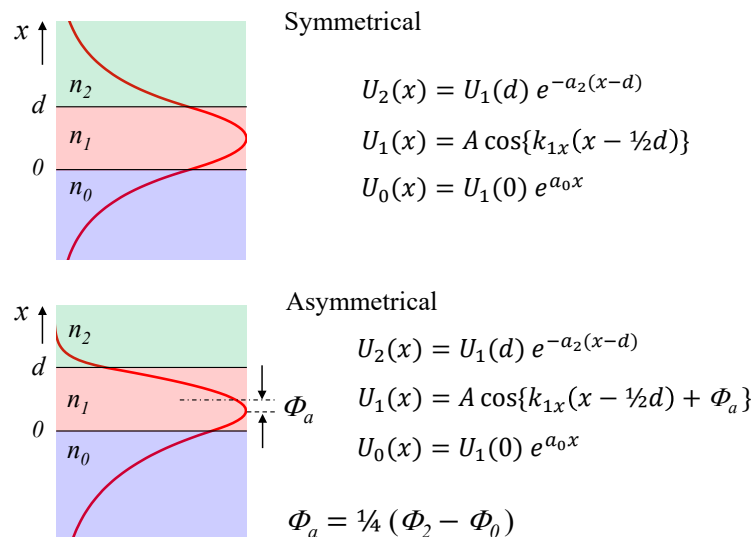


Figure 2.26: Mode profile in symmetric and asymmetric waveguides

Problem 2.15: Phase shift in asymmetrical waveguide.

Problem: Show that the phase shift Φ_a in equations 2.85 and 2.86 equals $\frac{1}{4}(\Phi_2 - \Phi_0)$

Solution: By choosing $U_1(0)$ and $U_1(d)$ as amplitudes for $U_0(x)$ and $U_2(x)$ the boundary conditions for the continuity of the field are automatically satisfied. We can then find Φ_a from the requirement for the continuity of its first derivative. For the derivatives of $U_1(x)$ and $U_0(x)$ at the lower boundary we find (with $A=1$):
 $\left[\frac{\partial}{\partial x} U_1(x)\right]_{x=0} = -k_{1x} \sin(-\frac{1}{2}k_{1x}d + \Phi_a)$ and $\left[\frac{\partial}{\partial x} U_0(x)\right]_{x=0} = \alpha_0 \cos(-\frac{1}{2}k_{1x}d + \Phi_a)$.
 By equating them we find $\tan(-\frac{1}{2}k_{1x}d + \Phi_a) = \frac{\alpha_0}{k_{1x}}$, or $-\frac{1}{2}k_{1x}d + \Phi_a = \arctan \frac{\alpha_0}{k_{1x}}$. In the same way we find for the upper boundary $+\frac{1}{2}k_{1x}d + \Phi_a = -\arctan \frac{\alpha_2}{k_{1x}}$. By remembering that $\arctan \frac{\alpha_0}{k_{1x}} = \frac{1}{2}\Phi_0$ and $\arctan \frac{\alpha_2}{k_{1x}} = \frac{1}{2}\Phi_2$ (equation 2.62) it follows by adding the two equations that $\Phi_a = \frac{1}{4}(\Phi_0 - \Phi_2)$

$$\text{Odd mode: } U(x) = \begin{cases} A \sin(k_{1x}(x - \frac{1}{2}d) + \Phi_a) & 0 \leq x \leq d \\ U_0(x) = U_1(0) \cdot e^{\alpha_0 x} & x \leq 0 \\ U_2(x) = U_1(d) \cdot e^{-\alpha_2(x-d)} & x \geq d \end{cases} \quad (2.86)$$

in which U is either E_y (TE-polarization) or H_y (TM-polarization) and the phase shift $\Phi_a = \frac{1}{4}(\Phi_0 - \Phi_2)$ is found from the boundary conditions, as shown in problem 2.15. For symmetric waveguides it is zero. The waveguide extends from $x = 0$ to $x = d$. For an asymmetrical waveguide the position of the maxima will shift by an amount $\Delta d = \Phi_a/k_{1x}$ from the center.

The other field components (H_x and H_z for TE-polarized modes, and E_x and E_z for TM-polarized modes) follow directly from E_y or H_y using equation(s) 2.35.

Figure 2.27 (left) shows the mode profiles of the fundamental TE and TM mode in the standard waveguides described in problem 2.14: a $0.5\mu\text{m}$ thick InGaAsP waveguide ($n = 3.36$) on an InP substrate ($n = 3.17$) with a thick InP top-cladding layer (symmetric case) or with air on top (asymmetric case). We see that for low contrast the difference between the TE- and the TM-mode is small, the TM-mode has a slightly lower effective index and a lower confinement in the waveguide layer. For high contrast on one side the difference between the TE- and TM-mode increases and both modes are pushed into the substrate.

The right figure shows the H_y - and the E_x -field of the TM-mode in the asymmetric (air-covered) waveguide. We see that the E-field perpendicular to the waveguide interface shows sharp peaks at the waveguide interfaces. The discontinuity in the field scales with the square of the index contrast.

2.4.8 Gaussian approximation of modes.

For a quick estimation of diffraction of modal fields in slab waveguides (in a star coupler, for example) or in free space it is convenient to approximate the modal field as a Gaussian beam. As can be seen from figure 2.28 modes with V-parameters in the range from 2 to 5, which is a customary range for waveguides in PICs, match quite well with Gaussian beams. The translation from the modal field to its Gaussian approximation

Gaussian approximation

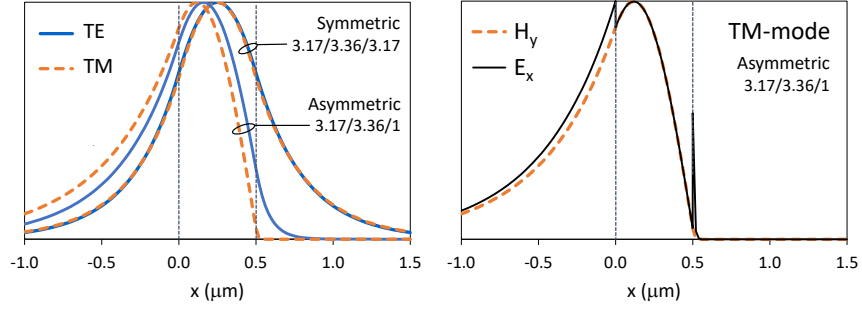


Figure 2.27: Mode profiles in a semiconductor waveguide

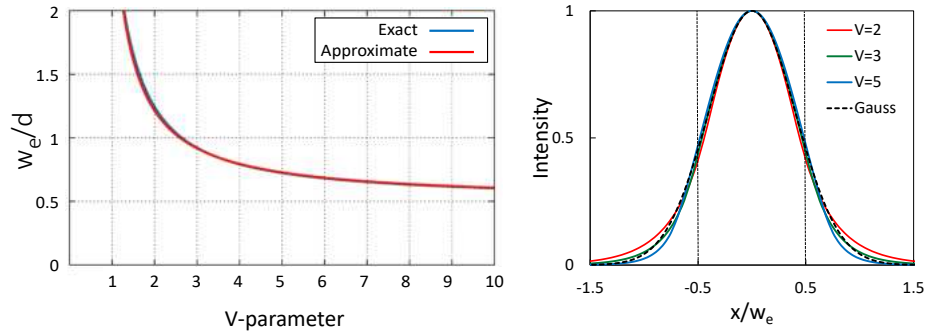


Figure 2.28: The effective width of the TE₀-mode.

is done using equation 2.30.

$$w_0 = w_e \sqrt{2/\pi}$$

Figure 2.28(left) shows the effective width (divided by the waveguide width) of the fundamental TE-mode as a function of the V -parameter for symmetrical waveguides. This curve can be approximated by the following empirical formula:

$$w_e \approx d \left(0.5 + \frac{1}{V - 0.6} \right) \quad (2.87)$$

which gives us an easy means of estimating the effective width and the corresponding Gaussian width of the mode. Figure 2.28(right) shows that the match between the Gaussian beam and the TE-polarized waveguide modes is quite good for V -parameters ranging from 2-5. For small index contrasts it is also useful for TM-modes.

2.4.9 Power carried by a mode.

power of a mode

The power carried by a mode (per unit length in the y -direction) is found from $P = \int E_y \times H_x dx$ (TE) or $P = \int E_x \times H_y dx$ (TM). Expressing the x -components of the field in terms of the y -components yields the following expressions:

$$\text{TE:} \quad P = \frac{\beta}{\omega_0} \int |E_y|^2 dx \quad (2.88)$$

$$\text{TM:} \quad P = \frac{\beta}{\omega \epsilon_0} \int \frac{1}{n^2} |H_y|^2 dx \quad (2.89)$$

Problem 2.16: Mode properties in a standard waveguide stack.

Problem a: Calculate for the waveguide structure described in problem 2.14 the relative field amplitudes at the interfaces of the film for both the symmetrical and the asymmetrical waveguide.

Solution:

The relative field amplitudes follow from equation 2.85 with $x = 0$ and $x = d$. For the symmetrical case $k_{1x} = k_0 \sqrt{n_1^2 - N^2} = \frac{2\pi}{1.55} \sqrt{3.36^2 - 3.27^2} = 3.13$, $\Phi_a = 0$ and $U_1(0) = U_1(d) = A \cos(-\frac{1}{2}k_{1x}d) = 0.709A$, using the effective index values found in problem 2.14. With a numerical mode solver we find $0.702A$, which matches the normalized approach very well. For the asymmetrical waveguide $N = 3.23$ and $k_{1x} = 3.75$. From $\alpha_i = k_0 \sqrt{N^2 - n_i^2}$ we find $\alpha_0 = 2.51$ and $\alpha_2 = 12.45$. From equation 2.62 we then find $\Phi_0 = 2 \arctan(2.51/3.75) = 1.18$ and $\Phi_2 = 2 \arctan(12.45/3.75) = 2.56$ so that $\Phi_a = \frac{1}{4}(\Phi_2 - \Phi_0) = 0.345$. With these values we find $U_1(0) = A \cos(-\frac{1}{2}k_{1x}d + \Phi_a) = 0.830A$ and $U_1(d) = A \cos(\frac{1}{2}k_{1x}d + \Phi_a) = 0.284A$. With a numerical mode solver we find $0.829A$ and $0.288A$, respectively.

Problem b: Estimate the equivalent Gaussian width of the mode of the symmetrical waveguide as described in problem 2.14 (use equation 2.22). Can we use the Gaussian approximation to estimate the divergence of the beam coupled out of this waveguide?

Solution: The equivalent Gaussian width of the fundamental mode can be estimated from the effective waveguide width according to equation 2.87. For the symmetrical waveguide we find $V = k_0 d \sqrt{n_1^2 - n_0^2} = 2.26$ from which we find $w_e \approx d(0.5 + \frac{1}{V-0.6}) = 0.55$ and $w_0 \approx w_e \sqrt{\frac{2}{\pi}} = 0.44 \mu\text{m}$. The divergence of the beam coupled out of the waveguide is calculated from $2\theta_0 = 2 \arcsin(\frac{\lambda}{\pi w_0})$. In this case $\frac{\lambda}{\pi w_0} = \frac{1.55}{0.44\pi} = 1.12$. This is larger than one, which means that the vertical divergence of the outcoupled beam is too large to be estimated with a Gaussian beam approximation. The reason is that because the wavelength in air is more than three times larger than the wavelength in the semiconductor, the width of the mode is smaller than the diffraction-limited spot size in air. In the lateral direction the mode width is usually sufficiently large for the Gaussian approximation to be used.

Substitution of the expressions for the mode profile brings us to the following formula for TE-polarization:

$$\text{TE: } P = |A|^2 \frac{\beta}{2\omega_0} \left(d + \frac{1}{\alpha_0} + \frac{1}{\alpha_2} \right) \quad (2.90)$$

For TM-polarization the formula is more complicated (see Appendix 2B on page 2-69).

Comparison of this formula with the formula for the effective width (equation 2.29) reveals that the formula for P equals w_e apart from a factor $\beta/(\omega_0 E_{\max}^2)$, so that the effective mode width (equation 2.29)

$$w_e = \frac{\int |E_y|^2 dx}{|E_y|_{\max}^2} \quad (2.91)$$

follows as:

$$w_e = \frac{1}{2} \left(d + \frac{1}{\alpha_0} + \frac{1}{\alpha_2} \right) = \frac{d_e}{2} \quad (\text{TE}) \quad (2.92)$$

in which d_e is the effective waveguide width (note that $1/\alpha_0$ and $1/\alpha_2$ are the effective penetration depths into the cladding media).

2.4.10 Phase and group velocity, group index

To calculate the velocity of a wave we have to distinguish between phase velocity and group velocity. The phase velocity v_{ph} of a wave is found as the speed of a plane with constant phase. The group velocity v_{gr} is the speed of the envelope of a wave packet, e.g. a pulse, as shown in Figure 2.29. It represents the speed with which the energy carried by the wave propagates. It cannot exceed the speed of light, whereas the phase velocity can become larger than the speed of light. For a plane wave v_{ph} equals v_{gr} . It is only dependent on the material properties ϵ and μ and their wavelength dependence (material dispersion). For other waves, such as waveguide modes, the phase velocity differs from the group velocity, and its dispersion will not only depend on the material properties but also on the waveguide properties (waveguide dispersion).

The group velocity is the velocity component of the photons in the z -direction: $v_{\text{gr}} = \frac{c}{n} \cos\theta$, in which c is the speed of light in vacuum and n is the refractive index of the waveguide material. From figure 2.4 it is seen that the phase velocity in the z -direction follows from $v_{\text{ph}} = \lambda_z f = \lambda_0 f / \cos(\theta) = \frac{c}{n} / \cos(\theta)$, and thus exceeds the speed of light $\frac{c}{n}$ in the waveguide medium.

More general the phase velocity v_{ph} of a mode is found by considering the phase of the modal field: $\Phi(z, t) = \omega t - \beta z$. The expression for v_{ph} is then found by considering the speed dz/dt of the planes for which $d\Phi/dt = 0$, i.e. $d\Phi/dt = \frac{d}{dt}(\omega t - \beta z) = \omega - \beta(dz/dt) = \omega - \beta v_{\text{ph}} = 0$. From this v_{ph} follows as

$$v_{\text{ph}} = \frac{\omega}{\beta} \quad (2.93)$$

In the same way, the group velocity v_{gr} can be found by considering the propagation speed of the energy contained in the wave, e.g. in the interference pattern of two waves $\exp(j\Phi)$ with $\Phi_1 = \omega_1 t - \beta_1 z$ and $\Phi_2 = \omega_2 t - \beta_2 z$, which can be conceived of as a train of cosine-shaped pulses. The maxima occur at $\Phi_1 = \Phi_2$. Their speed follows from $\frac{\partial}{\partial t}(\Phi_1 - \Phi_2) = 0$, or $\frac{\partial}{\partial t}[(\omega_1 - \omega_2)t - (\beta_1 - \beta_2)z] = \Delta\omega - \Delta\beta \frac{\partial z}{\partial t} = \Delta\omega - \Delta\beta v_{\text{gr}} = 0$, from which we find:

$$v_{\text{gr}} = \frac{dz}{dt} = \frac{d\omega}{d\beta} \quad (2.94)$$

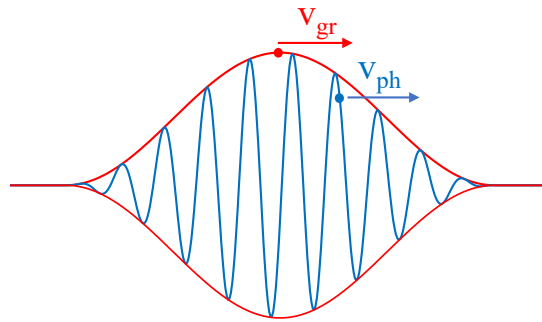


Figure 2.29: Phase and group velocity (1).

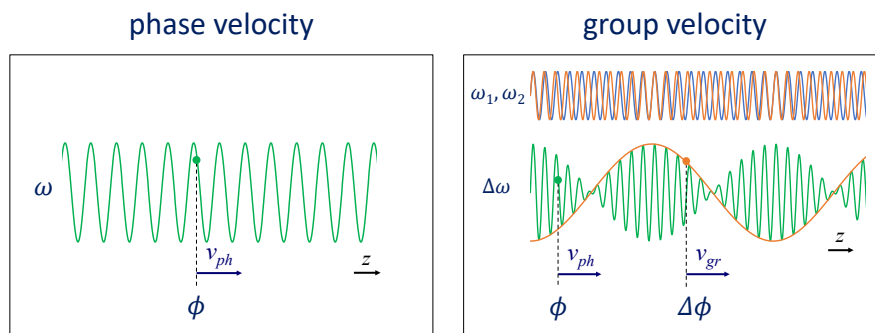


Figure 2.30: Phase and group velocity (2).

The effective index N is related to the phase velocity according to:

$$N = \frac{\beta}{k_0} = \frac{\beta}{\omega} c = \frac{c}{v_{ph}}$$

The phase shift $\Delta\phi$ which a plane wave traveling over a length L through a medium with a frequency-independent refractive index N experiences if we change the frequency with an amount $\Delta\omega = 2\pi\Delta f$ is $\Delta\phi = \Delta\beta L$ in which $\Delta\beta = \Delta\omega/v_{ph} = \Delta\omega N/c$. If the refractive index is frequency-dependent, which is the case for the effective index in a waveguide, then $\Delta\beta = \frac{d\beta}{d\omega}\Delta\omega$. Differentiating $\beta = \omega N/c$ gives $\frac{d\beta}{d\omega} = N/c + \frac{\omega}{c} \frac{dN}{d\omega} = N(1 + \frac{\omega}{N} \frac{dN}{d\omega})/c$. The second term describes the effect of the frequency-dependence of the effective index (dispersion). If we define a group index

group index

$$N_{gr} \equiv N + \omega \frac{dN}{d\omega} \tag{2.95}$$

we see that

$$N_{gr} = \frac{d\beta}{d\omega} c = \frac{c}{v_{gr}}$$

We can thus write the phase shift as $\Delta\phi = L\Delta\beta = L\Delta\omega N_{gr}/c = L\Delta\omega/v_{gr}$.

The change in the phase transmission of a waveguide due to a change in frequency is thus described by the group index N_{gr} rather than the effective (phase) index N , which does not account for the frequency dependence of the propagation constant. This frequency dependence (dispersion) has two components: the refractive index of

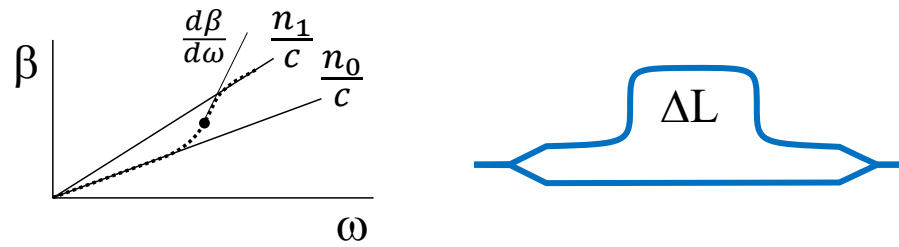


Figure 2.31: Waveguide dispersion

Problem 2.17: Group velocity and phase velocity.

Problem a: Sea waves with a wavelength of 16 m and a propagation speed of 5 m/s are incident on the beach at an angle of 10° relative to normal incidence. What is the wavelength and the phase velocity measured along the beach?

Solution: The component of the wave vector along the beach is $k_x = k \sin \theta = (2\pi/16) \sin 10^\circ = 0.068 \text{ m}^{-1}$, so that the wavelength along the direction of the beach is $2\pi/0.068 = 92 \text{ m}$. The velocity is $v_{ph,x} = v / \sin \theta = 5 / \sin 10^\circ = 28.8 \text{ m/s}$.

Problem b: A Mach-Zehnder Interferometer (MZI) is a circuit in which an input signal is equally divided over two branches, and combined again as depicted in Figure 2.31. If the two branches are equal all input power will be combined in the output waveguide, but if there is a phase difference the signals in the two branches may interfere destructively. The power in the output waveguide can be described as $P_{\text{out}} = \frac{1}{2} P_{\text{in}} (1 + \cos \Delta \phi)^2$. If the two branches have a different length then $\Delta \phi = \beta \Delta L$ and the output power will vary periodically between 100% and 0% when β is varied (by varying the frequency). Calculate the group-index if $\Delta L = 100 \mu\text{m}$ and the measured wavelength difference $\Delta \lambda$ between successive zeroes in the transmission is 6.8 nm.

Solution: $L \Delta \omega N_{\text{gr}} / c = 2\pi \implies \Delta f = \frac{c}{N_{\text{gr}} L}$. $N_{\text{gr}} = c / (L \Delta f) = \lambda^2 / (L \Delta \lambda) = 1.55^2 / (100 \cdot 6.8 \cdot 10^{-3}) = 3.53$

the waveguide material is usually frequency dependent (material dispersion), but even if it is frequency independent the propagation constant of the waveguide is frequency dependent (waveguide dispersion).

The latter is illustrated in figure 2.31. The straight lines show the frequency dependence of the propagation constant of a plane wave propagating through media with frequency-independent refractive indices n_1 and n_0 , respectively. In a symmetric waveguide with core index n_1 and cladding index n_0 the effective index $\beta = \omega N / c$ will be close to n_0 for low frequencies, because for these frequencies the mode is mainly propagating through the cladding. For high frequencies it will propagate mainly through the core. In the transition region it will change from the lower to the upper straight line, as indicated in the figure by the dotted line. We see that in this region the group index, which scales with $d\beta/d\omega$, will become larger than n_1 , contrary to the effective index, which will always remain smaller than n_1 . For a more detailed calculation of the group index of a semiconductor waveguide see Appendix 2C on page 2-70.

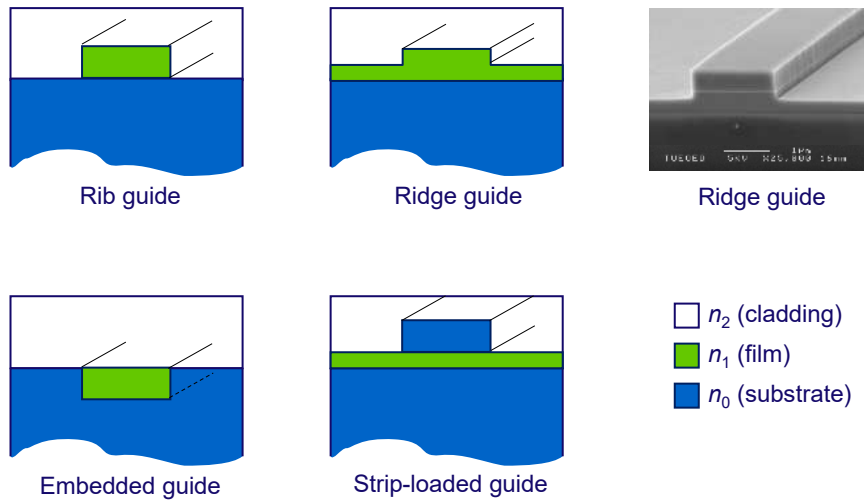


Figure 2.32: Commonly used waveguide types.

2.5 Wave propagation in three-dimensional waveguides

If we want to guide a wave both in the transverse (vertical) and the lateral (horizontal) direction, we need to introduce a lateral refractive index contrast, for example by local introduction (through a mask) of index increasing atoms or ions, or by removing some material besides the waveguide. Figure 2.32 shows some three-dimensional waveguide geometries. Etching processes mostly yield more or less rectangular structures, whereas diffusion processes yield graded structures.

For the computation of propagation constants and mode profiles in three-dimensional structures a number of methods have been developed. Finite Difference methods are most widely applied for computing both scalar and vectorial solutions with good accuracy. This and other methods will be described in chapter 6.

In this section we will discuss an approximate method which has found broad application because of its simplicity: the Effective Index Method (EIM). It was introduced by Knox and Toullos [98] for the analysis of microwave dielectric waveguides and has found wide application for the analysis of optical waveguide structures. For the calculation of mode properties of straight waveguides there is not much advantage in applying an EIM nowadays, because computers are fast enough to use 3d mode solvers. For the analysis of large non-cylindrical structures computation times can still become prohibitive, and the EIM can be applied with advantage to reduce the computation time.

effective index method

2.5.1 Effective Index Method (EIM)

In figure 2.23 it is illustrated that a vertically guided beam, which is incident on an interface between two films with effective refractive indices N_1 and N_2 , is reflected and refracted in the same way as a free propagating beam which is incident on an interface between two media with real refractive indices N_1 and N_2 . Such an interface can be created by covering half of the film with a mask and etching away part of the uncovered film. This observation suggests that the three-dimensional problem may be solved by reducing it to a two dimensional two-media structure in which the effective indices of

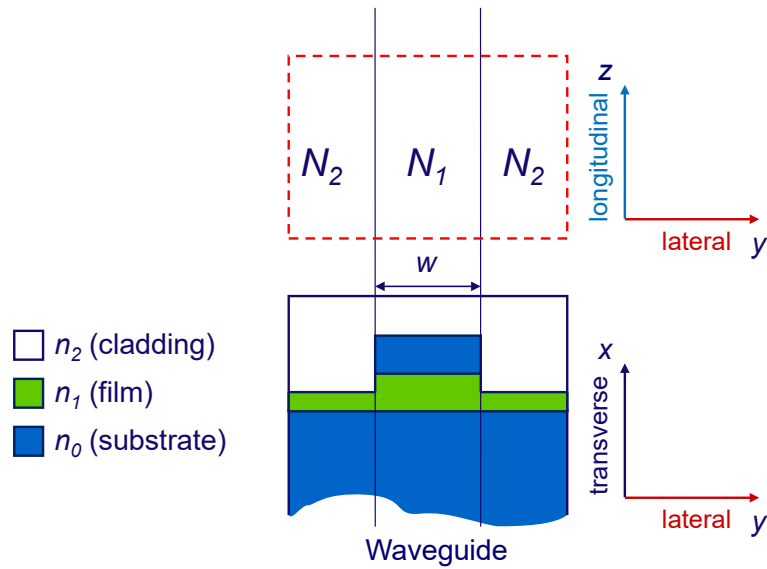


Figure 2.33: The effective index method.

the different regions are treated as the refractive indices of the two-dimensional semi-spaces. This method can also be applied to three-dimensional waveguide structures as illustrated in figure 2.33

The three-dimensional waveguide is first subdivided in three vertical columns, in which the structure is piecewise y -invariant. In each of these columns we compute the local *effective index* of the vertical mode which carries the signal, as if the column was infinitely wide. Next we compute the effective index of the whole structure by considering it as a two dimensional waveguide in which each column is represented by a layer with the column width as thickness and its effective index value as refractive index. Actually we split the three-dimensional problem into two two-dimensional ones: first we solve the transverse problem per column, and next the lateral one in which each column is treated as a layer with the column effective index as refractive index. Because the lateral and the transverse direction are perpendicular to each other, a mode which is TE or TM-polarized in the vertical structure, should be treated as a TM or TE-mode, respectively, in the lateral structure.

Calculating the propagation constant and the mode profile $U(x, y)$ of a mode in a 3-dimensional waveguide with the EIM thus comes to the following steps:

1. Calculate the effective indices N_1 and N_2 in the core and cladding regions using a numerical 2D mode solver and calculate the corresponding mode profiles $X_1(x)$ and $X_2(x)$ in the core and cladding region, respectively. Most practical components work with the fundamental mode in the vertical direction. The vertical waveguide structure is usually single mode. In deep etched waveguides we can use the refractive index of the cladding material (air or a dielectric passivation layer) for N_2 .
2. Calculate the effective index N or the propagation constant $\beta = Nk_0$ of the 3-d mode and the corresponding mode profile $Y(y)$, again with a 2D mode solver, but now with the values N_1 and N_2 as refractive indices of the core and cladding region, and the waveguide width w as film thickness. Note that if the field is

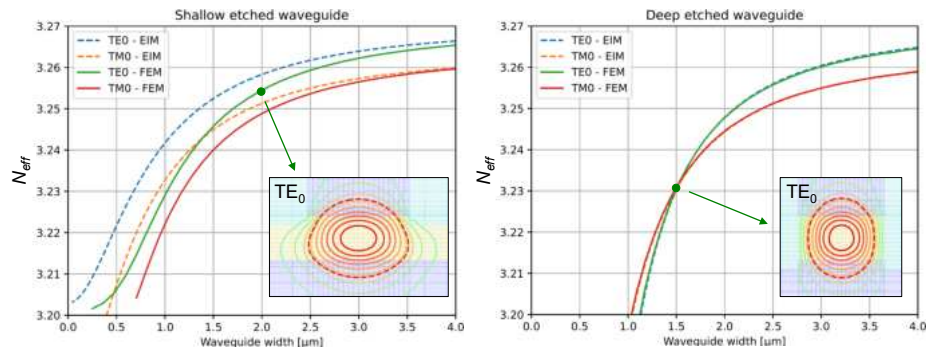


Figure 2.34: Effective indices of the fundamental modes in shallow and deep etched waveguides, calculated with the Effective Index Method and a two-dimensional mode solver.

TE-polarized, the lateral field $Y(y)$ will be TM-polarized.

3. The mode profile $U(x, y)$ follows as $U(x, y) = X_1(x) \cdot Y(y)$ in the core region, and $U(x, y) = X_2(x) \cdot Y(y)$ in the cladding regions. The field profile thus has a discontinuity at the interface between the core and the cladding regions, which is due to the approximative character of the EIM.

For a broad class of ridge-type waveguides the EIM may be applied with advantage. It can also be used for more complex waveguide structures, such as waveguide arrays, as long as they can be separated in piecewise y -invariant structures, which then translate into a multi-layer structure in the horizontal direction. The accuracy of the EIM depends on the specific waveguide configuration and will be discussed below. In Appendix 2D on page 2-71 we give a derivation of the EIM and the conditions under which it applies.

2.5.2 EIM accuracy in a standard waveguide

As shown in Appendix 2D the EIM gives good results if the mode profile $U(x, y)$ can be written as $X(x) \cdot Y(y)$, in which $X(x)$ is the vertical mode profile and $Y(y)$ is the lateral mode profile. This implies that the vertical mode profile is the same everywhere inside as well as next to the waveguide. In general, this holds for modes with circular or elliptical beam profiles.

Figure 2.34 shows the effective indices of the fundamental TE and TM modes in a standard shallow and deep etched waveguide, as computed with the approximative EIM and an exact 2d Finite-Element mode solver (FEM).

For the deep-etched waveguide we see an excellent fit between the EIM and an exact mode solver. For the shallow-edged waveguide the fit is much less: for wide waveguides it is good, but for narrow waveguides it is poor. The reason is the difference in the vertical mode profile at both sides of the waveguide edge. If there is a good match between the fields inside and outside the waveguide the mode shape is more or less elliptical, as we see for the $1.5 \mu\text{m}$ wide deep etched waveguide in 2.34-right. For the mode of the $2 \mu\text{m}$ wide shallow waveguide (left figure) we see a significant deviation from an elliptical shape, which is an indication that the EIM will be less accurate. Caution is thus required when applying the EIM.

Problem 2.18: Properties of a standard ridge guide.

Problem a: Calculate the effective index of the TE₀₀-mode in a 2 μm wide ridge guide, which is fabricated by etching the region besides the waveguide 100 nm into the waveguide film. The waveguide structure and the wavelength are as described in problem 2.14 ($\lambda = 1.55 \mu\text{m}$, $d=0.5 \mu\text{m}$, $n_1 = 3.36$, $n_2 = n_0 = 3.17$).

Solution: Normally, using the EIM, the problem is solved by first calculating the effective mode index inside and beside the ridge, using a numerical 2D mode-solver, and then repeating the procedure in the lateral direction, again with a 2D mode solver, but now with the effective indices inside and beside the ridge as input. In this example, we will use the normalized graphical method for didactical reasons. In problem 2.14 we found that in the unetched region $N_1 = 3.27$.

In the etched region the V-parameter is $V = k_0 d \sqrt{n_1^2 - n_0^2} = \frac{2\pi \cdot 0.4}{1.55} \sqrt{3.36^2 - 3.17^2} = 1.81$ while the asymmetry $a = 7.3$. From the b-V-diagram (figure 2.25) we read $b \approx 0.15$, so that $N_2 = \sqrt{n_0^2 + b(n_1^2 - n_0^2)} \approx 3.20$.

For the lateral waveguide structure the V-parameter is $V_y = k_0 w \sqrt{N_1^2 - N_2^2} = \frac{2\pi \cdot 2}{1.55} \sqrt{3.27^2 - 3.20^2} = 5.5$ while the asymmetry $a = 0$. From the b-V-diagram (figure 2.25) we read $b \approx 0.8$, so that $N = \sqrt{N_2^2 + b(N_1^2 - N_2^2)} \approx 3.256$.^a For comparison: with a numerical 3d mode solver we find $N = 3.2521$.

Problem b: As in 1, if the structure is covered with SiO₂ ($n = 1.44$)

Solution: If the etched region is covered with SiO₂ the asymmetry parameter $a = (n_0^2 - n_2^2)/(n_1^2 - n_0^2) = 6.4$. From the b-V graph we see that we find almost the same value for b which means that the effect of the SiO₂ cover is small. With a numerical mode solver we find: $N = 3.2524$.

Problem c: How many modes can propagate in the waveguide?

Solution: Vertically the waveguide is single-mode. The maximum lateral order follows from $m = 1 + \text{int}(5.5/\pi) = 2$. Because this waveguide also guides the TM₀-mode, the total number of guided modes is 4: TE₀₀, TM₀₀, TE₀₁, and TM₀₁. With a numerical mode-solver we find that also the modes TE₀₂ and TM₀₂ are guided. However, their effective index $N < 3.18$. which means that they are very close to cutoff ($N = 3.17$) and extend deep into the substrate.

^aBecause in the lateral direction the mode should be considered as a TM-mode, we should use the b-V-diagram for TM modes. For small index contrasts the difference with TE is small, however.

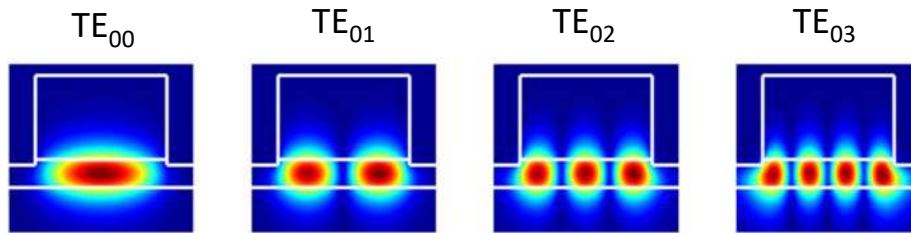


Figure 2.35: Mode profiles in 3-dimensional waveguides

2.5.3 Mode nomenclature

In three-dimensional waveguides, such as depicted in figure 2.32, a mode can be identified with two mode numbers: a transverse and a lateral one. The first number describes the transverse mode order, the second one the lateral order. A mode labeled TE_{mn} is a TE-polarized mode with transverse mode number m and lateral mode number n , i.e. in the transverse direction the mode profile has m zeroes and in the lateral one it has n zeroes. In most practical waveguide components the waveguides are vertically single mode, but in the lateral direction they can support more modes. Figure 2.35 illustrates the electric field of the first four TE-modes in a broad waveguide.

*mode
nomenclature*

Strictly speaking we can no longer talk about TE- or TM-modes because in a three-dimensional waveguide both types have an E_z -component and an H_z -component. Because the transverse confinement is usually much stronger than the lateral one, the magnitudes of these components will be different, so we can speak of a mode which is predominantly TE-polarized or predominantly TM-polarized. In this book the polarization of a mode will be determined as the polarization which the mode would assume when the waveguide width is infinite.

Another nomenclature which is often used in literature labels the modes after the field type of their z -component. So an H-mode is a TE-polarized slab mode, and an E-mode is a TM-polarized one. Modes of three-dimensional waveguides are labeled HE_{mn} or EH_{mn} modes: the first character indicates the transverse polarization state and the second one the lateral polarization state; the mode numbers are as described before. So a HE_{mn} -mode is a TE_{mn} -mode and an EH_{mn} -mode is a TM_{mn} -mode.

*HE mode
EH mode*

In waveguide structures with strong absorption or amplification (complex refractive index) the mode profile becomes complex and although the real and imaginary parts will have zeroes, the complex profile will have no zeroes any longer, so that the nomenclature described above will fail. Complex mode solvers usually do not label the modes but only return their complex effective index or propagation constant. The user has to determine the mode type from the corresponding (complex) mode profile.

2.6 Wave propagation in multi-layer and active waveguides

For finding solutions of the wave equation for multi-layer waveguides or active waveguides with gain or strong absorption, there are many methods, all of which require dedicated software as described in chapter 6. Almost all numerical mode solvers can handle multi-layer waveguides. In Appendix 2E on page 2-72 we describe as an exam-

*multi-layer
waveguides
active waveguides*

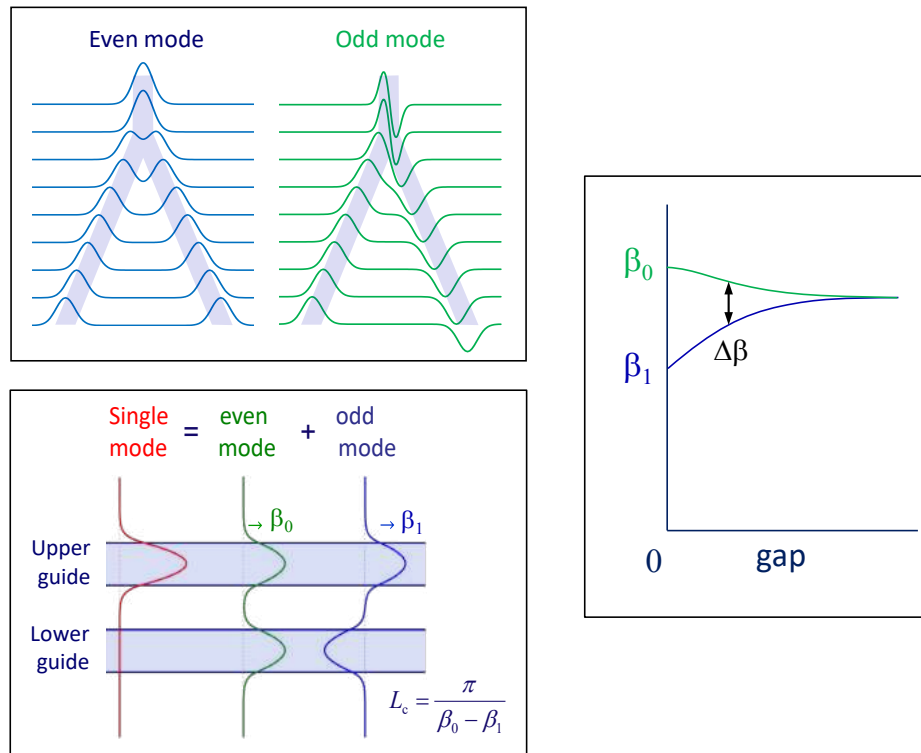


Figure 2.36: System modes in a Y-junction and a synchronous coupler

Transfer Matrix Method ple the Transfer Matrix Method, which is an extension of the method described in 2.4.4. It is also used in quantum mechanics.

In the following sections we will discuss a number of components which can be described as (two-dimensional) multi-layer structures.

2.6.1 System mode theory

system modes System modes are waveguide modes of complex waveguide systems. They do not differ from normal modes, the difference is in the complexity of the waveguide system. Figure 2.36 illustrates the concept of a system mode for a Y-junction and a coupler.

Y-junction A Y-junction can combine the power in two waveguides without loss if the waveguides approach each other sufficiently slowly, so that the modal field can adapt itself to the changing waveguide geometry (adiabatic mode adaptation). If the modes in both input waveguides have the same phase they will combine into the fundamental mode of the waveguide at the end of the junction. If the modes in the two input waveguides have opposite phase, they will combine into the (anti-symmetric) first order mode. So the two modes with equal phase at the input of the junction can be considered as the fundamental or even mode of the system of two uncoupled waveguides. And in the same way the modes with opposite phase can be considered as the first-order or odd mode of the system. Because of reciprocity the field evolution works in both propagation directions.

At any position along the Y-junction where there is a non-zero gap, the system can be considered as a five-layer waveguide: two waveguide layers, two cladding layers and a

gap-layer. Using a mode solver we can calculate the propagation constants of the first two system modes for different values of the gap width or thickness.

The right graph in figure 2.36 shows how the propagation constants of the two system modes depend on the gap between the waveguides. If the gap is large the waveguides (right in the figure) are not coupled and both modes have the same propagation constant, which equals the propagation constant of a single waveguide. If the waveguides come closer the propagation constant of the fundamental mode increases and becomes maximal when the two waveguides touch each other (zero gap); there the waveguide width and the V-parameter are twice as large as for the single waveguide, so the confinement in the waveguide is larger and consequently the effective index and the propagation constant are higher. If we taper the waveguide down to its original width, as shown in the top left figure, β_0 goes back to its original value. The anti-symmetric odd mode does not benefit from an increased coupling because the two parts of the mode extinguish each other in the coupling zone, so that its confinement, and consequently its propagation constant decreases. The behaviour of the coupled modes shows an analogy with the energy levels of atoms in quantum mechanics. If the atoms couple the energy levels split, and the stronger the coupling the larger the splitting.

Also synchronous couplers can be understood and analyzed using system mode theory. A synchronous coupler consists of two identical waveguides which couple over a length L . In two dimensions it can be described as a 5-layer structure. Just as we have seen in the Y-junction, the coupled waveguides support an even and an odd system mode. The modes of the single waveguides are no modes of the system. They can be conceived as the sum of the odd and the even system mode, as depicted in the lower left figure. With a single waveguide mode we excite thus the two system modes. Because they have a different propagation constant their phase relation will change with an amount $\Delta\beta_{01}z$. The length L_c after which the phase difference $\Delta\beta_{01}L_c = \pi$, is called the coupling length. At this length the two modes interfere constructively in the other waveguide, so it is the length after which all input power has coupled to the other waveguide. As the interference is periodic, after a length $2L_c$ all power will be back again in the original waveguide. The coupling length thus follows from:

*synchronous
coupler*

coupling length

$$L_c = \frac{\pi}{\Delta\beta_{01}} = \frac{\pi}{\beta_0 - \beta_1} \quad (2.96)$$

The propagation constants β_0 and β_1 of the first two system modes can easily be computed with a numerical mode solver. The formula also applies to three-dimensional structures.

2.6.2 Coupled mode theory

In the previous section we have seen how coupled waveguides can be analysed using system mode theory. Another approach, which is frequently used, is coupled mode theory. Both approaches have their advantages and disadvantages. In this section we will explain the coupled mode approach and how it is related to the system mode approach.

Synchronous coupling. The z-dependence of a waveguide mode is described by the factor $\exp(-j\beta z)$, so the change in amplitude is described by $\frac{dU}{dz} = -j\beta U$. If the waveguide is weakly coupled to another waveguide, as depicted in figure 2.37, the field in the other waveguide also contributes to the change in amplitude and vice versa:

*synchronous
coupling*

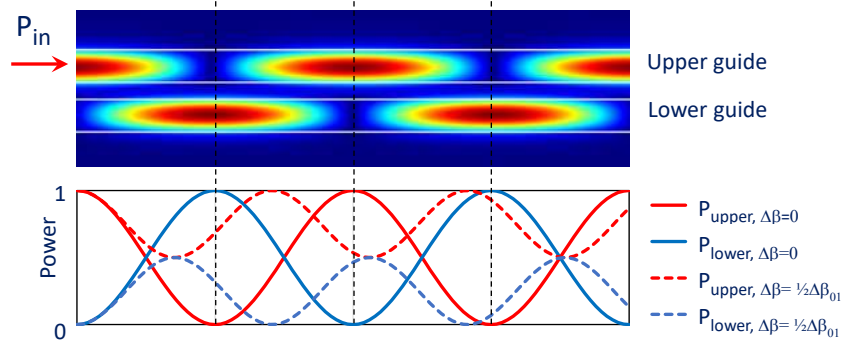


Figure 2.37: Power exchange in coupled waveguides

$$\begin{aligned}\frac{dU_a}{dz} &= -j\beta U_a - jcU_b \\ \frac{dU_b}{dz} &= -j\beta U_b - jcU_a\end{aligned}\quad (2.97)$$

in which the subscripts a and b refer to the upper and the lower waveguide, respectively, and the constant c is called the coupling coefficient. We start with the case of synchronous coupling, i.e. coupling between identical waveguide for which $\beta_a = \beta_b = \beta$. For the case of weak coupling, i.e. $c \ll \beta$, we can write the fields U_a and U_b in both waveguides as

$$U(z) = u(z)e^{-j\beta z}\quad (2.98)$$

in which $u(z)$ is a slowly varying function which describes the change in amplitude due to the coupling. If we substitute this expression into equation 2.97 we can eliminate the fast z -dependence $-j\beta U$:

$$\begin{aligned}\frac{du_a}{dz} &= -jcU_b \\ \frac{du_b}{dz} &= -jcU_a\end{aligned}\quad (2.99)$$

This set of coupled equations has solutions of the form:

$$\begin{aligned}u_a &= Ae^{-jcz} + Be^{+jcz} \\ u_b &= Ae^{-jcz} - Be^{+jcz}\end{aligned}\quad \begin{aligned}A &= B \\ &= 2A\cos(cz) \\ &= -2jA\sin(cz)\end{aligned}\quad \begin{aligned}\implies P_a &= C\cos^2(cz) \\ \implies P_b &= C\sin^2(cz)\end{aligned}\quad (2.100)$$

The coefficients A and B follow from the boundary conditions. With $U_a(0) = 2A$ and $U_b(0) = 0$ we find $A = B$ and the power couples periodically between the two waveguides as described in the above equation and depicted in figure 2.37. From equations 2.100 we see that the period L_{per} of the power exchange follows as $L_{\text{per}} = \pi/c$. The coupling length L_c , which is defined as the length over which all the power is transferred to the other waveguide, is half the period L_{per} :

$$L_c = \frac{\pi}{2c}\quad (2.101)$$

By comparing this formula with the expression that we found from the Systems Mode approach (equation 2.96) we find the following relation between c and the difference $\Delta\beta_{01}$ between the even and the odd system mode:

$$c = \frac{\Delta\beta_{01}}{2} = \frac{\beta_0 - \beta_1}{2}\quad (2.102)$$

Problem 2.19: Coupling between waveguides.

Problem a: For a waveguide coupler consisting of two standard waveguides as described in problem 2.18 with a gap of $1 \mu\text{m}$ between them we find for the propagation constants of the even and the odd system mode the values $13.18315 \text{ rad}/\mu\text{m}$ and $13.18067 \text{ rad}/\mu\text{m}$, respectively. Calculate the coupling coefficient and the coupling length of the coupler.

Solution: From formula 2.102 we find; $c = (13.18315 - 13.18067)/2 = 1.24 \cdot 10^{-3}$ and $L_c = \pi/2c = 1267 \mu\text{m}$.

Problem b: A 3-dB MMI-coupler is a compact coupler which divides the power in each of the two input ports equally over two closely spaced output waveguides with a phase difference of 90° . If the output waveguides are closely spaced they are coupled and power will be transferred between these waveguides following the description of equation 2.100. How much coupling (defined as $\phi_c = \int cz dz$) can we allow between the output waveguides if we want to keep the contribution of the coupling to the imbalance below 0.1 dB?

Solution: At the 3-dB point the coupling phase $\phi_c = \pi/4$, and the power in both waveguides develops as $P_a = \cos^2(\pi/4 + \Delta\phi_c)$ and $P_b = \sin^2(\pi/4 + \Delta\phi_c)$. An imbalance of 0.1 dB occurs at the point where $P_a = 0.5 \cdot 10^{-0.05/10}$ and $P_b = 0.5 \cdot 10^{+0.05/10}$, i.e. $P_a = 0.494$ and $P_b = 0.506$. From $P_a = \cos^2(\phi) = 0.494$ it follows that $\phi = 0.791$ and $\phi_c = \phi - \pi/4 = 0.006$.

Problem c: After which distance does the 0.1 dB imbalance occur if the output waveguides are standard waveguides with a gap of $1 \mu\text{m}$, as described in 1?

Solution: It occurs if $\phi_c = c\Delta z = 0.006$, from which we find: $\Delta z = 0.006/c = 0.006/1.24 \cdot 10^{-3} \approx 5 \mu\text{m}$. For a good performance of the coupler it is thus important to increase the gap between the output waveguides as fast as possible by using curved output waveguides.

The coupling coefficient can be calculated from the overlap between the fields of the coupled waveguides. However, this requires fairly complicated computations. For the case of weak coupling ($c \ll \Delta\beta$) an approximate analytical expression has been derived (Unger [96]). For high-index semiconductor waveguides this formula is not accurate enough, however, so the most easy and straightforward way for calculating the coupling coefficient is by using equation 2.102.

Asynchronous coupling In synchronous couplers 100% of the power is periodically transferred. If the two coupled waveguide do not have the same propagation constant, then there is still a periodic power transfer between the waveguides, but the transferred power is $<100\%$.

For this case the power transfer between the waveguides can be described as follows:

$$\begin{aligned} P_a(z) &= C(1 - \rho \sin^2 c'z) \\ P_b(z) &= C\rho \sin^2 c'z \end{aligned} \quad (2.103)$$

in which

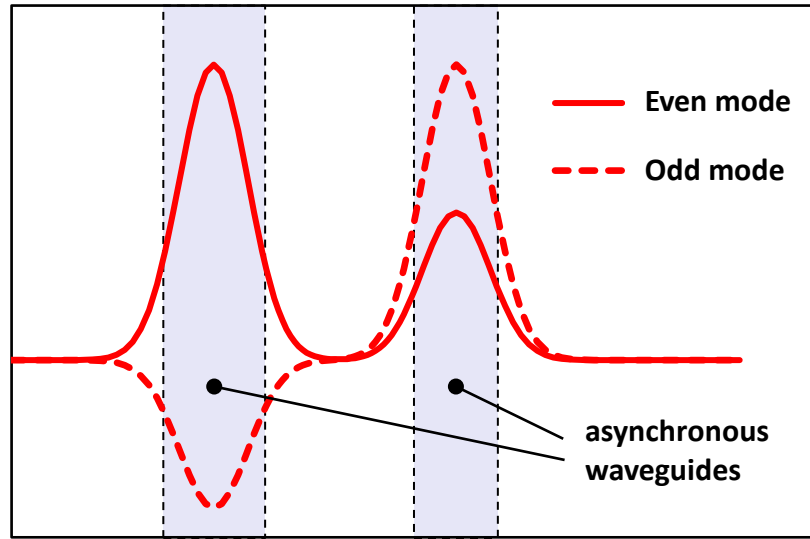


Figure 2.38: System modes in asynchronous coupled waveguides.

$$c' = \sqrt{c^2 + (\Delta\beta_{ab})^2}$$

$$\rho = c^2 / c'^2$$

Note that here $\Delta\beta_{ab}$ is the mismatch between the propagation constants of the fundamental modes in the two waveguides $\Delta\beta_{ab} = |\beta_a - \beta_b|$, different from $\Delta\beta_{01}$, which is the difference between the propagation constants of the even and the odd system mode. Due to the mismatch $\Delta\beta$ the coupling length becomes shorter and a smaller fraction of the power is transferred. Figure 2.37 illustrates the asynchronous coupling behavior for the case that $\Delta\beta = c = \frac{1}{2}\Delta\beta_{01}$.

The effect of asynchronism can also be studied with the System Mode approach. Figure 2.38 shows the mode profile of the odd and even system modes in two waveguides with different widths and, consequently, different propagation constants. Due to the different widths, the fundamental mode gets its highest peak in the widest waveguide, whereas the first order mode peaks in the narrower waveguide. Due to the difference in mode profile the two modes can no longer extinguish each other, and the power transfer between the waveguides becomes incomplete, as illustrated in Fig. 2.38.

2.6.3 Waveguides with absorption or gain

Photodetectors, optical amplifiers and lasers usually consist of a waveguide stack which contains an absorbing or an amplifying layer. The absorption or gain can be described with a complex refractive index n_c

$$n_c = n - j\kappa \quad (2.104)$$

in which n is the real part of the refractive index and the imaginary part κ is called the extinction coefficient, often denoted as k , but to avoid confusion with the wavenumber k we use the Greek letter κ . So a positive value of κ denotes absorption and a negative

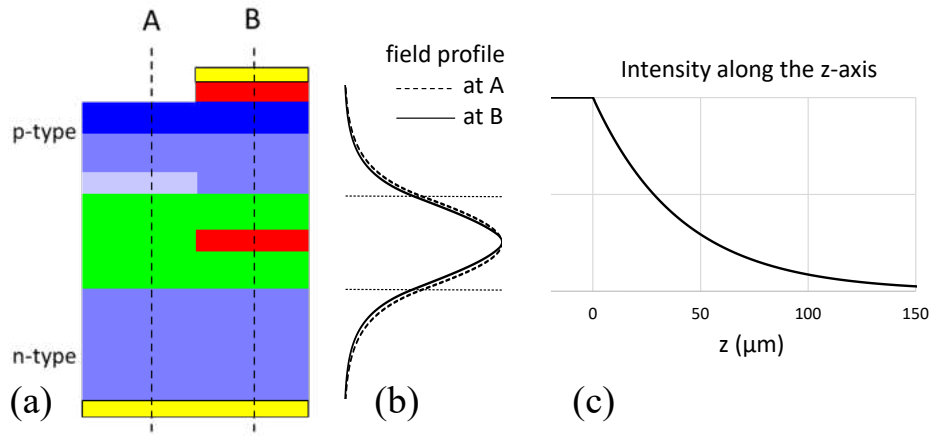


Figure 2.39: a) Waveguide detector structure. b) Mode profiles in the transparent (A) and the absorbing (B) waveguide region. c) Intensity profile along the propagation axis.

value gain. The propagation of a plane wave in a medium with refractive index n_c is described by:

$$\begin{aligned}
 U(z) &= U_0 e^{-jn_c \frac{2\pi}{\lambda_0} z} = U_0 e^{-\alpha z} e^{-j\beta z} \\
 \text{with } \alpha &= 2\pi\kappa/\lambda_0 \\
 \beta &= 2\pi n/\lambda_0 \\
 \text{and } I(z) &= |U(z)|^2 = U_0^2 e^{-2\alpha z}
 \end{aligned} \tag{2.105}$$

Note that α denotes here the extinction coefficient of the field and $1/\alpha$ the $1/e$ extinction length. The absorption coefficient for the modal power is 2α and the absorption length $1/2\alpha$ is the length over which the power decreases by a factor $1/e$. Typical values for an absorbing InGaAsP layer (@1550 nm) are $n = 3.6$ and $2\alpha = 4000 \text{ cm}^{-1}$, from which it follows that for bulk material $\kappa \approx 0.05$ and the absorption length is $z_{1/e} = 1/2\alpha = 2.5 \mu\text{m}$.

absorption length

For the computation of propagation constants and mode profiles in waveguides which contain absorbing or amplifying layers we need numerical mode solvers which can handle complex refractive indices, as described in chapter 6.

Detector Figure 2.39a illustrates the structure of a waveguide photodetector. It consists of an absorbing waveguide section (B) which is connected to a transparent waveguide (A). The transparent and the absorbing waveguide sections share the doped cladding layers, except for an undoped layer on top of the transparent waveguide which is included to keep the strongly absorbing p-dopant at some distance of the waveguide mode. In the detector section the dopant is extended to the waveguide layer in order to keep the electrical series resistance of the photodetector low.

The main difference between the transparent and the detector waveguide is an absorbing layer in the center of the waveguide. For a guided mode in the detector waveguide the absorption length will be significantly larger than in bulk material, because the confinement in the absorbing layer is small, typically less than 10%. We can calculate the absorption length by calculating the complex propagation constant

complex propagation constant

$$\beta_c = \beta + j\alpha \tag{2.106}$$

Problem 2.20: Detector absorption length.**Problem:**

Estimate the absorption length in a waveguide detector with an InGaAsP layer with refractive index $n = 3.6$ and extinction coefficient $2\alpha = 4000 \text{ cm}^{-1}$, if the confinement factor Γ of the mode in the absorbing layer is 5%.

Solution: With $\alpha_{mode} = \Gamma\alpha_{bulk}$ we find $\alpha = 0.05 \cdot 2000 \text{ cm}^{-1} = 0.01 \mu\text{m}^{-1}$, from which it follows that $z_{1/e^2} = 1/\alpha = 100 \mu\text{m}^{-1}$

absorption length with a complex mode solver, as described in chapter 6. The $1/e$ absorption length follows as

$$z_{1/e} = 1/2\alpha \quad (2.107)$$

Because for a detector waveguide the imaginary part κ of the complex effective index N_c , is much smaller than its real part N , the perturbation of the mode profile by the imaginary part will be small. So for most practical cases we can calculate the mode profile in the detector neglecting the imaginary part of the refractive index profile. Figure 2.39b shows the profile of the TE₀-mode in the transparent waveguide (A) and in the absorbing waveguide (B), for the foundry process of SMART Photonics. As we can see, there is a small mismatch between the two mode profiles. Because of the higher index of the active layer the field in the detector is slightly more confined in the waveguide layer. However, the difference is small; the theoretical coupling efficiency at the junction (for the vertical profile) is better than 99.9%, which corresponds to a coupling loss < 0.01 dB.

Figure 2.39c shows how the intensity of the mode decreases in the detector waveguide. We see that with a length of $150 \mu\text{m}$ more than 97 % of the modal power is absorbed, mainly by conversion into electrons and holes, which cause an electrical current in the external circuit.

Optical amplifier In the SMART Photonics platform the structure of the detector and the optical amplifier are identical. It is optimized for the amplifier, but it also provides a detector with good performance. As a consequence the vertical mode profile for the detector and the optical amplifier are identical (if they have the same widths). The gain in the active layer is caused by carrier injection. It can be modeled with a negative value of κ . Also in an optical amplifier $\kappa \ll N$ and the vertical mode profiles in the detector and the optical amplifier are in good approximation the same. A difference between the amplifier and the detector are the relatively high carrier concentrations in the active layer, which reduce the refractive index as described in chapter 3.

Waveguides with metal electrodes In a detector or an optical amplifier the absorption loss in the metal electrodes is usually small when compared to the absorption or the gain of the active layer. In phase modulators it is much more important because they are usually designed for low insertion loss. The structure of a phase modulator is very similar to that of a detector or an optical amplifier as shown in figure 2.39a, the main difference being that it does not have the active layer shown in section B. It often has the same structure as the transparent waveguide shown in section A, but with metal electrodes and an InGaAs contact layer.

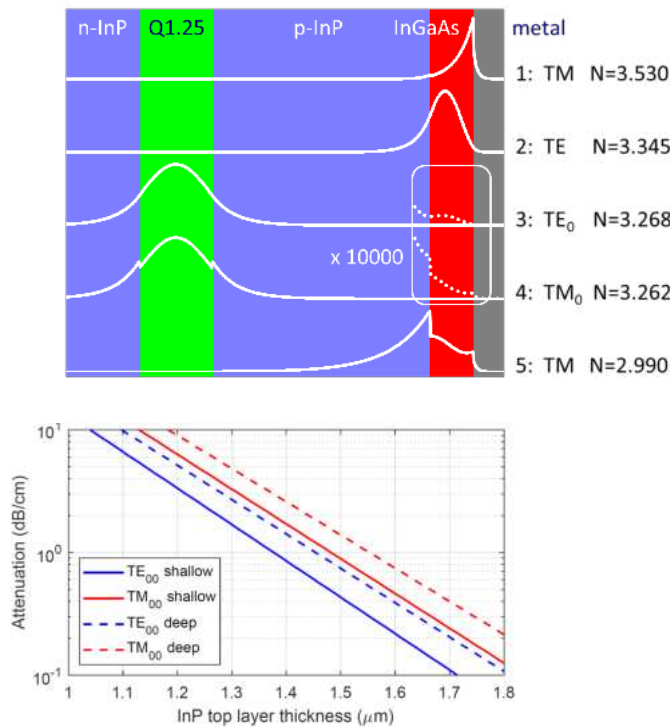


Figure 2.40: Upper: Mode profile in a waveguide with metal electrodes. Lower: Dependence on the top-layer thickness of the contributions of the InGaAs and the metal contact layer to the propagation loss.

Figure 2.40 (upper) shows the electric field intensity for five different modes in the waveguide stack of the SMART Photonics foundry, calculated with a complex mode solver. The stack consists of an n-doped substrate, a 0.5 μm thick Q1.25 waveguide layer, a 1.5 μm thick (p-doped) InP layer, a 0.3 μm thick InGaAs layer for lowering the contact resistance, and a metal contact consisting of a titanium-platinum-gold stack. A more detailed description is given in Appendix C. The waveguide layer itself has low losses, but the 300 nm thick InGaAs layer and the metal layers on top are highly absorptive and will contribute significantly to the propagation loss of the waveguide modes through the tail of the modal field, if the top InP-layer is not thick enough.

The metal electrode supports a so-called plasmonic mode, which has two exponentially decaying tails at both sides of the metal-semiconductor interface, as shown in the curve labeled 1. Its effective index is higher than that of the dielectric guided modes. In normal circumstances it will not be excited, but it will be returned as the mode with the highest effective index by a complex mode solver. The mode labeled 2 is the fundamental TE-polarized waveguide mode of the high-index InGaAs layer. Also this mode will normally not be excited.

The modes labeled 3 and 4 are the fundamental TE and TM modes of the Q1.25 waveguide layer. The main loss contribution for these modes comes from the p-doped InP-layer and is in the order of 1 dB/cm. Through the tails of the field distribution also the InGaAs layer and the metal layers contribute to the loss. The inset shows a 10,000x magnification of the intensity profile. Especially the InGaAs layer will contribute if the InP top-layer is not thick enough. The mode labeled 5 is the fundamental TM-

polarized waveguide mode of the InGaAs layer. Normally it will not be excited. It is seen that for these complex modes the effective index can become lower than the substrate index ($n_{\text{InP}}=3.17$).

Figure 2.40(lower) shows the contribution of the 300 nm thick InGaAs contact layer and the metal electrode to the propagation loss of standard shallow and deep etched waveguides as a function of the top layer thickness. From the figure it is seen that the TM-polarized modes experience a higher loss than the TE-polarized modes. For the standard top-layer thickness of 1.5 μm the contribution to the propagation loss in a shallow etched waveguide is < 1 dB/cm both for TE and TM-polarized modes. In a deep etched waveguide the loss of both modes is a bit higher but remains well below 2 dB/cm for the TM-mode and below 1 dB/cm for the TE-mode. In modulators, which are usually only a few millimeters long, this is not a big problem. However, in longer waveguides it is important to remove the InGaAs top layer.

2.6.4 Curved waveguides

curved waveguide Curved waveguides are most frequently used for creating connections between different components. We want to make them as short as possible, but if the radius becomes too small they introduce radiation loss. Figure 2.41 shows the basic geometry of a curved waveguide or waveguide bend.
waveguide bend

The radius of a waveguide bend can be defined in different ways. It is most practical to define it as the radius R_t of the outer edge of the waveguide, for reasons which will become clear later. In a bend the phase fronts of the mode follow the curvature and the propagation of the mode can be described with an angular propagation constant β_ϕ , which means that the wavelength $\lambda_\phi(r)$ in the tangential direction increases linearly with r : $\lambda_\phi(r) = r\beta_\phi$. The local wave number $k(r)$ is related to the angular wave number $k_\phi = \beta_\phi$ as

$$\begin{aligned} k(r) &= \sqrt{k_\phi^2 + k_r^2} \\ k_\phi(r) &= 2\pi/\lambda_\phi(r) = \beta_\phi/r \\ k_r(r) &= \sqrt{N_1^2 k_0^2 - k_\phi^2} \\ \alpha_r(r) &= \sqrt{k_\phi^2 - N_2^2 k_0^2} \end{aligned} \quad (2.108)$$

Just outside the waveguide the tangential wave number $k_\phi(r)$ is larger than $N_2 k_0$, in which N_2 is the local effective index outside the waveguide, so that α_r is real and the mode will decay exponentially³ in the radial direction. With increasing radius k_ϕ decreases and at a certain radius (R_c) it will equal $N_2 k_0$. For larger radii α_r becomes imaginary, which means that k_r becomes real and the mode starts radiating. Curved waveguide modes are, therefore, fundamentally lossy. If the field intensity at R_c is small, the radiation loss will be small. With decreasing bending radius R_t the critical radius R_c will come closer to the waveguide and at a certain value of R_t the radiation losses will get unacceptably high.

conformal transformation **Conformal transformation.** For a quantitative analysis of wave propagation in curved waveguides we will use the conformal transformation approach introduced by Heiblum and Harris [99]. In this approach the index profile $N(r)$ of the waveguide in the cartesian (y, z) coordinate system is transformed into a profile $N_t(r)$ in a cylindrical

³For curved waveguide modes the field in the waveguide and in the cladding region is described by a Bessel function and a modified Bessel function, respectively. If $R \gg w$ they look similar to the cosine and exponential profiles that we are used to for straight waveguides, in the vicinity of the waveguide.

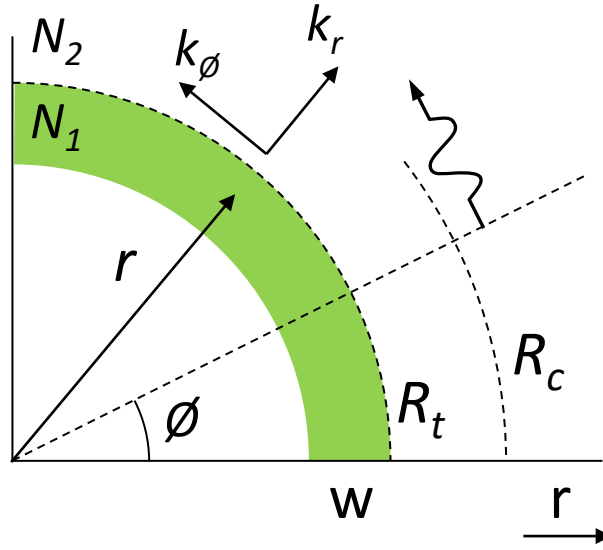


Figure 2.41: Curved waveguide geometry

(r, ϕ) coordinate system, such that the Helmholtz equation takes the same form as in the cartesian coordinate system. Figure 2.42 illustrates the approach. As a first step we transform the index profile as follows:

$$N_t(r) = N(r) \cdot e^{(r-R_t)/R_t} \quad (2.109)$$

Then we calculate the propagation constant β_t for the transformed index profile in a Cartesian coordinate system with coordinates r and ϕ . This can be done with a normal mode solver. Next we transform the so found propagation constant β_t into the radial propagation constant β_ϕ according to:

$$\beta_\phi = \beta_t R_t \quad (2.110)$$

The mode profile follows as:

$$U(r) = U_t(r) \cdot e^{-(r-R_t)/R_t} \approx U_t(r) \quad \text{if } (r - R_t) \ll R_t \quad (2.111)$$

The derivation of this approach goes as follows. In polar coordinates (r, ϕ) the Helmholtz equation looks as follows:

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} + N^2(r) k_0^2 \right] U(r, \phi) = 0$$

For solutions of the form $U(r, \phi) = U(r) e^{-j\beta_\phi \phi}$ it transforms into:

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \{N^2(r) k_0^2 - \frac{\beta_\phi^2}{r^2}\} \right] U(r) = 0$$

If we make the following substitutions: $u = R_t \ln \frac{r}{R_t}$ and $\beta_t = \beta_\phi / R_t$ it takes the following form:

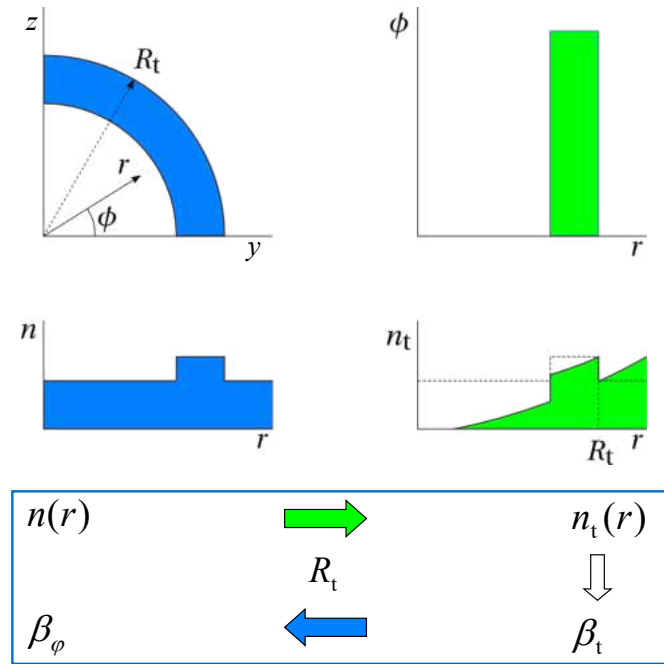


Figure 2.42: The conformal transformation approach

$$\left[\frac{\partial^2}{\partial u^2} + \{N_t^2(u)k_0^2 - \beta_t^2\} \right] U_u(u) = 0 \quad (2.112)$$

in which:

$$\begin{aligned} u &= R_t \ln \frac{r}{R_t} \approx r - R_t \\ N_t(u + R_t) &= N(r) e^{r/R_t} \approx N(r) \cdot (1 + u/R_t) \\ \beta_\phi &= \beta_t R_t \\ U(r) &= U_t(r) \end{aligned} \quad (2.113)$$

With the above transformation formulas the propagation constant and the mode profile of curved waveguides can be calculated using a normal 2d mode solver with a staircase approximation of the transformed effective index profile.⁴

Figure 2.43 shows the transformed index profile and the corresponding mode profile for the fundamental mode in four waveguides with increasing curvature. The thick red line shows the mode in a straight waveguide, the other curves show the effect of increasing curvature. In curved waveguides we have two loss mechanisms in addition to the absorption and scattering loss which occurs in all waveguides types.

⁴For application of this method to TM-modes there is a complication. In the boundary conditions for TM modes the refractive index plays a role: it requires continuity of $(1/n^2) dU/dr$. This condition applies to the original index profile, and not to the transformed one. If we apply a normal mode solver to the transformed profile, it will use the transformed index profile also for the boundary conditions. This will introduce an error. For small contrasts the error will be small. And for TE-modes this problem does not occur. However, the assumption that the fields in the real and the transformed domain are equal is also an approximation, because the difference in refractive index between the real and the transformed domain causes a difference in power density. However, if $u/R_t \ll 1$ this error will be small.

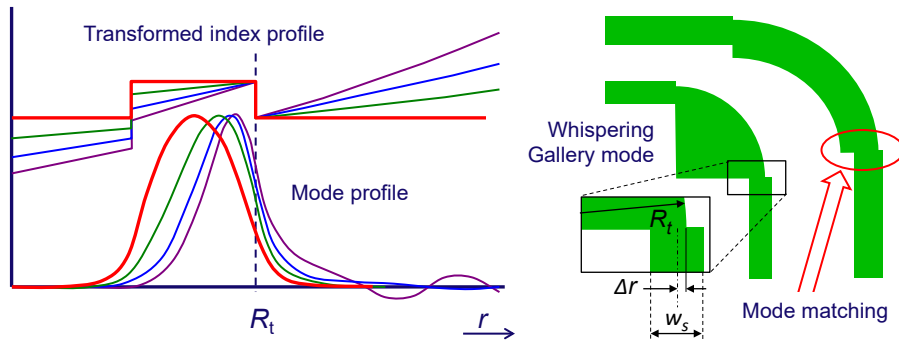


Figure 2.43: Modes in curved waveguides

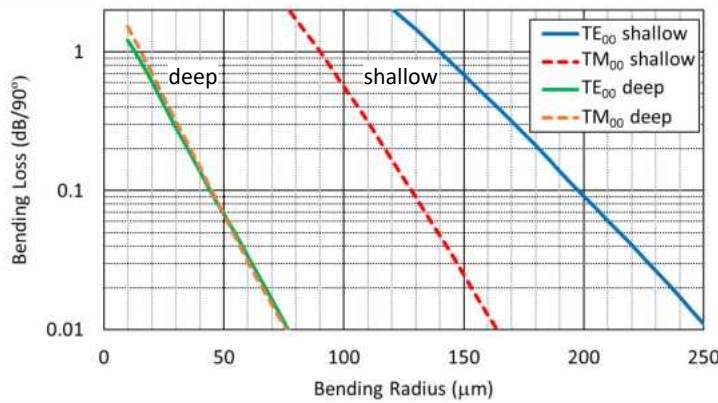


Figure 2.44: Bending loss for the fundamental modes in standard shallow and deep etched waveguides.

The first one is radiation loss. For the mode in the waveguide with the highest curvature we see a clearly radiating part in the outer cladding, such a mode has high radiation losses. The radiation starts at the radius R_c for which the refractive index in the transformed domain equals the effective index in that domain: $N_t(R_c) \geq N_e = \beta_t/k_0$. The radiation or bending loss follows from the complex part of the propagation constant which is found by the mode solver:

$$A_\phi = 20 \log \left(\beta_\phi'' \frac{\pi}{2} \right) \text{ dB/90}^\circ \tag{2.114}$$

in which β_ϕ'' is the imaginary part of the angular propagation constant $\beta_\phi = \beta_\phi' + j\beta_\phi''$. More information on radiation losses can be found in chapter 9.

The second loss mechanism is the mode mismatch loss at the junctions between straight and curved waveguides. We see that the mode profile is compressed and shifted towards the outer edge of the waveguide. The coupling loss at the junction between a straight and a curved waveguide is minimal if the mode profiles in the straight and curved waveguide match as good as possible. The match can be optimized by choosing the width of the straight waveguide such that the effective width of its mode is matched to that in the curved waveguide, and by applying an offset between the waveguides such that the modal fields are aligned, as indicated in figure 2.43. More information about optimization of waveguide junctions is provided in chapter 10.

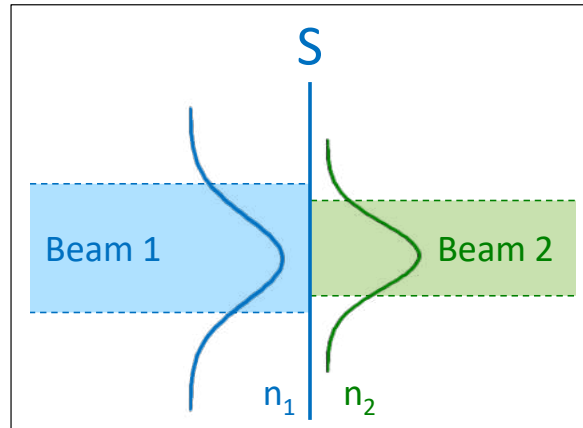


Figure 2.45: Two beams coupling through a surface S.

Figure 2.44 shows the radiation loss of the fundamental modes in shallow and deep etched standard waveguides as a function of the bending radius. As we see from the figure for shallow waveguides a bending radius of $200 \mu\text{m}$ is a safe design, for a deep-etched waveguide bending radii down to $50 \mu\text{m}$ can be applied.

Whispering Gallery Modes

Whispering Gallery Modes. From figure 2.43 it is seen that modes in waveguides with a very strong curvature are guided only by the outer edge of the waveguide: they do not 'feel' the inner edge any longer. This inner edge can, therefore, be omitted. So a circle-sector with a sufficiently small radius, as shown in the figure, can guide a mode. Such a mode is called a Whispering Gallery Mode (WGM), after its analogy with acoustic waves which can be guided over a large distance by a circular gallery. In reference [100] it is shown that on proper design the coupling loss between a straight waveguide and a WGM-bend can be well below 0.1 dB, so that a WGM-bend is an attractive solution for designing compact low-loss waveguide bends.

2.7 Beam coupling

In this section we describe the coupling of two beams through a surface S, as shown in Figure 2.45. Such coupling can occur on the chip, e.g. between two waveguides which are not perfectly aligned or matched, between a passive and an active waveguide, or between a waveguide and a diffracted beam in a free propagation region. And it can also occur at the edge of the chip, when coupling light from a (lensed) fibre or a microscope objective into a waveguide on the chip.

coupling efficiency overlap integral

The power coupling efficiency η between two beams can be estimated by calculating the overlap between the beams over a surface S using the overlap integral:

$$\eta_c = \frac{\left| \int^S U_1(\mathbf{r}) \cdot U_2(\mathbf{r}) ds \right|^2}{\int^S |U_1(\mathbf{r})|^2 ds \cdot \int^S |U_2(\mathbf{r})|^2 ds} \quad (2.115)$$

The surface S can be straight but it may also be curved.

For elliptical beams, which we can write as $U_1(\mathbf{r}) = U_1(x, y) = U_{1x}(x) \cdot U_{1y}(y)$, the coupling coefficient is the product of the 1d-overlap integrals in the x and y-direction:

Problem 2.21: Coupling efficiency between Gaussian beams.

Problem: Calculate the 1d-overlap between two Gaussian beams with different widths w_1 and w_2 and an offset Δy .

Solution: We substitute the following expressions for $U_1(y)$ and $U_2(y)$ in Eq. 2.115:

$U_1(y) = A_1 e^{-\frac{(y-\Delta y)^2}{2w_1^2}}$ and $U_2(y) = A_2 e^{-\frac{y^2}{2w_2^2}}$ with $A_i = \sqrt{\frac{1}{w_i\sqrt{\pi}}}$, $i = 1, 2$, so that w_1 and w_2 are the $1/e^2$ -widths of the intensity distribution $U^2(y)$. The normalisation factors A_1 and A_2 are chosen such that the integrals in the denominator equal unity. This gives us the following expression for the coupling efficiency: $\eta = \left| \int_{-\infty}^{\infty} A_1 \exp[-\frac{(y-\Delta y)^2}{2w_1^2}] \cdot A_2 \exp[-\frac{y^2}{2w_2^2}] dy \right|^2$. Using some calculus we can simplify this to the following expression:

$$\eta = \frac{2w_1 w_2}{w_2^2 + w_1^2} \exp \left[-\frac{\Delta y^2}{(w_2^2 + w_1^2)} \right]. \quad (2.118)$$

$$\eta_c = \eta_{cx} \cdot \eta_{cy} \quad (2.116)$$

If the refractive index at both sides of the surface is different, such as at the cleaved end face of the chip we have to account for the fact that part of the incident power will be reflected. The power coupled to the transmitted beam can then be estimated from

$$\eta_T = (1 - R)\eta_c, \quad (2.117)$$

where R is the (power) reflection coefficient.

For very weakly confined beams the Fresnel reflection coefficient (Eq. 2.47) can be used. However, for the standard waveguides as described in Appendix C, which have a strong vertical confinement, more rigorous (numerical) analysis, such as described in Chapter 13 is required. If the endface is AR (Anti-Reflection) coated, the reflection can usually be neglected for calculating the transmitted power.

Figure 2.46 illustrates the three most important forms of beam misalignment or mismatch: a lateral offset, a width mismatch or an angular mismatch. For all three cases simple analytical expressions are available for the coupling loss if both beams can be approximated as Gaussian beams.

Waveguide modes in standard waveguides can be described very well as Gaussian beams, as described in 2.4.8. If we substitute the formulas for Gaussian beams into Eq. 2.115, as described in Problem 2.21, we find for the three cases the following analytic expressions.

$$\eta = e^{-\frac{\Delta y^2}{2w_1^2}}, \quad L = 10 \log \eta \approx 2.2 \left(\frac{\Delta y}{w_1} \right)^2 \quad (2.119)$$

From Fig. 2.46(a) we see that the coupling tolerance is fairly tight on a lateral offset. With an offset of $0.25 \mu\text{m}$ the coupling loss is already 0.2 dB, and for $0.5 \mu\text{m}$ it is almost 1 dB. The estimates apply for a $2 \mu\text{m}$ wide standard waveguide, as described in

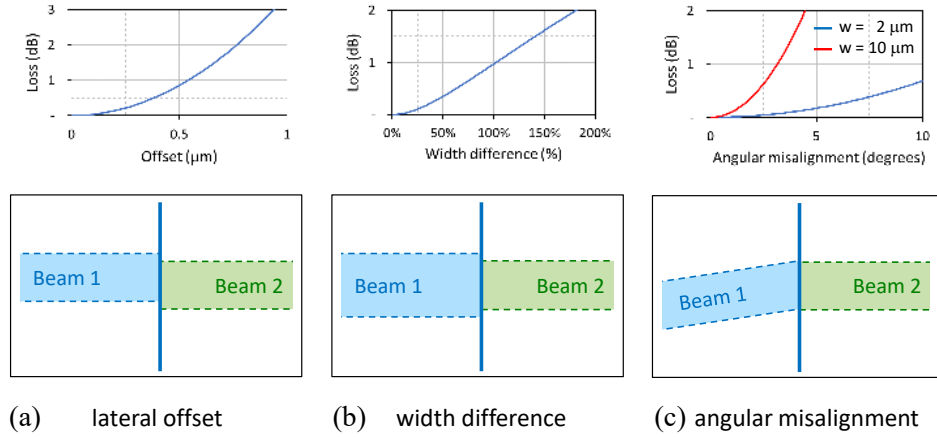


Figure 2.46: Three different cases of beam misalignment or mismatch. The graphs are calculated for lateral misalignment of a $2\ \mu\text{m}$ wide standard waveguide, and for 2 and $10\ \mu\text{m}$ wide waveguide in Fig. (c).

Appendix C, for example for coupling light from a lensed fiber with a matched spot diameter into the chip. Note that the example applies to the lateral direction only. For an estimate including both horizontal and vertical misalignment, we have to calculate the coupling loss in both directions and multiply them as described in Eq. 2.116.

For a width mismatch we find

$$\eta = 2 \frac{w_1 \cdot w_2}{w_1^2 + w_2^2}, \quad L = 10 \log \left(\frac{1}{2} \left\{ \frac{w_1}{w_2} + \frac{w_2}{w_1} \right\} \right) \quad (2.120)$$

From Fig. 2.46(b) we see that the coupling is very tolerant to width mismatch. Even with 25% width mismatch the coupling loss stays below 0.1 dB, and with 100% width mismatch, i.e. $w_1 = 2w_2$, it remains under 1 dB. So for coupling light into a chip from a lensed fibre or a microscope objective for measurement purposes, one may consider to use a larger input beam. This will lead to a significant increase in lateral alignment tolerance, at the cost of a small increase in loss.

Equations 2.120 and 2.120 are directly derived from Eq. 2.118.

For an angular offset ϕ : $U_1(y) = A e^{-\frac{y^2}{w_1^2}} e^{-jk_{1y}y}$, $k_{1y} = n_1 k_0 \sin \phi$ and we find

$$\eta = e^{-\frac{1}{4} k_{1y}^2 w_1^2}, \quad L = 1.1 k_{1y}^2 w_1^2 \approx 43 \frac{w_1^2}{\lambda_1^2} \sin^2 \phi \quad (2.121)$$

in which $\lambda_1 = \lambda_0/n_1$. From Fig. 2.46(c) we see that for the standard waveguide the coupling is very tolerant: even with a misalignment of a few degrees the coupling loss is well below 0.1 dB. For the case of a fibre matched waveguide width ($w = 10\ \mu\text{m}$) this is different: an angular offset of 1° introduces already 0.2 dB coupling loss. Note that if a high suppression of the first-order mode is required, the requirements on angular offset become much stricter.

2.8 Appendices

Appendix 2A Derivation of the Gaussian beam formula

We can calculate the diffraction field $U(x, z)$ corresponding to a Gaussian near field distribution $U(x) = A \exp(-x^2/w_0^2)$ using Fourier Optics.

For the spectral distribution equation 2.14 yields:

$$U_k(k_x) = Aw_0/(2\sqrt{\pi}) \exp[-k_x^2/(4w_0^2)] \quad (2.122)$$

Substituting this into equation 2.15 yields:

$$U(x, z) = Aw_0/(2\sqrt{\pi}) \int \exp(-\frac{1}{4}k_x^2w_0^2 - jk_x x - jk_z z) dk_x \quad (2.123)$$

If $k_z = \sqrt{n^2k_0^2 - k_x^2}$ is approximated as $k_z \approx nk_0[1 - \frac{1}{2}k_x^2/(n^2k_0^2)]$, then the exponent can be rewritten as $-u_c^2 - x^2/w_c^2 - jnk_0z$ with $u_c = [\frac{1}{2}k_x w_0 + jx/w_c]$ and $w_c = \sqrt{w_0^2 - 2jz/(nk_0)}$, in which w_c is a complex z -dependent beam width. Substitution into the above equation and remembering that $\int \exp(-u^2) du = \sqrt{\pi}$ brings us onto the following form:

$$U(x, z) = [Aw_0/w_c(z)] \exp[-x^2/w_c^2(z)] \exp(-jnk_0z) \quad (2.124)$$

The formula shows that along the z -axis the beam behaves like a Gaussian beam with a complex z -dependent beam waist $w_c(z)$. As will be shown below, the x -dependence $\exp(-x^2/w_c^2)$, can be written as

$$\exp[-x^2/w^2 - jnk_0x^2/(2R)] \quad (2.125)$$

in which the beam waist w is a real number. We recognize the latter expression as the phase $\Phi(x, z = R)$ of a spherical wave with radius R and origin $(0, 0)$

$$\begin{aligned} \Phi(x, z = R) &= nk_0|\vec{r} - \vec{r}_0| = nk_0\sqrt{R^2 + x^2} \\ &\approx nk_0R(1 + \frac{1}{2}x^2/R^2) = \Phi_{x=0} + nk_0x^2/(2R). \end{aligned} \quad (2.126)$$

Formula 2.125 can be derived as follows. With $u = z/z_0$ in which $z_0 = \frac{1}{2}nk_0w_0^2$ we find, after multiplying both the numerator and the denominator of the exponent in formula 2.124 with $(1 + ju)$:

$$\begin{aligned} x^2/w_c^2 &= x^2/w^2 + jx^2/[uw_0^2(1 + 1/u^2)] \\ &= x^2(1 + ju)/[w_0^2(1 + u^2)] \\ &= x^2/w^2 + jx^2nk_0/(2R) \end{aligned} \quad (2.127)$$

with $w = w_0\sqrt{1 + u^2}$ and $R = z(1 + 1/u^2)$. This is the exponent of a Gaussian beam with waist w and curvature radius R .

Appendix 2B Derivation of the normalized dispersion relation

In section 2.67 on page 2-34 we introduced the normalized transverse propagation constants (normalized with respect to d)

$$\begin{aligned} u &= k_{1x}d \\ v &= \alpha_0d \\ w &= \alpha_2d \end{aligned} \quad (2.128)$$

with which we can write the dispersion equation 2.61 as

$$u - \arctan \frac{v}{u} - \arctan \frac{w}{u} = m\pi \quad (2.129)$$

Note that the parameter w denotes here the normalized transverse propagation constant in the superstrate, and not the waveguide width. By combining equations 2.78 and 2.80 it is easily verified that

$$u = V\sqrt{1-b} \quad (2.130)$$

and by adding u^2 and v^2 we find $u^2 + v^2 = V^2$, from which it follows that

$$v = V\sqrt{b} \quad (2.131)$$

For w we find using equation 2.79 after some manipulation

$$w = V\sqrt{b+a} \quad (2.132)$$

If we substitute the expressions for u , v and w into equation 2.129 we arrive at the normalized dispersion relation of equation 2.77 on page 2-39.

$$V\sqrt{1-b} - \arctan \sqrt{\frac{b}{1-b}} - \arctan \sqrt{\frac{b+a}{1-b}} = m\pi$$

*normalization for
TM-modes*

Normalization for TM-modes For TM polarization the dispersion relation obtains the following shape in terms of the normalized parameters:

$$V\sqrt{1-b} - \arctan \left(\frac{n_1^2}{n_0^2} \sqrt{\frac{b}{1-b}} \right) - \arctan \left(\frac{n_1^2}{n_2^2} \sqrt{\frac{a+b}{1-b}} \right) = m\pi \quad (2.133)$$

in which the asymmetry parameter a_{TM} is defined as

$$a_{\text{TM}} = \frac{n_1^4}{n_2^4} a_{\text{TE}} \quad (2.134)$$

It is seen that the refractive indices do not disappear from the formula, which is prohibitive for a normalization as described for TE-polarized modes. This is due to the fact that they appear in the boundary conditions (relation 2.36) for TM-polarized waves. If $n_0/n_1 \approx 1$ and $n_2/n_1 \approx 1$, the solutions for TE- and TM-polarization are so close that for many applications the TE solutions may be used for TM polarization as well. If the index contrast is larger some adaptations can be made to apply the normalized approach also for TM-waves. However, they reduce the elegance of the method. Because today almost all calculations are done using numerical mode solvers, and we use the normalized approach here to provide physical insight in the properties of waveguide modes, most of which qualitatively also applies to TM-modes, we will pay no further attention to normalization for TM-modes.

Appendix 2C Group index

For calculating the group index N_{gr} of a waveguide we have to calculate the frequency dispersion

$$dN/d\omega = \frac{dN}{db} \frac{db}{dV} \frac{dV}{d\omega}.$$

Problem 2.22: Group velocity and phase velocity.

Problem: Estimate the group index of the TE₀-mode in the waveguide described in problem 2.14. Use the following numbers for the core and cladding layers: $n_0 = 3.17$, $dn_0/d\lambda = -0.12 \mu\text{m}^{-1}$, $n_1 = 3.36$, $dn_1/d\lambda = -0.23 \mu\text{m}^{-1}$.

Solution: The group index follows from equation 2.95. In problem 2.14 we found for the effective index $N = 3.27$. The dispersion follows from $dN/d\omega = \frac{dN}{db} \frac{db}{dV} \frac{dV}{d\omega}$. The first two factors follow as $\frac{dN}{db} \approx b(n_1 - n_0) = 0.51(3.36 - 3.17) \approx 0.1$ (from equation 2.82) and $\frac{db}{dV} \approx 0.2$ (from figure 2.25). The third one follows from equation 2.135, with $\partial V/\partial\omega = (d/c)\sqrt{n_1^2 - n_0^2} \approx 2 \cdot 10^{-15}$, $\partial V/\partial n_1 = k_0 dn_1/\sqrt{n_1^2 - n_0^2} \approx 6$, and $\partial V/\partial n_0 = -k_0 dn_0/\sqrt{n_1^2 - n_0^2} \approx -5.7$. The material dispersion $dn/d\omega = (dn/d\lambda)(d\lambda/d\omega) = (dn/d\lambda)(-\lambda/\omega)$. With $\lambda/\omega = \lambda^2/(2\pi c) = 1.3 \cdot 10^{-15} \mu\text{ms}$ we find $dn_1/d\omega = -0.23 \cdot (-1.3 \cdot 10^{-15}) = +0.3 \cdot 10^{-15}$ and $dn_2/d\omega = -0.12 \cdot (-1.3 \cdot 10^{-15}) = +0.16 \cdot 10^{-15}$. For the material dispersion contribution we thus find $(\partial V/\partial n_1)(dn_1/d\omega) + (\partial V/\partial n_2)(dn_2/d\omega) = 6 \cdot 0.3 \cdot 10^{-15} - 5.7 \cdot 0.16 \cdot 10^{-15} \approx 0.9 \cdot 10^{-15}$. The material dispersion is thus twice the waveguide dispersion and $dV/d\omega \approx 3 \cdot 10^{-15}$. This gives us $dN/d\omega = 6 \cdot 10^{-17}$ and $N_{\text{gr}} \equiv N + \omega \frac{dN}{d\omega} = 3.27(1 + 1.2 \cdot 10^{15} \cdot 6 \cdot 10^{-17}) = 3.27 \cdot 1.07 = 3.5$.

The first factor can be estimated as $\frac{dN}{db} \approx b(n_1 - n_0)$ (from equation 2.82) and the second one can be estimated from the b-V diagram (figure 2.25). The third factor follows as

$$dV/d\omega = \partial V/\partial\omega + (\partial V/\partial n_1)(dn_1/d\omega) + (\partial V/\partial n_2)(dn_2/d\omega) \quad (2.135)$$

in which the first term describes the waveguide dispersion and the two last terms the material dispersion.

With $V = k_0 d \sqrt{n_1^2 - n_0^2} = (\omega d/c) \sqrt{n_1^2 - n_0^2}$ we find $\partial V/\partial\omega = (d/c) \sqrt{n_1^2 - n_0^2}$, $\partial V/\partial n_1 = k_0 dn_1/\sqrt{n_1^2 - n_0^2}$, and $\partial V/\partial n_0 = -k_0 dn_0/\sqrt{n_1^2 - n_0^2}$. A numerical example of the calculation of the group index for the standard InP waveguide is given in problem 2.22

Appendix 2D Derivation of the Effective Index Method

In section 2.2.2 we have seen that in a two-dimensional waveguide the modal field in the waveguide layer can be described as the interference pattern of two plane waves propagating at mirrored angles θ^+ and θ^- , and non-uniform plane waves with an exponentially decaying field in the cladding layers. The solutions for the propagation constant β are found from the requirement that at each boundary the field and its derivative should be continuous. This approach is also valid for 3-dimensional waveguides, as depicted in figure 2.33. Here we can describe the mode as the interference pattern of 4 plane waves: two at an angle θ_y^+ in the y-z-plane and angles θ_x^+ and θ_x^- in the x-z-plane, and two at an angle θ_y^- in the y-z-plane and angles θ_x^+ and θ_x^- in the x-z-plane. And for guided modes the fields should be exponentially decaying everywhere outside the waveguide core. The possible values of β can then be found by applying the boundary conditions at all interfaces. However, for a 3d-waveguide this is fairly complex because of the large number of boundaries.

We will explain the EIM for the simplified structure of figure 2.33, although it is also valid for similar structures for which an effective index can be calculated in all three

lateral regions. We are looking for solutions of the form

$$U(x, y, z) = U(x, y) e^{-j\beta z} \quad (2.136)$$

which satisfy the boundary conditions at all interfaces. We assume that the mode profile $U(x, y)$ can be written as $X(x) \cdot Y(y)$, which means that the vertical mode profile $X(x)$ is the same in the (lateral) core and cladding regions. This restricts the validity of the EIM to waveguide structures where the vertical field profile in the cladding regions differs not too much from that in the core region.

If we substitute the field of equation 2.136 with $U(x, y) = X(x) \cdot Y(y)$ in the 3-dimensional Helmholtz equation $\nabla^2 U + n^2 k_0^2 U = 0$ we find:

$$Y \frac{\partial^2}{\partial x^2} X + X \frac{\partial^2}{\partial y^2} Y - (\beta^2 - n^2 k_0^2) XY = 0 \quad (2.137)$$

If we assume that $X(x)$ is a 2-dimensional solution of the Helmholtz equation $\frac{\partial^2}{\partial x^2} X + (n^2 k_0^2 - \beta_x^2) X$, with solutions $X(x) = ae^{jk_x x} + be^{-jk_x x}$, in which $k_x = \sqrt{n^2 k_0^2 - \beta_x^2}$, and we substitute this in equation 2.137, we find

$$\frac{\partial^2}{\partial y^2} Y + (\beta_x^2 - \beta^2) Y = \frac{\partial^2}{\partial y^2} Y + (N_{1X}^2 k_0^2 - \beta^2) Y = 0 \quad (2.138)$$

with solutions $Y(y) = ae^{jk_y y} + be^{-jk_y y}$ in which $k_y = \sqrt{N_{1X}^2 k_0^2 - \beta^2}$ in the central region, and $Y(y) = ce^{-|\alpha_y y|}$, in which $\alpha_y = \sqrt{\beta^2 - N_{2X}^2 k_0^2}$ in the cladding regions, and N_{1X} and N_{2X} are the effective indices in the core and the cladding regions, respectively. From equation 2.138 it is seen that the lateral mode profile $Y(y)$ is found as described in section 2.4 for a two-dimensional waveguide structure by substituting the effective indices N_0, N_1, N_2 and the waveguide width w for the real indices n_0, n_1, n_2 and the film thickness d . By squaring and adding k_x, k_y and β we find: $k_x^2 + k_y^2 + \beta^2 = n^2 k_0^2$ as expected: k_x, k_y and β are the x, y and z-components of the plane wave vector in the core medium.

Appendix 2E The Transfer Matrix Method (TMM)

The modes of an n -layer waveguide can be found in the same way as for a three-layer waveguide by applying the continuity conditions at each interface. This gives $2(n-1)$ equations (one for the field and one for its derivative at each interface) which are easily reduced to a set of two equations using a matrix formalism. The method works as follows.

We start with a first guess for the propagation constant β . Next, the complex transverse propagation constants in the different layers are computed from:

$$\gamma_{ix} = jk_{ix} = \sqrt{\beta^2 - n_i^2 k_0^2} \quad i = 0, 1, 2, \dots, n \quad (2.139)$$

We have introduced the complex transverse propagation constant γ_x instead of k_x because in the different layers the transverse propagation constant can assume both complex, real and imaginary values, whereas k_x is normally used for real-valued wave numbers. In each layer we search for solutions of the form:

$$U_i(x) = a_i e^{-\gamma_i x} + b_i e^{+\gamma_i x} \quad (2.140)$$

*transverse
propagation
constant*

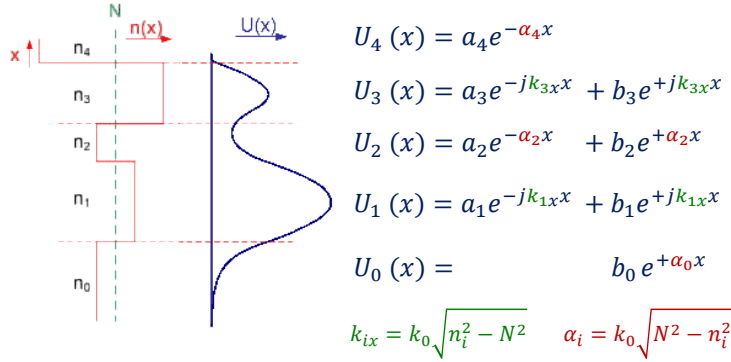


Figure 2.47: Transfer-Matrix method for a real-valued refractive index profile. The complex transverse propagation constant γ has been replaced by the real propagation constants k_x and α to make visible that the field has a cosine shape in layers where $N > n_i$ and a hyperbolic cosine shape where $N < n_i$.

The numbering ranges from $i = 0$ for the substrate to $i = n$ for the superstrate. For a guided mode, γ_{0x} will be real and $a_0 = 0$ (otherwise the mode will explode at infinity). For a radiating mode, γ_{0x} will be imaginary, but again $a_0 = 0$, because a_0 is the coefficient of a wave propagating towards the waveguide; we are only interested in those modes for which power is leaking away from the waveguide. For a given value of β , the mode profile in the substrate is now known: $U_0(x) = b_0 \exp(+\gamma_{0x}x)$, and so are the field and its derivative at the interface with the first layer. Through the continuity conditions the field and its derivative at the other side of the interface are fixed and, as a consequence, the whole field in the first layer and, by repeating the procedure, the fields in all the other layers are also fixed. For the same reasons for which $a_0 = 0$ in the substrate, the coefficient b_n of the incoming wave in the superstrate has to be zero. This gives us a dispersion equation for the multi layer waveguide structure: $b_n = 0$. If for the chosen starting value of β the value of $b_n \neq 0$, we have to try another value of β until a root of the dispersion relation is found. This can be done with a complex numerical root finder. For each trial of β we have to compute the value of b_n . An easy way of doing this is by multiplying the coefficient vector (a_0, b_0) of the substrate field with the (complex) 2×2 transfer matrices of the $(n - 2)$ films and the $(n - 1)$ interfaces in order to obtain the vector (a_n, b_n) of the superstrate field. This explains the name of the method: Transfer-Matrix Method. A mathematical description is given hereafter.

We start with TE polarization. To shorten the notation we will write the field and its derivative in layer i as the field vector $\Psi_i(x) = [U_i(x), (\partial/\partial x)U_i(x)]^T$. In layer i , x goes from 0 to d_i . For convenience, we will call the field $\Psi_i(0)$, at the bottom of layer i , Ψ_i . Now we try to find the transfer matrix T_i , so that $\Psi_{i+1} = T_i \Psi_i$. For layer i , it follows from equation 2.140 that:

$$\Psi_i(x) = \begin{pmatrix} e^{-\gamma_i x} & e^{\gamma_i x} \\ -\gamma_{ix} e^{-\gamma_i x} & \gamma_{ix} e^{\gamma_i x} \end{pmatrix} \begin{pmatrix} a_i \\ b_i \end{pmatrix} = C_i(x) A_i \quad (2.141)$$

in which A_i is the coefficient vector $(a_i, b_i)^T$. From this equation it follows immediately that:

$$A_i = C_i^{-1} \Psi_i; \quad C_i^{-1} = [C_i(0)]^{-1} = \frac{1}{2\gamma_i} \begin{pmatrix} \gamma_i & -1 \\ \gamma_i & 1 \end{pmatrix} \quad (2.142)$$

Because of the continuity conditions $\Psi_{i+1} = \Psi_i(d_i)$, and thus we find for Ψ_{i+1} :

$$\Psi_{i+1} = C_i(d_i) C_i^{-1} \Psi_i = T_i \Psi_i \quad (2.143)$$

The transfer matrix is then found from:

$$\begin{aligned} T_i &= C_i(d_i)C_i^{-1} \\ &= \cosh(\gamma_i d_i) \begin{pmatrix} 1 & -\frac{1}{\gamma_i} \tanh(\gamma_i d_i) \\ -\gamma_i \tanh(\gamma_i d_i) & 1 \end{pmatrix} \end{aligned} \quad (2.144)$$

The matrix is defined for both real and imaginary γ_i , with $\cosh(jy) = \cos(y)$, and $\tanh(jy) = j \tan(y)$. Knowing this, we can write A_n in terms of Ψ_1 as:

$$A_n = C_n^{-1} \Psi_n = C_n^{-1} \left(\prod_{i=1}^n T_i \right) \Psi_1$$

If we describe the field in the substrate as in equation 2.140, with $-\infty < x < 0$, $a_0 = 0$ and $b_0 = 1$, we find:

$$\Psi_1 = \Psi_0(0) = \begin{pmatrix} 1 \\ \gamma_0 \end{pmatrix}$$

from which b_n follows as the second element of A_n and the dispersion relation follows by requiring

$$b_n = 0.$$

For TM polarization we can write for $\Psi_i(x)$: $\Psi_i(x) = [U_i(x), (1/n_i^2)(\partial/\partial x)U_i(x)]^T$. The rest of the calculation is analogous to the one for TE polarization.

When we have found a value for β for which $b_n = 0$, the profile of the corresponding mode follows from the calculated coefficient vectors $A_i = (a_i, b_i)^T$ through equation 2.140.

Since the above description is valid for complex as well as real refractive indices, the method may be applied for analyzing modes in absorbing waveguide structures or waveguides with gain, as they occur in lasers. A difference with the three layer waveguide is that the mode number does not appear in the dispersion relation. Which mode will be found by the root finder is dependent on the starting value and has to be established by the user in retrospect. It is noted that for strongly absorbing waveguide structures or structures with a high gain the identification of the modes becomes very complicated because the zeroes in the mode profile disappear and different modes may come very close to each other. Further the effective index of a guided mode can become lower than the substrate index, as shown in curve 5 of figure 2.40. This is not possible in loss-free waveguides.

graded index profile With a TMM we can also analyze graded index profiles by approximating them by a stepwise uniform profile with a large number of steps (staircase approximation). If we have some idea about the maximal value of k_x in the mode spectrum, a reasonable estimate for the number of subdivisions is 10-20 layers per transverse wavelength $\lambda_x = 2\pi/k_x$. A worst-case guess can be found from $k_{x,\max} = k_0 \sqrt{n_{\max}^2 - n_0^2}$. The number of layers is sufficient if the value of β does not change significantly on further reduction of the step size. The speed of the TMM is such that profiles containing several hundreds of steps can be analyzed in a short time, which allows for a very good approximation of graded profiles.