Optical Waveguides and Fibers

Christian Koos,[1] and Wolfgang Freude
Institute of Photonics and Quantum Electronics (IPQ)
http://www.ipq.kit.edu
Karlsruhe Institute of Technology (KIT)
Winter Term 2011/2012

January 26, 2016

mailto:christian.koos@kit.edu

Contents

Chapter 1

Introduction

1.1 The nature of light

According to Maxwell,[1] light propagates as a wave with a wavelength λ. In vacuum, the speed of light is $c = 2.997\,924\,58 \times 10^8\,\mathrm{m/s}$. However, Planck[2] found that the energy of light radiated from a hot black body is emitted in quanta, the energy of which is in proportion to the observed frequency $f = c/\lambda$, so that each quantum or "photon" has an energy $W = hf$; Planck's constant is $h = 6.626\,075\,5 \times 10^{-34}\,\mathrm{Js}$. Further, it was shown by de Broglie[3] that each particle having momentum p may be associated with a wavelength $\lambda = h/p$. Obviously, the nature of light is ambiguous. Einstein[4] formulated[5]: "Light is like the French philosopher Voltaire[6]. Voltaire was born catholic, converted as a young man to Protestantism, and returned to Catholicism shortly before his death. Light is born as a particle, lives as a wave, and dies as a photon when being absorbed."

1.2 Communication with light

An optical communication system uses lightwaves in a vacuum wavelength range $0.6\,\mu\mathrm{m} \ldots 1.2\,\mu\mathrm{m} \leq \lambda \leq 1.6\,\mu\mathrm{m}$ corresponding to carrier frequencies $f = c/\lambda$ of $500\,\mathrm{THz} \ldots 250\,\mathrm{THz} \geq f \geq 190\,\mathrm{THz}$. A communication *system* is referred to as a point-to-point transmission link. When many transmission links are interconnected with multiplexing or switching functions, they are called a communication *network*.

The basic principle of an optical transmission link is shown in Fig. 1.1. At the transmitter, an electrical data signal is converted to the optical domain, where data can be encoded in the amplitude and/or the phase of the optical carrier wave. The optical signal is transmitted through optical fibers. For transmission over long distances, optical amplifiers are used to compensate for propagation loss. The transmission capacity of an optical fiber can be significantly increased by wavelength-division multiplexing (WDM) schemes, which exploit a multitude of optical carriers at different wavelengths, all of which can independently be used for data transmission.

[1] James Clerk Maxwell, mathematician and physicist, ⋆Edinburgh 13.6.1831, †Cambridge 5.11.1879. Professor in Cambridge, UK

[2] Max Planck, physicist, ⋆Kiel 23.4.1858, †Göttingen 4.10.1947. Professor in Kiel and Berlin. Nobel prize in physics 1918

[3] Louis Victor, 7. Duke of Broglie (since 1960), named Louis de Broglie, physicist, ⋆Dieppe 15.8.1892, †Louveciennes (Département Yvelines) 19.3.1987. Nobel prize in physics 1929 (together with O. W. Richardson)

[4] Albert Einstein, physicist, ⋆Ulm 14.3.1879, †Princeton (NJ) 18.4.1955. "Technical expert 3rd class" at the patent office in Bern (1902–09). Professor at the University of Zurich and Prague (1911/12) and at the Swiss Federal Institute of Technology (ETH) in Zurich. Emigration to the USA in 1933. Professor at the Institute for Advanced Study in Princeton (NJ). American citizen since 1940. Formulated in 1905 (1914–16) the special (general) theory of relativity. Nobel prize in physics 1921

[5] Jahns, J.: Photonik. Grundlagen, Komponenten und Systeme. München: Oldenbourg-Verlag 2001. Page 9

[6] Pseudonym or pen-name of François Marie Arouet, philosopher and writer, ⋆Paris 21.11.1694, †Paris 30.5.1778

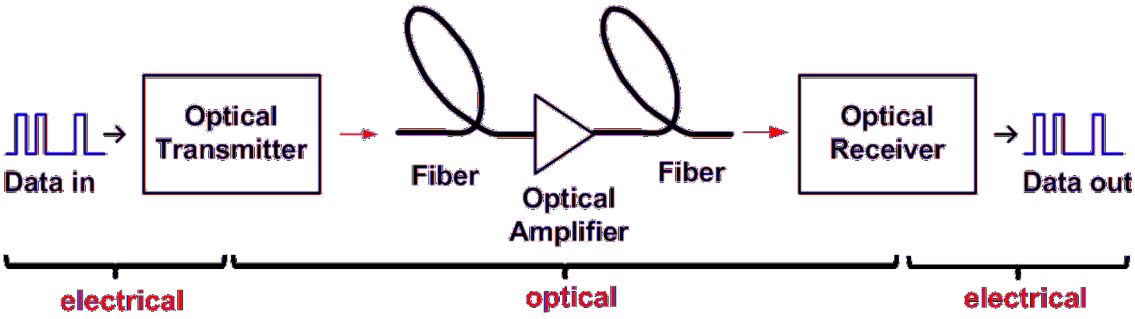

Figure 1.1: Optical point-to-point transmission link: An electrical signal is transformed into an optical data signal by an optical transceiver. The optical signal is transmitted through an optical fiber and amplified. At the receiver side, the optical signal is converted back to the electrical domain. Data can be encoded in the amplitude and the phase of the optical signal.

Advantages of optical communications

Obviously, optical communication systems can replace conventional electrical systems only if there is some advantage to be gained, which justifies the additional expenses of a twofold conversion current-light and light-current. Some important advantages of optical signal transport are:

- **Large transmission capacity** because of the large fiber bandwidth in the order of $(250 - 190)\,\mathrm{THz} = 60\,\mathrm{THz}$
- **Long transmission reach** due to low fiber loss, about $2.2, 0.35, 0.15\,\mathrm{dB/km}$ at $\lambda = 0.85, 1.3, 1.55\,\mu\mathrm{m}$, i. e., down to 3 dB loss for a fiber length of $L = 20\,\mathrm{km}$ corresponding to a power attenuation by a factor of only 2
- **Immunity to electromagnetic interference** because of the high carrier frequency, and because of the strong confinement of the light inside the fiber

Technological milestones in optical communications

The first commercial optical fiber communication system were deployed in the 1970s, with data rates of a few Mbit/s. Since then, optical communications has seen tremendous progress, and current record data rates that can be transmitted through single-mode fibers are of the order of tens of Tbit/s. Three milestones of lightwave technology are especially noteworthy:

- **Low-loss fibers**: Following an earlier suggestion,[7,8] the first low-loss fibers were produced [9] in 1970 reducing the loss from $1\,000\,\mathrm{dB/km}$ to below $20\,\mathrm{dB/km}$. Further progress[10] resulted by 1979 in a loss of only $0.2\,\mathrm{dB/km}$ near $\lambda = 1.55\,\mu\mathrm{m}$. The ultimate low loss[11] of $0.154\,\mathrm{dB/km}$ for fibers with a silica (SiO_2) core and a fluorine-doped cladding is limited only by the amorphous structure of silica (Rayleigh scattering) and was reached in 1986.
- **Semiconductor lasers:** Although semiconductor lasers were first made[12] in 1962, their use became practical only after 1970 when GaAs lasers operating continuously at room

[7] Kao, K. C.; Hockham, G. A.: Proc. IEE 113 (1966) 1151

[8] Werts, A.: Onde Electr. 45 (1966) 967

[9] Kapron, F. P.; Keck, D. .B; Maurer, A. D.: Appl. Phys. Lett. 17 (1970) 423

[10] Miya, T.; Terunuma, Y.; Hosaka, T.; Miyashita, T.: Ultimate low-loss single-mode fibre at 1.55 μm. Electron. Lett. 15 (1979) 106–108

[11] Kanamori, H.; Yokota, H.; Tanaka, G.; Watanabe, M.; Ishiguro, Y.; Yoshida, I.; Kakii, T.; Itoh, S.; Asano, Y.; Tanaka, S.: Transmission characteristics and reliability of pure-silica-core single-mode fibers. IEEE J. Lightwave Technol. LT-4 (1986) 1144–1149

[12] Nasledov, D. N.; Rogachev, A. A.; Ryvkin, S. M.; Tsarenkov, B. V.: Fiz. Tverd. Tela. 4 (1962) 1062 (Soviet Phys. Solid State 4 (1962) 782

temperature were available.[13]

- **Optical fiber amplifier**: Finally, it was only after the invention and perfection of the erbium-doped fiber amplifier (EDFA) [14] in 1986 that optical communication became as powerful as it is today.[15,16]

Short- and long-reach optical data transmission

Large-scale deployment of optical transmission systems started in the 1990s when broadband erbium-doped fiber amplifiers became available. Optical data transmission was first deployed in long-haul links, where increasing data rates and large transmission distances justified the additional technical effort. In the late 1990s several WDM systems were installed across the Pacific and the Atlantic oceans in response to the data traffic growth induced by the Internet, see Fig. 1.2. Today, 40 Gbit/s or even 100 Gbit/s of data are transmitted per wavelength channel, and modern WDM systems can have more than 100 wavelength channels. A record data rate of 101.7 Tbit/s was recently transmitted through a single-mode fiber in a laboratory experiment [28].

However, optical data transmission is no longer limited to long-haul transmission. Internet data traffic is currently increasing with annual rates of more than 30 %. Short- and medium-reach interconnects in data centers currently represent the capacity bottleneck in data networks. As a consequence, electrical interconnects are currently replaced by so-called active optical cables (AOC). AOC plug directly into standard copper cable ports; optical transmitters and receivers contained in the cable ends convert the signal back and forth for optical transport over fiber, see Fig. 1.3. Similarly, Optical systems for rack-to-rack data transmission are currently being installed in high-performance computers (HPC) and warehouse-scale computers (WSC), where the interconnect bandwidth limitations are currently the most severe obstacle to further increase performance while decreasing the power consumption.

In the future, it is anticipated that the advantages of optical techniques will pay out for transmissions over even shorter distances: Board-to-board interconnects will be used to transmit data between printed circuit boards that reside within the same rack, Fig. 1.4. These systems require compact optical transmitters and receivers, that are integrated with electronic circuitry on the same chip. In a long-term perspective, optical transmission might even be used for on-chip communication, Fig. 1.5. In this context, silicon nanophotonics and plasmonics are active areas of current research.

---

[13] Alferov, Z.: IEEE Sel. Topics Quantum Electron. 6 (2000) 832

[14] Poole, S. B.; Payne, D. N.; Mears, R. J.; Fermann, M. E.; Laming, R. E.: J. Lightwave Technol. 4 (1986) 870

[15] *Why need amplifiers be distributed along a transmission distance?* Due to attenuation in the transmitting fibre the optical signal decays exponentially with the transmission span. Practical spans without amplification are about 70 km. Why are the spans so short (see Footnote 16 on Page 3)?

A transatlantic transmission from New York to London experiences an attenuation of about 1 400 dB (7 000 km @ 0.2 dB / km). Thus, for receiving one photon in London we have to inject 10^{140} photons into the optical fibre end in New York. If all the mass of our sun ($m_{\text{sun}} = 3 \times 10^{33}$ g) having an energy equivalent of $W_{\text{sun}} = mc^2 = 1.8 \times 10^{47}$ J could be converted into photons with a photon energy $hf = 6 \times 10^{-34}$ Js $\times$ 200 THz $= 1.2 \times 10^{-19}$ J, we had generated 1.5×10^{66} photons at a wavelength of 1.55 μm ($f \approx 200$ THz), and could bridge a span with 660 dB loss, corresponding to a transmission distance of 3 300 km only. For a direct transmission New York – London we thus had to evaporate $10^{140}/10^{66} = 10^{74}$ suns.

This is quite a bit. The (observable) universe is estimated to have an extension of 1.4×10^{10} light years. Its mean density is supposed to be 3×10^{-30} g / cm^3 (`http://curious.astro.cornell.edu/question.php?number=342`). Therefore, the universe's mass (comprising not only suns) is $m_{\text{univ}} = 7 \times 10^{54}$ g, and its energy equivalent is $W_{\text{univ}} = m_{\text{univ}}c^2 = 6 \times 10^{68}$ J corresponding to 4.7×10^{87} photons at a wavelength of 1.55 μm. If we are able to receive one photon then the maximum span will be 877 dB /(0.2 dB / km) = 4 385 km. However, for bridging the distance New York – London in one go we had to burn $10^{140}/10^{87} = 10^{53}$ universes!

[16] Calculations stimulated by an oral presentation of N. J. Doran (S. K. Turitsyn, M. P. Fedoruk, N. J. Doran and W. Forysiak: Optical soliton transmission in fiber lines with short-scale dispersion management. 25th European Conference on Optical Communication (ECOC'99), Nice, France, September 26–30, 1999). — Universe's mass calculations and web address contributed by Dipl.-Phys. Jan Brückner, Karlsruhe, Germany, June 23, 2005

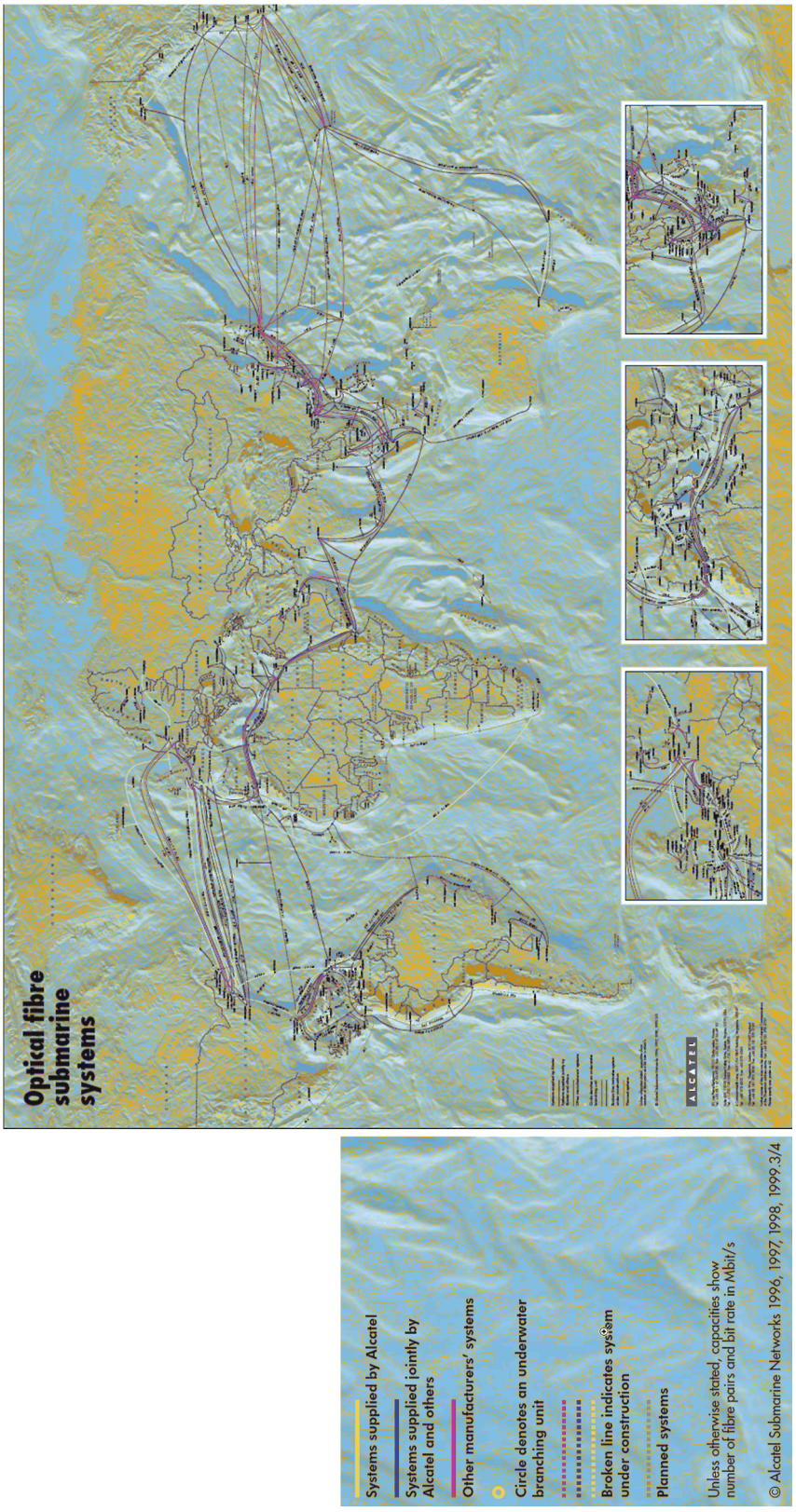

Figure 1.2: Global optical communication Networks: In the 1990s, several WDM systems were installed across the Pacific and the Atlantic oceans in response to the growth of Internet traffic, (Source: Alcatel-Lucent).

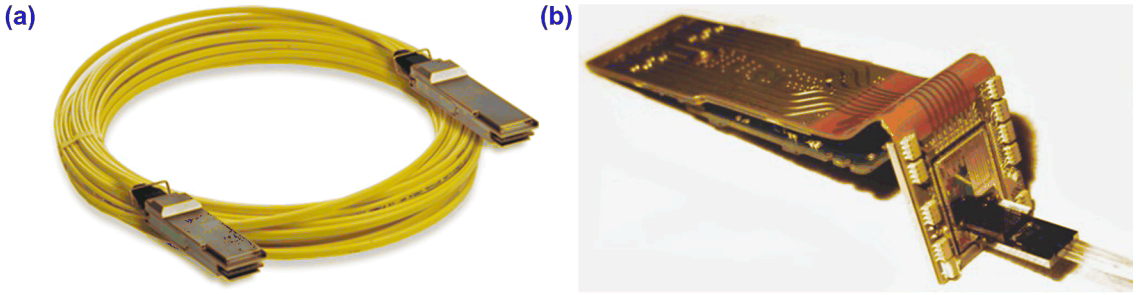

Figure 1.3: (a) Optical interconnects are playing an increasingly important role for communication networks in large-scale data centers. Communication is performed by so-called active optical cables (AOC). AOC have electric connectors, while data transmission in the cable is performed optically. (b) Conversion between optical and electrical signals is performed by optical transmitter and receiver chips that are integrated into the connector housing of the cable (Source: Luxtera).

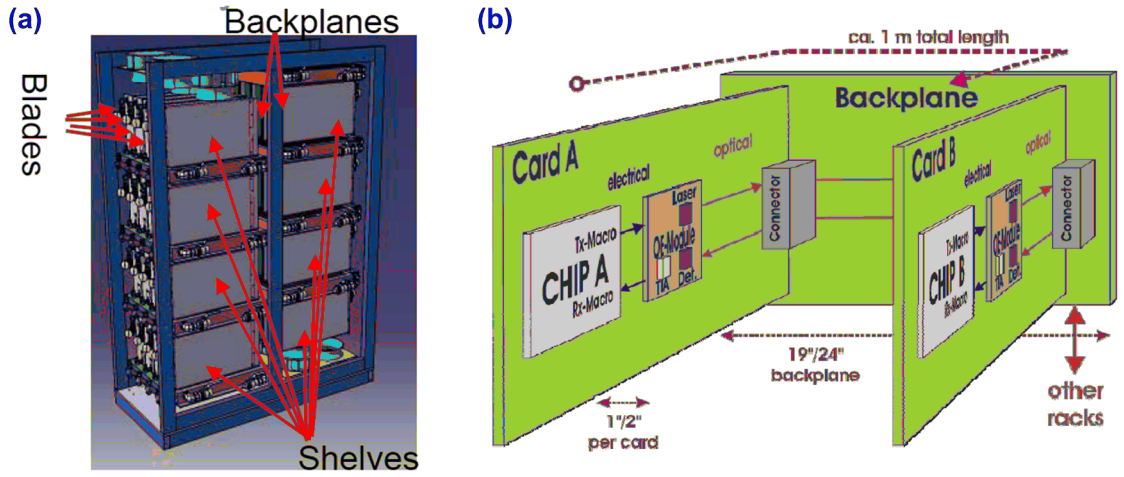

Figure 1.4: Optical board-to-board interconnects: (a) In data centers and high-performance computers, blade servers are organized in racks and connected to a common backplane for data exchange. To increase communication bandwidth, short-reach optical board-to-board interconnects are subject to current research. (b) Board-board optical interconnects require compact chip-scale transmitter and receiver modules as well as novel concepts for optical connections between printed circuit board (PCB) cards and optical backplanes (Source: IBM).

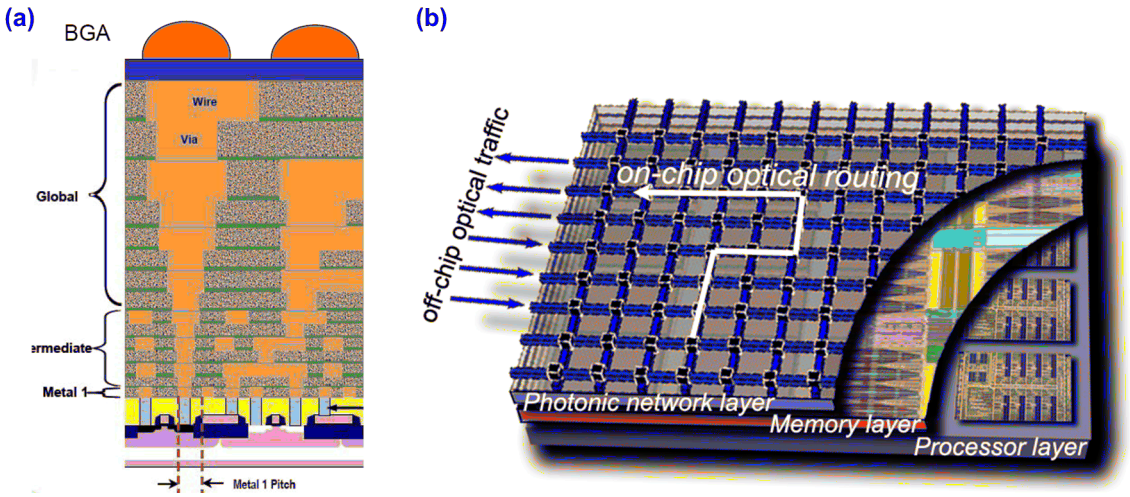

Figure 1.5: Optical on-chip communications. (a) Communication within microprocessor chips is currently handled by a stack of usually more than ten layers of metal interconnects. These interconnects are limited in bandwidth and prone to crosstalk. (b) Future concepts of optical on-chip data transmission are subject to ongoing research. Scenarios comprise 3D integration of dedicated processor, memory, and network layers. (Source: IBM).

Waveguides and fibers

Optical waveguides and fibers are essential building blocks of all aforementioned optical communication systems: Integrated optical waveguides are the basis of nearly any passive or active devices that are used in optical data communications. Fused silica fibers are the transmission medium in long-haul systems, spanning distances of thousands of kilometres. Polymer fibers and waveguides are used for short-reach interconnects, in-house communications and board-to-board links. Moreover, optical waveguides play an important role for other applications: Optical sensing, metrology and biophotonics are unthinkable without dedicated waveguide concepts.

In the following chapters, we will introduce the fundamental principles of optical waveguiding together with the corresponding mathematical models that are required for quantitative device analysis. An overview of current waveguide and fiber technologies will be given, and basic waveguide-based components will be discussed. The interested reader is referred to related textbooks in the field of photonics and optical waveguiding:

- B. E. A. Saleh, M. C. Teich, 'Fundamentals of Photonics', Second Edition; John Wiley & Sons, 2007, ISBN 978-0-471-35832-9
- K. Okamoto, 'Fundamentals of Optical Waveguides', Second Edition, Elsevier Academic Press 2006, ISBN 978-0-12-525096-2
- K. Iizuka, 'Elements of Photonics', Volume II; John Wiley & Sons, 2002, ISBN 0471408158

Chapter 2

Fundamentals of wave propagation in optics

In this chapter, we will review the fundamentals of optical wave propagation. Starting from Maxwell's equations, we will derive the wave equations for inhomogeneous media and simplify them for the case of weak inhomogeneities. We will further analyze the interaction of optical waves with media, and derive simple models to describe the frequency dependence of the refractive index for both dielectric and metallic materials. This leads to material dispersion and dispersive broadening of optical signals. We will quantitatively analyze these effects.

In the following sections, we will make use of some basic relations from vector differential calculus. For convenience, a short summary of the most relevant relations is given in Section A.2.2 in the Appendix.

2.1 Maxwell's equations in optical media

2.1.1 Maxwell's equations and constitutive relations

In the absence of any free carriers and currents, Maxwell's equations take the following form [14]:

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = 0 \tag{2.1}$$

$$\nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t} \tag{2.2}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r},t) = 0 \tag{2.3}$$

$$\nabla \times \mathbf{H}(\mathbf{r},t) = \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} \tag{2.4}$$

The vector $\mathbf{r} = (x, y, z)^{\mathrm{T}}$ defines a point in three-dimensional space. In optics, media are generally assumed to be nonmagnetic. The magnetic flux density $\mathbf{B}$ is then related to the magnetic field $\mathbf{H}$ by

$$\mathbf{B}(\mathbf{r},t) = \mu_0 \mathbf{H}(\mathbf{r},t), \tag{2.5}$$

where $\mu_0 = 1.25664 \times 10^{-6}\ \mathrm{V\,s\,/(A\,m)}$ is the magnetic permeability of vacuum. The relation between the electric field $\mathbf{E}$ and the electric displacement $\mathbf{D}$ can be expressed as

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t), \tag{2.6}$$

where $\epsilon_0 = 8.85419 \times 10^{-12}\ \mathrm{A\,s\,/(V\,m)}$ is the electric permittivity of vacuum, and where $\mathbf{P}$ denotes the electric polarization. These equations are also referred to as constitutive relations.

2.1.2 Monochromatic waves and frequency-domain analysis

In many cases, monochromatic electromagnetic fields are considered

$$\Psi(t) = \mathrm{Re}\left\{\underline{\Psi}(\omega)\exp\left(\mathrm{j}\,\omega t\right)\right\}, \tag{2.7}$$

where $\underline{\Psi}$ represents a complex wave amplitude, and $\omega = 2\pi f$ is the angular frequency. Maxwell's equations can then be written as

$$\nabla \cdot \underline{\mathbf{D}}(\mathbf{r},\omega) = 0 \tag{2.8}$$
$$\nabla \times \underline{\mathbf{E}}(\mathbf{r},\omega) = -\,\mathrm{j}\,\omega\underline{\mathbf{B}}(\mathbf{r},\omega) \tag{2.9}$$
$$\nabla \cdot \underline{\mathbf{B}}(\mathbf{r},\omega) = 0 \tag{2.10}$$
$$\nabla \times \underline{\mathbf{H}}(\mathbf{r},\omega) = \mathrm{j}\,\omega\underline{\mathbf{D}}(\mathbf{r},\omega), \tag{2.11}$$

and the constitutive relations are given by

$$\underline{\mathbf{B}}(\mathbf{r},\omega) = \mu_0\underline{\mathbf{H}}(\mathbf{r},\omega), \tag{2.12}$$
$$\underline{\mathbf{D}}(\mathbf{r},\omega) = \epsilon_0\underline{\mathbf{E}}(\mathbf{r},\omega) + \underline{\mathbf{P}}(\mathbf{r},\omega). \tag{2.13}$$

The same set of equations is obtained, when using Fourier transforms $\widetilde{\Psi}(\omega)$ of time-domain quantities $\Psi(t)$ rather than complex amplitudes $\underline{\Psi}$ of monochromatic waves. The Fourier transformation is given by

$$\widetilde{\Psi}(\omega) = \int_{-\infty}^{+\infty} \Psi(t)\exp\left(-\,\mathrm{j}\,\omega t\right)\mathrm{d}t \tag{2.14}$$

$$\Psi(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widetilde{\Psi}(\omega)\exp\left(\mathrm{j}\,\omega t\right)\mathrm{d}\omega \tag{2.15}$$

For a monochromatic electromagnetic field $\Psi(t) = \mathrm{Re}\left\{\underline{\Psi}(\omega_0)\exp\left(\mathrm{j}\,\omega_0 t\right)\right\}$, oscillating at an angular frequency ω_0, the Fourier transform and the complex time-domain amplitudes are related by

$$\widetilde{\Psi}(\omega) = \frac{1}{2}\left(\underline{\Psi}(\omega_0)\delta\left(\omega-\omega_0\right) + \underline{\Psi}^*(\omega_0)\delta\left(\omega+\omega_0\right)\right), \tag{2.16}$$

where δ represents the Dirac delta function. In linear optics, Fourier transforms and complex wave amplitudes are interchangeable, whereas, in nonlinear optics, the additional factor of $\frac{1}{2}$ is important and may lead to confusion if complex wave amplitudes and Fourier transforms are not clearly distinguished. For the remainder of this lecture, we will largely use the notation for complex wave amplitudes $\underline{\Psi}(\omega)$, keeping in mind that the corresponding quantity can also be understood as the Fourier transform $\widetilde{\Psi}(\omega)$ of a non-harmonic time-domain quantity $\Psi(t)$.

2.1.3 Eletric susceptibility and complex refractive index

The electric susceptibility represents the response of the electric polarization $\mathbf{P}(\mathbf{r},t)$ to the electric field $\mathbf{E}(\mathbf{r},t)$. In real media, the polarization follows the electric field with a time delay that can be described by a convolution with a causal influence function $\chi(\mathbf{r},t)$, where $\chi(\mathbf{r},t<0)=0$. In the frequency-domain, this corresponds to a multiplication with the corresponding complex transfer function $\underline{\chi}(\mathbf{r},\omega)$,

$$\mathbf{P}(\mathbf{r},t) = \epsilon_0\int_{-\infty}^{\infty}\chi(\mathbf{r},\tau)\mathbf{E}(\mathbf{r},t-\tau)\,\mathrm{d}\tau \tag{2.17}$$
$$\underline{\mathbf{P}}(\mathbf{r},\omega) = \epsilon_0\underline{\chi}(\mathbf{r},\omega)\underline{\mathbf{E}}(\mathbf{r},\omega) \tag{2.18}$$

The constitutive relation (2.13) in the frequency domain can then be reformulated,

$$\begin{aligned}\underline{\mathbf{D}}(\mathbf{r},\omega) &= \epsilon_0\underline{\mathbf{E}}(\mathbf{r},\omega) + \underline{\mathbf{P}}\,(\mathbf{r},\omega)\\ &= \epsilon_0\left(1+\underline{\chi}(\mathbf{r},\omega)\right)\underline{\mathbf{E}}(\mathbf{r},\omega)\\ &= \epsilon_0\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)\underline{\mathbf{E}}(\mathbf{r},\omega)\\ &= \epsilon_0\underline{n}^2(\mathbf{r},\omega)\underline{\mathbf{E}}(\mathbf{r},\omega).\end{aligned} \tag{2.19}$$

The complex relative dielectric constant $\underline{\epsilon}_{\mathrm{r}}$ and the complex refractive index $\underline{n}$ are related by

$$\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega) = 1+\underline{\chi}(\mathbf{r},\omega) = \underline{n}^2(\mathbf{r},\omega). \tag{2.20}$$

From this relation, we can derive various relationships of the real and imaginary parts of $\underline{\epsilon}_{\mathrm{r}}$ and $\underline{n}$,

$$\begin{array}{lll}\underline{n} = n - \mathrm{j}\,n_i, & & \underline{\epsilon}_r = \epsilon_r - \mathrm{j}\,\epsilon_{ri},\\ \epsilon_r = n^2 - n_i^2, & & \epsilon_{ri} = 2nn_i,\\ n^2 = \frac{1}{2}\epsilon_r\left(1+\sqrt{1+\epsilon_{ri}^2/\epsilon_r^2}\right), & & n_i = \epsilon_{ri}/(2n),\\ n \approx \sqrt{\epsilon_r} & (\text{for } |\epsilon_{ri}| \ll \epsilon_r) & n_i \approx \epsilon_{ri}/(2\sqrt{\epsilon_r}),\\ n \approx \sqrt{|\epsilon_{ri}|/2} & (\text{for } |\epsilon_{ri}| \gg \epsilon_r) & n_i \approx \mathrm{sgn}(\epsilon_{ri})\sqrt{|\epsilon_{ri}|/2}\,.\end{array} \tag{2.21}$$

Note the somewhat surprising negative signs for the imaginary parts in Eq. (2.21) result from the convention to assign positive values of n_i to lossy media, whereas negative values of n_i correspond to media with gain, see Section 2.2 for a more detailed discussion.

In general optics, the phenomenon of a frequency-dependent refractive index and hence a frequency-dependent phase velocity is also referred to as dispersion. This more general notion of dispersion differs slightly from the use of the term in optical communications, where dispersion is used as a synonym for "group velocity dispersion" or "group delay dispersion", thereby embracing all effects that lead to a frequency-dependent group velocity, see Section (2.4).

2.2 Wave equation and plane waves

2.2.1 Wave equations in the general form

Reshaping of Maxwell's equations 2.8 – 2.11 leads to the vector wave equations for the electric and the magnetic fields,

$$\nabla^2\underline{\mathbf{E}}(\mathbf{r},\omega) + \nabla\left(\frac{\nabla\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)}{\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)}\cdot\underline{\mathbf{E}}(\mathbf{r},\omega)\right) + k_0^2\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)\underline{\mathbf{E}}(\mathbf{r},\omega) = 0 \tag{2.22}$$

$$\nabla^2\underline{\mathbf{H}}(\mathbf{r},\omega) + \frac{\nabla\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)}{\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)}\times\left(\nabla\times\underline{\mathbf{H}}(\mathbf{r},\omega)\right) + k_0^2\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)\underline{\mathbf{H}}(\mathbf{r},\omega) = 0, \tag{2.23}$$

where $k_0 = \omega/c$ represents the free-space wavenumber. For arbitrary media with strong spatial variations of $\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)$, these equations cannot be solved analytically. Weakly inhomogeneous media, however, allow for some simplifications.

2.2.2 Wave equations for weakly inhomogeneous media

For homogeneous or weakly inhomogeneous media, where the dielectric constant $\underline{\epsilon}_{\mathrm{r}}(\mathbf{r},\omega)$ does not change significantly over an optical wavelength, we find that $|\nabla\underline{\epsilon}_{\mathrm{r}}/\underline{\epsilon}_{\mathrm{r}}| \ll |\underline{n}k_0|$ (see problem set), and we can neglect the second expressions on the left-hand sides of Eqs. (2.22) and (2.23) in comparison to the first ones. This leads to the so-called Helmholtz equations for the electric and magnetic fields,

$$\nabla^2 \underline{\mathbf{E}}(\mathbf{r},\omega) + k_0^2 \underline{\epsilon}_\mathrm{r}(\mathbf{r},\omega)\underline{\mathbf{E}}(\mathbf{r},\omega) = 0 \tag{2.24}$$

$$\nabla^2 \underline{\mathbf{H}}(\mathbf{r},\omega) + k_0^2 \underline{\epsilon}_\mathrm{r}(\mathbf{r},\omega)\underline{\mathbf{H}}(\mathbf{r},\omega) = 0 \tag{2.25}$$

Note that, in contrast to Eqs. (2.22) and (2.23), the vector components of the magnetic and electric fields are now decoupled, i.e., Eqs. (2.22) and (2.23) can each be decomposed into three scalar differential equations for three vector components.

2.2.3 Plane waves in isotropic homogeneous media

For isotropic homogenous media, $\underline{\epsilon}_\mathrm{r}(\mathbf{r},\omega) = \underline{\epsilon}_\mathrm{r}(\omega)$ is scalar, and the equations are solved by so-called plane waves of the form

$$\mathbf{E}(\mathbf{r},t) = \mathrm{Re}\left\{\underline{\mathbf{E}}(\mathbf{r},\omega)\,\mathrm{e}^{\,\mathrm{j}\,\omega t}\right\} = \mathrm{Re}\left\{\underline{\mathbf{E}}_0\,\mathrm{e}^{\,\mathrm{j}(\omega t - \underline{\mathbf{k}}\mathbf{r})}\right\}, \tag{2.26}$$

$$\mathbf{H}(\mathbf{r},t) = \mathrm{Re}\left\{\underline{\mathbf{H}}(\mathbf{r},\omega)\,\mathrm{e}^{\,\mathrm{j}\,\omega t}\right\} = \mathrm{Re}\left\{\underline{\mathbf{H}}_0\,\mathrm{e}^{\,\mathrm{j}(\omega t - \underline{\mathbf{k}}\mathbf{r})}\right\}, \tag{2.27}$$

where the wave vector $\underline{\mathbf{k}}$ defines the direction of propagation and obeys the relation

$$\underline{\mathbf{k}}^2 = k_0^2 \underline{\epsilon}_\mathrm{r}(\omega). \tag{2.28}$$

The surfaces of equal phase ("phase fronts") are planes that are normal to $\mathbf{k}$. Phase fronts travel with the phase velocity $v_p = c/n\,(\omega)$.

For a plane wave, the magnetic field can be derived from the electric field and vice versa. Introducing the plane-wave solution for the electric field, Eq. (2.26), into Maxwell's curl equations (2.2) and (2.4), we find

$$\underline{\mathbf{H}}_0 = \frac{1}{\omega\mu_0}\underline{\mathbf{k}} \times \underline{\mathbf{E}}_0, \tag{2.29}$$

$$\underline{\mathbf{E}}_0 = -\frac{1}{\omega\epsilon_0\epsilon_r}\underline{\mathbf{k}} \times \underline{\mathbf{H}}_0. \tag{2.30}$$

Maxwell's divergence equations (2.1) (2.3) lead to

$$\mathbf{k} \cdot \underline{\mathbf{E}}_0 = 0 \tag{2.31}$$

$$\mathbf{k} \cdot \underline{\mathbf{H}}_0 = 0. \tag{2.32}$$

In other words: The vectors ($\underline{\mathbf{k}}$, $\underline{\mathbf{E_0}}$, $\underline{\mathbf{H_0}}$) are mutually orthogonal and form a right-handed system.

Note that the wave vector $\underline{\mathbf{k}}$ is in general a complex number,

$$\underline{\mathbf{k}} = \mathbf{k} - \mathrm{j}\mathbf{k}_i = k_0 \underline{n}\mathbf{e}_k \tag{2.33}$$

where $\mathbf{e}_k$ denotes the unit vector in the direction of propagation. The time-averaged energy flux carried by the wave is obtained from the real part of the complex Poynting vector $\underline{\mathbf{S}}$ [14],

$$\mathrm{Re}\left\{\underline{\mathbf{S}}\right\} = \mathrm{Re}\left\{\frac{1}{2}\underline{\mathbf{E}}(\mathbf{r},\omega) \times \underline{\mathbf{H}}^\star(\mathbf{r},\omega)\right\} = \frac{\mathbf{k}}{2\omega\mu_0}\left|\underline{\mathbf{E}}_0\right|^2 e^{-\mathbf{2k}_i\mathbf{r}} \tag{2.34}$$

If we assume a plane wave propagating into positive z-direction, then $\mathbf{k} = k_0\underline{n}\mathbf{e}_z$ and the power decays as $e^{-\alpha z}$, where the power attenuation coefficient α is linked to the imaginary part of the refractive index by

$$\alpha = 2k_0 n_i. \tag{2.35}$$

Due to the convention to use negative signs for the imaginary parts in Eq. (2.21), a positive value of n_i corresponds to a positive attenuation coefficient α and therefore to optical loss, whereas negative values of n_i and α occur for media that show optical gain.

2.3 Material dispersion, Kramers-Kronig relations and Sellmeier equations

2.3.1 Causality and Kramers-Kronig relations

If an electric field **E** penetrates a dielectric medium, it generates an electric polarization **P**. In the time-domain this can be described by a convolution with an influence function $\chi(t)$, which corresponds to a multiplication with the Fourier transform $\underline{\chi}(f)$ of the influence function in the frequency domain, Eqs. (2.17) and (2.18). The Fourier transform is in general a complex function of frequency $f = \omega/(2\pi)$,

$$\underline{\chi}(f) = \chi(f) + \mathrm{j}\,\chi_i(f).$$

The time-domain influence function must be causal, i.e., $\chi(t < 0) = 0$, since no electric polarization can be induced before the electric field penetrates the medium. As a consequence, the real and the imaginary part of $\underline{\chi}(f)$ are related by the Hilbert transform,

$$\chi(f) = -\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{\chi_i(f_0)}{f_0 - f}\,\mathrm{d}f_0 \tag{2.36}$$

$$\chi_i(f) = \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{\chi(f_0)}{f_0 - f}\,\mathrm{d}f_0 \tag{2.37}$$

In these relations, $\mathcal{P}\int_{-\infty}^{\infty}\ldots\mathrm{d}f_0$ denotes the Cauchy principal value of the integral, i.e., the integral must be interpreted as a limiting value for which the integration limits approach the pole at $f_0 = f$ symmetrically from both sides,

$$\mathcal{P}\int_{-\infty}^{\infty}\frac{\chi(f_0)}{f_0 - f}\,\mathrm{d}f_0 = \lim_{\epsilon\to 0}\left(\int_{-\infty}^{f-\epsilon}\frac{\chi(f_0)}{f_0 - f}\,\mathrm{d}f_0 + \int_{f+\epsilon}^{\infty}\frac{\chi(f_0)}{f_0 - f}\,\mathrm{d}f_0\right). \tag{2.38}$$

Note that $\chi(t)$ is real and hence $\chi(f) = \chi(-f)$ and $\chi_i(f) = -\chi_i(-f)$. Inserting these relations in Eqs. (2.36) and (2.37), the so-called Kramers-Kronig relations can be derived:

$$\chi(f) = -\frac{2}{\pi}\mathcal{P}\int_{0}^{\infty}\frac{f_0\,\chi_i(f_0)}{f_0^2 - f^2}\,\mathrm{d}f_0 \tag{2.39}$$

$$\chi_i(f) = \frac{2}{\pi}\mathcal{P}\int_{0}^{\infty}\frac{f\,\chi(f_0)}{f_0^2 - f^2}\,\mathrm{d}f_0 \tag{2.40}$$

These relations allow us to calculate either the imaginary or the real part of the susceptibility if the respective other quantity is known. This implies that the absorption spectrum of a medium is linked to its frequency-dependent refractive index and vice versa. Absorption and dispersion are intimately related.
For a constant $\chi(f)$, i.e., $\chi(f) = \epsilon_r(f) - 1 = \mathrm{const}_f$, we find $\chi_i(f) = \epsilon_{ri}(f) = 0$ from Eq. (2.36), which implies $\chi(f) = 0$ and $\epsilon_r(f) = 1$, Eq. (2.37). There is hence no dispersionless medium with $n \neq 1$! Real media with refractive index $n \neq 1$ always have loss (or gain) in some frequency ranges, and the refractive index is always frequency-dependent. $\chi(f) \approx \mathrm{const}_f$ and $\chi_i \approx 0$ is only possible in certain frequency ranges.

2.3.2 The Lorentz oscillator model of dielectric media

To better understand the frequency dependence of the refractive index, let us consider a simple model for the interaction of an electric field **E** with a dielectric medium. Assuming that the medium does not contain any free charges, the external electric field leads to a displacement of charges within the individual atoms and molecules, thereby inducing an electric dipole moment,

Fig. 2.1. The electric polarization $\mathbf{P}$ is then given by the dipole moment that is induced per unit volume,

$$\mathbf{P} = N\mathbf{p}, \tag{2.41}$$

where N is the number of charges per unit volume. Assuming $\mathbf{E} = E_x\mathbf{e}_x$, and denoting the dis-

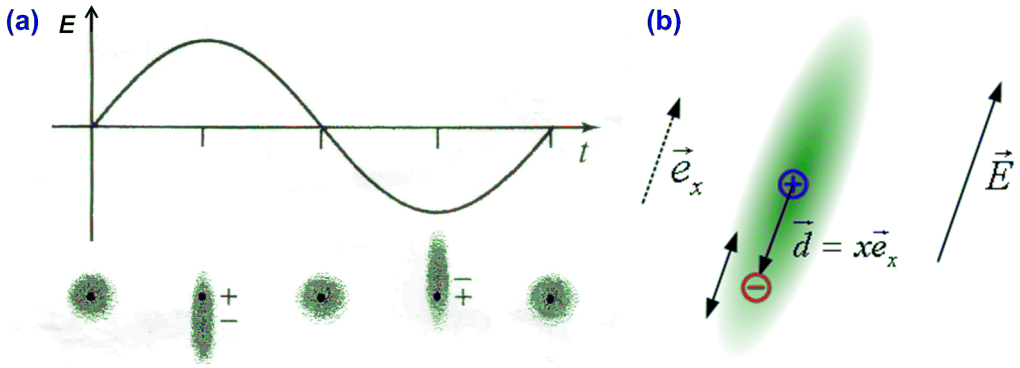

Figure 2.1: Lorentz model of electrons that are bound to a positively charged nucleus. (a) An external electric field E leads to a displacement of a bound electron and hence to a shift $\mathbf{d}$ of the center of negative charges away from the positively charged nucleus. This causes an electric dipole moment that oscillates with the frequency of the external field. (b) For an electric field oriented along the x-direction, the problem can be simplified to a one-dimensional equation of motion for the electron. (Figure adapted from [30]).

placement of a single bound electron with x, the dipole moment $\mathbf{p} = p_x\mathbf{e}_x$ is given by $p_x = -xe$. The displacement of the electron obeys Newton's second law, which states that the sum of forces (driving plus restoring force) equals electron mass m_e times acceleration,

$$m_e\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -eE_x - m_e\omega_r^2x - m_e\gamma_r\frac{\mathrm{d}x}{\mathrm{d}t}, \tag{2.42}$$

where $\gamma_r \ll \omega_r$. Solving this equation for a monochromatic excitation $E_x = \mathrm{Re}\{\underline{E}_x \exp(\mathrm{j}\,\omega t)\}$, the complex electric polarization is obtained, and the dielectric susceptibility can be derived,

$$\underline{\chi}(\omega) = \chi_0\frac{\omega_r^2}{\omega_r^2 - \omega^2 + j\omega\gamma_r}, \tag{2.43}$$

where

$$\chi_0 = \frac{Ne^2}{\epsilon_0 m_e\omega_r^2}. \tag{2.44}$$

The real and the imaginary parts of the complex electric susceptibility hence read

$$\chi(\omega) = \frac{\left(\omega_r^2 - \omega^2\right)\omega_r^2}{\left(\omega_r^2 - \omega^2\right)^2 + \omega^2\gamma_r^2}\chi_0 \tag{2.45}$$

$$\chi_i(\omega) = -\frac{\omega\gamma_r\omega_r^2}{\left(\omega_r^2 - \omega^2\right)^2 + \omega^2\gamma_r^2}\chi_0. \tag{2.46}$$

From these relations, we may deduce some very general properties of dielectric media:

- Well below resonance, i.e. for $\omega \ll \omega_r$ we find $\chi \approx \chi_0$ and $\chi_i \approx 0$. In the low-frequency limit, dielectric media hence exhibit only refraction, but no absorption.

- Well above resonance, i.e. for $\omega \gg \omega_r$ we find $\chi \approx \chi_i \approx 0$. The medium hence behaves essentially like free space; electrons cannot any more "follow" the electric field because it changes too fast. As a consequence, most materials appear transparent at X-ray frequencies and have a refractive index close to unity .

- Right at resonance, i.e. for $\omega = \omega_r$ we find $\chi \approx 0$ and $\chi_i \approx Q\chi_0$, where $Q = \omega_r/\gamma_r$. This implies strong absorption and a strongly frequency-dependent refractive index.

Eqs. (2.45) and (2.46) lead to a characteristic behavior of the refractive index near a resonance line, see Fig. (2.2). For real media, different electrons exhibit different resonance frequencies, Fig. (2.3), and the real and imaginary parts of the susceptibility are given by a sum over the various contributions,

$$\chi(\omega) = \sum_\nu \frac{\left(\omega_{r\nu}^2 - \omega^2\right)\omega_r^2}{\left(\omega_{r\nu}^2 - \omega^2\right)^2 + \omega^2\gamma_{r\nu}^2}\chi_{0\nu}, \tag{2.47}$$

$$\chi_i(\omega) = -\sum_\nu \frac{\omega\gamma_{r\nu}\omega_{r\nu}^2}{\left(\omega_{r\nu}^2 - \omega^2\right)^2 + \omega^2\gamma_{r\nu}^2}\chi_{0\nu}. \tag{2.48}$$

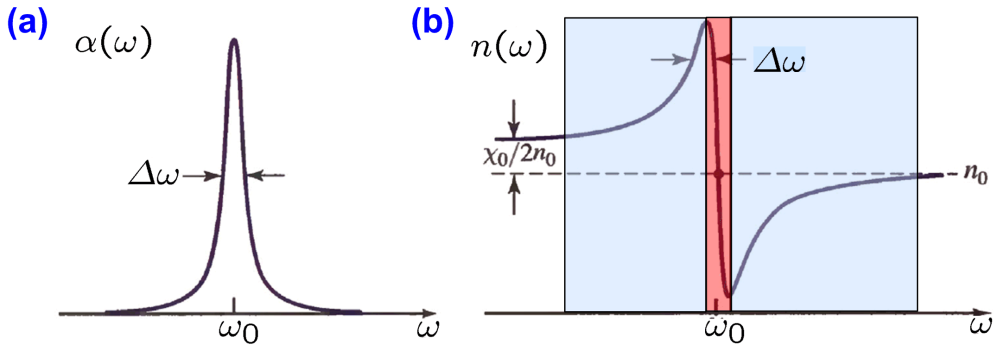

Figure 2.2: (a) Absorption coefficient α, and (b) refractive index n of a dielectric medium near resonance frequency ω_0. The absorption spectrum assumes the shape of a narrow line, whereas the refractive index undergoes a typical pattern which shows an increased refractive index for frequencies below resonance and a decreased refractive index for frequencies above resonance (Figure adapted from [30]).

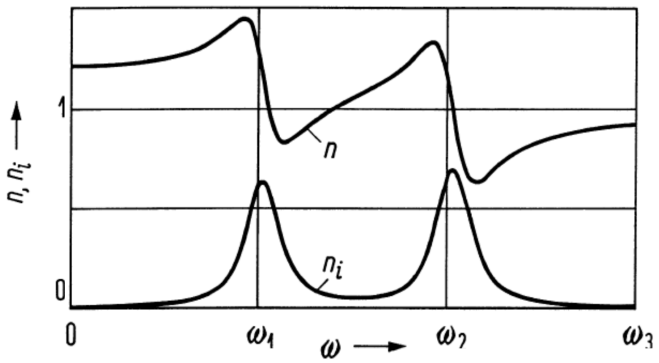

Figure 2.3: In real media, the real and imaginary parts of the susceptibility are given by a sum over the contributions of various electrons.

Example: Lenses for X-ray radiation

At X-ray frequencies, $\omega \gg \omega_r$, we find $\chi \approx \chi_i \approx 0$ and $\chi < 0$. Hence, materials are in essence transparent for these wavelengths, and X-rays are widely used for tomographic analysis of the interior structures of biological and technical samples. At the same time, the refractive indices at X-ray frequencies are close to unity, and it is therefore challenging to produce refractive optical components. X-ray lenses are therefore composed of arrays of single lens elements, Fig. (2.4). In contrast to normal optics, focusing X-ray lenses must have concave cross sections, since the refractive index is smaller than unity.

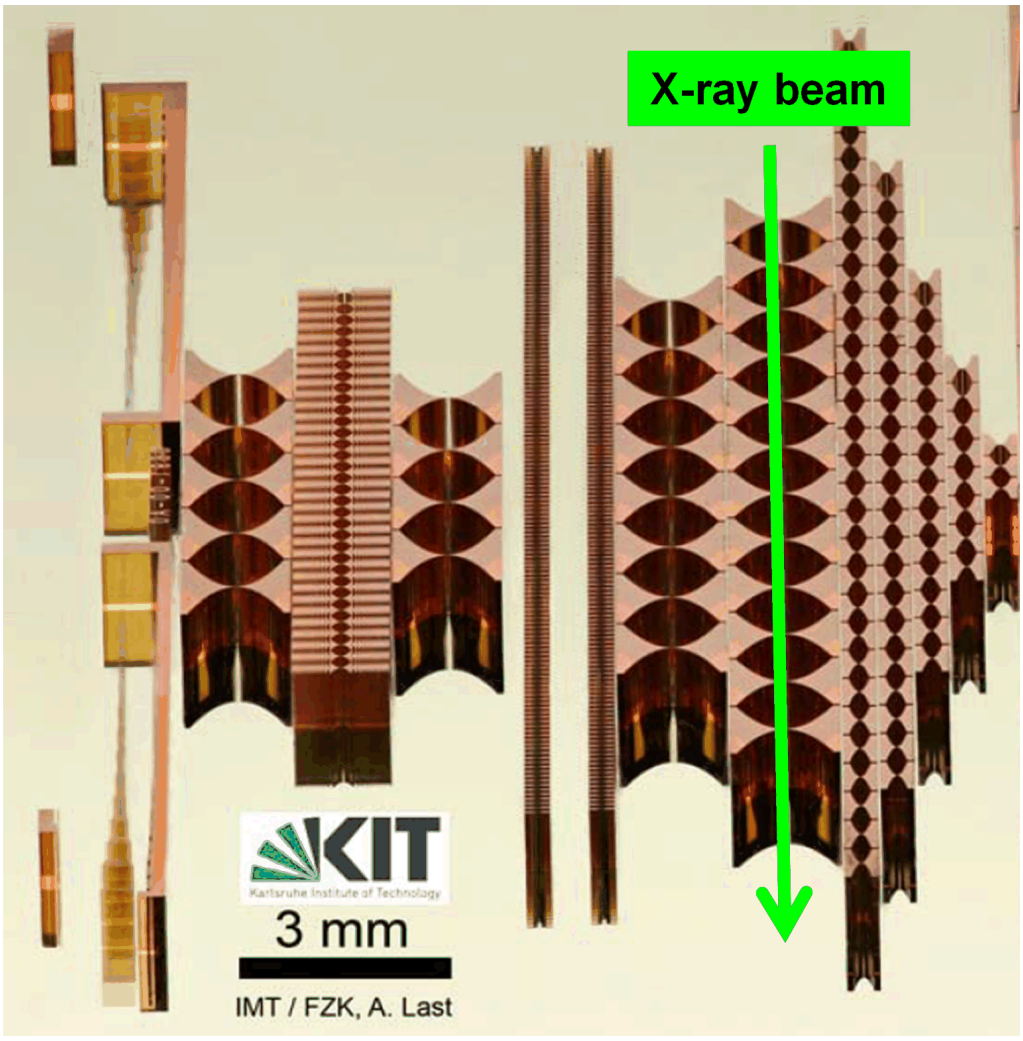

Figure 2.4: To obtain sufficient refractive power despite refractive indices n that are close to unity, X-ray lenses are composed of an array of individual lens elements. Since $n < 1$ for X-ray frequencies, focussing lenses must have concave shapes. (Source: A. Last, KIT-IMT)

2.3.3 The Drude model of conductive media

To model the optical properties of metallic media, we can perform an analysis similar to the one presented in Section 2.3.2. The only difference is that we are now dealing with free charges that are not bound to any nucleus. The restoring forces hence vanish and we may set $\omega_r = 0$ in Eq. (2.42).

The complex electric susceptibility can hence be written as

$$\underline{\chi}(\omega) = -\frac{\omega_p^2}{\omega^2}\frac{1}{1+\mathrm{j}\,\frac{\gamma_r}{\omega}} \approx -\frac{\omega_p^2}{\omega^2} \tag{2.49}$$

where the plasma frequency ω_p is given by

$$\omega_p^2 = \frac{Ne^2}{\epsilon_0 m_e}. \tag{2.50}$$

From these relations, we may again deduce some very general optical properties of metallic media:

- For $\omega < \omega_p$ we have $\underline{\epsilon}_\mathrm{r} = 1 - {}^{\omega_p^2}/_{\omega^2} < 0$. As a consequence, both $\underline{n}$ and $\underline{\mathbf{k}}$ are purely imaginary, which leads to attenuation of the wave without propagation ("forbidden band"). In addition, the wave impedance becomes imaginary, i.e., waves impinging on the material from outside are reflected at the surface.
- For $\omega > \omega_p$ we find $\underline{\epsilon}_\mathrm{r} = 1 - {}^{\omega_p^2}/_{\omega^2} > 0$. In this case, both $\underline{n}$ and $\mathbf{k}$ are purely real, i.e., the metal behaves like a lossless dielectric, but features unique dispersion characteristics. This band is referred to as the plasmonic band.
- For $\omega = \omega_p$ we have $\underline{\epsilon}_\mathrm{r} = 0$ and both $\underline{n}$ and $\mathbf{k}$ vanish. In this case, the wave does hence not travel in the conductive medium. The electric current density oscillates; the quantum quasi-particle associated with these oscillations is called a plasmon.

The special properties of metallic materials are important for plasmonic wave propagation, see Section 3.5 for a more detailed discussion.

2.3.4 Sellmeier equations

If dielectric media interact with an optical wave with a frequency far from any material resonances, we have $|\omega_r - \omega| \gg \gamma_r$ and the complex susceptibility can be approximated by

$$\underline{\chi}(\omega) \approx \frac{\omega_r^2}{\omega_r^2 - \omega^2}\chi_0 \tag{2.51}$$

In this case, $\underline{\chi}$ is approximately real and absorption is negligible. In real media, multiple resonances corresponding to different lattice and electronic vibrations contribute to the refractive index n. The dependence of n on frequency can hence be written as

$$n^2 = 1 + \sum_\nu \chi_{0\nu}\frac{f_\nu^2}{f_\nu^2 - f^2} = 1 + \sum_\nu \chi_{0\nu}\frac{\lambda^2}{\lambda^2 - \lambda_\nu^2} \tag{2.52}$$

The relation is known as the Sellmeier equation. The Sellmeier coefficients $\chi_{0\nu}$and $\lambda_\nu$are tabulated in reference books on optical materials [26] or material data sheets. The Sellmeier equations are a useful tool to calculate the wavelength-dependent refractive index of an optical material. Sellmeier coefficients of selected materials are listed in Fig. 2.5.

2.4 Signal propagation in dispersive media

Any time-dependent optical signal has nonzero bandwidth in the frequency domain, i.e. it consists of different frequency components. Due to the frequency dependence of the refractive index, these components travel at different speeds, which leads to a deformation of the optical signal during propagation. This phenomenon is referred to as group delay dispersion, group velocity dispersion, or just "dispersion". In optical telecommunications, "dispersion" embraces all effects, that lead to a spread in group delay time (or group velocity) of different spectral components. In this section, we will mathematically analyze the effect that a frequency-dependent refractive index has on the shape of an optical pulse traveling through a dispersive medium.

Material	Sellmeier Equation (Wavelength λ in μm)	Wavelength Range (μm)
Fused silica	$n^2 = 1 + \frac{0.6962\lambda^2}{\lambda^2 - (0.06840)^2} + \frac{0.4079\lambda^2}{\lambda^2 - (0.1162)^2} + \frac{0.8975\lambda^2}{\lambda^2 - (9.8962)^2}$	0.21–3.71
Si	$n^2 = 1 + \frac{10.6684\lambda^2}{\lambda^2 - (0.3015)^2} + \frac{0.0030\lambda^2}{\lambda^2 - (1.1347)^2} + \frac{1.5413\lambda^2}{\lambda^2 - (1104.0)^2}$	1.36–11
GaAs	$n^2 = 3.5 + \frac{7.4969\lambda^2}{\lambda^2 - (0.4082)^2} + \frac{1.9347\lambda^2}{\lambda^2 - (37.17)^2}$	1.4–11
BBO	$n_o^2 = 2.7359 + \frac{0.01878}{\lambda^2 - 0.01822} - 0.01354\lambda^2$	0.22–1.06
	$n_e^2 = 2.3753 + \frac{0.01224}{\lambda^2 - 0.01667} - 0.01516\lambda^2$	
KDP	$n_o^2 = 1 + \frac{1.2566\lambda^2}{\lambda^2 - (0.09191)^2} + \frac{33.8991\lambda^2}{\lambda^2 - (33.3752)^2}$	0.4–1.06
	$n_e^2 = 1 + \frac{1.1311\lambda^2}{\lambda^2 - (0.09026)^2} + \frac{5.7568\lambda^2}{\lambda^2 - (28.4913)^2}$	
$LiNbO_3$	$n_o^2 = 2.3920 + \frac{2.5112\lambda^2}{\lambda^2 - (0.217)^2} + \frac{7.1333\lambda^2}{\lambda^2 - (16.502)^2}$	0.4–3.1
	$n_e^2 = 2.3247 + \frac{2.2565\lambda^2}{\lambda^2 - (0.210)^2} + \frac{14.503\lambda^2}{\lambda^2 - (25.915)^2}$	

Figure 2.5: Sellmeier coefficients of various materials [30].

2.4.1 Frequency-domain representation

For simplicity, let us consider a plane wave with a carrier frequency $f_c = \omega_c/(2\pi)$ and a time-dependent optical amplitude propagating along the z-direction in a homogeneous dispersive dielectric material. The optical fields do not depend on x and y, so only the $z-$dependence is retained in the following equations.

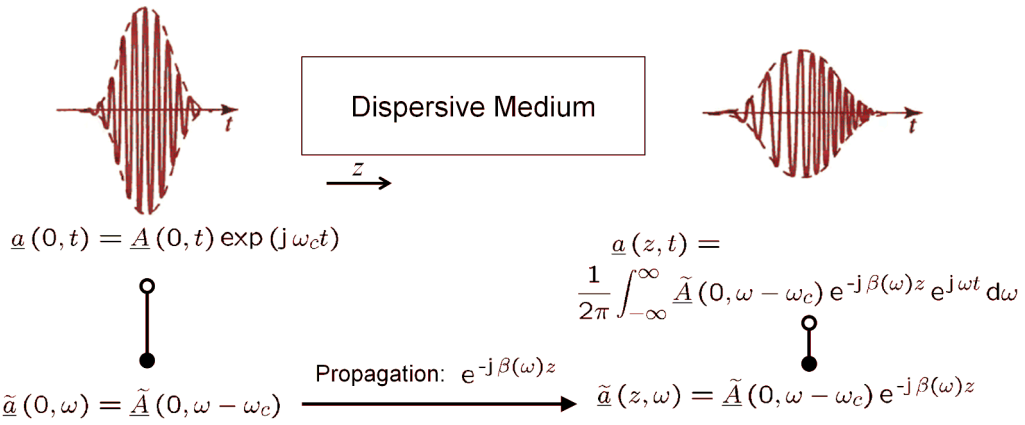

Figure 2.6: Analysis of pulse propagation in a dispersive medium. The dispersion properties of the medium are taken into account in frequency domain via the frequency-dependent propagation constant $\beta(\omega)$.

The basic principle of the analysis is illustrated in Fig. 2.6. For any given position along the $z-$axis, the time-dependence of the signal can be expressed by the carrier wave $\exp(\mathrm{j}\,\omega_c t)$ multiplied with a slowly varying complex envelope $\underline{A}(z,t)$,

$$\underline{a}(z,t) = \underline{A}(z,t)\exp(\mathrm{j}\,\omega_c t) \tag{2.53}$$

$$\underline{\tilde{a}}(z,\omega) = \underline{\tilde{A}}(z,\omega-\omega_c) \tag{2.54}$$

The signal is launched into the dispersive medium at $z = 0$,

$$\underline{a}(0,t) = \underline{A}(0,t)\exp(\mathrm{j}\,\omega_c t) \tag{2.55}$$

$$\underline{\tilde{a}}(0,\omega) = \underline{\tilde{A}}(0,\omega-\omega_c)\,. \tag{2.56}$$

The propagation along z is then most easily described in the frequency domain by multiplication with the frequency-dependent propagator $\exp(-\mathrm{j}\,\beta(\omega)z)$ in the Fourier domain,

$$\underline{\tilde{a}}(z,\omega) = \underline{\tilde{a}}(0,\omega)\exp(-\mathrm{j}\,\beta(\omega)z) \tag{2.57}$$

where

$$\beta(\omega) = \frac{\omega}{c}n(\omega) \tag{2.58}$$

corresponds to the real part of the $z-$component of the wavevector. We can now calculate the Fourier-domain representation of the signal at any $z > 0$ inside the medium

$$\begin{aligned}\underline{\tilde{a}}(z,\omega) &= \underline{\tilde{a}}(0,\omega)\exp(-\mathrm{j}\,\beta(\omega)z)\\ &= \underline{\tilde{A}}(0,\omega-\omega_c)\exp(-\mathrm{j}\,\beta(\omega)z)\end{aligned} \tag{2.59}$$

The corresponding time-domain signal is then obtained by inverse Fourier transformation,

$$\underline{a}\,(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{\underline{A}}\,(0,\omega-\omega_c)\exp\left(-\mathrm{j}\,\beta(\omega)z\right)\exp\left(\mathrm{j}\,\omega t\right)\mathrm{d}\omega. \tag{2.60}$$

For some simple cases, we can solve this integral analytically and derive important properties of signal propagation in dielectric media.

2.4.2 Taylor expansion of the dispersion relation and group velocity

For the simple case of narrow-band optical spectra it is useful to expand $\beta(\omega)$ in a Taylor series about the carrier frequency ω_c,

$$\beta(\omega) \approx \beta_c^{(0)} + (\omega-\omega_c)\beta_c^{(1)} + \frac{(\omega-\omega_c)^2}{2!}\,\beta_c^{(2)} + \frac{(\omega-\omega_c)^3}{3!}\,\beta_c^{(3)} + \dots\,, \tag{2.61}$$

where

$$\beta_c^{(i)} = \left.\frac{\mathrm{d}^i\beta(\omega)}{\mathrm{d}\omega^i}\right|_{\omega=\omega_c}. \tag{2.62}$$

Retaining only the first two terms $\beta_c^{(0)}$ and $\beta_c^{(1)}$, the signal in Eq. (2.60) can be written as

$$\underline{a}\,(z,t) = \underline{A}\left(0, t-\beta_c^{(1)}z\right)\exp\left(\mathrm{j}\left(\omega t-\beta_c^{(0)}z\right)\right). \tag{2.63}$$

This corresponds to a time-shifted copy of the input signal, modulated by a phase-shifted optical carrier, Fig. 2.7. The phase shift of the carrier wave is given by

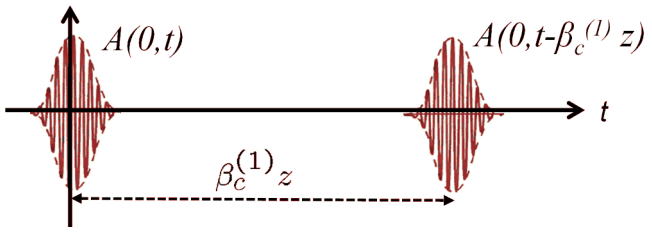

Figure 2.7: Signal propagation without consideration of group velocity dispersion. The signal envelope is delayed by $t_g = \beta_c^{(1)}z$; the carrier wave experiences a phase shift of $\beta_c^{(0)}z$. (Figure adapted from [30])

$$\beta_c^{(0)}z = \omega_c\frac{z}{v_p}\,, \qquad v_p = \frac{\omega_c}{\beta_c^{(0)}}\,, \qquad \beta_c^{(0)} = \frac{\omega_c}{c}n\left(\omega_c\right), \tag{2.64}$$

where v_p denotes the phase velocity of the symbol. The signal envelope travels at the group velocity v_g, and experiences a temporal group delay t_g ,

$$t_g = \frac{z}{v_g} = \beta_c^{(1)}z\,, \qquad v_g = \frac{1}{\beta_c^{(1)}} = \frac{c}{n_g}. \tag{2.65}$$

The group refractive index n_g depends on the center frequency ω_c (wavelength λ_c) and is given by

$$n_g\left(\omega_c\right) = n\left(\omega_c\right) + \omega_c\left.\frac{\mathrm{d}n(\omega)}{\mathrm{d}\omega}\right|_{\omega=\omega_c}, \tag{2.66}$$

$$n_g\left(\lambda_c\right) = n\left(\lambda_c\right) - \lambda_c\left.\frac{\mathrm{d}n(\lambda)}{\mathrm{d}\lambda}\right|_{\lambda=\lambda_c}. \tag{2.67}$$

2.4.3 Group velocity dispersion

The second- and higher-order terms of the Taylor expansion in Eq. (2.63) account for the frequency dependence of the group velocity. If any $\beta_c^{(i)} \neq 0$ for $i \geq 2$, different spectral components of the signal experience different group delays. This so-called group velocity dispersion (GVD) leads to a deformation of the signal envelope. In many cases of practical interest, the second-order dispersion related to $\beta_c^{(2)}$ dominates over all other effects. The temporal spread of two wave packets centered at ω_c and $\omega_c + \Delta\omega_c$ is then given by

$$\triangle t_g = \left.\frac{\mathrm{d}t_g(\omega)}{\mathrm{d}\omega}\right|_{\omega=\omega_c} \times \Delta\omega_c = \beta_c^{(2)}\,\Delta\omega_c\, z \tag{2.68}$$

It is therefore useful to define a material dispersion coefficient M_f, that relates the group delay spread $\triangle t_g$ of two spectral components to their frequency separation $\Delta f_c = \Delta\omega_c/(2\pi)$

$$\frac{\triangle t_g}{z} = M_f\,\Delta f_c\,, \qquad M_f = \frac{2\pi}{c}\frac{\mathrm{d}n_g(\omega)}{\mathrm{d}\omega} = 2\pi\beta_c^{(2)}. \tag{2.69}$$

For short pulses, the material dispersion coefficient M_f (unit: s/(m Hz)) is a measure of pulse-time broadening per unit distance and spectral width. The spectral separation can alternatively be expressed in terms of wavelength rather than frequency. The group delay spread of two spectral components centered around λ_c and $\lambda_c + \Delta\lambda_c$ is then given by

$$\frac{\triangle t_g}{z} = M_\lambda\,\Delta\lambda_c\,. \tag{2.70}$$

Using the wavelength-dependent group refractive index, Eq. (2.67), the material dispersion coefficient can be expressed as

$$M_\lambda = \frac{1}{c}\frac{\mathrm{d}n_g(\lambda)}{\mathrm{d}\lambda} = -\frac{2\pi c}{\lambda^2}M_f. \tag{2.71}$$

The material dispersion coefficient M_λ is usually given in units of ps/(km nm), where the group delay difference is given in ps, the propagation length in km, and the spectral separation in nm. Depending on the sign of M_λ, the dispersion properties of the material are referred to as normal ($\beta_c^{(2)} > 0$, $M_f > 0$, $M_\lambda < 0$) or anomalous ($\beta_c^{(2)} < 0$, $M_f < 0$, $M_\lambda > 0$) . The material dispersion of fused silica is plotted as a function of frequency in Fig. 2.8 along with the refractive index n and the group refractive index n_g. At the so-called zero-dispersion wavelength where the material dispersion coefficient vanishes ($\lambda_0 \approx 1.3\,\mu$m for fused silica), the group refractive index n_g reaches a minimum, and the frequency dependence of the refractive index n exhibits an inflection point.

2.4.4 Example: Dispersive broadening of a Gaussian impulse

As an example, let us consider a Gaussian impulse propagating in positive $z-$direction. At $z = 0$, the signal can be written as

$$\underline{a}\,(0,t) = \underline{A}_o \exp\left(-\frac{t^2}{2\sigma_t^2}\right)\exp\left(\mathrm{j}\,\omega_c t\right) \tag{2.72}$$

$$\widetilde{\underline{a}}\,(0,t) = \underline{A}_o\sqrt{2\pi\sigma_t^2}\exp\left(-\frac{\sigma_t^2\,(\omega-\omega_c)^2}{2}\right) \tag{2.73}$$

The spectral width is related to the impulse duration, and the time-bandwidth product assumes a minimum value at $z = 0$,

$$\sigma_\omega = 2\pi\sigma_f = \frac{1}{\sigma_t}; \qquad \sigma_t\sigma_\omega = 1. \tag{2.74}$$

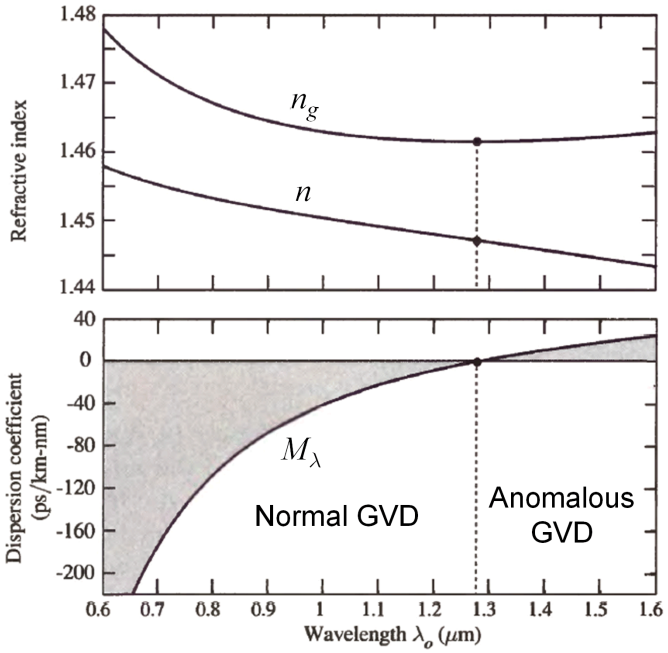

Figure 2.8: Wavelength dependence of the refractive index n, group refractive index n_g, and material dispersion coefficient M_λ of fused silica. The so-called zero-dispersion wavelength of $\lambda_0 \approx 1.3\,\mu$m is indicated by a dashed vertical line. At a wavelength of $1.55\,\mu$m, the material dispersion of fused silica amounts to $M_\lambda \approx 22 \frac{\text{ps}}{\text{km nm}}$ (Figure adapted from [30])

As shown on one of the problem sets, group delay dispersion leads to broadening of the signal envelope in the time domain, which results in an increase of the time-bandwidth product,

$$\sigma_t(z) = \sqrt{\sigma_t^2(0) + \frac{\left(\beta_c^{(2)} z\right)^2}{\sigma_t^2(0)}}, \tag{2.75}$$

$$\sigma_\omega(z) = \sigma_\omega(0), \tag{2.76}$$

$$\sigma_t(z)\,\sigma_\omega(z) = \sqrt{1 + \frac{\left(\beta_c^{(2)} z\right)^2}{\sigma_t^4(0)}}. \tag{2.77}$$

Note that in this example, the sign of $\beta_c^{(2)}$ does not play a role for the amount of dispersive pulse broadening. An intuitive explanation of impulse broadening is given in Fig. 2.9. For normal group velocity dispersion (GVD), we have $\beta_c^{(2)} > 0$ and $M_\lambda < 0$, i.e. longer wavelengths experience a smaller group delay, and will arrive first (at smaller times). The leading edge of the impulse will hence appear red-shifted. For anomalous GVD, in contrast, $\beta_c^{(2)} < 0$ and $M_\lambda > 0$, i.e. longer wavelengths experience a bigger group delay, and will arrive last (at bigger times). The leading edge of the impulse will appear blue-shifted.

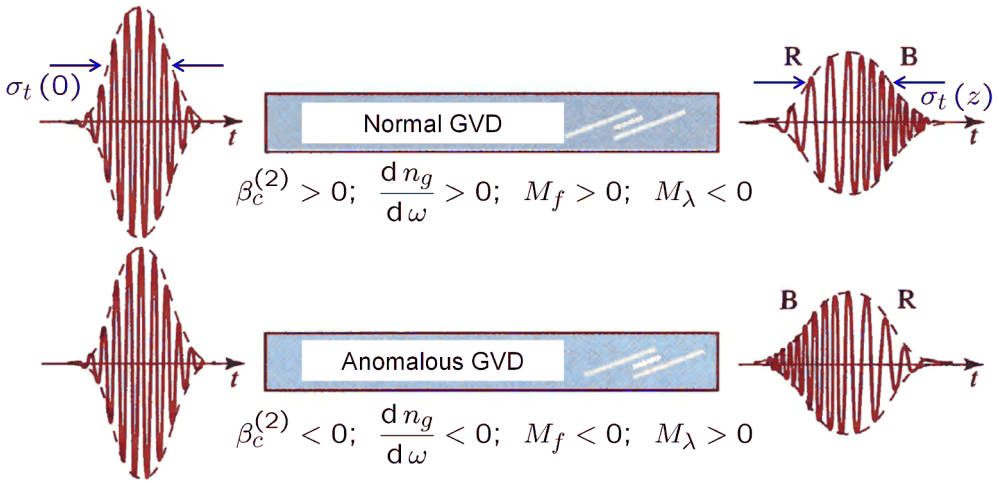

Figure 2.9: Principle of dispersive broadening of an optical impulse during propagation through media with normal and anomalous group velocity dispersion (GVD). Depending on the sign of the dispersion coefficient, the leading edge of the impulse will be red- or blue-shifted (R, B). (Figure adapted from [30])

Chapter 3

Slab waveguides

The most simple dielectric waveguide is the symmetric step-index slab waveguide, which consists of a waveguide core of refractive index n_1 embedded into a cladding material of refractive index n_2, see Fig. 3.1. The fundamental principle of waveguiding can be explained by a simple ray-optics picture: Light rays impinging on the core-cladding interface from inside with a sufficiently large angle of incidence (measured with respect to the normal of the interface plane) will undergo total internal reflection and thereby be confined to the waveguide core. This picture is qualitatively correct, but turns out inappropriate to explain a number of important effects such as the existence of waveguide modes or the origin of chromatic waveguide dispersion. In this section, we will derive an analytical electromagnetic model of the slab waveguide which allows us to study fundamental effects associated with signal propagation in waveguides. These insights are the basis for understanding more complex waveguide structures for which analytical descriptions do not exist.

We will approach the slab waveguide in two different ways, both leading to the same results. In a first approach, the reflection from the two core-cladding interfaces is first studied separately and then combined by enforcing lateral self-consistence of the reflected fields. This leads to a discrete set of guided modes. In a second approach, we formulate a plane-wave ansatz for the guided modes of an asymmetric slab waveguide.

3.1 Reflection from a plane dielectric boundary

Consider a plane wave impinging on a plane dielectric boundary between two media of refractive indices n_1 and n_2, see Fig. 3.2. For simplicity, we assume both media to be lossless, i.e., the associated refractive indices and the components of the wave vectors are real. As an ansatz, we use a superposition of three plane waves:

- Incident wave in region 1 ($x < 0$):

$$\underline{\mathbf{E}}_{\mathrm{i}}(\mathbf{r}) = \underline{\mathbf{E}}_1 \exp\left(-\mathrm{j}\,\mathbf{k}_1 \cdot \mathbf{r}\right), \qquad \mathbf{k}_1 = \begin{pmatrix} k_{1x} \\ 0 \\ k_{1z} \end{pmatrix} = n_1 k_0 \begin{pmatrix} \cos(\vartheta_1) \\ 0 \\ \sin(\vartheta_1) \end{pmatrix} \tag{3.1}$$

$$\underline{\mathbf{H}}_{\mathrm{i}}(\mathbf{r}) = \frac{1}{\omega\mu_0}\mathbf{k}_1 \times \underline{\mathbf{E}}_1 \exp\left(-\mathrm{j}\,\mathbf{k}_1 \cdot \mathbf{r}\right)$$

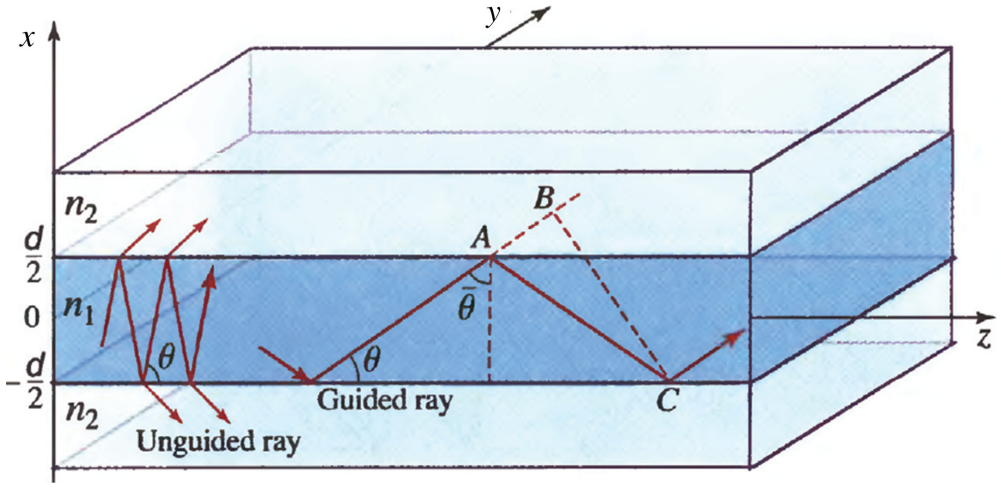

Figure 3.1: Basic principle of a symmetric slab waveguide, consisting of a waveguide core of refractive index n_1 embedded into a cladding material of refractive index n_2. In a simple picture, we can consider light "rays" that will undergo total internal reflection if they hit the core-cladding interface from inside with a sufficiently large angle of incidence $\bar{\Theta}$ (measured with respect to the normal of the interface plane). These rays will hence be confined to the waveguide core. Using more detailed analysis that takes into account the wave nature of light, we will show that the boundary conditions at both core-cladding interfaces can only be fulfilled simultaneously by discrete angles of $\bar{\Theta}$. (Figure adapted from [30])

- Transmitted wave in region 2 ($x > 0$):

$$\underline{\mathbf{E}}_{\mathrm{t}}(\mathbf{r}) = \underline{\mathbf{E}}_2 \exp(-\mathrm{j}\,\mathbf{k}_2 \cdot \mathbf{r}), \qquad \mathbf{k}_2 = \begin{pmatrix} k_{2x} \\ 0 \\ k_{2z} \end{pmatrix} = n_2 k_0 \begin{pmatrix} \cos(\vartheta_2) \\ 0 \\ \sin(\vartheta_2) \end{pmatrix} \tag{3.2}$$

$$\underline{\mathbf{H}}_{\mathrm{t}}(\mathbf{r}) = \frac{1}{\omega\mu_0}\mathbf{k}_2 \times \underline{\mathbf{E}}_2 \exp(-\mathrm{j}\,\mathbf{k}_2 \cdot \mathbf{r})$$

- Reflected wave in region 1 ($x < 0$):

$$\underline{\mathbf{E}}_{\mathrm{r}}(\mathbf{r}) = \underline{\mathbf{E}}_3 \exp(-\mathrm{j}\,\mathbf{k}_3 \cdot \mathbf{r}), \qquad \mathbf{k}_3 = \begin{pmatrix} k_{3x} \\ 0 \\ k_{3z} \end{pmatrix} = n_3 k_0 \begin{pmatrix} -\cos(\vartheta_3) \\ 0 \\ \sin(\vartheta_3) \end{pmatrix} \tag{3.3}$$

$$\underline{\mathbf{H}}_{\mathrm{r}}(\mathbf{r}) = \frac{1}{\omega\mu_0}\mathbf{k}_3 \times \underline{\mathbf{E}}_3 \exp(-\mathrm{j}\,\mathbf{k}_3 \cdot \mathbf{r})$$

For simplicity, we will omit the spacial arguments ($\mathbf{r}$) in the following. The corresponding dielectric displacements $\mathbf{D}$ and magnetic flux densities $\mathbf{B}$ can be obtained from the electric and magnetic field by using Eqs. (2.12) and (2.19) in the respective domains,

$$\underline{\mathbf{D}}_{\mathrm{i,r}} = \epsilon_0 \underline{n}_1^2 \underline{\mathbf{E}}_{\mathrm{i,r}} \tag{3.4}$$

$$\underline{\mathbf{D}}_{\mathrm{t}} = \epsilon_0 \underline{n}_2^2 \underline{\mathbf{E}}_{\mathrm{t}} \tag{3.5}$$

$$\underline{\mathbf{B}}_{\mathrm{i,r,t}} = \mu_0 \underline{\mathbf{H}}_{\mathrm{i,r,t}} \tag{3.6}$$

The plane-wave ansatz fulfills Maxwell's equations in the upper ($x < 0$) and in the lower half space ($x > 0$). To obtain a valid solution in the entire space, we have to additionally satisfy the following boundary conditions at $x = 0$:

- Normal components of $\mathbf{D}$ and $\mathbf{B}$ are continuous:

$$\left(n_1^2 \underline{\mathbf{E}}_\mathrm{i} + n_1^2 \underline{\mathbf{E}}_\mathrm{r} - n_2^2 \underline{\mathbf{E}}_\mathrm{t}\right) \cdot \mathbf{n} = 0 \tag{3.7}$$

$$\left(\mathbf{k}_1 \times \underline{\mathbf{E}}_\mathrm{i} + \mathbf{k}_3 \times \underline{\mathbf{E}}_\mathrm{r} - \mathbf{k}_2 \times \underline{\mathbf{E}}_\mathrm{t}\right) \cdot \mathbf{n} = 0 \tag{3.8}$$

- Tangential components of $\mathbf{E}$ and $\mathbf{H}$ are continuous:

$$\left(\underline{\mathbf{E}}_\mathrm{i} + \underline{\mathbf{E}}_\mathrm{r} - \underline{\mathbf{E}}_\mathrm{t}\right) \times \mathbf{n} = 0 \tag{3.9}$$

$$\left(\mathbf{k}_1 \times \underline{\mathbf{E}}_\mathrm{i} + \mathbf{k}_3 \times \underline{\mathbf{E}}_\mathrm{r} - \mathbf{k}_2 \times \underline{\mathbf{E}}_\mathrm{t}\right) \times \mathbf{n} = 0 \tag{3.10}$$

In these relations, $\mathbf{n}$ represents the surface normal, i.e., the unit vector that is perpendicular to the interface plane. The chosen plane wave ansatz is valid if and only if the boundary conditions according to Eqs. (3.7) - (3.10) can be fulfilled.

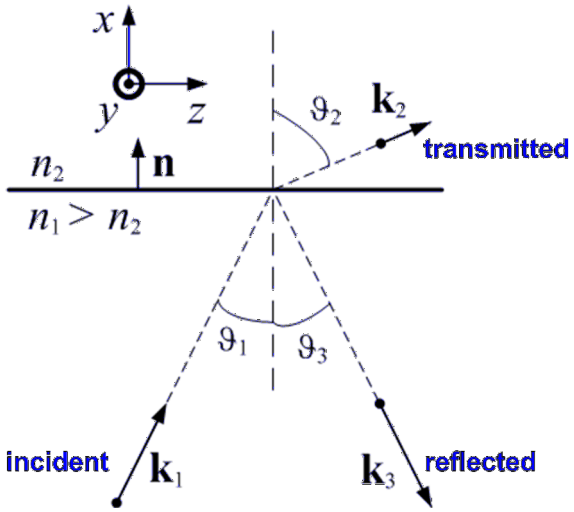

Figure 3.2: Reflection and transmission of light at a plane dielectric boundary. The propagation directions of the incident, reflected, and transmitted waves are given by the wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$, respectively. Using a plane-wave ansatz for each of the waves, we can fulfill Maxwell's equations in the lower ($x < 0$) and in the upper half space ($x > 0$), and we only need to match the boundary conditions at $x = 0$. The quantity $\mathbf{n}$ represents the surface normal, i.e., the unit vector that is perpendicular to the interface plane at $x = 0$.

3.1.1 Snell's law and law of reflection

The boundary conditions according to Eqs. (3.7) to (3.10) must be satisfied in all points of the (y, z)-plane simultaneously. Independent of the nature of the boundary conditions, the spatial variation of the fields in the (y, z)-plane must hence be the same,

$$k_{1y} = k_{2y} = k_{3y} = 0, \tag{3.11}$$

$$k_{1z} = k_{2z} = k_{3z}. \tag{3.12}$$

This leads to the law of reflection,

$$\vartheta_1 = \vartheta_3, \tag{3.13}$$

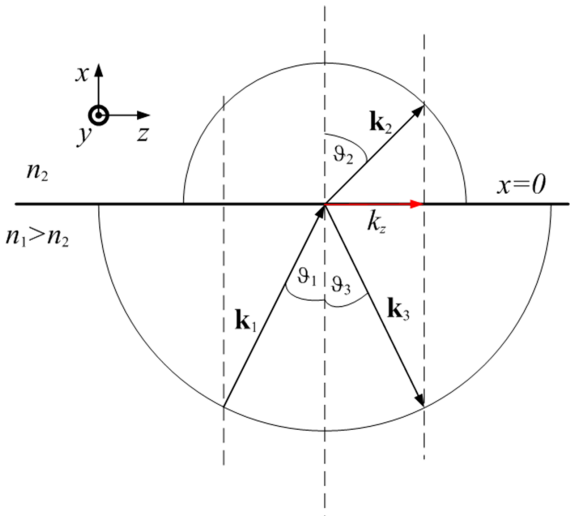

Figure 3.3: Graphical representation of Snell's law and the law of reflection. The boundary conditions must be satisfied in all points of the (y, z)-plane simultaneously. The spatial variation of the fields in the (y, z)-plane and hence the tangential component of the wave vectors (here: k_z) must be the same, whereas the length of the wave vectors is dictated by the refractive index, $|\mathbf{k}_1| = n_1 k_0$, $|\mathbf{k}_2| = n_2 k_0$, and $|\mathbf{k}_3| = n_3 k_0$. This leads directly to Eqs. (3.13) and (3.14).

and to Snell's law for the transmitted wave,

$$n_1 \sin(\vartheta_1) = n_2 \sin(\vartheta_2)\,. \tag{3.14}$$

For a graphical representation of these laws, see Fig. 3.3.

3.1.2 TE- and TM-polarization

To simplify further the analysis, we consider the boundary conditions separately for two different polarizations in the following, see Fig. 3.4. For a so-called transverse-electric wave (**TE-wave**), the electric field $\underline{\mathbf{E}}_\mathrm{i}$ is perpendicular to the plane of incidence whereas the magnetic field $\underline{\mathbf{H}}_\mathrm{i}$ is parallel to it,

$$\underline{\mathbf{E}}_\mathrm{i} = \begin{pmatrix} 0 \\ E_{\mathrm{i}y} \\ 0 \end{pmatrix}, \qquad \underline{\mathbf{H}}_\mathrm{i} = \begin{pmatrix} H_{\mathrm{i}x} \\ 0 \\ H_{\mathrm{i}z} \end{pmatrix}, \tag{3.15}$$

and we assume that this also applies to the reflected and the transmitted wave,

$$\underline{\mathbf{E}}_\mathrm{r} = \begin{pmatrix} 0 \\ E_{\mathrm{r}y} \\ 0 \end{pmatrix}, \qquad \underline{\mathbf{H}}_\mathrm{i} = \begin{pmatrix} H_{\mathrm{r}x} \\ 0 \\ H_{\mathrm{r}z} \end{pmatrix}, \tag{3.16}$$

$$\underline{\mathbf{E}}_\mathrm{t} = \begin{pmatrix} 0 \\ E_{\mathrm{t}y} \\ 0 \end{pmatrix}, \qquad \underline{\mathbf{H}}_\mathrm{i} = \begin{pmatrix} H_{\mathrm{t}x} \\ 0 \\ H_{\mathrm{t}z} \end{pmatrix}. \tag{3.17}$$

The validity of these assumptions is confirmed by showing that the boundary conditions at $x = 0$ can be fulfilled. TE-wave are also referred to as H-waves due to the longitudinal H_z-component of the magnetic field or s-polarized wave due to the orthogonal (German: senkrecht) electric field.

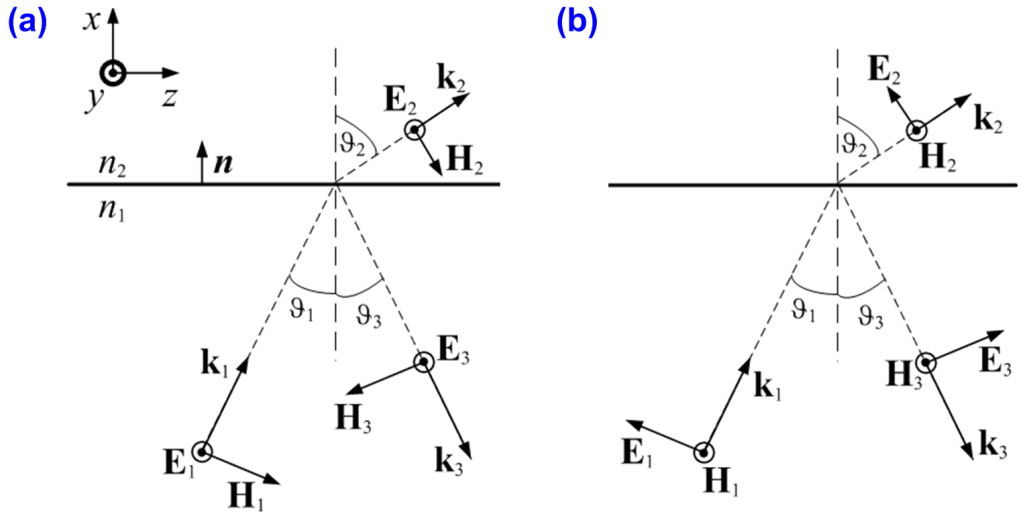

Figure 3.4: (a) Transverse-electric, and (b) transverse magnetic wave propagating in the (x, z)-plane and impinging on a plane dielectric boundary in the (y, z)-plane. Any arbitrary polarization state can be treated as a superposition of a TE and a TM wave.

Similarly, a so-called transverse-magnetic wave (**TM-wave**) has a magnetic field $\underline{\mathbf{H}}_\mathrm{i}$ which is perpendicular to the plane of incidence, whereas the electric field $\underline{\mathbf{E}}_\mathrm{i}$ is parallel to it.

$$\underline{\mathbf{H}}_\mathrm{i} = \begin{pmatrix} 0 \\ H_{\mathrm{i}y} \\ 0 \end{pmatrix}, \qquad \underline{\mathbf{E}}_\mathrm{i} = \begin{pmatrix} E_{\mathrm{i}x} \\ 0 \\ E_{\mathrm{i}z} \end{pmatrix}, \tag{3.18}$$

and the associated reflected and the transmitted waves can be written as,

$$\underline{\mathbf{H}}_\mathrm{r} = \begin{pmatrix} 0 \\ H_{\mathrm{i}y} \\ 0 \end{pmatrix}, \qquad \underline{\mathbf{E}}_\mathrm{r} = \begin{pmatrix} E_{\mathrm{r}x} \\ 0 \\ E_{\mathrm{r}z} \end{pmatrix}, \tag{3.19}$$

$$\underline{\mathbf{H}}_\mathrm{t} = \begin{pmatrix} 0 \\ H_{\mathrm{t}y} \\ 0 \end{pmatrix}, \qquad \underline{\mathbf{E}}_\mathrm{t} = \begin{pmatrix} E_{\mathrm{t}x} \\ 0 \\ E_{\mathrm{t}z} \end{pmatrix}. \tag{3.20}$$

TM-waves are also referred to as E-waves because of their longitudinal E_z-components or p-polarized wave due to the fact that the electric field is parallel to the plane of incidence. Any arbitrary polarization state can be treated as a superposition of a TE and a TM wave.

3.1.3 Reflection and transmission coefficients

Reflection and transmission coefficients are obtained by enforcing the boundary conditions for the electromagnetic field components. For TE-polarized fields, the continuity of the normal **D**-field component, Eq. (3.7), is inherently fulfilled by the field ansatz, and the continuity of the normal **B**-field component, Eq. (3.8), together with Snell's law, Eq. (3.14), implies the continuity of the tangential **E**-field, Eq. (3.9),

$$\underline{E}_{1y} + \underline{E}_{3y} = \underline{E}_{2y}. \tag{3.21}$$

The continuity of the tangential **H**-field leads to

$$\cos\vartheta_1 \underline{E}_{1y} - \cos\vartheta_1 \underline{E}_{3y} = \cos\vartheta_2 \underline{E}_{2y}. \tag{3.22}$$

From these relations, the amplitude reflection and transmission factors for TE-polarized waves are obtained,

$$R_{\mathrm{TE}} = \frac{E_{3y}}{E_{1y}} = \frac{k_{1x} - k_{2x}}{k_{1x} + k_{2x}} = \frac{n_1 \cos\vartheta_1 - n_2 \cos\vartheta_2}{n_1 \cos\vartheta_1 + n_2 \cos\vartheta_2} \tag{3.23}$$

$$T_{\mathrm{TE}} = \frac{E_{2y}}{E_{1y}} = 1 + R_{\mathrm{TE}} = \frac{2k_{1x}}{k_{1x} + k_{2x}} = \frac{2n_1 \cos\vartheta_1}{n_1 \cos\vartheta_1 + n_2 \cos\vartheta_2} \tag{3.24}$$

Similarly, the amplitude reflection and transmission factors for TM-polarization can be derived,

$$R_{\mathrm{TM}} = \frac{H_{3y}}{H_{1y}} = \frac{n_2^2 k_{1x} - n_1^2 k_{2x}}{n_2^2 k_{1x} + n_1^2 k_{2x}} = \frac{n_2 \cos\vartheta_1 - n_1 \cos\vartheta_2}{n_2 \cos\vartheta_1 + n_1 \cos\vartheta_2} \tag{3.25}$$

$$T_{\mathrm{TM}} = \frac{H_{2y}}{H_{1y}} = 1 + R_{\mathrm{TM}} = \frac{2n_2^2 k_{1x}}{n_2^2 k_{1x} + n_1^2 k_{2x}} = \frac{2n_1 \cos\vartheta_1}{n_2 \cos\vartheta_1 + n_1 \cos\vartheta_2} \tag{3.26}$$

The power reflection factors ρ are related to the amplitude reflection factors by

$$\rho_{\mathrm{TE}} = |R_{\mathrm{TE}}|^2 \tag{3.27}$$

$$\rho_{\mathrm{TM}} = |R_{\mathrm{TM}}|^2 \tag{3.28}$$

The corresponding power transmission factors τ can be calculated by exploiting the fact the interaction of the plane wave with the interface must be lossless, i.e., $\rho + \tau = 1$,

$$\tau_{\mathrm{TE}} = 1 - \rho_{\mathrm{TE}} \tag{3.29}$$

$$\tau_{\mathrm{TM}} = 1 - \rho_{\mathrm{TM}} \tag{3.30}$$

Note that $\tau_{\mathrm{TE}} \neq |T_{\mathrm{TE}}|^2$ ($\tau_{\mathrm{TM}} \neq |T_{\mathrm{TM}}|^2$) since the transmitted wave propagates in a different medium and under a different angle than the incoming wave, and the relation between the magnitude of the transverse electric (magnetic) field and the power flux in x-direction is hence different.

For a plane boundary between air ($n_1 = 1$) and GaAs ($n_2 = 3.6$), the power reflection factors are plotted as a function of incidence angle ϑ, see Fig. 3.5. For TE-polarization, the power reflection is minimum at normal incidence ($\vartheta = 0$), and it increases monotonically with ϑ, whereas for TM-polarization, the reflection vanishes for the so-called Brewster angle $\vartheta_{1\mathrm{B}}$ which is given by

$$\tan\vartheta_{1\mathrm{B}} = \frac{n_2}{n_1}. \tag{3.31}$$

Note that for $\vartheta_1 = \vartheta_{1\mathrm{B}}$, the propagation direction of the transmitted (refracted) wave is perpendicular to that of the reflected wave.

3.1.4 Total internal reflection

According to Snell's law, $\vartheta_2 > \vartheta_1$ if $n_1 > n_2$, and $\vartheta_2 = \pi/2$ if $\vartheta_1 = \vartheta_{1\mathrm{T}}$, where the limiting angle $\vartheta_{1\mathrm{T}}$ for total internal reflection is given by

$$\sin\vartheta_{1\mathrm{T}} = \frac{n_2}{n_1} \tag{3.32}$$

To understand what happens for $\vartheta_1 > \vartheta_{1\mathrm{T}}$, let us consider the longitudinal components of the wave vectors,

$$k_{2z} = k_{1z} = n_1 k_0 \sin\vartheta_1 > n_2 k_0, \tag{3.33}$$

see Fig. 3.6. Since $k_{2x}^2 + k_{2z}^2 = n_2^2 k_0^2$, the transverse wave vector component k_{2x} must be purely imaginary,

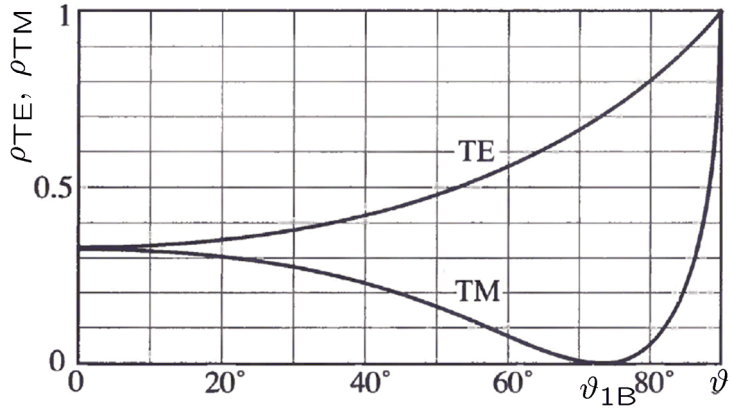

Figure 3.5: The power reflection factors as a function of incidence angle ϑ for a plane boundary between air ($n_1 = 1$) and GaAs ($n_2 = 3.6$). For TM-polarization, the reflection vanishes for the so-called Brewster angle ϑ_{1B}. (Figure adapted from [30])

$$k_{2x}^2 = n_2^2 k_0^2 - k_{2z}^2 < 0 \tag{3.34}$$

$$k_{2x} = \pm \mathrm{j}\, k_0 \sqrt{n_1^2 \sin^2 \vartheta_1 - n_2^2} = \pm \mathrm{j}\, k_{2x}^{(\mathrm{i})} \tag{3.35}$$

This leads to an evanescent "transmitted" field in region (2), which decays exponentially in the x-direction,

$$\underline{\mathbf{E}}_{\mathrm{t}} = \underline{\mathbf{E}}_2 \exp\left(-\mathrm{j}\, \mathbf{k}_2 \cdot \mathbf{r}\right) = \underline{\mathbf{E}}_2 \exp\left(-\mathrm{j}\, k_{2z} z\right) \exp\left(-k_{2x}^{(\mathrm{i})} x\right). \tag{3.36}$$

The refracted wave propagates along the surface and decays exponentially beyond the interface, i.e., no power is transmitted through the interface, which can be verified by calculating the Poynting vector of the evanescent field. The amplitude reflection factors for the incident wave are complex numbers with unit magnitude. For TE-polarization, we find

$$R_{\mathrm{TE}} = \frac{E_{3y}}{E_{1y}} = \frac{k_{1x} + \mathrm{j}\, k_{2x}^{(\mathrm{i})}}{k_{1x} - \mathrm{j}\, k_{2x}^{(\mathrm{i})}} = \exp\left(\mathrm{j}\, \varphi_{\mathrm{TE}}\right), \tag{3.37}$$

$$\varphi_{\mathrm{TE}} = 2 \arctan\left(\frac{k_{2x}^{(\mathrm{i})}}{k_{1x}}\right) = 2 \arctan\left(\frac{\sqrt{n_1^2 \sin^2 \vartheta_1 - n_2^2}}{n_1 \cos \vartheta_1}\right), \tag{3.38}$$

and similar relations hold for the TM case:

$$R_{\mathrm{TM}} = \frac{H_{3y}}{H_{1y}} = \frac{n_2^2 k_{1x} + \mathrm{j}\, n_1^2 k_{2x}^{(\mathrm{i})}}{n_2^2 k_{1x} - \mathrm{j}\, n_1^2 k_{2x}^{(\mathrm{i})}} = \exp\left(\mathrm{j}\, \varphi_{\mathrm{TM}}\right) \tag{3.39}$$

$$\varphi_{\mathrm{TM}} = 2 \arctan\left(\frac{n_1^2}{n_2^2} \frac{k_{2x}^{(\mathrm{i})}}{k_{1x}}\right) = 2 \arctan\left(\frac{n_1^2}{n_2^2} \frac{\sqrt{n_1^2 \sin^2 \vartheta_1 - n_2^2}}{n_1 \cos \vartheta_1}\right) \tag{3.40}$$

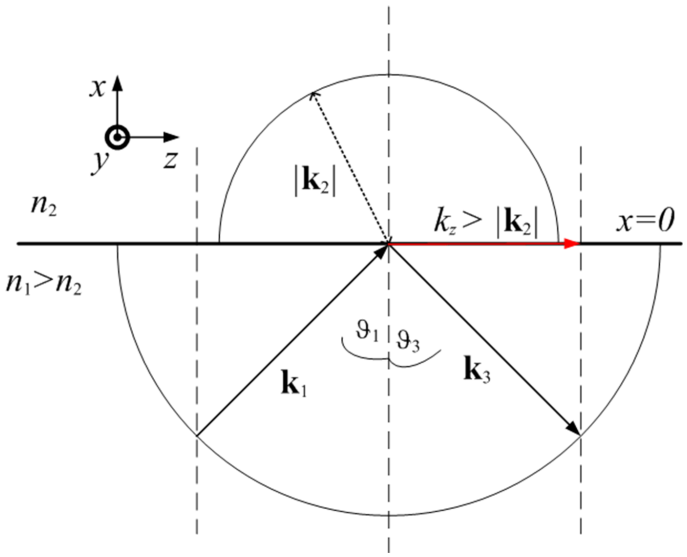

Figure 3.6: If the angle of incidence ϑ_1 exceeds the critical angle $\vartheta_{1\mathrm{T}}$, then the z-component k_z of the the wavevectors is larger than the total magnitude of the wavevector $\mathbf{k}_2$ in region 2. This results in a purely imaginary k_{2z} component and the wave is evanescent in positive $x-$direction.

Those two equations can be combined in a single relation,

$$R_\mathrm{p} = \exp\left(\mathrm{j}\,\varphi_p\right) \tag{3.41}$$

$$\varphi_\mathrm{p} = 2\arctan\left(\sigma_\mathrm{p}\,\frac{\sqrt{n_1^2\sin^2\vartheta_1 - n_2^2}}{n_1\cos\vartheta_1}\right) \tag{3.42}$$

$$\sigma_\mathrm{p} = \begin{cases} 1 & \text{for p=TE} \\ n_1^2/n_2^2 & \text{for p=TM} \end{cases} \tag{3.43}$$

Total internal reflection leads to a standing wave in the region of the incident field, see Fig. 3.7.

3.2 An intuitive approach to slab waveguide eigenmodes

3.2.1 Wave picture of the slab waveguide and eigenvalue equation

In the previous section, we have analyzed reflection and transmission of a plane wave at a plane dielectric interface. A slab waveguide consists of two such interfaces, Fig. 3.1, at each of which total internal reflection occurs, if the angle of incidence ϑ is bigger than the limiting angle $\vartheta_{1\mathrm{T}}$ for total internal reflection, Eq. (3.32). One might be tempted to conclude that wave patterns with arbitrary angles of incidence may propagate in the waveguide, provided that $\vartheta > \vartheta_{1\mathrm{T}}$. However, this would be an erroneous assumption; only discrete angles are permitted in reality. To understand this, we must remember that the rays shown in the Fig. 3.8 are only convenient representations of plane waves, each of which has a dedicated phase. The phase fronts are perpendicular to the propagation direction of the plane waves are indicated by dashed lines in Fig. 3.8. In addition, four points A, B, C, and D are marked in the figure. The ray from point A to point B does not undergo any reflection, whereas the longer ray from C to D is reflected twice. A and C lie on a common phase front, and the same applies to B and D. The phase difference accumulated along AB and CD must hence differ by an integer multiple of 2π,

$$-n_1 k_0 \overline{\mathrm{AB}} = -n_1 k_0 \overline{\mathrm{CD}} + 2\varphi_\mathrm{p} + m\,2\pi, \quad m \in \mathbb{N} \tag{3.44}$$

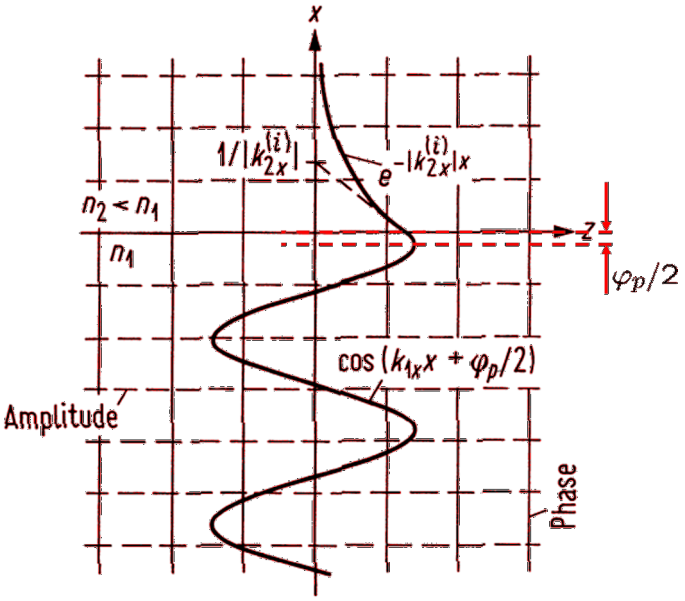

Figure 3.7: Field distribution for the case of total internal reflection at $x = 0$: We will find a standing wave in the region of the incident field, and an evanescent decay for $x > 0$. The phase shift φ_p depends on both the refractive indices n_1 and n_2 and on the angle of incidence ϑ_1.

From geometrical considerations we find that $\overline{\mathrm{AB}} = (h \tan\vartheta - h/\tan\vartheta)\sin\vartheta$ and $\overline{\mathrm{CD}} = h/\cos\vartheta$, and Eq. (3.44) can be transformed into

$$2h\, n_1 k_0 \cos\vartheta = 2\varphi_\mathrm{p} + m\, 2\pi \tag{3.45}$$

The equation has a discrete number of solutions for ϑ, each of which represents a so-called waveguide mode.

An equivalent relation can be derived from an alternative consideration. In the waveguide core, the field is represented by a superposition of an upward-propagating and a downward-propagating plane wave, whereas it is evanescent in the outer regions (waveguide cladding), Fig. 3.9 (a). A well-defined wave pattern within the waveguide core requires the partial waves reflected from the upper and lower interfaces to be consistent with each other or, in other words, the wave must reproduce itself after propagating through the waveguide thickness twice and after being reflected from the top and from the bottom interface, Fig. 3.9 (b). With these intuitive arguments, we can formulate a transverse self-consistence condition for a symmetric slab waveguide operated in TE-polarization by requiring the round-trip phase shift to be an integer multiple of 2π,

$$-2k_{1x}h + 2\varphi_\mathrm{TE} = -m \cdot 2\pi, \qquad m \in \mathbb{N}_0 \tag{3.46}$$

This relation is equivalent to Eq. (3.45) and can be solved by a discrete set of longitudinal propagation constants $\beta = k_{1z} = k_{2z}$ only. Each solution corresponds to a specific field pattern, a so-called waveguide mode. To simplifiy further analysis, we define some normalized quantities for the frequency and for the lateral and longitudinal components of the wave vectors in the core and

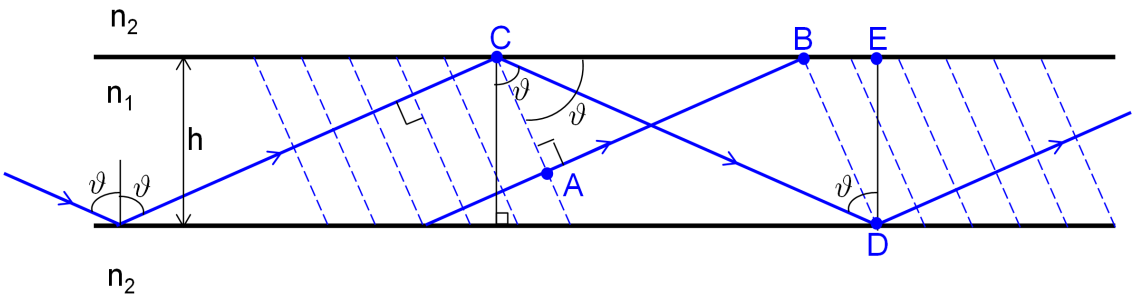

Figure 3.8: Rays and plane waves in slab waveguides: Each ray, represented by a solid blue line, represents a plane wave with a dedicated phase. The corresponding phase fronts are indicated by dashed lines. The ray from point A to point B does not undergo any reflection, whereas the longer ray from C to D is reflected twice. A and C lie on a common phase front, and the same applies to B and D. The phase difference accumulated along AB and CD must hence differ by an integer multiple of 2π, see Eq. (3.44).

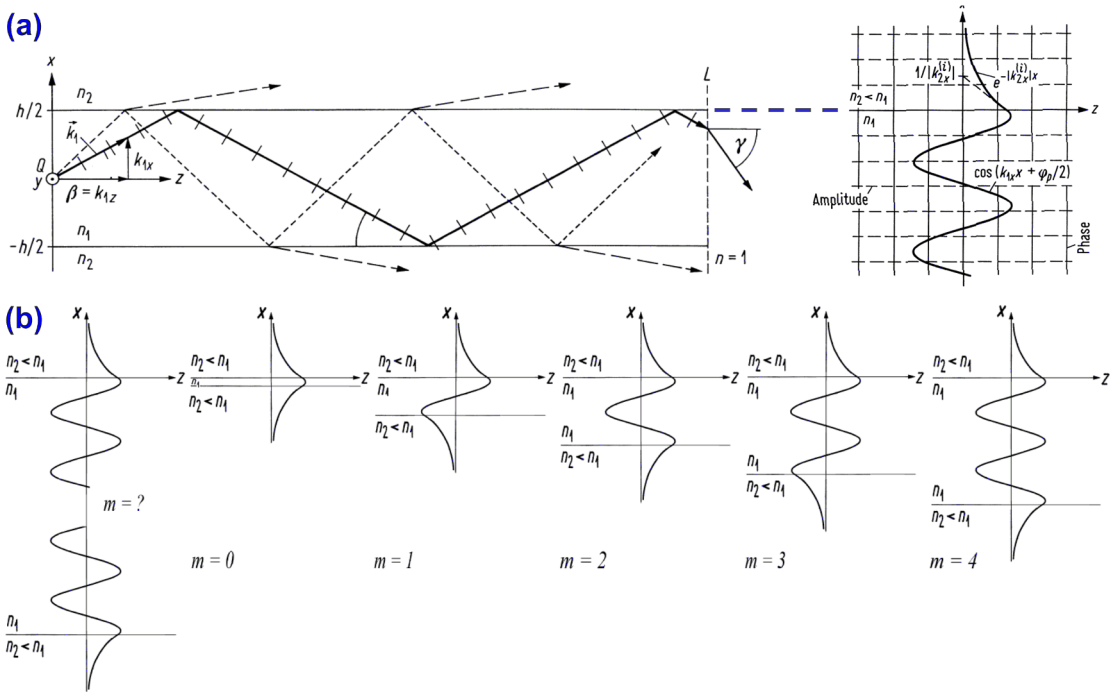

Figure 3.9: (a) An alternative approach to the slab waveguide: For a guided wave, the field is evanescent outside the waveguide core, and can be represented by a superposition of an upward-propagating and a downward-propagating plane wave within the core. (b) A well-defined wave pattern requires the partial waves reflected from the upper and lower interfaces to be consistent with each other. This leads to the formation of distinct field patterns that can propagate along the waveguide – so-called waveguide modes.

cladding regions,

$$u = \frac{h}{2} k_{1x} = \frac{h}{2} \sqrt{n_1^2 k_0^2 - \beta^2} \tag{3.47}$$

$$w = \frac{h}{2} k_{2x}^{(i)} = \frac{h}{2} \sqrt{\beta^2 - n_2^2 k_0^2} \tag{3.48}$$

$$V = \frac{h}{2} k_0 \sqrt{n_1^2 - n_2^2} = \frac{h}{2} k_0 A_N = \sqrt{u^2 + w^2} \tag{3.49}$$

$$A_N = \sqrt{n_1^2 - n_2^2} \tag{3.50}$$

$$\Delta = \frac{n_1^2 - n_2^2}{2 n_1^2} \approx \frac{n_1 - n_2}{n_1} \qquad \text{for} \qquad n_1 \approx n_2 \tag{3.51}$$

$$B = \frac{\beta^2 - n_2^2 k_0^2}{n_1^2 k_0^2 - n_2^2 k_0^2} = \frac{n_e^2 - n_2^2}{n_1^2 - n_2^2} = \frac{w^2}{V^2}, \qquad 0 < B < 1 \tag{3.52}$$

In these equations, u denotes the transverse core phase constant ("transversales Phasenmaß"), w is the transverse cladding attenuation ("transversales Dämpfungsmaß"), V is referred to as the normalized frequency, A_N is the numerical aperture of the waveguide, and Δ the relative refractive index difference. The quantity B represents the normalized propagation constant of the waveguide mode, and n_e is the effective refractive index of the mode, which can be thought of as an average index seen by the guided mode,

$$\beta = n_e k_0, \qquad n_2 < n_e < n_1, \qquad n_2 k_0 < \beta < n_1 k_0 \tag{3.53}$$

Inserting Eqs. (3.47) to (3.52) into Eq. (3.46) and using Eq. (3.38) for the phase of the reflection factor, we can derive the eigenvalue equation for TE modes,

$$u \tan\left(u - m\frac{\pi}{2}\right) = \sqrt{V^2 - u^2}. \tag{3.54}$$

Likewise, an equation for TM modes can be derived,

$$u \tan\left(u - m\frac{\pi}{2}\right) = \frac{n_1^2}{n_2^2} \sqrt{V^2 - u^2} \tag{3.55}$$

These relations can again be combined into a unified equation for TE and TM modes,

$$u \tan\left(u - m\frac{\pi}{2}\right) = \sigma_\mathrm{p} \sqrt{V^2 - u^2} \tag{3.56}$$

$$\sigma_\mathrm{p} = \begin{cases} 1 & \text{for p=TE} \\ n_1^2/n_2^2 & \text{for p=TM} \end{cases} \tag{3.57}$$

3.2.2 Graphical solution and discussion

To understand the implications of Eq. (3.54), we sketch the left-hand and the right-hand side as a function of transverse phase constant u, see Fig. 3.10 (a). The vertical axis corresponds to the transverse cladding attenuation w. The intersection points of the "stretched" branches of the tangent, $u \tan\left(u - m\frac{\pi}{2}\right)$, with the circle (ellipse), $\sigma_\mathrm{p} \sqrt{V^2 - u^2}$, correspond to guided TE$_m$-(TM$_m$-)modes of the slab waveguide. The radius (semiminor axis) of the circle (ellipse) is given by the normalized frequency V. From the intersection points, the eigenvalues u_{pm} and hence the propagation constants β_{pm} are found for a given normalized frequency V. The functional dependence of the (normalized) propagation constant β (B) on the (normalized) frequency ω (V) is also referred to as the dispersion relation of the respective mode, Fig. 3.10 (b). From these diagrams, a number of important conclusions can be drawn:

- The smaller V, the less modes are guided. The fundamental modes ($m = 0$) have the largest possible β, B (the smallest possible u) as compared to modes with higher mode index m.

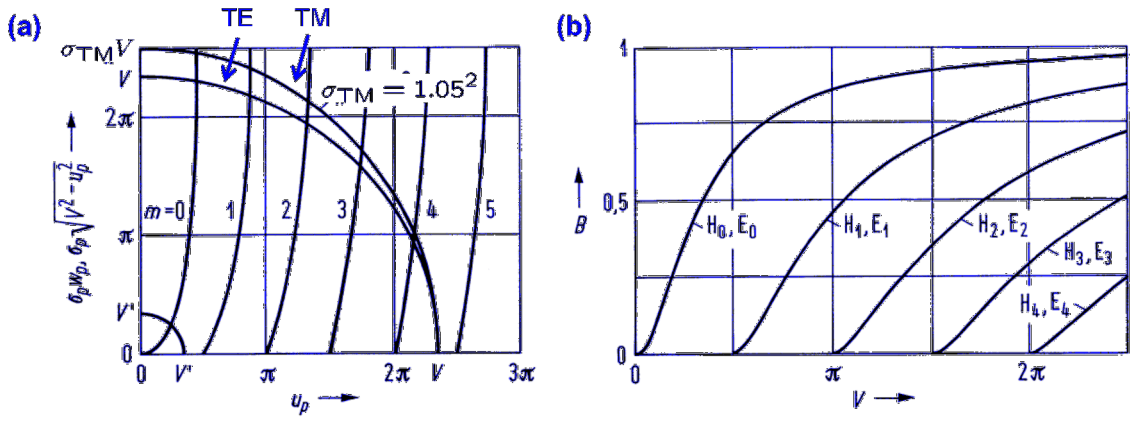

Figure 3.10: Graphical solution of the eigenvalue equation for TE and TM modes for a slab waveguide. (a) The two sets of graphs correspond to the left-hand and the right-hand side of Eq. (3.54) as a function of transverse phase constant u. The intersection points of the "stretched" branches of the tangent with the circle (ellipse) correspond to guided TE_m-(TM_m-)modes of the slab waveguide. The radius (semiminor axis) of the circle (ellipse) is given by the normalized frequency V. (b) (b) Dispersion relation $B(V)$ of various modes for the case of a low index-contrast waveguide (weak guidance, $\Delta \ll 1$). In this figure, TE_m- and TM_m-modes are denoted H_m and E_m, respectively. Since the index difference is small, TE- and TM-modes have virtually the same propagation constants.

- For $V < \frac{\pi}{2}$, there is only one guided TE and one guided TM mode. The waveguide is called single-moded.
- TE_m-modes have a always a larger β, B (smaller u) than the corresponding TM_m-modes of the same order m. This inequality can be also seen from the amplitude reflection coefficients for the two polarizations: The relation $r_{TE} > r_{TM}$ holds, therefore a TE_m-mode is always "more confined" to the high-index core region than the corresponding TM_m-mode. As a consequence, the propagation constant of a TE_m-mode must be closer to k_1 than in the TM_m case.
- In theory, the symmetric slab waveguide supports always at least one guided TE- and one guided TM-mode, no matter how low the frequency or how small the waveguide might be, i.e., there is no lower cut-off frequency. There are, however, practical limits: The lower the normalized frequency V, the further the evanescent tails of the guided modes extend into the cladding and the lower the contrast between the effective index n_e of the guided mode and the cladding index n_2 becomes. The wave is hence only weakly guided at low frequencies and may leak through the finite thickness of the cladding.
- For weak guidance $\Delta \ll 1$, the polarization parameter is nearly the same for both polarizations, $\sigma_{\mathrm{TM}} \approx \sigma_{\mathrm{TE}} = 1$. The propagation constants for TM_m-and TE_m-modes approach each other asymptotically, and both modes propagate with virtually the same β. Low-index contrast waveguides do not exhibit any appreciable polarization dependence, and their investigation can be restricted to a scalar analysis in many cases of practical interest.

Field patterns of guided modes ($m = 1 \ldots 3$) of a dielectric slab waveguide are depicted in Fig. 3.9 (b).

3.3 Mode ansatz for the asymmetric slab waveguide

In this section, we will use a more rigorous ansatz to calculate the eigenmodes of a slab waveguide. We will first introduce general properties of eigenmodes of longitudinally invariant structures, and then consider the special case of an asymmetric slab waveguide which is infinitely extended in one transverse direction.

3.3.1 Waveguide modes

A homogeneous waveguide in general is a dielectric structure which is invariant along the propagation direction of the optical power. Without loss of generality, we assume the light to propagate in positive z-direction. The refractive index profile can then be written as

$$n(\mathbf{r}) = n(x, y). \tag{3.58}$$

In the following, we assume that we have lossless waveguides, where $\mathrm{Im}\{\underline{n}\} = 0$ throughout the medium.

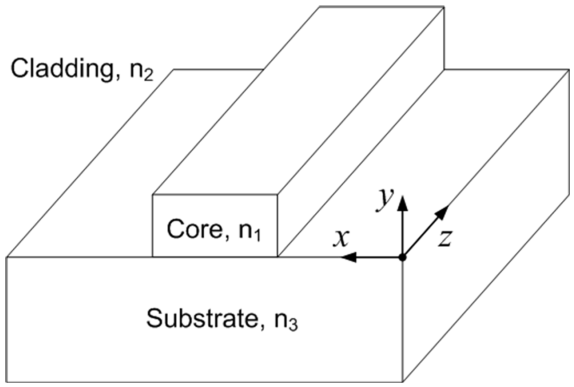

Figure 3.11: A homogeneous waveguide in general is a dielectric structure which is invariant along the propagation direction of the optical power. A lossless homogenous waveguide features a set of electromagnetic wave patterns which do not change their transverse shapes during propagation along z, so-called eigenmodes.

It can be shown that a lossless homogeneous waveguide possesses a set of eigenmodes, i.e., electromagnetic wave patterns which do not change their transverse shapes during propagation. The total field associated with a specific eigenmode is represented by

$$\underline{\mathbf{E}}(\mathbf{r}, t) = \underline{\mathcal{E}}(x, y) \exp(\mathrm{j}(\omega t - \beta z)), \tag{3.59}$$
$$\underline{\mathbf{H}}(\mathbf{r}, t) = \underline{\mathcal{H}}(x, y) \exp(\mathrm{j}(\omega t - \beta z)), \tag{3.60}$$
$$\beta = n_e k_0,$$

where the vectorial quantities $\underline{\mathcal{E}}(x, y)$ and $\underline{\mathcal{H}}(x, y)$ denote the electric and magnetic mode fields in the transverse plane, β corresponds to the propagation constant of the mode, and n_e is the effective refractive index which can be interpreted as an average refractive index which is "seen" by the wave. Eigenmodes with real propagation constant β (real effective refractive index n_e) are referred to as propagating eigenmodes. There are no propagating eigenmodes with a propagation constant larger than $n_1 k_0$, where n_1 is the maximum index in the waveguide cross section

$$|\beta| < n_1 k_0 \quad \text{for} \quad \beta \in \mathbb{R} \tag{3.61}$$

For so-called guided eigenmodes, the propagation constant β is real (or has at least a dominant real part) and the field profiles are confined to the core, i.e.,

$$\underline{\mathcal{E}}_m(x,y) \to 0 \text{ for } \left(x^2+y^2\right) \to \infty \tag{3.62}$$

$$\underline{\mathcal{H}}_m(x,y) \to 0 \text{ for } \left(x^2+y^2\right) \to \infty \tag{3.63}$$

Guided modes transport optical energy in positive z-direction along the waveguide; the mode fields decay evanescently outside the waveguide core. The guided modes form a discrete set; the corresponding propagation constants are in the range

$$n_2 k_0 < |\beta| < n_1 k_0, \tag{3.64}$$

where n_2 denotes the maximum refractive index in the cladding region.

In addition, there are so-called radiation modes, for which the fields extend to infinity. They transport power away from the waveguide core. Propagating radiation modes form continuous sets with propagation constants

$$|\beta| < n_2 k_0 \quad \text{for} \quad \beta \in \mathbb{R} \tag{3.65}$$

Radiation modes show an oscillatory behavior to at least one side of the waveguide structure. For general waveguide structures, calculation of radiation modes is difficult. Numerical solvers can in general not be used, because the transverse fields associated with a radiation mode extend to infinity and do not fit in any finite computational domain. Analytical solutions are only known for simple geometries such as slab waveguides. For waveguides with rectangular cross sections, semi-analytical methods can be used [27].

It can be shown that guided modes and radiation modes form a complete set. That means that every field distribution $\underline{\mathbf{E}}(\mathbf{r})$, $\underline{\mathbf{H}}(\mathbf{r})$ can be represented as a superposition of these modes

$$\underline{\mathbf{E}}(\mathbf{r}) = \sum_m a_m \underline{\mathcal{E}}_m(x,y) \exp\left(-\mathrm{j}\,\beta_m z\right) + \sum_\mu \int_\rho a_\mu(\rho)\, \underline{\mathcal{E}}_{\rho,\mu}(x,y) \exp\left(-\mathrm{j}\,\beta_\mu(\rho)\, z\right)\, \mathrm{d}\,\rho \tag{3.66}$$

$$\underline{\mathbf{H}}(\mathbf{r}) = \sum_m a_m \underline{\mathcal{H}}_m(x,y) \exp\left(-\mathrm{j}\,\beta_m z\right) + \sum_\mu \int_\rho a_\mu(\rho)\, \underline{\mathcal{H}}_{\rho,\mu}(x,y) \exp\left(-\mathrm{j}\,\beta_\mu(\rho)\, z\right)\, \mathrm{d}\,\rho \tag{3.67}$$

In addition, the transverse mode fields of the eigenmodes are orthogonal. We will come back to these properties when studying coupled-mode theory in Section 6.1.1.

3.3.2 Guided modes of slab waveguides

Maxwell's equations for waveguide modes

Inserting the mode ansatz, Eqs. (3.59) and (3.60) into Maxwell's equations, Eqs. (2.1) - (2.4), we find

$$\begin{aligned}
\frac{\partial \underline{\mathcal{E}}_z}{\partial y} + \mathrm{j}\,\beta \underline{\mathcal{E}}_y &= -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_x \\
-\frac{\partial \underline{\mathcal{E}}_z}{\partial x} - \mathrm{j}\,\beta \underline{\mathcal{E}}_x &= -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_y \\
\frac{\partial \underline{\mathcal{E}}_y}{\partial x} - \frac{\partial \underline{\mathcal{E}}_x}{\partial y} &= -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_z
\end{aligned} \tag{3.68}$$

and

$$\begin{aligned}
\frac{\partial \underline{\mathcal{H}}_z}{\partial y} + \mathrm{j}\,\beta \underline{\mathcal{H}}_y &= \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_x \\
-\frac{\partial \underline{\mathcal{H}}_z}{\partial x} - \mathrm{j}\,\beta \underline{\mathcal{H}}_x &= \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_y \\
\frac{\partial \underline{\mathcal{H}}_y}{\partial x} - \frac{\partial \underline{\mathcal{H}}_x}{\partial y} &= \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_z
\end{aligned} \tag{3.69}$$

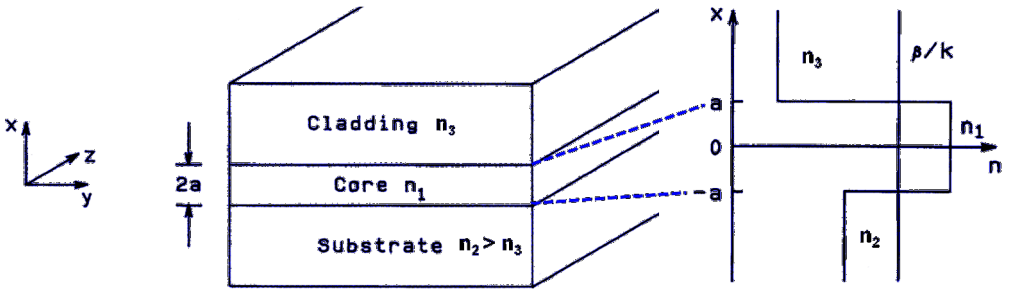

Figure 3.12: Asymmetric slab waveguide with core refractive index n_1, substrate index n_2, and cladding index n_3: Without loss of generality, we assume the structure to be infinitely extended in the y-direction, while light propagates in positive z-direction. (Figure adapted from [25])

Slab waveguide modes

In the following, we assume a slab waveguide which is infinitely extended in the y-direction, Fig. 3.12. In this case, all derivatives with respect to y vanish and we obtain

$$\begin{aligned} \mathrm{j}\,\beta\underline{\mathcal{E}}_y &= -\mathrm{j}\,\omega\mu_0\underline{\mathcal{H}}_x \\ -\frac{\partial\underline{\mathcal{E}}_z}{\partial x} - \mathrm{j}\,\beta\underline{\mathcal{E}}_x &= -\mathrm{j}\,\omega\mu_0\underline{\mathcal{H}}_y \\ \frac{\partial\underline{\mathcal{E}}_y}{\partial x} &= -\mathrm{j}\,\omega\mu_0\underline{\mathcal{H}}_z \end{aligned} \tag{3.70}$$

and

$$\begin{aligned} \mathrm{j}\,\beta\underline{\mathcal{H}}_y &= \mathrm{j}\,\omega\epsilon_0 n^2\underline{\mathcal{E}}_x \\ -\frac{\partial\underline{\mathcal{H}}_z}{\partial x} - \mathrm{j}\,\beta\underline{\mathcal{H}}_x &= \mathrm{j}\,\omega\epsilon_0 n^2\underline{\mathcal{E}}_y \\ \frac{\partial\underline{\mathcal{H}}_y}{\partial x} &= \mathrm{j}\,\omega\epsilon_0 n^2\underline{\mathcal{E}}_z \end{aligned} \tag{3.71}$$

We may separate these relations into a set of three equations for so-called TE-modes, for which the transverse mode field comprise nonzero $\underline{\mathcal{E}}_y$, $\underline{\mathcal{H}}_x$, $\underline{\mathcal{H}}_z$ modes only, and another set for so-called TM-modes, for which only the field components $\underline{\mathcal{H}}_y$, $\underline{\mathcal{E}}_x$, $\underline{\mathcal{E}}_z$ are different from zero. From these equations, we may derive wave equations for each of the mode families.

TE modes of the asymmetric slab waveguide

For the TE mode, we can derive the wave equation for the $\underline{\mathcal{E}}_y$-component,

$$\frac{\partial^2\underline{\mathcal{E}}_y(x)}{\partial x^2} + \left(\omega^2\mu_0\epsilon_0 n^2(x) - \beta^2\right)\underline{\mathcal{E}}_y(x) = 0. \tag{3.72}$$

We now use an ansatz of the form

$$\underline{\mathcal{E}}_y(x) = \begin{cases} A\cos(k_{1x}x - \varphi) & \text{for } -a \le x \le a \\ A\cos(-k_{1x}a - \varphi)\exp\left(k_{2x}^{(i)}(x+a)\right) & \text{for } x < -a \\ A\cos(k_{1x}a - \varphi)\exp\left(-k_{3x}^{(i)}(x-a)\right) & \text{for } x > a \end{cases}, \tag{3.73}$$

which fullfills the wave equation for $x \neq \pm a$ if

$$k_{1x} = \sqrt{n_1^2 k_0^2 - \beta^2}, \tag{3.74}$$

$$k_{2x}^{(i)} = \sqrt{\beta^2 - n_2^2 k_0^2}, \tag{3.75}$$

$$k_{3x}^{(i)} = \sqrt{\beta^2 - n_3^2 k_0^2}. \tag{3.76}$$

The continuity of $\underline{\mathcal{E}}_y$ and $\underline{\mathcal{H}}_{\mathbf{x}}$ at $x = \pm a$ are already included in the ansatz. In addition, we must fullfill the boundary conditions for the $\underline{\mathcal{H}}_z$-component. Using Eqs. (3.70) and (3.73), we find

$$\underline{\mathcal{H}}_z(x) = \frac{\mathrm{j}}{\omega\mu_0} \begin{cases} -k_{1x} A \sin(k_{1x} x - \varphi) & \text{for } -a \leq x \leq a \\ k_{2x}^{(i)} A \cos(-k_{1x} a - \varphi) \exp\left(k_{2x}^{(i)}(x+a)\right) & \text{for } x < -a \\ -k_{3x}^{(i)} A \cos(k_{1x} a - \varphi) \exp\left(-k_{3x}^{(i)}(x-a)\right) & \text{for } x > a \end{cases} \tag{3.77}$$

The continuity of $\underline{\mathcal{H}}_z$ at $x = \pm a$ now yields implicit equations for the mode propagation constant β and for the field asymmetry parameter φ,

$$\tan(u + \varphi) = \frac{w}{u} \tag{3.78}$$

$$\tan(u - \varphi) = \frac{w'}{u} \tag{3.79}$$

where

$$u = k_{1x} a = a\sqrt{n_1^2 k_0^2 - \beta^2}, \tag{3.80}$$

$$w = k_{2x}^{(i)} a = a\sqrt{\beta^2 - n_2^2 k_0^2}, \tag{3.81}$$

$$w' = k_{3x}^{(i)} a = a\sqrt{\beta^2 - n_3^2 k_0^2}. \tag{3.82}$$

Eqs. (3.78) and (3.79) can be rewritten as

$$u + \varphi = \arctan\left(\frac{w}{u}\right) + m_1 \pi, \tag{3.83}$$

$$u - \varphi = \arctan\left(\frac{w'}{u}\right) + m_2 \pi, \tag{3.84}$$

where m_1 and m_2 are integer numbers. These relations may be separated an implicit relation for the transverse core phase constant u and an equation that allows to calculate φ once u is known,

$$u = \frac{1}{2}\arctan\left(\frac{w}{u}\right) + \frac{1}{2}\arctan\left(\frac{w'}{u}\right) + \frac{m_u \pi}{2}, \tag{3.85}$$

$$\varphi = \frac{1}{2}\arctan\left(\frac{w}{u}\right) - \frac{1}{2}\arctan\left(\frac{w'}{u}\right) + \frac{m_\varphi \pi}{2}, \tag{3.86}$$

where $m_u = m_1 + m_2$ and $m_\varphi = m_1 - m_2$, which is equivalent to $m_\varphi = 2m_1 - m_u$. Hence, m_φ is even if m_u is even, and m_φ is odd if m_u is odd. Moreover, from Eq. (3.73), we know that values of φ which differ by an integer multiple of π lead to the same physical mode field. Therefore, for a given even m_u, all corresponding m_φ are even and belong to the same mode, and for a given odd m_u, all corresponding m_φ are odd and belong to the same mode. We may rewrite Eqs. (3.85) and (3.86) using a single quantity m instead of m_u and m_φ,

$$u = \frac{1}{2}\arctan\left(\frac{w}{u}\right) + \frac{1}{2}\arctan\left(\frac{w'}{u}\right) + \frac{m\pi}{2}, \tag{3.87}$$

$$\varphi = \frac{1}{2}\arctan\left(\frac{w}{u}\right) - \frac{1}{2}\arctan\left(\frac{w'}{u}\right) + \frac{m\pi}{2}. \tag{3.88}$$

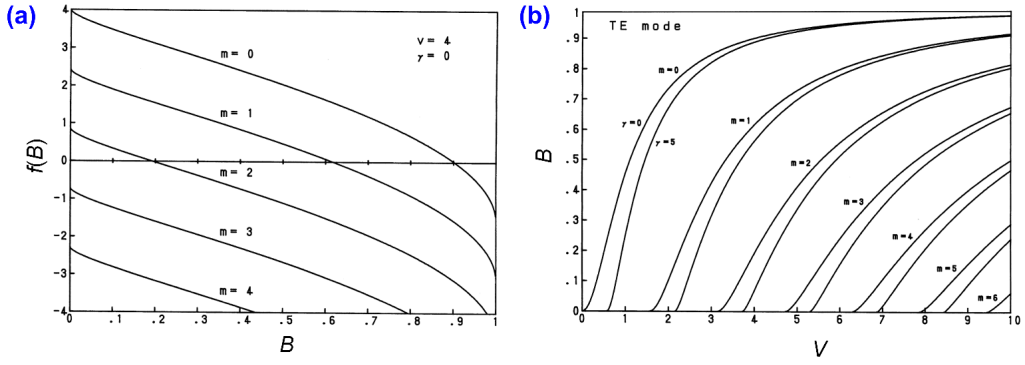

Figure 3.13: (a) Numerical solution of eigenvalue equations for TE-modes according to Eq. (3.92). (b) Dispersion relations for symmetric and asymmetric slab waveguides. There is no cutoff frequency for the fundamental waveguide mode of a symmetric waveguide ($\gamma = 0$), whereas a cutoff frequency exists for the fundamental mode of an asymmetric waveguide (Figure adapted from [25]).

In addition to the normalized frequency V, and the normalized propagation constant B, we introduce here a new parameter γ which is associated with the waveguide asymmetry,

$$V = ak_0\sqrt{n_1^2 - n_2^2} \tag{3.89}$$

$$\gamma = \frac{n_2^2 - n_3^2}{n_1^2 - n_2^2} \tag{3.90}$$

$$B = \frac{\beta^2 - n_2^2 k_0^2}{n_1^2 k_0^2 - n_2^2 k_0^2} = \frac{n_e^2 - n_2^2}{n_1^2 - n_2^2} \tag{3.91}$$

The eigenvalue equation for TE modes of the asymmetric slab waveguide can then be re-stated as

$$\underbrace{V\sqrt{1-B} - \frac{1}{2}\arctan\left(\sqrt{\frac{B}{1-B}}\right) - \frac{1}{2}\arctan\left(\sqrt{\frac{\gamma+B}{1-B}}\right) - \frac{m\pi}{2}}_{f(B)} = 0 \tag{3.92}$$

This relation can be solved numerically for the normalized propagation constant B by computing the zero of $f(B)$ for given parameters V, γ, and m, see Fig. 3.13 (a). Once B is known, we can calculate k_{1x} and φ as used in Eq. (3.73). For a symmetric slab waveguide, we have $\gamma = 0$, and Eq. (3.92) reproduces Eq. (3.54). The dispersion relations for symmetric and asymmetric slab waveguides are depicted in Fig. 3.13 (b). For a symmetrical waveguide, there is no cutoff frequency below which the fundamental waveguide mode ($\gamma = 0$) cannot propagate any more. In contrast to that, a lower cutoff frequency exists for the fundamental mode of an asymmetric waveguide.

TM modes of the asymmetric slab waveguide

For the TM modes, the wave equation for the transverse magnetic field component is found to be

$$n^2(x)\frac{\partial}{\partial x}\left(\frac{1}{n^2(x)}\frac{\partial}{\partial x}\underline{\mathcal{H}}_y(x)\right) + \left(\omega^2\mu_0\epsilon_0 n^2(x) - \beta^2\right)\underline{\mathcal{H}}_y(x) = 0 \tag{3.93}$$

A derivation similar to the one discussed for TE modes can be used here to obtain an eigenvalue equation for the transverse core phase constant of a guided TM mode,

$$u = \frac{1}{2}\arctan\left(\frac{n_1^2}{n_2^2}\frac{w}{u}\right) + \frac{1}{2}\arctan\left(\frac{n_1^2}{n_2^2}\frac{w'}{u}\right) + \frac{m\pi}{2}. \tag{3.94}$$

This relation can again be written in terms of the normalized frequency V, the normalized propagation constant B, and the waveguide asymmetry parameter γ as defined in Eqs. (3.89), (3.90), and (3.91),

$$\underbrace{V\sqrt{1-B} - \frac{1}{2}\arctan\left(\frac{n_1^2}{n_2^2}\sqrt{\frac{B}{1-B}}\right) - \frac{1}{2}\arctan\left(\frac{n_1^2}{n_3^2}\sqrt{\frac{\gamma+B}{1-B}}\right) - \frac{m\pi}{2}}_{f(B)} = 0 \tag{3.95}$$

Numerical solution of this relation allows to calculate the normalized propagation constant B. For $\gamma = 0$, Eq. (3.95) reproduces Eq. (3.55).

Examples of mode fields for asymmetric slab waveguides

For a strongly asymmetric slab waveguide ($n_1 = 3.38$, $n_2 = 3.17$, $n_3 = 1$) and a normalized frequency of $V = 4$, the three lowest-order TE mode fields are sketched in Fig. 3.14. When decreasing the normalized frequency, leakage into the substrate will occur while the mode field remains evanescent for $x > a$.

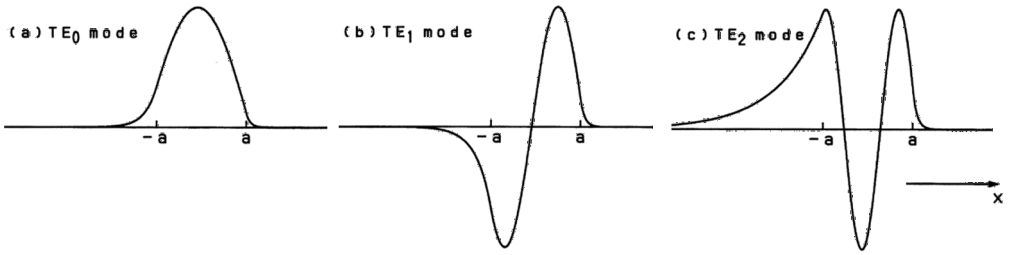

Figure 3.14: Examples of mode fields for a strongly asymmetric slab waveguide ($n_1 = 3.38$, $n_2 = 3.17$, $n_3 = 1$) and a normalized frequency of $V = 4$. The field penetrates much deeper into the high-index substrate ($x < -a$) than into the air cladding on top of the waveguide core ($x > a$). (Figure adapted from [25])

3.3.3 Radiation modes of slab waveguides

We will restrict our discussion to some qualitative remarks on radiation modes of slab waveguides. A rigorous analytical investigation for asymmetric slab waveguides can be found in [23].

A radiation mode of a slab waveguide can be thought of as a plane wave incident on the waveguide from the side, see Fig. 3.15. Radiation modes show an oscillatory behavior to at least one side of the waveguide structure. If, for an asymmetric slab waveguide, the refractive index n_2 of the substrate is larger than the refractive index n_3 of the cover material, so-called substrate modes can occur, which are evanescent in the cover region, but can propagate extended into the substrate, see Fig. 3.16. For these modes, the propagation constant β obeys the relation

$$n_3 k_0 < |\beta| < n_2 k_0. \tag{3.96}$$

For smaller propagation constants, so-called cover modes exist, which show oscillatory behavior in both the cover and the substrate,

$$|\beta| < n_3 k_0 \tag{3.97}$$

Moreover, even for lossless waveguides, there are eigenmodes with purely imaginary propagation constant β. These modes do not propagate in the longitudinal direction, but they decay evanescently. They can vary rapidly in the transverse direction with propagation constants

$$n_3 k_0 < k_x < \infty \tag{3.98}$$

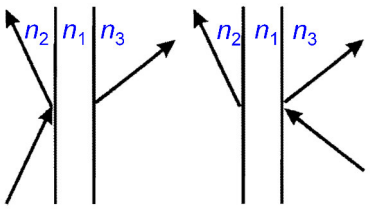

Figure 3.15: Radiation modes of slab waveguides: Propagating radiation modes can be thought of as plane waves impinging on the waveguide structure from outside. These modes show an oscillatory behavior to at least one side of the waveguide structure.

Such modes are needed to describe the fine structure of the field in the vicinity of a sub-wavelength waveguide imperfection. For a symmetric slab waveguide, the location of complex propagation constants β for different types of guided and radiation modes are sketched in Fig. 3.17.

3.4 Dispersion in dielectric optical waveguides

As in homogeneous media, any time-dependent signal traveling through a waveguide is subject to dispersion, i.e., to a spread of group delays for different spectral components of the signal. In this section, we will introduce different dispersion effects that occur within dielectric waveguides. We will come back to this in the context of optical fibers, where signals propagate over long distances and where dispersive broadening of the signals is therefore crucial, see section 5.6.2. As alluded to earlier, the notion of "dispersion" in optical telecommunications embraces all effects, which lead to a group delay spread (or a group velocity spread) of a signal propagating along a waveguide. Dispersion is also and more precisely referred to as "group delay dispersion" or "group velocity dispersion (GVD)".

3.4.1 Group and phase delay

Let us first consider a time-dependent signal propagating in a single waveguide mode. The signal has nonzero bandwidth in the frequency domain, i.e., it consists of different frequency components, which travel at different speeds. The signal propagation can be treated in a similar way as introduced in section 2.4 for plane waves in homogeneous media: First, all time-domain signals are transformed in the frequency domain. For each waveguide mode (mode index m), signal propagation is described by multiplication with the frequency-dependent propagator $\exp\left(-\mathrm{j}\,\beta_m(\omega)z\right)$. Taylor expansion of $\beta_m(\omega)$ about the carrier frequency ω_c allows to identify the phase delay and the group delay in analogy to section 2.4.

For a certain mode, the phase delay is given by $\beta_m\left(\omega_c\right)z$, where the modal propagation constant β_m is related to the mode's effective refractive index n_e by

$$\beta\left(\omega_c\right) = \frac{\omega_c}{c} n_e\left(\omega_c\right), \tag{3.99}$$

where the mode index m has been omitted for the sake of readability. Similarly, the group delay t_g and the group velocity v_g can be described by using the mode's effective refractive index n_e and the corresponding effective group refractive index n_{eg},

$$t_g = \frac{z}{v_g} = \beta_m^{(1)}\left(\omega_c\right) z\,, \qquad v_g = \frac{1}{\beta_m^{(1)}\left(\omega_c\right)} = \frac{c}{n_{eg}}. \tag{3.100}$$

The effective group refractive index n_{eg} depends on the center frequency ω_c (wavelength λ_c) and

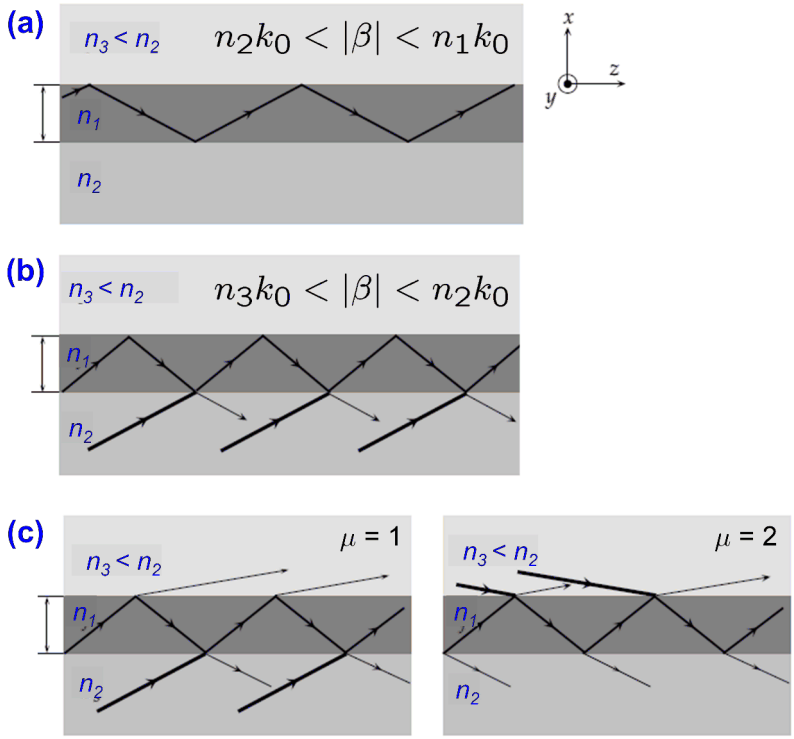

Figure 3.16: Ray pictures of slab waveguide eigenmodes: (a) Guided mode; (b) Radiation mode, which propagates in the substrate (refractive index n_2) only while it is evanescent in the cover region (refractive index $n_3 < n_2$); these modes are also referred to as substrate modes; (c) Radiation modes which propagate in both the substrate and the cover region, so-called cover modes. For a given propagation constant β two distinct radiation modes with different field patterns can be found, indicated by the mode indices $\mu = 1$ and $\mu = 2$. (Figure adapted from [5])

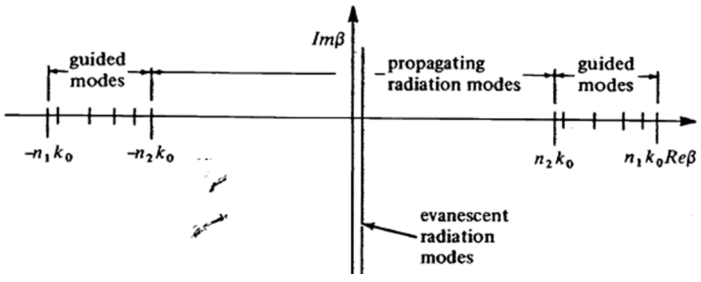

Figure 3.17: Location of mode propagation constants β of a symmetric slab waveguide in the complex plane. There exists a discrete set of guided modes with real propagation constants $n_2k_0 < |\beta| < n_1k_0$. Propagating radiation modes can be found on the real axis for $0 < |\beta| < n_2k_0$, whereas evanescent radiation modes occupy the imaginary axis. (Figure adapted from [22])

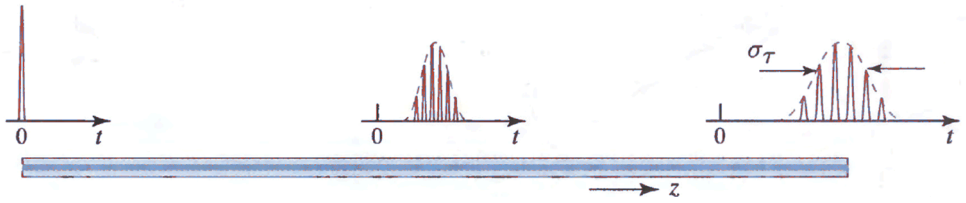

Figure 3.18: Intermodal dispersion in a multimode waveguide: A single optical impulse is launched into a multimode waveguide at $z = 0$, thereby exciting a multitude of different waveguide modes. Each of the modes experiences a distinct group delay, leading to a sequence of optical impulses at the output of the waveguide. (Figure adapted from [30])

is given by

$$n_{eg}(\omega_c) = n_e(\omega_c) + \omega_c \left.\frac{\mathrm{d}n_e(\omega)}{\mathrm{d}\omega}\right|_{\omega=\omega_c}, \tag{3.101}$$

$$n_{eg}(\lambda_c) = n_e(\lambda_c) - \lambda_c \left.\frac{\mathrm{d}n_e(\lambda)}{\mathrm{d}\lambda}\right|_{\lambda=\lambda_c}. \tag{3.102}$$

Higher-order contributions $\beta_m^{(\nu)}(\omega_c)$, $\nu = 2, 3, \ldots$ in the Taylor expansion of $\beta_m(\omega)$ lead to chromatic dispersion, i.e., to a frequency-dependent spread of the group delay, see section 3.4.3.

3.4.2 Intermodal dispersion and polarization mode dispersion

Note that the group delay as defined in Eq. (3.100) depends on the considered waveguide mode (mode index m), i.e., signals propagating in different waveguide modes experience different group delays,

$$t_{g,m} = \beta_m^{(1)}(\omega_c)\, L \tag{3.103}$$

When coupling light to a multimode waveguide, it is impossible to excite exactly one waveguide mode. As a consequence, a single impulse of light launched into a multimode waveguide will resolve into a sequence of impulses at the output of the waveguide, see Fig. 3.18. The corresponding group delay spread is referred to as **intermodal dispersion** or **modal dispersion**. For multimode waveguides, this is the dominating effect. It is usually much stronger than chromatic dispersion which will be discussed in the next section.

In so-called single-mode waveguides, two modes with different polarization states can still propagate. They usually have slightly different propagation constants and group velocities. The associated group delay spread is referred to as **polarization mode dispersion (PMD)**. PMD occurs also in optical fibers with nominally circular cross sections, see section 5.6.3: Under real environmental conditions, these fibers exhibit a small birefringence which varies randomly along the waveguide and which leads to a group delay spread for different polarizations.

3.4.3 Chromatic dispersion

If, in a single-mode waveguide, only one polarization is excited, the signal still experiences temporal broadening because different spectral components of the signal experience different group velocities. The associated spread of the wavelength-dependent group delays is referred to as chromatic dispersion and can be expanded into a power series about the center wavelength,

$$\frac{\triangle t_g}{z} = C_\lambda\, \Delta\lambda_c + D_\lambda\, \Delta\lambda_c^2, \tag{3.104}$$

$$C_\lambda = -\frac{2\pi c}{\lambda^2}\beta_c^{(2)}, \tag{3.105}$$

where $\beta_c^{(2)} = \beta_m^{(2)}(\omega_c)$, and where the mode index m has been omitted for the sake of clarity. The chromatic dispersion coefficient C_λ is usually given in units of ps/(km nm), where the group delay difference is given in ps, the propagation length in km, and the spectral separation in nm. The second-order expansion D_λ coefficient describes the wavelength-dependence of the chromatic dispersion and is therefore also referred to as the dispersion slope.

For signal propagation in a homogeneous medium, group delay dispersion is caused by the medium's material dispersion, i.e. the frequency dependence of the refractive index, see section 2.4. In a waveguide, however, both the material dispersion of the waveguide medium and the frequency-dependence of the waveguiding mechanism itself contributes to chromatic dispersion. In general, these two effects are intimately interlinked. For weakly guided modes, however, the two contributions can be explicitely separated,

$$\frac{t_g}{L} = \frac{\mathrm{d}\beta}{\mathrm{d}\omega} = \frac{1}{c}\frac{\mathrm{d}\beta}{\mathrm{d}k_0} \approx \left\{ \begin{array}{c} \Delta \ll 1 \\ \frac{\mathrm{d}V}{\mathrm{d}k_0} \approx \frac{V}{k_0} \\ n_{1g} - n_{2g} \\ \approx n_1 - n_2 \end{array} \right\} \approx \underbrace{\frac{n_{2g}}{c}}_{\text{material}} + \underbrace{\frac{n_{1g} - n_{2g}}{c} \overbrace{\frac{\mathrm{d}(VB)}{\mathrm{d}V}}^{\text{group delay factor}}}_{\text{waveguide}} . \tag{3.106}$$

The chromatic dispersion coefficient can then be written as

$$C_\lambda = M_\lambda + W_\lambda \tag{3.107}$$

where the material dispersion coefficient M_λ and the waveguide dispersion coefficient W_λ are given by

$$M_\lambda = \frac{1}{c}\frac{\mathrm{d}n_g(\lambda)}{\mathrm{d}\lambda}. \tag{3.108}$$

$$W_\lambda = -\frac{n_{1g} - n_{2g}}{c\lambda} \underbrace{V \frac{\mathrm{d}^2(VB)}{\mathrm{d}V^2}}_{\substack{\text{dispersion} \\ \text{factor}}} \tag{3.109}$$

Example (see Problem Set):

Compare two different slab waveguides with $n_2 = n_3 = 1.45$ and $n_1 = n_2 + \delta n$:

Waveguide 1: $\delta n = 0.005$; $2a = 4$ μm

Waveguide 2: $\delta n = 0.005 \times 10$; $2a = 4$ $\mu\mathrm{m}/\sqrt{10}$

For both waveguides, a certain real frequency f corresponds to the same normalized frequency:

$$V = ak_0 A_N = ak_0\sqrt{n_1^2 - n_2^2} \approx \frac{2\pi f}{c} \underbrace{a\sqrt{2n_2\delta n}}_{\text{identical for both waveguides}}$$

Since $\gamma = 0$ (symmetric waveguide), the eigenvalue equation for TE modes, Eq. (3.92), yields the same normalized propagation constant B for both waveguides, and since $n_1^2/n_2^2 \approx 1$, we obtain practically the same propagation constant for the TM modes, Eq. (3.95). However, the dispersion properties of the waveguides differ quite significantly, see Fig. 3.19. Both material dispersion and waveguide dispersion contribute to the overall chromatic dispersion, see Eqs. (3.107), (3.108), and (3.109). In our example, silica glass is used as a base material for both waveguides, and the material dispersion coefficient M_λ is negative for $\lambda < \lambda_0 \approx 1.3$ μm, where λ_0 denotes the zero material dispersion wavelength (normal group velocity dispersion), and positive for $\lambda > \lambda_0$. The waveguide dispersion W_λ is negative for all wavelengths in the considered range. For waveguide 2, its magnitude is approximately a factor of 10 larger than for waveguide 1. This shifts the

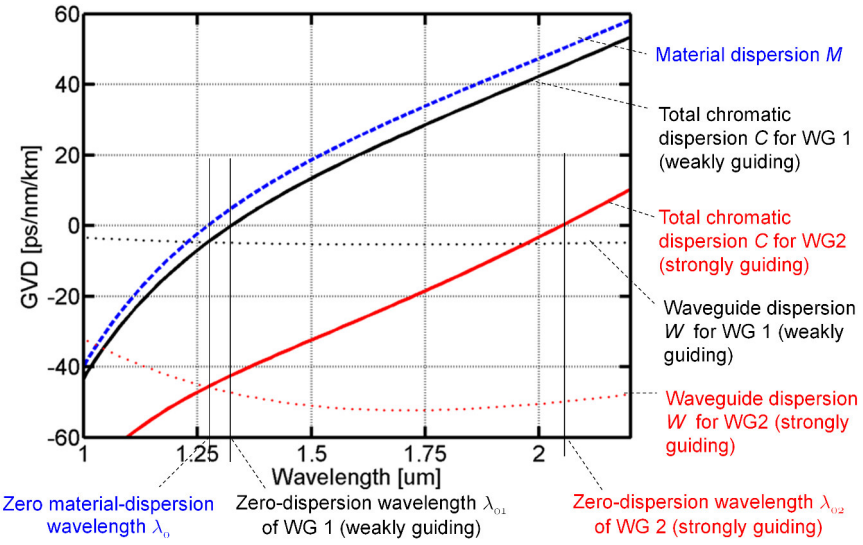

Figure 3.19: Interplay of material and waveguide dispersion. We consider two waveguides WG1 and WG2, which differ in both the refractive index contrast and the waveguide width in a way that a given real frequency ω translates into the same normalized frequency $V = a\frac{\omega}{c}\sqrt{n_1^2 - n_2^2}$. Both waveguides feature the same material dispersion M_λ. However, for the strongly guiding waveguide (higher index contrast, smaller waveguide width), the waveguide dispersion is much stronger than for the weakly guiding case. At a wavelength of 1.55 μm, this results in a strongly negative chromatic dispersion coefficient C_λ, which can be used to compensate the positive chromatic dispersion of a weakly guiding waveguide. This is the principle of dispersion compensation in fiber-optic communication links, see section 5.6.2.

zero-dispersion wavelength of the waveguide from approximately 1.3 μm to a value beyond 2 μm. For a fixed wavelength of $\lambda \approx 1.55$ μm, we find

$$C_\lambda\,(1.55\ \mu\text{m}) \approx \begin{cases} 18\ \frac{\text{ps}}{\text{km nm}} & \text{for waveguide 1} \\ -30\ \frac{\text{ps}}{\text{km nm}} & \text{for waveguide 2} \end{cases}.$$

By combining appropriate lengths of these two different waveguides, a net chromatic dispersion of zero can be achieved,

$$\Delta t_g = \Delta\lambda\,(C_{\lambda 1} L_1 + C_{\lambda 2} L_2) = 0 \qquad \text{for} \qquad C_{\lambda 1} L_1 + C_{\lambda 2} L_2 = 0.$$

This is the principle of dispersion compensation, which we will discuss again in the context of dispersion compensating fibers, see section 5.6.2.

3.5 Metal-dielectric waveguide structures and surface plasmons

In the previous sections, we have considered dieletric slab waveguides. Here, we will investigate optical waveguiding in structures that involve both dielectrics and metals. We will find that strongly localized modes can exist at a single interface between a metal and a dielectric provided that the optical frequency is smaller than the plasma frequency of the metal. Propagation of these modes involves density oscillations of the electron plasma just below the surface of the metal. The associated quasiparticles are called surface plasmon polaritons (SPP), or, in short, surface plasmons. Both the wavelength of SPP and the penetration depth to either side of the

TABLE 4.1 Refractive indices and extinction coefficients of selected metals and semiconductors

Material	λ (μm)	n	κ	$\varepsilon_r - j\varepsilon_i = (n - j\kappa)^2$	F
Au	0.633	0.17	3.0	−8.97−j1.02	
	0.653	0.166	3.15	−9.89−j1.05	
	1.55	0.550	11.5	−132−j12.6	
Ag	0.633	0.065	3.9	−15.2−j0.507	
	0.653	0.140	4.15	−17.2−j1.16	
	1.55	0.514	10.8	−116−j11.1	
Cu	0.633	0.14	3.15	−9.91−j0.88	
	0.653	0.214	3.67	−13.4−j1.57	
	1.55	0.606	8.26	−67.9−j10.0	
Al	0.633	1.2	7	−47.56−j16.8	
	0.653	1.49	7.82	−58.9−j23.3	
	1.55	1.44	16.0	−254−j46.1	
Cr	0.633	3.19	2.26	+5.07−j14.4	
	1.590	4.13	5.03	−8.24−j41.5	
Ge	0.633	4.5	1.7	+17.4−j15.3	
	0.653	5.294	0.638	+27.6−j6.76	
	1.55	4.275	0.00567	+18.3−j0.049	
GaAs	0.633	3.856	0.196	+14.8−j1.51	
	0.653	3.826	0.179	+14.6−j1.37	
	1.55	3.3737	—	+11.4	
Si	0.633	3.882	0.019	+15.07−j0.148	
	0.653	3.847	0.016	+15.0−j0.123	
	1.532	3.4784	—	+12.1	

Figure 3.20: Refractive indices and extinction coefficients of selected metals and semiconductors (Adapted from [5]).

interface can be much smaller than the vacuum wavelength of the corresponding photon. This can potentially enable ultracompact optical devices with subwavelength dimensions as well as highly sensitive optical detection schemes for single molecules adsorbed to the surface of a metal-dielectric interface. It is for these promises that plasmonics is currently a lively field of ongoing research.

3.5.1 Refractive indices and extinction coefficients of metals

For ideal metals, where damping of the plasma oscillations is neglected, the electric susceptibility is given by Eq. (2.49). The corresponding relative dielectric constant is then given by

$$\epsilon_r = 1 - \frac{\omega_p^2}{\omega^2} \tag{3.110}$$

where the plasma frequency of the metal can be estimated from Eq. (2.50). For real metals and semiconductors, however, the relative dielectric constant results from various contributions of both bound and free charges, and damping of plasma oscillations cannot be neglected. This leads to complex dielectric constant $\underline{\epsilon}_r$ and a complex refractive index $\underline{n}$ of the form

$$\underline{\epsilon}_r = \epsilon_r - \mathrm{j}\epsilon_{ri} \tag{3.111}$$

$$\underline{n} = n - \mathrm{j}n_i \tag{3.112}$$

The negative imaginary part n_i of the complex refractive index is often referred to as the extinction coefficient and denoted by κ. Refractive indices and extinction coefficients of selected metals and semiconductors are listed in Fig. 3.20 [5].

3.5.2 Metal-clad slab waveguides

Let us first consider a metal-clad slab waveguide where a high-index waveguide core is sandwiched between a metal top layer and low-index lower cladding, see Fig. 3.21. For calculating the modes,

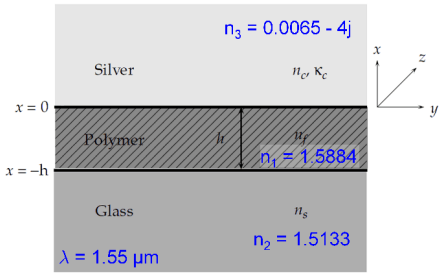

Figure 3.21: Metal-clad slab waveguide, consisting of a polymer waveguide core which is sandwiched between a silver top layer and low-index glass cladding. (Adapted from [5]).

we can use the same procedure as in Section 3.3.2, but with complex refractive indices $\underline{n}_1$, $\underline{n}_2$, and $\underline{n}_3$. This results in a complex propagation constant of the form

$$\underline{\beta} = \beta - \mathrm{j}\beta_i, \tag{3.113}$$

the imaginary part of which is related to the modal power attenuation constant α by

$$\alpha = 2\beta_i. \tag{3.114}$$

For a slab waveguide consisting of a glass substrate ($\underline{n}_2 = 1.5133$), a polymer core ($\underline{n}_1 = 1.5884$), and a silver cladding ($\underline{n}_3 = 0.0065 - 4\mathrm{j}$), the normalized propagation constant B according to Eq. (3.91) and the modal loss coefficient α/k_0 are plotted as a function of normalized frequency V in Fig. 3.22. The following facts may be observed from the plots:

- The $B-V-$curves have essentially the same shape as for dielectric slab waveguide, except for the TM_0-mode, for which the propagation constant is essentially independent of frequency.
- The loss of TM-modes is much larger than the loss of the corresponding TE-modes. For a given frequency, the attenuation increases with mode index.
- For all modes except TM_0, the loss coefficient α reaches a peak slightly above the cutoff frequency and then decreases monotonically as frequency increases, thereby changing by more than two orders of magnitude. In contrast to that, the attenuation of the TM_0-mode is much higher, but is essentially independent from frequency.

This behavior can be understood from looking at the mode profiles, see Fig. 3.23:

- For all modes except TM_0, the fields extend across the entire waveguide core and evanescently leak into the substrate. The interaction of the modes with the dielectric core and substrate is hence roughly comparable to the case of an asymmetric dielectric slab waveguide, and this results in similar dispersion curves.
- For the TE-modes, the electric field is entirely parallel to the (nearly) perfectly conducting metal surface and must hence (nearly) vanish. This leads to a very weak interaction of the mode with the metal top layer, and attenuation is hence weak.
- The TM-modes, in contrast, have a dielectric displacement component D_x (electric field component E_x) which is perpendicular to the metal surface, and which is hence continued by an electric current inside the metal. As a consequence, the TM mode fields penetrate deeper into the metal and hence experience considerable ohmic loss.
- Generally, for higher-order modes, a larger fraction of the mode field propagates in the metal top layer, and losses are therefore higher.

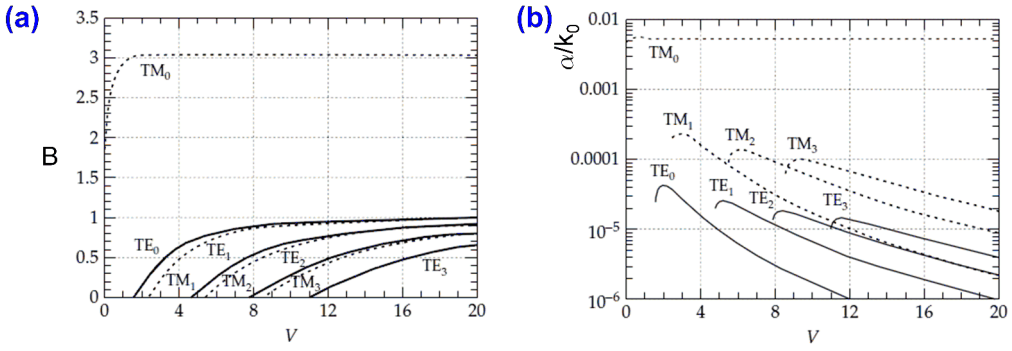

Figure 3.22: Modes of metal-clad slab waveguide: (a) The $B-V-$curves have essentially the same shape as for dielectric slab waveguide, except for the TM_0-mode, for which the propagation constant is essentially independent of frequency. (b) The attenuation of the TM_0-mode is much higher than that of all other waveguide modes, but essentially independent from frequency. (Adapted from [5]).

- For the special case of the TM_0-mode, the mode field is entirely confined to a thin layer near the metal-dielectric interface. The fraction of the mode field propagating in the dielectric substrate can be neglected, and the effective index seen by the mode is essentially frequency independent.

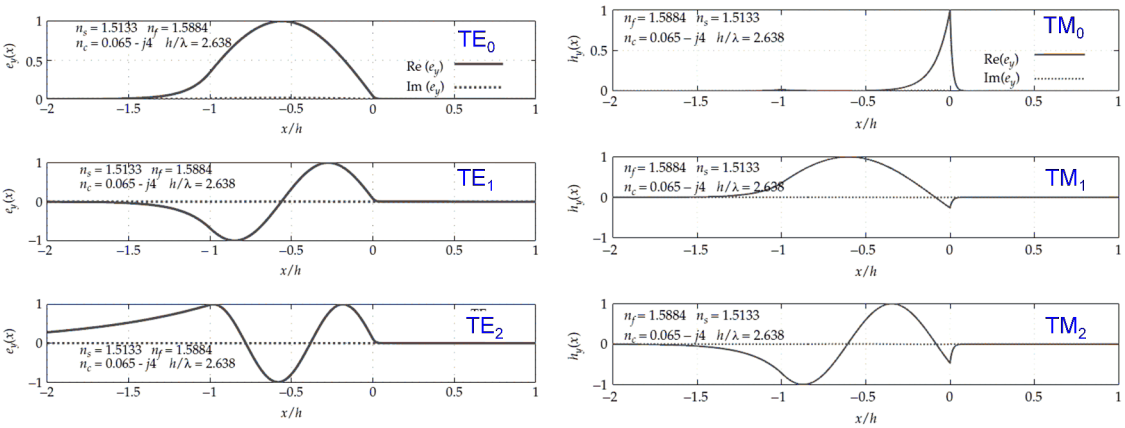

Figure 3.23: Mode fields of metal-clad slab waveguide: For all modes except TM_0, the fields extend across the entire waveguide core and evanescently leak into the substrate. For the special case of the TM_0-mode, however, the mode field is entirely confined to a thin layer near the metal-dielectric interface. The mode field does not interact with the dielectric substrate, and the effective index seen by the mode is essentially independent of frequency. (Adapted from [5]).

3.5.3 Surface plasmon polaritons

Derivation from Maxwell's Equations

Let us now focus on the TM_0 surface wave propagating along the boundary between a dielectric material with relative dielectric constant $\underline{\epsilon}_{rd}$, and a metal with relative dielectric constant $\underline{\epsilon}_{rm}$. This so-called surface plasmon polariton (SPP) wave consists of a transverse magnetic field of the

form

$$\underline{\mathbf{H}}(\mathbf{r},t) = \begin{pmatrix} 0 \\ \underline{\mathcal{H}}_y(x) \\ 0 \end{pmatrix} \exp\left(\mathrm{j}\left(\omega t - \beta z\right)\right), \tag{3.115}$$

where

$$\underline{\mathcal{H}}_y(x) = \begin{cases} H_0 e^{-k_{mx}^{(i)} x} & \text{for} \quad x > 0 \\ H_0 e^{k_{dx}^{(i)} x} & \text{for} \quad x < 0 \end{cases}. \tag{3.116}$$

The quantities $k_{mx}^{(i)}$ and $k_{dx}^{(i)}$ denote the negative and positive imaginary parts of the transverse wave vector components in the metal and the dielectric region. The corresponding electric mode field components are then given by

$$\underline{\mathcal{E}}_x(x) = \begin{cases} \frac{\beta \underline{\mathcal{H}}_y(x)}{\omega \epsilon_0 \underline{\epsilon}_{rm}} & \text{for} \quad x > 0 \\ \frac{\beta \underline{\mathcal{H}}_y(x)}{\omega \epsilon_0 \underline{\epsilon}_{rd}} & \text{for} \quad x < 0 \end{cases}, \tag{3.117}$$

$$\underline{\mathcal{E}}_z(x) = \begin{cases} \mathrm{j}\frac{k_{mx}^{(i)} \underline{\mathcal{H}}_y(x)}{\omega \epsilon_0 \underline{\epsilon}_{rm}} & \text{for} \quad x > 0 \\ -\mathrm{j}\frac{k_{dx}^{(i)} \underline{\mathcal{H}}_y(x)}{\omega \epsilon_0 \underline{\epsilon}_{rd}} & \text{for} \quad x < 0 \end{cases}. \tag{3.118}$$

The continuity of the $\underline{\mathcal{D}}_x$-component is inherently fulfilled by the ansatz[1]; the continuity of the E_z-component leads to

$$\frac{k_{mx}^{(i)}}{\underline{\epsilon}_{rm}} = -\frac{k_{dx}^{(i)}}{\underline{\epsilon}_{rd}} \tag{3.119}$$

For simplicity, let us assume real dielectric constants $\underline{\epsilon}_{rd} = \epsilon_{rd}$ and $\underline{\epsilon}_{rm} = \epsilon_{rm}$ for the moment. For a physically meaningful bound surface wave, we must require that the x-components of the wave vectors are purely imaginary ,

$$k_{mx}^{(i)} = \sqrt{\beta^2 - \epsilon_{rm} k_0^2}, \tag{3.120}$$

$$k_{dx}^{(i)} = \sqrt{\beta^2 - \epsilon_{rd} k_0^2}, \tag{3.121}$$

Inserting Eqs. (3.120) and (3.121) into Eq. (3.119), we obtain an analytic expression for the dispersion relation of the surface plasmon polariton,

$$\beta = k_0 \sqrt{\frac{\epsilon_{rm}\epsilon_{rd}}{\epsilon_{rm} + \epsilon_{rd}}} \tag{3.122}$$

The corresponding lateral decay constants are then given by

$$k_{mx}^{(i)} = k_0 \sqrt{\frac{-\epsilon_{rm}^2}{\epsilon_{rm} + \epsilon_{rd}}} \tag{3.123}$$

$$k_{dx}^{(i)} = k_0 \sqrt{\frac{-\epsilon_{rd}^2}{\epsilon_{rm} + \epsilon_{rd}}} \tag{3.124}$$

[1] Note that in many cases of classical electrodynamics, the surface-normal component of the dielectric displacement $\mathbf{D}$ is considered to be discontinuous at a metal-dielectric interface. In this consideration, the discontinuity of the normal $\mathbf{D}$-field component is associated with the surface charge density σ. In contrast to that, our consideration assumes a nonzero penetration depth of the fields into the metal. This penetration depth is typically a small fraction of the wavelength. The fields then obey the normal continuity rules for dielectric materials, but decay in the volume of the metal due to a nonzero volume charge density ρ.

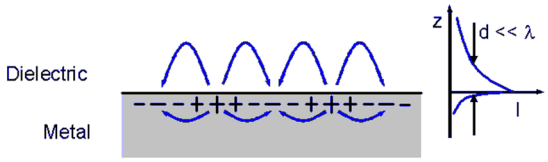

Figure 3.24: Electric field and charge distribution of a surface plasmon polariton (SPP).

A bound propagating solution exists if β, $k_{mx}^{(i)}$, and $k_{dx}^{(i)}$ are real. Assuming a positive ϵ_{rd}, this is only possible if

$$\epsilon_{rm} < -\epsilon_{rd} \tag{3.125}$$

Surface plasmon polariton formation requires at least one of the materials to have a sufficiently negative dielectric constant. This condition can be fulfilled by metals provided that the operation frequency is sufficiently far below the plasma frequency, see Eq. (3.110),

$$\omega < \frac{\omega_p}{\sqrt{1+\epsilon_{rd}}} \tag{3.126}$$

The corresponding mode fields are sketched in Fig. 3.24. The E_x-component changes its sign at $x = 0$ - this hints to regions of nonzero charge density, see sketch of the electric field lines and charge density distributions in Fig. 3.24.

Dispersion relation

Using the free-electron gas dispersion model for the metal, Eq. (3.110), we can sketch the dispersion relation of surface plasmon polaritons, see Fig. 3.25 (a). Note that dispersion relations in plasmonics are usually depicted by plotting the frequency ω (or the photon energy) as a function of propagation constant β or its magnitude $|\beta|$ in case of complex numbers. In such diagrams, bound modes lie below the so-called light line which is given by

$$\omega = c\beta\sqrt{\epsilon_{rd}} \tag{3.127}$$

and which corresponds to the dispersion relation of a plane wave propagating in the homogeneous dielectric medium of dielectric constant ϵ_{rd}. The dispersion relation of the plasmon can be subdivided in three regions:

- For $\omega < \omega_p/\sqrt{1+\epsilon_{r,d}}$ we find bound surface plasmons, for which the dispersion relation lies below the light line and which hence decays evanescently into the dielectric region.
- For $\omega_p/\sqrt{1+\epsilon_{r,d}} < \omega < \omega_p$, the propagation constant β is purely imaginary and decays evanescently along the direction of propagation
- For $\omega > \omega_p$, the wave can propagate both within the metal and within the dielectric. The associated modes do not decay in the transverse direction and are hence not bound to the surface.

When using real material data rather than the ideal electron gas model, the dispersion relation changes quantitatively without loosing its overall shape, see Fig. 3.25 (b).

Loss and penetration depth

In any cases of practical interest, $|\epsilon_{rm}| \gg \epsilon_{rd}$. Both the SPP propagation constant β and the lateral decay constants $k_{mx}^{(i)}$ and $k_{dx}^{(i)}$ are therefore much bigger than the associated vacuum wavenumber k_0, see Eqs. (3.120), (3.121), and (3.122). As a consequence, the effective wavelength $2\pi/\beta$ and

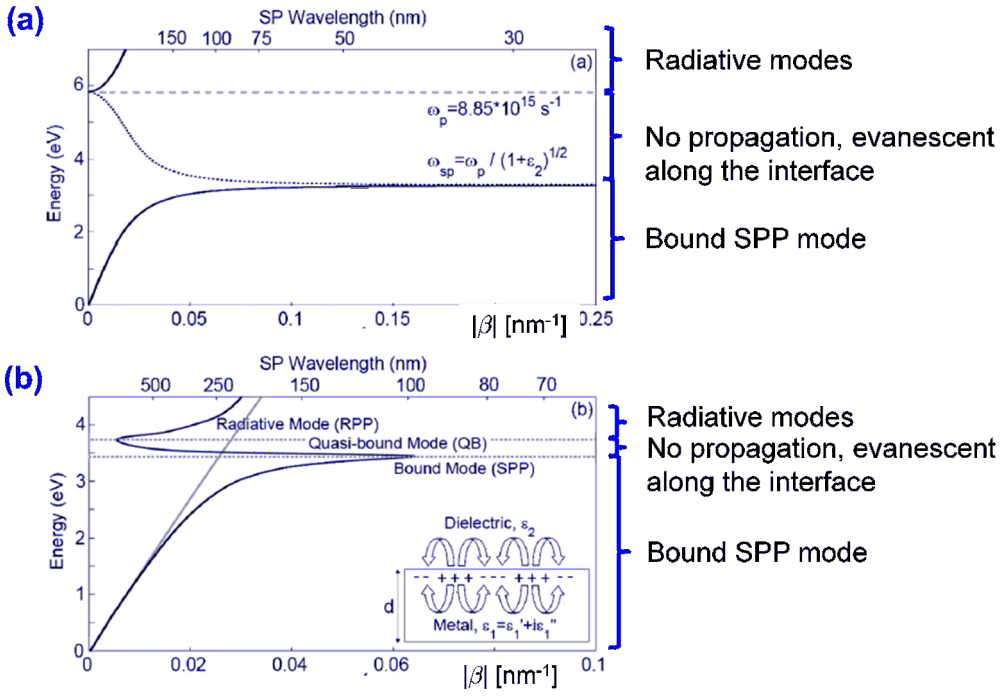

Figure 3.25: Dispersion relations of surface plasmon polaritons (SPP) for idealized and real material data; (a) Dispersion relation for an Ag/SiO_2 interface based on free-electron gas dispersion model. (b) Dispersion relation for an Ag/SiO_2 interface based on real material data for the complex refractive index of Ag. (Figures adapted from [5])

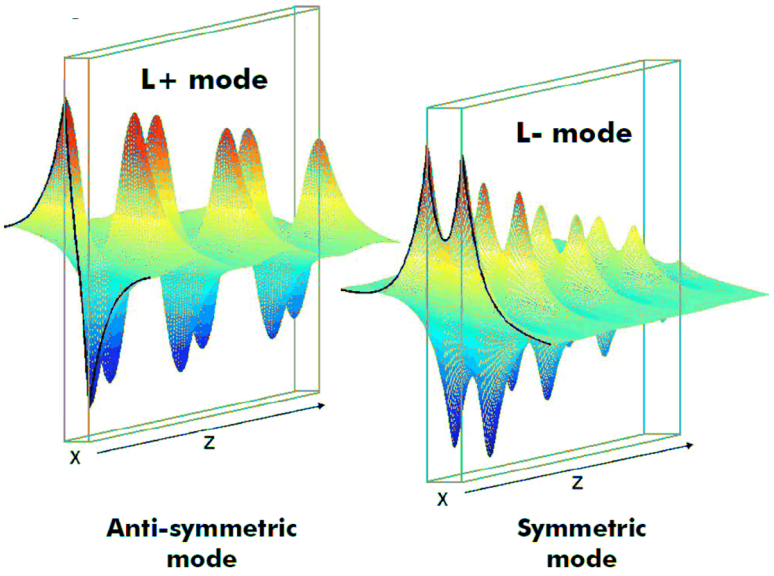

Figure 3.26: Electric field distribution of coupled surface plasmons supported by a thin metal film that is sandwiched between two identical dielectric layers. For the symmetric mode (L-), most of the light propagates within the metal film, leading to high losses. The antisymmetric mode has usually lower propagation loss, since a larger fraction of the mode field propagates outside the metal film. (Figures adapted from [8])

the lateral penetration depths $\delta_m = 1/k_{mx}^{(i)}$ and $\delta_d = 1/k_{dx}^{(i)}$ are much smaller than the vacuum wavelength λ. By exploiting the strong confinement of light, plasmonics structures can potentially enable ultra-compact photonic devices of sub-wavelength scales. Plasmonic devices are therefore a busy area of ongoing research.

One of the main drawbacks, however, is the fact that plasmonic waveguides still suffer from high losses. Assuming that the negative imaginary $\epsilon_{rm}^{(i)}$ part of the complex dielectric constant of metal $\underline{\epsilon}_{rm} = \epsilon_{rm} - \mathrm{j}\epsilon_{rm}^{(i)}$ is much smaller than the real part, $\epsilon_{rm}^{(i)} \ll \epsilon_{rm}$ we can calculate the real and the imaginary part of the plasmon's complex propagation constant $\underline{\beta} = \beta - \mathrm{j}\beta_i$ [8],

$$\beta = k_0 \sqrt{\frac{\epsilon_{rm}\epsilon_{rd}}{\epsilon_{rm} + \epsilon_{rd}}} \tag{3.128}$$

$$\beta_i = k_0 \left(\frac{\epsilon_{rm}\epsilon_{rd}}{\epsilon_{rm} + \epsilon_{rd}} \right)^{3/2} \left(\frac{\epsilon_{rm}^{(i)}}{2\epsilon_{rm}} \right) \tag{3.129}$$

Propagation distances $L_p = (2\beta_i)^{-1}$ are typically several tens of microns. Losses can be considerably decreased by using asymmetric coupled surface plasmons propagating along the surfaces of a thin metal film, see Fig. 3.26.

Chapter 4

Planar integrated waveguides

In the previous chapter, we have considered two-dimensional slab waveguides that are infinitely extended in the y- and the z-direction. Light was confined in the x-direction only, and it was possible to calculate the mode fields by a mathematically exact semi-analytical method[1]. We will now extend our analysis to the case of three-dimensional optical waveguides, where light is confined both in the vertical and in the horizontal direction. For these structures, only approximate semi-analytical solutions exist, and they can only be applied to a limited range of rather simple waveguide geometries. For many structures of practical interest, numerical modeling methods are therefore needed, the basics of which will be introduced as well. Nevertheless, approximated solutions remain very important for simplified design and modeling.

4.1 Basic structures of planar waveguides

Basic structures of integrated optical waveguides are depicted in Fig. 4.1. For a certain application, the waveguide structure has to be chosen based on the available material system, the propagation loss in the waveguide, the desired integration density or the connections to other waveguide devices or external fibers. The index difference $\Delta n = n_1 - n_2$ between the core and the high-index part of the cladding can range from a few percent for glass waveguides to values larger than unity, e.g., for the case of silicon-on-insulator structures. In the case of high index contrast, roughness of the core-cladding boundary leads to radiation and must be minimized by technological means.

Channel waveguides, Fig. 4.1 (a), consist of a rectangular core embedded in a uniform cladding material. These waveguides are characterized by the refractive index of the core and the cladding, n_1 and n_2, respectively. For a simplified mathematical analysis, we can exploit the symmetry of the waveguide in both the x and y directions. As for the symmetric slab waveguide, there is no cutoff of the fundamental modes if we assume that the cladding extends to infinity.

Strip waveguides, Fig. 4.1 (b), consist of a rectangular core structure of refractive index n_1 that is deposited directly on the substrate of a lower refractive index n_2 and covered by a medium of refractive index n_3. For large index differences $n_1 - n_3$, strip waveguides are prone to roughness-induced scattering loss. The cross section of the waveguide core does not have to be rectangular. Depending on the fabrication method, the sidewalls can, e.g., be inclined leading to a trapezoidal shape of the waveguide core. As for asymmetric slab waveguides, if $n_2 \neq n_3$, the fundamental modes run into cut-off for low frequencies.

Rib waveguides, Fig. 4.1 (c), comprise a film layer deposited on a substrate where the thickness of the film is increased along the waveguide. Interpreting the vertical layer stack as a slab

[1] The method is semianalytical in the sense that the mode fields could be formulated analytically whereas numerical techniques were needed to solve the resulting implicit equation for the propagation constant β.

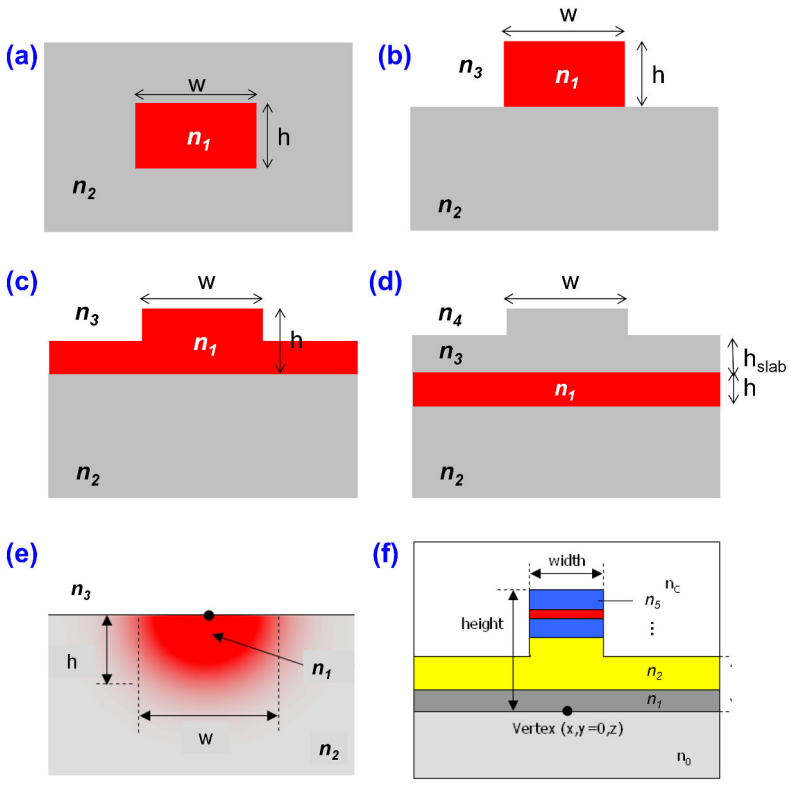

Figure 4.1: Basic structures of integrated optical waveguides; (a) Channel waveguide; (b) Strip waveguide; (c) Rib waveguide; (d) Ridge waveguide; (e) Diffused waveguide; (f) Multilayer waveguide.

waveguide, the increased thickness leads to an increase of the effective index in the region of the waveguide core and light is confined to this region. For large index differences $n_1 - n_3$ between the core and the top cladding, this structure is prone to roughness-induced scattering loss.

Ridge waveguides, Fig. 4.1 (d), consist of a continuous core layer of refractive index n_1 and a laterally structured top layer of lower refractive index n_3 which is in turn covered by the top cladding (refractive index n_4). As in rib waveguides, light is guided by a locally increased effective refractive index of slab waveguide that corresponds to the vertical layer stack. Since the rough sidewalls of the structured top layer are farther from the waveguide core, ridge waveguides are less prone to scattering loss.

Diffused waveguides, Fig. 4.1 (e), have a refractive index profile that gradually decreases from a high value n_1 in the center of the waveguide core to the lower refractive index n_2 of the substrate material. The index profile can be created by a diffusion process, where a dopant diffuses into a substrate material, thereby increasing the local refractive index. The index profile is approximated by analytical functions such as Gaussian or Error-Function (erf) distributions, where the width w and the height h correspond to the diffusion lengths in the horizontal and the vertical directions. The index difference $n_1 - n_2$ is usually rather low, and if air is used as a top cladding ($n_3 = 1$) light guidance in the lateral direction is rather weak.

Multilayer waveguides, Fig. 4.1 (f), can consist of complicated layer stacks (refractive indices n_1, $n_2 \ldots$) that are deposited on a substrate (refractive index n_0) and that confine light in

the vertical direction. Lateral guidance of light is achieved by partially etching the layer stack. Such structures are often used for active devices, where a double heterostructure *pn*-junction is integrated into the waveguide core. Depending on the etch depth, multilayer structures can be prone to scattering loss if rough etched sidewalls exist in a region of high field intensity.

4.2 Guided modes of a rectangular waveguide: The Marcatili method

In this section, we will consider a semi-analytical method that allows to approximately calculate guided modes of rectangular slab waveguides provided that the mode field is essentially confined to the waveguide core. This method has originally been proposed by Marcatili [21] and is therefore often referred to as the "Marcatili method".

Consider a rectangular channel waveguide as depicted in Fig. 4.2. In the following, we consider the case where the relative index difference between the core and the cladding region is small, i.e., $n_1/n_2 \approx 1$, and we assume that the electromagnetic field in the shaded areas can be neglected. This is true if the field is mostly confined to the core, i.e., if the waveguide is operated far from cut-off of the respective mode. We further assume that the guided fields can be separated in two modes, for each of which one of the transverse magnetic field components vanishes,

$$\begin{aligned} \underline{\mathcal{H}}_x = 0;\ \underline{\mathcal{H}}_y \text{ and } \underline{\mathcal{E}}_x \text{ dominate} \quad &\Rightarrow \quad \mathcal{E}_x - \text{mode} \\ \underline{\mathcal{H}}_y = 0;\ \underline{\mathcal{H}}_x \text{ and } \underline{\mathcal{E}}_y \text{ dominate} \quad &\Rightarrow \quad \mathcal{E}_y - \text{mode} \end{aligned}$$

We will analyze the $\mathcal{E}_x$-modes in more detail. For the $\mathcal{E}_y$-modes, the analysis can be derived in a similar way.

Basic equations for $\mathcal{E}_x$-modes

Using these assumptions, we can further simplify Maxwell's equations for waveguide modes, Eqs. (3.68) and (3.69). For $\mathcal{E}_x$-modes, $\underline{\mathcal{H}}_x = 0$ and we obtain

$$\frac{\partial \underline{\mathcal{E}}_z}{\partial y} + \mathrm{j}\,\beta \underline{\mathcal{E}}_y = 0 \tag{4.1}$$

$$-\frac{\partial \underline{\mathcal{E}}_z}{\partial x} - \mathrm{j}\,\beta \underline{\mathcal{E}}_x = -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_y \tag{4.2}$$

$$\frac{\partial \underline{\mathcal{E}}_y}{\partial x} - \frac{\partial \underline{\mathcal{E}}_x}{\partial y} = -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_z \tag{4.3}$$

and

$$\frac{\partial \underline{\mathcal{H}}_z}{\partial y} + \mathrm{j}\,\beta \underline{\mathcal{H}}_y = \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_x \tag{4.4}$$

$$-\frac{\partial \underline{\mathcal{H}}_z}{\partial x} = \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_y \tag{4.5}$$

$$\frac{\partial \underline{\mathcal{H}}_y}{\partial x} = \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_z \tag{4.6}$$

Using additionally the divergence equation for the magnetic field,

$$\nabla \cdot \mathbf{H} = 0 \tag{4.7}$$

$$\frac{\partial \underline{\mathcal{H}}_x}{\partial x} + \frac{\partial \underline{\mathcal{H}}_y}{\partial y} - \mathrm{j}\,\beta \underline{\mathcal{H}}_z = 0, \tag{4.8}$$

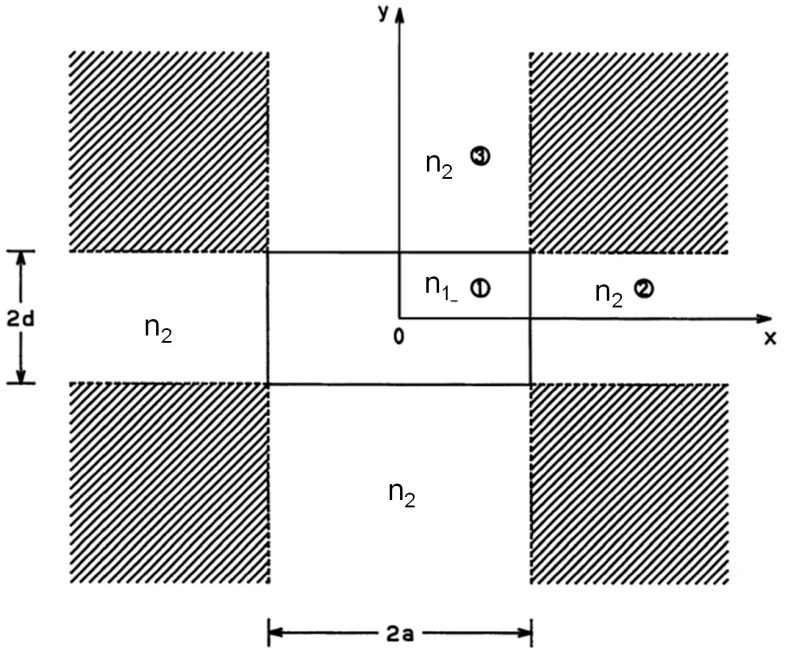

Figure 4.2: Simplified model of a rectangular channel waveguide: The relative index difference between the core and the cladding region is assumed to be small, i.e., $n_1/n_2 \approx 1$, and the electromagnetic field in the shaded areas is neglected. This approximation is valid if the field is mostly confined to the core, i.e., if the waveguide is operated far from cut-off of the respective mode (Figure adapted from [21]).

we can express all nonzero electromagnetic field components of the $\underline{\mathcal{E}}_x$-mode by $\underline{\mathcal{H}}_y$:

$$\underline{\mathcal{E}}_x = \frac{\omega\mu_0}{\beta}\underline{\mathcal{H}}_y - \frac{1}{\mathrm{j}\beta}\frac{\partial}{\partial x}\left(\frac{1}{\mathrm{j}\omega\epsilon_0 n^2}\frac{\partial \underline{\mathcal{H}}_y}{\partial x}\right) \tag{4.9}$$

$$\underline{\mathcal{E}}_y = \frac{1}{\omega\epsilon_0 n^2 \beta}\frac{\partial^2 \underline{\mathcal{H}}_y}{\partial x \partial y} \tag{4.10}$$

$$\underline{\mathcal{E}}_z = \frac{1}{\mathrm{j}\omega\epsilon_0 n^2}\frac{\partial \underline{\mathcal{H}}_y}{\partial x} \tag{4.11}$$

$$\underline{\mathcal{H}}_z = \frac{1}{\mathrm{j}\beta}\frac{\partial \underline{\mathcal{H}}_y}{\partial y} \tag{4.12}$$

Within the homogeneous core and cladding regions, the refractive index does not vary, $n(x,y) = n_1$ or $n(x,y) = n_2$, and can hence be taken out of the differential operator in Eq. (4.9). Moreover, the vectorial wave equation for the magnetic field, Eq. (2.25), can be separated into scalar equation for each vector component. Using the mode ansatz for the magnetic field, Eq. (3.60), we obtain the wave equation for the $\underline{\mathcal{H}}_y$-component,

$$\frac{\partial^2 \underline{\mathcal{H}}_y}{\partial x^2} + \frac{\partial^2 \underline{\mathcal{H}}_y}{\partial y^2} + \left(k_0^2 n^2 - \beta^2\right)\underline{\mathcal{H}}_y = 0 \tag{4.13}$$

Mode field ansatz for $\mathcal{E}_x$-modes

The further analysis proceeds in analogy to the slab waveguide: We will use an ansatz for the $\underline{\mathcal{H}}_y$-component ($\underline{\mathcal{H}}_x$-component) of the $\mathcal{E}_x$-mode ($\mathcal{E}_y$-mode) that fulfills the respective wave equation within the homogeneous core and cladding regions. In this ansatz, we use the propagation constant

β as a free parameter. From there, we will derive all other nonzero components of the mode fields. Enforcing the boundary conditions at the core-cladding interface, we obtain an implicit equation for β which can then be solved numerically.

The ansatz for the transverse magnetic field components represent guided electromagentic fields that are confined to the waveguide core and that decay exponentially in the cladding region while being continuous a the core-cladding interface. Since the waveguide is symmetric with respect to both the x- and the y-axis, it is sufficient to consider the first quadrant, i.e., the region $x > 0$, $y > 0$, see Fig. 4.2. For the $\mathcal{E}_x$-modes, the ansatz for the $\mathcal{H}_y$-field reads

$$\underline{\mathcal{H}}_y(x,y) = \begin{cases} A\cos(k_{1x}x - \Phi_x)\cos(k_{1y}y - \Phi_y) & \text{in region 1} \\ A\cos(k_{1x}a - \Phi_x)\exp\left(-k_{2x}^{(i)}(x-a)\right)\cos(k_{1y}y - \Phi_y) & \text{in region 2} \\ A\cos(k_{1x}x - \Phi_x)\cos(k_{1y}d - \Phi_y)\exp\left(-k_{3y}^{(i)}(y-d)\right) & \text{in region 3} \end{cases} \tag{4.14}$$

where the transverse wavenumbers and decay constants k_{1x}, k_{1y}, and $k_{2x}^{(i)}$, $k_{3y}^{(i)}$ are related to the propagation constant β by

$$-k_{1x}^2 - k_{1y}^2 - \beta^2 + n_1^2 k_0^2 = 0, \tag{4.15}$$

$$k_{2x}^{(i)\,2} - k_{1y}^2 - \beta^2 + n_2^2 k_0^2 = 0, \tag{4.16}$$

$$-k_{1x}^2 + k_{3y}^{(i)\,2} - \beta^2 + n_2^2 k_0^2 = 0. \tag{4.17}$$

Due to waveguide symmetry, the mode fields must be even or odd with respect to x and y, and the phases Φ_x and Φ_y are hence given by

$$\Phi_x = (p-1)\,\frac{\pi}{2} \qquad p = 1, 2\ldots \tag{4.18}$$

$$\Phi_y = (q-1)\,\frac{\pi}{2} \qquad q = 1, 2\ldots \tag{4.19}$$

Using the mode field ansatz of Eq. (4.14) together with Eqs. (4.9) – (4.12), we can estimate the order of magnitude of the various field components. Eliminating β from Eqs. (4.15) – (4.17), we find

$$k_{1x}^2 + k_{2x}^{(i)\,2} = \left(n_1^2 - n_2^2\right) k_0^2 \tag{4.20}$$

$$k_{1y}^2 + k_{3x}^{(i)\,2} = \left(n_1^2 - n_2^2\right) k_0^2 \tag{4.21}$$

According to our assumptions, the index difference between the core and the cladding region is small, i.e., $n_1^2 - n_2^2 \ll n_1^2$, and we obtain the following relations

$$k_{1x} \ll n_1 k_0, \qquad k_{1y} \ll n_1 k_0, \qquad \beta \approx n_1 k_0, \tag{4.22}$$

i.e. the mode field is predominantly propagating along the z-direction with relatively small transverse wavevector components. This is in full analogy to a weakly guiding slab waveguide, where plane waves must hit the core-cladding interface under a sufficiently large angle $\vartheta \approx 90°$ to maintain total internal reflection.

Inserting Eq. (4.14) into Eq. (4.9), we find that the magnitudes of the dominant components of the electromagnetic field are related by

$$|\underline{\mathcal{E}}_x| \approx \left|\frac{\omega\mu_0}{\beta}\right| |\underline{\mathcal{H}}_y| \tag{4.23}$$

Similarly, we can use Eq. (4.9) to show that the $\underline{\mathcal{E}}_y$-component is weaker than the $\underline{\mathcal{E}}_x$-component, by a factor of the order of δ^2,

$$|\underline{\mathcal{E}}_y| \approx \left|\frac{k_{1x}^2}{n_1^2 k_0^2}\right| |\underline{\mathcal{E}}_x| \approx O\left(\delta^2\right) |\underline{\mathcal{E}}_x| \tag{4.24}$$

where

$$\delta \approx \frac{k_{1x}}{n_1 k_0} \approx \frac{k_{1y}}{n_1 k_0} \ll 1. \tag{4.25}$$

The $\underline{\mathcal{E}}_y$-component can therefore be safely neglected when matching the boundary conditions at the core-cladding interfaces. In contrast to that, we find from Eqs. (4.11) and (4.12) that the longitudinal magnetic and electric field components are weaker than the dominant transverse components by only a factor of the order of δ,

$$|\underline{\mathcal{E}}_z| \approx \left|\frac{k_{1x}\beta}{n_1^2 k_0^2}\right| |\underline{\mathcal{H}}_y| \approx O(\delta)\ |\underline{\mathcal{E}}_x| \tag{4.26}$$

$$|\underline{\mathcal{H}}_z| \approx \left|\frac{k_{1y}}{\beta}\right| |\underline{\mathcal{H}}_y| \approx O(\delta)\ |\underline{\mathcal{H}}_y| \tag{4.27}$$

The continuity of the $\underline{\mathcal{E}}_z$- and the $\underline{\mathcal{H}}_z$-components at the core-cladding interface must hence be considered.

Boundary conditions and dispersion equation for $\mathcal{E}_x$-modes

The $\underline{\mathcal{E}}_z$-component in regions 1, 2 and 3 is obtained by inserting Eq. (4.14) in Eq. (4.11),

$$\underline{\mathcal{E}}_z(x,y) = \frac{A}{\mathrm{j}\,\omega\epsilon_0} \begin{cases} -\frac{k_{1x}}{n_1^2} \sin(k_{1x}x - \Phi_x) \cos(k_{1y}y - \Phi_y) & \text{in region 1} \\ -\frac{k_{2x}^{(i)}}{n_2^2} \cos(k_{1x}a - \Phi_x) \exp\left(-k_{2x}^{(i)}(x-a)\right) \cos(k_{1y}y - \Phi_y) & \text{in region 2} \\ -\frac{k_{1x}}{n_2^2} \sin(k_{1x}x - \Phi_x) \cos(k_{1y}d - \Phi_y) \exp\left(-k_{3y}^{(i)}(y-d)\right) & \text{in region 3} \end{cases}. \tag{4.28}$$

The $\underline{\mathcal{E}}_z$-component must be continuous both across the vertical ($x = a$) and the horizontal ($y = d$) core-cladding interface. The continuity condition at $x = a$ leads to the first dispersion equation for $\mathcal{E}_x$-modes

$$k_{1x}a = (p-1)\frac{\pi}{2} + \arctan\left(\frac{n_1^2\, k_{2x}^{(i)}}{n_2^2\, k_{1x}}\right) \tag{4.29}$$

The continuity of $\underline{\mathcal{E}}_z$ at $y = d$ cannot be fulfilled by the ansatz, since $n_1 \neq n_2$. However, for low index contrasts, $n_1 - n_2 \ll 1$, the associated error is negligible.

The $\underline{\mathcal{H}}_z$-component can be obtained from Eq. (4.12),

$$\underline{\mathcal{H}}_z(x,y) = \frac{A}{\mathrm{j}\,\beta} \begin{cases} -k_{1y} \cos(k_{1x}x - \Phi_x) \sin(k_{1y}y - \Phi_y) & \text{in region 1} \\ -k_{1y} \cos(k_{1x}a - \Phi_x) \exp\left(-k_{2x}^{(i)}(x-a)\right) \sin(k_{1y}y - \Phi_y) & \text{in region 2} \\ -k_{3y}^{(i)} \cos(k_{1x}x - \Phi_x) \cos(k_{1y}d - \Phi_y) \exp\left(-k_{3y}^{(i)}(y-d)\right) & \text{in region 3} \end{cases}. \tag{4.30}$$

For the $\underline{\mathcal{H}}_z$-component, the continuity at $x = a$ is inherently fulfilled, and the continuity condition at $y = d$ yields the second dispersion equation for $\mathcal{E}_x$-modes:

$$k_{1y}d = (q-1)\frac{\pi}{2} + \arctan\left(\frac{k_{3y}^{(i)}}{k_{1y}}\right) \tag{4.31}$$

From Eqs. (4.15), (4.16), and (4.17), we find

$$k_{2x}^{(i)\,2} = k_0^2\left(n_1^2 - n_2^2\right) - k_{1x}^2 \tag{4.32}$$

$$k_{3y}^{(i)\,2} = k_0^2\left(n_1^2 - n_2^2\right) - k_{1y}^2 \tag{4.33}$$

Inserting these relations into Eqs. (4.29) and (4.31), we obtain implicit equations for k_{1x} and k_{1y}. After numerical solution, the propagation constant β of the mode is obtained by

$$\beta^2 = k_0^2 n_1^2 - k_{1x}^2 - k_{1y}^2 \tag{4.34}$$

Analysis of $\mathcal{E}_y$-modes

A similar analysis can be performed for the$\mathcal{E}_y$-modes. Using again Eqs. (3.68) and (3.69) together with the assumption that $\underline{\mathcal{H}}_y = 0$ we can express all nonzero electromagnetic field components by $\underline{\mathcal{H}}_x$:

$$\underline{\mathcal{E}}_x = -\frac{1}{\omega\epsilon_0 n^2 \beta} \frac{\partial^2 \underline{\mathcal{H}}_x}{\partial x \partial y} \tag{4.35}$$

$$\underline{\mathcal{E}}_y = -\frac{\omega\mu_0}{\beta} \underline{\mathcal{H}}_x + \frac{1}{\mathrm{j}\beta} \frac{\partial}{\partial y} \left(\frac{1}{\mathrm{j}\omega\epsilon_0 n^2} \frac{\partial \underline{\mathcal{H}}_x}{\partial y} \right) \tag{4.36}$$

$$\underline{\mathcal{E}}_z = -\frac{1}{\mathrm{j}\omega\epsilon_0 n^2} \frac{\partial \underline{\mathcal{H}}_x}{\partial y} \tag{4.37}$$

$$\underline{\mathcal{H}}_z = \frac{1}{\mathrm{j}\beta} \frac{\partial \underline{\mathcal{H}}_x}{\partial x} \tag{4.38}$$

Within the homogeneous core and cladding regions, the wave equation for the $\underline{\mathcal{H}}_x$-component is given by

$$\frac{\partial^2 \underline{\mathcal{H}}_x}{\partial x^2} + \frac{\partial^2 \underline{\mathcal{H}}_x}{\partial y^2} + \left(k_0^2 n^2 - \beta^2\right) \underline{\mathcal{H}}_x = 0 \tag{4.39}$$

By using a field ansatz for $\underline{\mathcal{H}}_x$ in analogy to Eq. (4.14), and by matching again the boundary conditions of the $\underline{\mathcal{E}}_z$- and the $\underline{\mathcal{H}}_z$-component, we obtain the dispersion equation for $\mathcal{E}_y^{(p,q)}$-modes:

$$k_{1x} a = (p-1)\frac{\pi}{2} + \arctan\left(\frac{k_{2x}^{(i)}}{k_{1x}}\right) \tag{4.40}$$

$$k_{1y} d = (q-1)\frac{\pi}{2} + \arctan\left(\frac{n_1^2\, k_{3y}^{(i)}}{n_2^2\, k_{1y}}\right) \tag{4.41}$$

Classification of mode fields

The mode fields obtained from the Marcatili method can be classified as $\mathcal{E}_x^{(p,q)}$- and $\mathcal{E}_y^{(p,q)}$-modes. In this nomenclature, the subscript x or y denotes the dominant transverse component of the electric field ($\underline{\mathcal{E}}_x$ or $\underline{\mathcal{E}}_y$). The integer index $p = 1, 2\ldots$ ($q = 1, 2\ldots$) denotes the number of extrema that this field component has along the x-direction (y-direction), see Fig. 4.3 and 4.4. $\mathcal{E}_x^{(p,q)}$-modes are often also referred to as (quasi-)TE modes; $\mathcal{E}_y^{(p,q)}$-modes are often denoted as (quasi-)TM modes.

Limitations of the Marcatili Method

As stated above, the Marcatili Method is only a valid approximation if all of the following criteria are fulfilled:

- Rectangular cross section
- Index difference between the core an the cladding region is small, $n_1^2 - n_2^2 \ll n_1^2$
- The waveguide is operated far from the cut-off frequency of the respective mode and the field is confined to the waveguide core. This minimizes errors that arise from the fact that the field in the corner regions has been ignored.

In all other cases, dedicated analytical or numerical methods have to be used to calculate the waveguide modes.

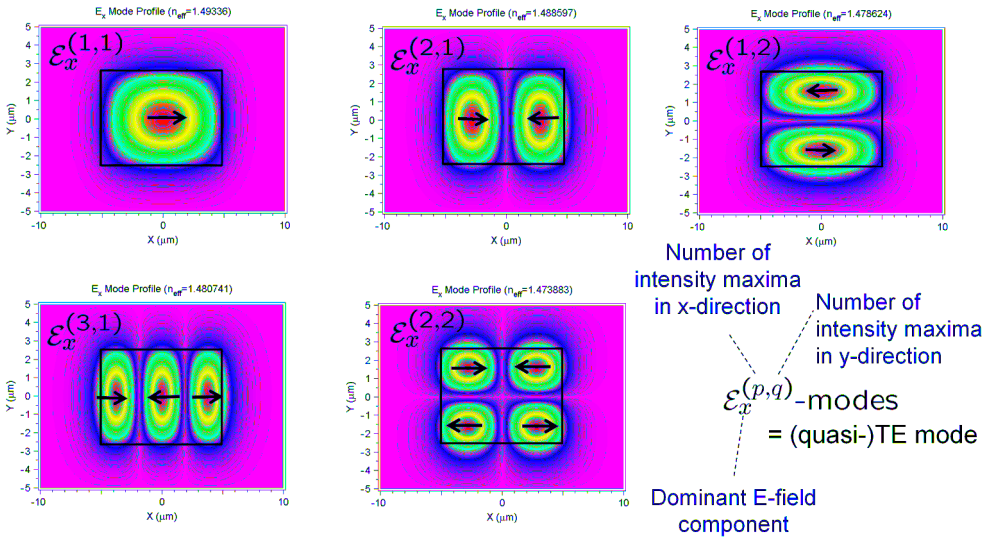

Figure 4.3: $\mathcal{E}_x^{(p,q)}$-mode fields of rectangular channel waveguides. The colors correspond to the optical intensity. Arrows indicate the relative orientation of the dominant transverse electric field component ($\underline{\mathcal{E}}_x$). The integer index $p = 1, 2\ldots$ ($q = 1, 2\ldots$) denotes the number of extrema that this field component has along the x-direction (y-direction).

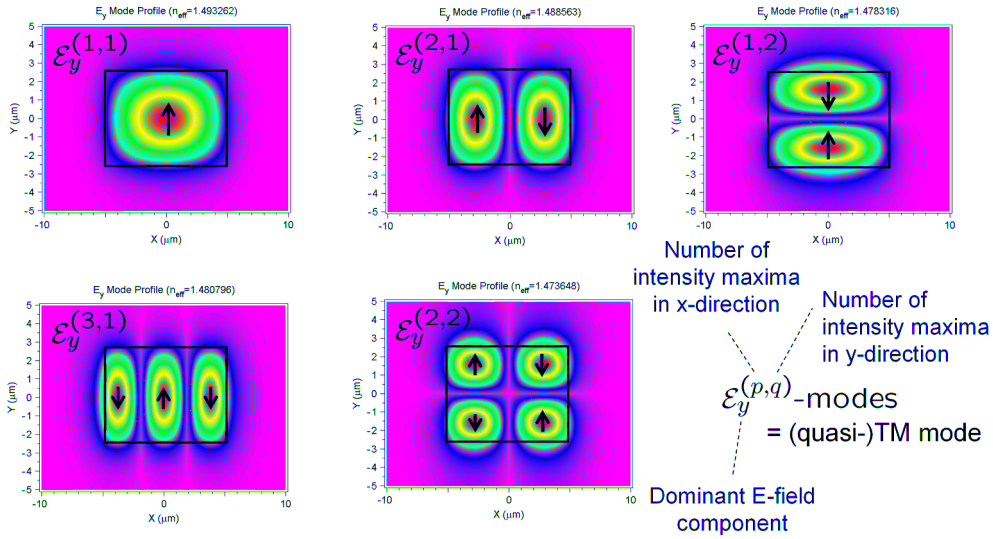

Figure 4.4: $\mathcal{E}_y^{(p,q)}$-mode fields of rectangular channel waveguides. The colors correspond to the optical intensity, and arrows indicate the relative orientation of the dominant transverse electric field component ($\underline{\mathcal{E}}_y$). The integer index $p = 1, 2\ldots$ ($q = 1, 2\ldots$) denotes again the number of extrema that this field component has along the x-direction (y-direction).

Slab waveguide interpretation of the Marcatili Method

The Marcatili method can be interpreted as a replacement of the three-dimensional (3D) channel waveguide by two "equivalent" two-dimensional (2D) slab structures, which are defined by the waveguide's layer sequence in the horizontal and the vertical direction, see Fig. 4.5. An $\mathcal{E}_x^{(p,q)}$-mode then translates into a TM mode for the equivalent vertical slab waveguide (slab thickness $2a$) and into a TE mode for the equivalent horizontal slab waveguide (slab thickness $2d$). The dispersion equations (4.40) and (4.41) correspond to the eigenvalue equations for the two slab waveguides. Equivalent considerations can be applied to the $\mathcal{E}_y^{(p,q)}$-modes, Fig. 4.6. In the Marcatili method, a unified ansatz is used to simultaneously fulfill the boundary conditions of both slab waveguides. We will later see that this leads to an underestimation of the effective refractive index.

Kumar's method

Marcatili's method can be extended by taking into account the mode fields in the corner regions of the waveguide. Such a method was proposed by Kumar et al. [19], see Ref. [24, 25] for more details.

4.3 The effective-index method

The basic idea of the Marcatili method is to replace a 3D waveguide by two mutually independent 2D slab waveguides whose parameters come directly from the 3D structure. In the course of the solution, a unified mode ansatz is used to fulfill the boundary conditions of both waveguides. In a similar way, the effective-index method approximates a 3D waveguide by two related slab structures. However, only the parameters of the first slab waveguide are taken directly from the 3D structure, whereas the second slab waveguide is defined based on the mode solution for the first one, i.e., the two structures are solved individually. We will later see that the effective index method slightly overestimates the effective refractive index of a 3D waveguide structure, whereas the Marcatili method tends to underestimate it.

Basic idea

To understand the basic principle of the effective-index method let us consider the waveguide depicted in Fig. 4.7. To apply the effective-index method, two basic assumptions must be fulfilled:

- The medium is only weakly inhomogeneous, i.e., the refractive index contrast is small ($n_1 - n_2 \ll n_1$, $n_1 - n_3 \ll n_1$), and the scalar wave equations (2.24) and (2.25) (Helmholtz equations) can be used.
- The horizontal waveguide dimensions are significantly larger than the vertical dimensions ($w \gg h_1$, $w \gg h_2$), i.e., we are considering waveguide structures that are "flat and wide".

Under these assumptions we obtain a good approximation of the 3D waveguide structure if we split the mode calculation in two steps: In a first step, we calculate the y-dependence of the fields by modeling the different regions (slab 1, slab 2, slab 3) as individual, infinitely extended extended slab waveguides. For each position x, this yields a y-dependent mode field $F(x, y)$ and an effective index $n_{es}(x)$ of the local slab waveguide stack. In the example depicted in Fig. 4.7 (a), we obtain refractive indices $n_{\mathrm{es}2}$, $n_{\mathrm{es}1}$, and $n_{\mathrm{es}3}$ for the core region (slab 2) and for the adjacent regions (slab 1, slab 3), respectively. In a second step, the resulting effective indices are used to define a 2D slab waveguide with vertically oriented core and cladding layers, see Fig. 4.7 (b), which can again be solved with the standard semianalytical methods introduced in Section 3.3. This method can be generalized to a wide range of low index-contrast planar 3D waveguide geometries, including structures which are nonuniform in the direction of propagation. By calculating the effective index for the vertical layer stack at each position (x, z), the 3D index profile $n(x, y, z)$ is the reduced to an equivalent 2D profile $n_{es}(x, z)$, and the resulting 2D structure can be analyzed with analytical or simplified numerical methods.

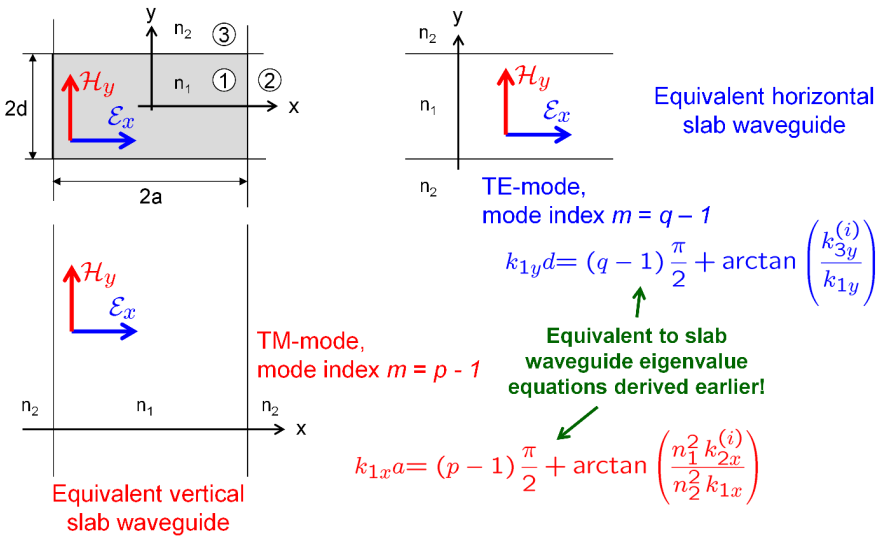

Figure 4.5: Slab-waveguide interpretation of an $\mathcal{E}_x^{(p,q)}$-mode. The mode field can be considered as a TM mode for the equivalent vertical slab waveguide (slab thickness $2a$) and as a TE mode for the equivalent horizontal slab waveguide (slab thickness $2d$). The dispersion equations (4.29) and (4.31) correspond to the eigenvalue equations for the two slab waveguides.

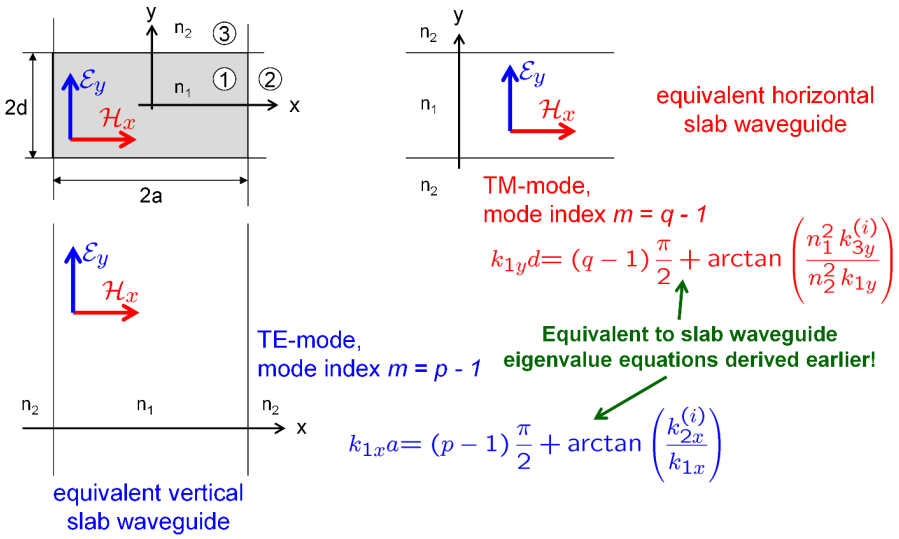

Figure 4.6: Slab-waveguide interpretation of an $\mathcal{E}_y^{(p,q)}$-mode. The mode field can be considered as a TE mode for the equivalent vertical slab waveguide (slab thickness $2a$) and as a TM mode for the equivalent horizontal slab waveguide (slab thickness $2d$). The dispersion equations (4.40) and (4.41) correspond to the eigenvalue equations for the two slab waveguides.

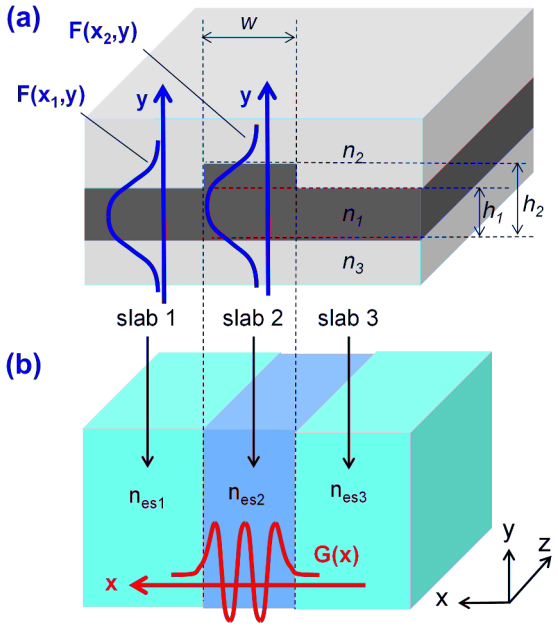

Figure 4.7: (a) Typical waveguide structure that can be approximated by using the effective-index method. The refractive index contrast is assumed to be small ($n_1 - n_2 \ll n_1$, $n_1 - n_3 \ll n_1$), such that the scalar wave equation for weakly inhomogeneous media can be used. In addition, the lateral waveguide dimensions are assumed much larger than the vertical dimensions ($w \gg h_1$, $w \gg h_2$). First, the mode fields $F(x_1, y)$ and $F(x_2, y)$ are calculated together with the effective indices n_{es1}, n_{es2}, and n_{es2} within and outside the core region, respectively. (b) The result is then used to define a 2D slab waveguide with vertically oriented core and cladding layers and refractive indices n_{es1}, n_{es2}, and n_{es3}. Both slab waveguides can be solved with standard semianalytical methods.

Mathematical derivation

The intuitive approach described in the previous section can be formulated mathematically by using a separation ansatz for the wave equation. The derivation starts from the wave equation for weakly inhomogeneous media, Eqs. (2.24) and (2.25). By inserting the mode ansatz, Eqs. (3.59) and (3.60), we obtain the wave equation for the mode field,

$$\frac{\partial^2 \underline{\Psi}(x,y)}{\partial x^2} + \frac{\partial^2 \underline{\Psi}(x,y)}{\partial y^2} + \left(k_0^2 n^2(x,y) - \beta^2\right) \underline{\Psi}(x,y) = 0, \tag{4.42}$$

where $\underline{\Psi}(x,y)$ can be any component of the mode fields $\underline{\mathcal{E}}(x,y)$ and $\underline{\mathcal{H}}(x,y)$. The basic assumption of the effective index method is that the rapidly varying x-dependence of the mode fields can be factored out, i.e.,

$$\underline{\Psi}(x,y) = F(x,y)\, G(x), \tag{4.43}$$

Making use of the fact that the horizontal waveguide dimensions are significantly larger than the vertical dimensions, we can assume that $F(x,y)$ is only a slowly varying function of x, whereas $G(x)$ dominates the x-dependence of $\underline{\Psi}(x,y)$ and of its derivatives,

$$\frac{\partial^2 \underline{\Psi}(x,y)}{\partial x^2} \approx F(x,y)\, \frac{\partial^2 G(x)}{\partial x^2}. \tag{4.44}$$

Inserting Eq. (4.43) in Eq. (4.42) and using Eq. (4.44), we obtain

$$F(x,y)\,\frac{\partial^2 G(x)}{\partial x^2} + G(x)\,\frac{\partial^2 F(x,y)}{\partial y^2} + \left(k_0^2 n^2(x,y) - \beta^2\right) F(x,y)\, G(x) = 0 \tag{4.45}$$

$$\underbrace{\frac{1}{G(x)}\frac{\partial^2 G(x)}{\partial x^2} - \beta^2}_{\text{purely } x\text{-dependent}} + \underbrace{\frac{1}{F(x,y)}\frac{\partial^2 F(x,y)}{\partial y^2} + k_0^2 n^2(x,y)}_{y-\text{dependence dominates, only weakly } x\text{-dependent}} = 0 \tag{4.46}$$

Eq. (4.46) must be fulfilled for all values of y. This is only possible if the right part on the left-hand side does not depend on y which allows us to separate the two expressions. We introduce the effective index $n_{es}(x)$ of the horizontal slab waveguide structure as an x-dependent separation variable,

$$\frac{1}{F(x,y)}\frac{\partial^2 F(x,y)}{\partial y^2} + k_0^2 n^2(x,y) = k_0^2 n_{es}^2(x) \tag{4.47}$$

$$\frac{1}{G(x)}\frac{\partial^2 G(x)}{\partial x^2} - \beta^2 = -k_0^2 n_{es}^2(x) \tag{4.48}$$

Eq. (4.46) has hence been separated in two relations for the expression $F(x,y)$ and $G(x)$,

$$\frac{\partial^2 F(x,y)}{\partial y^2} + \left(k_0^2 n^2(x,y) - k_0^2 n_{es}^2(x)\right) F(x,y) = 0 \tag{4.49}$$

$$\frac{\partial^2 G(x)}{\partial x^2} + \left(k_0^2 n_{es}^2(x) - \beta^2\right) G(x) = 0 \tag{4.50}$$

For a given x, Eq. (4.49) corresponds to the wave equation for a horizontal slab waveguide. By solving the associated eigenvalue problem at this specific x-position, we obtain the local y-dependence $F(x,y)$ of the electromagnetic field together with the effective refractive index $n_{es}(x)$ of the local slab structure. By using $n_{es}(x)$ in Eq. (4.50), the wave equation for the equivalent vertical slab waveguide is obtained and can be solved for the eigenfunction $G(x)$ and the propagation constant β. The full spatial dependence of the field is then given by Eq. (4.43).

Examples: E_x-modes of rib and channel waveguides

As an example, let us consider the E_x-modes of low index-contrast rib waveguide, see Fig. 4.8. For the E_x-modes, the dominant electric field component is oriented horizontally, i.e., parallel to the dielectric interfaces at the top and at the bottom of the waveguide. This corresponds to a TE-polarized mode field within the horizontal slab waveguides. In a first step, we calculate the effective indices of the slab sections[2],

$$n_{es}(x) = \begin{cases} 1.4809 & \text{for } -1.5\,\mu\text{m} < x < 1.5\,\mu\text{m} \\ 1.4633 & \text{elsewhere} \end{cases}$$

These values define the related vertical slab waveguide, for which a dominant E_x-component now corresponds to a TM-mode. This leads to an effective refractive index of $n_e = 1.4737$, which differs quite significantly from the numerical result of $\bar{n}_e = 1.4846$ that was obtained from a finite-element mode solver [12].

For a channel waveguide, it is impossible to define a vertical layer stack outside the core region, see Fig. 4.9. In this case, the refractive index of the cladding material may be directly used for the region outside the waveguide core to define the vertically oriented equivalent slab waveguide. The obtained effective refractive index still differs from the result of a numerical calculation, but the deviation is smaller. If exact numbers are needed, findings obtained from approximate effective index solutions should always be cross-checked with numerical tools.

[2] A particularly useful mode solver for slab waveguides can be found at http://www.computational-photonics.eu/oms.html.

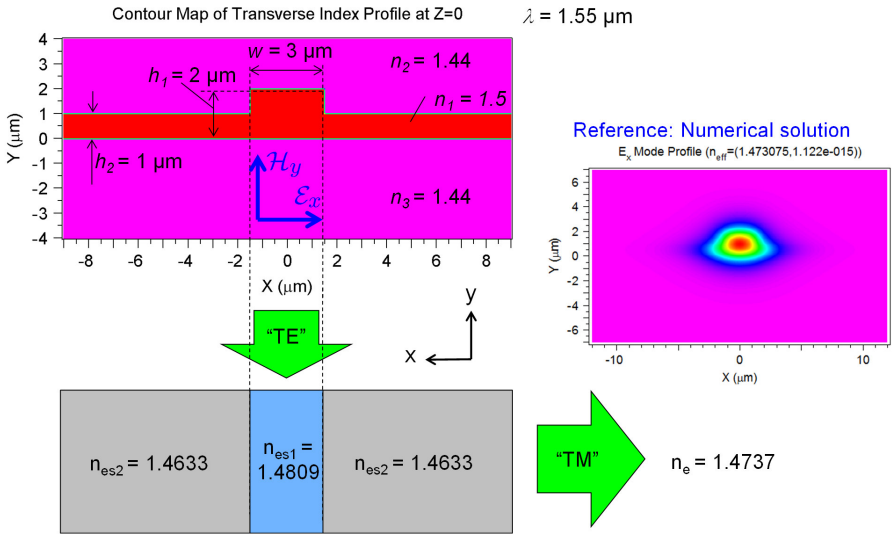

Figure 4.8: Calculation of E_x-modes of a rib waveguide using the effective-index method. For the E_x-modes, the dominant electric field component is oriented horizontally, i.e., parallel to the dielectric interfaces at the top and at the bottom of the waveguide. This corresponds to a TE-polarized mode field within the horizontal slab waveguides, leading to effective indices of 1.4809 (1.4633) within (outside) the core region. These values define the related vertical slab waveguide, for which a dominant E_x-component now corresponds to a TM-mode of effective refractive index $n_e = 1.4737$, which differs quite significantly from the numerical result of $\bar{n}_e = 1.4846$.

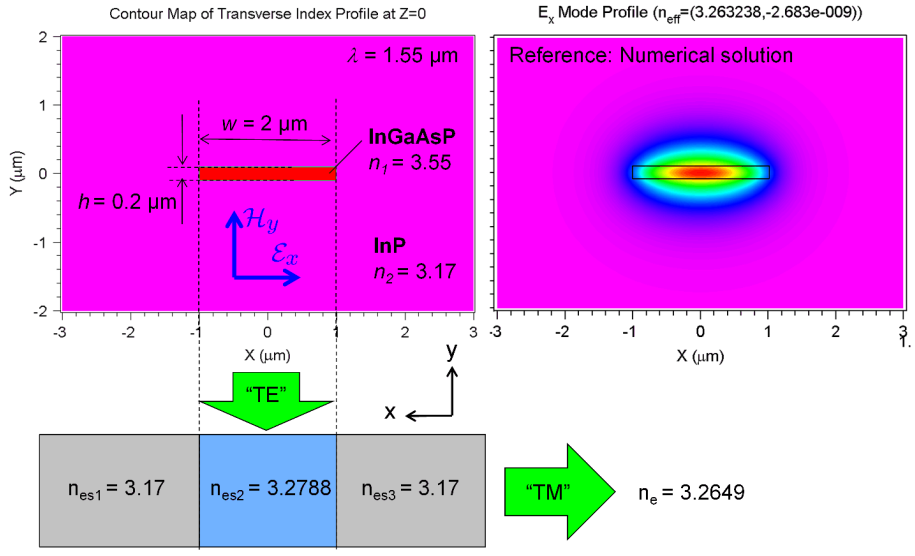

Figure 4.9: Calculation of E_x-modes of a channel waveguide using the effective-index method. For the region outside the waveguide core, the refractive index of the cladding material is directly used to define the vertically oriented equivalent slab waveguide. The dominant electric field component is oriented horizontally, i.e., parallel to the dielectric interfaces at the top and at the bottom of the waveguide, which corresponds to a TE-polarized mode field within the horizontal slab waveguides, leading to effective indices of 3.2788 within the core region. These values define the related vertical slab waveguide, for which a dominant E_x-component now corresponds to a TM-mode of effective refractive index $n_e = 3.2649$, which still differs a bit from the numerically calculated value of $\bar{n}_e = 3.2632$.

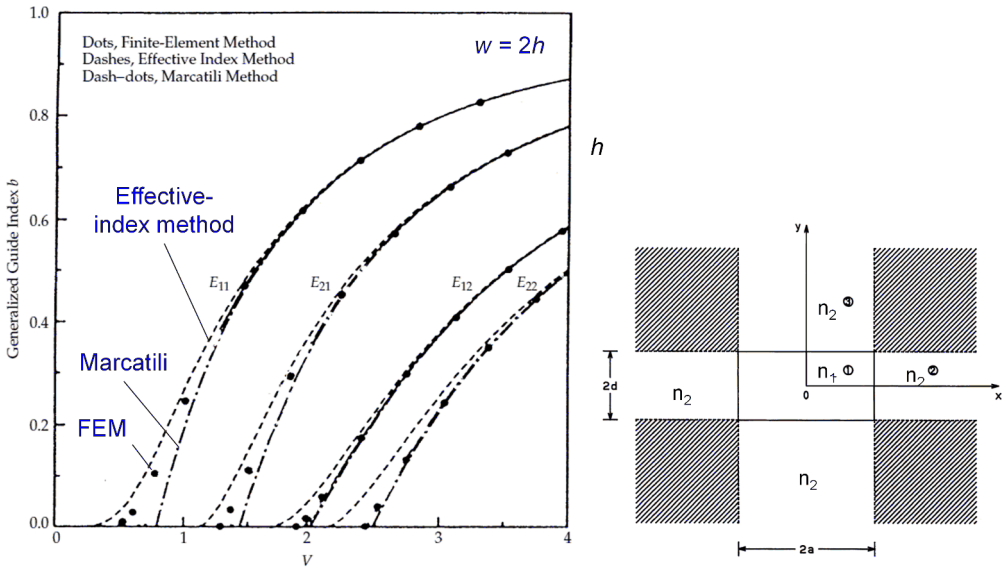

Figure 4.10: Comparison of effective refractive indices obtained from different approximation methods for the case of a rectangular channel waveguide with $w = 2h$. The effective parameters are defined by $V = k_0 h\sqrt{n_1^2 - n_2^2}$ and $B = \frac{\beta^2 - k_0^2 n_2^2}{k_0^2 n_1^2 - k_0^2 n_2^2}$. Both the effective index method and the Marcatili method give accurate results for modes far above cutoff, where the mode fields are well confined to the waveguide core. Near cutoff, the effective index method tends to overestimate the propagation constant of the mode field, whereas the Marcatili underestimates it. (Figure adapted from [5]).

Accuracy of approximation methods

A comparison of effective refractive indices obtained from different approximation methods is depicted in Fig. (4.10) for the case of a rectangular channel waveguide with $w = 2h$ [5]. In this analysis, the effective parameters are defined by $V = k_0 h\sqrt{n_1^2 - n_2^2}$ and $B = \frac{\beta^2 - k_0^2 n_2^2}{k_0^2 n_1^2 - k_0^2 n_2^2}$. It is found that both the effective index method and the Marcatili method give accurate results for modes far above cutoff, where the mode fields are well confined to the waveguide core. Near cutoff, however, these methods have to be used with care. As stated earlier, the effective index method tends to overestimate the propagation constant of the mode field, whereas the Marcatili underestimates it. A deeper discussion of the error sources can be found in Ref. [5].

4.4 Numerical methods for mode calculation

4.4.1 General principles

For high-index contrast waveguides or complicated waveguide cross sections, analytical approximations cannot be used. In these cases, numerical methods are needed to calculate the guided mode fields. In this section we will introduce the basic principle of finite-difference and finite-element mode solvers. For simplicity, we assume weakly inhomogeneous media for which the scalar Helmholtz equation can be used. Advanced mode solvers, that account for the vectorial nature of the mode fields, rely on the same basic principles.

The general procedure of numerical mode solvers is depicted in Fig. 4.11. Irrespective of the exact implementation, numerical solution procedures rely on the discretization of the refractive index profile within a finite computational domain ("meshing"). The computational domain has to be chosen sufficiently large to make sure that the electromagnetic fields are close to zero near the

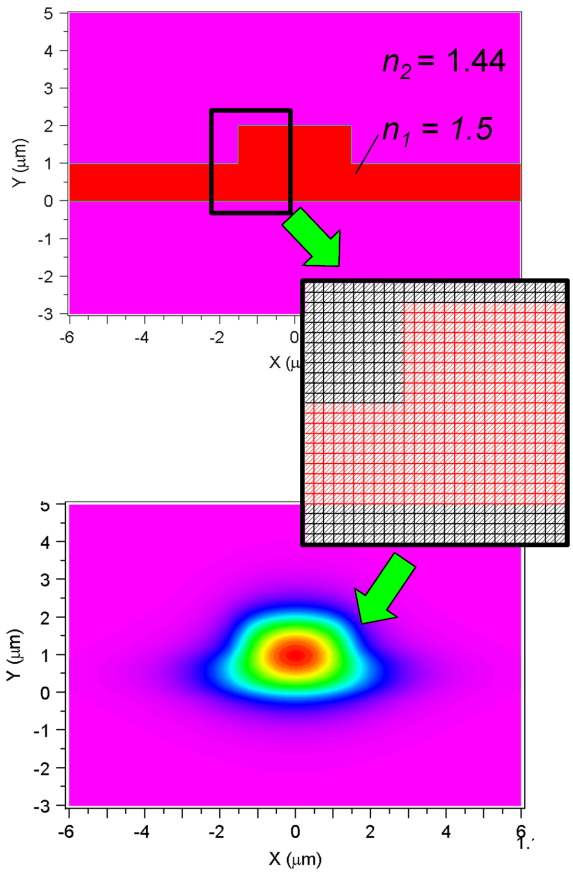

Figure 4.11: General procedure of numerical mode solvers: The refractive index profile is discretized within a finite computational domain ("meshing") and the electromagnetic field quantities are represented by a discrete set of field values at the respective mesh points. Maxwell's equations or the corresponding wave equations can be then approximated by a set of linear equations for the unknown field quantities which can be solved numerically.

domain boundary. Within this domain, the electromagnetic field quantities are represented by a discrete set of field values at the respective mesh points. Maxwell's equations or the corresponding wave equations can be then approximated by a set of linear equations for the unknown field quantities which can be solved numerically. Numerous effects can lead to errors in numerical schemes, the most common of which will be shortly discussed in the following.

Numerical mode solvers are contained in most commercial software packages for electromagnetic field calculations. For photonic applications, widely used software packages are, e.g., the following:

- BEAMProp by Rsoft Inc., Ossining, NY, http://www.rsoftdesign.com
- Microwave Studio by CST, Darmstadt, http://www.cst.de
- FIMMwave by Photon Design, Oxford, UK, http://www.photond.com/
- HFSS by Ansoft, Pittsburgh http://www.ansoft.com

4.4.2 Finite difference method

In finite difference methods, the refractive index profile is sampled at discrete grid points. Depending on the meshing algorithm, the grid points will or will not be equidistant in space, Fig. 4.12.

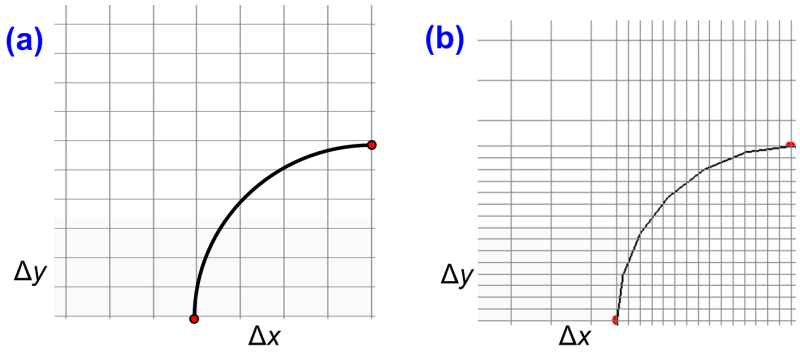

Figure 4.12: In finite difference methods, both the refractive index profile and the electromagnetic fields are sampled at discrete mesh points. (a) Uniform grid: Δx and Δy do not vary throughout the computational domain. (b) Nonuniform grid: Δx and Δy are locally adapted to the structure.

The field equations are then discretized by replacing the derivatives of a function $f(x)$ by their finite difference representation. For a uniform discretization along x with step size $\Delta x = x_{p+1} - x_p$, the first and the second derivative of a function $f(x)$ is given by

$$f'(x_p) \approx \frac{f(x_{p+1}) - f(x_{p-1})}{\Delta x} \tag{4.51}$$

$$f''(x_p) \approx \frac{f(x_{p+1}) + f(x_{p-1}) - 2f(x_p)}{\Delta x^2} \tag{4.52}$$

In this way the set of partial differential equations is replaced by a linear set of equations which can be solved using standard algebraic methods. This shall now be demonstrated by formulating a finite-difference version of the scalar wave equation.

Scalar mode solver

To derive the fundamental relations of a scalar mode solver, we insert the mode ansatz, Eqs. (3.59) and (3.60) into the wave equations for weakly inhomogeneous media, Eqs. (2.24) and (2.25). This yields

$$\frac{\partial^2 \underline{\Psi}(x,y)}{\partial x^2} + \frac{\partial^2 \underline{\Psi}(x,y)}{\partial y^2} + \left(k_0^2 n^2(x,y) - \beta^2\right) \underline{\Psi}(x,y) = 0 \tag{4.53}$$

where $\underline{\Psi}(x,y)$ can again be any component of the mode fields $\underline{\mathcal{E}}(x,y)$ and $\underline{\mathcal{H}}(x,y)$.
To discretize the computational domain, we assume an equidistant mesh with grid points $(x_p, y_q)^T$, where

$$\Delta x = x_{p+1} - x_p, \tag{4.54}$$

$$\Delta y = y_{q+1} - y_q, \tag{4.55}$$

see Fig. 4.13. Both the refractive index profile $n(x,y)$ and the field quantity $\underline{\Psi}(x,y)$ are represented by the respective values in the discrete grid points,

$$n_{p,q} = n(x_p, y_q), \tag{4.56}$$

$$\underline{\Psi}_{p,q} = \underline{\Psi}(x_p, y_q), \tag{4.57}$$

where $p = 1 \ldots P$ and $q = 1 \ldots Q$. The second-order derivatives in Eq. (4.53) are approximated

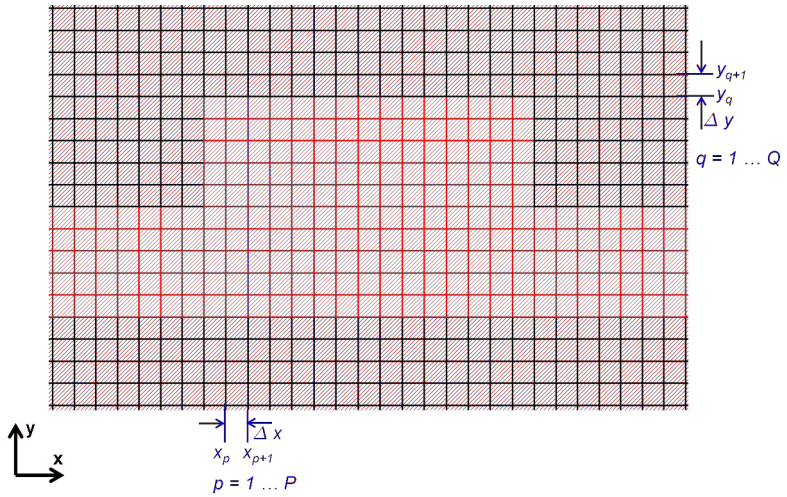

Figure 4.13: Finite-difference representation of a rib waveguide. In this example, a uniform grid is used, i.e., the grid points $(x_p, y_q)^T$, $p = 1 \ldots P$ and $q = 1 \ldots Q$, are equidistant.

by finite differences,

$$\left.\frac{\partial^2 \underline{\Psi}(x,y)}{\partial x^2}\right|_{x=x_p,\, y=y_q} = \frac{\underline{\Psi}_{p+1,q} + \underline{\Psi}_{p-1,q} - 2\underline{\Psi}_{p,q}}{\Delta x^2}, \tag{4.58}$$

$$\left.\frac{\partial^2 \underline{\Psi}(x,y)}{\partial y^2}\right|_{x=x_p,\, y=y_q} = \frac{\underline{\Psi}_{p,q+1} + \underline{\Psi}_{p,q-1} - 2\underline{\Psi}_{p,q}}{\Delta y^2}. \tag{4.59}$$

Eq. (4.53) can hence be rewritten as

$$\frac{\underline{\Psi}_{p+1,q} + \underline{\Psi}_{p-1,q} - 2\underline{\Psi}_{p,q}}{\Delta x^2} + \frac{\underline{\Psi}_{p,q+1} + \underline{\Psi}_{p,q-1} - 2\underline{\Psi}_{p,q}}{\Delta y^2} + k_0^2 n_{p,q}^2 \underline{\Psi}_{p,q} = \beta^2 \underline{\Psi}_{p,q}. \tag{4.60}$$

Re-writing Eq. (4.60) for each node $(x_p, y_q)^T$, we obtain a matrix eigenvalue equation for the field values $\underline{\Psi}_{p,q}$ and the propagation constant β,

$$\mathbf{A}\widehat{\underline{\Psi}} = \beta^2 \widehat{\underline{\Psi}}, \tag{4.61}$$

where $\widehat{\underline{\Psi}} = \left(\underline{\Psi}_{1,1} \ldots \underline{\Psi}_{P,Q}\right)$ is a $(PQ, 1)$-vector that contains all field values in the mesh points, and $\mathbf{A}$ is a sparse (PQ, PQ)-matrix that represents the coupling between these field values. Equation (4.61) can be solved numerically to obtain the propagation constant β and the field distribution $\underline{\Psi}_{p,q}$.

This procedure can also be extended to vectorial mode solvers. The vector $\widehat{\underline{\Psi}}$ then contains different field components, the coupling of which is given by a still sparse, but larger matrix $\mathbf{A}$.

Boundary conditions and termination of computational domain

In any numerical solver, the electromagnetic field is only represented within the finite computational domain. Problems arise when the finite difference equations are applied at the boundary points $p = 1$, $p = P$, $q = 1$ or $q = Q$, since field values outside the computational domain ($p = 0$, $p = P + 1$, $q = 0$ or $q = Q+1$) are required. Two main methods are commonly used to overcome these problems:

Perfectly matched layers (PML) A straightforward solution is to set the field values outside the computational domain to zero. This can be interpreted as putting a perfect metallic conductor (perfect mirror) at the domain boundary. To prevent perturbations from

back-reflection of radiation into the computational domain, a layer of absorbing material is artificially added inside the domain boundaries. The wave impedance of this layer must be matched to the impedance of the computational domain to prevent reflection from the surface of the absorbing material. However, adjusting the parameters of the absorber to minimize reflection is commonly cumbersome, and artificial reflections persist in many cases. In addition, PML often require to significantly extend the computational domain, thereby increasing computation time. For these reasons, PML are rarely used in conjunction with finite difference methods.

Transparent boundary conditions (TBC) Field values outside the computational domain can also be guessed in a way that back-reflections from the domain boundaries are minimized. To this end, it is usually assumed that the field behaves like an outgoing plane wave whose characteristics (amplitude, phase) are estimated by a heuristic algorithm. By using the plane wave assumption, the field outside the domain boundary can be related to the interior points, thus completing the set of equations. TBC are generally very effective in allowing radiation to freely escape the computational domain, without the need to extend the domain. Nevertheless, there are problems for which they do not perform well. In these cases, PML can still be an attractive alternative.

In any case, numerical convergence studies must be performed to ensure that reflection of light by the boundary of the computational domain does not impair the numerical results.

4.4.3 Finite-element method

In the finite-element method (FEM), the dielectric structure is split into basic geometrical elements (e.g., triangles or tetrahedra), so-called finite elements, within which the refractive index is assumed to be constant, see Fig. 4.14. The Finite Element Method (FEM) does not solve the wave equation directly, but a related variational problem that is constructed from the operator of the differential equation. For the two-dimensional scalar wave equation, Eq. (4.53), the corresponding functional is given by [16]

$$\begin{aligned} I =& \frac{1}{2} \iint_\Omega \left[\left(\frac{\partial \underline{\Psi}(x,y)}{\partial x} \right)^2 + \left(\frac{\partial \underline{\Psi}(x,y)}{\partial y} \right)^2 + \left(k_0^2 n^2(x,y) - \beta^2 \right) \underline{\Psi}^2(x,y) \right] \mathrm{d}x\, \mathrm{d}y \\ & - \int_{\partial\Omega} \left[\underline{\Psi}(x,y) \frac{\partial \underline{\Psi}(x,y)}{\partial n} \right] \mathrm{d}s \end{aligned} \tag{4.62}$$

where Ω denotes the computational domain. Solving the wave equation is equivalent to minimizing this integral. To this end, the integral is first expressed as a sum over the finite elements, which are labeled by a subscript $p = 1 \ldots P$. In each of the elements, the electromagnetic field is expanded into a set of $q = 1 \ldots Q$ basis functions $\underline{\Psi}_{pq}(x,y)$,

$$\underline{\Psi}_p(x,y) = \sum_q c_{pq}\, \underline{\Psi}_{pq}(x,y)\,, \tag{4.63}$$

where c_{pq} denote the expansion coefficients. The functional can then be written as a quadratic matrix equation,

$$I = \frac{1}{2} \mathbf{c}^{\mathbf{T}} \left(\mathbf{A} - \left(k_0^2 n^2 - \beta^2 \right) \mathbf{B} \right) \mathbf{c} \tag{4.64}$$

where $\mathbf{c}$ denotes a $(PQ, 1)$-vector that contains all expansion coefficients c_{pq}, and where $\mathbf{A}$ and $\mathbf{B}$ are (PQ, PQ)-matrices that are obtained by integrating the basis functions $\underline{\Psi}_{pq}(x,y)$ and their derivatives over the various finite elements. A necessary condition for the functional I to assume a minimal value is stationarity, i.e., the local gradient must vanish,

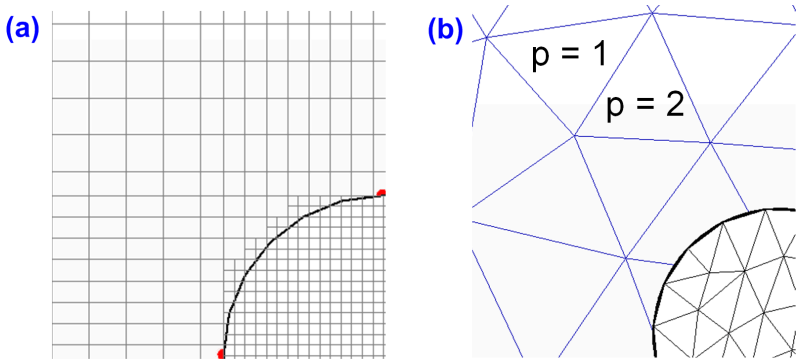

Figure 4.14: In finite-element methods, the computational domain is split into basic geometrical elements, so-called finite elements, within which the refractive index is assumed to be constant. (a) Finite-element mesh with rectangular geometry elements. (b) Finite-element mesh with triangular geometry elements.

$$\nabla_{\mathbf{c}}^T I = \left(\mathbf{A} - \left(k_0^2 n^2 - \beta^2\right) \mathbf{B}\right) \mathbf{c} = 0. \tag{4.65}$$

This corresponds to an eigenvalue matrix equation that can be solved numerically to obtain the expansion coefficients c_{pq} and the mode propagation constant β,

$$\mathbf{B^{-1}Ac} = \left(k_0^2 n^2 - \beta^2\right) \mathbf{c}. \tag{4.66}$$

Using Eq. (4.63), the electromagnetic field can now be reconstructed from the expansion coefficients c_{pq}.

4.4.4 Sources of errors in numerical mode solvers

There are several sources of errors in numerical solvers which should be carefully checked:

Model errors are due to basic assumptions in the underlying algorithm that may or may not apply to the modeled structure. As an example, methods that have been developed for low-index contrast fibers cannot necessarily be applied to high-index contrast integrated waveguides, and scalar methods should only be used in weakly guiding waveguides. It is therefore essential to check the intrinsic limitations of a numerical solver before applying it to a specific problem.

Discretization errors are caused by the approximation of the refractive index profile by discrete grid points and by the finite-difference approximations of the derivatives or the finite-element approximation of an integral expression. For any simulation, a convergence study should be performed, in which the discretization steps are reduced continuously until the results do not change any more.

Finite computational domain: Errors can also occur if the computational domain is too small for the problem, i.e., if the electromagnetic field interacts with the domain boundaries. This can be avoided by a convergence study in which the computational domain is continuously extended until the results do not change any more.

Note that there is a fundamental difference between the errors due to an approximate method like the effective index method and a numerical method like the finite difference method: With exact numerical methods we can start from Maxwell's equations and the error can always be reduced by, e.g., refining the discretization parameters or by extending the computational domain. This

is of course at the expense of larger calculation times. In approximate methods like the effective-index methods, the equations are solved rigorously, but the equations are only approximations of Maxwell's equations. It is generally important to critically check the results of numerical methods by, e.g., looking carefully at the field plots and doing plausibility tests.

4.5 Waveguide technologies and fabrication methods

Integrated optical waveguides can be realized in various material systems, each of which has specific strengths and weaknesses. The advantages, drawbacks and some important material systems are discussed in the following sections.

4.5.1 Glass waveguides

Glass waveguides are fabricated on large-area glass substrates or oxidized silicon wafers. The refractive index contrasts are rather small (usually $\Delta n < 0.1$). Therefore confinement of the optical wave is weak, which leads to large waveguide cross sections and requires large bending radii to minimize radiation losses. This results in large devices, but also allows for low-loss coupling to standard single-mode fibers, which also have low index contrasts and hence large mode diameters. Glass waveguides are currently the preferred technology for high-quality passive devices such as optical filters and arrayed waveguide gratings. Currently two different fabrication methods are commonly used: Diffusive ion exchange or deposition and structuring of the core material. These methods shall be shortly discussed in the following.

Ion exchange waveguides

In special types of glasses, the refractive index can be increased by exchanging the ions that are usually embedded into the glass matrix. Substituting, e.g., Na by Ag within the speciality glass IAG4[3] leads to an increase of the refractive index by up to 0.1. Diffusion can be purely thermal or assisted by electric fields that drive the ions into the bulk substrate material. This results in waveguiding channels just below the surface or deeply within the substrate, see Fig. 4.15. In such structures, the refractive index varies continuously across the waveguide core.

Deposition and etching, Silicon optical bench (SOB)

The second technique is based on the deposition of glass layers by means of chemical vapour deposition (CVD) or flame hydrolysis (FHD) on a substrate. In many cases, a silicon wafer is used as substrate material, and the integration platform is then often referred to as silicon optical bench (SOB) [13]. Waveguide cores are then structured by lithography and dry etching processes and overclad by additional low-index glass layers. Amorphous silica is often used as base material for the core and the cladding layers, and the refractive index of the core layer is increased by doping with phosphorous (P) or germanium (Ge). Depending on the doping material, the relative index difference $\Delta = \left(n_1^2 - n_2^2\right)/2n_1^2$ ranges from a fraction of a percent to a few percent. The relatively large waveguide cross-sections allow easy coupling with standard single-mode fibers, having a core diameter of approximately 9 µm. A schematic cross section and a fabricated SOB waveguide are depicted in Fig. 4.16.

4.5.2 Lithium niobate waveguides

Lithium niobate ($LiNbO_3$) is an anisotropic crystal which features strong electro-optical and acousto-optical effects and is commonly used to realize efficient optical switches and broadband electro-optic modulators. Waveguides on $LiNbO_3$ substrates are fabricated using diffusion processes, where two methods are commonly used, see Fig. 4.17.

[3] IAG4 consists of SiO_2, Al_2O_3, B_2O_3, Na_2O, TiO_2, Na^+, and F^-

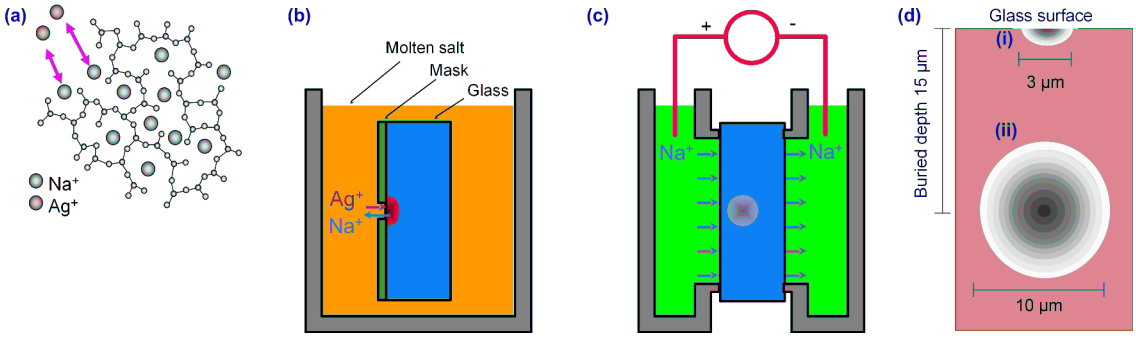

Figure 4.15: Glass waveguide fabrication based on ion exchange; (a) The refractive index of certain glasses can be increased by exchanging the ions that are usually embedded into the glass matrix. Substituting Na^+ by Ag^+ within the speciality glass IAG4 leads to an increase of the refractive index by up to 0.1. (b) Ag^+ ions can be moved into the glass matrix by thermal indiffusion from the surface. The lateral shape of the waveguide is defined by a mask. The refractive index varies continuously across the waveguide core. (c) To produce buried waveguides, field-assisted ion exchange is used in a second step. In this process, the ions near drift under the influence of an externally applied electric field. Diffusion continues during drift such that the cross section of the waveguide increases as it penetrates the glass substrate (d) Cross-section of an ion-exchange waveguide (i) near the surface and (ii) in the volume of the substrate (Figure source: Leoni fiber optics)

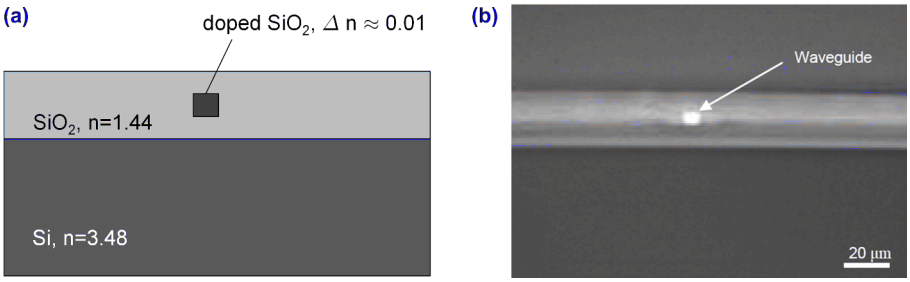

Figure 4.16: Glass waveguides fabricated by deposition and etching of doped glass layers. (a) Schematic cross section of a glass waveguide of doped SiO_2 in a silicon base wafer ("Silicon optical bench, SOB"); The relative index difference $\Delta = \left(n_1^2 - n_2^2\right)/2n_1^2 \approx \Delta n/n$ ranges typically between 0.3 % and 1.5 %. (b) Cross sectional view of a fabricated SOB waveguide. (Picture source ?)

Ti-indiffused waveguides

For indiffusion of Ti, the metal is deposited on the surface of the wafer and patterned by lithography and etching. A subsequent thermal treatment at around $1000 - 1050°C$ for several hours leads to indiffusion of Ti into the $LiNbO_3$ substrate, thereby causing a refractive index increase of around $\Delta n \approx 0.002 \ldots 0.01$. Typical waveguide widths are $10\,\mu$m. Due to the anisotropy of the Lithium niobate crystal, the behavior of resulting integrated optical components shows a large polarization dependence.

Proton-exchanged waveguides

For proton exchange, an Al-mask is patterned on the surface of the substrate, and the wafer is then immersed in boiling benzoic acid ($200 - 240°C$) for 2-4 hours. This this leads to a substitution of Li^+- by H^+-ions. For the so-called extra-ordinary polarization, the resulting $Li_{1-x}H_xNbO_3$ has a refractive index which is higher than that of the surrounding $LiNbO_3$ by$\Delta n \approx 0.02 \ldots 0.05$, whereas for the ordinary polarization, the index is decreased, $\Delta n \approx -0.04$.

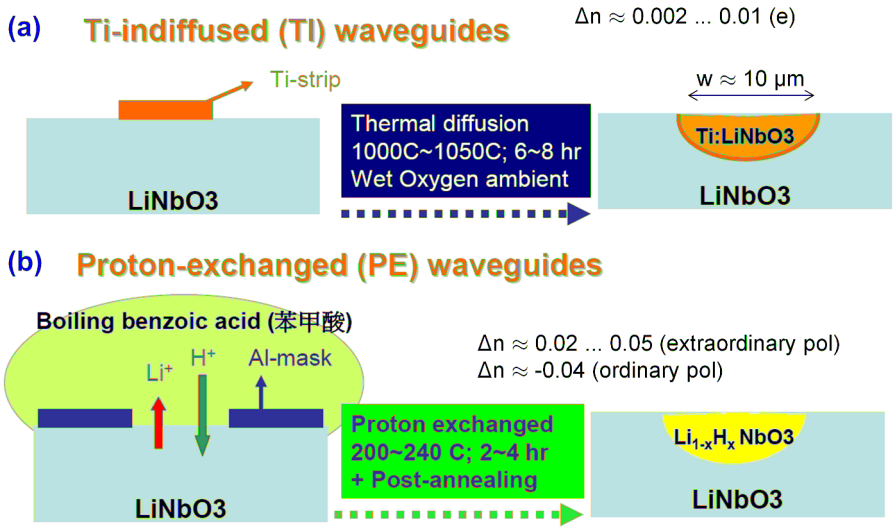

Figure 4.17: Fabrication of lithium niobate waveguides: (a) Indiffusion of titanium (Ti) under high temperatures; (b) "Proton-exchange", substituting Li^+ by H^+-ions. This leads to a refractive index increase for one polarization (extraordinary) and to a decrease for the other polarization (ordinary). (Picture Source ?)

4.5.3 Polymer waveguides

Polymers represent a broad set of materials. Polycarbonate (also used for compact disks) and PMMA are the preferred materials for purely passive applications. Other types of polymers show functional properties such as, e.g., a large electro-optical or non-linear coefficient, but this is mostly at the expense of a reduced long term stability.

Optical losses are relatively low around 800 nm, but increase considerably at infrared telecommunication wavelengths between 1300 nm and 1500 nm. This is due to overtones of C-H-bond oscillations. Losses can be decreased if hydrogen H is substituted by deuterium (heavy hydrogen), chlorine or fluorine since the larger atom mass of these materials shifts oscillations further into the IR, see Fig. 4.18 (a). The fundamental oscillation of the C-H-bond lies near 3390 nm, for deuterium it is 4484 nm and for Fluorine, the oscillation can be shifted to 8000 nm. Typical polymers such as poly(methyl-methacrylate) (PMMA) or polycarbonate (PC) exhibit losses of typically several dB per meter, which increase significantly for longer wavelengths, see Fig. 4.18 (b) and (c).

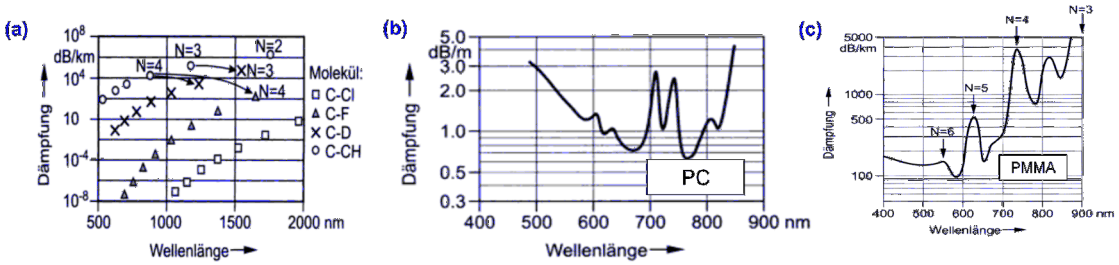

Figure 4.18: Losses of polymer materials; (a) Attenuation contributions from different overtones (overtone order N) of C-H-, C-D-, C-F-, and C-Cl-bonds (D = deuterium, Dämpfung = attenuation; Wellenlänge = wavelength; Molekül = molecule). The larger the atom mass of the involved elements, the further the oscillations are shifted into the infrared. (b) Attenuation profile of polycarbonate (PC); (c) Attenuation profile of poly(methyl-methacrylate) (PMMA). The attenuation is dominated by overtone oscillations of C-H-bonds. It can still be below 0.1 dB/m at around 570 nm and is significantly lower than that of PC. (Figures adapted from [32])

Polymer waveguides can be fabricated on various substrates such as silicon wafers or special polymer sheets. Two mainstream processes to fabricate polymer waveguides rely on ridge subsequent overcladding or trench formation and overcladding, see Fig. 4.19. Polymer waveguides are, e.g., used for optical board-level interconnects. To facilitate fabrication and coupling, typical waveguide widths and heights amount to several tens of microns, which leads to highly multimode structures, see Fig. 4.20.

4.5.4 Silicon nitride waveguides

Silicon nitride waveguides usually use silicon dioxide (SiO_2, $n = 1.44$ at 1550 nm) as a cladding material in combination with waveguide cores that consist of alternating layers of Si_3N_4 ($n = 2.1$ at 1550 nm) and SiO_2, see Fig. 4.21. By varying the thickness of the layers, this technology allows for medium and high index-contrast waveguides that exhibit low propagation loss and are suitable for wavelengths from below 500 nm to 2 µm. Silicon nitride waveguides are, e.g., commercially offered by LioniX[4] under the label "TriPleX". Applications are not only limited to telecommunications but also comprise biochemical sensing and co-integration of optical waveguides and microfluidic systems.

4.5.5 Waveguides based on III-V compound semiconductors

III-V compound semiconductors feature direct bandgaps and are therefore ideally suited for active optical devices. The bandgap can be adjusted by the composition of the material. This allows to integrate laser diodes and photodiodes together with passive components. Two material systems are commonly used.

InGaAsP/InP waveguides

This material system allows for monolithic integration of optical devices in the 1.3 μm and 1.5 μm wavelength range, where fused-silica optical fibers have the lowest loss. High index-contrast waveguides are formed by epitaxial layer growth on an InP substrates and by etching. Waveguide structures are typically small (1-2 μm lateral dimensions), which can lead to problems when coupling light from optical fibers to on-chip waveguides. Typical cross sections of active InGaAsP waveguides are depicted in Fig. 4.22.

[4] http://www.lionixbv.nl/

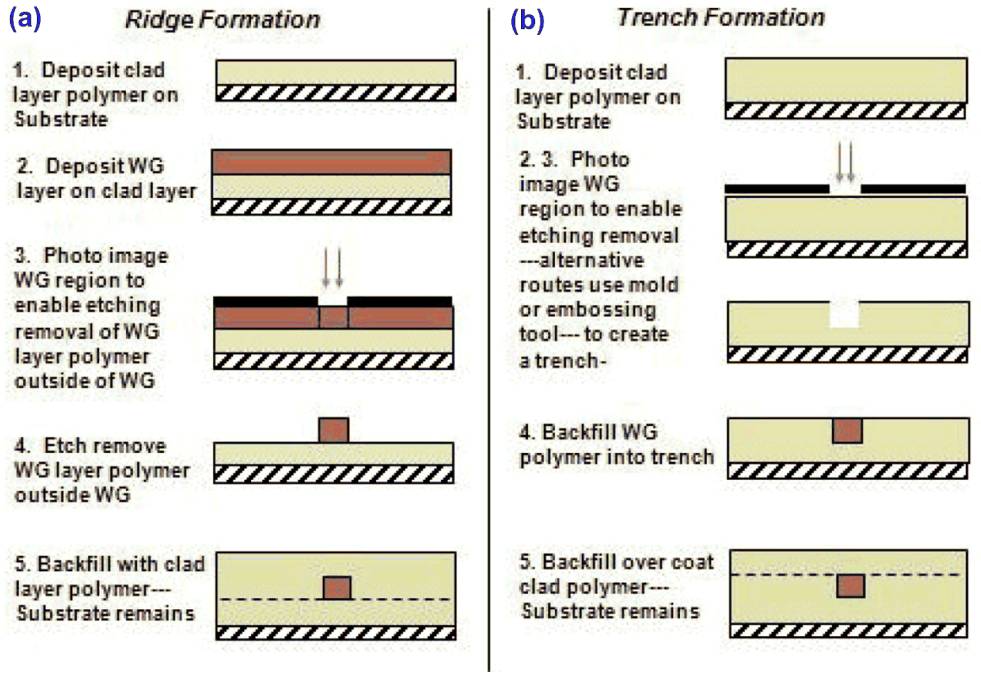

Figure 4.19: Polymer waveguide fabrication processes. (a) Ridge formation and subsequent overcladding; (b) Trench formation and backfilling (Figures adapted from http://www.opticalinterlinks.com/pdf/Paper.Poly.WG.Conn.pdf)

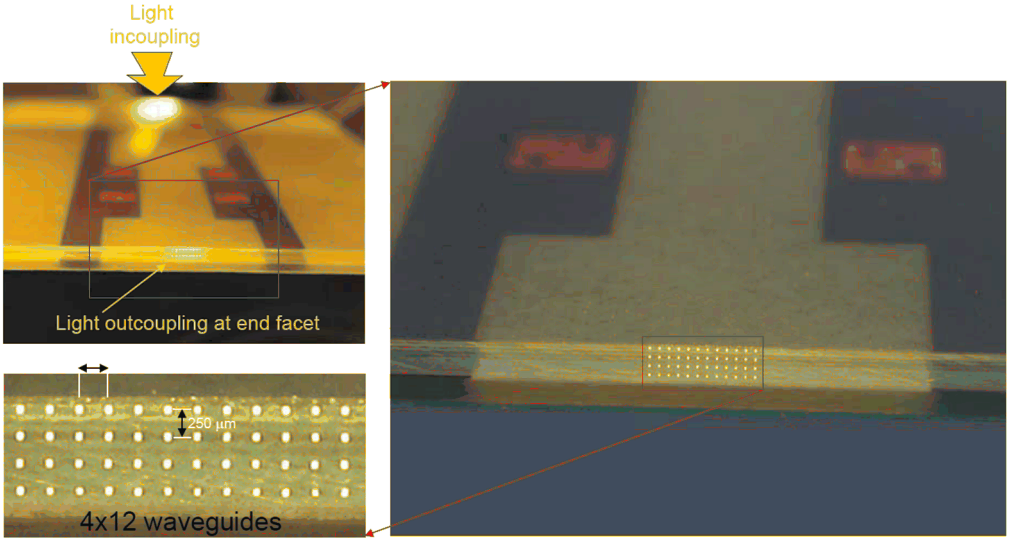

Figure 4.20: Multilayer optical waveguide in a printed-circuit board (PCB): 4×12 waveguides are arranged on a $250\,\mu\text{m} \times 250\,\mu\text{m}$ grid. Each of the waveguides has a cross section of approximately $50\,\mu\text{m} \times 50\,\mu\text{m}$ and is highly multimode. (Figures adapted from http://researcher.watson.ibm.com/researcher/files/us-yvlasov//vlasov%20ECOC%20tutorial%202008.pdf)

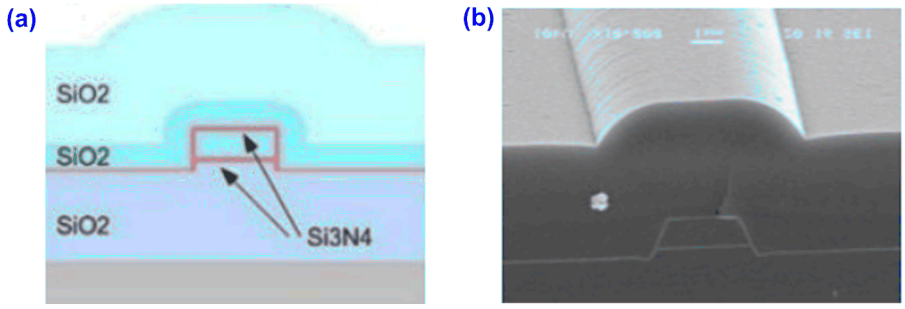

Figure 4.21: Waveguides based on silicon nitride (Si_3N_4, $n = 2.1$ at 1550 nm) and silicon dioxide (SiO_2, $n = 1.44$ at 1550 nm). These materials have a wide transparency window ranging from wavelengths from below 500 nm to 2 µm. Silicon nitride technology is commercially offered under the label "TriPleX". (a) Schematic waveguide cross section; (b) Cross-sectional view of fabricated waveguide; (Figures adapted from http://www.lionixbv.nl/)

AlGaAs/GaAs waveguides

GaAs-based semiconductor materials allow for monolithic integration of devices for the 0.5 μm wavelength range. Such components are used for short-distance interconnects, in scanners and in CD-players. As in the InP/InGaAsP material system, waveguide structures are typically small (1-2 μm lateral dimensions) and fiber-chip coupling can hence be challenging.

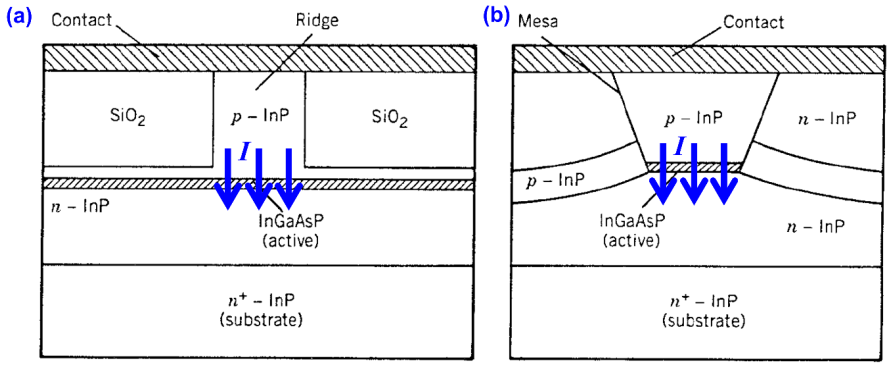

Figure 4.22: Active waveguides made of III-V-compound semiconductors. The active regions are electrically pumped; the pump current I is depicted in blue. (a) Ridge-waveguide structure; (b) Channel waveguide structure; (Figures adapted from [2])

4.5.6 Silicon-on-insulator (SOI) waveguides

Silicon is transparent at infrared telecommunication wavelengths and offers the potential to process optical devices in standard CMOS electronic fabs. This makes the material system particularly attractive for the co-integration of electronic and photonic devices, and substantial research effort is currently dedicated to silicon photonics. Waveguides are fabricated in a thin silicon layer ($n = 3.48$ at 1550 nm, thickness ~ 250 nm) which is optically isolated from the substrate material by an intermediate silica buffer layer ($n = 1.44$ at 1550 nm, thickness $\approx 2\,\mu$m). The fabrication process is illustrated in Fig. 4.23. The most crucial step is the definition of the waveguide structures. This requires highly sophisticated optical lithography systems operating at deep-UV wavelengths (typically 193 nm), thereby enabling resolutions smaller than 50 nm. After development, the resist structure can be used as an etch mask to structure the underlying silicon layer with reactive ion etching (RIE). The transfer of the waveguide structure to the silicon layer can be done either

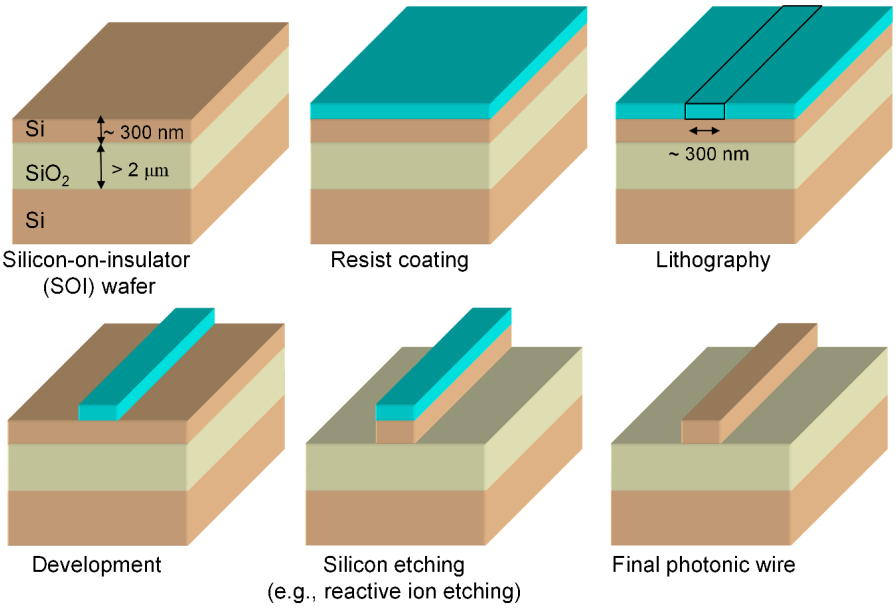

Figure 4.23: Fabrication of silicon-on-insulator waveguides by lithography and reactive ion etching.

directly, as shown in Fig. 4.23, or via an intermediate hard mask of, e.g., silicon nitride. Due to the high refractive index contrast between the core and the cladding material, silicon-on-insulator nanowire waveguides can have radii of curvature as small as 1 μm without appreciable radiation loss. On the other hand, the high index difference leads to strong scattering of light at the rough waveguide sidewalls ans hence to appreciable optical propagation loss.

Chapter 5

Optical fibers

Optical fibers are flexible filaments made of a transparent materials such as fused silica, speciality glasses or polymers. They are often not much thicker than a human hair ($\sim 100\,\mu$m diameter), and in the case of single-mode fibers, light is only carried in a core region of less than $10\,\mu$m diameter in the center of the fiber. Today, many different kinds of optical fibers for various applications are commercially available, covering wavelengths from the ultraviolet to the infrared. They are used in optical communications, but also in fiber-optical sensors, optical metrology and for spectroscopic applications. In this section, we will introduce the basic concepts of wave propagation in fibers, we will discuss different types of optical fibers, and we will give an overview on current fiber technologies.

5.1 Optical fiber basics

5.1.1 Silica fibers - a historical perspective

Silica-based optical fibers are the most important transmission medium for long-distance large-capacity optical communication systems. One of the most important characteristic of an optical fiber is its loss as a function of the wavelength, see Fig. 5.1. Early glass fibers in the 1960's showed transmission losses of typically 1000 dB/km - by far too much to be a viable solution for optical communications. In 1966, however, Charles K. Kao and George A. Hockham showed that these losses were mainly caused by impurities in the glass, not by a fundamental problem with the technology itself. They estimated that, by removing the impurities, optical attenuations below 20 dB/km can be achieved in silica fibers, which would make them a practical communication medium. For this discovery, Kao was awarded the Nobel Prize in Physics in 2009. Early optical communications in the 1970's focused on the first transmission window between 800 nm to 900 nm, since fibers had a local minimum in the attenuation curve and GaAs-based optical sources and photodetectors were available at these wavelengths. In the following decades, losses in the region between 1100 nm and 1600 nm could be reduced by several orders of magnitude by diminishing the concentration of hydroxyl ions and metallic ion impurities in the fiber material. This region is referred to as the "long-wavelength-region". Two optical communication windows are defined here: The second window, centered around 1300 nm, and the third window around 1550 nm. The second generation of fiber-optic communications was developed in the early 1980s, operating at around 1300 nm and using InGaAsP lasers. In the 1990s, third-generation fiber-optic systems appeared, operating at 1550 nm, where fiber losses could be reduced to below 0.2 dB/km. The lowest transmission loss ever demonstrated in a silica fiber is 0.154 dB/km [33], which corresponds to a power loss of 50 % after 20 km. However, this value was achieved for a fiber with a pure silica core and fluourine-doped cladding with a slightly lower refractive index. In prevalent single-mode fibers the core region is doped with germanium to increase the refractive index, which leads to

slightly higher loss figures of approximately[1] 0.20 dB/km.

With decreasing propagation loss, longer transmission distances became possible, and dispersion-induced pulse spreading became a limiting factor. These problems were overcome by using narrowband (longitudinally single-moded) laser sources and dispersion-shifted fibers or combinations of standard single-mode fibers and dispersion-compensating fibers. The principles of dispersion compensation will be presented in Section 5.6.2.

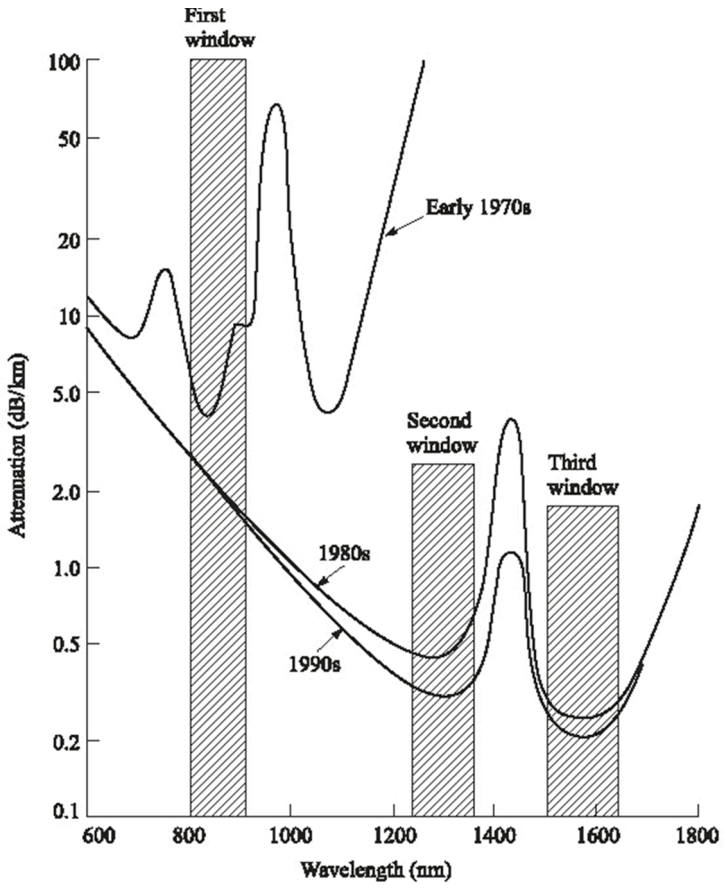

Figure 5.1: Propagation loss and transmission windows of optical fibers made of silica. In the early 1970s, fiber losses were of the order of 10 dB/km, with a local attenuation minimum between 800 nm and 900 nm. Early fiber links relied on this so-called first transmission window, using GaAs-based optical sources and photodetectors. In the following decades, losses in the region between 1100 nm and 1600 nm could be reduced by several orders of magnitude by diminishing the concentration of hydroxyl ions and metallic ion impurities in the fiber material. In combination with InGaAsP lasers, this led to communication systems that use the second window around 1300 nm and the third window around 1550 nm. (Figure adapted from [17])

5.1.2 Various types of optical fibers

Various kinds of optical fibers with different cross-sectional refractive index distributions exist, see Fig. 5.2. Each fiber type has its own advantages and disadvantages and is therefore suited for certain application fields only. Optical fibers can be subdivided in three groups:

[1] See, e.g., data sheet of Corning SMF28 standard single-mode fiber, http://www.focenter.com/Fiber%20Optic%20Center%5Ccorning%5Cdatasheets%5Csmf-28.pdf

Step-index fibers are characterized by an abrupt transition between the core and the cladding region. The index difference is often specified by the relative refractive index difference Δ which relates the core index n_1 and the cladding index n_2,

$$\Delta = \frac{n_1^2 - n_2^2}{2n_1^2}. \tag{5.1}$$

For silica-based step-index fibers, relative index differences are usually smaller than 1 %. The index difference is achieved by either doping the core (e.g., with germanium) or the cladding (e.g., with flourine), see Fig. 5.2 (d). Single-mode step-index fibers allow for low-loss broadband transmission, but coupling to and from these fibers is challenging due to the small core diameter of typically less than 10 μm. Multi-mode step-index fibers with core diameters of typically 50 μm facilitate coupling, but suffer from large intermodal dispersion and are therefore only used for short transmission distances. In extreme cases, core diameters of plastic step-index fibers can be as large as 1 mm with refractive index differences of, e.g., 7 %. Such fibers are mainly used in industrial environments, where robustness and easy coupling are important and data rates are rather low.

In **graded-index fibers**, the refractive index decreases continuously from the maximum value n_1 at the fiber axis to a constant value n_2 in the cladding region. The relative index difference in silica-based graded-index fibers is usually less than 1 %. Large core diameters of ~ 50 μm still allow for easy coupling, while dispersion is smaller than in comparable step-index fibers[2].

For special applications, **fibers with non-rotationally symmetric index profiles** are used. Holey fibers, e.g., consist of a single material with a periodic or an aperiodic array of air holes that are running in parallel to the fiber axis. The presence of air holes reduces the average refractive index in the cladding region. In a similar arrangement, multiple reflections from dielectric interfaces between the air holes and the bulk material lead to the formation of an optical bandgap and hence to confinement of light to the core of the fiber. In polarization-preserving fibers, the effective index of refraction is increased in two regions that are positioned diametrically with respect to the waveguide core, Fig. 5.2 (g). This leads to the formation of two eigenmodes that are predominantly linearly polarized - one parallel to the direction in which the refractive index is raised, and one perpendicular to it.

5.1.3 Cylindrical coordinates and refractive index profile representation

For rotationally symmetric index profiles, cylindrical coordinates (polar coordinates) (r, φ, z) are used to simplify the analysis. For more details on the relation between cartesian and cylindrical coordinates and on the use of differential operators in polar coordinates, see Appendix A.3. To simplify the analysis, we make use of the fact that for usual operating conditions, the fields are well confined to the waveguide core and do not reach far into the cladding. We assume, as an approximation, that the cladding is infinitely thick. For a uniform waveguide, the refractive index profile is independent of φ, z. For a generalized representation of the refractive index profile, we use a model function of the form

$$n^2(r) = \begin{cases} n_1^2 \left[1 - 2\Delta\, g\left(\frac{r}{a}\right)\right] & \text{for } 0 \le r < a, \\ n_1^2 \left[1 - 2\Delta\right] = n_2^2 & \text{for } a \le r < \infty, \end{cases} \tag{5.2}$$

where $g(r/a)$ represents the so-called profile function fulfilling $g(0) = 0$ and $g(1) = 1$. In this equation, the refractive index on the fiber axis is denoted as n_1. The relative refractive index difference Δ was already defined in Eq. (5.1). Note that, in this representation, the index profile is represented by a function that is formally continuous at the core-cladding interface. Step-index profiles are then represented by a profile function $g(r/a)$ that has an infinitely large slope. An important class of profile functions is represented by the so-called power law profiles,

$$g\left(\frac{r}{a}\right) = \left(\frac{r}{a}\right)^q, \qquad 0 \le q < \infty. \tag{5.3}$$

[2] This is explained in more detail in the lecture "Field Propagation and Coherence" by W. Freude.

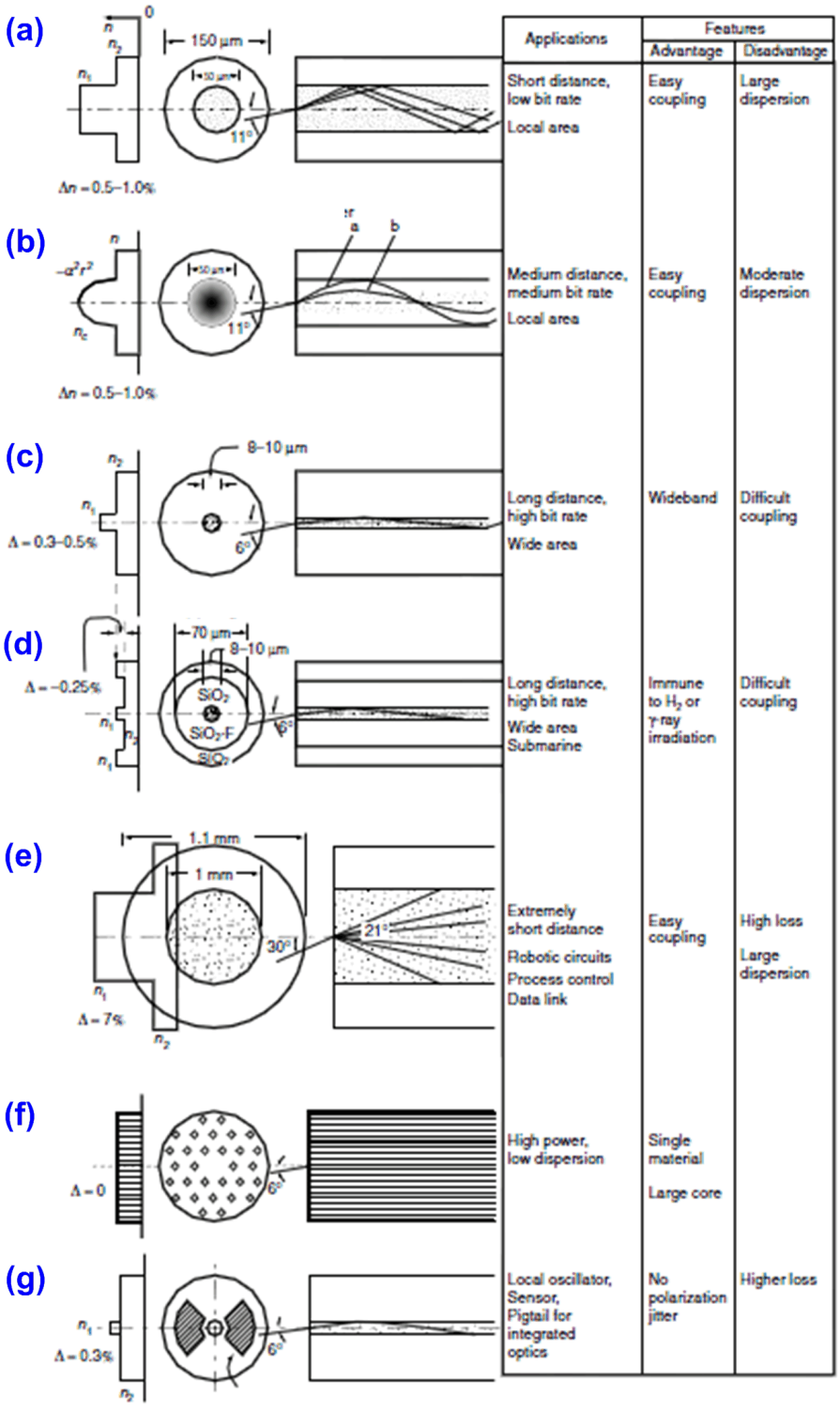

Figure 5.2: Optical fibers exist in a wide variety of materials and cross-sectional refractive index profiles. Each fiber type has its own advantages and disadvantages, which makes it suited for certain application fields only. (a) Multi-mode step-index fiber; (b) Graded-index fiber; (c) Single-mode fiber; (d) Silica-core fiber with fluorine-added cladding; (e) Plastic fiber; (f) Holey fiber; (g) Polarization-preserving fiber (Figure adapted from [11])

For $q \to \infty$ a step-index profile results ($n = n_1$ for $0 \leq r < a$, and $n = n_2$ for $r \geq a$), whereas finite q leads to various types of graded-index profiles. The profile exponent ("grading parameter") q may be chosen such that intermodal dispersion becomes minimum. The choice of $q = 2$ defines a so-called parabolic index profile.

5.1.4 Maxwell's equations in cylindrical coordinates

For the mode fields, an ansatz equivalent to Eqs. (3.59) and (3.60) can be formulated in cylindrical coordinates:

$$\underline{\mathbf{E}}(\mathbf{r}, t) = \underline{\mathcal{E}}(r, \varphi) \exp\left(\mathrm{j}\left(\omega t - \beta z\right)\right), \tag{5.4}$$

$$\underline{\mathbf{H}}(\mathbf{r}, t) = \underline{\mathcal{H}}(r, \varphi) \exp\left(\mathrm{j}\left(\omega t - \beta z\right)\right), \tag{5.5}$$

Writing Maxwell's curl equations in cylindrical coordinates (see Appendix A.3), we obtain the following relations for the mode field components:

$$\frac{1}{r}\frac{\partial \underline{\mathcal{E}}_z}{\partial \varphi} + \mathrm{j}\,\beta \underline{\mathcal{E}}_\varphi = -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_r \tag{5.6}$$

$$-\mathrm{j}\,\beta \underline{\mathcal{E}}_r - \frac{\partial \underline{\mathcal{E}}_z}{\partial r} = -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_\varphi \tag{5.7}$$

$$\frac{1}{r}\left(\frac{\partial\left(r\underline{\mathcal{E}}_\varphi\right)}{\partial r} - \frac{\partial \underline{\mathcal{E}}_r}{\partial \varphi}\right) = -\mathrm{j}\,\omega\mu_0 \underline{\mathcal{H}}_z \tag{5.8}$$

and

$$\frac{1}{r}\frac{\partial \underline{\mathcal{H}}_z}{\partial \varphi} + \mathrm{j}\,\beta \underline{\mathcal{H}}_\varphi = \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_r \tag{5.9}$$

$$-\mathrm{j}\,\beta \underline{\mathcal{H}}_r - \frac{\partial \underline{\mathcal{H}}_z}{\partial r} = \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_\varphi \tag{5.10}$$

$$\frac{1}{r}\left(\frac{\partial\left(r\underline{\mathcal{H}}_\varphi\right)}{\partial r} - \frac{\partial \underline{\mathcal{H}}_r}{\partial \varphi}\right) = \mathrm{j}\,\omega\epsilon_0 n^2 \underline{\mathcal{E}}_z \tag{5.11}$$

Using these relations, the transverse field components can be expressed by the longitudinal components,

$$\underline{\mathcal{E}}_r = -\frac{\mathrm{j}}{k_0^2 n^2 - \beta^2}\left(\beta\frac{\partial \underline{\mathcal{E}}_z}{\partial r} + \frac{\omega\mu_0}{r}\frac{\partial \underline{\mathcal{H}}_z}{\partial \varphi}\right) \tag{5.12}$$

$$\underline{\mathcal{E}}_\varphi = -\frac{\mathrm{j}}{k_0^2 n^2 - \beta^2}\left(\frac{\beta}{r}\frac{\partial \underline{\mathcal{E}}_z}{\partial \varphi} - \omega\mu_0\frac{\partial \underline{\mathcal{H}}_z}{\partial r}\right) \tag{5.13}$$

$$\underline{H}_r = -\frac{\mathrm{j}}{k_0^2 n^2 - \beta^2}\left(\beta\frac{\partial \underline{\mathcal{H}}_z}{\partial r} - \frac{\omega\epsilon_0 n^2}{r}\frac{\partial \underline{E}_z}{\partial \varphi}\right) \tag{5.14}$$

$$\underline{H}_\varphi = -\frac{\mathrm{j}}{k_0^2 n^2 - \beta^2}\left(\frac{\beta}{r}\frac{\partial \underline{\mathcal{H}}_z}{\partial \varphi} + \omega\epsilon_0 n^2\frac{\partial \underline{\mathcal{E}}_z}{\partial r}\right) \tag{5.15}$$

These relations will be used in the following to first formulate the mode field ansatz in terms of $\underline{\mathcal{E}}_z$ and $\underline{\mathcal{H}}_z$. The transverse mode field components can finally be derived from the solution for the $z-$components.

5.2 Step-index fibers

5.2.1 Wave theory of step-index fibers

The analysis of guided modes in step-index fibers proceeds in a similar way as for the slab waveguide. First we use a mode ansatz for the longitudinal mode field components $\underline{\mathcal{E}}_z$ and $\underline{\mathcal{H}}_z$, and we

formulate a solution based on the mode propagation constant β as an unknown free parameter. An implicit equation for the mode propagation constant β can then be obtained by enforcing continuity of the appropriate field components at the core-cladding interface. By numerically solving the implicit equation for β, we can calculate the mode field components $\underline{\mathcal{E}}_z$ and $\underline{\mathcal{H}}_z$, from which all other components can be derived.

Wave equation and separation of variables

Within the homomgeneous core and cladding regions, the different components of the vectorial wave equations, Eqs. (2.24) and (2.25), are decoupled. Expressing the ∇^2-operator in cylindrical coordinates and using the mode ansatzes Eqs. (5.4) and (5.5), the wave equations can be written as

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \underline{\mathcal{E}}_z}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 \underline{\mathcal{E}}_z}{\partial \varphi^2}+\left(k_0^2 n^2-\beta^2\right)\underline{\mathcal{E}}_z=0 \tag{5.16}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \underline{\mathcal{H}}_z}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 \underline{\mathcal{H}}_z}{\partial \varphi^2}+\left(k_0^2 n^2-\beta^2\right)\underline{\mathcal{H}}_z=0 \tag{5.17}$$

We may represent these relations by a single equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \underline{\Psi}(r,\varphi)}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 \underline{\Psi}(r,\varphi)}{\partial \varphi^2}+\left(k_0^2 n^2-\beta^2\right)\underline{\Psi}(r,\varphi)=0, \tag{5.18}$$

where $\underline{\Psi}$ can stand for $\underline{\mathcal{E}}_z$ or $\underline{\mathcal{H}}_z$. For the solution we assume that the r-dependence can be separated from the φ-dependence by an ansatz of the form

$$\underline{\Psi}(r,\varphi)=g(r)\;h(\varphi). \tag{5.19}$$

Inserting Eq. (5.19) into Eq. (5.18), we obtain

$$\underbrace{\frac{r}{g(r)}\frac{\partial}{\partial r}\left(r\frac{\partial g(r)}{\partial r}\right)+r^2\left(k_0^2 n^2-\beta^2\right)}_{C_1}+\underbrace{\frac{1}{h(\varphi)}\frac{\partial^2 h(\varphi)}{\partial \varphi^2}}_{C_2}=0 \tag{5.20}$$

The left part depends on r only, whereas the right part is purely φ-dependent. Since the relation must hold for all r and φ, both parts must be constant and $C_1=-C_2$.

Let us first consider the φ-dependence and the corresponding boundary conditions. As a matter of fact, $h(\varphi)$ must be periodic in φ,

$$h(\varphi+2\pi)=h(\varphi) \tag{5.21}$$

The basic solution for h is hence given by

$$h(\varphi)=\begin{cases}\cos(\nu\varphi)\\ \text{or}\\ \sin(\nu\varphi)\end{cases}\qquad \text{for } \nu=0,1,2\ldots, \tag{5.22}$$

Note that for $\nu=0$, we obtain the trivial solution $h=0$ and hence $\underline{\Psi}=0$ for the case of the $\sin(\nu\varphi)$-dependence. As we will see later, this corresponds to the case where one of the longitudinal field components $\underline{\mathcal{E}}_z$ or $\underline{\mathcal{H}}_z$ vanishes, while the other is nonzero. The corresponding modes are denoted as TE or TM modes.

Inserting the basic solution for $h(\varphi)$, Eq. (5.22) in Eq. (5.20), we obtain a differential equation for the r-dependence,

$$r^2\frac{\partial^2 g(r)}{\partial r^2}+r\frac{\partial g(r)}{\partial r}+\left[\left(k_0^2 n^2-\beta^2\right)r^2-\nu^2\right]g(r)=0. \tag{5.23}$$

This relation corresponds to the Bessel differential equation which is solved by Bessel functions and modified Bessel functions, see Appendix A.3 [29]. Two cases have to be distinguished:

1. $k_{1r}^2 = k_0^2 n_1^2 - \beta^2 > 0$: This is the case inside the waveguide core, i.e. for $0 \leq r \leq a$ and $n = n_1$. The field oscillates as a function of r. The general solution reads

$$g(r) = C_4 \, \mathrm{J}_\nu \left(u \frac{r}{a} \right) + C_5 \, \mathrm{Y}_\nu \left(u \frac{r}{a} \right) \tag{5.24}$$

where J_ν (Y_ν) denote Bessel functions of the first (second) kind of order ν, and where the transverse phase constant is given by

$$u = a \sqrt{n_1^2 k_0^2 - \beta^2} \tag{5.25}$$

2. $k_{2r}^{(i)2} = \beta^2 - k_0^2 n_2^2 > 0$: This is the case in the cladding outside the waveguide core, i.e. for $r > a$ and $n = n_2$. The field behaves monotonously as a function of r. The general solution reads

$$g(r) = C_6 \, \mathrm{I}_\nu \left(u \frac{r}{a} \right) + C_7 \, \mathrm{K}_\nu \left(u \frac{r}{a} \right) \tag{5.26}$$

where I_ν and K_ν denote modified Bessel functions. The transverse cladding constant is given by

$$w = a \sqrt{\beta^2 - n_2^2 k_0^2} \tag{5.27}$$

The set of general solutions according to Eqs. (5.24) and (5.26) can be further restricted by requiring "physically meaningful" field profiles. For the Bessel function of the second kind, we observe that $\left| \mathrm{Y}_\nu \left(u \frac{r}{a} \right) \right| \to \infty$ for $r \to 0$, see Fig. 5.3 (a). We may hence conclude that $C_5 = 0$ - otherwise the mode field would have a singularity at $r = 0$. Likewise, we find for the modified Bessel functions that $\left| \mathrm{I}_\nu \left(u \frac{r}{a} \right) \right| \to \infty$ for $r \to \infty$, Fig. 5.3 (b). To ensure that the field is localized to the waveguide core, we must hence require $C_7 = 0$. The complete solution for $\underline{\Psi}(r, \varphi)$ can then be written as:

$$\underline{\Psi}(r, \varphi) = \begin{cases} A \, \mathrm{J}_\nu \left(u \frac{r}{a} \right) \cos(\nu\varphi + \psi) & \text{for } 0 \leq r \leq a \\ A \, \frac{\mathrm{J}_\nu(u)}{\mathrm{K}_\nu(w)} \, \mathrm{K}_\nu \left(u \frac{r}{a} \right) \cos(\nu\varphi + \psi) & \text{for } r > a \end{cases} \tag{5.28}$$

where $\nu = 0, 1, \ldots$ and $\psi \in \left\{ 0; \frac{\pi}{2} \right\}$.

Solutions for $\underline{\mathcal{E}}_z$ and $\underline{\mathcal{H}}_z$

From Eq. (5.13) we find that $\partial \underline{\mathcal{H}}_z / \partial r$ and $\partial \underline{\mathcal{E}}_z / \partial \varphi$ must have the same φ-dependence. We conclude that, if $\underline{\mathcal{H}}_z$ has a $\cos(\nu\varphi)$-dependency, $\underline{\mathcal{E}}_z$ must have a $\sin(\nu\varphi)$ and vice versa. The z-components of the electromagnetic field can hence be written as

$$\underline{\mathcal{E}}_z(r, \varphi) = \begin{cases} A \, \mathrm{J}_\nu \left(u \frac{r}{a} \right) \cos(\nu\varphi + \psi) & \text{for } 0 \leq r \leq a \\ A \, \frac{\mathrm{J}_\nu(u)}{\mathrm{K}_\nu(w)} \, \mathrm{K}_\nu \left(u \frac{r}{a} \right) \cos(\nu\varphi + \psi) & \text{for } r > a \end{cases} \tag{5.29}$$

$$\underline{\mathcal{H}}_z(r, \varphi) = \begin{cases} B \, \mathrm{J}_\nu \left(u \frac{r}{a} \right) \sin(\nu\varphi + \psi) & \text{for } 0 \leq r \leq a \\ B \, \frac{\mathrm{J}_\nu(u)}{\mathrm{K}_\nu(w)} \, \mathrm{K}_\nu \left(u \frac{r}{a} \right) \sin(\nu\varphi + \psi) & \text{for } r > a \end{cases} \tag{5.30}$$

where $\nu = 0, 1, \ldots$ and $\psi \in \left\{ 0; \frac{\pi}{2} \right\}$.

5.2.2 TE-, TM- and hybrid modes of step-index fibers

Transverse-electric (TE) modes

For $\nu = 0$ and $\psi = \frac{\pi}{2}$ the longitudinal electric field component vanishes,

$$\underline{\mathcal{E}}_z(r, \varphi) = 0 \tag{5.31}$$

$$\underline{\mathcal{H}}_z(r, \varphi) = \begin{cases} B \, \mathrm{J}_0 \left(u \frac{r}{a} \right) & \text{for } 0 \leq r \leq a \\ B \, \frac{\mathrm{J}_0(u)}{\mathrm{K}_0(w)} \, \mathrm{K}_0 \left(u \frac{r}{a} \right) & \text{for } r > a \end{cases} \tag{5.32}$$

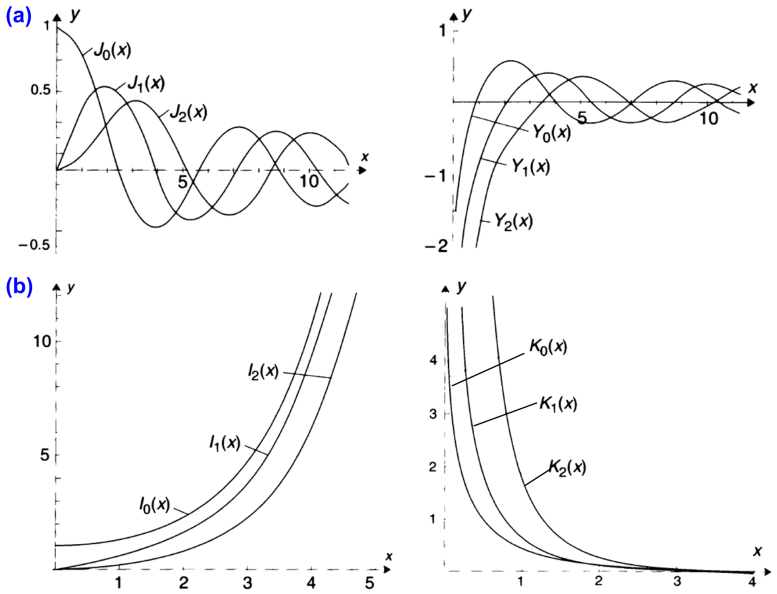

Figure 5.3: Plots of Bessel functions and modified Bessel functions. (a) J_ν and Y_ν denote ν-th order Bessel functions of the first and second kind, respectively. Note that $\left|\mathrm{Y}_\nu\left(u\frac{r}{a}\right)\right| \to \infty$ for $r \to 0$, and that the functions are hence not suited for representation of a finite electromagnetic field component within the waveguide core (near $r = 0$). (b) I_ν and K_ν denote ν-th order modified Bessel functions. Note that $\left|\mathrm{I}_\nu\left(u\frac{r}{a}\right)\right| \to \infty$ for $r \to \infty$, and that the functions are hence not suited for representation of a finite electromagnetic field component in the waveguide cladding (for $r \to \infty$). (Adapted from [29])

This leads to the so-called transverse-electric (TE) modes, which have only three nonzero electromagnetic field components,

$$\underline{\mathcal{E}}_r = 0;\ \underline{\mathcal{E}}_\varphi \neq 0;\ \underline{\mathcal{E}}_z = 0 \tag{5.33}$$

$$\underline{\mathcal{H}}_r \neq 0;\ \underline{\mathcal{H}}_\varphi = 0;\ \underline{\mathcal{H}}_z \neq 0 \tag{5.34}$$

Inserting Eqs. (5.31) and (5.32) in Eqs. (5.12) and (5.13), we can derive the other electric field components,

$$\underline{\mathcal{E}}_r = 0 \tag{5.35}$$

$$\underline{\mathcal{E}}_\varphi = \mathrm{j}\,\omega\mu_0 \begin{cases} -\frac{Ba}{u}\,\mathrm{J}_1\left(u\frac{r}{a}\right) & \text{for } 0 \leq r \leq a \\ \frac{Ba}{w}\,\frac{\mathrm{J}_0(u)}{\mathrm{K}_0(w)}\,\mathrm{K}_1\left(w\frac{r}{a}\right) & \text{for } r > a \end{cases} \tag{5.36}$$

The continuity of $\underline{\mathcal{E}}_\varphi$ at $r = a$ yields the dispersion equation for TE modes:

$$\frac{\mathrm{J}_1(u)}{u\,\mathrm{J}_0(u)} = -\frac{\mathrm{K}_1(w)}{w\,\mathrm{K}_0(w)} \tag{5.37}$$

where

$$u^2 + w^2 = V^2, \tag{5.38}$$

and where the normalized frequency V for optical fibers is defined in analogy to Eq. (3.49),

$$V = ak_0\sqrt{n_1^2 - n_2^2}. \tag{5.39}$$

Note that the continuity of $\underline{\mathcal{H}}_r$ at $r = a$ reproduces Eq. (5.37). Eqs. (5.37) and (5.38) must be solved numerically to obtain u, w, and hence β.

Transverse-magnetic (TM) modes

Likewise, the $\underline{\mathcal{H}}_z$-component vanishes for $\nu = 0$ and $\psi = 0$,

$$\underline{\mathcal{E}}_z(r,\varphi) = \begin{cases} A\ \mathrm{J}_0\left(u\frac{r}{a}\right) & \text{for } 0 \le r \le a \\ A\ \frac{\mathrm{J}_0(u)}{\mathrm{K}_0(w)}\ \mathrm{K}_0\left(u\frac{r}{a}\right) & \text{for } r > a \end{cases} \tag{5.40}$$

$$\underline{\mathcal{H}}_z(r,\varphi) = 0 \tag{5.41}$$

For the resulting transverse-magnetic (TM) modes, there are again only three nonzero electromagnetic field components,

$$\underline{\mathcal{E}}_r \neq 0;\ \underline{\mathcal{E}}_\varphi = 0;\ \underline{\mathcal{E}}_z \neq 0 \tag{5.42}$$

$$\underline{\mathcal{H}}_r = 0;\ \underline{\mathcal{H}}_\varphi \neq 0;\ \underline{\mathcal{H}}_z = 0 \tag{5.43}$$

Deriving the $\underline{\mathcal{H}}_\varphi$-component by means of Eq. (5.15) and enforcing continuity at $r = a$, we obtain the dispersion equation for TM modes:

$$\frac{\mathrm{J}_1(u)}{u\,\mathrm{J}_0(u)} = -\left(\frac{n_2}{n_1}\right)^2 \frac{\mathrm{K}_1(w)}{w\,\mathrm{K}_0(w)} \tag{5.44}$$

where

$$u^2 + w^2 = V^2. \tag{5.45}$$

The continuity of $n^2\underline{\mathcal{E}}_r$ at $r = a$ reproduces Eq. (5.44). Numerical solution of Eqs. (5.44) and (5.45) yields u, w, and hence β.

Hybrid modes

In the general case, both $\underline{\mathcal{E}}_z$ and $\underline{\mathcal{H}}_z$ are nonzero. The resulting modes are referred to as hybrid modes. By inserting Eqs. (5.29) and (5.30) into Eqs. (5.12) to (5.15), we can derive all transverse field components. The field components must satisfy the usual boundary conditions at $r = a$ which, after some calculation [25], leads to the dispersion equation for hybrid modes,

$$\left[\frac{\mathrm{J}'_\nu(u)}{u\,\mathrm{J}_\nu(u)} + \frac{\mathrm{K}'_\nu(w)}{w\,\mathrm{K}_\nu(w)}\right]\left[\frac{\mathrm{J}'_\nu(u)}{u\,\mathrm{J}_\nu(u)} + \left(\frac{n_2}{n_1}\right)^2 \frac{\mathrm{K}'_\nu(w)}{w\,\mathrm{K}_\nu(w)}\right] \tag{5.46}$$

$$= \nu^2\left[\frac{1}{u^2} + \frac{1}{w^2}\right]\left[\frac{1}{u^2} + \left(\frac{n_2}{n_1}\right)^2 \frac{1}{w^2}\right] \tag{5.47}$$

where

$$u^2 + w^2 = V^2 \tag{5.48}$$

As in the previous cases, numerical solution of Eqs. (5.44) and (5.45) yields u, w, and hence β.

Dispersion relation, single-mode condition, mode designation and fields

Solving the dispersion relations for different normalized frequencies V, we can derive the various dispersion relations of different step-index fiber modes, see Fig. 5.4. Two integer subscripts ν and μ are used to describe the mode fields. The first subscript ν as used in the previous sections relates to the azimuthal variation of the fields and is therefore also called the azimuthal mode index. TE-modes exist only for $\nu = 0$ and are hence rotationally symmetric. They are denoted as $\mathrm{TE}_{0\mu}$-modes, where the nonzero mode index μ denotes the number of extremal points of the azimutal electric field component $\underline{\mathcal{E}}_\varphi$ along the radial direction r. Likewise, the $\mathrm{TM}_{0\mu}$ is rotationally symmetric,

and the number of extremal points of the azimutal magentic field component $\underline{\mathcal{H}}_\varphi$ along the radial direction r is given by the nonzero mode index μ. The hybrid $\mathrm{EH}_{\nu\mu}$ and $\mathrm{HE}_{\nu\mu}$-modes exist for $\nu \geq 1$ and are twofold degenerate with a $\sin(\nu\varphi)$ and $\cos(\nu\varphi)$-dependence. The designation as $\mathrm{EH}_{\nu\mu}$ or $\mathrm{HE}_{\nu\mu}$-modes is somewhat involved and based on the field with dominant longitudinal component: For a dominant $\underline{\mathcal{H}}_z$-component, the mode is referred to as an $\mathrm{EH}_{\nu\mu}$, whereas a dominant $\underline{\mathcal{E}}_z$-component leads to a $\mathrm{HE}_{\nu\mu}$-mode.

The fundamental mode turns out to be the hybrid HE_{11}-mode. Assuming an infinitely extended cladding region, this mode does not have a lower cutoff frequency, i.e. there is always at least one guided mode (with a two-fold polarization degeneracy) in an ideal step-index fiber. The upper limit of the single-mode region can be derived from the cut-off frequency of the next higher-order mode set. Observing that $\beta = n_2 k_0$, $w = 0$ and hence $V = u$ at cut-off, we find that no higher-order modes exist if

$$V = V_{01} \approx 2.405, \tag{5.49}$$

where V_{01} denotes the first zero of the 0th-order Bessel function $J_0(V)$.

From Fig. 5.4 we can observe that the modes of a step-index fiber form groups with similar dispersion characteristics and cutoff-frequencies. This applies, e.g., to the TE_{01}-, TM_{01}-, and HE_{21}-modes, or to the HE_{12}-, HE_{31}-, and EH_{11}-modes. For low index-contrast fibers, these mode families can be represented by a set of linearly polarized modes, see Section 5.2.3.

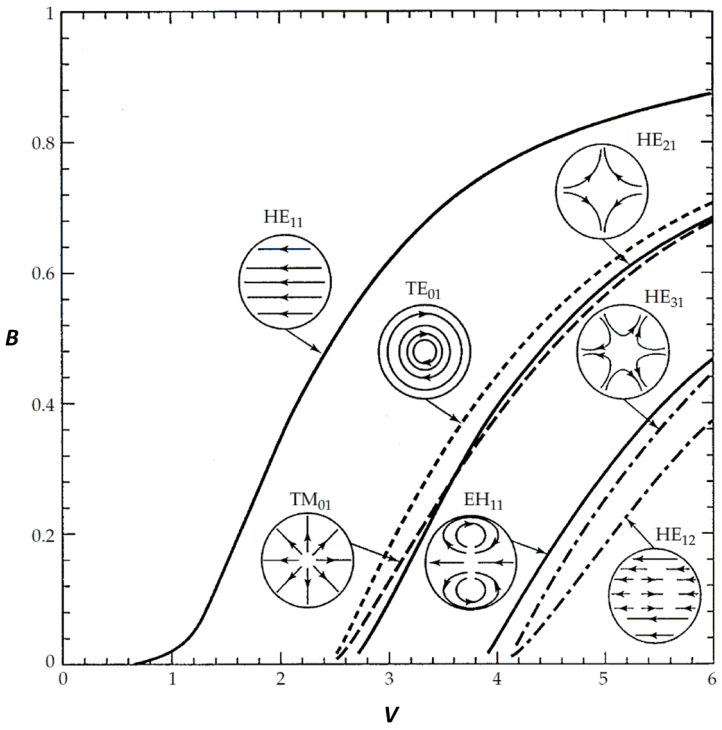

Figure 5.4: Dispersion relations of the various modes of a step-index fiber. The fundamental mode of a step index fiber is the hybrid HE_{11}-mode. The normalized cutoff frequency $V_{01} \approx 2.405$ for first higher-order mode is given by the first zero of the 0th-order Bessel function $J_0(V)$. Note that the higher-order modes form groups with similar dispersion characteristics and cutoff-frequencies, see, e.g., the TE_{01}-, TM_{01}-, and HE_{21}-modes, or the HE_{12}-, HE_{31}-, and EH_{11}-modes. (Adapted from [6])

5.2.3 Linearly polarized (LP) modes

In the last section, a rigorous analysis for the TE, TM and hybrid modes of a step-index fiber has been described. No special assumptions had to be made concerning the index difference between the core and the cladding material, and the results are valid for any step-index waveguide with rotational symmetry. For practical fibers, however, the relative refractive index difference Δ is very small, typically 0.3%. This allows drastic simplifications of the mode analysis by using so-called linearly polarized (LP) modes. To this end, we impose two fundamental assumptions, the validity of which will be confirmed with hindsight by comparing the results to those of the rigorous analysis.

First, we assume that the wave equations for weakly inhomogeneous materials, Eqs. (2.24) and (2.25), can be used, where the equations for the different Cartesian coordinates are decoupled,

$$\nabla^2 \underline{\mathcal{E}}_x + \left(k_0^2 n^2 - \beta^2\right) \underline{\mathcal{E}}_x = 0 \tag{5.50}$$

Second, we assume that one transverse Cartesian electric field component is usually much stronger than the other electric field components. This is in analogy to Marcatili's treatment of a rectangular waveguide. Without loss of generality, we assume that the $\underline{\mathcal{E}}_x$-component of the modal field dominates whereas $\underline{\mathcal{E}}_y$ vanishes, $|\underline{\mathcal{E}}_x| \gg |\underline{\mathcal{E}}_z|$ and $|\underline{\mathcal{E}}_y| = 0$. Following this approach, the equation for the $\mathcal{E}_x$-component is formulated in analogy to Eq. (5.16) in polar coordinates,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\,\frac{\partial \underline{\Psi}(r,\varphi)}{\partial r}\right) + \frac{1}{r^2}\,\frac{\partial^2 \underline{\Psi}(r,\varphi)}{\partial \varphi^2} + \left(k_0^2 n^2 - \beta^2\right)\underline{\Psi}(r,\varphi) = 0, \tag{5.51}$$

where $\underline{\Psi}(r,\varphi)$ represents the $\underline{\mathcal{E}}_x$-component of the modal field. As discussed in Section (5.2.2), this equation is solved by a separation ansatz which leads to a Bessel differential equation for the dependence on r. The Bessel functions J_ν (modified Bessel functions K_ν) are again identified as physically reasonable solutions in the core (cladding) region.

To match the boundary conditions at the core-cladding interface $r = a$ exactly, the $\underline{\mathcal{E}}_x$-component of the modal field must be decomposed in an $\mathcal{E}_\varphi$-component and an $\mathcal{E}_r$-component that obey different boundary conditions: The $\mathcal{E}_\varphi$-component is tangential to the dielectric boundary and must hence be continuous, whereas the $\mathcal{E}_r$-component that is normal to the boundary with $n^2\mathcal{E}_r$ being the continuous quantity. However, since the index contrast is low, $n_1^2/n_2^2 \approx 1$, both conditions are approximately fulfilled if we require $\underline{\Psi}(r,\varphi)$ to be continuous across the boundary. Using the same procedure as in Section (5.2.2), the solution of Eq. (5.51) can be written as

$$\underline{\Psi}(r,\varphi) = \begin{cases} A\ \mathrm{J}_\nu\left(u\frac{r}{a}\right)\cos\left(\nu\varphi + \psi\right) & \text{for } 0 \le r \le a \\ A\ \frac{\mathrm{J}_\nu(u)}{\mathrm{K}_\nu(w)}\,\mathrm{K}_\nu\left(u\frac{r}{a}\right)\cos\left(\nu\varphi + \psi\right) & \text{for } r > a \end{cases}, \tag{5.52}$$

where $\nu = 0, 1, 2, \ldots$ and $\psi \in \left\{0; \frac{\pi}{2}\right\}$, and where

$$u^2 + w^2 = V^2$$

The parameters u and w have to be determined from the boundary conditions of the longitudinal field components at $r = a$. To this end, we calculate the $\underline{\mathcal{E}}_z$- component from Maxwell's divergence equation

$$\begin{aligned} \nabla \cdot \underline{\mathbf{D}}(\mathbf{r},t) &= \nabla \cdot \left(\epsilon_0 n^2(\mathbf{r})\underline{\mathbf{E}}(\mathbf{r},t)\right) \approx \epsilon_0 n^2(\mathbf{r})\,\nabla \cdot \underline{\mathbf{E}}(\mathbf{r},t) = \\ &= \epsilon_0 n^2(\mathbf{r})\left(\frac{\partial \underline{\mathcal{E}}_x}{\partial x} + \frac{\partial \underline{\mathcal{E}}_y}{\partial y} - \mathrm{j}\,\beta \underline{\mathcal{E}}_z\right)\mathrm{e}^{\mathrm{j}(\omega t - \beta z)} = 0, \end{aligned} \tag{5.53}$$

which for $\underline{\mathcal{E}}_y = 0$ leads to

$$\underline{\mathcal{E}}_z = -\frac{\mathrm{j}}{\beta}\,\frac{\partial \underline{\mathcal{E}}_x}{\partial x} = -\frac{\mathrm{j}}{\beta}\left(\cos\varphi\frac{\partial \underline{\Psi}}{\partial r} - \frac{\sin\varphi}{r}\,\frac{\partial \underline{\Psi}}{\partial \varphi}\right) \tag{5.54}$$

Note that $\underline{\Psi}(r,\varphi)$ is assumed to be continuous across $r = a$ and that $\underline{\Psi}(r,\varphi)$ has identical φ-dependencies inside and outside the core, i.e., $\partial\underline{\Psi}/\partial\varphi$ must also be continuous. Eq. (5.54) hence implies that $\partial\underline{\Psi}/\partial r$ must be continuous as well at the boundary. This yields the characteristic equation for the $\mathrm{LP}_{\nu\mu}$-modes:

$$\frac{u\,\mathrm{J}'_\nu(u)}{\mathrm{J}_\nu(u)} = \frac{w\,\mathrm{K}'_\nu(w)}{\mathrm{K}_\nu(w)} \tag{5.55}$$

where

$$u^2 + w^2 = V^2 \tag{5.56}$$

Using the recurrence relations of the Bessel functions for $\nu > 0$,

$$\mathrm{J}'_\nu(u) = \mathrm{J}_{\nu-1}(u) - \frac{\nu}{u}\,\mathrm{J}_\nu(u)\,, \tag{5.57}$$

$$\mathrm{K}'_\nu(w) = -\,\mathrm{K}_{\nu-1}(w) - \frac{\nu}{w}\,\mathrm{K}_\nu(w)\,, \tag{5.58}$$

Eq. (5.55) can be rewritten,

$$\frac{u\,\mathrm{J}_{\nu-1}(u)}{\mathrm{J}_\nu(u)} = -\frac{w\,\mathrm{K}_{\nu-1}(w)}{\mathrm{K}_\nu(w)}, \tag{5.59}$$

where

$$u^2 + w^2 = V^2. \tag{5.60}$$

In the special case $\nu = 0$, we can use the indentities

$$\mathrm{J}_{-\nu}(u) = (-1)^\nu\,\mathrm{J}_\nu(u) \tag{5.61}$$

$$\mathrm{K}_{-\nu}(w) = \mathrm{K}_\nu(w) \tag{5.62}$$

to obtain the dispersion relation for the $\mathrm{LP}_{0\mu}$-modes,

$$\frac{u\,\mathrm{J}_1(u)}{\mathrm{J}_0(u)} = \frac{w\,\mathrm{K}_1(w)}{\mathrm{K}_0(w)}. \tag{5.63}$$

The relations (5.59) and (5.63) can again be solved numerically, see Fig. 5.5, where the numerical solution of Eq. (5.59) is plotted along with the circular arc that represents the additional condition $u^2+w^2 = V^2$, Eq. (5.60). The intersection points define the individual $\mathrm{LP}_{\nu\mu}$-modes. The first index $\nu = 0, 1, 2, 3\ldots$ relates to the azimuthal dependence of the dominant transverse field component: There are 2ν intensity maxima in azimuthal direction. The second index $\mu = 1, 2, 3, \ldots$ labels the various solutions for a given ν and relates to the radial field dependence: The number of intensity maxima in radial direction can be identified with μ where, for $\nu = 0$, the intensity maximum at $r = 0$ has to be counted as well. The corresponding field patterns are plotted in Fig. 5.6. Note that each $\mathrm{LP}_{0\mu}$-mode is twofold degenerate since it represents two orthogonal polarizations (dominant $\underline{\mathcal{E}}_x$- and dominant $\underline{\mathcal{E}}_y$-component). For $\nu > 1$, the $\mathrm{LP}_{\nu\mu}$-modes are four-fold degenerate: Two azimuthal dependencies $\cos\nu\varphi$, $\sin\nu\varphi$ (rotated against each other by 90 °) in two orthogonal linear polarizations each.

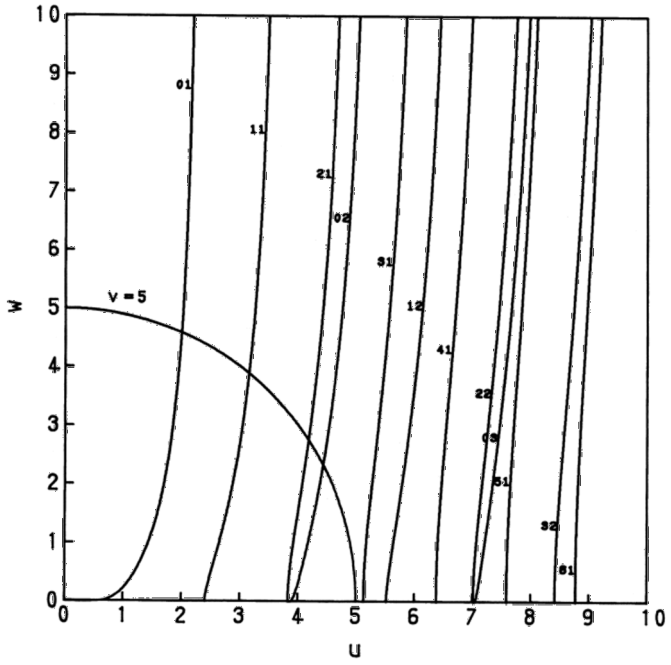

Figure 5.5: Graphical solution of the characteristic equation for LP-modes. The vertical branches correspond to the numerical solution of Eq. 5.59, whereas the circular arc represents the additional condition $u^2 + w^2 = V^2$, Eq. 5.60. (Adapted from [6])

Relationship between $LP_{\nu\mu}$-modes and hybrid, TE- and TM-modes

The linearly polarized $LP_{\nu\mu}$-modes can be constructed by appropriate superpositions of hybrid, TE- and TM-modes, see Fig. 5.7. The fundamental LP_{01}-mode is equivalent to the HE_{11}-mode of the step-index fiber. Both modes have no lower cut-off frequency and are hence always guided. The $LP_{1\mu}$-modes correspond to a superposition of the $HE_{2\mu}$-, $TE_{0\mu}$- and $TM_{0\mu}$-modes, which have similar dispersion characteristics, Fig. 5.4. For $\nu > 1$, the $LP_{\nu\mu}$-modes can be represented by a combination of $HE_{\nu+1,\mu}$- and $EH_{\nu-1,\mu}$-modes. The correspondence between $LP_{\nu\mu}$-modes and hybrid modes can hence be summarized as

$$
\begin{aligned}
&LP_{0\mu} \;\Leftrightarrow\; HE_{1\mu},\\
&LP_{1\mu} \;\Leftrightarrow\; HE_{2\mu},\ TE_{0\mu},\ TM_{0\mu},\\
&LP_{\nu\mu} \;\Leftrightarrow\; HE_{\nu+1,\mu},\ EH_{\nu-1,\mu} \text{ for } \nu > 1.
\end{aligned}
$$

Cut-off frequencies

Consider the behavior of a certain waveguide mode as the frequency is decreased. The waveguiding limit is achieved when the field extends into the whole cladding region, i.e., when $\beta \to n_2 k_0$. This implies:

$$w = a\sqrt{\beta^2 - n_2^2 k_0^2} \to 0, \tag{5.64}$$

$$u = \sqrt{V^2 - w^2} \to V, \tag{5.65}$$

$$w\frac{\mathrm{K}_{\nu-1}(w)}{\mathrm{K}_{\nu}(w)} \to 0. \tag{5.66}$$

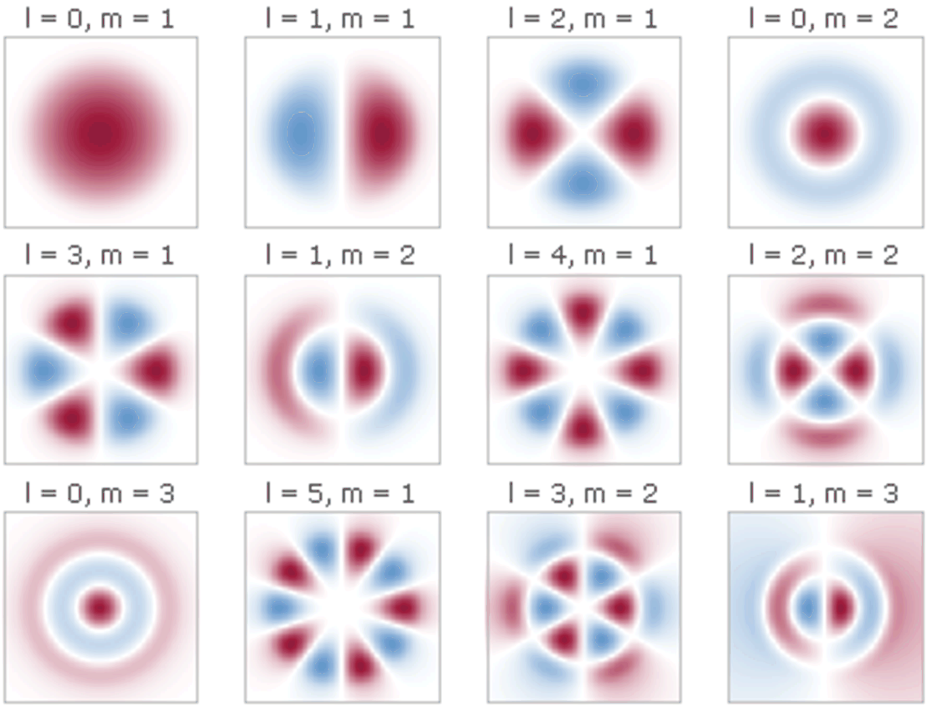

Figure 5.6: Mode fields of $LP_{\nu\mu}$-modes. The first mode index $\nu = 0, 1, 2, 3 \ldots$ (denoted as l in this figure) relates to the azimuthal dependence of the dominant transverse field component (2ν intensity maxima in azimuthal direction). The second index $\mu = 1, 2, 3, \ldots$ (denoted as m) gives the number of intensity maxima in radial direction, where, for $\nu = 0$, the intensity maximum at $r = 0$ has to be counted as well. (Adapted from http://www.rp-photonics.com/multimode_fibers.html)

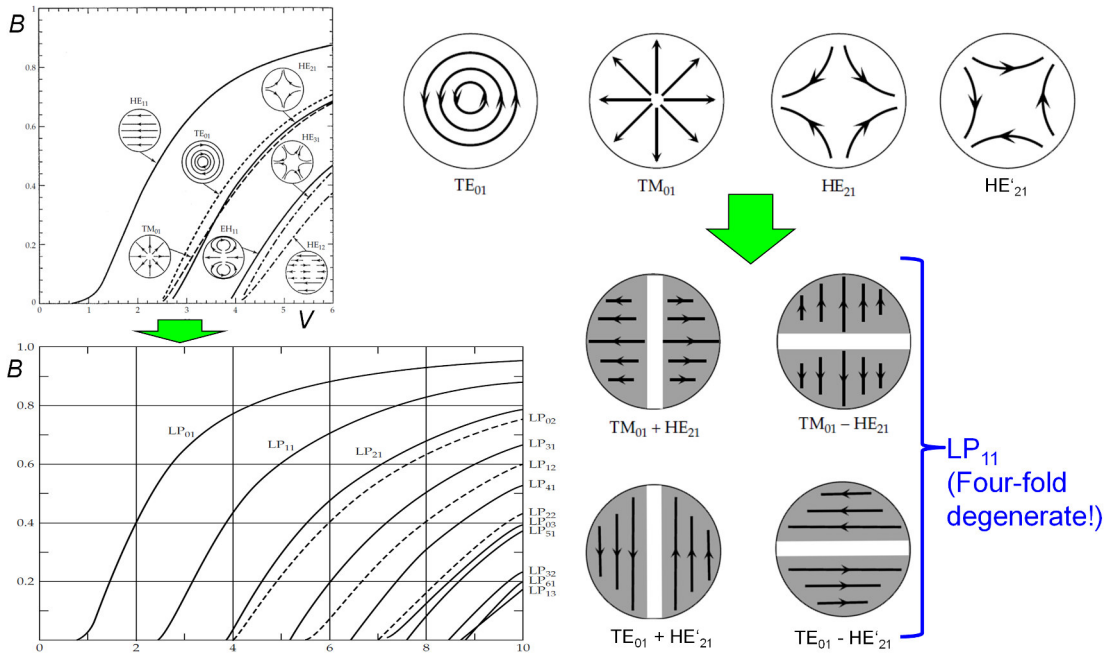

Figure 5.7: Linearly polarized (LP) and hybrid modes: The fundamental LP_{01}-mode is equivalent to the HE_{11}-mode. The LP_{11}-modes correspond to a superposition of the HE_{21}-, TE_{01}- and TM_{01}-modes. Note that the $\mathrm{LP}_{0\mu}$-modes and the $\mathrm{LP}_{2,\mu-1}$-modes have identical cut-off frequencies. (Adapted from [6])

From Eq. 5.63, we may hence conclude that $\mathrm{J}_{\nu-1}(V) = 0$ at cut-off. In other words: The normalized cut-off frequency $V_{\nu\mu,\mathrm{c}}$ of the $\mathrm{LP}_{\nu\mu}$-mode is given by the μ-th positive zero $j_{\nu-1,\mu}$ of the $(\nu-1)$-th order Bessel function $\mathrm{J}_{\nu-1}$,

$$V_{\nu\mu,\mathrm{c}} = j_{\nu-1,\mu}, \tag{5.67}$$

where $u = 0$ is counted as a zero of $\mathrm{J}_{\nu-1}(u)$ only for $\nu = 0$,

$$V_{01,\mathrm{c}} = j_{-1,1} = 0, \tag{5.68}$$

$$V_{0\mu,\mathrm{c}} = j_{-1,\mu} = j_{1,\mu-1} = V_{2,\mu-1,\mathrm{c}}. \tag{5.69}$$

This implies that the $\mathrm{LP}_{0\mu}$-modes and the $\mathrm{LP}_{2,\mu-1}$-modes have identical cut-off frequencies, see Fig. 5.7.

Number of modes in multimode fibers

If a multimode fiber is operated at a given normalized frequency V, the number of modes is proportional to the number of Bessel zeros $j_{\nu-1,\mu}$ that fulfill $j_{\nu-1,\mu} < V$, where $\nu = 0, 1, 2\ldots$ and $\mu = 1, 2\ldots$ For a coarse estimation of the number of guided modes, we use the fact that large zeros of the Bessel functions can be approximated by [1]

$$j_{\nu-1,\mu} \approx \left(\mu + \frac{1}{2}\nu - \frac{3}{4}\right)\pi < V. \tag{5.70}$$

In a $\nu-\mu-$diagram, the guided modes occupy a triangular area below the line defined by $j_{\nu-1,\mu} = \left(\mu + \frac{1}{2}\nu - \frac{3}{4}\right)\pi$, see Fig. 5.8. The majority of the points correspond to modes that are four-fold degenerate - each of them represents two azimuthal dependencies and two orthogonal linear

polarizations. The number of guided modes can hence be estimated by

$$M_g \approx 4 \cdot \frac{1}{2} \cdot 2 \left(\frac{V}{\pi} - \frac{1}{4} \right) \cdot \left(\frac{V}{\pi} - \frac{1}{4} \right) \approx \frac{4}{\pi^2} V^2 \approx \frac{V^2}{2}. \tag{5.71}$$

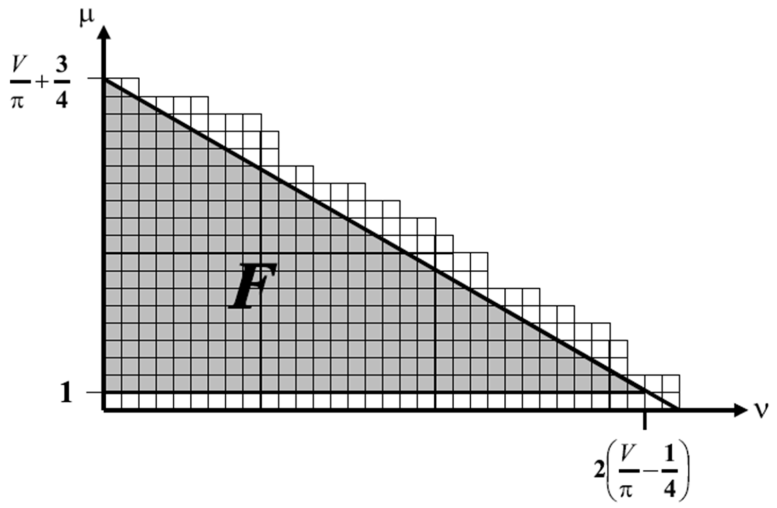

Figure 5.8: Number of guided $\mathrm{LP}_{\nu\mu}$-modes in a step-index fiber. The guided modes are represented by a triangular area below the line defined by $j_{\nu-1,\mu} = \left(\mu + \frac{1}{2}\nu - \frac{3}{4}\right)\pi$ where $\nu = 0, 1, 2, 3\ldots$ and $\mu = 1, 2, 3\ldots$. The majority of the points correspond to modes that are four-fold degenerate.

5.3 Graded-index fibers

In the last section, we have considered ideal step-index fibers with homogeneous core and cladding regions, for which closed-form expressions could be obtained. However, many fibers of practical interest do not have a step-index profile - this is either by design or due to the manufacturing process. So-called graded-index fibers ("GRIN-lenses") are designed to have a parabolic index profile, for which n^2 depends quadratically on the distance r from the fiber axis, i.e., $q = 2$ in Eq. (5.2). Other fibers are designed to have step-index profiles, but diffusion of dopants during the manufacturing process leads to deviations from the ideal index distribution. In most cases, however, the index profile can be assumed to vary monotonously in r and to be independent of the azimuthal variable φ. We can then use Eqs. (5.2) and (5.3) to represent $n^2(r)$.

In this section, we will only give a short outline of the analysis of graded-index fibers. We will sketch how to calculate the modes of parabolic-index profiles and how to extend the analysis to general power-law profiles. A more detailed analysis can be found in the lecture "Field Propagation and Coherence (FPC)" [9] and in textbooks [5].

5.3.1 Infinitely extended parabolic profile

The parabolic index profile is represented by an exponent of $q = 2$ in Eqs. (5.2) and (5.3). To enable a closed-form solution, we may further assume that the index profile is infinitely extended, i.e.,

$$n^2(r) = n_1^2 \left[1 - 2\Delta g(r/a)\right] \tag{5.72}$$

$$g\left(\frac{r}{a}\right) = \left(\frac{r}{a}\right)^2, \text{ in } 0 \le r < \infty \tag{5.73}$$

Strictly speaking, this is an unphysical assumption, because $n^2(r)$ would become negative for $r \to \infty$, but errors are negligible for modes that are confined to the core and do not extend far into the cladding. Under this assumption, closed-form solutions for the LP-modes can be calculated by a similar procedure as the one applied to step index fibers.

First we assume that the refractive index profile is only weakly inhomogeneous and that we may hence use the scalar form of the wave equations, for which the different Cartesian coordinates are decoupled, see Eqs. (2.24) and (2.25). In addition, we assume again that one transverse Cartesian electric field component is much stronger than the other electric field components. For the dominant field component, we again choose an ansatz of the form $\underline{\Psi}(\mathbf{r}) = \underline{\Psi}(r,\varphi)\exp(-\mathrm{j}\beta z)$, for which $\underline{\Psi}(r,\varphi)$ must fulfill the wave equation for mode fields,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\,\frac{\partial \underline{\Psi}(r,\varphi)}{\partial r}\right) + \frac{1}{r^2}\,\frac{\partial^2 \underline{\Psi}(r,\varphi)}{\partial \varphi^2} + \left(k_0^2 n^2(r) - \beta^2\right)\underline{\Psi}(r,\varphi) = 0, \tag{5.74}$$

This relation can be solved by a separation ansatz for which the radial and the angular dependencies are written as a product,

$$\underline{\Psi}(r,\varphi) = Q(r)\,\cos(\nu\varphi + \psi)\,, \tag{5.75}$$

where $\nu = 0, 1, 2, \ldots$ and $\psi \in \left\{0; \frac{\pi}{2}\right\}$. A differential equation for $Q(r)$ can be derived [31] . The solution yields the so-called Gauss-Laguerre modes

$$Q_{\nu\mu}(r) = \sqrt{\tfrac{2/n_1}{w_0^2\pi}}\sqrt{\tfrac{(\mu-1)!}{(\nu+\mu-1)!}}\,\left(\tfrac{2r^2}{w_0^2}\right)^{\nu/2}\mathrm{e}^{-r^2/w_0^2}\,\mathrm{L}_{\mu-1}^{(\nu)}\left(\tfrac{2r^2}{w_0^2}\right), \tag{5.76}$$

where the Gaussian field radius w_0 of the fundamental mode is given by

$$w_0^2 = \frac{a^2}{V/2} \tag{5.77}$$

$\mathrm{L}_{\bar\mu}^{(\nu)}$ denotes generalized, $\mathrm{L}_{\bar\mu}^{(0)} \equiv \mathrm{L}_{\bar\mu}$ ordinary Laguerre polynomials of degree $\bar\mu$ and order ν:

$$\begin{aligned}
&\mathrm{L}_{\bar\mu}^{(\nu)}(x) = \sum_{i=0}^{\bar\mu}\tbinom{\bar\mu+\nu}{\bar\mu-i}\tfrac{(-x)^i}{i!}, \quad \lim_{x\to\infty}\tfrac{\mathrm{L}_{\bar\mu}^{(\nu)}(x)}{x^{\bar\mu}} = \tfrac{(-1)^{\bar\mu}}{\bar\mu!}, \quad \nu,\bar\mu = 0,1,2,\ldots,\\
&\mathrm{L}_0^{(\nu)}(x) = 1, \qquad \mathrm{L}_1^{(\nu)}(x) = -x+\nu+1,\\
&\mathrm{L}_2^{(\nu)}(x) = \tfrac{1}{2}\left[x^2 - 2(\nu+2)x + (\nu+1)(\nu+2)\right].
\end{aligned} \tag{5.78}$$

The propagation constants of the various modes can be calculated analytically,

$$\beta = k_1\sqrt{1-2\delta}, \quad \frac{\delta}{\Delta} = \frac{m}{m_{\max}}, \quad m = \nu + 2\mu - 1, \quad m_{\max} = \frac{V}{2}; \quad w_0^2 = \frac{a^2}{V/2}. \tag{5.79}$$

where the normalized index difference Δ is given by

$$\Delta = \frac{n_1^2 - n_2^2}{2n_1^2} \approx \frac{n_1 - n_2}{n_1}. \tag{5.80}$$

The fundamental mode has the transverse dependency of a rotationally symmetric Gaussian with field radius w_0. All modes with principal mode number $m = \text{const}$ belong to the same propagation constant β and are therefore degenerate. As mentioned, the solutions of Eq. (5.76) describe the modes of an infinitely extended parabolic fiber, see Eqs. (5.72) and (5.73). They are equivalent to the real (finite) index profile as long as the resulting fields do not reach significantly into the cladding beyond the physical core-cladding interface at $r = a$. This is true for most of the modes if $V \gg V_{\nu\mu G}$.

A few measured mode fields are shown in Fig. 5.9. The mode index ν is related to the azimuthal field dependence $\cos(\nu\varphi)$ or $\sin(\nu\varphi)$, where 2ν is the number of intensity maxima along the circumference of the core. The mode index μ denotes the various solutions for a given ν and is equal to the number of intensity maxima in radial direction, where, for $\nu = 0$, the maximum at $r = 0$ has to be counted as well.

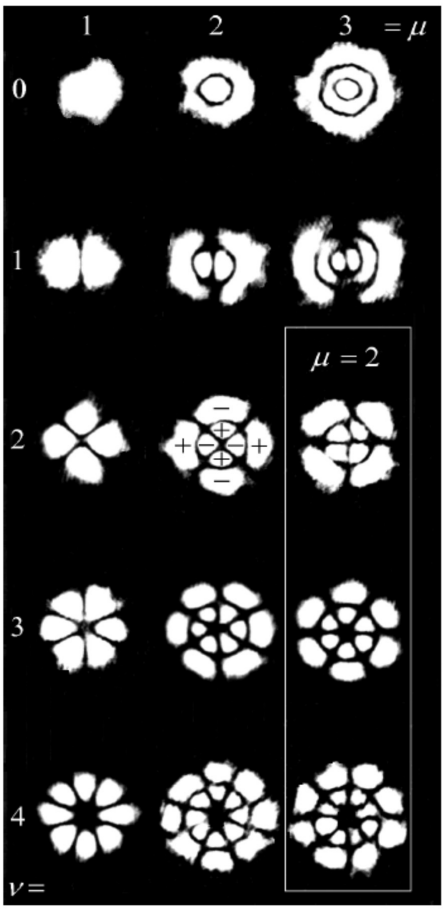

Figure 5.9: Gauss-Laguerre mode fields of a parabolic-index fiber ($a = 23\,\mu$m; $A_N = 0.2$; $V = 46$; $\lambda = 0.6328\,\mu$m). The mode index ν is related to the azimuthal field dependence $\cos(\nu\varphi)$ or $\sin(\nu\varphi)$, where 2ν is the number of intensity maxima along the circumference of the core. The mode index μ beam denotes the various solutions for a given ν and is equal to the number of intensity maxima in radial direction (including the maximum at $r = 0$ for $\nu = 0$). For this specific fiber, the radius of the fundamental mode ($\nu = 0$, $\mu = 1$) is $4.8\,\mu$m.

Number of guided modes

The cutoff frequency $V_{\nu\mu G}$ for the mode with index (ν, μ) may be estimated by observing that $\beta = k_2$, i.e., $\delta = \Delta$ and $m = m_{max} = \frac{V}{2}$, at cutoff, see Eq. (5.79). Guided modes must hence fulfill the relation

$$\nu + 2\mu - 1 < \frac{V}{2}. \tag{5.81}$$

This corresponds again to a triangle in the (ν, μ)-plane with corner points $\left(\frac{V}{2} - 1, 0\right)$ and $\left(0, \frac{V}{4} + \frac{1}{2}\right)$. Within this rectangle, most of the pairs (ν, μ) correspond to a four-fold degenerate mode (two

polarizations, $\cos(\nu\varphi)$- and $\sin(\nu\varphi)$-dependence). The number of guided modes can hence be estimated to be

$$M_g \approx 4 \cdot \frac{1}{2} \cdot \left(\frac{V}{2} - 1\right) \cdot \left(\frac{V}{4} + \frac{1}{2}\right) \approx \frac{V^2}{4} \tag{5.82}$$

Comparing Eqs. (5.71) and (5.82), we find that a parabolic index profile accepts approximately only half the number of guided modes as a step-index fiber having the same relative refractive index difference Δ.

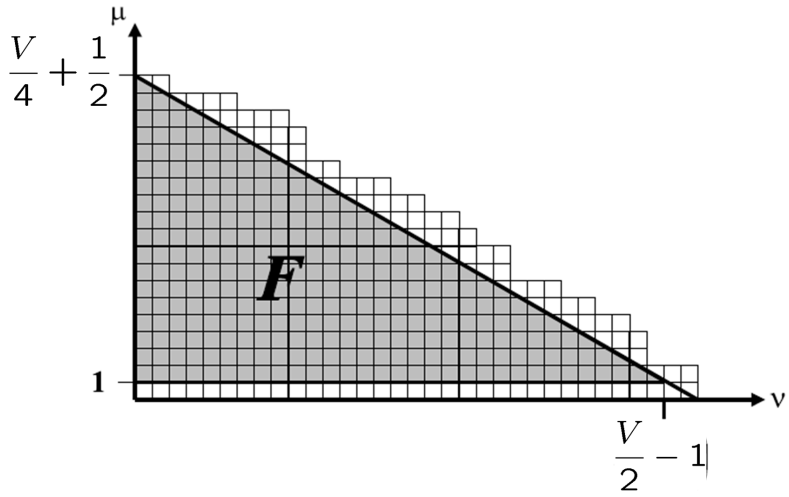

Figure 5.10: Number of guided modes in an infinitely extended parabolic index profile: The guided modes are represented by a triangular area below the line defined by $\nu+2\mu-1 = V/2$. The majority of the points correspond to modes that are four-fold degenerate.

5.3.2 General power-law index profiles

For arbitrary values of index profile exponent q, it is not possible to express the mode fields as closed-form solutions. Instead, a semi-analytical approach based on a power series approximation is usually used to represent the dominant transverse mode field component $\underline{\Psi}(r,\varphi)$ within the core region $r < a$,

$$\underline{\Psi}(r,\varphi) = A\left(\frac{r}{a}\right)^{\nu} \sum_{i=0}^{\infty} c_i \left(\frac{r}{a}\right)^{i} \cos(\nu\varphi + \psi), \tag{5.83}$$

where $\nu = 0, 1, 2, \ldots$ and $\psi \in \left\{0; \frac{\pi}{2}\right\}$. In the homogeneous cladding region, the field is represented by the corresponding modified Bessel function $K_\nu(w)$ of order ν. This leads to a recurrence relation for the coefficients c_i and to a dispersion relation of the form

$$\nu + \frac{\sum_{i=0}^{\infty} i c_i}{\sum_{i=0}^{\infty} c_i} = \frac{w K_\nu'(w)}{K_\nu(w)}, \tag{5.84}$$

where $w = a\sqrt{\beta^2 - n_2^2 k_0^2}$. A more detailed description of the analysis can be found in [5].

5.4 Microstructured fibers

In traditional optical fibers, light is guided by total internal reflection in a fiber core whose refractive index is larger than that of the cladding area. In optical communications, the vast majority

of fibers are built according to this concept. There are, however, other fiber structures in which guidance of light relies on internal micro- and nanostructures. This enables dedicated fiber designs for special applications. In this section, we will give a short overview on the underlying concepts.

Hole-assisted fibers

A large refractive index contrast between the core and the cladding region can be achieved if air holes are used in the cladding region of the fiber. This allows for tight confinement of light to the core and is often used in so-called highly nonlinear fibers (HNLF), see Fig. 5.11 (a) and (b). In these fibers, large optical intensities lead to nonlinear interaction and thereby to the generation of new spectral components. When pumped with pulsed laser light, HNLF can be used for supercontinuum generation, Fig. 5.11 (c).

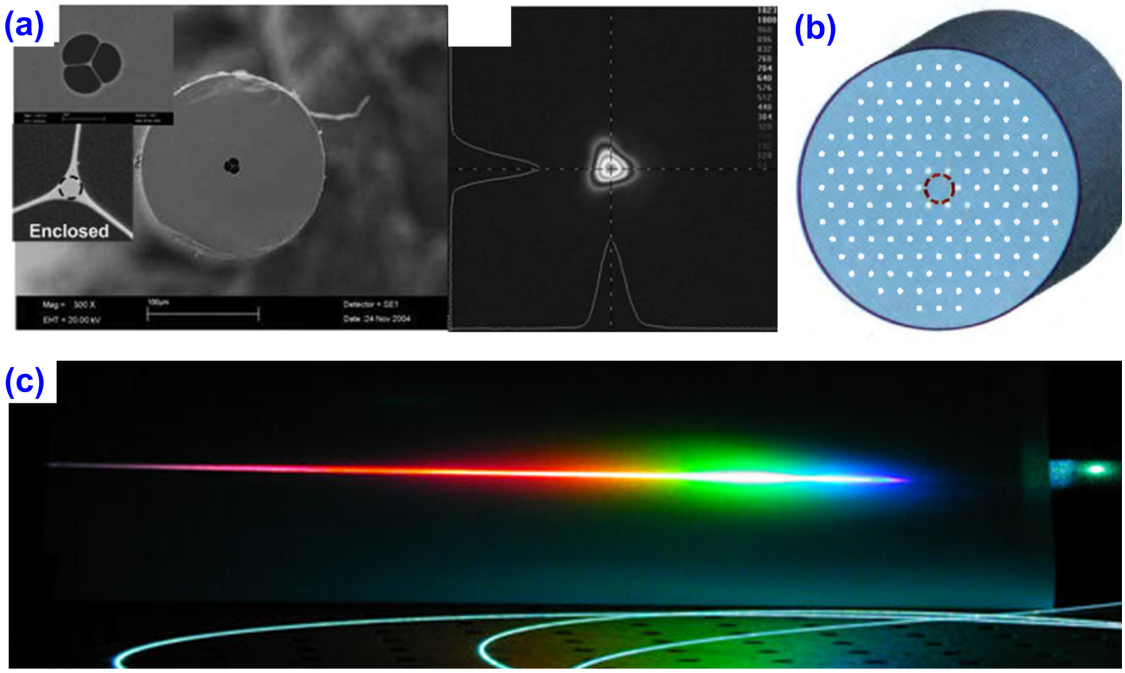

Figure 5.11: Highly-nonlinear microstructured fibers; (a) High index-contrast small-core fiber consisting of a small triangular core which is held by two suspended membranes. Tight confinement results from the fact that the high-index glass-core is predominantly surrounded by air. (Figure adapted from [20]) (b) Sketch of a hole-assisted fiber ("holey fiber") consisting of a solid core and a cladding containing a large volume fraction of air holes. The diameter of the holes can be much smaller than the wavelength such that the average refractive index of the core is lower than that of the cladding (Figure adapted from [30]) (c) Supercontinuum generation using a microstructured fiber: The strong confinement of light and the large interaction lengths lead to strong third-order nonlinear interaction. High-power laser pulses launched into the fiber at one end are spectrally broadened by a combination of four-wave mixing (FWM) and self-phase modulation (SPM). (Figure adapted from http://www.bath.ac.uk/physics/groups/cppm/nonlinear_pcf.php)

Bragg fibers

In so-called Bragg fibers, concentric rings of high- and low-index materials are arranged around a central core, see Fig. 5.12 (a). At each dielectric interface, a small part of the light is reflected. If the layer thicknesses are chosen appropriately, constructive interference occurs for the back-

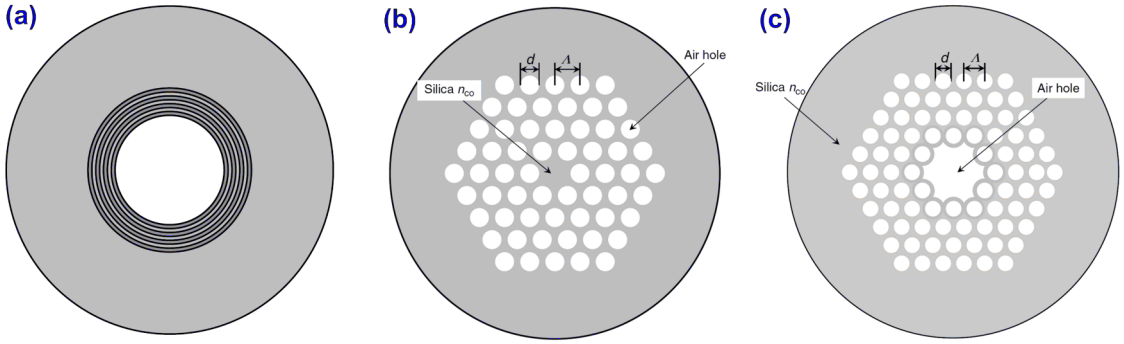

Figure 5.12: Photonic-bandgap fibers. (a) Bragg fiber, consisting of concentric rings of high- and low-index materials that are arranged around a central core. The core may be solid or hollow (hollow-core fiber). (b) Photonic-crystal fiber with solid core. The cladding consists of a regular pattern of air holes. Multiple reflections from these air holes form a photonic bandgap. (c) Hollow-core photonic crystal fiber. The existence of a photonic bandgap in the cladding can lead to confinement of light in a low-index hollow core region. (Figures adapted from [25])

reflected portions, and light cannot propagate in the cladding region[3]. Light injected in one end of the Bragg fiber propagates along the fiber axis even if the refractive index of the waveguide core is lower than that of the cladding. For some applications, it is even beneficial to use a hollow waveguide core, e.g., for fibers that carry large optical powers as in laser-based material processing.

Bragg fibers are a special type of photonic-bandgap fibers, for which light propagation in the cladding is prohibited by multiple reflections from periodically arranged objects. In lieu of alternating layers, a regular pattern of air holes can also be used to form a photonic bandgap. In this case, the fiber is referred to as photonic-crystal fiber, see next section.

Photonic-crystal fibers

Just like in Bragg fibers, where a one-dimensional transverse periodicity of the refractive index profile leads to the confinement of light to the waveguide core, a two-dimensional arrangement of air holes can also exhibit a photonic band gap. These fibers are referred to as photonic crystal fibers (PCF) or holey fibers, Fig. 5.12 (b) and (c). PCF can have solid or hollow core regions. Hollow-core photonic crystal fibers are often used for transmission of ultra-short high-power pulses or for sensing and spectroscopy applications.

Mode analysis of microstructured fibers

The index profiles of microstructured fibers are not necessarily circular symmetric, see Fig. 5.12 and 5.11. As a consequence, numerical methods are generally needed to calculate the mode fields. Such methods have been discussed in Section (4.4). An example of a numerical mode calculation is shown in Figs. 5.13 and 5.14. Two modes are obtained - an E_x-mode and an E_y-mode. The two modes are degenerate due to symmetry of the waveguide, i.e., they have the same effective refractive index ($n_e = 1.4455$), and any linear combination of the modes can again be considered as a mode field, which does not change its shape during propagation along z.

[3]The phenomenon is also referred to as a photonic bandgap in analogy to semiconductor physics: In a semiconductor, the interaction of the electron probability density functions with regularly arranged atoms leads to the formation of forbidden energy bands for electrons. In a so-called photonic crystal, a regular arrangement of microstructures leads to the formation of forbidden energies bands for photons. The corresponding wavelengths are reflected from the surface of the structure and can not penetrate the photonic crystal. See, e.g., ref. [15] for more details on photonic crystals.

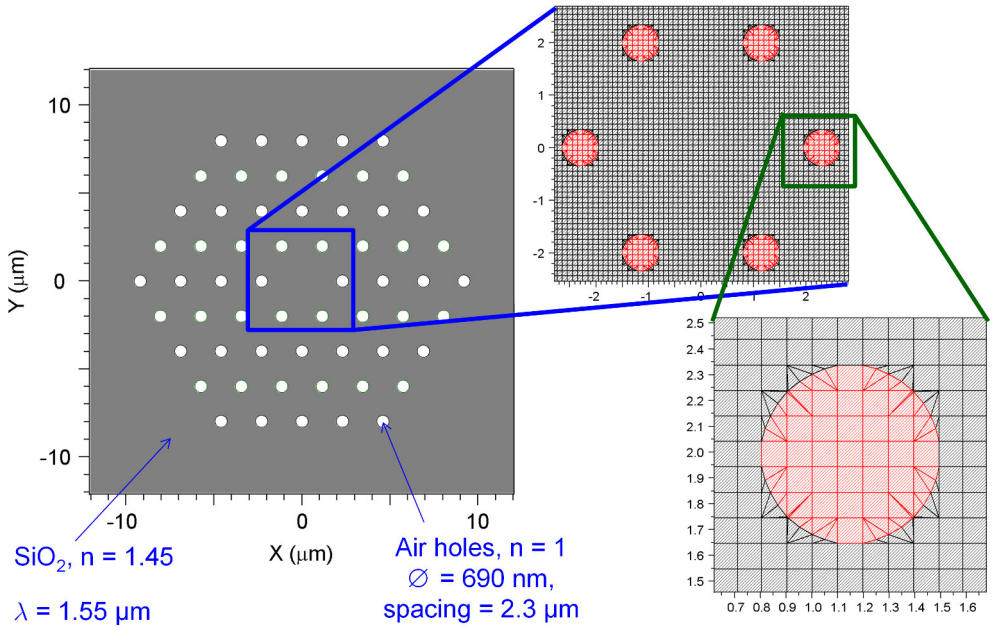

Figure 5.13: Numerical mode calculation of a photonic-crystal fiber. The structure is represented by finite elements, the which can be adapted to the shape of the of the air holes.

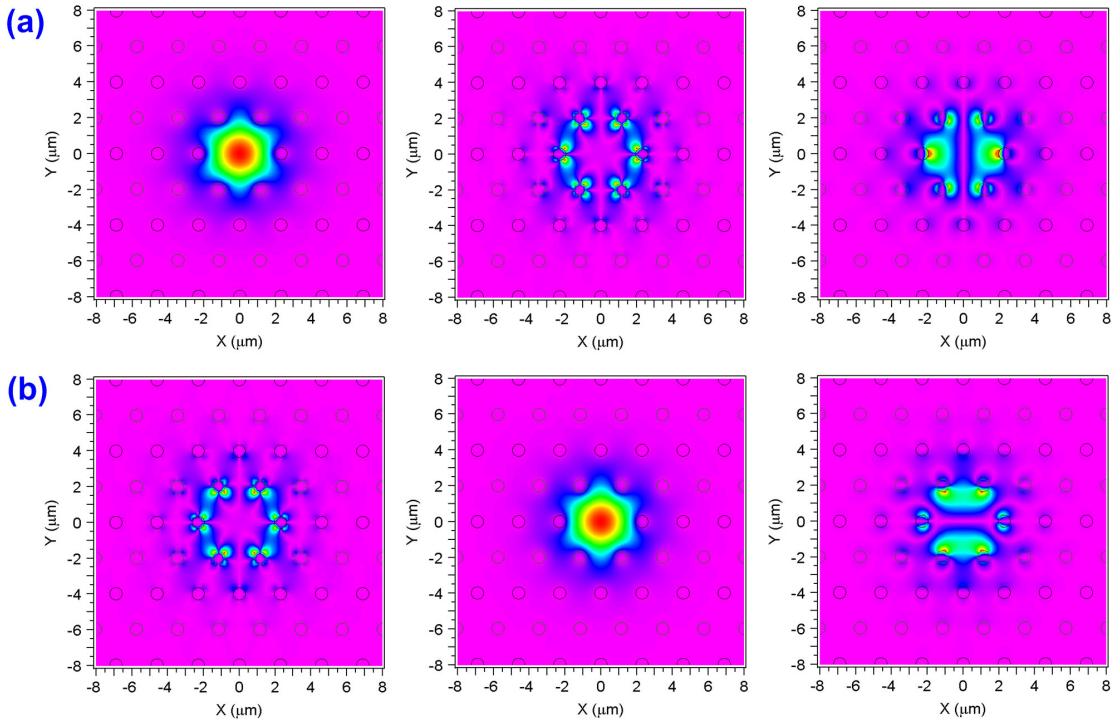

Figure 5.14: Fundamental modes of the photonic-crystal fiber as obtained from a finite-element calculation. (a) E_x-modes. (b) E_y-modes. Due to symmetry, the two modes are degenerate, i.e., they have the same effective refractive index ($n_e = 1.4455$), and any linear combination of the modes can again be considered as a mode field, which does not change its shape during propagation along z.

5.5 Fiber technologies and fabrication methods

In this section, we will give a short overview on the prevalent fiber technologies and the corresponding fabrication methods. We will first introduce the basic principle of fiber drawing and then discuss some important material systems. More detailed information can be found in textbooks such as [11], [10], or [32].

5.5.1 Fabrication of optical fibers

Fiber fabrication techniques can be divided in two basic categories: One-stage processes, where the fiber is directly drawn from the liquefied base material, and two-stage processes, where a solid preform rod is fabricated and then drawn into a thin thread, see Fig. 5.15.

One-stage processes

One-stage techniques are clearly simpler than their two-stage counterparts, and they allow for fast fabrication and unlimited fiber lengths. As an example of a one-stage process, we consider the so-called double-crucible method. This technique is often used to fabricate fibers from multicomponent glasses with low melting points, which are sometimes also referred to as "soft glasses".

A sketch of the double-crucible method is depicted in Fig. 5.15 (a). It comprises two heated crucibles with a nozzle at the bottom. The inner crucible contains a melt of high-index core material, whereas the outer one is filled with molten low-index cladding material. The melts are combined in the nozzle, where the molten core material is exposed to the molten cladding. Diffusion processes at the interface lead to graded-index profiles, which can be controlled by the temperature in the diffusion region. The fiber diameter is continuously measured during the drawing process and controlled by adjusting the drawing speed. The fiber is coated by resin which is cured by UV radiation in a confocal reflector unit. Drawing speed is controlled by a so-called capstan[4] drive, and the fiber is finally coiled onto a bobbin.

A similar method can be used to fabricate so-called plastic-clad silica fibers or polymer-clad silica fibers (PCS), i.e., optical fibers that have a silica-based core and a plastic cladding. The core of the fiber is then drawn from the crucible of molten glass and directly coated with a polymer cladding. This results in a simple and low-cost fabrication process but high transmission loss, such that PCS fibers are only useful for short-distance transmission.

For fabricating plastic optical fibers, simple extrusion methods can be used in a one-stage process, which is very similar to the double-crucible method. This allows for cost-efficient mass production, but the resulting fibers do not reach top-level performance in terms of uniformity and propagation loss and are only used in short-distance communications.

Two-stage processes

Two-stage processes, in contrast, are more complex and allow for drawing of finite fiber sections only. However, they allow for better control of material composition, purity and fiber geometry and are hence used for fabrication of low-loss single-mode silica fibers. An apparatus for drawing of optical fibers in a two-stage process is depicted in Fig. 5.15 (b). Drawing starts from a solid preform rod, which has a cross-sectional profile similar to that of the fiber. The lower end of the preform is heated by a graphite filament heater. For silica glasses, typical temperatures amount to 2000 - 2100 °C. A protective Ar gas atmosphere is used to prevent oxidation of the filament. The fiber is then drawn from the molten end of the preform in a similar way as in the double-crucible method. The fabrication technique for the preform rod depends strongly on the involved materials and will be discussed in more detail in the following sections.

[4] A capstan drive usually consist of a rotating spindle which is in contact with one or more (rubber) wheels, so-called pinch rollers. Capstan drives were originally developed to move tapes in various kinds of tape recorders.

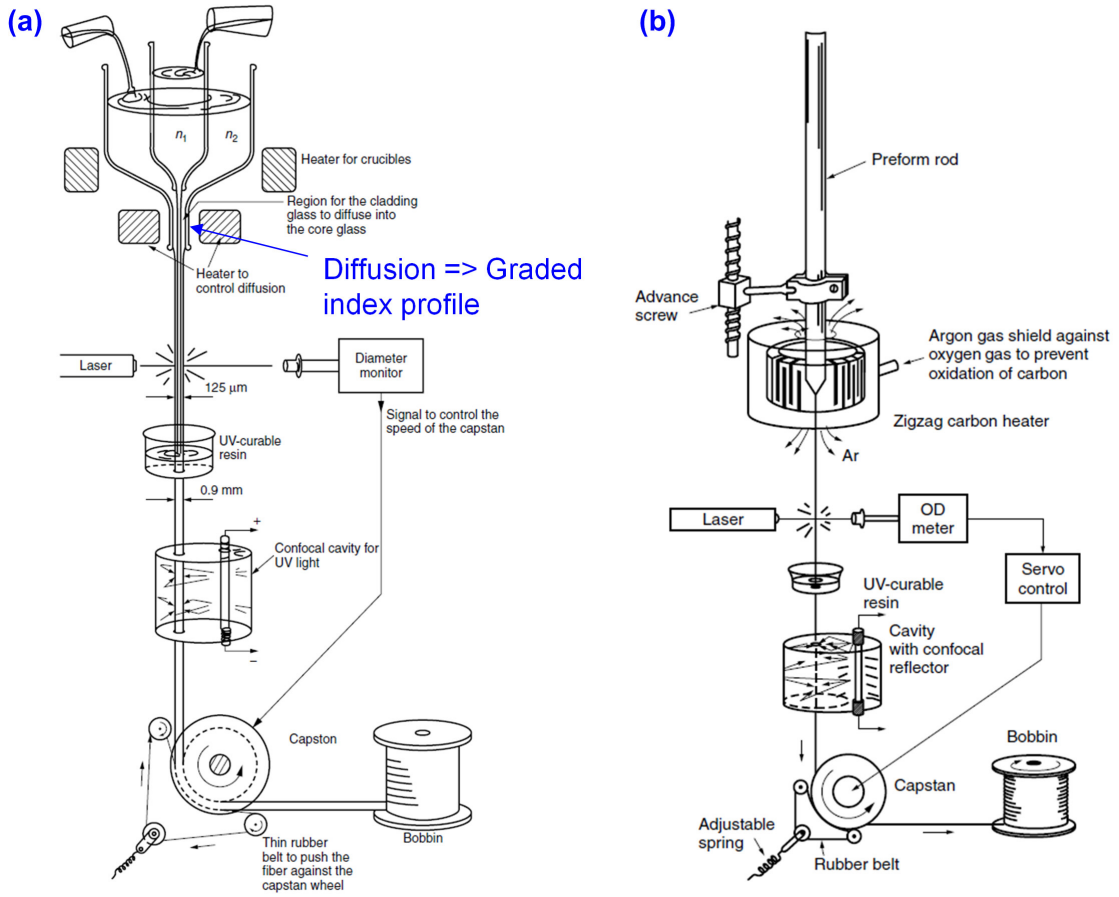

Figure 5.15: Drawing of optical fibers. (a) One-stage process: The fiber is directly drawn from liquefied base materials, which are provided by a set of two crucibles (double-crucible method). (b) Two-stage process: The fiber is drawn from a solid preform, which can, e.g., be fabricated by chemical vapor deposition (CVD), see Section 5.5.2. (Figure adapted from [11])

5.5.2 Fiber technologies

Glass is by far the most common material used for optical fibers, but not all glasses are useful for fiber fabrication. Glasses in general are noncrystalline solids, in which the molecules are arranged randomly. Glasses can be thought of as liquids with the molecules frozen in place by very fast cooling. The most prevalent type of glass in everyday's life is the so-called soda-lime-silica glass (German: Kalk-Natron-Glas) which consists mainly of SiO_2 (71-75%), NaO_2 (12-16%), and CaO (10-15%) and is used for window panes, bottles, jars etc. However, due to high absorption, soda-lime-silica glass is not usable for optical fibers. Instead, special optical glass compositions or highly pure silica glass with lower attenuation is used.

Fibers made from optical glasses

Optical glasses are available in different compositions covering a large range of refractive indices ranging between 1.44 and 1.8 at visible wavelengths[5]. Attenuation coefficients typically amount to 1 dB/m, which is too high for long-distance optical communications, but sufficiently low for, e.g., fiber bundles for illumination and image-transmission in endoscopic applications.

Glass fibers are either drawn from solid preforms, which are usually fabricated by the so-called rod-in-tube method, see Fig. 5.16. In this technique, a high-index glass rod is inserted into a low-index jacket tube, which is then collapsed around the rod by heating. To minimize scattering loss, the core cladding interface must be very smooth and clean. This is usually achieved by fire-polishing. There are also other methods where the molten core glass is poured into the cladding tube or sucked into the tube using a vacuum pump. Alternatively, soft-glass fibers can be directly drawn from liquefied base materials in a one-stage process, see Section 5.5.1.

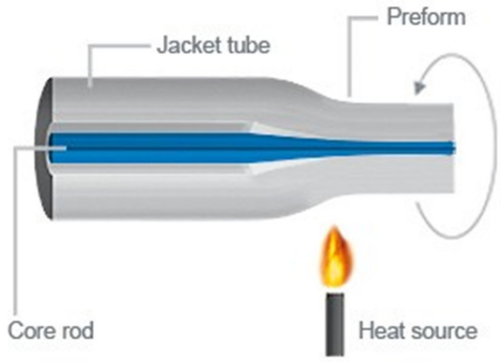

Figure 5.16: Rod-in-tube technique: A core rod of glass with high refractive index is inserted into a glass tube with lower refractive index, and the two pieces are fused together. (Adapted from http://heraeus-quarzglas.com)

Low-loss silica glass fibers

Transmission fibers for long-distance optical communications consist of highly-pure fused silica (amorphous SiO_2, German: Quarzglas). Even smallest impurities of metal ions such as Cr, Mn, V, Fe or Cu lead to considerable absorption loss such that purity grades of the order of 0.1 ppb are needed. At the same time, a highly uniform fiber geometry is essential to keep scattering losses small. This can only be achieved in an advanced two-stage fiber drawing process. State-of-the-art silica fibers have propagation losses of approximately 0.2 dB/km.

[5] See, e.g., the Schott glass catalogue, http://www.schott.com/advanced_optics/english/tools_downloads/download/index.html#Optical%20Glass.

The highly pure preform material is made synthetically by oxidizing silicon tetrachloride ($SiCl_4$) in a chemical vapor deposition (CVD) process yielding highly pure small glass particles, see Fig. 5.17. This so-called silica soot can then be melted into a compact preform. To increase or decrease the refractive index, small concentrations of dopants are added. Adding small amounts of, e.g., $GeCl_4$ or $POCl_3$ to the gas mixture yields GeO_2 or P_2O_2 and leads to an increased refractive index. In a similar way, the addition of Flourine (F) leads to a decreased refractive index. To minimize the impurity of the deposited materials, the CVD processes exploits the fact that the vapor pressure of liquid $SiCl_4$ and $GeCl_4$ is much higher than that of impurity metal halides such as VCl_4 and $FeCl_3$, see Fig. 5.18. Hence, if $SiCl_4$ and $GeCl_4$ are evaporated at, e.g., 1 atm pressure and 65 °C, VCl_4 and $FeCl_3$ stay in the liquid and are not deposited on the preform.

To fabricate the preform out of the material gases, three basic methods can be used. In the **inside vapor deposition method** (IVD), the material vapors are mixed with purified oxygen and led into a silica tube which is heated from outside, e.g., by an oxyhydrogen flame, see Fig. 5.17 (a). The material vapors are oxidized and form the silica soot, which is fused to the glass wall as the flame passes. Instead of a flame, a microwave plasma can also be used to heat the tube and to invoke oxidation of the material gases; this technique is also referred to as plasma-activated chemical vapor deposition (PCVD) .

In contrast to that, the **outside vapor deposition** (OVD) deposits the soot on a rotating rod of carbon or ceramics by injecting the material gases directly into the oxyhydrogen flame, Fig. 5.17 (b). After removing the starting rod, the soot is fused into a compact preform in a furnace. In the OVD technique, the deposited soot contains OH^--ions, which would lead to increased fiber loss and must be removed in a drying process by flushing the preform with Cl_2 at elevated temperatures before fusing it into a preform. Still, outside vapour deposition suffers from the fact that the starting rod must be removed, which may lead to irregularities in the refractive index profile of the core.

Similarly, the material gases are injected into the torch in the **vapor axial deposition** (VAD) technique, see Fig. 5.17 (c). The preform is rotated and slowly pulled upwards as the glass soot is deposited. As in OVD, OH^--ions must be removed from the the deposited soot, and the soot root must be fused into a solid preform. Multilayer structures can be achieved by a suitable arrangement of torches and careful adjustment of the proportion of material gases. In contrast to IVD and OVD, VAD allows for preforms that can in principle be infinitely long.

Photonic-crystal and microstructured fibers

Photonic-crystal and microstructured fibers are drawn from solid preforms. The preforms are fabricated from a stack of hollow glass tubes and rods, which is then enclosed into an outer tube, see Fig. 5.19. Transmission losses are high when compared to conventional silica glass fibers, such that microstructured fibers are mainly used for niche applications such as high-power transmission, optical sensing or supercontinuum generation.

Plastic optical fibers (POF)

Compared to glass fibers, polymer fibers show a poor performance in terms of propagation loss. As in polymer waveguides (see Section 4.5.3) the main source of losses in plastic optical fibers is absorption due to overtones of C-H-bond oscillations, see Fig. 4.18. Losses can be decreased if hydrogen is substituted by deuterium (heavy hydrogen), chlorine or fluorine since the larger atom mass of these materials shifts oscillations further into the IR. The fundamental oscillation of the C-H-bond lies near 3390 nm, for deuterium it is 4484 nm and for Fluorine, the oscillation can be shifted to 8000 nm.

Nevertheless, best laboratory POF still have losses of around 50 dB/km at best, and attenuation of commercially available plastic fibers amounts to typically more than 100 dB/km, even at visible wavelengths. Despite their low fabrication costs, polymer fibers can hence only be used for a limited range of applications comprising, e.g., image transmission bundles or short-reach data links in industrial or automotive applications, optical interconnects, or home installations.

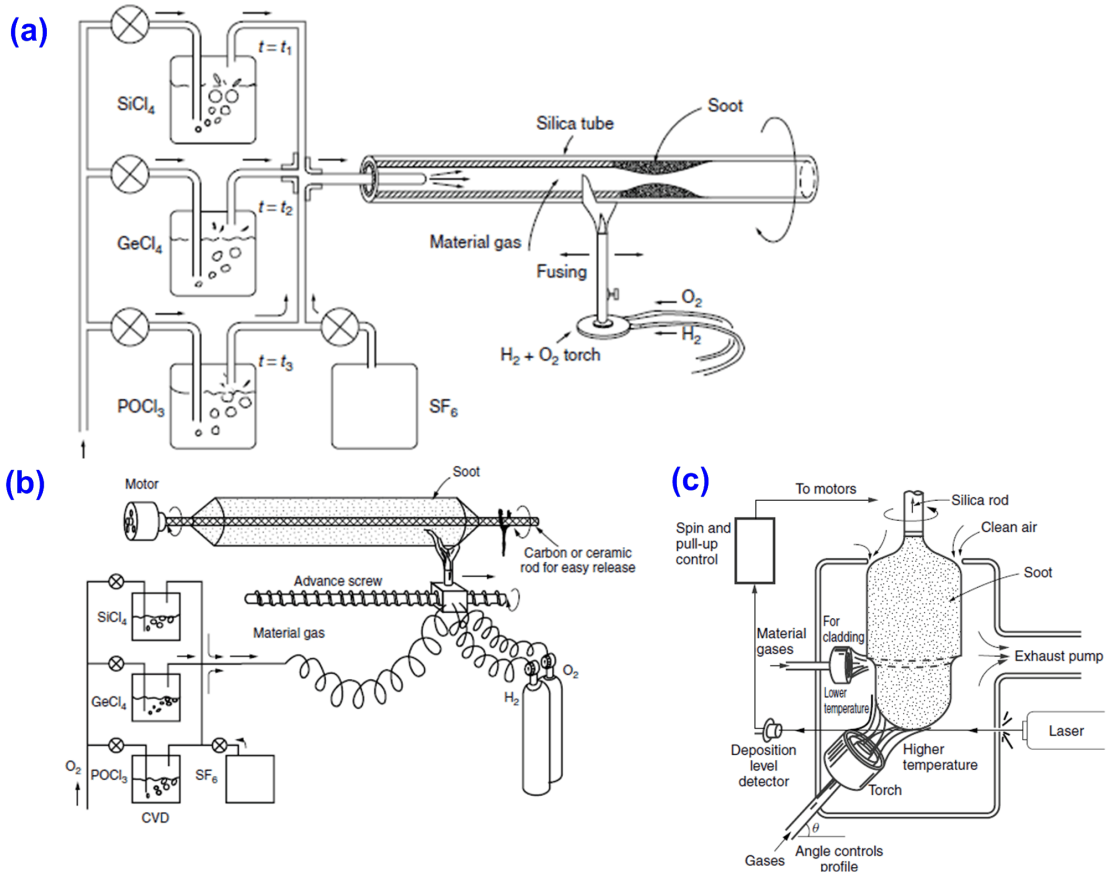

Figure 5.17: Silica preform fabrication using chemical vapor deposition (CVD) techniques. Highly pure silica is made synthetically by oxidizing silicon tetrachloride ($SiCl_4$) in a CVD process. This yields highly pure small glass particles, so-called silica soot, which is fused into a compact preform. To increase or decrease the refractive index, small concentrations of dopants are added. (a) Inside vapor deposition (IVD); (b) Outside vapor deposition (OVD); (c) Vapor axial deposition (VAD); (Figures adapted from [11])

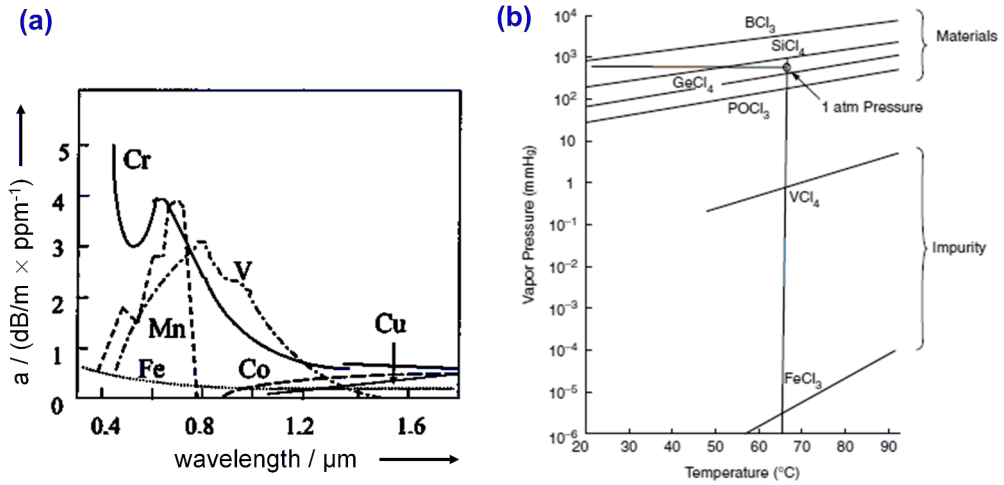

Figure 5.18: Absorption loss due to metal impurities and fabrication of low-loss silica glass. (a) Absorption losses caused by metal impurities in silica glass amount to typically a few dB/m for purity grades in the ppm range. To achieve losses below 1 dB/km, purity grades of the order of 0.1 ppb are needed (Figure adapted from [10]). (b) Synthesis of highly pure silica glass: The vapor pressures of impurity-metal halides such as VCl_4 or $FeCl_4$ are significantly lower than that of fiber materials and dopants (BCl_4, $SiCl_4$, $GeCl_4$, $POCl_3$). Evaporation of, e.g., $SiCl_4$ at 65° and normal pressure leaves unwanted VCl_4 or $FeCl_4$ in the liquid phase and allows to synthesize highly pure SiO_2. (Figure adapted from [11])

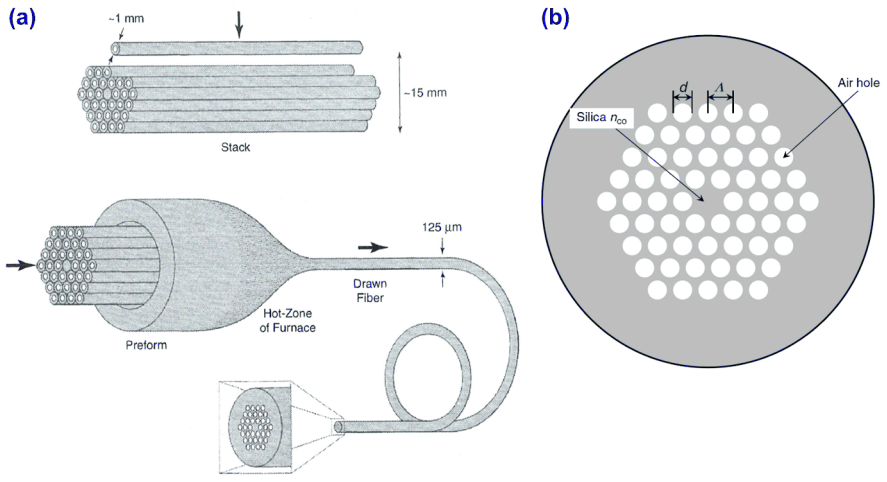

Figure 5.19: Fabrication of microstructured fibers. (a) The preform is obtained from a stack of hollow glass tubes and rods, enclosed by an outer tube. The axial holes of the preform are transferred to the fiber during the drawing process. (b) Fiber cross section. (Figures adapted from [10])

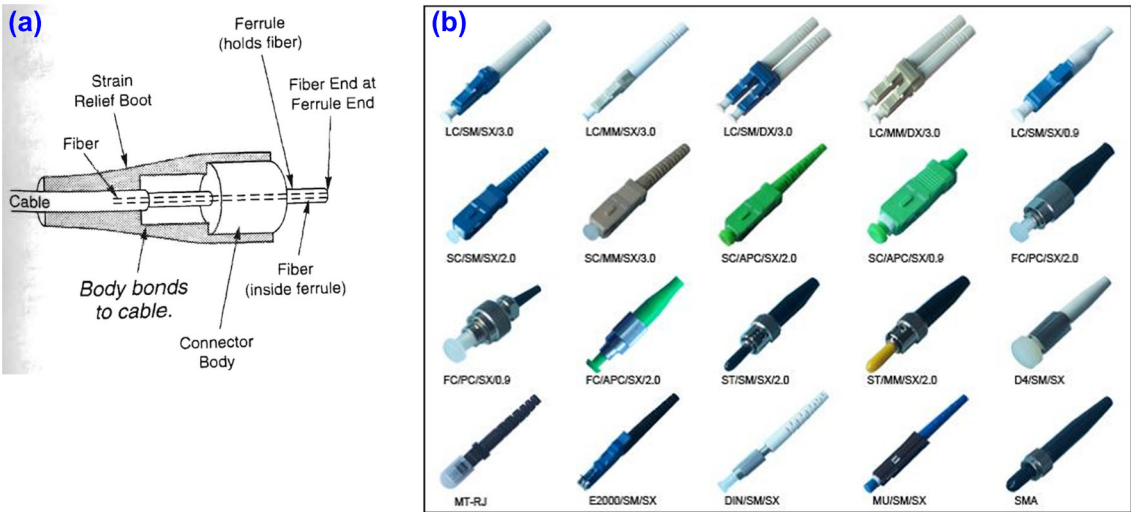

Figure 5.20: (a) "Generic-ferrule" fiber connector: The fiber is held in the center of the ferrule, the outer diameter is precisely matched to the inner diameter of the guide sleeve of the connector. The connector body is directly attached to the outer tube of the fiber cable to provide strain relief. (Figure adapted from [10]) (b) Different kinds of fiber connectors. Green outer plastic sleeves indicate fibers with angled physical contact (APC), i.e., the end of the ferrule is polished at an 8°-angle to avoid reflections from an open fiber end. To avoid damage of fiber facets, it should never be tried to plug an APC connector (e.g., FC/APC/SX/2.0) to a PC connector (e.g., FC/PC/SX/2.0) (Figure adapted from http://www.lotuscope.com).

Polymer fibers can be fabricated either in a two-stage process by drawing from a solid preform, or by a one-stage extrusion process [32]. The latter allows for cheap mass production of plastic optical fibers.

5.5.3 Fiber-optic connectors and cables

Fiber optic connectors must enable reproducible low-loss connections of fibers with low back-reflection. In the case of single-mode fibers with core diameters below $10\,\mu$m, this requires quite severe machining tolerances. An overview of common fiber connectors is depicted in Fig. 5.20. To avoid back reflections from the fiber facet, some of them have the ferrule end-polished at an 8°-angle. These fibers are referred to as angled physical contact (APC) fibers, in contrast to so-called physical-contact (PC) fibers with a flat facet. APC fibers are usually indicated by a green plastic sleeve, whereas PC fibers usually have black sleeves. To avoid damage of fiber facets, it should never be tried to connect an APC to a PC fiber.

For practical applications, optical fibers need to be bundled and/or protected against environmental influences. This is done by different fiber cables, an overview of which is shown in Fig. 5.21. When used within protective plastic tubes, optical fibers are often embedded into a jelly which allows the fiber to move freely within the tube. This avoids microbending and strain of the fiber which could, e.g., be caused by differences in thermal expansion coefficients of the fiber and of the tube material. When combining several fibers to multi-fiber cables or ribbons, so-called strength members are integrated to avoid stretching of the cable. In addition, the individual fibers are often stranded around a central steel wire. Stranding provides flexible stretching of the cable (similar to a telephone cord). Multiple fiber cables can be combined to multi-unit fiber cables which may contain additional strands such as power supply wires and gas pipes for keeping moisture off the inside of the cable.

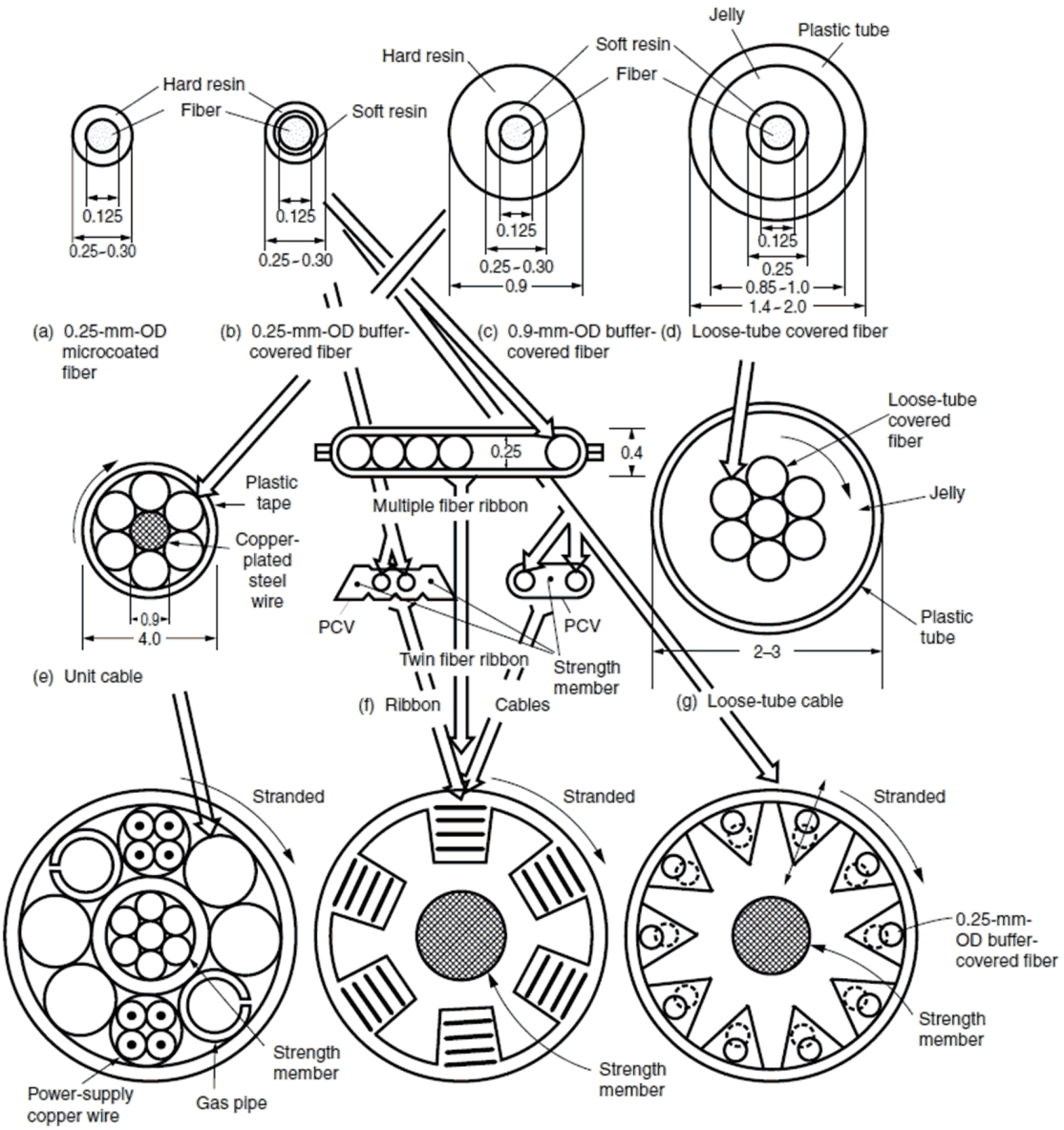

Figure 5.21: Optical fiber cables (all dimensions in millimeters). (a, b, c, d) Single fibers with different coating and buffer tubes. To avoid microbending and strain, optical fibers are often embedded into a jelly which allows the fiber to move freely within the protective plastic tube. (e, f, g) Various fiber-optic cables, comprising several fibers each. (h, i, j) Multi-unit cables with additional strength members, electrical wires, or gas pipes that keep moisture off the inside of the cable. The individual fibers are often stranded around a central steel wire, which enables more flexible stretching of the cable. (Figure adapted from [11])

5.6 Signal propagation in single-mode fibers

Data signals propagating along optical fibers are usually impaired by two major effects: Fiber attenuation and dispersion. If the propagating power is very high, nonlinear-optical effects occur as an additional source of impairment. These effects comprise, e.g., stimulated Brillouin scattering, stimulated Raman scattering, and third-order nonlinear interaction (Kerr-effect) leading to cross-phase modulation, self-phase modulation and four-wave mixing. Nonlinear-optical effects are analyzed in more detail in the lecture "Nonlinear Optics" and will not be considered further here. In addition, we will focus our consideration to single-mode fibers, where dispersion is governed by intramodal dispersion (chromatic dispersion) and polarization mode dispersion, see Section 3.4.2. Multi-mode fibers and intermodal dispersion is treated in more detail in the lecture "Field Propagation and Coherence".

5.6.1 Fiber attenuation

Fiber attenuation can be caused by various loss mechanisms and limits the length of the optical link. For linear loss mechanisms, the amount of power lost per distance is proportional to the local power propagating in the fiber,

$$\frac{\mathrm{d}P(z)}{\mathrm{d}z} = -\alpha P(z)\,, \tag{5.85}$$

where α denotes the attenuation coefficient of the fiber. This relation is solved by an exponential power decay along the fiber,

$$P(z) = P_0\,\mathrm{e}^{-\alpha z}, \qquad \frac{a}{z} = 10\lg\frac{P_0}{P(z)} = \alpha\,10\lg\mathrm{e} = 4.34\,\alpha. \tag{5.86}$$

The power attenuation constant α (unit km^{-1}) is usually expressed by specifying the attenuation a per distance z (unit dB / km).

The loss spectrum of a single-mode fiber is depicted in Fig. 5.22. In the following, we will discuss different sources of fiber losses.

Material absorption (adapted from [2])

Material absorption can be broadly subdivided into intrinsic and extrinsic absorption. Intrinsic absorption corresponds to the unavoidable absorption of the fiber's base material, whereas extrinsic absorption is due to impurities.

As explained in Section 2.3, any material with a nonzero dielectric susceptibility (refractive index $n \neq 1$) must have intrinsic absorption in some frequency range. For silica (SiO_2) molecules, electronic resonances lie in the UV region ($\lambda < 0.4\ \mu$m) whereas vibrational resonances occur in the infrared wavelength range ($\lambda > 7\ \mu$m). Because of the amorphous structure of silica, these resonances form broad absorption bands rather than narrow spectral lines. The overtones and the decaying tails of these bands extend into the visible and near-infrared region and are denoted as "UV absorption" and "IR absorption" in the loss spectrum depicted in Fig. 5.22. The intrinsic material loss of fused silica is below 0.1 dB/km (0.03 dB/km) in the wavelength range 0.8 μm – 1.6 μm (1.3 μm –1.6 μm).

Extrinsic absorption, in contrast, results from impurities within the fiber. As mentioned in the section 5.5.2, transition metals such as Fe, Cu, Co, Ni, Mn and Cr absorb strongly in the wavelength range between 0.6 μm – 1.6 μm. By synthesizing highly pure silica glass from chemical vapor deposition (CVD) processes (Section 5.5.2), the concentrations of these impurities can be kept below 1 ppb leading to loss levels of less than 1 dB/km. The main source of extrinsic absorption in silica fibers is then caused by the presence of OH^--ions in the glass matrix. The vibrational resonance of these ions occurs near 2.73 μm; its harmonic and combination tones with silica resonances produce absorption peaks around 0.95 μm, 1.24 μm, and 1.39 μm, Fig. 5.22. A

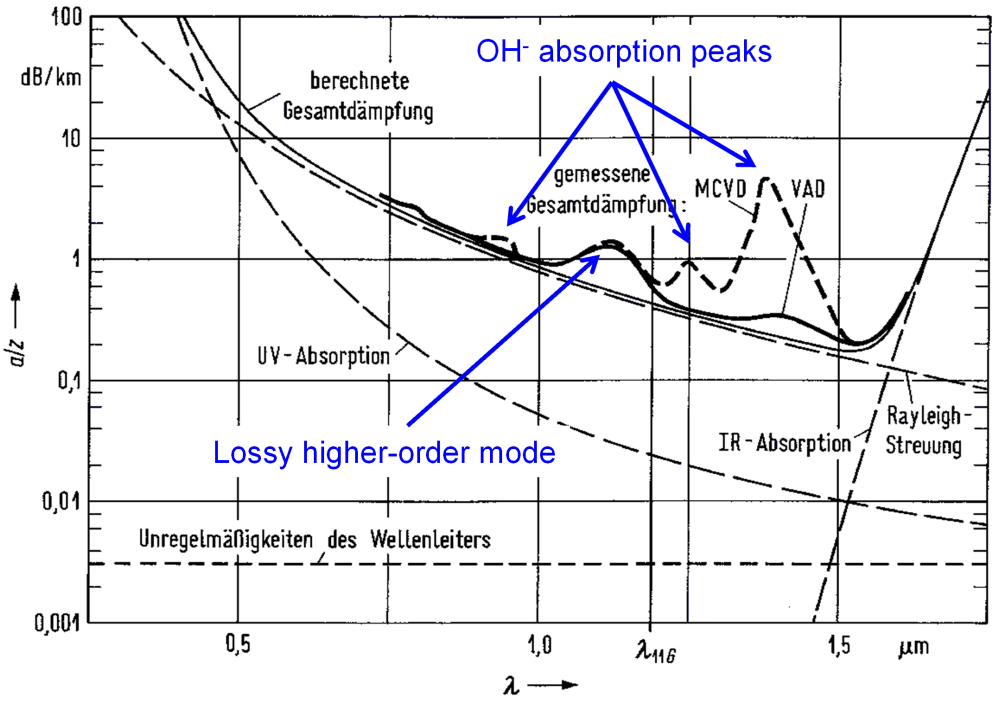

Figure 5.22: Attenuation spectrum of a Ge-doped singlemode silica fiber ($\Delta = 0.25\,\%$): For large wavelengths, fiber losses are dominated by infrared (IR) absorption caused by vibrational resonances of the molecules. For short wavelengths, the dominant loss mechanisms are Rayleigh scattering caused by random density fluctuations of the amorphous glass matrix, and ultra-violet (UV) absorption due to electronic resonances of the SiO_2 molecules. The measured total attenuation ("gemessene Gesamtdämpfung") is displayed for two fiber types, one of which was fabricated with vapor axial deposition (VAD), whereas the other was fabricated with modified chemical vapor deposition (MCVD). Apparently, this specific MCVD process resulted in elevated concentrations of hydroxyl (OH^-) ions, which is clearly visible from the characteristic absorption peaks at around 0.95 μm, 1.24 μm, and 1.39 μm. Another loss peak occurs for wavelengths slightly below the cutoff of the LP_{11} mode (around 1200 nm). This is caused by coupling of power from the fundamental to the next higher-order mode. Waveguide irregularities ("Unregelmäßigkeiten des Wellenleiters") do not play a significant role for fiber losses. This is in sharp contrast to integrated optical high index-contrast waveguides, where surface roughness is the main loss mechanism.

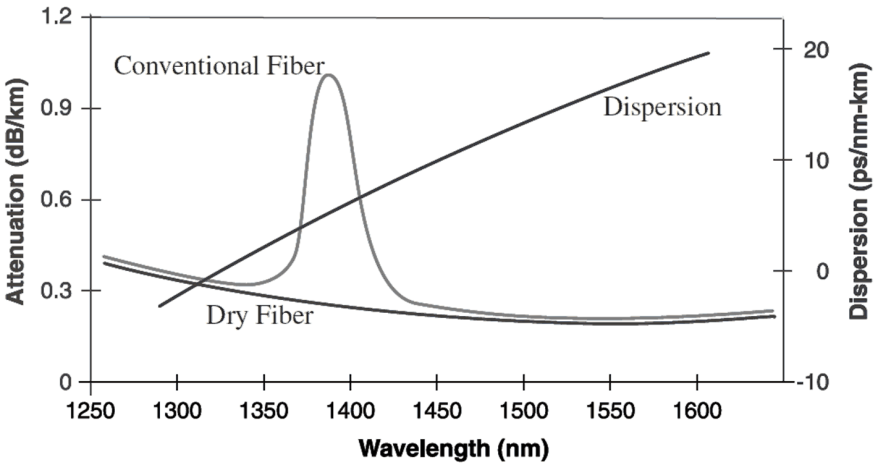

Figure 5.23: Loss spectrum of conventional fiber and so-called all-wave fiber ("dry fiber"). (Figure adapted from [2])

concentration of 1 ppm of OH^-can cause a loss of (50 dB/km) at 1.39 μm. In modern fibers, the OH^-concentration can be kept below 0.01 ppm, which results in an absorption peak of less than 1 dB/km at 1.39 μm. In so-called "dry fiber", the OH^-concentration is further reduced and the 1.39 μm absorption peak almost disappears. Such fibers can be used to transmit signals over the entire band between 1.3 μm and 1.6 μm and are, e.g., marketed under the trade name "AllWave" fiber.

Rayleigh scattering (adapted from [2])

Rayleigh scattering is a fundamental loss mechanism that arises from spatial fluctuations of the refractive index. The amorphous nature of fused silica leads to random density fluctuations and hence random fluctuations of the refractive index on a scale much smaller than the wavelength. The scattering loss varies as[6] λ^{-4}, and the corresponding fiber attenuation coefficient can be written as

$$\alpha_R = \frac{C}{\lambda^4}. \tag{5.87}$$

At $\lambda = 1.55$ μm Rayleigh scattering leads to losses of around $0.12 - 0.16$ dB/km and represents the dominant loss mechanism, Fig. 5.22.

Rayleigh scattering can be reduced by increasing the wavelength. This needs, however, materials with low IR absorption for wavelengths beyond 2 μm. Fluorozirconate (ZrF_4) has an intrinsic material absorption of about 0.01 dB/km at 2.55 μm, but extrinsic losses are still at around 1 dB/km. Chalcogenide glasses exhibit intrinsic minimum losses at wavelengths of about 10 μm, with predicted attenuation of 10^{-3} dB/km. Practical loss levels, however, are still larger than those of silica fibers. Fused silica fibers therefore remain the mainstay of optical communication systems despite considerable effort towards finding alternative materials.

Waveguide imperfections (adapted from [2])

An ideal perfectly straight single-mode fiber would guide the optical mode without any energy leakage. Real fibers, however, can exhibit additional losses due to imperfections of the waveguide geometry. This applies especially to multimode fibers, where power is coupled between different

[6]The the optical inhomgeneity can be considered as a Hertzian dipole with a constant dipole moment. The total power radiated by such a dipole is propotional to ω^4 and hence to λ^{-4}.

modes by waveguide imperfections. Three main types of waveguide imperfections may usually occur in single-mode fused silica fibers:

- Imperfections at the core-cladding interface, e.g., random core radius variations, can be considered as index variations on a scale longer than the wavelength and lead to additional propagation loss by so-called Mie scattering. These losses can be minimized by keeping the relative core radius variations very small, typically below 1%, which results in a scattering loss of typically less than 0.03 dB/km.

- Microbending, i.e., random axial distortions ("bends") can occur when the fiber is pressed against a rough surface, e.g. during cabling or inside a cable assembly that is pressing onto the fiber. In extreme cases, this can lead to extremely large additional losses (100 dB/km). Microbending losses can be reduced by appropriate cable assemblies and by keeping the normalized frequency V of the fiber as large as possible (typically in the range between 2.0 and 2.4) so that mode energy is primarily confined to the core.

- So-called macrobends with bend radii of typically a few millimeters can constitute another source of propagation loss, especially in multi-mode fibers. For a local radius of curvature of R, the bend-induced fiber attenuation is proportional to $\exp(-R/R_c)$, where $R_c = a/\left(n_1^2 - n_2^2\right)$. For single-mode fibers, we typically have $R_c = 0.2 - 0.5\mu$m. For macroscopic bends, we usually have $R > 5$ mm , and the associated losses are negligible in practice.

The sum of all aforementioned effects leads to the typical loss spectrum depicted in Fig. 5.22. Single-mode fused silica fibers usually have minimum propagation losses around a wavelength of 1.55 μm, and long-distance optical telecommunication systems usually operate in this wavelength range. Other sources of optical loss in fiber-optic transmission systems comprise imperfect splices or connectors. For extended fiber-optic systems, these localized losses are sometimes treated as part of the entire cable loss.

5.6.2 Chromatic dispersion and dispersion compensation

In single-mode fibers, the fundamental mode can exist in two orthogonal polarizations and is hence two-fold degenerate. Signal propagation is hence governed by polarization-mode dispersion (PMD) and chromatic dispersion (CD) as the dominant effects.

As discussed in Section 3.4.3, chromatic dispersion (CD) results from the wavelength-dependence of group velocity. The associated spread group delays can be expanded into a power series about the center wavelength,

$$\frac{\triangle t_g}{z} = C_\lambda \, \Delta\lambda_c + D_\lambda \, \Delta\lambda_c^2 \tag{5.88}$$

$$C_\lambda = -\frac{2\pi c}{\lambda^2}\beta_c^{(2)} \tag{5.89}$$

where C_λ denotes the chromatic dispersion coefficient (unit: $\frac{\text{ps}}{\text{km nm}}$), and where D_λ describes the so-called dispersion slope. For waveguides with small index contrasts between the core and the cladding, the chromatic dispersion coefficient can be written as a sum of the material dispersion coefficient M_λ and the waveguide dispersion coefficient W_λ,

$$C_\lambda = M_\lambda + W_\lambda. \tag{5.90}$$

The waveguide dispersion W_λ can be influenced by suitable choice of waveguide geometry. In the following sections, we will investigate signal distortions that are caused by chromatic dispersion, and we will discuss how they can be mitigated by dispersion engineering and dispersion compensation techniques.

Dispersion-induced signal distortions

For a quantitative analysis of dispersion-induced signal distortion, let us again consider a time-dependent signal propagating along an optical fiber in the fundamental LP_{01} mode, which, without loss of generality, is assumed to be polarized along the x-direction. The electric field can then be expressed as

$$\underline{E}_x(\mathbf{r},t) = \underbrace{\underline{\Psi}(x,y)}_{\text{lateral mode field}} \times \underbrace{\underline{A}(z,t)}_{\text{slowly varying envelope}} \times \underbrace{\exp\left(\mathrm{j}\left(\omega_c t - \beta_c z\right)\right)}_{\text{optical carrier}} \tag{5.91}$$

$$= \underline{\Psi}(x,y)\ \underline{a}(z,t) \tag{5.92}$$

where $\underline{\Psi}(x,y)$ denotes the lateral profile of the fundamental mode's $\underline{E}_x$-component, and where $\underline{A}(z,t)$ denotes the slowly varying complex envelope of the signal. The z- and t-dependence of the signal is given by the complex mode amplitude $\underline{a}(z,t)$ and its Fourier transform $\widetilde{\underline{a}}(z,\omega)$,

$$\underline{a}(z,t) = \underline{A}(z,t)\exp\left(\mathrm{j}\left(\omega_c t - \beta_c z\right)\right) \tag{5.93}$$

$$\widetilde{\underline{a}}(z,\omega) = \widetilde{\underline{A}}(z,\omega-\omega_c)\exp\left(-\mathrm{j}\,\beta_c z\right) \tag{5.94}$$

$\underline{A}(z,t)$ is a "slowly varying envelope" in the sense that it changes much slower than the underlying carrier signal, i.e., the bandwidth $\Delta\omega$ of $\widetilde{\underline{A}}(z,\omega)$ is much smaller that the carrier frequency, $\Delta\omega \ll \omega_c$. This approximation is often referred to as the slowly varying envelope approximation (SVEA).

The propagation along the fiber is most conveniently described by a complex frequency-dependent propagator $\exp\left(-\mathrm{j}\,\beta\left(\omega\right)z\right)$ in the frequency domain,

$$\widetilde{\underline{a}}(z,\omega) = \widetilde{\underline{a}}(0,\omega)\exp\left(-\mathrm{j}\,\beta\left(\omega\right)z\right) \tag{5.95}$$

For narrow-band optical spectra $\widetilde{\underline{a}}(z,\omega)$ it is useful to expand $\beta(\omega)$ in a Taylor series about the carrier frequency ω_c,

$$\beta(\omega) \approx \beta_c + (\omega-\omega_c)\beta_c^{(1)} + \frac{(\omega-\omega_c)^2}{2!}\,\beta_c^{(2)} + \frac{(\omega-\omega_c)^3}{3!}\,\beta_c^{(3)} + \dots\ , \tag{5.96}$$

Inserting Eqs. (5.96) and (5.94) in Eq. (5.95), we can derive an expression for the evolution of the slowly varying envelope along z,

$$\underline{A}(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\widetilde{\underline{A}}(0,\omega)\exp\left(-\mathrm{j}\left(\beta_c^{(1)}\omega z + \frac{\beta_c^{(2)}}{2}\omega^2 z + \frac{\beta_c^{(3)}}{6}\omega^3 z + \dots\right)\right)\exp\left(\mathrm{j}\,\omega t\right)\,\mathrm{d}\,\omega \tag{5.97}$$

A time-domain equivalent of this relation can be found by considering $\partial\underline{A}(z,t)/\partial z$ and by using basic properties of the Fourier transformation. This leads to the basic propagation equation of a signal inside a dispersive single-mode fiber,

$$\frac{\partial\underline{A}(z,t)}{\partial z} + \beta_c^{(1)}\frac{\partial\underline{A}(z,t)}{\partial t} - \mathrm{j}\,\frac{1}{2}\beta_c^{(2)}\frac{\partial^2\underline{A}(z,t)}{\partial t^2} - \frac{1}{6}\beta_c^{(3)}\frac{\partial^3\underline{A}(z,t)}{\partial t^3} - \dots = 0 \tag{5.98}$$

The quantity $\beta_c^{(1)}$ describes the group delay of the pulse envelope. In many cases, it is convenient to eliminate the group delay from the equation by introducing a retarded time frame,

$$t' = t - \beta_c^{(1)}z \tag{5.99}$$

$$z' = z \tag{5.100}$$

$$\underline{A}(z,t) = \underline{A}'\left(z, t-\beta_c^{(1)}z\right) \tag{5.101}$$

which allows to express the partial derivatives of $\underline{A}(z,t)$ by those of $\underline{A}'\left(z, t-\beta_c^{(1)} z\right)$,

$$\frac{\partial \underline{A}}{\partial z} = \frac{\partial \underline{A}'}{\partial z'} - \beta_1 \frac{\partial \underline{A}'}{\partial t'}; \qquad \frac{\partial^n \underline{A}}{\partial t^n} = \frac{\partial^n \underline{A}'}{\partial t'^n}. \tag{5.102}$$

The quantity $\beta_c^{(1)}$ can hence be eliminated from Eq. (5.98),

$$\frac{\partial \underline{A}'(z',t')}{\partial z'} - \mathrm{j}\,\frac{1}{2}\beta_c^{(2)} \frac{\partial^2 \underline{A}'(z',t')}{\partial t'^2} - \frac{1}{6}\beta_c^{(3)} \frac{\partial^3 \underline{A}'(z',t')}{\partial t'^3} - \ldots = 0 \tag{5.103}$$

Note that the primes are often omitted without further notice. If we neglect third- and higher-order dispersion terms, $\beta_c^{(3)} = \beta_c^{(4)} = \cdots = 0$, Eq. (5.104) can be rewritten as

$$\underline{A}'(z',t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{\underline{A}}'(0,\omega) \exp\left(-\mathrm{j}\,\frac{\beta_c^{(2)}}{2}\omega^2 z'\right) \exp(\mathrm{j}\,\omega t')\,\mathrm{d}\,\omega \tag{5.104}$$

Propagation of a chirped Gaussian impulse

Let us now consider the propagation of a so-called chirped Gaussian impulse, which may serve as a model for an optical data signal. A "chirped" Gaussian impulse is a waveform with a Gaussian envelope and a time-dependent frequency[7] and can be written as

$$\underline{A}(0,t) = \underline{A}_o \exp\left(-\frac{(1-\mathrm{j}\,\alpha)\,t^2}{2\sigma_t^2}\right), \tag{5.105}$$

where α denotes the so-called chirp parameter[8]. The instantaneous frequency of the complete signal $\underline{a}(0,t) = \underline{A}(0,t)\exp(\mathrm{j}\,\omega_c t)$ can be defined by the rate at which the phase of the complex signal changes,

$$\omega(t) = \frac{\mathrm{d}}{\mathrm{d}t} \arg\{\underline{a}(0,t)\} = \omega_c + \frac{\alpha}{\sigma_t^2} t \tag{5.106}$$

For $\alpha > 0$, the leading edge of the signal has a lower frequency ("red shift") than the trailing edge ("blue shift"); for $\alpha < 0$, the situation is reversed.

The frequency-domain representation of the chirped Gaussian impulse is obtained from the Fourier transform of Eq. (5.105),

$$\widetilde{\underline{A}}(0,\omega) = \underline{A}_o \sqrt{\frac{2\pi\sigma_t^2}{(1-\mathrm{j}\,\alpha)}} \exp\left(-\frac{\sigma_t^2\omega^2}{2\,(1-\mathrm{j}\,\alpha)}\right) \tag{5.107}$$

$$= \underline{A}_o \sqrt{\frac{2\pi\sigma_t^2}{(1-\mathrm{j}\,\alpha)}} \exp\left(-\frac{(1+\mathrm{j}\,\alpha)\,\omega^2}{2\sigma_\omega^2}\right) \tag{5.108}$$

where the variance of the Gaussian spectrum is given by

$$\sigma_\omega^2 = \frac{\left(1+\alpha^2\right)}{\sigma_t^2} \tag{5.109}$$

The corresponding time-bandwidth product at $z = 0$ is then given by

$$\sigma_t \sigma_\omega = \sqrt{1+\alpha^2} \tag{5.110}$$

[7] to chirp = zirpen, zwitschern

[8] Note that in the last section, α was used for the attenuation coefficient of the fiber and do not confuse the two quantities.

It is minimal for an unchirped Gaussian impulse ($\alpha = 0$). Any chirp, regardless of sign, leads to an increased spectral bandwidth (for the same temporal width) and hence increases the time-bandwidth product.

Using Eq (5.104), neglecting third- and higher-order dispersion terms, and omitting the primes denoting the retarded time frame, we obtain the complex signal envelope after a propagation distance z,

$$\underline{A}(z,t) = \frac{\underline{A}_o}{\sqrt{Q(z)}} \exp\left(-\frac{(1-\mathrm{j}\,\alpha)\,t^2}{2\sigma_t^2 Q(z)}\right) \tag{5.111}$$

where

$$Q(z) = 1 + (\mathrm{j} + \alpha)\,\frac{\beta_c^{(2)} z}{\sigma_t^2} \tag{5.112}$$

The pulse width $\sigma_t(z)$ is then given by

$$\frac{\sigma_t(z)}{\sigma_t(0)} = \sqrt{\left(1 + \alpha\frac{\beta_c^{(2)} z}{\sigma_t^2(0)}\right)^2 + \left(\frac{\beta_c^{(2)} z}{\sigma_t^2(0)}\right)^2} \tag{5.113}$$

For an impulse which is initially unchirped, $\alpha = 0$ and the envelope broadens as

$$\frac{\sigma_t(z)}{\sigma_t(0)} = \sqrt{1 + \left(\frac{z}{L_D}\right)^2}, \tag{5.114}$$

where the dispersion length L_D is given by

$$L_D = \frac{\sigma_t^2(0)}{\left|\beta_c^{(2)}\right|}. \tag{5.115}$$

L_D denotes the distance after which the width of an initially unchirped Gaussian pulse has increased by a factor of $\sqrt{2}$.

For a chirped input pulse ($\alpha \neq 0$), the pulse width may increase or decrease during propagation, depending on the relative signs of β_2 and α, see Fig. 5.24. If, e.g., $\beta_2 > 0$, then the chromatic dispersion coefficient $C_\lambda = -2\pi C/\lambda^2 < 0$, and the high-frequency ("blue") components of the signal spectrum are delayed with respect to the low-frequency ("red") parts. An initial pulse with a positive chirp parameter $\alpha > 0$ and hence a red-shifted leading edge will further broaden during propagation, whereas an initial pulse with a negative chirp parameter $\alpha < 0$ will first be compressed, before it also broadens for large propagation distances z.

Limitations of dispersive broadening on the data rate

Dispersive pulse broadening leads to limitations of the symbol rate that can be transmitted through the optical channel. The spectral broadening depends strongly on the spectral width of the data signal, which depends on the linewidth of the optical carrier and on the modulation.

In the case of narrowband sources, the linewidth of the carrier is much smaller than the bandwidth of the modulation. The spectral width of the signal is hence solely defined by the modulation speed, i.e., the duration of the signal pulses. If unchirped Gaussian impulses are launched into a fiber of length L, the pulse width at the fiber end can be calculated from Eq. (5.114),

$$\sigma_t^2(L) = \sigma_t^2(0) + \left(\frac{\beta_c^{(2)} L}{\sigma_t(0)}\right)^2 \tag{5.116}$$

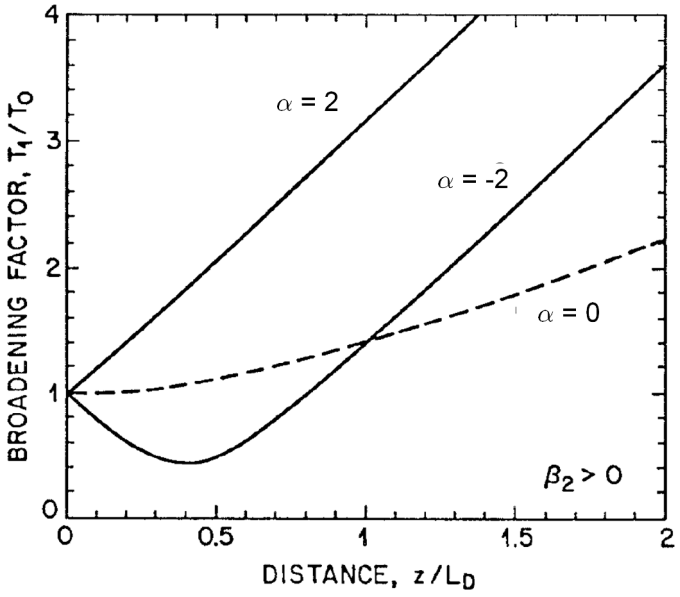

Figure 5.24: Evolution of pulse width during propagation through a dispersive fiber. The broadening factor gives the ratio of local pulse duration $\sigma_t(z) \hat{=} T_1$ and initial pulse duration $\sigma_t(0) \hat{=} T_0$. For $\beta_2 > 0$, an initial pulse with a positive chirp parameter $\alpha > 0$ will further broaden during propagation, whereas an initial pulse with a negative chirp parameter $\alpha < 0$ will first be compressed, before it also broadens for large propagation distances z. (Figure adapted from [2]).

For very big initial pulse widths, $\sigma_t^2(0) \to \infty$, the $\sigma_t^2(L)$ will be very large, too. The same is true for very short initial pulses $\sigma_t^2(0) \to 0$: In this case, the bandwidth of the signal is large, hence leading to a large spread of group velocities for the different spectral components. For a given fiber length L, there is obviously an optimum input pulse width $\sigma_{t,\mathrm{opt}}(0)$ which leads to a minimum pulse width $\sigma_{t,\ \mathrm{min}}(L)$ at the output,

$$\sigma_{t,\ \mathrm{min}}(L) = \sqrt{2\left|\beta_c^{(2)}\right| L} \quad \text{for} \quad \sigma_{t,\mathrm{opt}}(0) = \sqrt{\left|\beta_c^{(2)}\right| L}\ . \tag{5.117}$$

For a simple on-off-keying (OOK) data transmission, the minimum pulse width $\sigma_{t,\ \mathrm{min}}(L)$ dictates the minimum duration of a bit slot, $T_{\mathrm{min}} \approx 2\sigma_{t,\mathrm{min}}$, and the maximum achievable bit rate $B_{\mathrm{max}} = 1/T_{\mathrm{min}}$ must obey the relation

$$B_{\mathrm{max}} \sqrt{\left|\beta_c^{(2)}\right| L} \leq \frac{1}{2\sqrt{2}} \tag{5.118}$$

For narrowband sources, the maximum transmission bit rate B_{max} hence decreases as $1/\sqrt{L}$,

$$B_{\mathrm{max}} \propto \frac{1}{\sqrt{L}} \tag{5.119}$$

If the bandwidth is increased by a factor of 2, then the maximum duration of the output pulse reduces by a factor of 2 as well, while the spectral bandwidth and hence the dispersion-induced broadening increases. The length of the fiber must then be reduced by a factor of 4.

For broadband optical sources, the linewidth of the carrier $\sigma_{\omega,\mathrm{c}}$ is much larger than the bandwidth of the modulation and hence dictates the overall spectral width of the signal. For sufficiently short input pulse widths $\sigma^2(0)$, spectral broadening is solely dictated by $\sigma_{\omega,\mathrm{c}}$,

$$\sigma_t^2(L) = \sigma^2(0) + \left(\beta_c^{(2)} L \sigma_{\omega,\mathrm{c}}\right)^2 \approx \left(\beta_c^{(2)} L \sigma_{\omega,\mathrm{c}}\right)^2 \tag{5.120}$$

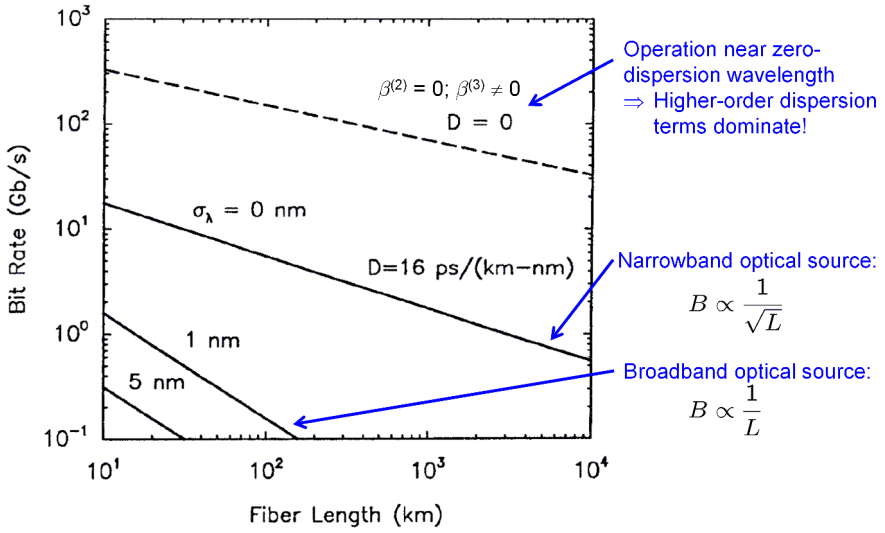

Figure 5.25: Dispersion-related limitations of the symbol rate for narrowband ($\sigma_\lambda = 0\,\mathrm{nm}$) and broadband ($\sigma_\lambda = 1\,\mathrm{nm}$, $\sigma_\lambda = 5\,\mathrm{nm}$) optical sources. For narrowband sources, $B_{\max} \propto \frac{1}{\sqrt{L}}$, whereas for broadband sources we find $B_{\max} \propto \frac{1}{L}$. When operating the fiber near its zero-dispersion wavelength ($D = 0$), higher-order dispersion terms dominate, which also leads to distance-dependent bandwidth limitations. (Figure adapted from [2]).

The maximum achievable bit rate $B_{\max} = 1/2\sigma_t\,(L)$ is then given by

$$B_{\max}\beta_c^{(2)} L \sigma_{\omega,\mathrm{c}} \leq \frac{1}{2}, \tag{5.121}$$

i.e., the bit rate now scales inversely to the fiber length,

$$B_{\max} \propto \frac{1}{L}. \tag{5.122}$$

If the fiber is operated near its zero-dispersion wavelength, $\beta_c^{(2)} \approx 0$ in Eq. (5.97) and $\beta_c^{(3)}$ has to be taken into account to describe the dominant effects of dispersive pulse broadening. For narrowband sources, a relation similar to Eq. (5.116) can be derived for third-order dispersion [2]. Dispersion-induced bit-rate limitations for different cases are depicted in Fig. 5.25.

Dispersion engineering and dispersion compensation

Dispersion-induced limitations of fiber-optic transmission systems can be mitigated either by operating close to the zero-dispersion wavelength of the fiber or by using a dispersion compensation technique. Long-distance optical telecommunications is mostly performed at wavelengths near $1.55\,\mu\mathrm{m}$, where silica fibers have minimum attenuation. For conventional single-mode fibers, material dispersion M_λ amounts to approximately $22\,\frac{\mathrm{ps}}{\mathrm{km\ nm}}$, whereas waveguide dispersion is slightly negative and assumes values of $W_\lambda \approx -6\,\frac{\mathrm{ps}}{\mathrm{km\ nm}}$. This leads to a total chromatic dispersion of the waveguide of $C_\lambda \approx 16\,\frac{\mathrm{ps}}{\mathrm{km\ nm}}$, see Fig. 5.26 (a). In so-called dispersion-shifted fibers, the diameter of the fiber core is decreased while the refractive index contrast between the core and the cladding region is increased. In this way, waveguide dispersion can be made more negative such that it approximately compensates the material dispersion. When used in wavelength division multiplexing (WDM) systems, dispersion shifted fibers can suffer from nonlinear interactions which cause crosstalk between different wavelength channels. To avoid this, nonzero dispersion-shifted fibers are often used, where chromatic dispersion still has a small nonzero value of, e.g., $C \approx 2\,\frac{\mathrm{ps}}{\mathrm{km\ nm}}$, see Fig. 5.26 (b).

If the index contrast is further increased while the fiber core diameter is reduced, the total chromatic dispersion can assume large negative values, Fig. 5.26 (c). Such fibers can, e.g., be used for compensating the positive chromatic dispersion of standard singlemode fibers. To this end, different optical fibers with opposite signs of $\beta_c^{(2)}$ are concatenated. For simplicity, we assume two fiber sections of lengths L_1 and L_2 and dispersion coefficients $\beta_{c,1}^{(2)}$ and $\beta_{c,2}^{(2)}$. Applying Eq. (5.104) consecutively to both fiber sections, we find the pulse envelope at $z = L = L_1 + L_2$,

$$\underline{A}\,(L,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{\underline{A}}\,(0,\omega)\exp\left(-\mathrm{j}\,\frac{1}{2}\left(\beta_{c,1}^{(2)}L_1 + \beta_{c,2}^{(2)}L_2\right)\omega^2\right)\exp\left(\mathrm{j}\,\omega t\right)\,\mathrm{d}\,\omega, \tag{5.123}$$

Neglecting third-order dispersion, the original pulse shape will be reproduced for

$$\beta_{c,1}^{(2)}L_1 + \beta_{c,2}^{(2)}L_2 = 0 \tag{5.124}$$

$$C_{\lambda 1}L_1 + C_{\lambda 2}L_2 = 0 \tag{5.125}$$

As an example let us assume that the first fiber is a standard single-mode fiber of length $L_1 = 100\,\mathrm{km}$ and dispersion coefficient $C_{\lambda 1} = 16\frac{\mathrm{ps}}{\mathrm{km\ nm}}$, whereas the second fiber is a dispersion-compensating fiber with negative dispersion coefficient $C_{\lambda 2} = -50\frac{\mathrm{ps}}{\mathrm{km\ nm}}$. Proper dispersion compensation then requires the length of the second fiber to be chosen according to $L_2 = -\frac{C_{\lambda 1}}{C_{\lambda 2}}L_1 = 32\,\mathrm{km}$.

5.6.3 Polarization-mode dispersion (PMD)

In ideal rotationally symmetric fibers, the fundamental mode is twofold degenerate and can exist in two orthogonal polarization states. Real fibers, however, have small deviations from perfect cylindrical symmetry, which lead to birefringence, i.e., the two orthogonal polarizations of the fundamental mode have slightly different mode profiles and hence experience different group delays. As a consequence, if an input pulse excites both polarizations, it will broaden during propagation along the fiber. This phenomenon is referred to as polarization mode dispersion.

The birefringence varies along the fiber in a random fashion, which leads to random changes of the polarization state. In addition, the polarization state will be different for different spectral components of the signal. The analytical treatment of polarization mode dispersion is quite complex and will not be derived here. For large fiber lengths L (typically several km), the variance σ_T of the differential group delays can be expressed as

$$\sigma_T \approx D_p\sqrt{L}, \tag{5.126}$$

where D_p is the so-called PMD parameter (unit: $\mathrm{ps}/\sqrt{\mathrm{km}}$). State-of-the-art fibers typically have PMD parameters of less than 0.1 $\mathrm{ps}/\sqrt{\mathrm{km}}$. Because of the $\sqrt{L}$-dependence, PMD-induced broadening is often small compared to chromatic disperision. It becomes relevant for long-reach high-speed transmission systems with well-compensated chromatic dispersion. Several schemes have been developed to compensate PMD, comprising optical and electrical techniques. More information on the stochastic anlysis of PMD and on compensation techniques can be found in [2] and the references therein.

(a)

$\lambda/\,\mu\mathrm{m}$	**1.1**	**1.3**	**1.56**
V	2.497	2.113	1.761
$V\,\mathrm{d}^2(VB)/\,\mathrm{d}V^2$	0.150	0.370	0.710
$M\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−23.18	+1.58	+21.93
$W\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−1.78	−3.72	−5.94
$C\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−24.96	−2.14	+15.99
$D(\lambda_C = 1.325\,\mu\mathrm{m}) = 0.0415\,\mathrm{ps}\,/\,(\mathrm{km\,nm^2})$			

(b)

$\lambda/\,\mu\mathrm{m}$	**1.1**	**1.3**	**1.56**
V	2.346	1.985	1.654
$V\,\mathrm{d}^2(VB)/\,\mathrm{d}V^2$	0.223	0.476	0.845
$M\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−25.04	+0.58	+21.5
$W\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−7.33	−13.24	−19.59
$C\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−32.37	−12.65	+1.94
$D(\lambda_C = 1.523\,\mu\mathrm{m}) = 0.024\,\mathrm{ps}\,/\,(\mathrm{km\,nm^2})$			

(c)

$\lambda/\,\mu\mathrm{m}$	**1.1**	**1.3**	**1.56**
V	2.531	2.141	1.784
$C\,/\,\frac{\mathrm{ps}}{\mathrm{km\,nm}}$	−97.0	−67.0	−50.0

Figure 5.26: Dispersion characteristics of single-mode fibers. (a) Conventional single-mode fiber (SMF) with $n_1 = 1.450840$, $n_2 = 1.446918$, $a = 4.1\,\mu\mathrm{m}$ and $\Delta = 0.27\,\%$. Wavelength of zero chromatic dispersion $C(\lambda_C) = 0$ is $\lambda_C = 1.325\,\mu\mathrm{m}$, cutoff at $\lambda_{11G} = 1.142\,\mu\mathrm{m}$ (b) Dispersion-shifted fiber (DSF) with $n_1 = 1.457893$, $n_2 = 1.446918$, $a = 2.3\,\mu\mathrm{m}$ and $\Delta = 0.75\,\%$. Wavelength of zero chromatic dispersion $C(\lambda_C) = 0$ is $\lambda_C = 1.523\,\mu\mathrm{m}$, cutoff at $\lambda_{11G} = 1.073\,\mu\mathrm{m}$ (c) Dispersion-compensating fiber (DCF) with $n_1 = 1.476754$, $n_2 = 1.446918$, $a = 1.5\,\mu\mathrm{m}$ and $\Delta = 2\,\%$, cutoff at $\lambda_{11G} = 1.158\,\mu\mathrm{m}$

Chapter 6

Waveguide-based devices

In this chapter, we will introduce basic devices that consist of optical waveguide as the fundamental building blocks. To provide a basic understanding of light propagation within such systems, we will first introduce the mode expansion method, which will later on be expanded to coupled-mode theory. Based on these methods, we will consider the fundamental working principles of directional couplers, multi-mode interference (MMI) devices, waveguide gratings, and optical loss and gain in dielectric waveguides. Moreover, we will analyze fundamental principles of linear optical systems by considering their scattering-matrix representations.

6.1 Propagation through dielectric waveguide structures

In the previous chapters, we have discussed methods to calculate the eigenmodes of dielectric waveguide structures which are uniform in the direction of propagation. In this section, we will expand our consideration to optical signals that propagate through dielectric waveguide structures for which the refractive index varies along the propagation direction. In some cases, semi-analytical approximations may be used that are based on the mode expansion method and coupled-mode theory. In other cases, a numerical treatment will be necessary, that can, e.g., be based on a finite-difference beam propagation algorithm.

6.1.1 Mode expansion method

Eigenmode expansion

In many cases of practical interest, dielectric waveguide structures consist of transverse index distributions which are piecewise invariant in the propagation direction, see Fig. 6.1. For each of these sections, an eigenmode expansion can be used to describe the field propagation,

$$\underline{\mathbf{E}}(\mathbf{r}) = \sum_m a_m \underline{\mathcal{E}}_m(x,y) \exp\left(-\mathrm{j}\,\beta_\mu\left(\rho\right) z\right) + \sum_\mu \int_\rho a_\mu\left(\rho\right) \underline{\mathcal{E}}_{\rho,\mu}(x,y) \exp\left(-\mathrm{j}\,\beta_\mu\left(\rho\right) z\right)\,\mathrm{d}\,\rho \quad (6.1)$$

$$\underline{\mathbf{H}}(\mathbf{r}) = \sum_m a_m \underline{\mathcal{H}}_m(x,y) \exp\left(-\mathrm{j}\,\beta_\mu\left(\rho\right) z\right) + \sum_\mu \int_\rho a_\mu\left(\rho\right) \underline{\mathcal{H}}_{\rho,\mu}(x,y) \exp\left(-\mathrm{j}\,\beta_\mu\left(\rho\right) z\right)\,\mathrm{d}\,\rho \quad (6.2)$$

In these equations, $\underline{\mathcal{E}}_m(x,y)$ and $\underline{\mathcal{H}}_m(x,y)$ represent a discrete set of guided eigenmodes that are labeled by an integer index m. Similarly, $\underline{\mathcal{E}}_{\rho,\mu}(x,y)$ and $\underline{\mathcal{H}}_{\rho,\mu}(x,y)$ correspond to the radiation modes of the waveguide. The discrete mode index μ denotes different families of radiation modes (polarizations, symmetry properties), each of which comprises a continuous set of modes which are denoted by the continuous parameter ρ.

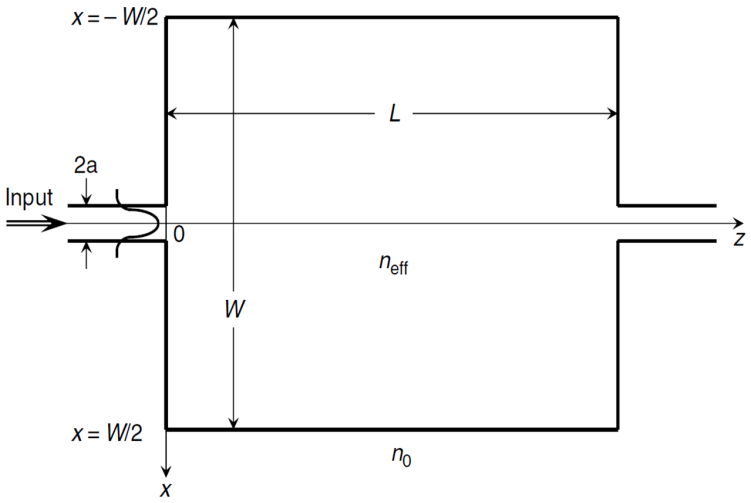

Figure 6.1: Many dielectric waveguide structures of practical interest are piecewise invariant along the z-direction. For each of these sections, an eigenmode expansion according to Eqs. (6.1) and (6.2) can be used to describe the field propagation. (Figure adapted from [25])

The continuous nature of the radiation mode spectrum can be understood when recalling the discussion of Section 3.3.3: A radiation mode of a simple slab waveguide can be thought of as a plane wave incident on the waveguide from the outside. The angle of propagation with respect to the waveguide axis can be arbitrary, and the propagation constant β can hence assume any value satisfying the relation $0 < |\beta| < n_2 k_0$, where, without loss of generality, n_2 denotes the biggest refractive index that can be found in the cladding region. The continuous mode parameter ρ can then be chosen to be the transverse propagation constant of the plane wave. For a given value of ρ, several such plane waves exist, which may, e.g., impinge from different directions, and they are indexed by the integer μ. For three-dimensional waveguides, calculating the radiation modes is somewhat more involved [27], but the basic principles stay the same.

It can be shown that guided modes and radiation modes form a complete set. That means that any arbitrary field distribution $\underline{\mathbf{E}}(\mathbf{r})$, $\underline{\mathbf{H}}(\mathbf{r})$ can be represented as a superposition of modes according to Eqs. (6.1) and (6.2), given that $\underline{\mathbf{E}}(\mathbf{r})$ and $\underline{\mathbf{H}}(\mathbf{r})$ fulfill Maxwell's equations.

For low index-contrast waveguides, waveguide modes have one dominant linearly polarized transverse field component, which is denoted as $\underline{\Psi}(x,y)$. In this case, mode field considerations can usually be restricted to this component, leading to a simplified scalar mode expansion of the dominant component $\Phi(x,y,z)$ of a weakly guided field,

$$\Phi(x,y,z) = \sum_m a_m \underline{\Psi}_m(x,y)\,\mathrm{e}^{-\mathrm{j}\,\beta_m z} + \sum_\mu \int_\rho a_\mu(\rho)\,\underline{\Psi}_{\rho,\mu}(x,y) e^{-\mathrm{j}\,\beta_\mu(\rho) z}\,\mathrm{d}\,\rho \tag{6.3}$$

Orthogonality relations

To calculate the coefficients a_m and $a_\mu(\rho)$ in Eqs. (6.1) and (6.2) from given field distributions $\underline{\mathbf{E}}(x,y,z=z_0)$ and $\underline{\mathbf{H}}(x,y,z=z_0)$, so-called orthogonality relations must be used. These relations can be derived from Maxwell's curl equations [22],

$$\nabla \times \underline{\mathbf{E}}(\mathbf{r},\omega) = -\mathrm{j}\,\omega\mu_0 \underline{\mathbf{H}}(\mathbf{r},\omega) \tag{6.4}$$

$$\nabla \times \underline{\mathbf{H}}(\mathbf{r},\omega) = \mathrm{j}\,\omega\epsilon \underline{\mathbf{E}}(\mathbf{r},\omega), \tag{6.5}$$

where $\epsilon = \epsilon_0 \epsilon_r$. Let us now consider the E- and the H-fields of two different modes of the structure,

which are denoted by subscripts ν and μ,

$$\underline{\mathbf{E}}_\nu(\mathbf{r}) = \underline{\mathcal{E}}_\nu(x,y)\exp\left(-\mathrm{j}\,\beta_\nu z\right), \tag{6.6}$$

$$\underline{\mathbf{H}}_\mu(\mathbf{r}) = \underline{\mathcal{H}}_\mu(x,y)\exp\left(-\mathrm{j}\,\beta_\mu z\right), \tag{6.7}$$

We apply Eq. (6.4) to mode ν and dot-multiply the resulting relation with $\underline{\mathbf{H}}_\mu^*(\mathbf{r})$. Similarly, we apply Eq. (6.5) to mode μ and dot-multiply the complex conjugate of resulting relation with $\underline{\mathbf{E}}_\nu(\mathbf{r})$. Subtracting the two expressions, we obtain the relation

$$\left(\underline{\mathbf{H}}_\mu^* \cdot (\nabla \times \underline{\mathbf{E}}_\nu) - \underline{\mathbf{E}}_\nu \cdot \left(\nabla \times \underline{\mathbf{H}}_\mu^*\right)\right) = \mathrm{j}\omega\left(\epsilon \underline{\mathbf{E}}_\nu \underline{\mathbf{E}}_\mu^* - \mu_0 \epsilon \underline{\mathbf{H}}_\nu \underline{\mathbf{H}}_\mu^*\right) \tag{6.8}$$

The expression left-hand side can be simplified by using the identity

$$\nabla \cdot \left(\underline{\mathbf{E}}_\nu \times \underline{\mathbf{H}}_\mu^*\right) = \underline{\mathbf{H}}_\mu^* \cdot (\nabla \times \underline{\mathbf{E}}_\nu) - \underline{\mathbf{E}}_\nu \cdot \left(\nabla \times \underline{\mathbf{H}}_\mu^*\right). \tag{6.9}$$

We now insert Eqs. (6.6) and (6.7) and apply the identity $\nabla \times (\Phi \mathbf{F}) = \Phi\,(\nabla \times \mathbf{F}) + (\nabla \Phi) \times \mathbf{F}$, where $\mathbf{F}$ denotes a vector field and Φ a scalar function, see Section A.2.2. Integrating over the entire cross section, we obtain

$$\underbrace{\iint_{-\infty}^{\infty} \nabla_t \cdot \left(\underline{\mathcal{E}}_\nu \times \underline{\mathcal{H}}_\mu^*\right) \mathrm{d}\,x\,\mathrm{d}\,y}_{I} - \mathrm{j}\left(\beta_\nu - \beta_\mu\right) \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_\nu \times \underline{\mathcal{H}}_\mu^*\right) \cdot \mathbf{e}_z \,\mathrm{d}\,x\,\mathrm{d}\,y \tag{6.10}$$

$$= \mathrm{j}\omega \iint_{-\infty}^{\infty} \left(\epsilon \underline{\mathcal{E}}_\nu \underline{\mathcal{E}}_\mu^* - \mu_0 \epsilon \underline{\mathcal{H}}_\nu \underline{\mathcal{H}}_\mu^*\right) \mathrm{d}\,x\,\mathrm{d}\,y, \tag{6.11}$$

where $\nabla_t = \mathbf{e}_x \frac{\partial}{\partial x} + \mathbf{e}_y \frac{\partial}{\partial y}$ denotes the nabla operator with respect to the transverse directions, and where $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are the unit vectors in x-, y-, and z-direction. It can be shown that the integral I on the left-hand side vanishes, and we will make this fact plausible with a few arguments: First, we restrict the integration to a finite region A in the transverse plane, and we use Gauss' divergence theorem to transform the area integral into a line integral along the boundary ∂A of the integration region,

$$\iint_A \nabla_t \cdot \left(\underline{\mathcal{E}}_\nu \times \underline{\mathcal{H}}_\mu^*\right) \mathrm{d}\,x\,\mathrm{d}\,y = \int_{\partial A} \left(\underline{\mathcal{E}}_\nu \times \underline{\mathcal{H}}_\mu^*\right) \cdot \mathbf{n}\,\mathrm{d}\,x\,\mathrm{d}\,y, \tag{6.12}$$

where $\mathbf{n}$ is the outward unit normal vector of the boundary ∂A. If at least one of the modes ν and μ is guided, the field strength decays exponentially as $x^2 + y^2 \to \infty$, and the line integral vanishes as the boundary of the integration area is shifted to infinity. If both modes are radiation modes, the field strengths do not decay for $x^2 + y^2 \to \infty$, but the line integral still vanishes due to the oscillatory behavior of the radiation modes along the boundary [22]. We may hence rewrite Eq. (6.12) as

$$\left(\beta_\nu - \beta_\mu\right) \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_\nu \times \underline{\mathcal{H}}_\mu^*\right) \cdot \mathbf{e}_z \,\mathrm{d}\,x\,\mathrm{d}\,y = -\omega \iint_{-\infty}^{\infty} \left(\epsilon \underline{\mathcal{E}}_\nu \underline{\mathcal{E}}_\mu^* - \mu_0 \epsilon \underline{\mathcal{H}}_\nu \underline{\mathcal{H}}_\mu^*\right) \mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.13}$$

Taking the complex conjugate of Eq. (6.13), interchanging the indices ν and μ, and subtracting the resulting relation from Eq. (6.13) leads to

$$\left(\beta_\nu - \beta_\mu\right) \iint_{-\infty}^{\infty} \left[\left(\underline{\mathcal{E}}_\nu \times \underline{\mathcal{H}}_\mu^*\right) + \left(\underline{\mathcal{E}}_\mu^* \times \underline{\mathcal{H}}_\nu\right)\right] \cdot \mathbf{e}_z \,\mathrm{d}\,x\,\mathrm{d}\,y = 0 \tag{6.14}$$

For nondegenerate modes we have $\beta_\nu \neq \beta_\mu$, and the integral on the left-hand side must vanish. This yields the so-called orthogonality relations which, for guided modes ν and μ, can be written as

$$\frac{1}{4} \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_\nu(x,y) \times \underline{\mathcal{H}}_\mu^\star(x,y) + \underline{\mathcal{E}}_\mu^\star(x,y) \times \underline{\mathcal{H}}_\nu(x,y)\right) \cdot \mathbf{e}_z \,\mathrm{d}\,x\,\mathrm{d}\,y = \mathcal{P}_\mu \delta_{\nu\mu}, \tag{6.15}$$

where $\delta_{\nu\mu}$ denotes the Kronecker delta, and where $\mathcal{P}_\mu$ represents the power that is associated with the mode field,

$$\mathcal{P}_\mu = \frac{1}{2} \iint_{-\infty}^{\infty} \mathrm{Re}\left\{ \underline{\mathcal{E}}_\mu(x,y) \times \underline{\mathcal{H}}_\mu^\star(x,y) \right\} \cdot \mathbf{e}_z \, \mathrm{d}\, x \, \mathrm{d}\, y. \tag{6.16}$$

For radiation modes, the corresponding integral does not converge, and the power associated with the mode fields cannot be directly calculated. The expression must then be interpreted as a power spectral density of the radiation modes,

$$\frac{1}{4} \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_{\rho,\mu}(x,y) \times \underline{\mathcal{H}}_{\rho',\mu'}^\star(x,y) + \underline{\mathcal{E}}_{\rho',\mu'}^\star(x,y) \times \underline{\mathcal{H}}_{\rho,\mu}(x,y) \right) \cdot \mathbf{e}_z \, \mathrm{d}\, x \, \mathrm{d}\, y = \mathcal{P}_{\rho,\mu} \delta_{\mu\mu'} \delta\left(\rho - \rho'\right), \tag{6.17}$$

where the mode fields are denoted by a continuous mode parameter ρ, and where $\delta\left(\rho - \rho'\right)$ is the Dirac delta function. The normalization coefficient $\mathcal{P}_{\rho,\mu}$ is then defined by the relation

$$\frac{1}{2} \iint_{-\infty}^{\infty} \mathrm{Re}\left\{ \underline{\mathcal{E}}_{\rho,\mu}(x,y) \times \underline{\mathcal{H}}_{\rho',\mu}^\star(x,y) \right\} \cdot \mathbf{e}_z \, \mathrm{d}\, x \, \mathrm{d}\, y = \mathcal{P}_{\rho,\mu} \delta\left(\rho - \rho'\right) \tag{6.18}$$

Note that guided modes and radiation modes are always orthogonal with respect to each other,

$$\frac{1}{4} \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_{\nu}(x,y) \times \underline{\mathcal{H}}_{\rho,\mu}^\star(x,y) + \underline{\mathcal{E}}_{\rho,\mu}^\star(x,y) \times \underline{\mathcal{H}}_{\nu}(x,y) \right) \cdot \mathbf{e}_z \, \mathrm{d}\, x \, \mathrm{d}\, y = 0 \tag{6.19}$$

In the case of degenerate guided modes with identical propagation constants, $\beta_\nu = \beta_\mu$, it cannot be concluded that the integral on the left-hand side of Eq. (6.14) must vanish. We may then still use linear combinations of the mode field solutions to construct a new set of mode fields which is orthogonal in the sense of Eqs. (6.17), (6.18), or (6.19).

In many cases of practical interest, simplified orthogonality relations can be used. Maxwell's equations for guided modes, Eqs. (3.68) and (3.69), are invariant with respect to the following transformation

$$\begin{aligned} \beta &\to -\beta \\ \mathcal{E}_x &\to \mathcal{E}_x \\ \mathcal{E}_y &\to \mathcal{E}_y \\ \mathcal{E}_z &\to -\mathcal{E}_z \\ \mathcal{H}_x &\to -\mathcal{H}_x \\ \mathcal{H}_y &\to -\mathcal{H}_y \\ \mathcal{H}_z &\to \mathcal{H}_z \end{aligned} \tag{6.20}$$

This transformation changes any solution of Maxwell's equations into another solution thereof; more precisely: It transforms the mode field of a forward propagating mode (propagation constant β) into the corresponding backward-propagating mode (propagation constant $-\beta$) and vice versa. Let us now consider two guided modes ν and μ that are neither identical ($\beta_\nu \neq \beta_\mu$) nor are they counterpropagating modes of the same propagation constant ($\beta_\nu \neq -\beta_\mu$) , and Eq. (6.15) hence reads

$$\frac{1}{4} \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_{\nu}(x,y) \times \underline{\mathcal{H}}_{\mu}^\star(x,y) + \underline{\mathcal{E}}_{\mu}^\star(x,y) \times \underline{\mathcal{H}}_{\nu}(x,y) \right) \cdot \mathbf{e}_z \, \mathrm{d}\, x \, \mathrm{d}\, y = 0, \tag{6.21}$$

We apply the transformations (6.20) to mode ν and rewrite Eq. (6.15),

$$\frac{1}{4} \iint_{-\infty}^{\infty} \left(\underline{\mathcal{E}}_{\nu}(x,y) \times \underline{\mathcal{H}}_{\mu}^\star(x,y) - \underline{\mathcal{E}}_{\mu}^\star(x,y) \times \underline{\mathcal{H}}_{\nu}(x,y) \right) \cdot \mathbf{e}_z \, \mathrm{d}\, x \, \mathrm{d}\, y = 0, \tag{6.22}$$

Adding Eqs. (6.15) and (6.21) and taking the real part of the resulting relation we obtain

$$\frac{1}{2}\iint_{-\infty}^{\infty} \mathrm{Re}\left\{\underline{\mathcal{E}}_\nu(x,y)\times\underline{\mathcal{H}}_\mu^\star(x,y)\right\}\cdot\mathbf{e}_z\,\mathrm{d}\,x\,\mathrm{d}\,y = 0 \quad \text{for} \quad \beta_\nu \neq \beta_\mu \text{ and } \beta_\nu \neq -\beta_\mu \tag{6.23}$$

This yields a simplified orthogonality relation for co-propagating modes,

$$\frac{1}{2}\iint_{-\infty}^{\infty} \mathrm{Re}\left\{\underline{\mathcal{E}}_\nu(x,y)\times\underline{\mathcal{H}}_\mu^\star(x,y)\right\}\cdot\mathbf{e}_z\,\mathrm{d}\,x\,\mathrm{d}\,y = \mathcal{P}_\mu\delta_{\nu\mu} \tag{6.24}$$

Note that this orthogonality relation should only be applied to modes that are all propagating in the same direction! Eq. (6.24) cannot "distinguish" counter-propagating fields that belong to the same mode ($\beta_\nu = -\beta_\mu$).

For modes of weakly guiding low-index-contrast waveguides, one transverse Cartesian field component is usually much stronger than the other one. This allows to further simplify the vectorial orthogonality relation according to Eq. (6.24). For an $\underline{\mathcal{E}}_x$-mode, we have $|\underline{\mathcal{E}}_x| \gg |\underline{\mathcal{E}}_z| \gg |\underline{\mathcal{E}}_y|$ and $|\underline{\mathcal{H}}_y| \gg |\underline{\mathcal{H}}_z| \gg |\underline{\mathcal{H}}_x|$, see Section 4.2. The relation between the magnetic field and the electric can then be approximated by

$$\underline{\mathcal{H}} \approx \frac{\beta}{\omega\mu_0}\begin{pmatrix} 0 \\ \underline{\mathcal{E}}_x \\ \frac{1}{j\beta}\frac{\partial\underline{\mathcal{E}}_x}{\partial y} \end{pmatrix} \tag{6.25}$$

Inserting Eq. (6.25) in Eq. (6.24) and replacing $\underline{\mathcal{E}}_x(x,y)$ by a scalar function $\underline{\Psi}(x,y)$, we obtain the orthogonality relation for scalar mode fields of weakly guided low-index-contrast waveguides,

$$\frac{\beta}{2\omega\mu_0}\iint_{-\infty}^{\infty} \underline{\Psi}_\nu(x,y)\,\underline{\Psi}_\mu^\star(x,y)\,\mathrm{d}\,x\,\mathrm{d}\,y = \mathcal{P}_\mu\delta_{\nu\mu}, \tag{6.26}$$

where $\mathcal{P}_\mu$ is given by

$$\mathcal{P}_\mu = \frac{\beta}{2\omega\mu_0}\iint_{-\infty}^{\infty} \left|\underline{\Psi}_\mu^\star(x,y)\right|^2\mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.27}$$

Note that the scalar orthogonality relation, Eq. (6.26), is again only valid if the modes are propagating in the same direction, since it was derived from Eq. (6.24). It cannot "distinguish" between counter-propagating fields that belong to the same mode ($\beta_\nu = -\beta_\mu$).

Example: Coupling efficiency to fundamental mode

To illustrate the use of the orthogonality relations, let us consider coupling of light into the front facet of a waveguide, see Fig. 6.2. Let $\underline{\mathbf{E}}(x,y,0)$ and $\underline{\mathbf{H}}(x,y,0)$ denote the launched electric and magnetic fields at $z = 0$, "just behind" the waveguide facet, i.e., just after passing the dielectric interface such that we do not need to take into account reflection from the waveguide facet. To calculate the power that is coupled into the fundamental mode of the waveguide, we expand $\underline{\mathbf{E}}(x,y,0)$ and $\underline{\mathbf{H}}(x,y,0)$ by the complete set of waveguide modes according to Eqs. (6.1) and (6.2),

$$\underline{\mathbf{E}}(x,y,0) = \sum_m a_m\underline{\mathcal{E}}_m(x,y) + \sum_\mu\int_\rho a_\mu(\rho)\,\underline{\mathcal{E}}_{\rho,\mu}(x,y)\;\mathrm{d}\,\rho \tag{6.28}$$

$$\underline{\mathbf{H}}(x,y,0) = \sum_m a_m\underline{\mathcal{H}}_m(x,y) + \sum_\mu\int_\rho a_\mu(\rho)\,\underline{\mathcal{H}}_{\rho,\mu}(x,y)\;\mathrm{d}\,\rho \tag{6.29}$$

The total power of the incident field is calculated by integrating the real part of the complex Poynting vector over the (x,y)-plane,

$$P_{\mathrm{tot}} = \frac{1}{2}\iint_{-\infty}^{\infty} \mathrm{Re}\left\{\underline{\mathbf{E}}(x,y,0)\times\underline{\mathbf{H}}^\star(x,y,0)\right\}\cdot\mathbf{e}_z\,\mathrm{d}\,x\,\mathrm{d}\,y \tag{6.30}$$

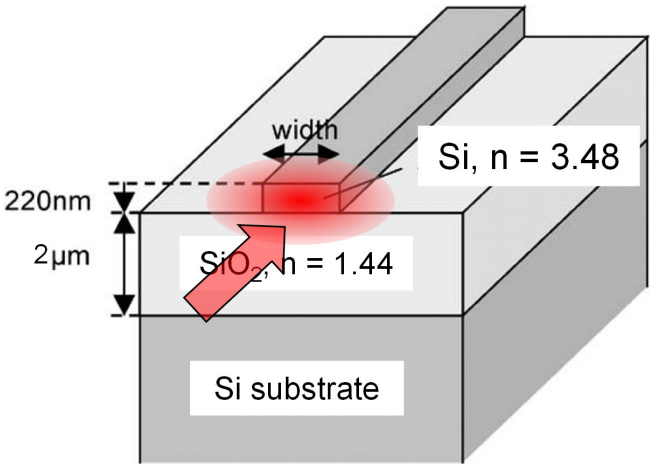

Figure 6.2: Coupling of a free-space beam to the facet of an integrated optical waveguide.

The power P_{fund} coupled into the fundamental mode depends on the mode coefficient a_0 and the power associated with the mode field,

$$P_{\text{fund}} = |a_0|^2 \cdot \frac{1}{2} \iint_{-\infty}^{\infty} \text{Re}\left\{\underline{\mathcal{E}}_0(x,y) \times \underline{\mathcal{H}}_0^\star(x,y)\right\} \cdot \mathbf{e}_z \,\mathrm{d}\,x \,\mathrm{d}\,y \tag{6.31}$$

The amplitude a_0 of the fundamental mode can be calculated by using the orthogonality relations. To this end, Eq. (6.28) is cross-multiplied with $\underline{\mathcal{H}}_0^\star(x,y)$ and the result is added to the cross-product of $\underline{\mathcal{E}}_0^\star(x,y)$ and Eq.(6.29). We then integrate both sides of the resulting relation over the entire (x,y)-plane and apply Eq. (6.15). After some manipulations we find the mode amplitude of the fundamental mode,

$$a_0 = \frac{1}{4\mathcal{P}_0} \cdot \iint_{-\infty}^{\infty} \left(\underline{\mathbf{E}}(x,y,0) \times \underline{\mathcal{H}}_0^\star(x,y) + \underline{\mathcal{E}}_0^\star(x,y) \times \underline{\mathbf{H}}(x,y,0)\right) \cdot \mathbf{e}_z \,\mathrm{d}\,x \,\mathrm{d}\,y \tag{6.32}$$

The power coupling efficiency $\eta = P_{\text{fund}}/P_{\text{tot}}$ is obtained by relating the power in the fundamental mode to the total power,

$$\eta = \frac{\left|\frac{1}{4} \cdot \iint_{-\infty}^{\infty} \left(\underline{\mathbf{E}}(x,y,0) \times \underline{\mathcal{H}}_0^\star(x,y) + \underline{\mathcal{E}}_0^\star(x,y) \times \underline{\mathbf{H}}(x,y,0)\right) \cdot \mathbf{e}_z \,\mathrm{d}\,x \,\mathrm{d}\,y\right|^2}{\frac{1}{2} \iint_{-\infty}^{\infty} \text{Re}\left\{\underline{\mathcal{E}}_0(x,y) \times \underline{\mathcal{H}}_0^\star(x,y)\right\} \cdot \mathbf{e}_z \,\mathrm{d}\,x \,\mathrm{d}\,y \cdot \frac{1}{2} \iint_{-\infty}^{\infty} \text{Re}\left\{\underline{\mathbf{E}}(x,y,0) \times \underline{\mathbf{H}}^\star(x,y,0)\right\} \cdot \mathbf{e}_z \,\mathrm{d}\,x \,\mathrm{d}\,y} \tag{6.33}$$

If a weakly guiding waveguide such as an optical fiber is illuminated with a linearly polarized excitation field $\underline{\mathbf{E}}(x,y,0) = \Phi(x,y,0)\ \mathbf{e}_x$, we may express the Poynting vector in terms of $\Phi(x,y,0)$ and use the simplified orthogonality relation, Eq. (6.26). The coupling efficiency can then be rewritten as:

$$\eta = \frac{P_{\text{fund}}}{P_{\text{tot}}} = \frac{\left|\iint_{-\infty}^{\infty} \Phi(x,y,0)\, \underline{\Psi}_0^\star(x,y) \,\mathrm{d}\,x \,\mathrm{d}\,y\right|^2}{\iint_{-\infty}^{\infty} |\underline{\Psi}_0(x,y)|^2 \,\mathrm{d}\,x \,\mathrm{d}\,y \cdot \iint_{-\infty}^{\infty} |\Phi(x,y,0)|^2 \,\mathrm{d}\,x \,\mathrm{d}\,y}. \tag{6.34}$$

The numerator of this fraction corresponds to the overlap of the mode field $\underline{\Psi}_0(x,y)$ and the excitation field $\Phi(x,y,0)$, whereas the denominator corresponds to the power contained in the two fields. The more "similar" $\underline{\Psi}_0(x,y)$ and $\Phi(x,y,0)$, the larger the coupling efficiency becomes. Ideal coupling ($\eta = 1$) is achieved, if the excitation field has the same shape as the mode field.

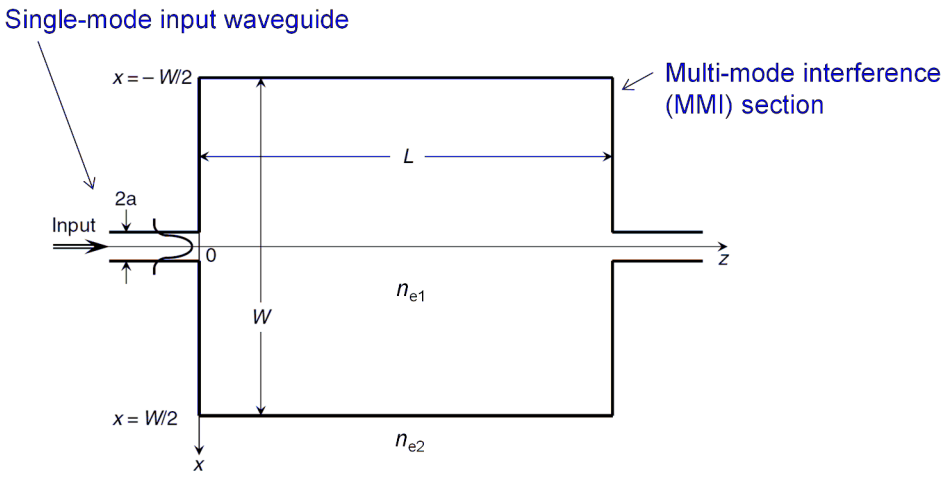

Figure 6.3: Multimode interference (MMI) devices consist of a broad multimode waveguide (MMI section), which is connected to several narrow feed waveguides. In the simplistic example shown here, only a single input and a single output waveguide is used. (Figure adapted from [25])

6.2 Multimode interference couplers

Multimode interference (MMI) couplers are used to split optical signals inside planar lightwave circuits. The central part of an MMI coupler is a broad multimode waveguide, connected to several narrow feed waveguides, which are normally single-moded, see Fig. 6.3. The operation principle of the devices is based on the so-called self-imaging effect, i.e., the fact that an input field is reproduced in single and multiple images at periodic intervals along the propagation direction.

The operating principle of MMI couplers can be analyzed using the mode expansion method. For simplicity, we assume low index-contrast waveguides which can be described in scalar approximation. Let $\Phi(x, y, 0)$ denote the excitation field which is launched from the input waveguides into the multimode sections at $z = 0$. This field can be expanded into guided modes $\underline{\Psi}_m(x, y)$ of the MMI section,

$$\Phi(x, y, 0) = \sum_m a_m \underline{\Psi}_m(x, y) \tag{6.35}$$

where the mode coefficients a_m are obtained from the orthogonality relation for scalar mode fields, Eq. (6.26),

$$a_m = \frac{\beta}{2\mathcal{P}_m \omega \mu_0} \iint_{-\infty}^{\infty} \Phi(x, y, 0)\, \underline{\Psi}_m^{\star}(x, y)\, \mathrm{d}\, x\, \mathrm{d}\, y. \tag{6.36}$$

Once the mode amplitudes are known, the field inside the MMI section can be calculated at any position z,

$$\Phi(x, y, z) = \sum_m a_m \underline{\Psi}_m(x, y) \exp\left(-\mathrm{j}\, \beta_m z\right). \tag{6.37}$$

The evolution of the field $\Phi(x, y, z)$ for $z > 0$ hence depends on the modal propagation constants β_m. We use the effective-index method to find an approximation for β_m, see section 4.3: The MMI section is represented by a broad multimode slab waveguide of width w and refractive indices n_{e1} in the core and n_{e2} in the cladding, see Fig. 4.7. The quantity n_{1e} denotes the effective index of the layer stack in the core region of the MMI section; n_{e2} corresponds to the effective index in the cladding region. The transverse phase constants of the various guided modes are then given by

$$u_m = \frac{w}{2} \sqrt{n_{e1}^2 k_0^2 - \beta^2}, \tag{6.38}$$

For modes far from cut-off, the transverse phase constants can be approximated by

$$u_m \approx (m+1)\,\frac{\pi}{2} \tag{6.39}$$

This leads to an expression for the difference between the propagation constant β_0 of the fundamental mode and that of higher-order modes,

$$\beta_0 - \beta_m = \frac{m\,(m+2)\,\pi}{3L_\pi}, \tag{6.40}$$

where the beat length L_π is defined to be:

$$L_\pi = \frac{4n_{1e}w^2}{3\lambda} \tag{6.41}$$

The field inside the MMI section can hence be written as

$$\Phi(x,y,z) = \sum_m a_m \underline{\Psi}_m(x,y) \exp\left(\mathrm{j}\,\frac{m\,(m+2)\,\pi}{3L_\pi} z\right) \exp\left(-\mathrm{j}\,\beta_0 z\right) \tag{6.42}$$

From this relation, we can derive some interesting properties of the MMI. For $z = 6L_\pi$, the initial field is reproduced,

$$\begin{aligned}\Phi(x,y,6L_\pi) &= \sum_m a_m \underline{\Psi}_m(x,y) \exp\left(\mathrm{j}\,m\,(m+2)\cdot 2\pi\right) \exp\left(-\mathrm{j}\,\beta_0 \cdot 6L_\pi\right) \\ &= \Phi(x,y,0)\ \mathrm{e}^{-\mathrm{j}\,\beta_0 \cdot 6L_\pi},\end{aligned} \tag{6.43}$$

and the same applies to $z = 12L_\pi$, $18L_\pi \ldots$, i.e., the initial image is reproduced repeatedly along the direction of propagation. To derive the field configuration for $z = 3L_\pi$, we have to make use of the fact that waveguide modes with even mode index m are symmetric with respect to the center line of the MMI section ($x = 0$), whereas odd mode indices belong to asymmetric modes,

$$\underline{\Psi}_m(-x,y) = \begin{cases} \underline{\Psi}_m(x,y) & \text{for } m \text{ even} \\ -\underline{\Psi}_m(x,y) & \text{for } m \text{ odd} \end{cases}. \tag{6.44}$$

As a consequence, the initial field is mirrored about the symmetry plane $x = 0$ of the MMI,

$$\Phi(x,y,3L_\pi) = \Phi(-x,y,0)\,\mathrm{e}^{-\mathrm{j}\,\beta_0 \cdot 3L_\pi} \tag{6.45}$$

see Fig. 6.4. In a similar way, we can show that at $z = \frac{3L_\pi}{2}$ two images are formed, one of which is mirrored about the symmetry plane $x = 0$ and delayed by $\pi/2$ with respect to the other,

$$\Phi\left(x,y,\frac{3L_\pi}{2}\right) = \left[\frac{1-\mathrm{j}}{2}\Phi(x,y,0) + \frac{1+\mathrm{j}}{2}\Phi(-x,y,0)\right] \mathrm{e}^{-\mathrm{j}\,\beta_0 \cdot \frac{3L_\pi}{2}} \tag{6.46}$$

see Fig. 6.4. More generally, one can show that N images are found after propagation distances

$$z = \frac{3pL_\pi}{N}, \tag{6.47}$$

where p and N are integers without a common divider. An in-depth analysis of MMI couplers can be found in [3] and the references therein. Note that the analytical treatment shown here is usually used for a first coarse design of MMI devices. Numerical methods are then needed to fine-tune the device geometry.

For light input at the center of the MMI section, the intensity distributions are depicted in Fig. 6.5. After a length of $3L_\pi/8$, the original light input at the center of the MMI is split in 8 equal images. If output waveguides are connected to the multimode section at the respective positions, the device can be used as an 1-to-8 power splitter. Note that the intensity pattern shown in Fig. 6.5 cannot be explained by Eq. (6.46): If light is input at the center of the MMI section, only even waveguide modes are excited. This case is not covered by Eq. (6.46).

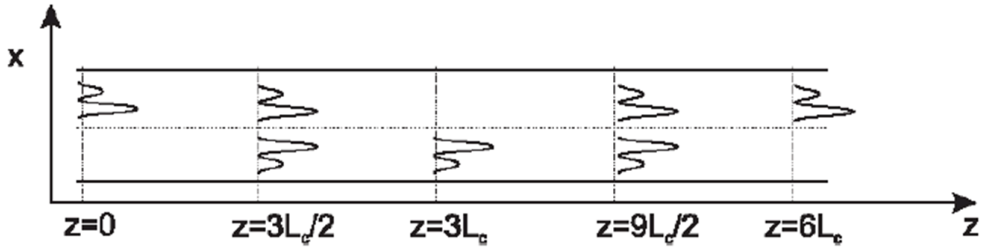

Figure 6.4: Image formation within an MMI section: For general excitation off the waveguide center, the transverse image repeats after $6L_\pi$, a mirrored version occurs after $3L_\pi$, and two copies of the input are obtained after $3L_\pi/2$. (Figure adapted from [4])

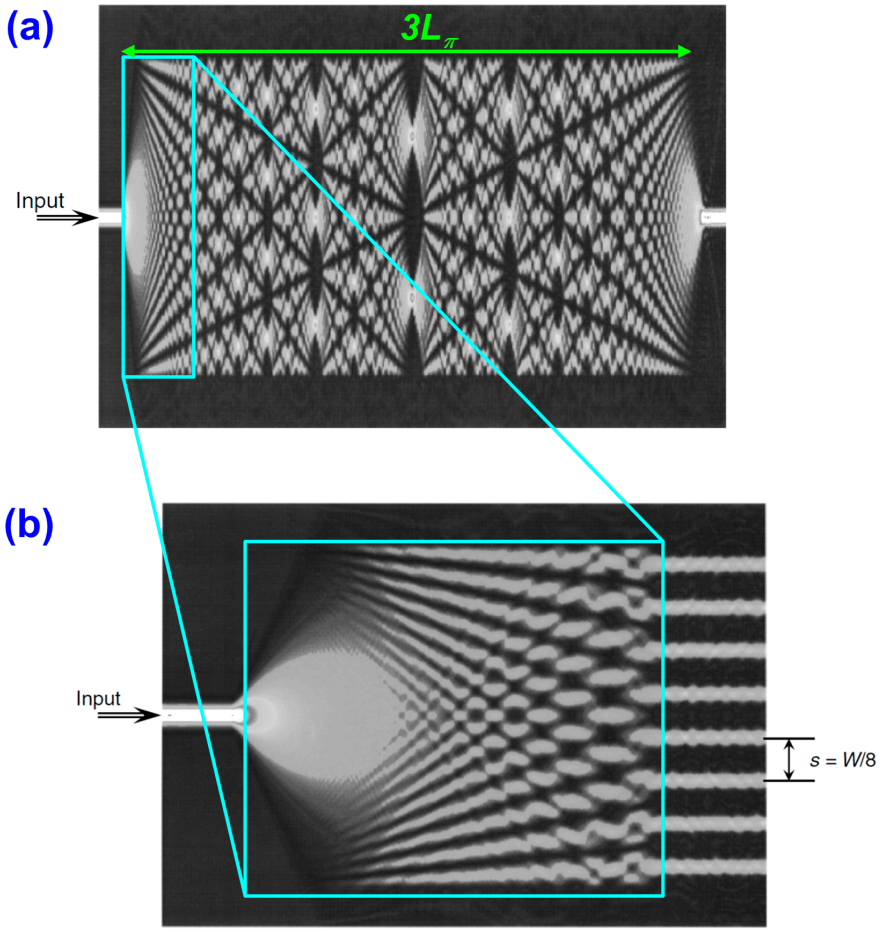

Figure 6.5: Intensity distribution for light input at the center of the MMI section. (a) Since $\Phi(x, y, 0) = \Phi(-x, y, 0)$, the initial image in the (x, y)-plane is already repeated after a propagation length of $3L_\pi$ rather than $6L_\pi$ as suggested by Eq. (6.43). (b) After a length of $3L_\pi/8$, light input at the center of the MMI is split in 8 equal images. By connecting output waveguides to the multimode section at the respective positions, the device can be used as an 1-to-8 power splitter (Figure adapted from [25]).

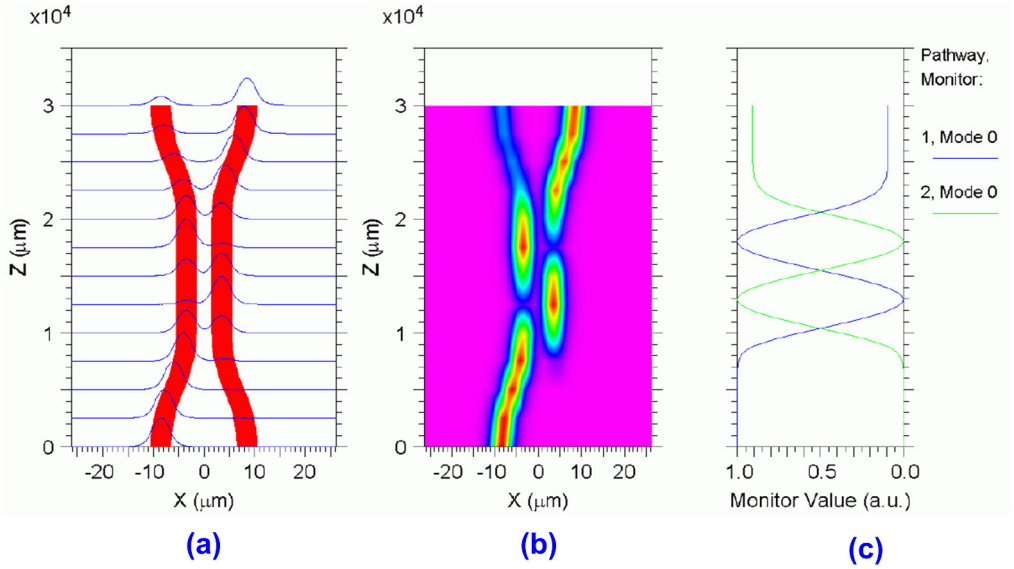

Figure 6.6: Basic principle of a directional coupler: Closely spaced waveguides are running in parallel to each other such that the evanescent tails of individual eigenmodes interact with the neighboring waveguide. This leads to coupling of the guided modes, and the power oscillates back and forth between the waveguides. (a) Optical intensity distribution (blue lines) at various positions along the propagation direction z for the case where light is launched into the left waveguide only; (b) Corresponding color-coded intensity distribution in the (x, y)-plane; (c) Power in the fundamental mode of the left ("1, Mode 0") and the right ("2, Mode 0") waveguide.

6.3 Directional couplers

A directional coupler consists of two closely spaced waveguides running in parallel to each other such that the evanescent tails of the first waveguide's eigenmodes interact with the second waveguide, see Fig. 6.6. As a consequence, the modes of the two waveguides are coupled, and the optical power is oscillating back and forth between the waveguides with a typical beat length L_b. In this section, we will first introduce the general idea of mode coupling in parallel waveguides, and then use this formalism to quantitatively analyze directional couplers.

6.3.1 Mode coupling in parallel waveguides

In isolated z-independent waveguides, all eigenmodes are orthogonal and propagate independently of each other. The complete field within the waveguide can then be written as a linear superposition of eigenmodes with constant amplitudes. This idea can be transferred to the situation where two closely spaced waveguides are coupled by evanescent fields. The total dielectric profile $\epsilon(x, y)$ of the two waveguides can be written as the sum of a background dielectric profile $\epsilon_b(x, y)$ and local perturbations $\Delta\epsilon_1(x, y)$ and $\Delta\epsilon_2(x, y)$ that correspond to the cores of the first and the second waveguide, see Fig. 6.7,

$$\epsilon(x, y) = \epsilon_b(x, y) + \Delta\epsilon_1(x, y) + \Delta\epsilon_2(x, y) \tag{6.48}$$

Let us now assume that the presence of the second waveguide represents only a small perturbation to the mode propagation in the first waveguide and vice versa. In this case, the electromagnetic fields may still be approximated by the modes of the two individual waveguide structures, but the mode amplitudes will change during propagation due to the coupling. This is expressed by z-dependent mode amplitudes $\underline{A}(z)$. Assuming both waveguides to be single-moded and neglecting radiation modes, we may restrict our consideration to the two fundamental modes of the two waveguides, which shall be denoted by subscripts 1 and 2,

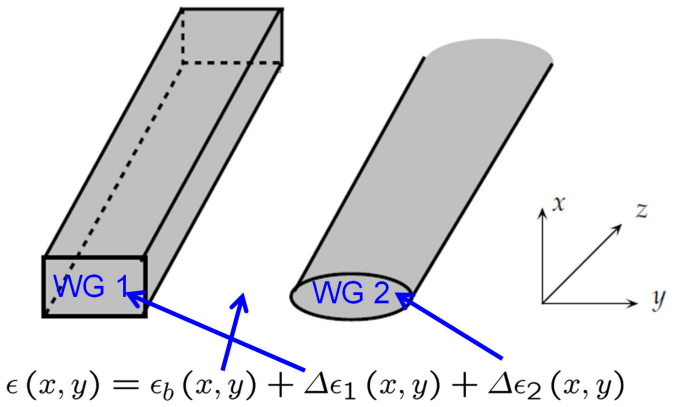

Figure 6.7: Basic principle of the perturbation approach that is used for the analysis of co-directional coupling: The total dielectric profile $\epsilon(x,y)$ of the two waveguides is represented by a sum of the background dielectric profile $\epsilon_b(x,y)$ and local perturbations $\Delta\epsilon_1(x,y)$ and $\Delta\epsilon_2(x,y)$ that correspond to the cores of the two waveguides. (Figure adapted from [5])

$$\underline{\mathbf{E}}(x,y,z) = \sum_{\nu=1}^{2} \underline{A}_\nu(z)\,\underline{\mathcal{E}}_\nu(x,y)\,\mathrm{e}^{-\mathrm{j}\,\beta_\nu z}, \tag{6.49}$$

$$\underline{\mathbf{H}}(x,y,z) = \sum_{\nu=1}^{2} \underline{A}_\nu(z)\,\underline{\mathcal{H}}_\nu(x,y)\,\mathrm{e}^{-\mathrm{j}\,\beta_\nu z}, \tag{6.50}$$

In the following, we use the orthogonality relation, Eq. (6.17) to project out the changes of the mode amplitudes $\partial\underline{A}_\nu/\partial z$ during propagation along z, i.e., we assume that the mode field pairs $\underline{\mathcal{E}}_1$, $\underline{\mathcal{H}}_1$ and $\underline{\mathcal{E}}_2$, $\underline{\mathcal{H}}_2$ are orthogonal to each other. This needs some explanation: Recall that Eq. (6.17) refers to two guided modes which belong to the same waveguide and which are indicated by subscripts 1 and 2. For these modes, orthogonality was proven in Section 6.1.1. In contrast to that, the mode field pairs $\underline{\mathcal{E}}_1$, $\underline{\mathcal{H}}_1$ and $\underline{\mathcal{E}}_2$, $\underline{\mathcal{H}}_2$ in Eqs. (6.49) and (6.50) refer to the fundamental modes of two distinct waveguides, which are denoted by 1 and 2. In a strict sense, they are not orthogonal. However, we may assume that Eq. (6.17) still holds approximately for $\underline{\mathcal{E}}_1$, $\underline{\mathcal{H}}_1$ and $\underline{\mathcal{E}}_2$, $\underline{\mathcal{H}}_2$. To understand this, let us assume that we expand the fundamental mode fields $\underline{\mathcal{E}}_2$, $\underline{\mathcal{H}}_2$ of waveguide 2 by mode fields of waveguide 1. This expansion would mainly consist of high-order guided modes and radiation modes of waveguide 1, since the low-order guided modes of waveguide 1 are close to zero at the position of waveguide 2 and can hence not contribute to the expansion. As a consequence, nearly all modes of the expansion are therefore orthogonal to the fundamental guided mode of waveguide 1, and this also applies to the entire expansion, such that Eq. (6.17) is approximately fulfilled. A more rigorous analysis taking into account the imperfect orthogonality of the mode field pairs $\underline{\mathcal{E}}_1$, $\underline{\mathcal{H}}_1$ and $\underline{\mathcal{E}}_2$, $\underline{\mathcal{H}}_2$ can be found in [25].

The superposition of the fields according to Eqs. (6.49) and (6.50) should satisfy Maxwell's equations for the total dielectric profile,

$$\nabla \times \underline{\mathbf{E}}(x,y,z) = -\,\mathrm{j}\,\omega\mu_0\underline{\mathbf{H}}(x,y,z), \tag{6.51}$$

$$\nabla \times \underline{\mathbf{H}}(x,y,z) = \mathrm{j}\,\omega\left(\epsilon(x,y) + \Delta\epsilon_1(x,y,z) + \Delta\epsilon_2(x,y,z)\right)\underline{\mathbf{E}}(x,y,z). \tag{6.52}$$

Inserting Eq. (6.49) and (6.50) in Eq. (6.51) and using the identity $\nabla \times (\Phi\mathbf{F}) = \Phi(\nabla \times \mathbf{F}) +$

$(\nabla\Phi \times \mathbf{F})$, we obtain

$$\sum_{\nu=1}^{2}\left[\underline{A}_\nu\left(\nabla\times\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right)+\frac{\partial\underline{A}_\nu}{\partial z}\mathbf{e}_z\times\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right] \tag{6.53}$$
$$=-\mathrm{j}\,\omega\mu_0\sum_{\nu=1}^{2}\underline{A}_\nu\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right),$$

where we have omitted the arguments x, y, and z for the sake of readability. Likewise, by inserting Eq. (6.49) and (6.50) in Eq. (6.52), we obtain a similar relation for the magnetic field,

$$\sum_{\nu=1}^{2}\left[\underline{A}_\nu\left(\nabla\times\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right)+\frac{\partial\underline{A}_\nu}{\partial z}\mathbf{e}_z\times\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right] \tag{6.54}$$
$$=\mathrm{j}\,\omega\left(\epsilon_b+\Delta\epsilon_1+\Delta\epsilon_2\right)\sum_{\nu=1}^{2}\underline{A}_\nu\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)$$

Eqs. (6.53) and (6.54) can be simplified by making use of the fact that $\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)$ and $\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)$ must fulfill Maxwell's equations for the respective waveguide,

$$\nabla\times\left(\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right)=-\mathrm{j}\,\omega\mu_0\left(\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right), \tag{6.55}$$
$$\nabla\times\left(\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right)=\mathrm{j}\,\omega\left(\epsilon_b+\Delta\epsilon_\nu\right)\left(\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)\right). \tag{6.56}$$

This leads to the relations

$$\sum_{\nu=1}^{2}\frac{\partial\underline{A}_\nu}{\partial z}\mathbf{e}_z\times\underline{\mathcal{E}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)=0 \tag{6.57}$$

$$\sum_{\nu=1}^{2}\frac{\partial\underline{A}_\nu}{\partial z}\mathbf{e}_z\times\underline{\mathcal{H}}_\nu\exp\left(-\mathrm{j}\,\beta_\nu z\right)=\mathrm{j}\,\omega\left(\Delta\epsilon_2\underline{A}_1\underline{\mathcal{E}}_1\exp\left(-\mathrm{j}\,\beta_1 z\right)+\Delta\epsilon_1\underline{A}_2\underline{\mathcal{E}}_2\exp\left(-\mathrm{j}\,\beta_2 z\right)\right) \tag{6.58}$$

Dot-mutliplying Eq. (6.57) with $\underline{\mathcal{H}}_1^\star(x,y)$ and Eq. (6.58) with $\underline{\mathcal{E}}_1^\star(x,y)$ and subtracting the resulting relations, we obtain

$$\sum_{\nu=1}^{2}\frac{\partial\underline{A}_\nu}{\partial z}\left[\left(\mathbf{e}_z\times\underline{\mathcal{E}}_\nu\right)\cdot\underline{\mathcal{H}}_1^\star-\left(\mathbf{e}_z\times\underline{\mathcal{H}}_\nu\right)\cdot\underline{\mathcal{E}}_1^\star\right]\exp\left(-\mathrm{j}\,\beta_\nu z\right) \tag{6.59}$$
$$=-\mathrm{j}\,\omega\left(\Delta\epsilon_2\underline{A}_1\underline{\mathcal{E}}_1\exp\left(-\mathrm{j}\,\beta_1 z\right)+\Delta\epsilon_1\underline{A}_2\underline{\mathcal{E}}_2\exp\left(-\mathrm{j}\,\beta_2 z\right)\right)$$

We now use the property $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ of the scalar triple products (German: 'Spatprodukt') on the left-hand side of Eq. (6.59), integrate over the entire cross (x,y)-plane, and use the orthogonality relation, Eq. (6.15). This leads to

$$\frac{\partial\underline{A}_1(z)}{\partial z}=-\frac{\mathrm{j}\,\omega}{4\mathcal{P}_1}\left(\iint_{-\infty}^{\infty}\Delta\epsilon_2\underline{A}_1\left|\underline{\mathcal{E}}_1\right|^2\mathrm{d}\,x\,\mathrm{d}\,y,+\iint_{-\infty}^{\infty}\Delta\epsilon_1\underline{A}_2\underline{\mathcal{E}}_2\underline{\mathcal{E}}_1^\star\exp\left(-\mathrm{j}\left(\beta_2-\beta_1\right)z\right)\right)\mathrm{d}\,x\,\mathrm{d}\,y, \tag{6.60}$$

The electric field $\underline{\mathcal{E}}_1(x,y)$ of the first mode is much stronger in the core region of the first waveguide ($\Delta\epsilon_1\neq 0$) than at the position of the second waveguide ($\Delta\epsilon_2\neq 0$) and vice versa. The first expression on the right-hand side of Eq. (6.60) can therefore often be neglected with respect to the second one provided that the index perturbations $\Delta\epsilon_1$ and $\Delta\epsilon_2$ are comparable and that the two waveguide modes have similar shapes. The evolution of $\underline{A}_1(z)$ is hence governed by the relation

$$\frac{\partial\underline{A}_1(z)}{\partial z}=-\frac{\mathrm{j}\,\omega}{4\mathcal{P}_1}\iint_{-\infty}^{\infty}\Delta\epsilon_1\underline{A}_2\underline{\mathcal{E}}_2\underline{\mathcal{E}}_1^\star\exp\left(-\mathrm{j}\left(\beta_2-\beta_1\right)z\right)\mathrm{d}\,x\,\mathrm{d}\,y, \tag{6.61}$$

A similar derivation can be performed for the second waveguide, and we obtain a set of two coupled-mode equations for the interaction of the two waveguides,

$$\frac{\partial \underline{A}_1(z)}{\partial z} = -\mathrm{j}\,\kappa_{12}\underline{A}_2(z)\,\mathrm{e}^{-\mathrm{j}(\beta_2-\beta_1)z}, \tag{6.62}$$

$$\frac{\partial \underline{A}_2(z)}{\partial z} = -\mathrm{j}\,\kappa_{21}\underline{A}_1(z)\,\mathrm{e}^{-\mathrm{j}(\beta_1-\beta_2)z}, \tag{6.63}$$

where the coupling coefficients $\kappa_{\nu\mu}$ are given by

$$\kappa_{\nu\mu} = \frac{\omega}{4\mathcal{P}_\nu}\iint_{-\infty}^{\infty} \Delta\epsilon_\nu(x,y)\,\underline{\mathcal{E}}_\mu(x,y)\cdot\underline{\mathcal{E}}_\nu^\star(x,y)\,\mathrm{d}\,x\,\mathrm{d}\,y \tag{6.64}$$

The quantity $\mathcal{P}_\mu$ is used for power-normalization of the numerically calculated mode fields $\underline{\mathcal{E}}_\nu(x,y)$,

$$\mathcal{P}_\mu = \frac{1}{2}\iint_{-\infty}^{\infty} \mathrm{Re}\left\{\underline{\mathcal{E}}_\mu(x,y)\times\underline{\mathcal{H}}_\mu^\star(x,y)\right\}\cdot\mathbf{e}_z\,\mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.65}$$

Note that $\mathcal{P}_\mu > 0$ in Eq. (6.65) corresponds to a mode field that is propagating in positive $z-$direction, whereas $\mathcal{P}_\mu < 0$ represents propagation in the negative $z-$direction. In this definition, the mode amplitudes $\underline{A}_\mu(z)$ are dimensionless quantities, and the physical power $P_\mu(z)$ carried by a waveguide mode of amplitude $\underline{A}_\mu(z)$ along the propagation direction is given by

$$P_\mu(z) = \left|\underline{A}_\mu(z)\right|^2\mathcal{P}_\mu, \tag{6.66}$$

where, again, $P_\mu(z) > 0$ ($P_\mu(z) < 0$) if power is flowing in positive (negative) $z-$direction. It is useful to normalize all mode fields to the same power $\mathcal{P} = \mathcal{P}_1 = \mathcal{P}_2$. Note that we can always define the phases of the mode fields $\underline{\mathcal{E}}_1(x,y)$ and $\underline{\mathcal{E}}_2(x,y)$ such both κ_{12} and κ_{21} are real quantities.

6.3.2 Analysis of directional couplers

Coupling of identical waveguides

We will now use these equations to describe the working principle of a directional coupler, Fig. 6.6. For simplicity, let us assume that both waveguides have the same cross section and that the mode fields are normalized to the same power, i.e., $\beta_1 = \beta_2 = \beta$ and $\kappa_{12} = \kappa_{21} = \kappa$. The coupled-mode equations can then be written as

$$\frac{\partial \underline{A}_1(z)}{\partial z} = -\mathrm{j}\,\kappa\underline{A}_2(z) \tag{6.67}$$

$$\frac{\partial \underline{A}_2(z)}{\partial z} = -\mathrm{j}\,\kappa\underline{A}_1(z) \tag{6.68}$$

Eliminating one of the two amplitudes, we obtain identical equations for both amplitudes,

$$\frac{\partial^2 \underline{A}_\nu(z)}{\partial z^2} = -\kappa^2\underline{A}_\nu(z), \quad \nu = 1,2 \tag{6.69}$$

If power is coupled only into the first waveguide at $z = 0$, $\underline{A}_1(0) = A_0$ and $\underline{A}_2(0) = 0$, the evolution of the mode amplitudes is given by

$$\underline{A}_1(z) = A_0\cos(\kappa z), \tag{6.70}$$

$$\underline{A}_2(z) = -\mathrm{j}\,A_0\sin(\kappa z), \tag{6.71}$$

i.e., the power is oscillating back and forth between the waveguides, where the phase of the second waveguide mode is lagging behind that of the first driving mode by $\pi/2$. This situation is comparable to that of coupled mechanical pendula with identical resonance frequencies, where power is oscillating back an forth between the two mechanical systems, and where a $\pi/2$ phase shift exists between the oscillations of the two pendula. This can be visualized by an animation linked under [1].

[1] http://www.theorphys.science.ru.nl/people/fasolino/sub_java/pendula/laboratory-en.shtml

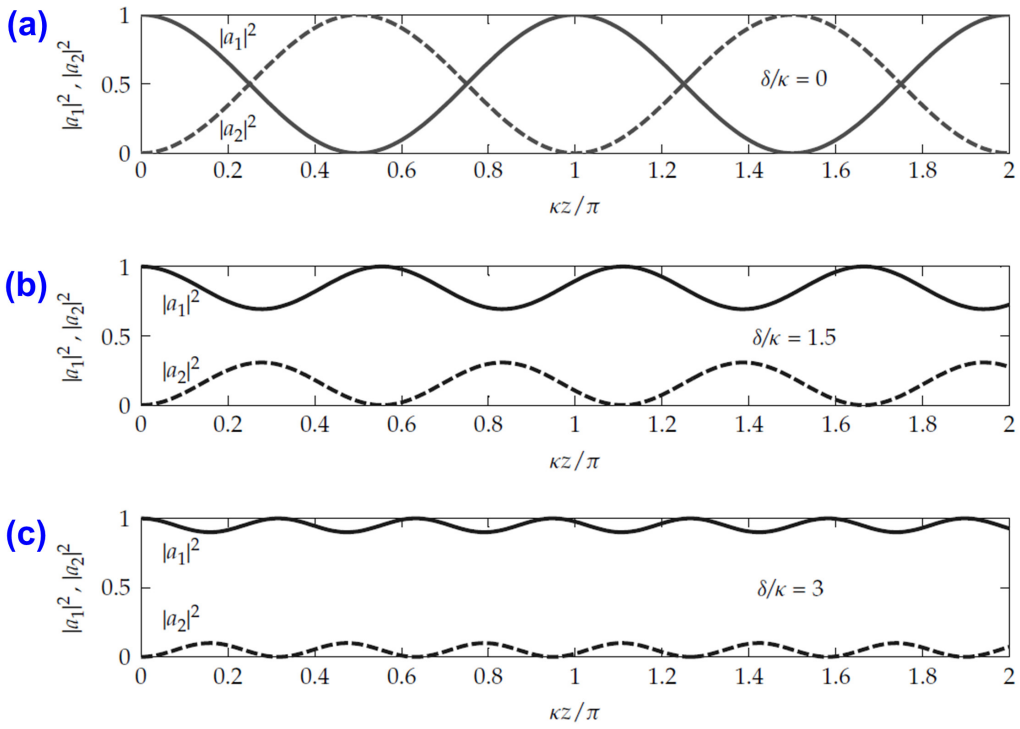

Figure 6.8: Power oscillations in directional couplers for different ratios δ/κ of propagation constant mismatch and coupling strength. (a) Zero propagation constant mismatch leads to a complete transfer of power from one waveguide to the other. (b, c) With increasing propagation constant mismatch, the power transfer between the waveguides becomes smaller. This corresponds to the mechanical analogon of coupled pendula that have different resonance frequencies. (Figure adapted from [5])

Coupling of waveguides with different cross sections

The solution of the coupled-mode equations can be generalized to the case of coupled waveguides that differ in shape and hence propagation constant. For given initial conditions $\underline{A}_1(0)$ and $\underline{A}_2(0)$, the general solution of Eqs. (6.62) and (6.63) can be written as [5]

$$\underline{A}_1(z) = \left[\left(\cos(\sigma z) - \mathrm{j}\,\frac{\delta}{\sigma}\sin(\sigma z)\right)\underline{A}_1(0) - \mathrm{j}\,\frac{\kappa_{12}}{\sigma}\sin(\sigma z)\,\underline{A}_2(0)\right]\mathrm{e}^{\mathrm{j}\,\delta z}, \tag{6.72}$$

$$\underline{A}_2(z) = \left[-\mathrm{j}\,\frac{\kappa_{21}}{\sigma}\sin(\sigma z)\,\underline{A}_1(0) + \left(\cos(\sigma z) + \mathrm{j}\,\frac{\delta}{\sigma}\sin(\sigma z)\right)\underline{A}_2(0)\right]\mathrm{e}^{-\mathrm{j}\,\delta z}, \tag{6.73}$$

where the auxiliary quantities δ and σ are given by

$$\sigma^2 = \kappa^2 + \delta^2, \tag{6.74}$$

$$\delta = \frac{\beta_1 - \beta_2}{2}. \tag{6.75}$$

The quantity δ corresponds to the mismatch of the propagation constants. The degree to which power is transferred between the waveguides depends strongly on the ratio δ/κ of propagation constant mismatch and coupling strength, see Fig. 6.8. Zero propagation constant mismatch leads to a complete transfer of power from one waveguide to the other, whereas for nonzero mismatch, only a fraction of the power is exchanged.

Example: 3dB-splitter based on coupled waveguides

For a 3 dB-coupler, we want equal power levels in both waveguides. For simplicity let us again assume that both waveguides have the same cross section and that the mode fields are normalized to the same power, $\mathcal{P} = \mathcal{P}_1 = \mathcal{P}_2$, $\beta_1 = \beta_2 = \beta$ and $\kappa_{12} = \kappa_{21} = \kappa$. From Eqs. (6.66), (6.70), and (6.71) we find

$$P_1(z) = |A_1(z)|^2 \mathcal{P} = |A_0|^2 \mathcal{P} \cos(\kappa z) \tag{6.76}$$

$$P_2(z) = |A_2(z)|^2 \mathcal{P} = |A_0|^2 \mathcal{P} \sin(\kappa z) \tag{6.77}$$

The two power levels are equal for $\kappa z = \pi/4$, and the corresponding length of the directional coupler then given by

$$L_{3\mathrm{dB}} = \frac{\pi}{4\kappa}. \tag{6.78}$$

6.4 Waveguide gratings

In the last section we have seen that co-propagating guided modes of two waveguides can be coupled by bringing the corresponding waveguides in close proximity to each other. The presence of one waveguide can then be treated as a small perturbation of the other waveguide's index profile. In this section, we will extend this idea to the case of mode coupling within a single waveguide which is subject to perturbations of the dielectric profile. We will first introduce the general idea of mode coupling by refractive index perturbations, and then focus on periodic index perturbations which form the basis of waveguide gratings. Such structures are, e.g., used in fiber Bragg gratings, grating-assisted fiber-chip couplers, distributed feedback (DFB) lasers or distributed Bragg reflector (DBR) lasers.

6.4.1 Perturbation theory and mode coupling

As in Section 6.3.1, we assume a dielectric structure which deviates only slightly from an ideal waveguide. The deviation is treated as a small perturbation $\Delta\epsilon(x,y,z)$ of a z-independent index profile $\epsilon(x,y)$,

$$\epsilon_p(x,y,z) = \epsilon(x,y) + \Delta\epsilon(x,y,z) \tag{6.79}$$

For "small" perturbations, the field can still be approximated by the modes of the unperturbed waveguide structure, but mode amplitudes will change during propagation due to coupling. This is expressed by z-dependent mode amplitudes $\underline{A}_\nu(z)$, and the field expansion for the perturbed waveguide can be written as

$$\underline{\mathbf{E}}(x,y,z) = \sum_\nu \underline{A}_\nu(z)\, \underline{\mathcal{E}}_\nu(x,y) \exp(-\mathrm{j}\,\beta_\nu z) + \sum_\mu \int_\rho \underline{A}_\mu(\rho,z)\, \underline{\mathcal{E}}_{\rho,\mu}(x,y) \exp(-\mathrm{j}\,\beta_\mu(\rho)\, z)\, \mathrm{d}\,\rho, \tag{6.80}$$

$$\underline{\mathbf{H}}(x,y,z) = \sum_\nu \underline{A}_\nu(z)\, \underline{\mathcal{H}}_\nu(x,y) \exp(-\mathrm{j}\,\beta_\nu z) + \sum_\mu \int_\rho \underline{A}_\mu(\rho,z)\, \underline{\mathcal{H}}_{\rho,\mu}(x,y) \exp(-\mathrm{j}\,\beta_\mu(\rho)\, z)\, \mathrm{d}\,\rho, \tag{6.81}$$

where $\underline{A}_\nu(z)$ ($\underline{A}_\mu(\rho,z)$) denote the guided (radiation) mode amplitudes of the waveguide. Clearly, the evolution of the various mode amplitudes depends strongly on the waveguide geometry $\epsilon(x,y)$ and the perturbation $\Delta\epsilon(x,y,z)$. This shall now be analyzed in more detail.

We start again from Maxwell's equations for the perturbed waveguide,

$$\nabla \times \underline{\mathbf{E}}(x,y,z) = -\mathrm{j}\,\omega\mu_0 \underline{\mathbf{H}}(x,y,z), \tag{6.82}$$

$$\nabla \times \underline{\mathbf{H}}(x,y,z) = \mathrm{j}\,\omega\left(\epsilon(x,y) + \Delta\epsilon(x,y,z)\right)\underline{\mathbf{E}}(x,y,z). \tag{6.83}$$

For the sake of simplicity, let us consider only guided modes in this derivation,

$$\underline{\mathbf{E}}(x,y,z) = \sum_\nu \underline{A}_\nu(z)\,\underline{\mathcal{E}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z) \tag{6.84}$$

$$\underline{\mathbf{H}}(x,y,z) = \sum_\nu \underline{A}_\nu(z)\,\underline{\mathcal{H}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z) \tag{6.85}$$

where $\underline{\mathcal{E}}_\nu(x,y)$ and $\underline{\mathcal{H}}_\nu(x,y)$ denote the eigenmode fields of the unperturbed waveguide structure, fulfilling Maxwell's equations for the unperturbed index profile $\epsilon(x,y)$,

$$\nabla \times \left(\underline{\mathcal{E}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z)\right) = -\mathrm{j}\,\omega\mu_0\left(\underline{\mathcal{H}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z)\right), \tag{6.86}$$

$$\nabla \times \left(\underline{\mathcal{H}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z)\right) = \mathrm{j}\,\omega\epsilon(x,y)\left(\underline{\mathcal{E}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z)\right). \tag{6.87}$$

As in Section 6.3.1, we insert Eq. (6.84) and (6.85) in Eq. (6.82) and use the identity $\nabla \times (\Phi\mathbf{F}) = \Phi(\nabla \times \mathbf{F}) + (\nabla\Phi \times \mathbf{F})$. This yields

$$\sum_\nu \left[\underline{A}_\nu \nabla\left(\times\underline{\mathcal{E}}_\nu \exp(-\mathrm{j}\,\beta_\nu z)\right) + \frac{\partial \underline{A}_\nu}{\partial z}\mathbf{e}_z \times \underline{\mathcal{E}}_\nu \exp(-\mathrm{j}\,\beta_\nu z)\right] \tag{6.88}$$
$$= -\mathrm{j}\,\omega\mu_0 \sum_\nu \underline{A}_\nu \underline{\mathcal{H}}_\nu \exp(-\mathrm{j}\,\beta_\nu z),$$

where we have omitted the arguments x, y, and z for the sake of readibility. Likewise, by inserting Eq. (6.84) and (6.85) in Eq. (6.83), we obtain a similar relation for the magnetic field,

$$\sum_\nu \left[\underline{A}_\nu\left(\nabla \times \underline{\mathcal{H}}_\nu \exp(-\mathrm{j}\,\beta_\nu z)\right) + \frac{\partial \underline{A}_\nu}{\partial z}\mathbf{e}_z \times \underline{\mathcal{H}}_\nu \exp(-\mathrm{j}\,\beta_\nu z)\right] \tag{6.89}$$
$$= \mathrm{j}\,\omega\left(\epsilon + \Delta\epsilon\right)\sum_\nu \underline{A}_\nu \underline{\mathcal{E}}_\nu \exp(-\mathrm{j}\,\beta_\nu z)$$

Making use of Eqs. (6.86) and (6.87), we can simplify Eqs. (6.88) and (6.89),

$$\sum_\nu \frac{\partial \underline{A}_\nu(z)}{\partial z}\mathbf{e}_z \times \underline{\mathcal{E}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z) = 0 \tag{6.90}$$

$$\sum_\nu \frac{\partial \underline{A}_\nu(z)}{\partial z}\mathbf{e}_z \times \underline{\mathcal{H}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z) = \mathrm{j}\,\omega\Delta\epsilon\sum_\nu \underline{A}_\nu(z)\,\underline{\mathcal{E}}_\nu(x,y)\exp(-\mathrm{j}\,\beta_\nu z). \tag{6.91}$$

Dot-mutliplying Eq. (6.91) with $\underline{\mathcal{H}}^\star_\mu(x,y)$ and Eq. (6.91) with $\underline{\mathcal{E}}^\star_\mu(x,y)$ and subtracting the resulting relations, we obtain

$$\sum_\nu \frac{\partial \underline{A}_\nu}{\partial z}\left[(\mathbf{e}_z \times \underline{\mathcal{E}}_\nu)\cdot\underline{\mathcal{H}}^\star_\mu - (\mathbf{e}_z \times \underline{\mathcal{H}}_\nu)\cdot\underline{\mathcal{E}}^\star_\mu\right]\exp(-\mathrm{j}\,\beta_\nu z) = -\mathrm{j}\,\omega\Delta\epsilon\sum_\nu \underline{A}_\nu\underline{\mathcal{E}}_\nu\exp(-\mathrm{j}\,\beta_\nu z) \tag{6.92}$$

We again use the property $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ of the scalar triple products on the left hand side of Eq. (6.92), integrate over the entire cross (x,y)-plane, and use the orthogonality relation, Eq. (6.15). This leads to

$$\frac{\partial \underline{A}_\mu(z)}{\partial z} = -\frac{\mathrm{j}\,\omega}{4\mathcal{P}_\mu} \sum_\nu \iint_{-\infty}^{\infty} \Delta\epsilon \underline{\mathcal{E}}_\nu \cdot \underline{\mathcal{E}}_\mu^\star \,\mathrm{d}\,x\,\mathrm{d}\,y\,\underline{A}_\nu\,\mathrm{e}^{-\mathrm{j}(\beta_\nu-\beta_\mu)z}, \tag{6.93}$$

which can be rewritten in the more familiar form of a mode-coupling equation,

$$\frac{\partial \underline{A}_\mu(z)}{\partial z} = -\mathrm{j}\sum_\nu \kappa_{\mu\nu}(z)\,\underline{A}_\nu(z)\,\mathrm{e}^{-\mathrm{j}(\beta_\nu-\beta_\mu)z}, \tag{6.94}$$

In this relation, the coupling coefficients $\kappa_{\mu\nu}(z)$ are given by

$$\kappa_{\mu\nu} = \frac{\omega}{4\mathcal{P}_\mu} \iint_{-\infty}^{\infty} \Delta\epsilon(x,y,z)\,\underline{\mathcal{E}}_\nu(x,y)\cdot\underline{\mathcal{E}}_\mu^\star(x,y)\,\mathrm{d}\,x\,\mathrm{d}\,y, \tag{6.95}$$

and $\mathcal{P}_\mu$ is used for power-normalization of the numerically calculated mode fields $\underline{\mathcal{E}}_\mu(x,y)$ and $\underline{\mathcal{H}}_\mu(x,y)$,

$$\mathcal{P}_\mu = \frac{1}{2}\iint_{-\infty}^{\infty} \mathrm{Re}\left\{\underline{\mathcal{E}}_\mu(x,y)\times\underline{\mathcal{H}}_\mu^\star(x,y)\right\}\cdot\mathbf{e}_z\,\mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.96}$$

Eq. (6.94) can be interpreted in an intuitive way: The perturbation $\Delta\epsilon(x,y,z)$ is "weighted" by the electric fields of the mode. Perturbations are hence more "effective" if they occur in a region where the electric fields are strong. Note that $\mathcal{P}_\mu > 0$ in Eq. (6.96) corresponds to a mode field that is propagating in positive $z-$direction, whereas $\mathcal{P}_\mu < 0$ represents propagation in the negative $z-$direction. As before, the mode amplitudes $\underline{A}_\mu(z)$ are dimensionless quantities, and the physical power $P_\mu(z)$ carried by a waveguide mode of amplitude $\underline{A}_\mu(z)$ along the propagation direction is given by

$$P_\mu(z) = \left|\underline{A}_\mu(z)\right|^2 \mathcal{P}_\mu, \tag{6.97}$$

where $P_\mu(z) > 0$ ($P_\mu(z) < 0$) if power is flowing in positive (negative) $z-$direction. If all mode fields are normalized to the same power $\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_3 = \ldots$, and if $\Delta\epsilon$ is real (no loss or gain), then the coupling coefficients $\kappa_{\nu\mu}$ obey the following symmetry relations:

$$\kappa_{\nu\mu} = \begin{cases} -\kappa_{\mu\nu}^\star & \text{for counterpropagating modes} \\ \kappa_{\mu\nu}^\star & \text{for co-propagating modes} \end{cases} \tag{6.98}$$

Note that we can always adjust the phases of the mode fields $\underline{\mathcal{E}}_\nu(x,y)$ and $\underline{\mathcal{E}}_\mu(x,y)$ such that a certain coupling coefficient $\kappa_{\nu\mu}$ becomes real.

6.4.2 Mode coupling by periodic perturbations

Let us now consider perturbations of the refractive index profile $\Delta\epsilon(x,y,z)$ that are periodic in z-direction with periodicity Λ, Fig. 6.9. We can expand the z-dependence in a Fourier series

$$\Delta\epsilon(x,y,z) = \sum_{q=-\infty}^{\infty} \Delta\epsilon_q(x,y)\,\mathrm{e}^{-\mathrm{j}\,qKz}, \tag{6.99}$$

where the periodicity is given by the "spatial wavenumber"

$$K = \frac{2\pi}{\Lambda}, \tag{6.100}$$

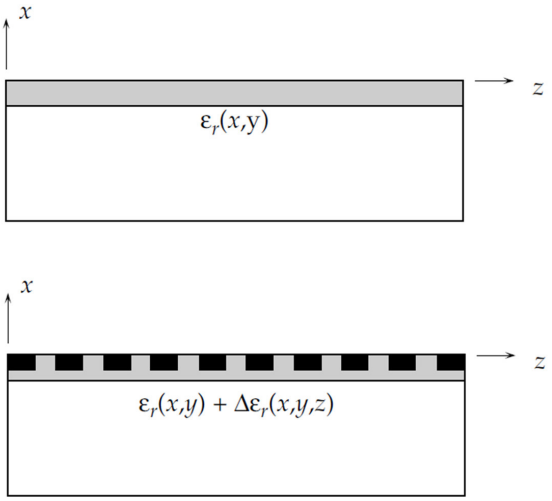

Figure 6.9: A waveguide grating consists of a periodic perturbation $\Delta\epsilon(x,y,z)$ of an optical waveguide with permittivity $\epsilon(x,y)$ which is otherwise homogenous along the propagation direction (z-direction). Two guided modes with propagation constants β_ν and β_μ are coupled by the grating if the difference of the propagation constants corresponds to a spatial frequency component of the index perturbation. (Figure adapted from [5])

and where the coefficients $\Delta\epsilon_q(x,y)$ define the shape of the grating elements. The coupled mode equations, Eq. (6.94), can then be written as

$$\frac{\partial \underline{A}_\mu(z)}{\partial z} = -\mathrm{j}\sum_\nu \sum_q \kappa_{\mu\nu,q}\underline{A}_\nu(z)\,\mathrm{e}^{-\mathrm{j}(\beta_\nu-\beta_\mu+qK)z} \tag{6.101}$$

where

$$\kappa_{\mu\nu,q} = \frac{\omega}{4\mathcal{P}_\mu}\iint_{-\infty}^{\infty}\Delta\epsilon_q(x,y)\,\underline{\mathcal{E}}_\nu(x,y)\cdot\underline{\mathcal{E}}_\mu^\star(x,y)\,\mathrm{d}\,x\,\mathrm{d}\,y \tag{6.102}$$

The right-hand side of Eq. (6.101) contains a sum over all contributions that originate from the interaction of the waveguide modes with the spatial harmonic components of the grating. To solve the relation, we need to integrate over z. For $\beta_\nu - \beta_\mu + qK \neq 0$, we will see periodic oscillations of the right-hand side as a function of z, which will not contribute significantly to the integral. Significant contributions to the change of the mode amplitudes $\underline{A}_\mu(z)$ originate therefore only from combinations of β_ν, β_μ and qK for which

$$\beta_\nu - \beta_\mu + qK \approx 0. \tag{6.103}$$

In other words: Two modes with propagation constants β_ν and β_μ are coupled by the spatial frequency component qK that corresponds to the difference $\beta_\nu - \beta_\mu$.

Example 1: Fiber-chip-coupling via gratings

A periodic grating etched into the surface of an integrated optical waveguide can be used to couple guided modes to radiation modes and hence to couple light into and out of the waveguide, see Figure 6.10. In this case, a guided mode with propagation constant β_{guid} shall be coupled to a radiation mode of propagation constant β_{rad}. If the radiated field propagates under an angle Θ with respect to the vertical direction, the propagation constants along the z-direction can be expressed as

$$\beta_{\mathrm{rad}} = n_3 k_0 \sin\Theta \tag{6.104}$$

$$\beta_{\mathrm{guid}} = n_e k_0 \tag{6.105}$$

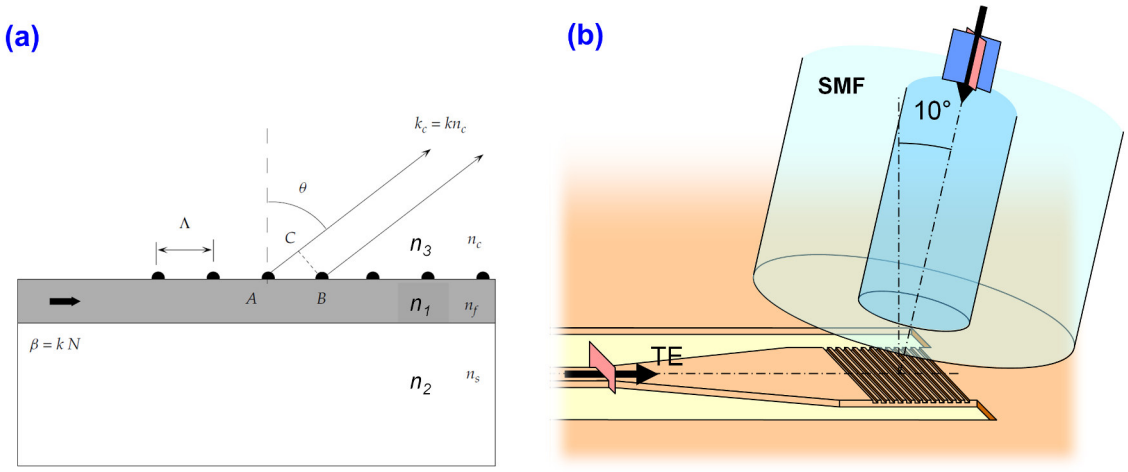

Figure 6.10: Grating assisted fiber-chip coupling (a) Principle of coupling guided modes to radiation modes by a periodic perturbation: The angle Θ under which the field is radiated with respect to the vertical direction is dictated by the fact that the spatial frequency $K = 2\pi/\Lambda$ of the grating must account for the difference $\beta_{\text{guid}} - \beta_{\text{rad}}$ of the guided and the radiation mode's propagation constant along the z-direction. (Figure adapted from [5]) (b) Schematic of a fiber-chip coupler that connects a standard single-mode fiber (SMF) to a grating that is etched into the surface of a silicon-on-insulator waveguide. (Figure adapted from http://www.helios-project.eu)

Using Eq. (6.103) and assuming a first-order grating ($|q| = 1$), we can calculate the fundamental spatial wavenumber and the period of the grating structure,

$$K = \beta_{\text{guid}} - \beta_{\text{rad}} \tag{6.106}$$

$$\Lambda = \frac{\lambda}{n_e - n_3 \sin\Theta} \tag{6.107}$$

For a typical case of a silicon-on-insulator (SOI) waveguide with effective refractive index $n_e = 2.5$ operated at a wavelength of $\lambda = 1.55\,\mu$m and a targeted propagation angle $\Theta = 10^\circ$, we obtain a grating period of 670 nm. In certain cases, diffraction gratings can radiate light in more directions that correspond to different so-called diffraction orders $|q|$ of the grating, Fig. 6.11.

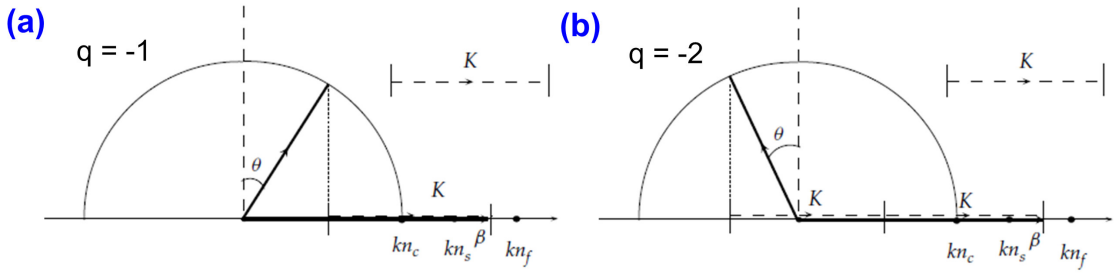

Figure 6.11: Graphical illustration of the grating equation 6.103 for $\beta_\nu = \beta_{\text{guid}}$ and $\beta_\mu = \beta_{\text{rad}}$ for (a) $q = -1$ and (b) $q = -2$. Waves radiated in the directions that correspond to $|q| = 1$ ($|q| = 2$) are referred to as first- and second-order diffractions. (Figure adapted from [5])

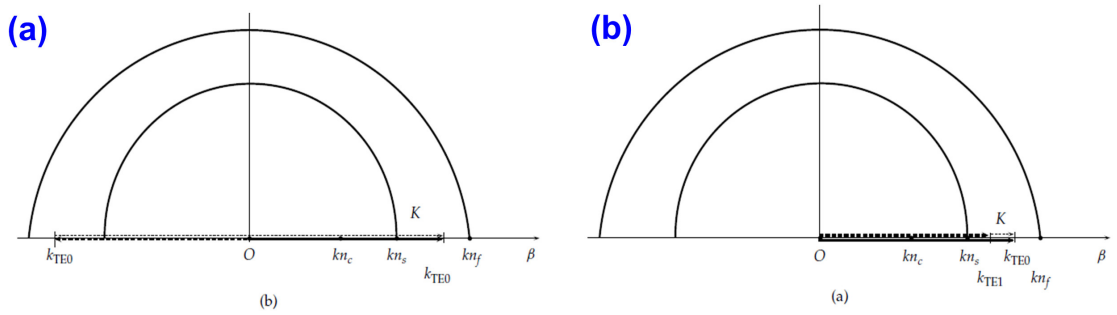

Figure 6.12: Graphical illustration of grating assisted mode conversion. (a) Conversion of the fundamental transverse-electric mode ("TE0") to the next higher-order TE mode ("TE1") requires a grating with a small wavenumber (big period), that just accounts for the difference of the propagation constants. (b) Coupling of a forward-propagating TE mode to its backward propagating counterpart. The wavenumber of the grating must be equal to twice the propagation constant of the mode. (Figure adapted from [5])

Example 2: Grating-assisted mode conversion

Waveguide gratings are also useful to convert one guided mode into a different guided mode. The grating must then be designed such that the fundamental spatial wavenumber matches the difference of the propagation constant of the two modes. An example is sketched as a wave vector diagram in Fig. 6.12 (a). To couple two co-propagating modes, a sufficiently large grating period (small K) is needed. Higher-order harmonics ($q = 2, 3., \dots$) of the grating might then lead to losses by coupling the guided modes to radiation modes.

Example 3: Contra-directional coupling

In the same way, we can couple two counter-propagating modes of the same waveguide by a sufficiently large spatial wavenumber K, see wavevector diagram in Fig. 6.12 (b). We will now analyze this case in more detail. For simplicity, let us consider two counter-propagating modes with "identical" field distributions and hence propagation constants of equal magnitude but opposite sign,

$$\beta_r = -\beta_l = \beta \tag{6.108}$$

$$\mathcal{P}_r = -\mathcal{P}_l \tag{6.109}$$

$$\kappa_{rl} = -\kappa_{lr} = \kappa \tag{6.110}$$

The two counter-propagating modes are most efficiently coupled by a grating with a fundamental wavenumber $K \approx 2\beta$ or $\Lambda \approx \pi/\beta$. We assume further that the perturbation $\Delta\epsilon\,(x, y, z)$ has zero mean value along z, i.e., the zero-order Fourier coefficient vanishes, $\Delta\epsilon_0\,(x, y) = 0$. Considering only the dominant contributions, i.e., the expressions that do not vary rapidly with z on the left-hand side of Eq. (6.101), we can write down the coupled-mode equations for this case,

$$\frac{\partial \underline{A}_\mathrm{l}\,(z)}{\partial z} = \mathrm{j}\,\kappa \underline{A}_\mathrm{r}\,(z)\,\mathrm{e}^{-\mathrm{j}(2\beta-K)z}, \tag{6.111}$$

$$\frac{\partial \underline{A}_\mathrm{r}\,(z)}{\partial z} = -\mathrm{j}\,\kappa \underline{A}_\mathrm{l}\,(z)\,\mathrm{e}^{-\mathrm{j}(-2\beta+K)z}, \tag{6.112}$$

where the subscript l (r) denotes the mode propagating to the left (right), Fig. 6.13. The coupling coefficient κ is related to the first harmonic $\Delta\epsilon_1\,(x, y)$ of the periodic dielectric profile perturbation,

$$\kappa = \frac{\omega}{4\mathcal{P}_\mu} \iint_{-\infty}^{\infty} \Delta\epsilon_1\,(x, y)\,\underline{\mathcal{E}}_\mathrm{l}(x, y) \cdot \underline{\mathcal{E}}_\mathrm{r}^{\star}(x, y)\,\mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.113}$$

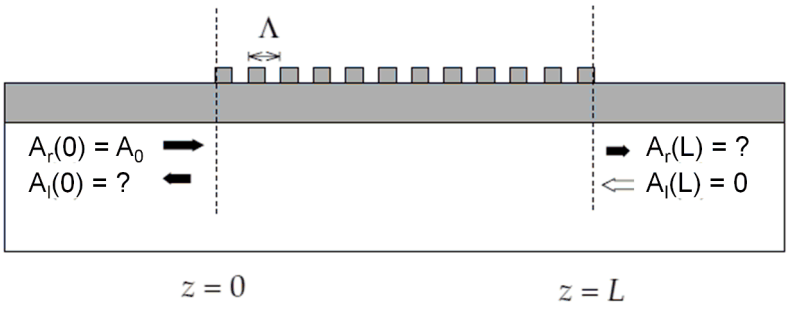

Figure 6.13: Coupling of counterpropagating modes in a waveguide Bragg grating. $\underline{A}_r(z)$ ($\underline{A}_l(z)$) denotes the local mode amplitude propagating to the right (left). Strong coupling of modes with propagation constants $\pm\beta$ is obtained for $\Lambda \approx \pi/\beta$. (Figure adapted from [5])

Defining the so-called detuning parameter,

$$\delta = \beta - \frac{K}{2} \tag{6.114}$$

we can state the coupled-mode equations for contra-directional coupling in the following form:

$$\frac{\partial \underline{A}_l(z)}{\partial z} = \mathrm{j}\,\kappa \underline{A}_r(z)\,\mathrm{e}^{-\mathrm{j}\,2\delta z} \tag{6.115}$$

$$\frac{\partial \underline{A}_r(z)}{\partial z} = -\mathrm{j}\,\kappa \underline{A}_l(z)\,\mathrm{e}^{\mathrm{j}\,2\delta z} \tag{6.116}$$

In many cases of practical interest, light is coupled to the grating from one side only, $\underline{A}_r(z=0) = A_0$ and $\underline{A}_l(z=L) = 0$. The general solution of Eqs. (6.115) and (6.116) is then given by [5]

$$\underline{A}_r(z) = A_0 \frac{\sigma\cosh(\sigma(z-L)) - \mathrm{j}\,\delta\sinh(\sigma(L-z))}{\sigma\cosh(\sigma L) + \mathrm{j}\,\delta\sinh(\sigma L)} \tag{6.117}$$

$$\underline{A}_l(z) = A_0 \frac{\mathrm{j}\,\kappa\sinh(\sigma(z-L))}{\sigma\cosh(\sigma L) + \mathrm{j}\,\delta\sinh(\sigma L)} \tag{6.118}$$

where the auxiliary parameter σ is defined by

$$\sigma^2 = \kappa^2 - \delta^2. \tag{6.119}$$

The power levels of the forward and backward propagating modes within the waveguide grating are depicted in Fig. 6.14 for operation at the Bragg frequency ($\delta = 0$) and for a normalized detuning of $\delta/\kappa = 1.2$. From Eq. (6.117), we can derive the peak reflectance for zero detuning ($\delta = 0$):

$$|\Gamma|^2_{\max} = \frac{|\underline{A}_l(0)|^2}{|\underline{A}_r(0)|^2} = \tanh^2(\kappa L)\,. \tag{6.120}$$

The reflectance asymptotically approaches 1 as the grating length increases. Reflection and transmission spectra of waveguide gratings with different lengths and coupling strengths are depicted in Fig. 6.15. Note that the horizontal axis (detuning) is normalized by the coupling coefficient κ. To obtain a narrow-band reflection with a peak reflection factor close to 1, we hence need a long grating (κL large) with a small coupling factor κ. Such gratings can, e.g., be realized by UV exposure of optical fibers, which creates a very small refractive index perturbation (small κ) over a length of centimeters.

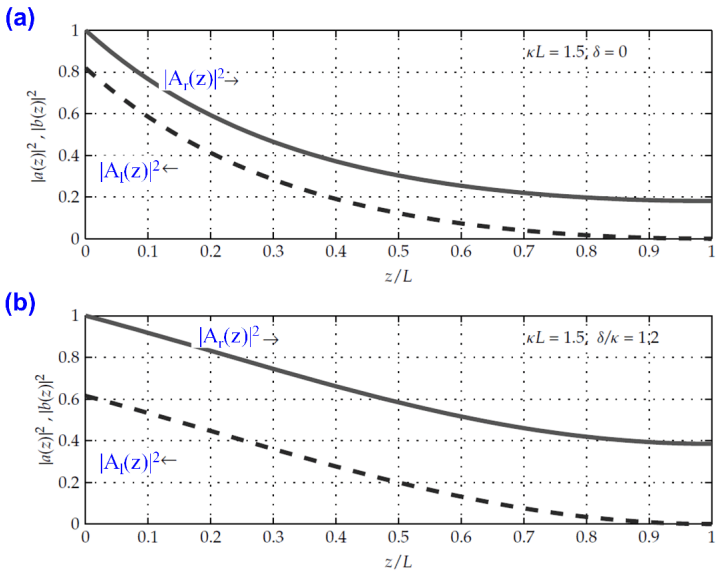

Figure 6.14: Normalized power distributions of modes propagating to the right ($|\underline{A}_r(z)|^2$) and to the left ($|\underline{A}_l(z)|^2$). Light with unity amplitude is launched to the grating from the right, $\underline{A}_r(0) = 1$, $\underline{A}_l(L) = 0$. The power transmittance and reflectance are given by $T = |\underline{A}_r(L)|^2 / |\underline{A}_r(0)|^2$ and $\Gamma = |\underline{A}_l(0)|^2 / |\underline{A}_r(0)|^2$, respectively. (a) Bragg grating operated at the Bragg frequency with zero detuning ($\delta = 0$); (b) Normalized detuning of $\delta/\kappa = 1.2$. (Figure adapted from [5])

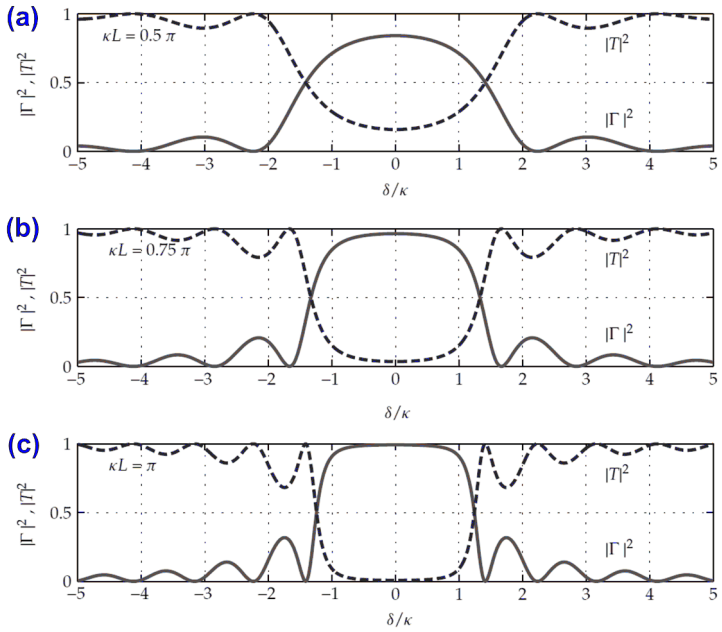

Figure 6.15: Reflectance and transmittance of different Bragg gratings as a function of normalized detuning δ/κ from the Bragg frequency. (a) Weak coupling, $\kappa L = 0.5\pi$; (b) Medium coupling, $\kappa L = 0.75\pi$; (c) Strong coupling, $\kappa L = \pi$. Note that the horizontal axis (detuning) is normalized by the coupling coefficient κ. To obtain a narrow-band reflection with a peak reflection factor close to 1, we hence need a long grating (κL large) with a small coupling factor κ. (Figure adapted from [5])

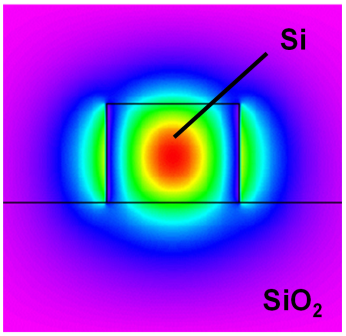

Figure 6.16: Mode field distribution of the fundamental $\underline{\mathcal{E}}_x^{(1,1)}$-mode (quasi-TE mode) in a silicon-on-insulator (SOI) waveguide. If silicon is affected by absorption loss, the complex refractive index has a nonzero imaginary part in the core region. This can be interpreted as a perturbation of the refractive index profile of an otherwise real refractive index profile.

6.5 Material gain and absorption in optical waveguides

Material absorption and gain are common phenomena in optical waveguides. Gain or loss is represented by the imaginary part n_i of the complex refractive index $\underline{n}$. However, in many cases of practical interest, n_i varies considerably over the cross section of the waveguide. In semiconductor optical amplifiers, some parts of the waveguide cross section might e.g. exhibit optical gain, whereas others might be lossy. It is intuitively clear that material gain or loss is more significant if it occurs in a region of high optical intensity. In this section we will quantitatively analyze the relation between modal loss or gain and the underlying material properties.

Absorption loss

As an example, let us consider a silicon-on-insulator (SOI) waveguide for which the waveguide core is affected by absorption loss, see Fig. 6.16 for a sketch of the mode field. As discussed in Section (2.3), the complex refractive index of a material depends generally on contributions from both bound and free charges. These contributions can in principle be taken into account by appropriate Sellmeier equations. However, in many cases of practical interest, empirical equations are used to describe the effect of free carriers. For silicon, the contribution of free carriers to the real and the imaginary part of the complex refractive index at 1550 nm can be modeled by [?]

$$\Delta n_r = -\left(8.8\times10^{-4}\frac{N_e}{\mathrm{cm}^{-3}} + 8.5\left(\frac{N_h}{\mathrm{cm}^{-3}}\right)^{0.8}\right)\times10^{-18}, \tag{6.121}$$

$$\Delta\alpha_m = 2k_0\Delta n_i = \left(8.5\frac{N_e}{\mathrm{cm}^{-3}} + 6.0\frac{N_h}{\mathrm{cm}^{-3}}\right)\times10^{-18}\mathrm{cm}^{-1}. \tag{6.122}$$

The influence of free-carrier absorption on the waveguide mode can be estimated by applying the coupled-mode theory. Assume that only the core of the SOI waveguide is affected by material absorption and hence features a nonzero imaginary part n_i of the complex refractive index. If we interpret the imaginary part as a small perturbation of an ideal lossless waveguide, we can calculate the corresponding change of the dielectric permittivity $\underline{\epsilon} = \epsilon_0\underline{n}^2 = \epsilon_0\left(n - \mathrm{j}\,n_i\right)^2 \approx \epsilon_0 n^2 - \mathrm{j}\,\epsilon_0\,2nn_i$,

$$\Delta\underline{\epsilon} = -\,\mathrm{j}\,\epsilon_0\,2nn_i. \tag{6.123}$$

Inserting this perturbation in Eq. (6.93), we can derive an equation for the power decay of the ν-th waveguide mode,

$$\frac{\partial A_\nu}{\partial z} = -\frac{1}{2}\alpha_\nu A_\nu \tag{6.124}$$

where the modal power attenuation coefficient α_ν is given by

$$\alpha_\nu = \frac{\omega\epsilon_0}{\mathcal{P}_\nu} \iint_{-\infty}^{\infty} n\,(x,y)\, n_i\,(x,y)\, |\underline{\mathcal{E}}_\nu\,(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.125}$$

Eq. (6.124) is solved by an exponentially decaying power profile,

$$|A_\nu(z)|^2 = |A_\nu(0)|^2\, \mathrm{e}^{-\alpha_\nu z} \tag{6.126}$$

If only a certain region the waveguide core is subject to material absorption ("loss region", l.r. in the following), and if the imaginary part of the complex refractive index is constant within this part ($n_i\,(x,y) = n_i$ within the lossy region, $n_i\,(x,y) = 0$ elsewhere), the modal power attenuation coefficient α_ν is related to the material absorption $\alpha_m = 2k_0 n_i$ by

$$\alpha_\nu = \Gamma_\nu \alpha_m, \tag{6.127}$$

where Γ_ν is the so-called field confinement factor,

$$\Gamma_\nu = \frac{\omega\epsilon_0 n_{\mathrm{l.r.}}}{2k_0\mathcal{P}_\nu} \iint_{\mathrm{l.r.}} |\underline{\mathcal{E}}_\nu(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y, \tag{6.128}$$

where $\iint_{\mathrm{l.r.}} \ldots \mathrm{d}\,x\,\mathrm{d}\,y$ indicates an integral over the lossy region of the waveguide, and where $n_{\mathrm{l.r.}}$ is the real part of the refractive index within this region. To obtain the modal loss, the material loss has hence to be weighted by the field confinement factor Γ_ν of the mode field to the lossy region. Γ_ν can hence be thought of as the fraction of power propagating in the region which is affected by optical loss. For the scalar fields of low index-contrast waveguides, the field confinement factor can be approximated by

$$\Gamma_\nu \approx \frac{\iint_{\mathrm{l.r.}} |\underline{\Psi}_\nu(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y}{\iint_{-\infty}^{\infty} |\underline{\Psi}_\nu(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y} \tag{6.129}$$

Active waveguides

The same principle can be applied to active waveguides, where only a portion of the cross section shows optical gain. This is, e.g., the case in lasers and semiconductor optical amplifiers where inversion is only achieved in a small quantum well region, which covers only a fraction of the waveguide cross section, Fig. 6.17. Interpreting again the imaginary part of the complex refractive index as a perturbation of the dielectric profile of the waveguide, we obtain from Eq. (6.93) a differential equation for the mode amplitudes,

$$\frac{\partial A_\nu}{\partial z} = \frac{1}{2} g_\nu A_\nu, \tag{6.130}$$

which leads to an exponential increase of power along z,

$$\left|A_\nu(z)^2\right| = \left|A_\nu(0)^2\right| \mathrm{e}^{g_\nu z} \tag{6.131}$$

The modal gain coefficient g_ν is linked to the material gain g_m by

$$g_\nu = \Gamma_\nu g_m, \tag{6.132}$$

where the field confinement factor Γ_ν can be interpreted as the fraction of optical power carried in the gain region (g.r.),

$$\Gamma_\nu = \frac{\omega\epsilon_0 n_{\mathrm{g.r.}}}{2k_0\mathcal{P}_\nu} \iint_{\mathrm{g.r.}} |\underline{\mathcal{E}}_\nu(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y. \tag{6.133}$$

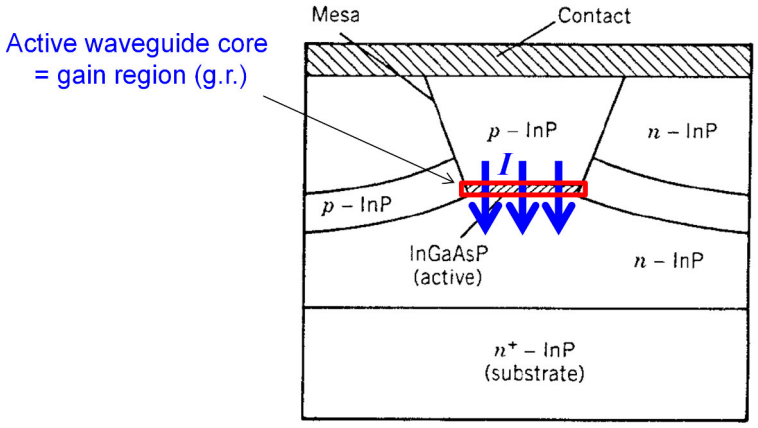

Figure 6.17: Modal gain in an active optical waveguide: The waveguide consists of a core region made from InGaAsP, which is part of a pn-junction operated in forward direction. Within the waveguide core, population inversion is obtained by sufficiently large external currents. The modal gain g_ν is obtained by weighting the material gain g_m of the core region with the corresponding field confinement factor of the mode, Eqs. (6.132), (6.133), and (6.134). (Figure adapted from [2])

For scalar mode fields of low-index contrast waveguides, the field confinement factor can be approximated by

$$\Gamma_\nu \approx \frac{\iint_{\text{g.r.}} |\underline{\Psi}_\nu(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y}{\iint_{-\infty}^{\infty} |\underline{\Psi}_\nu(x,y)|^2 \,\mathrm{d}\,x\,\mathrm{d}\,y}. \tag{6.134}$$

In lasers or optical amplifiers, a typical number for the field confinement factor amounts to $\Gamma \approx 20\%$.

6.6 Bent waveguides

Bent waveguides are essential building blocks of integrated optical circuits. In contrast to straight waveguides, bent waveguide show fundamental radiation loss. An intuitive explanation is given in Fig. 6.18: The phase front of the guided mode is rotating around the center of the bend. Because the group velocity of the phase fronts cannot exceed the speed of light (c/n), the phase fronts must "bend back" and cause radiation. Radiation losses increase exponentially with decreasing bend radius.

For a theoretical analysis of bent waveguides, we assume weakly inhomogeneous media and reduce the problem to two dimensions by applying the effective index method, see Section 4.3. We can then write the scalar wave equation for the dominant electric field component $\underline{\Psi}$ in polar coordinates (r, φ),

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \underline{\Psi}(r,\varphi)}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 \underline{\Psi}(r,\varphi)}{\partial \varphi^2} + k_0^2 n^2(r)\,\underline{\Psi}(r,\varphi) = 0 \tag{6.135}$$

A separation ansatz is used to split up the r- and the φ-dependence,

$$\underline{\Psi}(r,\varphi) = g(r)\,h(\varphi) \tag{6.136}$$

This leads to the relation

$$\frac{r^2}{g(r)}\frac{\partial^2 g(r)}{\partial r^2} + \frac{r}{g(r)}\frac{\partial g(r)}{\partial r} + k_0^2 n^2(r)\,r^2 = -\frac{1}{h(\varphi)}\frac{\partial^2 h(\varphi)}{\partial \varphi^2}, \tag{6.137}$$

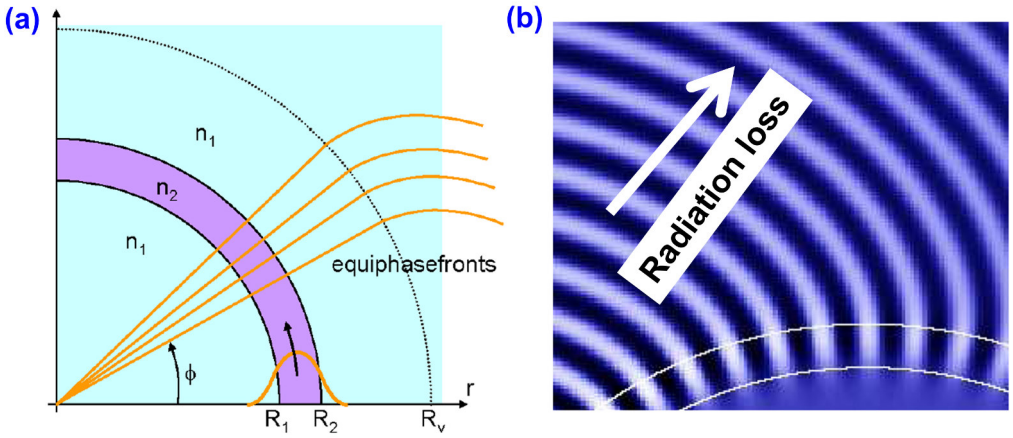

Figure 6.18: Phase fronts of light propagating along a waveguide bend: The group velocity of the phase fronts cannot exceed the speed of light (c/n). Therefore the outer parts of the phase fronts "bend back" and cause radiation. Radiation losses increase exponentially with decreasing bend radius. (Figure adapted from [4])

where both sides must be constant. For the φ-dependence we obtain the general solution

$$h(\varphi) = C_1 \, \mathrm{e}^{\pm \mathrm{j} \, \beta_\varphi \varphi}, \tag{6.138}$$

which corresponds to wave propagation in the azimuthal direction, i.e., along the bend. Inserting Eq. (6.138) in Eq. (6.137), we obtain a differential equation for the associated lateral field profile $g(r)$:

$$r^2 \frac{\partial^2 g(r)}{\partial r^2} + r \frac{\partial g(r)}{\partial r} + \left(k_0^2 n^2(r) \, r^2 - \beta_\varphi^2\right) g(r) = 0 \tag{6.139}$$

To solve this equation for arbitrary index profiles $n(r)$, we use a transformation of the radial coordinate r,

$$r = R_t \, \mathrm{e}^{\frac{u}{R_t}} \qquad u = R_t \ln\left(\frac{r}{R_t}\right) \qquad u, r > 0, \tag{6.140}$$

where R_t denotes a free parameter of the transformation which should advantageously be chosen close to the bend radius. The associated mode profile and index profile are denoted by a bar,

$$g(r) = \overline{g}(u(r)) \tag{6.141}$$

$$n(r) = \overline{n}(u(r)) \tag{6.142}$$

Inserting Eqs. (6.141) and (6.142) in Eq. (6.139), the transformed form of the Helmholtz equation can be written as

$$\frac{\partial^2 \overline{g}(u)}{\partial u^2} + \left(k_0^2 n_t^2(u) - \beta_t^2\right) \overline{g}(u) = 0 \tag{6.143}$$

where the transformed index profile $n_t(u)$ and the transformed propagation constant β_t are given by

$$n_t(u) = n\left(R_t \, \mathrm{e}^{\frac{u}{R_t}}\right) R_t \, \mathrm{e}^{\frac{u}{R_t}}, \tag{6.144}$$

$$\beta_t = \frac{\beta_\varphi}{R_t}. \tag{6.145}$$

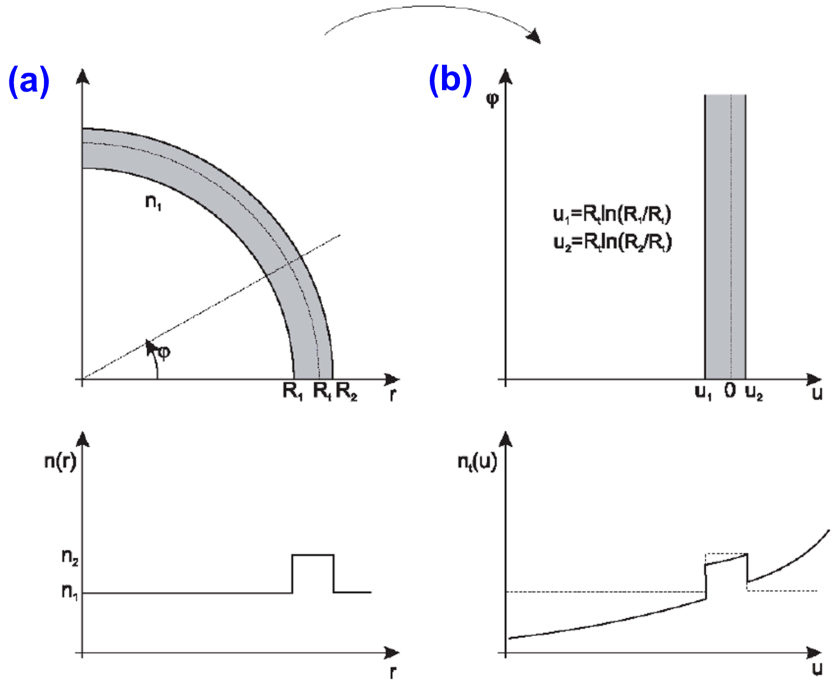

Figure 6.19: Coordinate transformation for a bend waveguide: (a) The original refractive index profile consists of a circular bend in a step-index waveguide. The radial coordinate r is transformed to a coordinate u by using Eq. (6.140). (b) After transformation, the Helmholtz equation in the (u, φ)-coordinate system assumes the same form as in a Cartesian coordinate system if the refractive index profile $n(r)$ is replaced by the transformed profile $n_t(u)$. Propagation in bent waveguides can therefore be calculated by solving an equivalent straight waveguide with a "deformed" refractive index profile. The refractive index of the waveguide is increased for large values of u, and the respective mode fields are "pressed" towards the outer side of the bend. (Figure adapted from [4]).

The transformed Helmholtz equation in the (u, φ)-coordinate system has exactly the same form as in a Cartesian coordinate system if the refractive index profile is replaced by the transformed profile $n_t(u)$. Propagation in bent waveguides can therefore be calculated by solving the equivalent straight waveguide, see Fig. 6.19.

To a first-order approximation, the transformation of the index profile according to Eq. (6.144) corresponds to a local "tilt" of the index profile of the waveguide. Power leakage will occur from the guided mode to the region defined by

$$\beta_t = n_{e,t} k_0 < n_t(u)\, k_0 \;\Leftrightarrow\; n_t(u) > n_{e,t}, \tag{6.146}$$

where $n_{e,t}$ denotes effective index of the mode in the transformed waveguide profile. This loss can be interpreted as „tunneling" of photons through the „potential barrier" defined by the curvature and the index contrast, see Fig. 6.20. Decreasing the radius of curvature makes the potential barrier narrower, and power leakage increases exponentially.

Since the mode field will concentrate in the region with highest refractive index, i.e., the mode will be „pressed" towards the outer side of the bend by the tilt of the index profile. This leads to adaptation losses at the transition between the straight and the bent section, which can be mitigated by a lateral offset at the transition, see Fig. 6.21 (a). Bend radii as small as 1 μm can be realized with high index-contrast silicon-on-insulator (SOI) waveguides. In this case, numerically optimized curve designs are needed to avoid losses at the transition of the straight waveguide to the bend, Fig. 6.21 (b). For wide waveguides and strong curvatures, the mode will entirely be guided

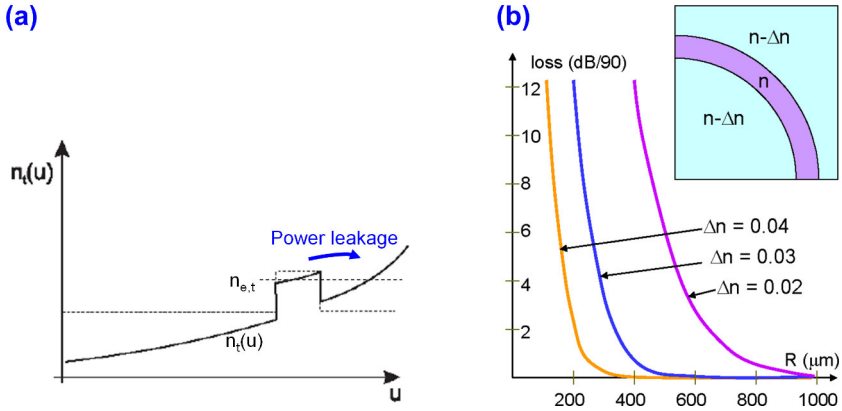

Figure 6.20: Power leakage in bent waveguides; (a) The transformation of the index profile according to Eq. (6.144) corresponds to a local "tilt" of the index profile. Power is lost by „tunneling" of photons from the core region to the region defined by $n_t\,(u) > n_{e,t}$, where $n_{e,t}$ denotes effective index of the mode in the transformed waveguide profile. (b) Decreasing the radius of curvature narrows the "barrier" that confines the photons, and power leakage increases exponentially. (Figure adapted from [4]).

by the outer contour of the waveguide and the tilt of the index profile, and the inner contour will not play a role. Such modes are called whispering gallery modes. The width of the waveguide does no longer play a role in the whispering-gallery regime. If the outer waveguide surfaces are smooth, whispering-gallery modes feature extremely low losses. They can, e.g., be used to realize high-Q resonators, see Fig. 6.21 (c).

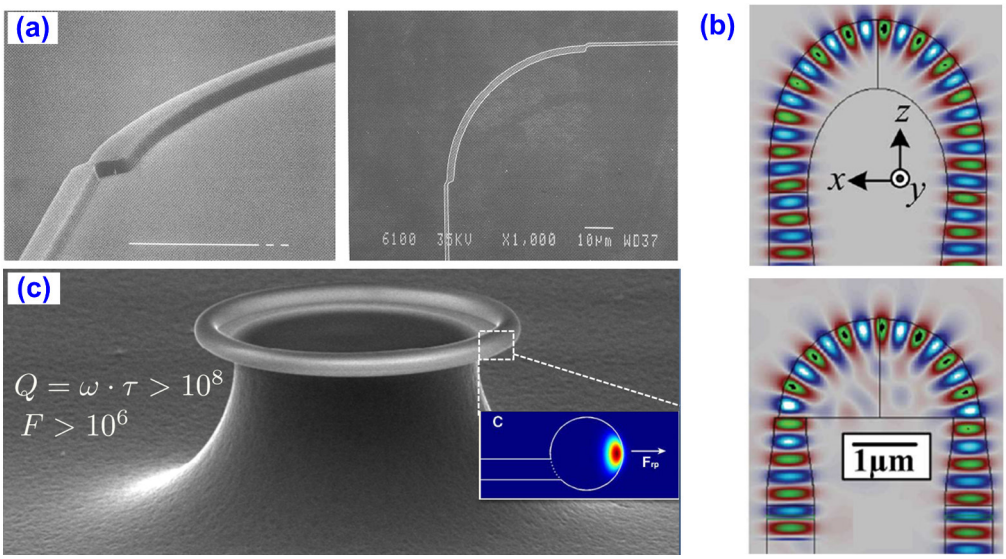

Figure 6.21: Various bends of integrated optical waveguides; (a) Transition losses between the straight waveguide section and the bend can be mitigated by a lateral offset of the waveguides (Figure adapted from [4]) (b) High index-contrast silicon-on-insulator (SOI) waveguides allow for bend radii as small as $1\,\mu$m. Numerically optimized curve designs are the needed to avoid losses at the transition of the straight waveguide to the bend (Figure adapted from [18]); (c) For wide waveguides and strong curvatures, the mode will entirely be guided by the outer contour of the waveguide, leading to so-called whispering gallery modes. Such modes may exhibit very low propagation losses, since interaction with a potentially rough sidewall are reduced to a minimum. Whispering-gallery modes are often used to realize optical resonators with extremely high quality factors (Figure adapted from [7]).

Appendix A

Mathematical Definitions and Conventions

A.1 Time- and Frequency-Domain Quantities

If not otherwise stated, t is the independent time domain variable, and ω denotes the corresponding frequency domain variable. The quantities $u(t)$, $v(t)$, and $h(t)$ are functions in the time domain, and $\widetilde{u}(\omega)$, $\widetilde{v}(\omega)$, and $\widetilde{h}(\omega)$ are the corresponding frequency domain spectra.

A.1.1 Fourier Transformation

The Fourier transform of a function $u(t)$ with respect to the independent variable t is denoted as $\mathfrak{F}_t\{u(t)\}$. Accordingly, $\mathfrak{F}_\omega^{-1}\{\widetilde{u}(\omega)\}$ refers to the inverse Fourier transform of a function $\widetilde{u}(\omega)$ with respect to the independent variable ω,

$$\mathfrak{F}_t\{u(t)\} = \widetilde{u}(\omega) = \int_{-\infty}^{+\infty} u(t)\,\mathrm{e}^{-\mathrm{j}\,\omega t}\,\mathrm{d}t, \tag{A.1}$$

$$\mathfrak{F}_\omega^{-1}\{\widetilde{u}(\omega)\} = u(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widetilde{u}(\omega)\,\mathrm{e}^{\mathrm{j}\,\omega t}\,\mathrm{d}\omega. \tag{A.2}$$

The independent variable t usually represents the time, and ω is angular frequency.

A.2 Vector calculus

A.2.1 The Nabla operator

Gradient

$$\nabla\psi = \operatorname{grad}\psi = \begin{pmatrix}\partial_x\\ \partial_y\\ \partial_z\end{pmatrix}\psi = \begin{pmatrix}\partial_x\psi\\ \partial_y\psi\\ \partial_z\psi\end{pmatrix} \tag{A.3}$$

Example: $\nabla e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}} = -\mathrm{j}\,\mathbf{k}e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}}$

Curl

$$\nabla \times \mathbf{A} = \operatorname{rot} \mathbf{A} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \times \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} \partial_y A_z - \partial_z A_y \\ \partial_z A_x - \partial_x A_z \\ \partial_x A_y - \partial_y A_x \end{pmatrix} \tag{A.4}$$

Example: $\nabla \times \mathbf{A}_0 e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}} = -\mathrm{j}\,\mathbf{k} \times \mathbf{A}_0 e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}}$

Divergence

$$\nabla \cdot \mathbf{A} = \operatorname{div} \mathbf{A} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \cdot \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \partial_x A_x + \partial_y A_y + \partial_z A_z \tag{A.5}$$

Example: $\nabla \cdot \mathbf{A}_0 e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}} = -\mathrm{j}\,\mathbf{k} \cdot \mathbf{A}_0 e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}}$

Laplacian

$$\nabla^2 \mathbf{A} = \triangle \mathbf{A} = \left(\partial_x^2 + \partial_y^2 + \partial_z^2\right) \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} \left(\partial_x^2 + \partial_y^2 + \partial_z^2\right) A_x \\ \left(\partial_x^2 + \partial_y^2 + \partial_z^2\right) A_y \\ \left(\partial_x^2 + \partial_y^2 + \partial_z^2\right) A_z \end{pmatrix} \tag{A.6}$$

Example: $\nabla^2 \mathbf{A}_0 e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}} = (-\mathrm{j})^2 (\mathbf{k} \cdot \mathbf{k}) \mathbf{A}_0 e^{-\mathrm{j}\,\mathbf{k}\cdot\mathbf{r}}$

A.2.2 Basic formulae of vector differential operators

A short summary of basic relations from vector differential calculus is given in Fig. A.1.

Linearität	
1. $\nabla(\alpha\Phi + \beta\Psi) = \alpha\,\nabla\Phi + \beta\,\nabla\Psi$	$\operatorname{grad}(\alpha\Phi + \beta\Psi) = \alpha\,\operatorname{grad}\Phi + \beta\,\operatorname{grad}\Psi$
2. $\nabla\cdot(\alpha\,\mathbf{F} + \beta\,\mathbf{G}) = \alpha\,\nabla\cdot\mathbf{F} + \beta\,\nabla\cdot\mathbf{G}$	$\operatorname{div}(\alpha\,\mathbf{F} + \beta\,\mathbf{G}) = \alpha\,\operatorname{div}\mathbf{F} + \beta\,\operatorname{div}\mathbf{G}$
3. $\nabla\times(\alpha\,\mathbf{F} + \beta\,\mathbf{G}) = \alpha\,\nabla\times\mathbf{F} + \beta\,\nabla\times\mathbf{G}$	$\operatorname{rot}(\alpha\,\mathbf{F} + \beta\,\mathbf{G}) = \alpha\,\operatorname{rot}\mathbf{F} + \beta\,\operatorname{rot}\mathbf{G}$
Operation auf Produkten	
4. $\nabla(\Phi\Psi) = \Phi\,\nabla\Psi + \Psi\,\nabla\Phi$	$\operatorname{grad}(\Phi\Psi) = \Phi\,\operatorname{grad}\Psi + \Psi\,\operatorname{grad}\Phi$
5. $\nabla(\mathbf{F}\cdot\mathbf{G}) = (\mathbf{F}\cdot\nabla)\mathbf{G} + (\mathbf{G}\cdot\nabla)\mathbf{F} + \mathbf{F}\times(\nabla\times\mathbf{G}) + \mathbf{G}\times(\nabla\times\mathbf{F})$	$\operatorname{grad}(\mathbf{F}\cdot\mathbf{G}) = (\mathbf{F}\cdot\operatorname{grad})\mathbf{G} + (\mathbf{G}\cdot\operatorname{grad})\mathbf{F} + \mathbf{F}\times\operatorname{rot}\mathbf{G} + \mathbf{G}\times\operatorname{rot}\mathbf{F}$
6. $\nabla\cdot(\Phi\,\mathbf{F}) = \Phi\,\nabla\cdot\mathbf{F} + (\nabla\Phi)\cdot\mathbf{F}$	$\operatorname{div}(\Phi\,\mathbf{F}) = \Phi\,\operatorname{div}\mathbf{F} + \mathbf{F}\cdot\operatorname{grad}\Phi$
7. $\nabla\cdot(\mathbf{F}\times\mathbf{G}) = \mathbf{G}\cdot\nabla\times\mathbf{F} - \mathbf{F}\cdot\nabla\times\mathbf{G}$	$\operatorname{div}(\mathbf{F}\times\mathbf{G}) = \mathbf{G}\cdot\operatorname{rot}\mathbf{F} - \mathbf{F}\cdot\operatorname{rot}\mathbf{G}$
8. $\nabla\times(\Phi\,\mathbf{F}) = \Phi\,\nabla\times\mathbf{F} + (\nabla\Phi)\times\mathbf{F}$	$\operatorname{rot}(\Phi\,\mathbf{F}) = \Phi\,\operatorname{rot}\mathbf{F} + (\operatorname{grad}\Phi)\times\mathbf{F}$
9. $\nabla\times(\mathbf{F}\times\mathbf{G}) = (\mathbf{G}\cdot\nabla)\mathbf{F} - (\mathbf{F}\cdot\nabla)\mathbf{G} + \mathbf{F}(\nabla\cdot\mathbf{G}) - \mathbf{G}(\nabla\cdot\mathbf{F})$	$\operatorname{rot}(\mathbf{F}\times\mathbf{G}) = (\mathbf{G}\cdot\operatorname{grad})\mathbf{F} - (\mathbf{F}\cdot\operatorname{grad})\mathbf{G} + \mathbf{F}\operatorname{div}\mathbf{G} - \mathbf{G}\operatorname{div}\mathbf{F}$
Zweifache Anwendung von ∇	
10. $\nabla\cdot(\nabla\times\mathbf{F}) = 0$	$\operatorname{div}\operatorname{rot}\mathbf{F} = 0$
11. $\nabla\times(\nabla\Phi) = \mathbf{0}$	$\operatorname{rot}\operatorname{grad}\Phi = \mathbf{0}$
12. $\nabla\times(\nabla\times\mathbf{F}) = \nabla(\nabla\cdot\mathbf{F}) - \nabla^2\mathbf{F}$	$\operatorname{rot}\operatorname{rot}\mathbf{F} = \operatorname{grad}\operatorname{div}\mathbf{F} - \Delta\mathbf{F}$

Figure A.1: Basic relations of vector differential operators (Adapted from [29])

A.3 Cylindrical coordinates and Bessel differential equations

Koordinatentransformation

$x = \rho\cos\varphi,\ y = \rho\sin\varphi,\ z = z$

$\rho = \sqrt{x^2 + y^2},\ \varphi = \tan^{-1}\frac{y}{x}$ (passender Zweig), $z = z$

$h_1 = 1,\ h_2 = \rho,\ h_3 = 1$

Beziehung zwischen Basisvektoren

$$\begin{cases} \mathbf{e}_x = \mathbf{e}_\rho\cos\varphi - \mathbf{e}_\varphi\sin\varphi \\ \mathbf{e}_y = \mathbf{e}_\rho\sin\varphi + \mathbf{e}_\varphi\cos\varphi \\ \mathbf{e}_z = \mathbf{e}_z \end{cases}$$

$$\begin{cases} \mathbf{e}_\rho = \mathbf{e}_x\cos\varphi + \mathbf{e}_y\sin\varphi \\ \mathbf{e}_\varphi = -\mathbf{e}_x\sin\varphi + \mathbf{e}_y\cos\varphi \\ \mathbf{e}_z = \mathbf{e}_z \end{cases}$$

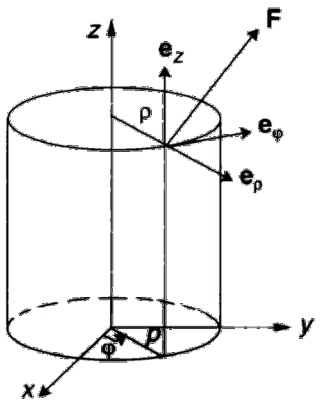

Figure A.2: Basic relations between cartesian coordinates (x, y, z) and cylindrical coordinates (ρ, φ, z). For cartesian coordinates, the quantities $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ denote the unit vector in x-, y-, and z-direction. Similarly, for cylindrical coordinates $\mathbf{e}_\rho$, $\mathbf{e}_\varphi$, and $\mathbf{e}_z$ denote the local unit vector in ρ-, φ-, and z-direction. Note that in the remainder of these lecture notes, r is used in lieu of ρ to denote the radial coordinate. (Adapted from [29])

$$\text{grad}\, u = \nabla u = \frac{\partial u}{\partial \rho}\mathbf{e}_\rho + \frac{1}{\rho}\frac{\partial u}{\partial \varphi}\mathbf{e}_\varphi + \frac{du}{dz}\mathbf{e}_z$$

$$\text{div}\, \mathbf{F} = \nabla\cdot\mathbf{F} = \frac{1}{\rho}\frac{\partial(\rho F_\rho)}{\partial \rho} + \frac{1}{\rho}\frac{\partial F_\varphi}{\partial \varphi} + \frac{\partial F_z}{\partial z}$$

$$\text{rot}\, \mathbf{F} = \nabla\times\mathbf{F} = \left(\frac{1}{\rho}\frac{\partial F_z}{\partial \varphi} - \frac{\partial F_\varphi}{\partial z}\right)\mathbf{e}_\rho + \left(\frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial \rho}\right)\mathbf{e}_\varphi + \\ + \frac{1}{\rho}\left(\frac{\partial(\rho F_\varphi)}{\partial \rho} - \frac{\partial F_\rho}{\partial \varphi}\right)\mathbf{e}_z$$

$$\Delta u = \nabla^2 u = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\frac{\partial u}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\partial^2 u}{\partial z^2} = u_{\rho\rho} + \frac{1}{\rho}u_\rho + \frac{1}{\rho^2}u_{\varphi\varphi} + u_{zz}$$

Figure A.3: Differential operators in cylindrical coordinates. Note that in the remainder of these lecture notes, r is used in lieu of ρ to denote the radial coordinate. (Adapted from [29])

1. $x^2y''+xy'+(a^2x^2-p^2)y=0 \Leftrightarrow y''+\frac{y'}{x}+\left(a^2-\frac{p^2}{x^2}\right)y=0 \Leftrightarrow$

$\Leftrightarrow \frac{1}{x}(xy')'+\left(a^2-\frac{p^2}{x^2}\right)y=0$

Lösung: $y=AJ_p(ax)+BY_p(ax)$ [Beachte. $|Y_p(x)|\to\infty$ mit $x\to 0+$]

2. $x^2y''+xy'-(a^2x^2+n^2)y=0 \Leftrightarrow y''+\frac{y'}{x}-\left(a^2+\frac{n^2}{x^2}\right)y=0 \Leftrightarrow$

$\Leftrightarrow \frac{1}{x}(xy')'-\left(a^2+\frac{n^2}{x^2}\right)y=0$

Lösung: $y=AI_n(ax)+BK_n(ax)$

Figure A.4: Bessel differential equations and the corresponding solutions. The symbol y represents a function of x, and the prime indicates a derivative with respect to the arguments, $y'(x)=\frac{\partial y}{\partial x}$. J_ν and Y_ν denote ν-th order Bessel functions of the first and second kind, respectively. Likewise, I_ν and K_ν denote ν-th order modified Bessel functions. (Adapted from [29])

Bibliography

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions.* Dover publications, 1972.

[2] Govind P. Agrawal. *Fiber-optic communcations systems.* John Wiley & Sons, 2002.

[3] M. Bachmann, P. A. Besse, and H. Melchior. General self-imaging properties in n × n multimode interference couplers including phase relations. *Appl. Opt.*, 33(18):3905–3911, Jun 1994.

[4] R. Baets. 'dielectric waveguides'. Lecture Notes "Microphotonics", Ghent University, 2009. http://photonics.intec.ugent.be.

[5] Chin-Lin Chen. *Foundations for guided-wave optics.* Wiley, Hoboken, New Jersey, 2007.

[6] D. Chen, H. R. Fetterman, A. Chen, W. H. Steier, L. R. Dalton, W. Wang, and Y. Shi. Demonstration of 110 GHz electro-optical polymer modulators. *Appl. Phys. Lett.*, 70(25):3335–3337, June 1997.

[7] P. Del'Haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, and T. J. Kippenberg. Optical frequency comb generation from a monolithic microresonator. *Nature*, 450(7173):1214–1217, December 2007.

[8] J. A. Dionne, L. A. Sweatlock, H. A. Atwater, and A. Polman. Planar metal plasmon waveguides: frequency-dependent dispersion, propagation, localization, and loss beyond the free electron model. *Phys. Rev. B*, 72:075405, Aug 2005.

[9] W. Freude. *Field Propagation and Coherence.* Karlsruhe Instiute of Technology (KIT), Institute of Photonics and Quantum Electronics (IPQ), 2010. Lecture Notes.

[10] Jeff Hecht. *Understanding Fiber Optics.* Prentice Hall International, 5 edition, 2005.

[11] K. Iizuka. *Elements of Photonics*, volume 2. John Wiley & Sons, 1 edition, 2002.

[12] RSoft Design Group Inc. *FemSIM 3.3 User Guide.* http://www.rsoftdesign.com, 2011.

[13] Thomas L. Koch Ivan P. Kaminov. *Optical Fiber Telecommunications III*, volume 2. Academic Press, 1997.

[14] J. D. Jackson. *Classical Electrodynamics.* John Wiley and Sons, New York, 1998.

[15] J. D. Joannopoulos, R. D. Maede, and J. N. Winn. *Photonic Crystals.* Princeton University Press, 1995.

[16] Kenji Kawano and Tsutomu Kitoh. *Optical Waveguide Analysis.* Wiley, 2001.

[17] Gerd Keiser. *Optical fiber communcations.* McGraw-Hill, 2nd edition, 1991.

[18] C. Koos, C. G. Poulton, L. Zimmermann, L. Jacome, J. Leuthold, and W. Freude. Ideal bend contour trajectories for single-mode operation of low-loss overmoded waveguides. *Photonics Technology Letters, IEEE*, 19(11):819 –821, June 1, 2007.

[19] Arun Kumar, K. Thyagarajan, and A. K. Ghatak. Analysis of rectangular-core dielectric waveguides: an accurate perturbation approach. *Opt. Lett.*, 8(1):63–65, Jan 1983.

[20] J. Y. Y. Leong, P. Petropoulos, J. H. V. Price, Heike Ebendorff-Heidepriem, S. Asimakis, R. C. Moore, K. E. Frampton, V. Finazzi, X. Feng, T. M. Monro, and D. J. Richardson. High-nonlinearity dispersion-shifted lead-silicate holey fibers for efficient 1-µm pumped supercontinuum generation. *J. Lightwave Technol.*, 24(1):183, Jan 2006.

[21] E. A. Marcatili. Dielectric rectangular waveguide and directional coupler for integrated optics. *Bell Syste. Tech. J.*, 48:2071–2102, 1969.

[22] D. Marcuse. *Light Transmission Optics*. Van Nostrand Reinhold, New York, 1972.

[23] D. Marcuse. *Theory of Dielectric Optical Waveguides*. Academic Press, London, 1974.

[24] K. Okamoto. *Fundamentals of Optical Waveguides*. Academic Press, San Diego, 2000.

[25] Katsunari Okamoto. *Fundamentals of Optical Waveguides*. Academic Press, 2006.

[26] E. D. Palik. *Handbook of Optical Constants of Solids*. Academic Press, San Diego, CA, 1998.

[27] C. G. Poulton, C. Koos, M. Fujii, A. Pfrang, Th. Schimmel, J. Leuthold, and W. Freude. Radiation modes and roughness loss in high index-contrast waveguides. *IEEE J. Sel. Top. Quantum Electron.*, 16(6), November/December 2006.

[28] Dayou Qian, Ming-Fang Huang, Ezra Ip, Yue-Kai Huang, Yin Shao, Junqiang Hu, and Ting Wang. 101.7-tb/s (370x294-gb/s) pdm-128qam-ofdm transmission over 3x55-km ssmf using pilot-based phase noise mitigation. In *Optical Fiber Communication Conference*, page PDPB5. Optical Society of America, 2011.

[29] L. Rade and B. Westergren. *Mathematische Formeln*. Springer, 1997.

[30] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*. Wiley, 2007.

[31] H.-G. Unger. *Planar optical waveguides and fibres*. Clarendon Press, Oxford, 1977.

[32] E. Voges and K. Petermann. *Optische Kommunikationstechnik*. Springer, 2002.

[33] H. Yokota, H. Kanamori, Y. Ishiguro, G. Tanaka, S. Tanaka, H. Takada, M. Watanabe, S. Suzuki, K. Yano, M. Hoshikawa, and H. Shimba. Ultra-low-loss pure-silica-core single-mode fiber and transmission experiment. In *Optical Fiber Communication*, page PD3. Optical Society of America, 1986.