Technical University of Denmark

DTU

Optical methods for characterization of surface structures on a nanometer scale

Gregersen, Niels; Hanson, Steen Grüner; Mørk, Jesper; Tromborg, Bjarne

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

Optical methods for characterization of surface structures on a nanometer scale.

Ph.D. thesis

Niels Gregersen
October 2006

Department of Communications, Optics & Materials
Nano•DTU
Technical University of Denmark

Optical methods for characterization of surface structures on a nanometer scale.

Department of Communications, Optics & Materials

Nano•DTU, Technical University of Denmark

Supervisors:

Bjarne Tromborg[i]	Department of Communications, Optics & Materials
Jesper Mørk[ii]	Department of Communications, Optics & Materials
Jørgen Garnæs	Danish Institute for Fundamental Metrology
Steen G. Hanson	Risø National Laboratory

[i] From November 1st, 2003 to June 30th, 2006. Now retired.
[ii] From July 1st, 2006.

Niels Gregersen   Kgs. Lyngby, October 31st, 2006

_____

Abstract

The topic of this Ph.D. thesis is optical characterization of surface structures on a nanometer scale. When studying a sample with subwavelength features, the diffraction limit sets a lower bound to the resolution achievable using far-field microscopy. In this work the possiblity of circumventing the diffraction limit by employing a scanning near-field optical microscope (SNOM) to perform the characterization is investigated.

Experimental SNOM images of the optical field distribution above a deep grating are analyzed with the purpose of identifying the grating topography. The inverse scattering problem is treated using an inversion procedure based on an optimization routine. The advantages and limitations of this method are examined.

The core of this work deals with the relation between the free-space optical field distribution in the vicinity of a nano-object and the experimentally measured SNOM image. Transfer functions (TFs) describing the coupling of the free-space field to the guided mode of the SNOM fiber are determined using numerical simulations. The TFs of uncoated tapered SNOM fiber tip of varying opening angle are analyzed and information about the relation between the free-space optical field and the SNOM image is obtained. The optical resolution of uncoated SNOM fiber tips is quantitized, and the effective plane of detection approximation, which is commonly used to interpret a SNOM image, is examined. The experimental SNOM measurement from the topography characterization is simulated, where the field-image relation is taken into account using the computed TFs, and the influence of imperfections in the tip geometry on the correspondence between modeling and experiment is analyzed.

The numerical simulations are performed using the eigenmode expansion technique, and its performance in computing the TFs for both uncoated and metal-coated fiber tips is examined.

Finally, a short independent study of the possibility of calculating the local density of states in a micropillar is presented.

Resumé

Emnet for denne Ph.D. afhandling er optisk karakterisering af overflade strukturer på nanometer skala. Ved et studie af et objekt med sub-bølgelængde karakteristika sætter diffraktionsgrænsen en nedre grænse for den opløsning, der kan opnås med fjernfelts-mikroskopi. I dette værk undersøges muligheden for at omgå diffraktionsgrænsen ved at benytte et skannende nærfelt optisk mikroskop (SNOM) til at gennemføre karakteriseringen.

Eksperimentelle SNOM billeder af den optiske feltfordeling over et dybt gitter bliver analyserede med henblik på at identificere gitterets topografi. Det inverse spredningsproblem bliver behandlet vha. en inversionsprocedure baseret på en optimeringsrutine. Fordelene og ulemperne ved denne metode bliver belyst.

Hovedparten af værket omhandler relationen mellem det optiske felt i frit rum i nærheden af et nano-objekt og det eksperimentelt målte SNOM billede. Overførselsfunktioner (OFer) beskrivende koblingen af fritrumsfeltet til den bundne mode i en SNOM fiber bliver bestemt vha. numerisk simulation. OF'erne for ubelagte graduerede SNOM fibertippe med varierende åbningsvinkel analyseres og information omkring relationen mellem det optiske fritrumsfelt og SNOM billedet opnås. Den optiske opløsning for ubelagte SNOM fibertippe bliver kvantiseret, og den effektive detektionsplan-approksimation som almindeligvis benyttes til at fortolke et SNOM billede bliver undersøgt. Den eksperimentelle SNOM måling fra topografi karakteriseringen bliver simuleret, hvor relationen mellem felt og billede bliver taget i betragtning vha. de beregnede OF'ere, og indflydelsen af imperfektioner i tip geometrien på overensstemmelsen mellem modelering og eksperiment bliver analyseret.

De numeriske simulationer bliver gennemført vha. eigenmode ekspansions-teknikken, og dens evne til at beregne OF'ere for både ubelagte og metalbelagte fibertippe bliver undersøgt.

Til sidst bliver et kort uafhængigt studie af muligheden for at udregne lokal tilstandstæthed i en mikrosøjle præsenteret.

Preface

This Ph.D. project has taken place in the framework of the Center for Micro-Optical Structures (CEMOST). The CEMOST was a consortium formed by six Danish companies, three institutes from the Danish advanced technology group (GTS) and two universities, and the objective of CEMOST was to meet the increasing demand for techniques allowing the characterization of structures having characteristic length scales in the nanometer scale.

The possibilities for performing optical characterization on a subwavelength scale were explored by the universities. The experimental work were performed at Aalborg University, and I was hired as a Ph.D. student at the Department of Communications, Optics & Materials (COM) at the Technical University of Denmark to investigate the theoretical and numerical issues.

As it is often the case in research, the project did not quite follow the original plan. It was initially expected that the main effort would go into topography reconstruction, but during the project it became clear that, before high-quality topography reconstructions could be performed, another issue needed addressing: The relation between the images obtained using near-field microscopy and the actual optical field distribution had to be investigated, and the understanding of this relation is of fundamental interest in near-field optics (NFO). For this reason the core of the Ph.D. thesis is devoted to the image-field relation.

Even though a close collaboration with the CEMOST partners, in particular Aalborg University, was maintained, the project was somewhat isolated from the core activities at COM. At a late stage, however, a collaboration with COM collegues was formed with the objective of calculating local density of states (LDOS) in a micropillar. This activity has nothing to do with optical characterization, but was initiated as we realized that the numerical modeling tools used in my Ph.D. work were capable of performing this type of computation. For this reason, a short independent chapter is included about micropillars.

Even though the project was somewhat isolated, I received help and support from numerous people. I would like to express my gratitude to my supervisor Jørgen Garnæs for his support and for selecting me for the position. I am grateful to my external supervisor Steen G. Hanson, who always claimed not to understand what I

v

List of acronyms

a-SNOM	Apertureless scanning near-field optical microscope
AD	Average deviation
AFM	Atomic force microscope
CEMOST	Center for Micro-Optical Structures
COM	Department of Communications, Optics & Materials
DBR	Distributed Bragg reflector
EET	Eigenmode expansion technique
EPD	Effective plane of detection
FD	Finite difference
FDTD	Finite-difference time-domain
FEM	Finite elements method
GTS	Godkendt teknologisk service
ISP	Inverse scattering problem
ITF	Intensity transfer function
LDOS	Local density of states
MMP	Multiple multipole
NFO	Near-field optics
NSOM	Near-field scanning optical microscope
QD	Quantom dot
PML	Perfectly matched layer
PSTM	Photon scanning tunneling microscope
PW	Plane-wave
RIE	Reactive ion etching
SA	Semi-analytical
SEM	Scanning electron microscope
SNOM	Scanning near-field optical microscope
STOM	Scanning tunnelling optical microscope
TE	Transverse electric
TF	Transfer function
TM	Transverse magnetic
VCSEL	Vertical-cavity surface-emitting laser

Table of contents

1 Introduction

During the past decades, nano-technology has received ever increasing attention. Even though the length scale involved in order for a physical system to classify as nano-technology is not well-defined, the tendency towards reducing the size of existing components is clear. This is particularly true in the electronics industry, as example the demand for increasing computer power has lead to a downscaling of gate lengths in Pentium CPUs to well below 100 nm[1]. The interest in nano-technology, however, is not only due to the need to reduce sizes: Within the field of quantum information, the length scale of a system of qubits may be tens of nanometers or less in order to achieve the quantum effects of interest. In the field of optics, efforts are made to move data processing from the electrical into the optical domain. Examples of devices with this functionality are Silicon-On-Insulator based Photonic-Crystal components. The length scale of a design may be on the order of 100 nm or more, however, fabrication tolerances of 10 nm or less are required to keep scattering losses at an acceptable level. In nano-technology, the design of a device with new functionalities requires, as in all fields, the ability to perform a topography characterization to verify that the device adheres to production specifications. There is thus an increasing demand for characterization of devices and structures with lengths on the sub-micron scale.

The objective of this Ph.D. thesis is to develop methodology for characterization of surface structures on a nanometer scale using optical microscopy. Today, no technology exists that allows optical imaging of geometric features of length scales shorter than approximately the wavelength of light. Subwavelength information is available in the optical field scattered by the structure, but it is contained in the evanescent part of the field, which is localized to the immediate surroundings of the sample and cannot be detected using conventional microscopy. Measurement of the optical field at nanometer distances from the surface, the so-called near field, has been possible since the invention of the scanning near-field optical microscope (SNOM) twenty years ago. However, the near field does generally not directly reflect the sample topography, and inversion routines must be employed to extract topographical information from the optical field distribution. Much research activity[2-4] has gone into the subject of topography reconstruction in the microwave regime, and the subject has also been treated in near-field optics. The topography reconstruction

techniques available today, however, are limited to geometries for which certain restrictive assumptions can be made, and a well-established method which can be used to identify a completely general structure from knowledge about the optical field does not yet exist. In this Ph.D. thesis a recently suggested method allowing topography characterization of a completely general sample has ben tested. To my knowledge, this is the first time near-field imaging has been used to identify the topography of structure, here a deep grating, that violates the assumptions of existing perturbative methods.

Even though detection of an evanescent field today is possible and subwavelength optical resolution can be achieved, it remains unclear what the optical image obtained using the near-field microscope actually represents. The image is related to the optical field distribution near the sample under study, but the exact nature of the relation is currently not well understood. The theoretical framework necessary to describe this relation has been around for some time, but, to my knowledge, the calculation of the transfer function (TF) necessary to correctly relate a near-field image to the optical field distribution has not previously been performed. In this work, a technique to numerically determine the TF is presented, and the method is used to compute TFs of uncoated fiber tips. An in-depth analysis of the TFs is made revealing new information about the relation between the free-space optical field and the optical SNOM images.

The final activity, unrelated to the main objective of the Ph.D. thesis, is the determination of the local density of states (LDOS) in a micropillar. In quantum optics, the LDOS is an important quantity as it describes the coupling strength of atomic transitions to the electromagnetic field. The numerical determination of the LDOS, however, is very numerically demanding. It has been performed in photonic crystals, but I am not aware of any calculations in the literature of the LDOS in a micropillar. In this thesis, it is described how the LDOS in a micropillar can be computed, and a preliminary calculation is presented.

The numerical method used in this work is the eigenmode expansion technique (EET). The EET is well-known, but it is only in recent years, that it has been put to use to simulate the optical field in large geometries. In this work, its performance in handling a geometry including a metal layer is evaluated, and, to my knowledge, it is the first time the technique is used to simulate the scattering of light on a high-loss material.

This Ph.D. thesis is organized as follows: An introduction to near-field microscopy is given in chapter 2. The basics of optical imaging are discussed, and it is explained how the diffraction limit sets a lower boundary to the optical resolution achievable using conventional microscopy. It is shown that topographical information about the structure on a subwavelength scale is available in the near-field, and that this can be detected using the SNOM. The main classes of SNOM configurations are presented and the challenges in optical characterization using near-field microscopy are summarized.

The subject of chapter 3 is the numerical simulation of the scattering of light on microscopic objects. An overview of the modeling techniques, which are most widely used in electromagnetics, is given, and the choice of the EET is explained. The details of this method, including three different schemes for determining eigenmodes, are described in detail.

In chapter 4, the topic of topography characterization using optical imaging is presented. A method for identifying the topography of a diffraction grating from experimental SNOM measurements is described, and the technique is used to determine the sample geometry. The strengths and weaknesses of the method are discussed.

The relation between the free-space optical field and experimental image obtained with the SNOM is analyzed in chapter 5. The theoretical framework necessary to describe this relation is presented, and the concept of transfer function (TF) is introduced. The ability of the EET to compute TFs of uncoated fiber tips is demonstrated in a convergence study, and the challenges related to the handling of complex refractive index profiles are examined. TFs of uncoated fiber tips are presented, and an in-depth analysis of the TFs is performed with particular focus on optical resolution and on the effective plane of detection approximation, which is widely used in the interpretation of SNOM images throughout the near-field community. A simulation of the experimental SNOM measurements from chapter 4 is made, and the influence of imperfections in the fiber tip geometry on the field-image relation is analyzed.

The possibility of calculating the local density of states (LDOS) in a micropillar is addressed in chapter 6. The concept is defined, and the EET is used to determine the LDOS in a test structure.

Finally, the work is summarized and conclusions are given in chapter 7.

2 Near-field microscopy

Optical characterization is an easy, reliable, cheap and non-destructive way to obtain geometric information about an object under study. Unfortunately, as the diffraction limit sets a fundamental lower boundary of approximately the wavelength of the illuminating light to the achievable resolution, we cannot visualise objects of length scales below ~ 1 μm using conventional far-field microscopy.

Alternative techniques, such as atomic force microscopy (AFM) and scanning electron microscopy (SEM), exist that allow characterization of objects on a nanometer scale, but these do not feature all the advantages inherent of regular optical characterization. It is therefore of interest to examine whether optical microscopy can be pushed into the sub-wavelength nanometer region without compromising the many advantages of regular optical characterization.

2.1 The diffraction limit

The most basic imaging system, illustrated in Fig. 2.1, consists of a light source, an object, a lens and an image. The light source illuminates the object, and light is reflected in all directions from the object.
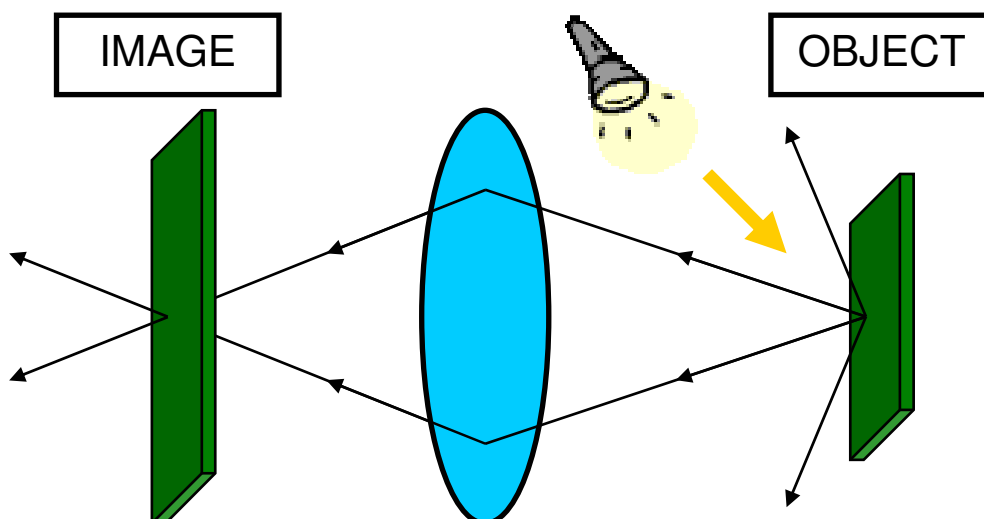
Fig. 2.1: Sketch of the basic imaging system.

The lens system refocuses the light resulting in a depiction of the object in the image plane. The size of the image may be enlarged (or reduced) compared to the object, depending on the geometry of the lens system.

For large objects, with length scales above ~ 1 mm, no sophisticated technology is needed; the human eye acts as the lens. When structures become smaller, we turn to conventional optical microscopy, and with a few lenses objects of length scales down to ~ 1 μm can effectively be visualized.

These imaging systems can be described entirely using geometrical optics. Here the wavelength of light is set to zero, an approximation which is very good in optical systems, where all characteristic lengths are large compared to the wavelength. The diffraction limit, however, sets in when the characteristic lengths are comparable to the wavelength, and thus it cannot be explained using geometrical optics. Rather, an exact wave theory is necessary.

The origin of the diffraction limit can be understood by examining the scattering of laser light of wavelength λ on the diffraction grating. For simplicity we consider a 2D geometry with uniformity along the y axis and study the propagation of the light polarized along this axis in absence of free charges and currents. In this case, the component E_y of the electrical field strength $\mathbf{E}$ is described by the Helmholtz equation valid in the frequency domain:

$$\nabla^2 E_y(x,z) + \varepsilon_r k_0^2 E_y(x,z) = 0 . \tag{2.1}$$

Here, $k_0 \equiv \omega/c \equiv 2\pi/\lambda$, ε_r is the dielectric constant, ω is the angular frequency, c is the speed of light, and the time dependence is given by $\exp(-i\omega t)$. The general free-space solution to the Helmholtz equation is then of the form

$$E_y(x,z) = C \exp(ik_x x + ik_z z) , \tag{2.2}$$

where the coefficients of the wave vector $\mathbf{k}$ are related by $k_x^2 + k_z^2 = k_0^2$. We can now identify two distinct classes of solutions depending on the value of k_x: When $|k_x| < k_0$, k_z is real-valued and the wave is propagating along the z axis. On the other hand, when $|k_x| > k_0$, k_z takes an imaginary value, the z dependence is of the form $\exp(-|k_z|z)$, and the wave is decaying along the z axis. These two classes of solutions belong to the propagating (or homogenous) regime and the evanescent (or inhomogeneous) regime, respectively.

Now, the diffraction condition describing the scattering of an incoming wave of k vector $\mathbf{k}_i$ to an outgoing wave of k vector $\mathbf{k}_d$ on a grating is given by

$$k_{x,d} = k_{x,i} \pm n2\pi / P . \qquad (2.3)$$

In this equation n is the diffraction order, and P is the grating period.
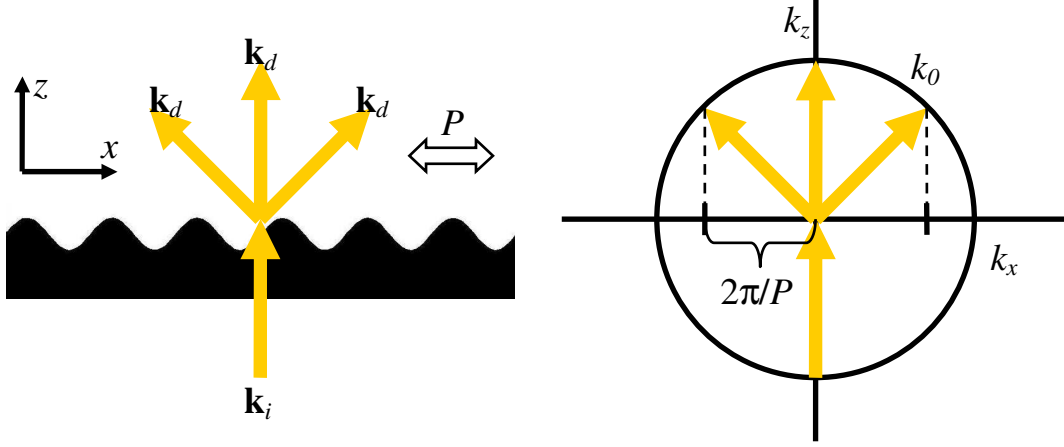
Fig. 2.2: Illustration of the diffraction of light on a grating of period P.
P is longer than the wavelength.

The scattering of an incoming wave with normal incidence is sketched in Fig. 2.2 when P is longer than the wavelength. The x components of the diffracted wave vectors are given by Eq. (2.3), and since $2\pi / P < k_0$ the orders -1 and 1 have $|k_x| < k_0$. The diffracted waves of first order will thus be propagating along the z axis.
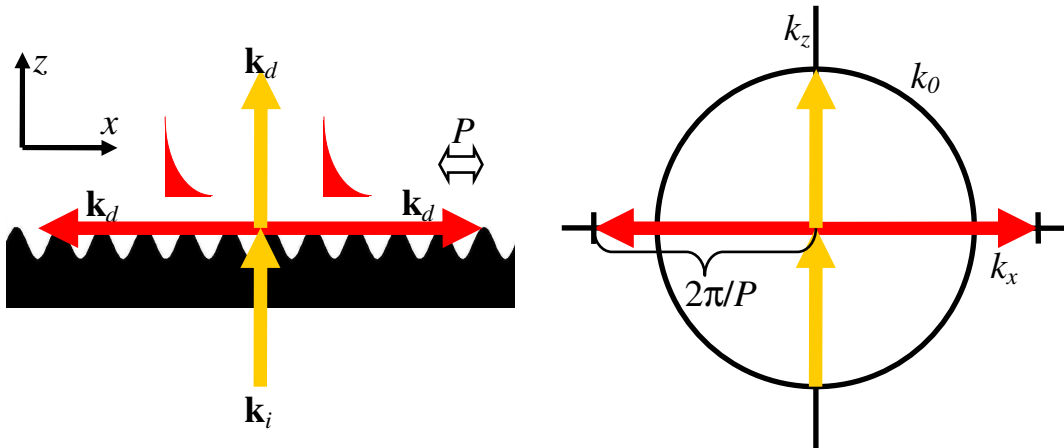
Fig. 2.3: Illustration of the diffraction of light on a grating of period P.
Here, P is shorter than the wavelength.

7

In Fig. 2.3, the situation, when P is shorter than the wavelength, is illustrated. This time $2\pi / P > k_0$, the first order diffracted waves have $|k_x| > k_0$, and these are thus evanescent along the z axis.

The diffraction limit can now be understood by comparing the optical fields detected by an observer positioned far from the grating. When $P > \lambda$, the diffracted waves are propagating, and they will be detected by the observer even though, he is positioned many wavelengths from the scatterer. The waves carry the necessary information about the grating geometry for the observer to perform a characterization. However, when $P < \lambda$, the diffracted waves are all evanescent, and they do not reach the remote observer. The optical field close to the scatterer is perturbed, but the field detected by the observer is identical to that in absence of corrugation, and the observer thus cannot detect the sub-wavelength period grating.

We conclude that information about the topography of the sub-wavelength period grating is available in the optical field, but it decays exponentially and can thus only be detected in close vicinity of the scatterer.

2.2 Near-field optics

The branch of physics that focuses on localized perturbations of the optical field, that can only be detected in close proximity of the object under study, is near-field optics (NFO). The near field is the optical field in close proximity of the object. In the near field, the evanescent waves have not yet decayed below the noise floor, whereas in the far field only propagating waves can be detected.

Historically, the pursuit of subwavelength optical resolution or "super-resolution" was the driving force behind NFO becoming a major field in optics. The first time, a principle of improving the resolution of an optical system beyond the diffraction limit was suggested, was by Synge[5] in 1928. His basic idea consists of discarding the traditional microscope objective and instead illuminating the sample under study with a spot of subwavelength diameter. The setup is sketched in Fig. 2.4. The spot is obtained by illuminating a thin metal film with an aperture in it using a strong light source. The aperture couples the propagating waves of the illuminating light to both homogenous and inhomogeneous waves with the evanescent waves quickly decaying away from the aperture.
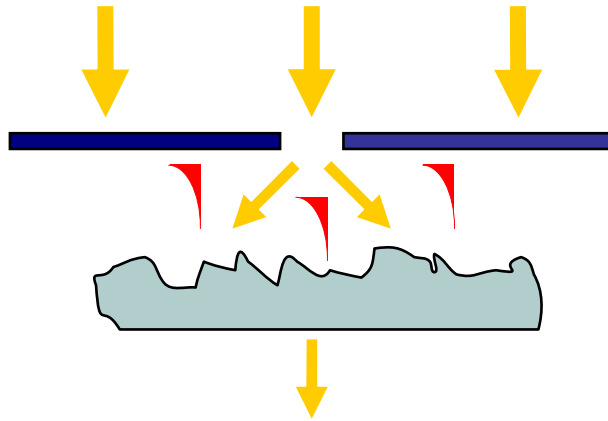
Fig. 2.4: Original design to achieve super-resolution by Synge.

However, if the sample is placed in close vicinity of the film, within the near-field generated by the hole, the sample scatters the evanescent components and coupled them into propagating waves, that can be observed in the far field. The optical resolution of the setup then only depends on the aperture diameter, and, by decreasing the diameter below the wavelength, super-resolution is achieved.

It should be emphasized that, the important physical mechanism allowing subwavelength optical resolution is the interaction of evanescent waves between a sample and a nano-probe of sub-wavelength size placed in close proximity of each other: Super-resolution is only achieved, when a propagating wave is scattered on a first object and generates an evanescent field, which reaches the second object and then couples to a propagating field that can be detected in the far field. A few years after his original publication, Synge, for practical reasons, actually proposed another scheme, where the spot in the aperture in the metal film was replaced by a far-field light source, but this setup would never have resulted in super-resolution as no scatterer is present in close proximity of the sample under study.

Unfortunately, when Synge presented his ideas, the technology necessary to build the experimental setups was not available, and his work was until recently largely ignored.

2.3 Scanning Near-Field Optical Microscopy

The first near-field microscope built by Ash and Nicholls[6] in 1972 operated in the microwave regime. Realizations of super-resolution at optical frequencies, however,

were performed for the first time simultaneously and independently in 1984 by Pohl[7] et al. in Switzerland and by Lewis[8] et al. in the United States. Pohl described his setup as an "optical stethoscope", but today the names scanning near-field optical microscope (SNOM) or near-field scanning optical microscope (NSOM) are generally used for optical microscopes capable of achieving subwavelength resolution.

SNOM microscopes exist in varying configurations, but they all feature a nano-probe, which may act as an emitter, a collector or simply a scatterer, scanning in close proximity above the sample under study. Two main classes of configurations are the apertureless SNOM and the transmission SNOM:

2.3.1 Apertureless SNOM

The operating principle of the apertureless SNOM[9], or a-SNOM, is sketched in Fig. 2.5(a). As the focus of this Ph.D. thesis is topography characterization on a nanometer scale, the near-field in this figure and in the following is illustrated as being generated by illumination of a sample with sub-wavelength features. However, it should be stressed that in the a-SNOM and the collection-mode SNOM setups (see next paragraph) the near-field could equally well be e.g. the evanescent tail of a surface plasmon polariton[10] or of the guided mode in a photonic crystal waveguide[11].

(a)   (b)
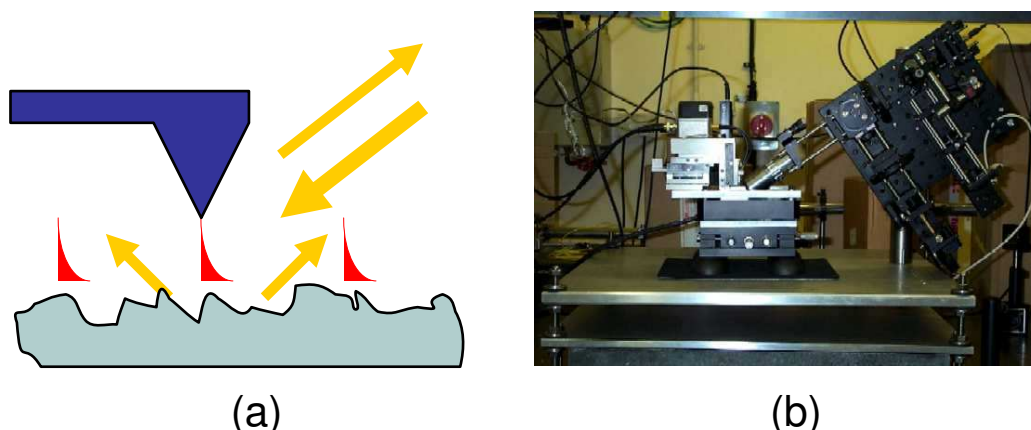
Fig. 2.5: The a-SNOM is sketched in (a). An experimental setup from King's College, London is depicted in (b).

As the name implies, the a-SNOM does not feature the small aperture of the original design by Synge, but rather a sharp metallic tip, usually a standard AFM tip, acting as the nano-probe. The evanescent components of the near-field are scattered by the

metallic tip and are coupled to propagating waves, that can be detected in the far field. A photo of an experimental a-SNOM setup is shown in Fig. 2.5(b).

An advantage of the a-SNOM setup is the possibility of using very sharp metal tips and thus potentially reaching a very high resolution. However, when the detection takes place in the far field, the signal measured may contain a background field,[12] which is not related to the near field under study, even when separate external light sources are completely suppressed. A possible source[12] of this background field is e.g. the propagating waves generated by the scattering of illuminating light on the metal tip, which do not interact with the sample and thus contain no near field information.

2.3.2 Transmission SNOM

This class of microscopes[9] features a tapered tip of an optical single-mode fiber acting as the nano-probe. The transmission SNOM may be operating either in illumination-mode or in collection-mode as sketched in Fig. 2.6.
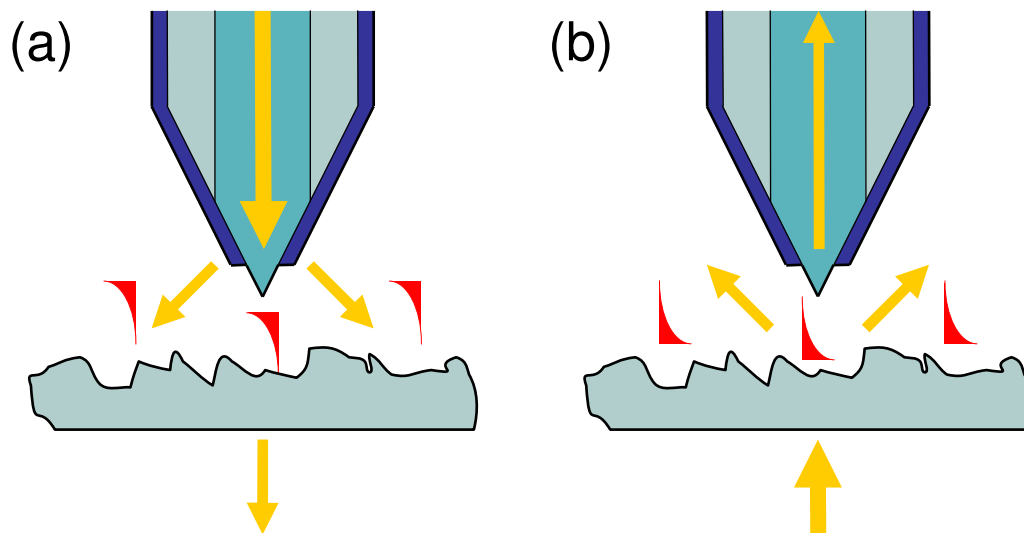
Fig. 2.6: Illustrations of transmission SNOM microscopes in illumination-mode (a) and collection-mode (b).

The illumination-mode SNOM setup was the first to appear[7]. Here the fiber tip is used as a nano-source of radiation. The tip is usually coated with a metal except at the very apex, where light can pass through a subwavelength sized hole, effectively reproducing the tiny illuminating spot from Synge's original idea. The propagating

guided mode of the fiber is scattered by the fiber tip and couples into evanescent waves. These interact with the sample under study and are scattered into propagating waves, that can be detected in the far-field.

The transmission SNOM in collection-mode works in a similar fashion, except the other way around. The near field is generated by illumination of the sample from the far field. The near field is scattered by the fiber tip, acting as a nano-collector, and is coupled to the guided fiber mode. The signal transmitted by the fiber is then detected remotely. The experimental collection SNOM setup used at Aalborg University, Denmark, to image the propagation of guided modes in photonic crystal waveguides is illustrated in Fig. 2.7. Today, new heterodyne experimental setups also exist,[13] that allow detection of not only the amplitude, but also the phase of the guided mode.
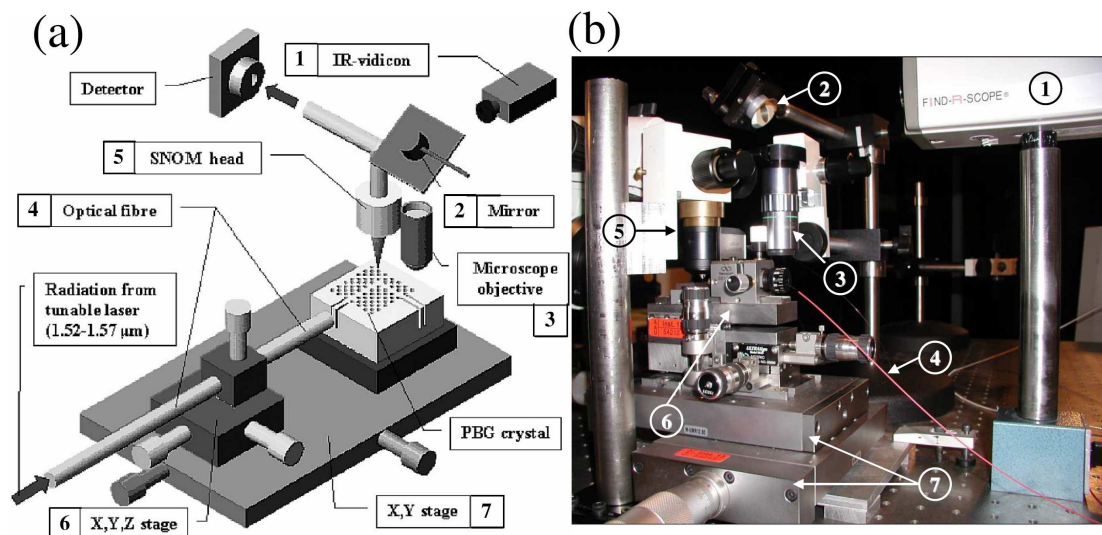
Fig. 2.7: **The collection-mode SNOM at Aalborg University is sketched in (a) and depicted in (b). Sketch and photo by Valentyn S. Volkov.**

Although the illumination-mode setup has been popular for historic reasons, the operating principle of near-field imaging of the collection-mode SNOM is perhaps more straight-forward: In illumination-mode, the detected signal consists of propagating waves generated by the scattering of the near field on the sample. The relations between the near field, the sample properties and the output signal are all highly nontrivial, whereas in the collection-mode, the signal measured can be almost[a] directly related to the local optical field intensity at the position of the fiber tip.

---

[a] The relation between the SNOM image in collection-mode and the free-space optical field intensity distribution is actually more complex as will be explained in chapter 5.

An important sub-class of the transmission setups that deserves attention is the photon scanning tunnelling microscope (PSTM) or scanning tunnelling optical microscope (STOM). The PSTM is similar to the collection-mode SNOM, except for the illuminating light being an evanescent field generated by illuminating a prism above the angle of total internal reflection. In the standard collection-mode setup, the SNOM image is easily dominated by the propagating illuminating field in particular when studying weakly scattering structures, but in the PSTM, the illuminating field is evanescent and does not spoil the near field of interest.

As the experimental SNOM images analyzed in this thesis are obtained from the collection SNOM at Aalborg University operating in the PSTM configuration, we will in the following focus on the properties of the collection SNOM.

2.4 Subwavelength optical imaging

Even though subwavelength components of the optical field can be detected using modern SNOM microscopy, several challenges remain to be overcome before optical imaging of structures with subwavelength resolution can be achieved.

First of all, even in the straight-forward collection SNOM configuration, the metal-coated fiber tip does not act as a lens in a geometrical optics sense. The imaging system of Fig. 2.1 focuses the light emitted from the sample under study producing an image directly related to the original object. But the collection SNOM simply detects the optical field distribution in a plane above the scatterer, a field distribution which, in general, does not directly reflect the topography of the object. For this reason, it is necessary to distinguish between *optical resolution*, which describes how well the optical field in a detection plane can be resolved, and *image resolution*, that is a measure of the size of the details in the image topography. In conventional far-field microscopy, there is no distinction between these two terms, but in NFO, the measured quantity is the optical field, which is only indirectly related to the object, and the optical and the image resolution become separated. The relation between the measured optical field and the object topography is further described in chapter 4.

Another issue is the influence of the nano-probe on the optical field. The close proximity of the probe assures that evanescent components of the optical field are detected. However, the highly conducting metal introduces a perturbation of the free-

space field distribution that would exist in absence of probe, a perturbation that will appear in the SNOM image and cannot immediately be separated from the near-field of interest. For this reason, the nano-probe in a-SNOM and collection SNOM cannot, in general, simply be considered a passive detector of the optical field, and when performing calculations on the optical field distribution in all configurations the nano-probe and the scatterer must, in general, be treated as one global system.

Finally, even if the interaction between the nano-probe and the sample is ignored, the relation between the measured signal and the free-space optical field at the position of the nano-collector is not trivial. In collection SNOM, the SNOM image is clearly related to the optical field intensity in the vicinity of the fiber tip, but the exact nature of the relation is unclear. We will address this issue in chapter 5.

3 Numerical modeling of the optical field

When simulating the scattering of light in a geometry that is large compared to the wavelength, a description based on geometrical optics can be used. In geometrical optics, light is essentially described as mechanical particles following straight paths in homogenous space and subject to reflection and refraction at surfaces. Computations on such a system can be performed efficiently, and today, powerful ray tracing techniques allow high-quality simulations of light propagating from various sources and being scattered by numerous objects at a speed which can be admired by inspecting the latest 3D computer games.

Unfortunately, the ray tracing algorithms cannot be used to model the scattering of light when the length scales are on the order of or below the wavelength of light. In this case, the zero-wavelength approximation of geometrical optics breaks down, and a complete wave description becomes necessary. In a wave description, light cannot be considered mechanical particles, and a full description of the vectorial nature of the electromagnetic field is necessary. Even though high-quality ray tracing simulations in large geometries can be performed efficiently, the numerical modeling of the optical field in nano-scale geometries remains a challenge. The reason can be understood by inspecting the wave equation describing the electric field in the frequency domain in absence of free charges and currents:

$$\frac{1}{\varepsilon_r(\mathbf{r})}\nabla\times\nabla\times\mathbf{E}(\mathbf{r}) = k_0^2\mathbf{E}(\mathbf{r}), \tag{3.1}$$

From a numerical point of view, all that is needed in geometrical optics to describe the propagation of a light particle are the coordinate sets of the scattering points in its path, whereas a wave description requires a complete solution to Eq. (3.1). To solve this equation a discretization of the geometry on a subwavelength scale is usually necessary. However, storing just one copy of the complex three-component electric field strength in a 10 x 10 x 10 µm geometry with a 25 nm discretization length requires 3 GB of RAM in Matlab, so even with the impressive computing power available in modern PCs, performing numerical simulations of the optical field on a nano-scale geometry remains a difficult task.

3.1 Modeling techniques

Today, a single optimal numerical method for performing nano-scale optical field calculations has not yet been identified. Rather, a number of techniques exists, each with their strengths and weaknesses. The techniques which are most widely used today in NFO include the following:

Finite-difference time-domain (FDTD) method[14]

This technique is probably the most popular in electromagnetics. The computational domain is discretized with a typical discretization length of 50 or 25 nm when performing calculations at optical frequencies. The governing equations are simply finite-difference versions of Maxwell's equations and the optical field is calculated in the time domain by determining the electric and magnetic fields at consecutive points in time. The FDTD algorithm is straight-forward and easy to implement. The initial condition is usually an excitation from a broadband pulse, and thus resonances and eigenmodes over a large bandwidth can be determined from a single computation using Fourier transformation. The major disadvantage of the FDTD technique is the necessity to discretize the entire computational domain on a subwavelength scale, and this sets restrictions to the size of the geometry that the FDTD method can handle. Another weakness is the necessity to introduce artificial boundaries in order to keep the computational domain finite. Advanced boundary conditions such as Perfectly Matched Layers (PMLs) have been developed to reduce the parasitic reflections from the boundary walls, but they are still not completely suppressed.

Finite element method (FEM)[15]

In this procedure, the computational domain is discretized into finite elements in a manner similar to that of the FDTD method. The computation is performed in the frequency domain, and the eigenvalue problem of Eq. (3.1) is solved by discretizing the rotation operators. The FEM is very generic and is implemented in powerful commercial software such as Femlab. It suffers from similar weaknesses as the FDTD technique; a discretization of the entire computational domain is necessary and the domain must be finite, which again introduces parasitic reflections at the boundary walls.

<u>Green's function method</u>[16]

In this frequency domain method, Eq. (3.1) is transformed into an integral equation and a Green's function or propagator function describing the propagation of the electric field is calculated. Once the propagator for a specific refractive index profile is calculated, the scattering of any incoming field can be determined directly using the propagator. A major advantage of this technique is that only the domain where the refractive index profile differs from a background profile, for which the Green's function has already been determined, needs to be discretized. The propagator for a small structure can thus quickly be determined from the analytical propagator describing free space. And when the Green's function for an advanced system has been determined, the propagator for a slightly perturbed system is easily determined, as only the perturbation region needs to be discretized. Another advantage of this technique is that no artificial boundary is needed to limit the computational domain. In spite of these advantages, when performing a calculation from scratch, the necessity to discretize the entire structure remains, and this puts a limitation on the size of the structure that can be studied.

<u>Multiple Multipoles (MMP) technique</u>[17]

The MMP technique is a modal method. The computational domain is divided into regions with constant refractive index and the fields in each region are expanded on basis functions that are the analytical free-space solutions of Eq. (3.1). By writing the field as a sum of analytical solutions, only the boundaries between the adjacent domains need to be discretized. The boundary conditions lead to a series of linear equations that are solved to obtain the optical field profile, and the expansion of the field on eigenfunctions of the wave equation usually results in rapid convergence. Many basis sets for the field expansion are possible, but multipole functions are generally found to be the most efficient. A weakness of the MMP technique, however, is that the convergence speed depends on the set of multipole functions used, and prior knowledge about the field profile is thus necessary in order to choose the optimal basis set.

The fiber tip structures studied in this Ph.D. project are large, and it was estimated that if the calculations were to be performed using the FDTD, the FEM or the Green's function technique, the computations would be too slow due to the large discretization

domain. The author therefore decided to investigate the Eigenmode Expansion Technique (EET), a modal method similar to the MMP method.

3.2 Eigenmode Expansion Technique

In this frequency domain technique, the geometry is divided into layers of uniform refractive index profile along a propagation axis, usually chosen as the z axis. For each layer, eigenmodes are determined for the particular index profile assuming uniformity along the entire z axis. As the name implies, the optical field in each layer is expanded on the corresponding eigenmodes, and the fields at each side of the interface between adjacent layers are connected using the scattering matrix formalism. The determination of eigenmodes, the expansion of the field on the eigenmodes and the scattering matrix formalism will be discussed in detail in the following.

The difference between the discretizations used in techniques such as the FDTD method and the EET is illustrated in the following figure. The majority of methods introduces a subwavelength discretization, illustrated in Fig. 3.1(a), of the computational domain leading to huge memory requirements. However, in the EET the geometry is only divided into layers of uniform refractive index profile along the z axis as sketched in Fig. 3.1(b).
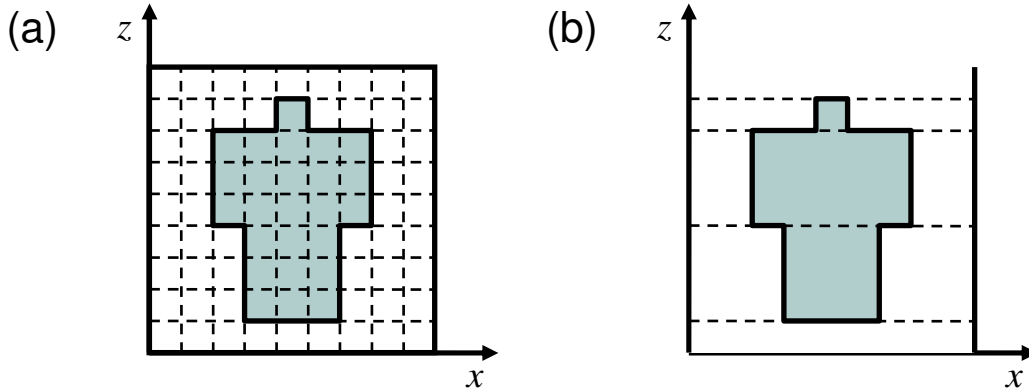
Fig. 3.1: A computational domain with subwavelength discretization is shown in (a). The corresponding EET layer division is illustrated in (b).

In the EET, no subwavelength discretization is necessary,[b] and for each layer only the eigenmode profiles and a set of expansion coefficients need to be stored in memory.

---

[b] The determination of eigenmodes using the finite difference technique is an exception; here a discretization is made in the *x-y* plane.

When performing calculations on a structure featuring long sections that are uniform along one axis, the EET thus has a huge advantage compared to e.g. the FDTD technique, as considerably less memory is required to store the optical field. Furthermore, when describing the field in a particular layer, the basis set in the field expansion consists of eigenmodes that are already exact[c] solutions to the wave equation. For this reason, it is expected that relatively few eigenmodes are needed to describe the optical field leading to an additional reduction in the memory requirement of the EET compared to other techniques. The convergence in the computation of transfer functions (TFs) is described in detail in section 5.2. However, a weakness of the EET is the necessity to expand the optical field on a finite set of eigenmodes. This means that the geometry under study must be limited by artificial boundaries in the *x-y* plane, introducing the usual problem of parasitic reflections from the boundary walls of this plane. No boundaries, however, are necessary to limit the geometry along the *z* axis.

Procedures[18,19] to determine eigenmodes in layers that are uniform along an axis and the mode matching technique[20,21] used to connect the fields at the interfaces have been known for many years. However, it is only recently that the EET has been used to perform calculations in large structures, where its advantages with respect to memory requirements and computation speed are clearly manifested: The EET has been used by Bienstman[22-24] to calculate the optical field in photonic crystal waveguides and in vertical-cavity surface-emitting lasers (VCSELs) in the presence of advanced boundary conditions such as PMLs to avoid parasitic reflections. And the performance of single-mode bottom-emitting VCSELs was optimized using the EET in Ref. 25.

All calculations of the optical field in this Ph.D. thesis are made using the EET. To the author's knowledge, the EET has not yet been implemented in any commercial software, and even though the code developed by Bienstman for his calculations is available[26], this code is limited to structures with a single refractive index step in the radial direction. For this reason, the EET calculations have been performed with home-made software, written in Matlab. The details of the method and the

---

[c] The eigenmodes are virtually exact when they are determined using the semi-analytical method described in section 3.4.3. If they are calculated using the finite difference or the plane-wave techniques described in sections 3.4.1 and 0, their precision depends on the number of points or plane waves used in the calculation.

considerations made during the development of the software are described in the following two sections.

3.3 Eigenmodes

The starting point of the EET calculation is the determination of eigenmodes that serve as a basis set for the expansion of the optical field. To calculate these, the refractive index profile is sliced into layers having a uniform index profile along the z axis. In Fig. 3.2(a), the index profile of a random glass sample is illustrated. The sample itself can be separated into three layers, II to IV, uniform along the z axis. Also including the air layers I and V, the structure consists of five layers in total. Since layers I and V are both air layers, there are only four different index profiles.
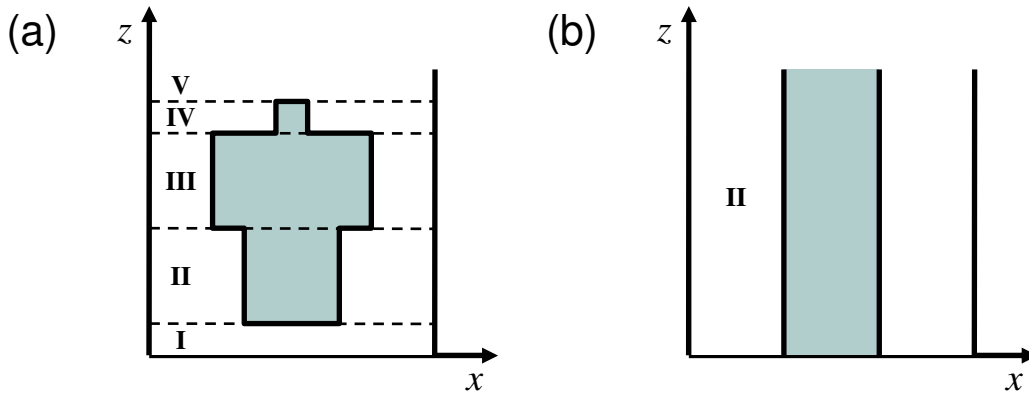
Fig. 3.2: The division of the geometry into five layers is sketched in (a). In (b), the *x-z* profile for zone II with uniformity along the *z* axis is illustrated.

For each layer, we now consider the geometry with index profile in the *x-y* plane corresponding to the original geometry, but with uniformity along the z axis. In the following, we refer to such a geometry as a *zone*. In Fig. 3.2(b), the *x-z* profile of zone II corresponding to layer II in Fig. 3.2(a) is depicted. Zones are infinite along the z axis and the zone profile of this example is identical to that of a glass waveguide, guiding light along the z axis.

We now turn our attention to the 2D *x-y* profile of each zone. The EET can, in principle, handle any index profile, such as the rectangular waveguide illustrated in Fig. 3.3(a).
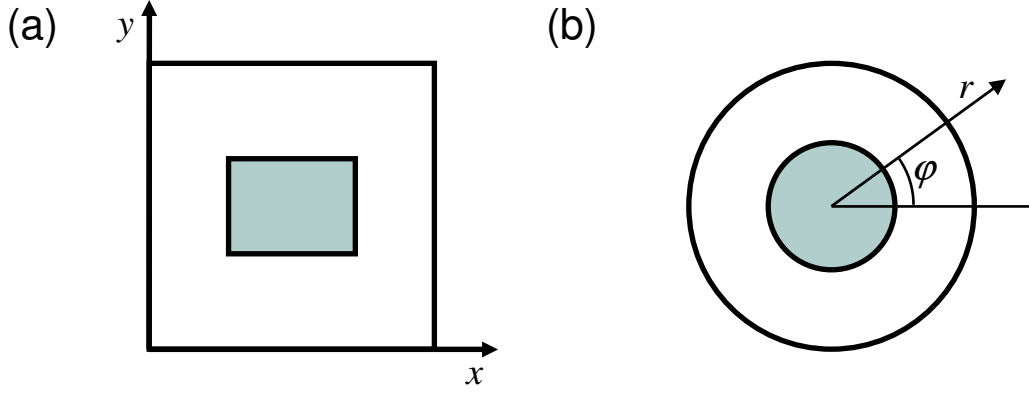
Fig. 3.3: A rectangular profile in the *x-y* plane is illustrated in (a). A profile in the same plane featuring cylindrical symmetry is sketched in (b).

Determining a single or a few precise eigenmodes in rectangular waveguides is not completely trivial due to the lack of analytic solutions and because of the divergences in the electric field that occur at the corners of the waveguide[27], but it is certainly possible[18]. However, the problem with eigenmode expansion in two-dimensional systems is that each dimension usually requires expansion on a minimum of 50 to 100 modes in a computation to ensure convergence. In two dimensions, somewhere between 2500 and 10000 modes are thus required, and the determination of so many eigenmodes is extremely numerically demanding. For this reason, the structures studied in this work either feature uniformity along the *y* axis or rotational symmetry as in the waveguide illustrated in Fig. 3.3(b). In these geometries, the effective number of dimensions of the eigenvalue problem is reduced to just one, and thus only a few hundred eigenmodes are needed in the computations.

The boundary conditions used in the computation of the transfer functions are perfectly conducting metal walls. The inclusion of advanced boundary conditions such as PMLs was estimated to be unnecessary as described further in section 5.2.1. In the topography reconstruction of the grating, periodic boundary conditions were employed.

3.3.1 The eigenvalue problem

Having defined zones that are uniform along one axis, we now inspect the wave equation from Eq. (3.1) for *z*-independent index profiles with the purpose of deriving the eigenvalue problem:

21

$$\nabla^2 \mathbf{E}(\mathbf{r}) - \nabla(\nabla \cdot \mathbf{E}(\mathbf{r})) + \varepsilon_r(\mathbf{r}_\perp)k_0^2 \mathbf{E}(\mathbf{r}) = 0 \qquad (3.2)$$

with time dependence given by $\exp(-i\omega t)$. $\mathbf{r}_\perp$ refers to the coordinates (x, y) of the plane normal to the z axis, and the vector identity $\nabla \times \nabla \times \mathbf{F} = -\nabla^2 \mathbf{F} + \nabla(\nabla \cdot \mathbf{F})$ has been used. z independence invites a Fourier transformation and we thus write the electric field as $\mathbf{E}(\mathbf{r}) = \mathbf{e}(\mathbf{r}_\perp)\exp(i\beta z)$, where β is the propagation constant. Furthermore, from the divergence relation

$$\nabla \cdot (\varepsilon_r \mathbf{E}) = 0 \qquad (3.3)$$

we obtain the expression $\nabla \cdot \mathbf{E} = -\mathbf{E} \cdot (\nabla \ln(\varepsilon_r(\mathbf{r}_\perp)))$. Insertion of the above expressions into Eq. (3.2) gives:

$$\nabla_\perp^2 \mathbf{e} + (\nabla_\perp + i\beta \mathbf{u}_z)(\mathbf{e} \cdot \nabla \ln \varepsilon_r(\mathbf{r}_\perp)) + \varepsilon_r(\mathbf{r}_\perp)k_0^2 \mathbf{e} = \beta^2 \mathbf{e} \qquad (3.4)$$

Here, $\mathbf{u}_z$ is the unit vector along the z axis. Eq. (3.4) is a second-order eigenvalue problem of the form $\hat{O}_0 \mathbf{e} + \hat{O}_1 \lambda \mathbf{e} + \hat{O}_2 \lambda^2 \mathbf{e} = 0$ as we have a β on the left side. Second-order eigenvalue problems are much more demanding to solve numerically than standard eigenvalue problems. Fortunately, the scalar equations for the in-plane components $\mathbf{e}_\perp$ in Eq. (3.4) are not coupled to e_z, and it thus suffices to solve the equation

$$\nabla_\perp^2 \mathbf{e}_\perp + \nabla_\perp(\mathbf{e}_\perp \cdot \nabla \ln \varepsilon_r(\mathbf{r}_\perp)) + \varepsilon_r(\mathbf{r}_\perp)k_0^2 \mathbf{e}_\perp = \beta^2 \mathbf{e}_\perp, \qquad (3.5)$$

which is a standard eigenvalue problem of the form $\hat{O}\mathbf{e}_\perp = \lambda \mathbf{e}_\perp$. Eq. (3.5) is the central equation. By solving it, we obtain eigenmodes $\mathbf{e}_{\perp;n}(\mathbf{r}_\perp)$ and eigenvalues that are the squares of the propagation constants β_n. These eigenmodes constitute the basis set upon which the optical field in the zone is expanded, and we can then describe an arbitrary optical field using

$$\mathbf{E}_\perp(\mathbf{r}) = \sum_n a_n^+ \mathbf{e}_{\perp;n}(\mathbf{r}_\perp)\exp(i\beta_n z) + \sum_n a_n^- \mathbf{e}_{\perp;n}(\mathbf{r}_\perp)\exp(-i\beta_n z), \qquad (3.6)$$

where the expansion coefficients a_n^+ and a_n^- refer to the forward- and backward-propagating parts of the field, respectively.

Even though Eqs. (3.5) and (3.6) only involve the in-plane electric field components $\mathbf{e}_\perp$, the complete six-component electromagnetic field description is readily available: When an in-plane profile has been calculated, we can determine the z component using the identity

$$e_z = -\nabla_\perp \cdot (\varepsilon_r \mathbf{e}_\perp)/(i\beta), \tag{3.7}$$

derived from the divergence relation from Eq. (3.3). Subsequently, the magnetic field strength $\mathbf{H}(\mathbf{r}_\perp)$ can be calculated from the Maxwell equation $\nabla \times \mathbf{E} = i\omega_0 \mu \mathbf{H}$.

In Cartesian coordinates (x, y), the matrix form of Eq. (3.5) is:

$$\begin{bmatrix} \nabla_\perp^2 + \varepsilon_r k_0^2 + \dfrac{\partial}{\partial x}\left(\dfrac{\partial}{\partial x}\ln\varepsilon_r\right) & \dfrac{\partial}{\partial x}\left(\dfrac{\partial}{\partial y}\ln\varepsilon_r\right) \\ \dfrac{\partial}{\partial y}\left(\dfrac{\partial}{\partial x}\ln\varepsilon_r\right) & \nabla_\perp^2 + \varepsilon_r k_0^2 + \dfrac{\partial}{\partial y}\left(\dfrac{\partial}{\partial y}\ln\varepsilon_r\right) \end{bmatrix}\begin{bmatrix} e_x \\ e_y \end{bmatrix} = \beta^2 \begin{bmatrix} e_x \\ e_y \end{bmatrix}. \tag{3.8}$$

From this equation we observe how the two field components are coupled by discontinuities in the dielectric constant. In chapter 4, we study a grating with uniformity along the y axis illuminated with incident light normal to the y axis. In this case the off-diagonal matrix elements vanish and the eigenvalue problem can be solved for each polarization separately. The possibility of decoupling two field components is important as it effectively allows a doubling of the precision with which an eigenmode can be determined.

In cylindrical coordinates we can write the electric field as a sum of contributions of different angular momenta:

$$\mathbf{e}_\perp(\mathbf{r}) = \sum_l \mathbf{e}_{\perp;l}(r)\exp(il\theta), \tag{3.9}$$

where l is the angular momentum. In a rotationally symmetric structure, the various contributions are not coupled, and it is thus necessary to perform calculations for only one angular momentum at a time. This effectively reduces the number of dimensions from 3 to 2, making exact 3D vectorial calculations feasible with only modest computing power. The calculation of the transfer function is performed for the angular momentum $l = 1$ corresponding to the guided mode of a single-mode fiber, and the index of angular momentum is skipped for simplicity.

When expressed in cylindrical coordinates (r, θ) the matrix form of Eq. (3.5) for a rotationally symmetric index profile becomes:

$$\begin{bmatrix} \nabla_\perp^2 - \dfrac{1}{r^2} + \varepsilon_r k_0^2 + \dfrac{\partial}{\partial r}\left(\dfrac{\partial}{\partial r}\ln\varepsilon_r\right) & -\dfrac{2il}{r^2} \\ \dfrac{2il}{r^2} + \dfrac{il}{r}\left(\dfrac{\partial}{\partial r}\ln\varepsilon_r\right) & \nabla_\perp^2 - \dfrac{1}{r^2} + \varepsilon_r k_0^2 \end{bmatrix}\begin{bmatrix} e_r \\ e_\theta \end{bmatrix} = \beta^2 \begin{bmatrix} e_r \\ e_\theta \end{bmatrix}. \tag{3.10}$$

In this case, the e_r and e_θ components are coupled and the equation must be solved for both components simultaneously. The exception is when the angular momentum l is equal to zero.

3.3.2 Eigenmode classes

The eigenmode profile describes the mode in the *x-y* plane, however, its *z* dependence is governed by its propagation constant β alone. It is therefore useful to classify different types of eigenmodes according their eigenvalues β^2.
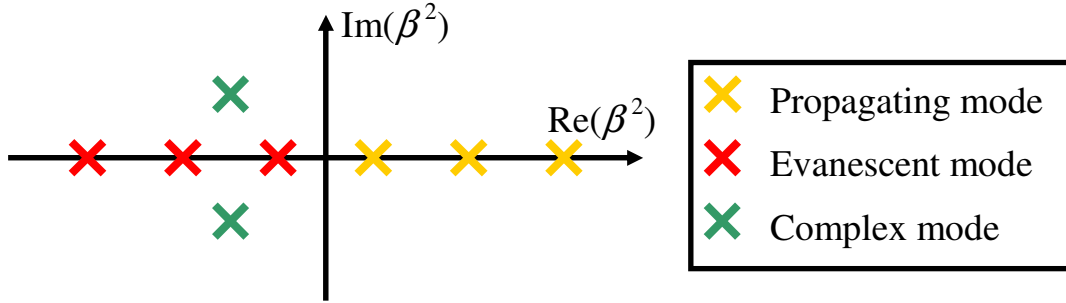
Fig. 3.4: The distribution of the eigenvalues $\boldsymbol{\beta^2}$ in the complex plane is sketched.

To do this, we consider the eigenvalues obtained in an arbitrary geometry, but with the restriction that the dielectric constant should be real-valued. In Fig. 3.4, three classes of eigenvalues are illustrated. When β^2 is real and $\beta^2 > 0$ or $\beta^2 < 0$, the propagation constant β is real or purely imaginary, respectively, and the mode is propagating or decaying along the *z* axis as described in section 2.1. However, there exists a third class of modes: Eigenmodes with complex eigenvalues can exist[28] even in real-valued refractive index profiles. These modes occur in complex conjugated pairs and they feature an oscillating decay along the *z* axis as their propagation constant is complex. The existence of these complex modes do not pose a problem for the eigenmode expansion, but the determination of eigenmodes using the semi-analytical approach is compromised[d] by their existence.

When studying optical waveguides in geometries without boundaries, one usually makes the distinction between guided and radiative modes. Guided modes are a discrete set of modes confined to the waveguide, whereas radiative modes extend to

---

[d] The weakness of the semi-analytical approach in handling complex modes will be discussed in section 3.4.3.

infinity and form a continuous set. In our geometries limited by artificial boundaries in the *x-y* plane, we can define semi-radiative modes as the discrete set of modes that extend to the entire geometry. However, unlike the class of complex modes that pose a problem for the semi-analytical approach[d], there is no practical reason in the EET for distinguishing between guided and semi-radiative modes.

3.3.3 Orthogonality

Even when considering the case where the dielectric constant is real-valued, the operator of the wave equation of Eq. (3.1) is not Hermitian. However, by setting $\mathbf{G}(\mathbf{r}) = \sqrt{\varepsilon_r(\mathbf{r}_\perp)}\mathbf{E}(\mathbf{r})$ we obtain the slightly modified form of the same equation:

$$\frac{1}{\sqrt{\varepsilon_r(\mathbf{r}_\perp)}}\nabla\times\nabla\times\frac{1}{\sqrt{\varepsilon_r(\mathbf{r}_\perp)}}\mathbf{G}(\mathbf{r}) = k_0^2\mathbf{G}(\mathbf{r}),\qquad(3.11)$$

where the operator on the left side is Hermitian. The basis set $\mathbf{G}_n(\mathbf{r}) = \sqrt{\varepsilon_r(\mathbf{r}_\perp)}\mathbf{e}_n(\mathbf{r}_\perp)\exp(i\beta_n z)$ defined in 3D space is a solution to Eq. (3.11) and is thus orthogonal with respect to the inner product $\langle\mathbf{G}_m|\mathbf{G}_n\rangle = \int\mathbf{G}_m^*\cdot\mathbf{G}_n d\mathbf{r}$. However, the orthogonality is assured by integrating in all three dimensions. The eigenvalue problem of Eq. (3.5) is only two-dimensional and the inner product is given by $\langle\mathbf{f}_m|\mathbf{f}_n\rangle = \int\mathbf{f}_m^*\cdot\mathbf{f}_n d\mathbf{r}_\perp$. The operator is not Hermitian even when introducing the expression $\mathbf{g}_n = \sqrt{\varepsilon_r(\mathbf{r}_\perp)}\mathbf{e}_n$, and neither of the basis sets $\mathbf{e}_n$ or $\mathbf{g}_n$ are orthogonal. However, another orthogonality relation derived from the reciprocity theorem[19] can be used. In absence of free currents the reciprocity theorem states that, given two mode profiles $(\mathbf{E}_1,\mathbf{H}_1)$ and $(\mathbf{E}_2,\mathbf{H}_2)$, the relation

$$\frac{\partial}{\partial z}\int(\mathbf{E}_1\times\mathbf{H}_2 - \mathbf{E}_2\times\mathbf{H}_1)\cdot\mathbf{u}_z d\mathbf{r}_\perp = 0\qquad(3.12)$$

holds true. Our eigenmodes are of the form

$$\begin{bmatrix}\mathbf{E}_{\perp,1}(\mathbf{r})\\E_{z,1}(\mathbf{r})\end{bmatrix}=\begin{bmatrix}\mathbf{e}_{\perp,1}(\mathbf{r}_\perp)\\e_{z,1}(\mathbf{r}_\perp)\end{bmatrix}\exp(i\beta_1 z)\quad\begin{bmatrix}\mathbf{H}_{\perp,1}(\mathbf{r})\\H_{z,1}(\mathbf{r})\end{bmatrix}=\begin{bmatrix}\mathbf{h}_{\perp,1}(\mathbf{r}_\perp)\\h_{z,1}(\mathbf{r}_\perp)\end{bmatrix}\exp(i\beta_1 z)\qquad(3.13)$$

$$\begin{bmatrix}\mathbf{E}_{\perp,2}(\mathbf{r})\\E_{z,2}(\mathbf{r})\end{bmatrix}=\begin{bmatrix}\mathbf{e}_{\perp,2}(\mathbf{r}_\perp)\\e_{z,2}(\mathbf{r}_\perp)\end{bmatrix}\exp(i\beta_2 z)\quad\begin{bmatrix}\mathbf{H}_{\perp,2}(\mathbf{r})\\H_{z,2}(\mathbf{r})\end{bmatrix}=\begin{bmatrix}\mathbf{h}_{\perp,2}(\mathbf{r}_\perp)\\h_{z,2}(\mathbf{r}_\perp)\end{bmatrix}\exp(i\beta_2 z),\qquad(3.14)$$

and insertion into Eq. (3.12) gives:

$$(\beta_2 + \beta_1)\int (\mathbf{e}_1 \times \mathbf{h}_2 - \mathbf{e}_2 \times \mathbf{h}_1) \cdot \mathbf{u}_z d\mathbf{r}_\perp = 0. \tag{3.15}$$

We now consider the transverse mode profile $(\mathbf{E}_1, \mathbf{H}_1)$ from before and the mode profile $(\mathbf{E}_2', \mathbf{H}_2')$, which is the backward propagating version of $(\mathbf{E}_2, \mathbf{H}_2)$ given by:

$$\begin{bmatrix} \mathbf{E}_{\perp,2}'(\mathbf{r}) \\ E_{z,2}'(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{\perp,2}(\mathbf{r}_\perp) \\ -e_{z,2}(\mathbf{r}_\perp) \end{bmatrix} \exp(-i\beta_2 z) \quad \begin{bmatrix} \mathbf{H}_{\perp,2}'(\mathbf{r}) \\ H_{z,2}'(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} -\mathbf{h}_{\perp,2}(\mathbf{r}_\perp) \\ h_{z,2}(\mathbf{r}_\perp) \end{bmatrix} \exp(-i\beta_2 z). \tag{3.16}$$

The components of the mode profile are identical to those in Eq. (3.14), except for the minus sign in the expression for the E_z and $\mathbf{H}_\perp$ the components. It appears since the sign of these components depends on the propagation constant as can be seen by inspecting Eq. (3.7) or the mode profiles given in appendix 8.1. Insertion of $(\mathbf{E}_1, \mathbf{H}_1)$ and $(\mathbf{E}_2', \mathbf{H}_2')$ into Eq. (3.12) results in:

$$(\beta_2 - \beta_1)\int (\mathbf{e}_1 \times \mathbf{h}_2 + \mathbf{e}_2 \times \mathbf{h}_1) \cdot \mathbf{u}_z d\mathbf{r}_\perp = 0. \tag{3.17}$$

Now, if $\beta_1 \neq \beta_2$ the integrals in Eqs. (3.15) and (3.17) must equal zero. Taking their sum, we obtain the orthogonality relation:

$$\int (\mathbf{e}_1 \times \mathbf{h}_2) \cdot \mathbf{u}_z d\mathbf{r}_\perp = 0. \tag{3.18}$$

This orthogonality relation is used in the mode-matching taking place at the interfaces between adjacent layers, as described further in section 3.5.1. The orthogonality relation is valid in arbitrary geometries, including those with complex index profiles.

3.3.4 Completeness

In order to represent any optical field using eigenmode expansion, the basis set must be complete. For geometries featuring a real-valued refractive index profile and uniformity along the y axis, it has been proven[29] that the eigenmodes obtained by solving Eq. (3.5) form a complete set. It is unclear whether the completeness extends to general geometries with complex index profiles, but so far no evidence has indicated that it is not the case and the completeness of the eigenmodes is thus generally assumed in spite of the lack of formal proof.

3.4 Determination of eigenmodes

In this section, the methods used in this Ph.D. thesis to determine eigenmodes are described. Two different optical field simulation tools have been programmed: The first is for generic 3D structures described using Cartesian coordinates. The computations in the topography reconstruction described in chapter 4 are performed using this solver. The second is for 3D structures featuring rotational symmetry, where a description using cylindrical coordinates is advantageous. The calculations of transfer functions of SNOM fiber tips described in chapter 5 are made using this code.

For the generic 3D structures an eigenmode solver based on plane-wave expansion was used. The Cartesian coordinate geometries studied in this work were glass gratings, which are simple structures that were handled efficiently using the plane wave code.

However, when calculating the transfer function of various SNOM fiber tips, obtaining proper convergence was a challenge. To determine which method would result in the best convergence, three different techniques for solving the eigenmode problem in cylindrical coordinates were investigated. A code based on a finite-difference approach to determining eigenmodes was made. Furthermore, a code, where the eigenmode is expanded on the eigenmodes of the empty metal cylinder, was developed. This can be considered a plane-wave expansion except for the basis set consisting of Bessel functions due to the cylindrical coordinate description rather than the trigonometric functions used when employing Cartesians coordinates. Finally, a code implementing a semi-analytical approach to determine mode profiles was programmed. Here, the eigenmodes are given as analytic solutions to the wave equation and are thus virtually exact.

Each method for determining eigenmodes has strength and weaknesses, and the method of choice will depend on the geometry under study.

3.4.1 Finite-difference technique

The finite-difference (FD) technique is probably the oldest and most well-known technique in computational electromagnetics and is certainly one of the simplest to

understand. The technique can be implemented in both Cartesian and cylindrical coordinates, but in this work the technique was used only for cylindrical coordinates.
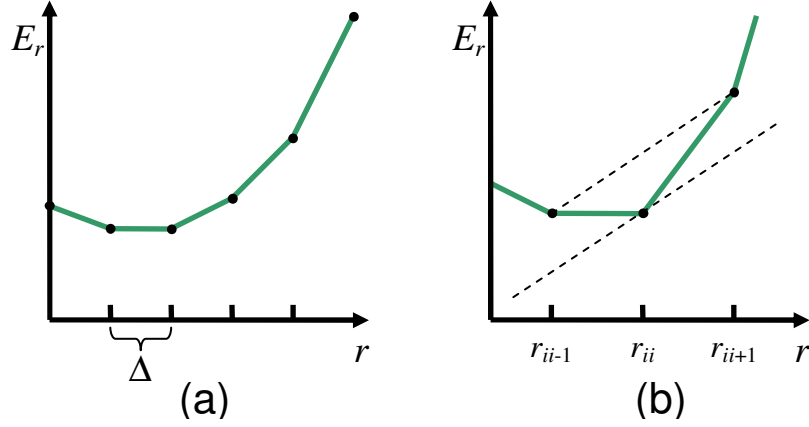
Fig. 3.5: In (a), a mode profile discretized in real space is depicted. In (b), the finite difference scheme used to perform differentiation at the point r_{ii} is illustrated.

A discretization grid (r_{ii}) is chosen in real space with separation distance Δ between nearest neighbours, and the field is evaluated at the discretization points as illustrated in Fig. 3.5(a). Numerically, the field components are represented by vectors as

$$e_r(r) \approx \begin{bmatrix} ... \\ e_r(r_{ii-1}) \\ e_r(r_{ii}) \\ e_r(r_{ii+1}) \\ ... \end{bmatrix} \quad \text{and} \quad e_\theta(r) \approx \begin{bmatrix} ... \\ e_\theta(r_{ii-1}) \\ e_\theta(r_{ii}) \\ e_\theta(r_{ii+1}) \\ ... \end{bmatrix}.$$

The field is known only at these points and the regions between points are described using interpolation.

The operator elements of Eq. (3.10) are then expressed in their finite-difference form and the eigenvalue problem is solved directly. The multiplicative elements are represented by matrices, where all off-diagonal elements are zero. The differential operator is expressed using the FD scheme illustrated in Fig. 3.5(b). Differentiation D_{ii} at a point r_{ii} is evaluated by taking the difference between the field values at the neighbouring points r_{ii+1} and r_{ii-1} and dividing it with their separation in space. Matrices representing the multiplicative operator $\varepsilon_r(r)$ and the differential operator $\partial/\partial r$ in the FD scheme for equidistant points are shown in the following figure:

$$(a) \quad \begin{bmatrix} \ddots & & & & 0 \\ & \varepsilon_r(r_{ii-1}) & & & \\ & & \varepsilon_r(r_{ii}) & & \\ & & & \varepsilon_r(r_{ii+1}) & \\ 0 & & & & \ddots \end{bmatrix} \qquad (b) \quad \begin{bmatrix} \ddots & & & & 0 \\ -\dfrac{1}{2\Delta} & 0 & \dfrac{1}{2\Delta} & & \\ & -\dfrac{1}{2\Delta} & 0 & \dfrac{1}{2\Delta} & \\ & & -\dfrac{1}{2\Delta} & 0 & \dfrac{1}{2\Delta} \\ 0 & & & & \ddots \end{bmatrix}$$

Fig. 3.6: Operator matrices are sketched representing the multiplicative operator $\varepsilon_r(r)$ in (a) and the differential operator for equidistant points in (b).

In this work, only piece-wise constant refractive index profiles along the radial axis were considered, but the generic form of the matrix representing $\varepsilon_r(r)$ in Fig. 3.6(a) can handle an arbitrary index profile, e.g. one with gradients.

The operator matrix for $\partial^2/\partial r^2$ is also required to express the operator of Eq. (3.11) and is obtained simply by squaring the matrix in Fig. 3.6(b). One might imagine that particular modifications to the matrices representing differential operators would be necessary to assure the correct handling of the discontinuities of the E_r component. However, this is not the case: Except for a correction necessary when using non-equidistant points, which is explained below, the matrices are used in their raw form and the discontinuity of the E_r component appears automatically.

An advantage of the finite difference scheme is that the elements of the operator matrix are zero except near the diagonal row. The operator matrix can thus be represented as a sparse matrix in Matlab requiring much less memory than a full matrix. Also, less calculation time is spent in the determination of eigenvectors of a sparse matrix than of a full one of same size. And it should be noted that in the FD technique, the computation time necessary to calculate modes is independent of the refractive index profile of the geometry.

When solving an eigenvalue problem involving N free variables, a maximum of N eigenvectors can be determined, and inversely, when requesting N eigenvectors, at least N free variables should included. This means that, the more eigenmodes are included in a computation, the more discretization points are used to represent these modes, and the more accurate the mode profiles become. However, we also have the

option of increasing the point resolution without changing the number *N* of eigenmodes by setting up the operator matrix corresponding to the increased resolution and then simply requesting only *N* eigenvectors from the eigenvalue problem solver routine. In this work, the point resolution was increased with a factor of 2 in all FD calculations to improve the quality of the mode profiles.

An example of an eigenmode profile determined using the FD technique is depicted in Fig. 3.7. The refractive index geometry is a glass waveguide of 500 nm radius with refractive index of ~ 1.46. The calculation is performed at the wavelength λ equal to 633 nm, and the fundamental mode obtained using a fairly coarse discretization is shown.
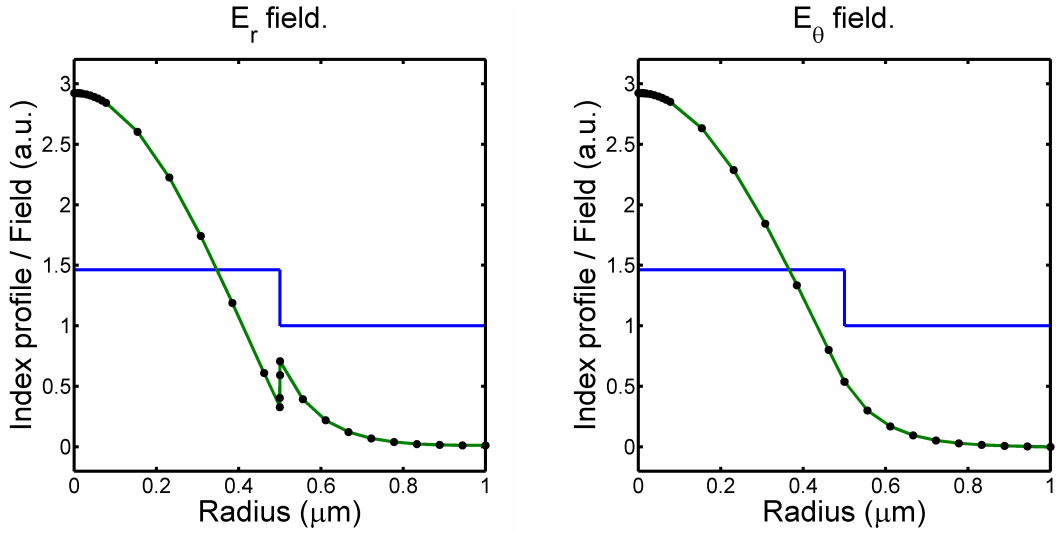
Fig. 3.7: **The refractive index profile (blue) together with the in-plane components (green) of a cylindrically symmetric mode profile obtained using the FD method are depicted.**

The mode is clearly confined to the waveguide, and we observe the discontinuity of the radial component at the glass-air interface. As illustrated in the figure, equidistant discretization points are generally used. However, it is convenient to include extra points at particular positions on the *r* axis:

The finite difference technique has difficulty in describing the mode correctly near $r = 0$, most likely due to the appearance of $1/r^2$ in the off-diagonal elements of the operator matrix in Eq. (3.10). For this reason, additional points were added with a separation of 1/8 of the original discretization distance as illustrated in Fig. 3.8(a).
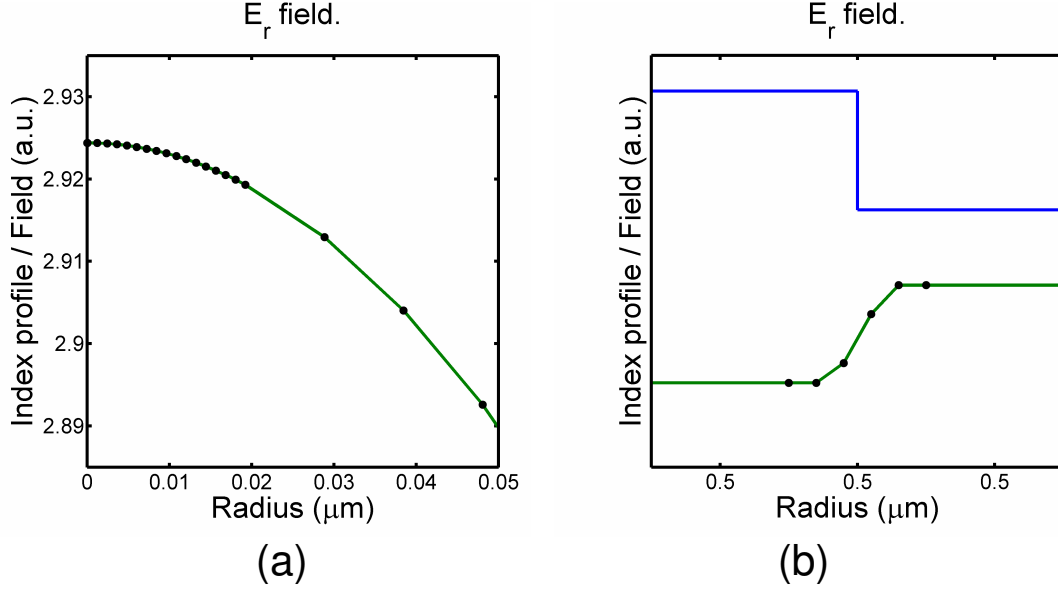
Fig. 3.8: **The inclusion of additional discretization points is illustrated near** $r = 0$ **(a)
and at the glass-air interface (b).**

Furthermore, additional points were added at each interface between rings of different refractive index. As illustrated in Fig. 3.8(b), the finite difference technique handles the discontinuity of the radial field component well. The jump, however, of the E_r component is distributed over four discretization points. If only equidistant points were used at the interfaces, the center of the continuous four-point jump of the E_r component could be at a distance of up to $\Delta/2$ from the interface and the discontinuous nature of the jump would not be obvious. We thus included additional points to assure that the field discontinuity occurs exactly at the interface and set the separation distance to 2 pm to make the jump sharp compared to the standard point separation distance Δ, which usually is ~ 10 nm.

When expressing the differential operator for non-equidistant points, the FD matrix of Fig. 3.6(b) should be modified in order to correctly handle the discontinuities of the E_r component at the interfaces: Correction coefficients that take into account the different separation length to the neighbouring points must be included:

31
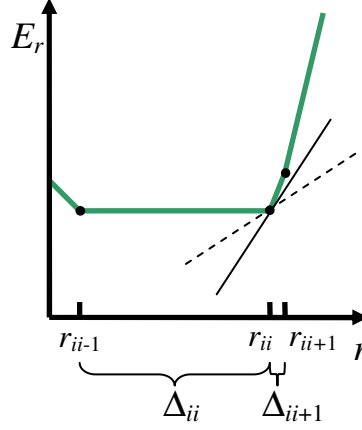
Fig. 3.9: Illustration of the standard gradient (dashed line) and the corrected gradient using weighing coefficients (full line) for non-equidistant points.

Referring to Fig. 3.9, the unmodified FD gradient D_{ii} at the discretization point r_{ii} can be expressed as the average of the gradient between points r_{ii-1} and r_{ii} and the gradient between points r_{ii} and r_{ii+1}:

$$D_{ii} = \frac{E_r(r_{ii}) - E_r(r_{ii-1})}{2\Delta_{ii}} + \frac{E_r(r_{ii+1}) - E_r(r_{ii})}{2\Delta_{ii+1}}. \tag{3.19}$$

We now introduce the correction coefficients $\Delta_{ii+1}/(\Delta_{ii} + \Delta_{ii+1})$ and $\Delta_{ii+1}/(\Delta_{ii} + \Delta_{ii+1})$, that take into account the unequal separation distances. The modified finite difference D_{ii} gradient becomes:

$$D_{ii} = \frac{\Delta_{ii+1}}{\Delta_{ii} + \Delta_{ii+1}} \frac{E_r(r_{ii}) - E_r(r_{ii-1})}{\Delta_{ii}} + \frac{\Delta_{ii}}{\Delta_{ii} + \Delta_{ii+1}} \frac{E_r(r_{ii+1}) - E_r(r_{ii})}{\Delta_{ii+1}}. \tag{3.20}$$

When $\Delta_{ii} = \Delta_{ii+1} = \Delta$, this reduces $D_{ii} = \left(E_r(r_{ii+1}) - E_r(r_{ii-1})\right)/(2\Delta)$. The standard gradient obtained using Eq. (3.19) and the corrected gradient of Eq. (3.20) are illustrated in Fig. 3.9.

3.4.2 Plane-wave expansion

In the plane-wave (PW) expansion technique, an eigenmode $\mathbf{e}_{\perp;n}$ is represented by a linear combination of basis modes:

$$\mathbf{e}_{\perp;n}(\mathbf{r}_\perp) = \sum_{j=1}^{N} c_{n,j} \mathbf{b}_{\perp;j}(\mathbf{r}_\perp), \tag{3.21}$$

where the basis modes $\mathbf{b}_{\perp;j}$ is any complete set of modes. For numerical reasons we are forced to consider a limited number N of basis modes.
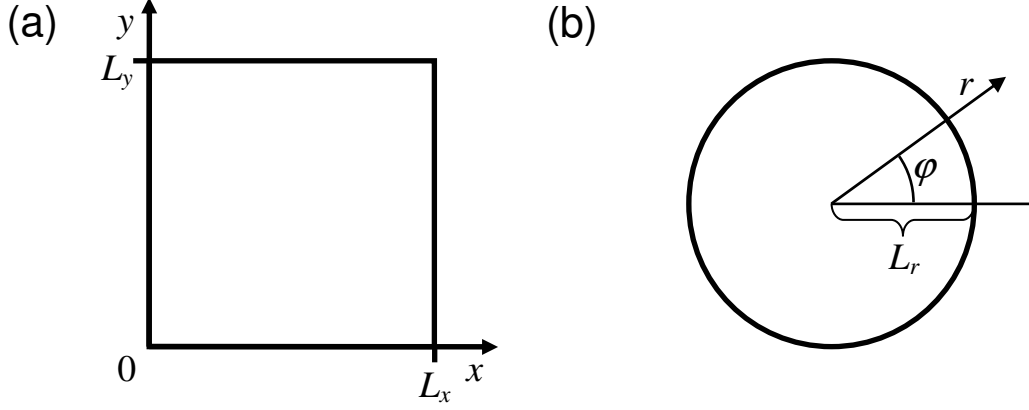
Fig. 3.10: The rectangular (a) and rotationally symmetric (b) vacuum geometries, whose solutions are used as basis sets, are illustrated.

As basis set, it is convenient to chose the solutions to the governing Eq. (3.5) for the vacuum geometries illustrated in Fig. 3.10. When the dielectric constant is equal to unity, the equation becomes:

$$\nabla^2_\perp \mathbf{b}_\perp + k_0^2 \mathbf{b}_\perp = \beta^2 \mathbf{b}_\perp . \tag{3.22}$$

The operator in Eq. (3.22) is Hermitian and the solutions are orthogonal plane waves, that can be described analytically. In general, all six field components of a solution to the wave equation are non-zero. However in the particular case of the vacuum geometry, the solutions can be classified as either transverse electric (TE) and transverse magnetic (TM). The complete solutions in Cartesian and cylindrical coordinates to the vacuum geometries of Fig. 3.10, including the z component and the magnetic field, are given in appendix 8.1.

When solving Eq. (3.5), the $\mathbf{e}_\perp$ components of the solutions to the vacuum geometry listed in the tables of the appendix are used as a basis set. When subsequently determining the z component and the magnetic field, the corresponding components from the tables are used as basis sets for the e_z and $\mathbf{h}$ components. Since the basis modes used to expand the various field components are chosen to satisfy the boundary conditions of the geometry, the boundary conditions are automatically respected by the eigenmodes as well.

Now, to calculate eigenmodes in geometries with non-uniform refractive index profile, we must represent the operator $\hat{O}$ of the eigenvalue problem in Eq. (3.5) as a matrix. To do this we use the inner product

$$\left\langle \mathbf{b}_{\perp;m} \middle| \mathbf{b}_{\perp;n} \right\rangle = \int \mathbf{b}_{\perp;m}^* \cdot \mathbf{b}_{\perp;n} d\mathbf{r}_\perp \, , \tag{3.23}$$

and the elements $O_{m,n}$ of the operator matrix are then given by:

$$O_{m,n} = \left\langle \mathbf{b}_{\perp;m} \middle| \hat{O} \mathbf{b}_{\perp;n} \right\rangle , \tag{3.24}$$

assuming that the basis modes are properly normalized such that $\left\langle \mathbf{b}_{\perp;m} \middle| \mathbf{b}_{\perp;n} \right\rangle = \delta_{mn}$. The δ_{mn} is a Kronecker delta, since our geometry is limited in the $\mathbf{r}_\perp$ plane. We now split the operator $\hat{O}$ of Eq. (3.5) into three operators as:

$$\left(\hat{O}_0 + \hat{O}_1 + \hat{O}_2 \right) \mathbf{e}_\perp = \beta^2 \mathbf{e}_\perp \, , \tag{3.25}$$

where $\hat{O}_0 = \nabla_\perp^2 + k_0^2$, $\hat{O}_1 = \left(\varepsilon_r(\mathbf{r}_\perp) - 1 \right) k_0^2$ and $\hat{O}_2 \mathbf{e}_\perp = \nabla_\perp \left(\mathbf{e}_\perp \cdot \nabla \ln \varepsilon_r(\mathbf{r}_\perp) \right)$. The advantage of using basis modes, that are solutions to the vacuum geometry, now becomes apparent. When evaluating the operator element

$$O_{m,n} = \left\langle \mathbf{b}_{\perp;m} \middle| \hat{O}_0 \mathbf{b}_{\perp;n} \right\rangle + \left\langle \mathbf{b}_{\perp;m} \middle| \hat{O}_1 \mathbf{b}_{\perp;n} \right\rangle + \left\langle \mathbf{b}_{\perp;m} \middle| \hat{O}_2 \mathbf{b}_{\perp;n} \right\rangle \equiv O_{0;m,n} + O_{1;m,n} + O_{2;m,n} \, , \tag{3.26}$$

the basis modes are eigenmodes of $\hat{O}_0$. The elements $O_{0;m,n}$ are thus given directly by $\left(k_{m,\perp}^2 + k_0^2 \right) \delta_{mn}$, where $k_{m,\perp}^2 = a_m^2 + b_m^2$ when using the basis modes in Cartesian coordinates described in appendix 8.1.1, and $k_{m,\perp}^2 = a_m^2$ when using those in cylindrical coordinates given in appendix 8.1.2. This is particularly useful in cylindrical coordinates, where the evaluation of nasty integrals involving $1/r^2$ is completely avoided. All that remains to be evaluated are then the elements $O_{1;m,n}$ and $O_{2;m,n}$.

In Cartesian coordinates, the calculation of the element $O_{1;m,n}$ for a piece-wise constant refractive index profile is straightforward. The integrals to be evaluated are of the general form $\int \sin(a_m x) \sin(a_n x) dx$, the analytic primitive function of this integral is known, and the operator elements are thus quickly determined.

When studying a piece-wise constant index profile in cylindrically symmetric coordinates, one runs into trouble when trying to evaluate the $O_{1;m,n}$ elements directly:

This involves determining the indefinite integral $\int \frac{\partial}{\partial r} J_l(a_m r) \frac{\partial}{\partial r} J_l(a_n r) r dr$, where J_l is the Bessel function of the first kind of order l, and no analytic primitive function for this integral is known. A useful trick, however, is to employ the transformations $f_+ \equiv (f_r + f_\theta / i)/\sqrt{2}$ and $f_- \equiv (f_r - f_\theta / i)/\sqrt{2}$ and the mode profiles in these "sum and difference" coordinates are given in appendix 8.1.2. When using the transformed coordinates, the operator matrix $\overline{\overline{O}}_1$ representing the operator $\hat{O}_1$ remains unchanged:

$$\begin{bmatrix} b_{r;m}^* & b_{\theta;m}^* \end{bmatrix} \begin{bmatrix} (\varepsilon_r - 1)k_0^2 & 0 \\ 0 & (\varepsilon_r - 1)k_0^2 \end{bmatrix} \begin{bmatrix} b_{r;n} \\ b_{\theta;n} \end{bmatrix} = \begin{bmatrix} b_{+;m}^* & b_{-;m}^* \end{bmatrix} \begin{bmatrix} (\varepsilon_r - 1)k_0^2 & 0 \\ 0 & (\varepsilon_r - 1)k_0^2 \end{bmatrix} \begin{bmatrix} b_{+;n} \\ b_{-;n} \end{bmatrix}.$$

However, using the expression on the right side simplifies the evaluation of the element $O_{1;m,n}$, as this only involves an integral of the general form $\int J_l(a_m r) J_l(a_n r) r dr$, and an analytic primitive function to this integral does exist.

The calculation of the element $O_{2;m,n}$ for piece-wise constant refractive index profiles in both coordinate sets is a bit tricky, as it involves the derivative of the dielectric constant, that may be discontinuous along a curve C. The procedure is to write the element as:

$$O_{2;m,n} = \int \mathbf{b}_{\perp;m}^* \cdot \nabla_\perp (\mathbf{b}_{n,\perp} \cdot \nabla \ln \varepsilon_r) d\mathbf{r}_\perp = -\int (\nabla_\perp \cdot \mathbf{b}_{\perp;m}^*)(\mathbf{b}_{n,\perp} \cdot \nabla \ln \varepsilon_r) d\mathbf{r}_\perp. \quad (3.27)$$

The contribution from a discontinuity of ε_r is then given by the curve integral

$$-\int_C (\nabla_\perp \cdot \mathbf{b}_{\perp;m}^*) \Delta \ln \varepsilon_r |\mathbf{b}_{n,\perp} \cdot d\mathbf{s}.$$

Here, $\Delta \ln \varepsilon_r$ is the discontinuity of $\ln \varepsilon_r$, and $d\mathbf{s}$ is a line element perpendicular to the curve in the direction of increasing ε_r.

So far, only the evaluation of operator matrix elements for piece-wise constant refractive index profiles is described. However, in the final stages of the Ph.D. work, graded index profiles of the form $\varepsilon_r(x) = ax + b$ in Cartesian coordinates were considered. The element $O_{1;m,n}$ is evaluated by calculating integrals of the general form $\int \sin(a_m x) \sin(a_n x) x dx$, which has an analytic primitive function that can be evaluated directly. The integral in the determination of the element $O_{2;m,n}$ is of the

form $\int \sin(a_m x)\sin(a_n x)/(x+c)dx$, and this integral also has a known antiderivative, however, the antiderivative includes the exponential integral function

$$\mathrm{expint}(x) = \int_x^\infty \frac{\exp(-t)}{t}dt\,,$$

and unfortunately, the expint function in Matlab is evaluated only using an iterative formula, making the calculation of the $\overline{\overline{O_2}}$ matrix somewhat slow.

The computation time of the eigenmode problem for a given number of basis modes is independent on the refractive index profile, however, the time necessary to calculate the operator matrix $\overline{\overline{O}}$ depends linearly on the number of regions of the index profile. As in the FD technique the quality of the mode profiles depends on the number N of eigenmodes used in a computation, and as previously we have the option of increasing the quality of the mode profiles without changing N by increasing the number of basis modes. In this work, $2N$ basis modes were used in the representation of an eigenmode.
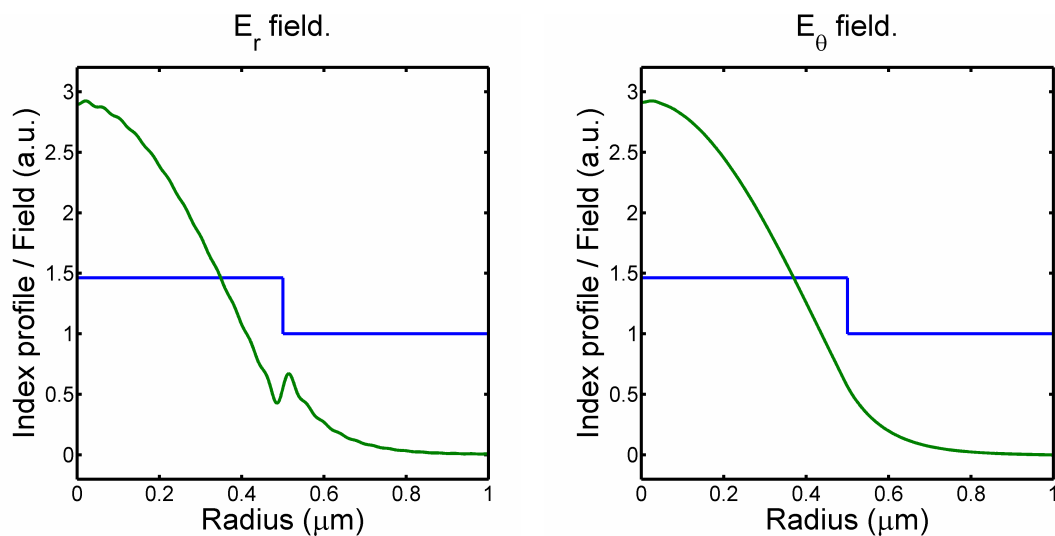
Fig. 3.11: The refractive index profile (blue) together with the in-plane components (green) of a cylindrically symmetric mode profile obtained using plane-wave expansion are depicted.

The in-plane field components of the fundamental mode profile obtained by solving the eigenvalue problem for the rotationally symmetric geometry of Fig. 3.7 using the plane-wave expansion method are illustrated in Fig. 3.11.

Comparing the profile of the radial component with that depicted in Fig. 3.7, we observe that the jump at the interface is less pronounced. The difficulty in describing

discontinuities in the field components is a general weakness of the plane-wave expansion technique. The problem is related to properties of Fourier transformation: The Fourier transform (FT) of a discontinuous function has very broad support and similarly, our discontinuous e_r component requires a large number of plane waves to be described correctly. As the computer can only handle a limited number of basis modes, the fine details of the sharp jump are lost.

3.4.3 The semi-analytical approach

When using the semi-analytical (SA) method there is, unlike the techniques described in the previous two sections, no expansion of the field on any grid or basis set, and no eigenvalue problem is solved in this technique. Rather, the field profiles are described by exact analytical solutions to the wave equation. A mode-matching is performed at the interfaces between rings of different dielectric constant, and this leads to a numerical problem of determining propagation constants of the modes. It has been shown that analytic solutions do not exist in rectangular waveguides, and for this reason the technique can only be used for rectangular geometries uniform along one axis or for the rotationally symmetric geometries described in the following.

As explained in detail in appendix 8.1.2, analytic solutions to the wave equation in a region of constant refractive index are available. Separating the angular and the radial dependence, the solutions are written as

$$\begin{bmatrix} e_z(\mathbf{r}_\perp) \\ h_z(\mathbf{r}_\perp) \end{bmatrix} = \exp(il\theta) \begin{bmatrix} e_z(r) \\ h_z(r) \end{bmatrix}, \tag{3.28}$$

and the modes are either TE or TM. Their radial profile is given by Bessel functions and depend on the sign of $k^2 - \beta^2$, where $k^2 = \varepsilon_r k_0^2$. We now introduce the notation $\mathbf{f} = (\mathbf{e}, \mathbf{h})$ for the complete optical field profile. When $k^2 - \beta^2 \equiv h^2 > 0$, the solutions $\mathbf{f}_{J;TE}$, $\mathbf{f}_{J;TM}$, $\mathbf{f}_{Y;TE}$ and $\mathbf{f}_{Y;TM}$ are allowed, and these solutions represent standing oscillating waves. When $k^2 - \beta^2 \equiv -q^2 < 0$, the solutions $\mathbf{f}_{I;TE}$, $\mathbf{f}_{I;TM}$, $\mathbf{f}_{K;TE}$ and $\mathbf{f}_{K;TM}$ are valid, and these modes are exponentially increasing or decaying. The radial dependence of z components of these functions is given explicitly in the following table:

Field component	$\mathbf{f}_{J;TE}$	$\mathbf{f}_{J;TM}$	$\mathbf{f}_{Y;TE}$	$\mathbf{f}_{Y;TM}$
e_z	0	$J_l(hr)$	0	$Y_l(hr)$
h_z	$J_l(hr)$	0	$Y_l(hr)$	0
	$\mathbf{f}_{I;TE}$	$\mathbf{f}_{I;TM}$	$\mathbf{f}_{K;TE}$	$\mathbf{f}_{K;TM}$
e_z	0	$I_l(qr)$	0	$K_l(qr)$
h_z	$I_l(qr)$	0	$K_l(qr)$	0

Table 3.1: The radial dependence of the z components of the eight possible solutions in a region of constant ε_r.

The solutions only depend on the quantities h and q which are given by k_0, the dielectric constant ε_r and the propagation constant β. The complete six component mode profile for TE and TM modes for the $\mathbf{f}_J$ functions are given in appendix 8.1.2. We now consider the rotationally symmetric refractive index profile sketched in the following figure:
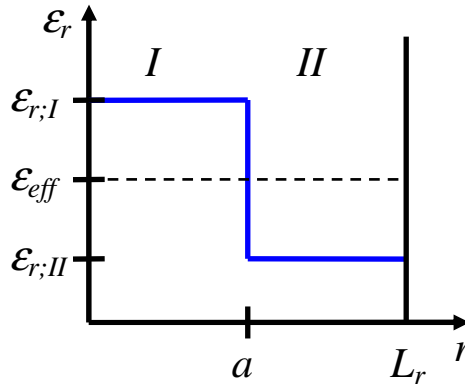
Fig. 3.12: Illustration of a geometry consisting of two rings I and II.

The geometry consists of two regions I and II, each featuring constant dielectric constants $\varepsilon_{r;I}$ and $\varepsilon_{r;II}$. We introduce $k_I^2 \equiv \varepsilon_{r;I}k_0^2$, $k_{II}^2 \equiv \varepsilon_{r;II}k_0^2$ and $\varepsilon_{eff} \equiv \beta^2/k_0^2$ and look for mode confined to the region I, which means that its propagation constant should be such that $\varepsilon_{r;I} > \varepsilon_{eff} > \varepsilon_{r;II}$. In region I we then have $k_I^2 - \beta^2 \equiv h_I^2 > 0$, and the mode profile $\mathbf{f}_I$ valid for $r \leq a$ is given by the linear combination:

$$\mathbf{f}_I = c_{I,1}\mathbf{f}_{J;TE} + c_{I,2}\mathbf{f}_{J;TM} + c_{I,3}\mathbf{f}_{Y;TE} + c_{I,4}\mathbf{f}_{Y;TM}. \tag{3.29}$$

In region II, $k_{II}^2 - \beta^2 \equiv -q_{II}^2 < 0$ and the mode profile $\mathbf{f}_{II}$ valid when $r \geq a$ is given by:

$$\mathbf{f}_{II} = c_{II,1}\mathbf{f}_{I;TE} + c_{II,2}\mathbf{f}_{I;TM} + c_{II,3}\mathbf{f}_{K;TE} + c_{II,4}\mathbf{f}_{K;TM}. \tag{3.30}$$

We now require that the mode profile respects the boundary conditions at $r = 0$, $r = a$ and $r = L_r$. As the Bessel function $Y_l(r)$ diverges when $r \rightarrow 0$, this sets the requirement $c_{I,3} = c_{I,4} = 0$. At $r = a$, we require that the tangential components of electric and magnetic fields are continuous, that is $e_{\theta;I}(a) = e_{\theta;II}(a)$, $e_{z;I}(a) = e_{z;II}(a)$, $h_{\theta;I}(a) = h_{\theta;II}(a)$ and, $h_{z;I}(a) = h_{z;II}(a)$. And finally, at $r = L_r$ the conditions are $e_{\theta;II}(L_r) = e_{z;II}(L_r) = 0$. The boundary conditions can be described by the matrix relation $\overline{\overline{B}}\,\overline{c} = \overline{0}$, where $\overline{c} = \left[c_{I,1}; c_{I,2}; c_{II,1}; c_{II,2}; c_{II,3}; c_{II,4}\right]$ and

$$
\overline{\overline{B}} =
\begin{bmatrix}
e_{J;TE;\theta}(h_I,a) & e_{J;TM;\theta}(h_I,a) & -e_{I;TE;\theta}(q_{II},a) & -e_{I;TM;\theta}(q_{II},a) & -e_{K;TE;\theta}(q_{II},a) & -e_{K;TM;\theta}(q_{II},a) \\
e_{J;TE;z}(h_I,a) & e_{J;TM;z}(h_I,a) & -e_{I;TE;z}(q_{II},a) & -e_{I;TM;z}(q_{II},a) & -e_{K;TE;z}(q_{II},a) & -e_{K;TM;z}(q_{II},a) \\
h_{J;TE;\theta}(h_I,a) & h_{J;TM;\theta}(h_I,a) & -h_{I;TE;\theta}(q_{II},a) & -h_{I;TM;\theta}(q_{II},a) & -h_{K;TE;\theta}(q_{II},a) & -h_{K;TM;\theta}(q_{II},a) \\
h_{J;TE;z}(h_I,a) & h_{J;TM;z}(h_I,a) & -h_{I;TE;z}(q_{II},a) & -h_{I;TM;z}(q_{II},a) & -h_{K;TE;z}(q_{II},a) & -h_{K;TM;z}(q_{II},a) \\
0 & 0 & e_{I;TE;\theta}(q_{II},L_r) & e_{I;TM;\theta}(q_{II},L_r) & e_{K;TE;\theta}(q_{II},L_r) & e_{K;TM;\theta}(q_{II},L_r) \\
0 & 0 & e_{I;TE;z}(q_{II},L_r) & e_{I;TM;z}(q_{II},L_r) & e_{K;TE;z}(q_{II},L_r) & e_{K;TM;z}(q_{II},L_r)
\end{bmatrix}.
$$

The boundary conditions can thus be expressed as a set of linear equations. The elements of $\overline{\overline{B}}$ depend on the quantities h_I and q_{II}, which in turn are given by the propagation constant β. The remaining quantities are all fixed.

The linear system $\overline{\overline{B}}\,\overline{c} = \overline{0}$ allows a non-trivial solution $\overline{c}$ when $\det \overline{\overline{B}} = 0$. We thus consider the function $\det \overline{\overline{B(\beta)}}$ and search for the zero points of this function. For each zero point a solution $\overline{c}$ exists with a corresponding propagation constant β, that describes an optical mode through Eqs. (3.29) and (3.30).

An example of a mode profile obtained using the semi-analytical approach is given in Fig. 3.13 below. The geometry is identical to that depicted in Fig. 3.7 and Fig. 3.11. Inspecting the figure, we observe the perfect shape of the mode profile. Unlike the solution obtained using the FD technique, the mode here is known in an arbitrary point in space, and unlike the mode profile given by PW technique, the discontinuity of the radial component is described exactly.
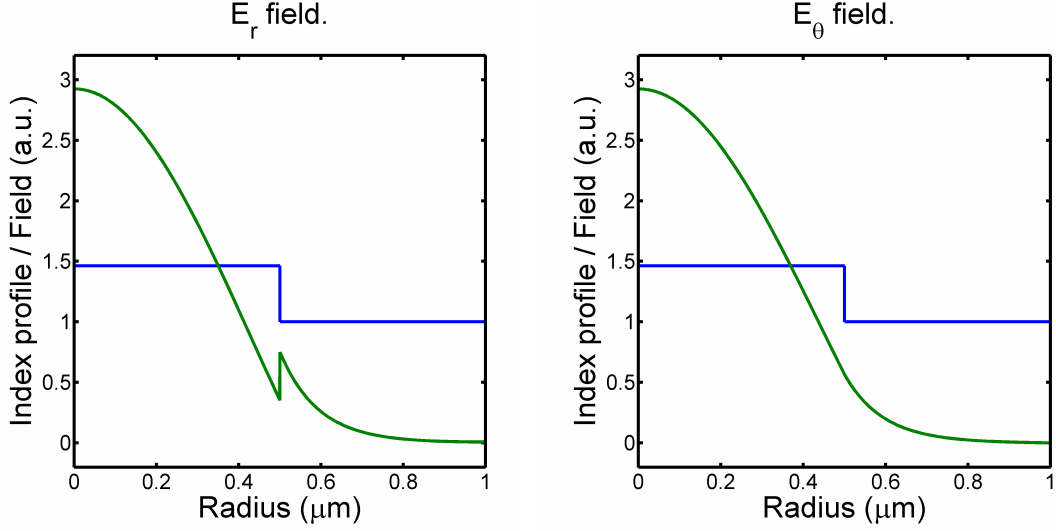
Fig. 3.13: The refractive index profile (blue) together with the in-plane components (green) of a cylindrically symmetric mode profile obtained using the SA approach are depicted.

Also, as the field above is described using only six variables (plus the propagation constant), very little memory is necessary to store a mode profile obtained using the SA approach.

In the example, we have considered a two-region refractive index geometry, and we have looked for a mode confined to region *I*. The procedure described above, however, is directly extended to handle an arbitrary piece-wise constant refractive index profile with an arbitrary number of regions, and any mode, localized or not, can be described. From a numerical point of view, every new region adds four extra coefficients to the vector $\bar{c}$, and the size of the matrix $\bar{\bar{B}}$ increases correspondingly. It is thus more computationally demanding to evaluate $\det \overline{\overline{B(\beta)}}$ in an advanced geometry with many regions compared to a simple two-region geometry.

Even though modes are completely determined once the zero points of the function $\det \overline{\overline{B(\beta)}}$ are known, the challenge of finding the zero points remains. The problem is that of determining zero points of an, in principle, completely unknown function.

In a lossless geometry with real-valued index profile, we usually expect the squares of the propagation constants β^2 to lie on the real axis as it is the case for the propagating and evanescent modes depicted in Fig. 3.4. A cut-off $\varepsilon_{r;\max} k_0^2$ exists where $\varepsilon_{r;\max}$ is the maximum value of the dielectric constant of the geometry. The zero-points β^2 will lie on the real half-axis given by $\beta^2 \leq \varepsilon_{r;\max} k_0^2$ and searching this half-axis for

40

zero-points can be done using a standard search routine. The oscillations of the function $\det \overline{\overline{B(\beta)}}$ around 0 with corresponding zero-points are illustrated in Fig. 3.14(a). As modes are determined simply by following the real axis, the computation time for finding the zero points increases only linearly with the number of modes.
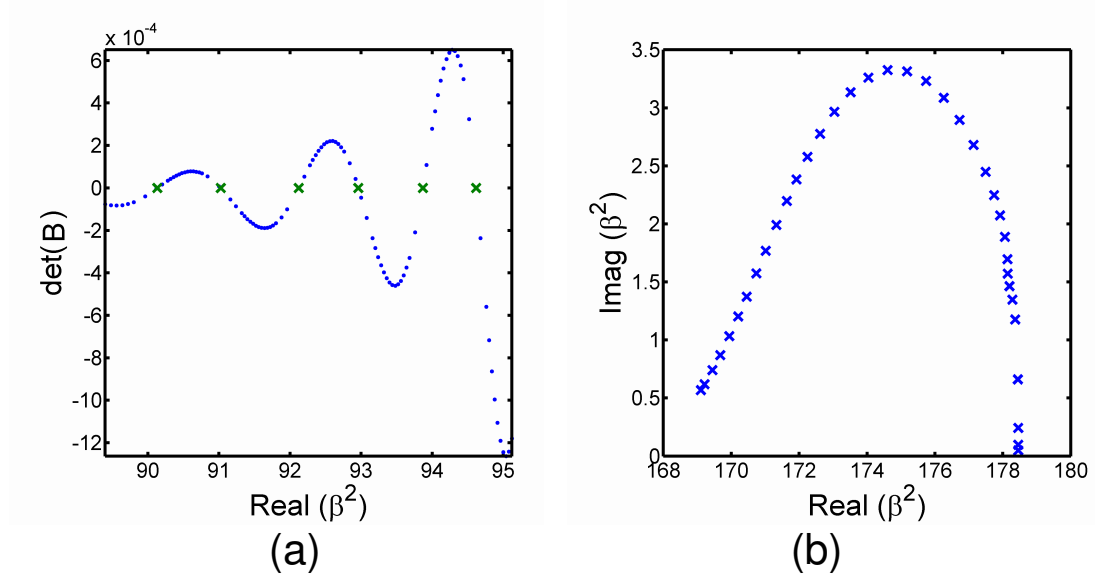
Fig. 3.14: In (a), the oscillations of the function det(B) are depicted (blue) with marked zero-points (green). The path of a zero-point in complex space when the imaginary part of the refractive index profile is increased is illustrated in (b).

In a geometry with loss, the squares of the propagation constants β^2 of the solutions no longer lie on the real axis, but are found in the complex plane. $\det \overline{\overline{B(\beta)}}$ is an analytic function, and zero points in the complex plane of an analytic function can be determined using a technique based on contour integration[30]. However, when a large number of zero-points has to be determined, this technique becomes very cumbersome. For this reason, the perturbative technique suggested in Ref. 23 has been used in this Ph.D. work instead: The method consists of initially determining the zero-points of the geometry, where the imaginary part of the refractive index profile has been set to zero, and this can be done using the technique determined above. We then increase the imaginary part of the index profile and calculate the value of the function $\det \overline{\overline{B(\beta)}}$ in the vicinity of our original zero-points. If the imaginary part is increased gently, the zero-points only move a little and we can follow the zero-points as they move around in the complex plane. The path of a zero-point as the imaginary part of a complex refractive index profile is increased is illustrated in Fig. 3.14(b).

41

The computation time necessary for determining complex propagation constants depends directly on the magnitude of the imaginary part of the index profile. When considering weak losses in dielectric materials, the imaginary part is small and the computation is fairly fast. However, when introducing metals with large imaginary parts into the geometry, the path of the zero-points becomes correspondingly longer and the computation time increases.

In the technique presented above to determine propagation constants, two assumptions are made: In the case of a real-valued index profile, all β^2 zero-points should lie on the real axis. And in the case of a complex index profile, when the imaginary part of the refractive index is increased, no new zero-points, unrelated to the original points determined on the real axis, should appear. Unfortunately, at least the first assumption does not generally hold: As explained in Section 3.3.2, a complex mode having its β^2 in the complex plane can exist even in the case of a real-valued geometry. These modes, which are necessary to form the complete basis set, will not be found by the search routine outlined above.

(a)   (b)
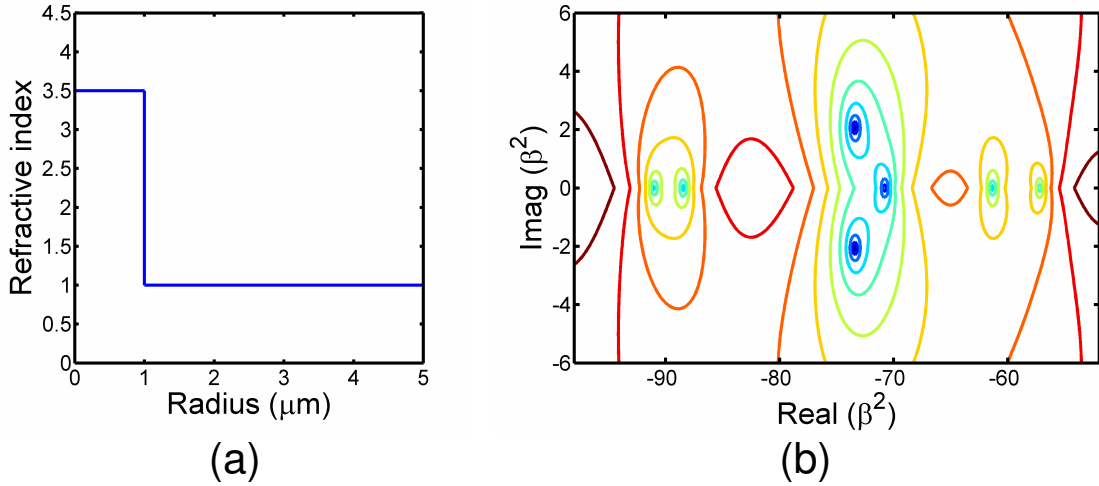
Fig. 3.15: **The refractive index profile (a) and a contour plot of ln(|det(B)|) with seven zero-points (b) are shown.**

In Fig. 3.15, an example of the occurrence of complex modes is illustrated. We consider modes of the real-valued index profile depicted in Fig. 3.15(a) for the wavelength $\lambda = 950$ nm and the angular momentum l equal to 8. The function $\det \overline{\overline{B(\beta)}}$ is mapped in the complex plane in Fig. 3.15(b). We observe a pair of zero-

points at $\beta^2 \approx -73.5 \pm 2.1i$, corresponding to complex modes. The occurrence of complex modes is described as "rare" in Ref. 24, however, in this work they were systematically observed in waveguide geometries when the index of the waveguide and/or the angular momentum are increased.

3.5 Scattering matrix formalism

In the previous sections, we have determined eigenmodes in geometries featuring uniformity along the z axis. We will now describe the mode-matching procedure applied in geometries consisting of two or more layers of different refractive index profiles. Two different procedures, the transfer matrix formalism and the scattering matrix formalism, exist, which can handle the mode-matching. The transfer matrix formalism has an advantage when handling stacks of layers with periodicity, but unfortunately it also suffers from numerical instability[22,24]. For this reason the scattering matrix formalism was employed in this work.

3.5.1 Reflection and transmission at an interface

Let us consider the two-layer geometry illustrated in Fig. 3.16(a).
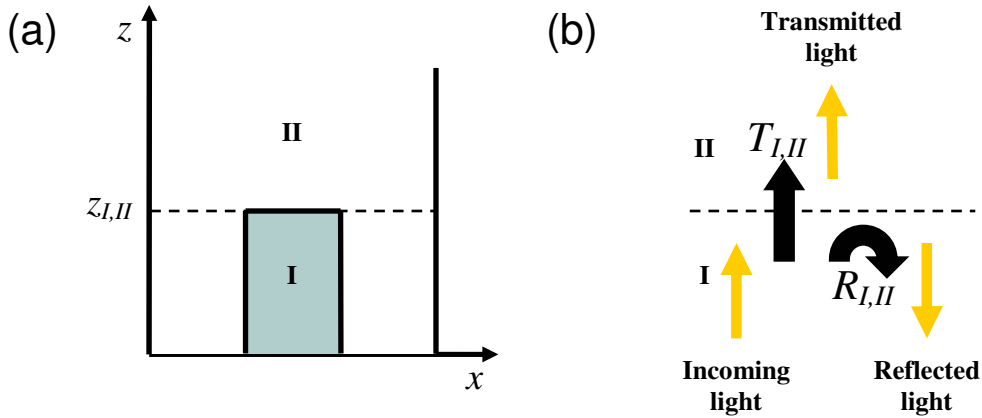

Fig. 3.16: A two-layer geometry (a) and the reflection and transmission of light at the zone interface (b) are illustrated.

In the two layers, the field is expanded on the eigenmodes of the corresponding zone as in Eq. (3.6). We now consider the situation, in which an incoming mode of index m is propagating forward in layer I, sketched in Fig. 3.16(a). It reaches the layer

interface and is partly reflected giving rise to a backward-propagating field in layer I. The incoming mode, however, is also partly transmitted, and this results in a forward propagating field in layer II. We then write the electric field in the two zones as follows:

$$\mathbf{E}_{I;\perp}(\mathbf{r}) = a_{I;m}^{+}{}' \mathbf{e}_{I;\perp;m}(\mathbf{r}_{\perp})e^{i\beta_{I;m}(z-z_{I,II})} + \sum_{n} a_{I;n}^{-} \mathbf{e}_{I;\perp;n}(\mathbf{r}_{\perp})e^{-i\beta_{I;n}(z-z_{I,II})} \qquad (3.31)$$

$$\mathbf{E}_{II;\perp}(\mathbf{r}) = \sum_{n} a_{II;n}^{+} \mathbf{e}_{II;\perp;n}(\mathbf{r}_{\perp})e^{i\beta_{II;n}(z-z_{I,II})} \, , \qquad (3.32)$$

where the expansion coefficients are defined relative to the position $z = z_{I,II}$ of the interface for convenience. The electric field is then given by $\mathbf{E}_{\perp}(\mathbf{r}) = \mathbf{E}_{I;\perp}(\mathbf{r})$ when $z \leq z_{I,II}$ and $\mathbf{E}_{\perp}(\mathbf{r}) = \mathbf{E}_{II;\perp}(\mathbf{r})$ when $z \geq z_{I,II}$. We have fixed $a_{I;m}^{+}{}'$, however, the coefficients $a_{I;n}^{-}$ and $a_{II;n}^{+}$ are yet unknown. We now introduce the reflection and transmission matrices $\overline{\overline{R_{I,II}}}$ and $\overline{\overline{T_{I;II}}}$ defined by $a_{I;n}^{-} = R_{I,II;nm}a_{I;m}^{+}{}'$ and $a_{II;n}^{+} = T_{I,II;nm}a_{I;m}^{+}{}'$. The coefficients of the matrices $\overline{\overline{R_{I,II}}}$ and $\overline{\overline{T_{I;II}}}$ are obtained by requiring that the boundary conditions are fulfilled at the interface between layers I and II, meaning that the tangential components of the electric and magnetic field should be continuous. Referring to appendix 8.2 for the details of the derivation, the defining expressions are given by:

$$\overline{\overline{T_{I,II}}} = 2\left(\overline{\overline{\langle \mathbf{e}_{II} | \mathbf{h}_{I}\rangle}}^{T} + \overline{\overline{\langle \mathbf{e}_{I} | \mathbf{h}_{II}\rangle}}\right)^{-1} \qquad \overline{\overline{R_{I,II}}} = \frac{1}{2}\left(\overline{\overline{\langle \mathbf{e}_{II} | \mathbf{h}_{I}\rangle}}^{T} - \overline{\overline{\langle \mathbf{e}_{I} | \mathbf{h}_{II}\rangle}}\right)\overline{\overline{T_{I,II}}} \, ,$$

where the inner product used above is $\langle \mathbf{e}_{m} | \mathbf{h}_{n}\rangle = \int(\mathbf{e}_{m} \times \mathbf{h}_{n}) \cdot \mathbf{u}_{z} d\mathbf{r}_{\perp}$.

When the reflection and transmission matrices are known, the coefficients $a_{I;n}^{-}$ and $a_{II;n}^{+}$ can be determined, and the field is well-defined by Eqs. (3.31) and (3.32). In Fig. 3.16, the structure is illuminated from below. However, we could equally well illuminate the geometry from above, and in this case we would use the matrices $\overline{\overline{R_{II,I}}}$ and $\overline{\overline{T_{II,I}}}$ to describe the field.

3.5.2 Scattering matrices

We can describe a two-layer geometry using the reflection and transmission matrices introduced above. Now, let us consider the three-layer geometry illustrated in Fig. 3.17(a).
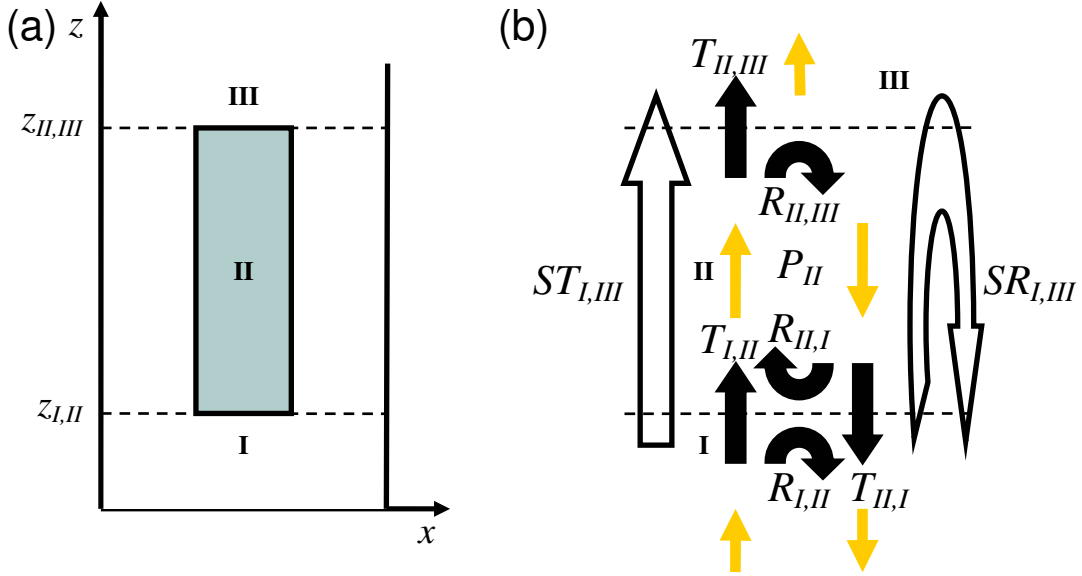

Fig. 3.17: A three-layer geometry is depicted in (a). The reflection and transmission of light at the two interfaces are illustrated in (b).

As previously, we illuminate the structure from below with light described by the eigenmode m. The field is given by:

$$\mathbf{E}_{I;\perp}(\mathbf{r}) = a^{+}_{I;m}{}'\mathbf{e}_{I;\perp;m}(\mathbf{r}_{\perp})e^{i\beta_{I;m}(z-z_{I,II})} + \sum_{n} a^{-}_{I;n}\mathbf{e}_{I;\perp;n}(\mathbf{r}_{\perp})e^{-i\beta_{I;n}(z-z_{I,II})} \qquad (3.33)$$

$$\mathbf{E}_{II;\perp}(\mathbf{r}) = \sum_{n} a^{+}_{II;n}\mathbf{e}_{II;\perp;n}(\mathbf{r}_{\perp})e^{i\beta_{II;n}(z-z_{I,II})} + \sum_{n} a^{-}_{II;n}\mathbf{e}_{II;\perp;n}(\mathbf{r}_{\perp})e^{-i\beta_{II;n}(z-z_{II,III})} \qquad (3.34)$$

$$\mathbf{E}_{III;\perp}(\mathbf{r}) = \sum_{n} a^{+}_{III;n}\mathbf{e}_{III;\perp;n}(\mathbf{r}_{\perp})e^{i\beta_{III;n}(z-z_{II,III})} , \qquad (3.35)$$

For the layer II we define coefficients $a^{\pm}_{II}{}'$ by $a^{+}_{II;m}{}' = P_{II,m}a^{+}_{I;m}$ and $a^{-}_{II;m}{}' = P_{II,m}a^{-}_{II;m}$, where $P_{II,m} = \exp(i\beta_{m}L_{II})$ and $L_{II} = z_{II,III} - z_{I,II}$. The coefficients $a^{\pm}$ can be thought of as "having their back against the interface", where the coefficients $a^{\pm}{}'$ are "facing the interface".

Referring to Fig. 3.17(b), we now consider the reflection of light by the three-layer geometry. The light of mode m travels forward in layer I. At the interface it is partly

reflected back into layer I with reflection described by the matrix $\overline{\overline{R_{I,II}}}$. However, part of the light is also transmitted into layer II. It propagates towards the interface between zone II and II and is partly reflected. It then travels backwards towards the first interface and is partly transmitted at the interface into layer I. This journey can be described by the matrix $\overline{\overline{T_{II,I}}}\,\overline{\overline{P_{II}}}\,\overline{\overline{R_{II,III}}}\,\overline{\overline{P_{II}}}\,\overline{\overline{T_{I,II}}}$, where $\overline{\overline{P_{II}}}$ is the propagation matrix; a diagonal matrix consisting of the elements $P_{II,m}$. This journey consisting of one round-trip inside layer II adds an additional reflection to that described by $\overline{\overline{R_{I,II}}}$. There is, however, also contributions from light making two, three etc. round-trips in layer II. The sum of these contributions is given by the scattering reflection matrix $\overline{\overline{SR_{I,III}}}$ describing reflection at the interface between layer I and II with multiple reflections inside layer II included:

$$SR_{I,III} = R_{I,II} + T_{II,I}P_{II}R_{II,III}P_{II}\left(1 - R_{II,I}P_{II}R_{II,III}P_{II}\right)^{-1}T_{I,II}, \qquad (3.36)$$

where the double bars have been skipped for simplicity. The details of this derivation are found in appendix 8.3. In a similar manner, one can derive the scattering transmission matrix $\overline{\overline{ST_{I,III}}}$, which describes transmission of light from layer I to layer III, again considering the multiple reflections in layer II. The relations $a_{I;n}^{-} = SR_{I,III;nm}a_{I;m}^{+}{}'$ and $a_{III;n}^{+} = ST_{I,III;nm}a_{I;m}^{+}{}'$ hold, and the fields in layer I and III are thus well-determined. The expansion coefficients in layer II are also readily determined using expressions given in the appendix.

When scattering reflection and transmission matrices can be determined in a three-layer structure, we can determine the field in a structure of a given number of layers recursively. The recursive procedure can be understood by inspecting the four-layer geometry of Fig. 3.18. We first calculate the scattering matrices $\overline{\overline{SR_{I,III}}}$, $\overline{\overline{ST_{I,III}}}$, $\overline{\overline{SR_{III,I}}}$ and $\overline{\overline{ST_{III,I}}}$ using the procedure outlined above. We can then treat the four-layer structure as a three-layer geometry, where reflection and transmission between layer I and III are described using the scattering matrices listed above.
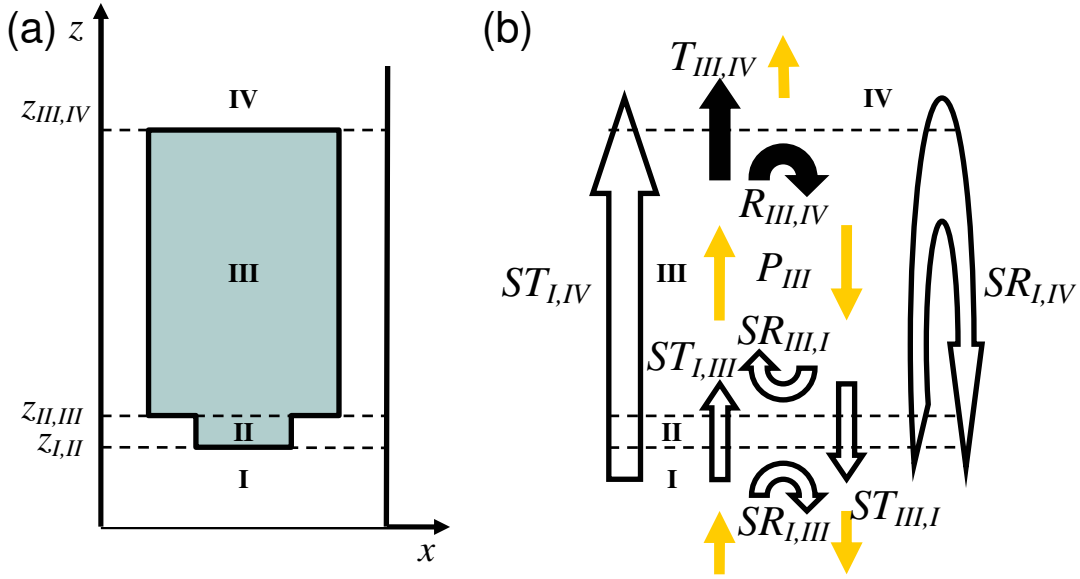
Fig. 3.18: A four-layer geometry is depicted in (a). The reflection and transmission of light are illustrated in (b).

The scattering reflection matrix $\overline{\overline{SR_{I,IV}}}$ is then given by an expression similar to Eq. (3.36) where the matrices $\overline{\overline{R_{I,II}}}$, $\overline{\overline{T_{I,II}}}$ etc. are simply replaced by $\overline{\overline{SR_{I,III}}}$, $\overline{\overline{ST_{I,III}}}$ etc. The complete expressions for all scatterings matrices are given in appendix 8.3.

Using the scattering matrix formalism, we can illuminate the structure from below or from above (or both simultaneously), and from these boundary conditions we can then completely determine the optical field in the entire structure.

3.6 Discussion

The heart of the EET is the computation of eigenmodes. Using the semi-analytical approach, exact mode profiles and propagation constants are obtained and the high quality of the eigenmodes leads to fast convergence[e] when the number of included modes is increased. The procedure to determine propagation constants in the semi-analytical approach is not perfect, however. Propagation constants appear in the complex plane even when real-valued refractive index profiles are considered, and a fast and reliable method of identifying these is currently lacking.

The finite-difference and the plane-wave expansion techniques systematically find propagation constants, but they do so at the expense of the quality of the eigenmodes

---

[e] The convergence of the EET is further addressed in section 5.2.

and with increased computation time. A future challenge for EET-users is thus the development of an efficient method of identifying propagation constants in the complex plane. When this is accomplished, the EET will represent an extremely attractive candidate in the simulation of light scattering on 2D or 3D cylindrically symmetric microscopic structures.

4 Topography characterization using near-field imaging

When studying a microscopic object using conventional far-field optical microscopy, the achievable resolution is limited to approximately half of the wavelength. We saw in chapter 1 that subwavelength information is also present in the field scattered by an object, but it is contained in the evanescent part of the field and does not reach a remote observer. The SNOM, however, allows access to the near field scattered by the object and thereby acquirement of information about its evanescent components before these decay beyond the noise floor. With this information, spatial resolutions beyond the diffraction limit can be achieved, and SNOM thus offers a non-destructive way of characterizing microscopic structures with subwavelength features.

Even though detection of evanescent field components is now possible, a connection between the detected field and the geometry of a sample studied is far from trivial. Well-established techniques exist to model the field produced by scattering on a small object, but determination of the topography from the field measured, the inverse scattering problem (ISP), remains a challenge. In initial papers[31,32], a surface profile reconstruction method based on the Rayleigh hypothesis was demonstrated. Inversion schemes for 3D structures under the validity of the Born approximation have also been proposed[33,34], but a general inversion procedure for arbitrary structures has not yet been identified.

Recently, a method based on an optimization procedure[35] was suggested and numerically tested. In this chapter, we employ this procedure to characterize topography of a deep grating from SNOM images. Such a structure cannot be described using the Rayleigh hypothesis or the Born approximation and is therefore interesting for examination of the advantages and limitations of the procedure developed.

4.1 Experimental SNOM measurements

The sample used in this work is a transmission diffraction grating of fused silica glass, that has been fabricated by Ibsen Photonics. The grating is linear and uniform with a constant period of 1013.6 nm. The grating fabrication involved holographic exposure

of a photoresist layer deposited on the top surface of a glass substrate. Hereafter the grating pattern was etched by reactive ion etching (RIE) into the surface to a depth of approximately 2.0 µm.
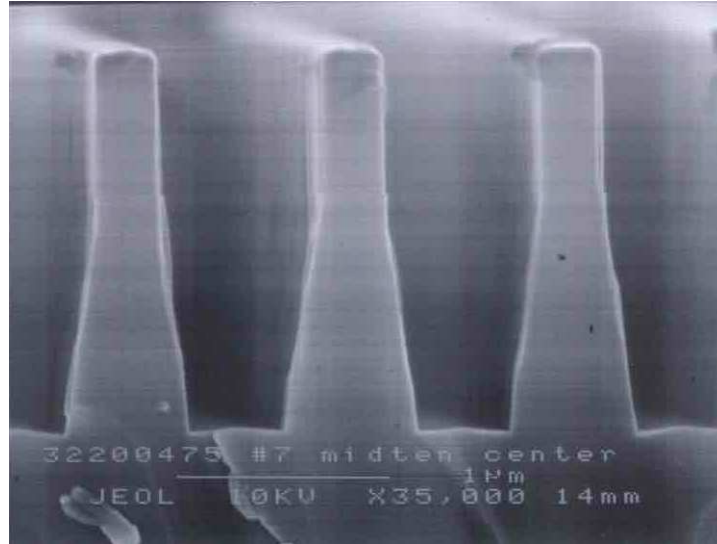

Fig. 4.1: A SEM image of the grating profile is shown.

The sidewall profile, illustrated in Fig. 4.1, was intended to be nearly vertical. After the RIE, the residual photoresist was removed, and the grating was cleaned in acid solutions.

The experimental setup consists of a collection SNOM with an uncoated fiber tip and an arrangement for sample illumination under the condition of total internal reflection. The setup, which is rather similar to the setup used in experiments reported in Ref. 36, is operating as a PSTM. In Fig. 4.2, the setup is shown schematically. A fiber tip of a stand-alone SNOM[37] was scanned close to the illuminated grating, collecting a tiny part of diffracted optical field. The light source was a HeNe laser with wavelength $\lambda \approx 633$ nm and power output of ~ 5 mW and the illumination beam was *s*-polarized with electric field vector directed along the *y* axis. The beam was weakly focused with focal length of ~ 500 mm and spot size of ~ 300 µm. The grating was attached with immersion oil to the base of a glass prism. The light intensity was modulated by an optical chopper with modulation frequency ~ 350 Hz in order to enable its synchronous detection. The intensity distribution near the grating surface was probed with an uncoated sharpened tip of a single-mode optical fiber fabricated by chemical etching in 40 % hydrofluoric acid.[38]
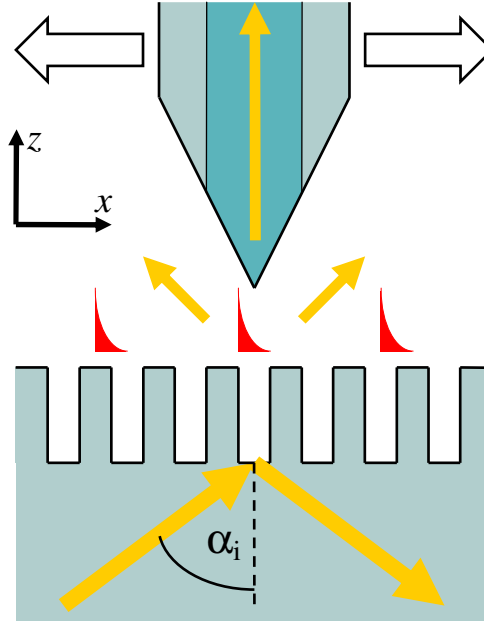
Fig. 4.2: The main part of the experimental SNOM setup.

The resulting fiber tip has a cone-angle of ~ 24° and curvature radius of less than 80 nm. The other end of the fiber is optically coupled to a photo-detector, whose signal at the modulation frequency is synchronously detected by using a lock-in amplifier. The fiber tip was scanned over the sample surface in constant height mode by using the shear-force feedback.[39] The experimental measurements were performed by Valentyn S. Volkov at Aalborg University.

Now, the angle of beam incidence $\alpha_i \approx 44.5°$ was slightly above the angle of total internal reflection, but the grating diffracted and transmitted light up to the third order. Even though the SNOM was operating in the PSTM configuration, both propagating and evanescent wave components were thus present above the grating.
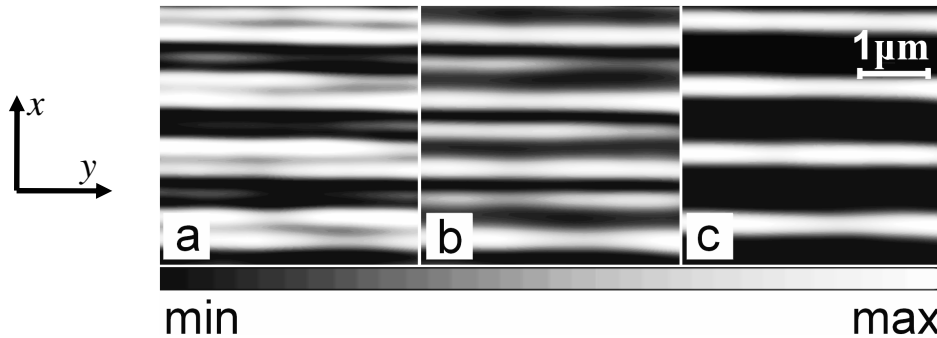
Fig. 4.3: Near-field optical images for tip-surface distances of 100 nm (a), 500 nm (b) and 900 nm (c) are shown. The scanning area is 3.9 x 3.9 μm^2.

SNOM optical images were obtained by scanning in the *x-y* plane at different constant heights above the top surface varying from 100 nm to 900 nm in steps of 100 nm, and three such scans are illustrated in Fig. 4.3. Even though the grating is uniform along the *y* axis, slight variations in the signal along this axis were observed in the SNOM images. As transpires from the following section, we are interested in the resulting *x-z* intensity profile, which was constructed by averaging along the *y* axis and using interpolation. The aforementioned signal variations along the *y* axis were included in the uncertainty of this profile.
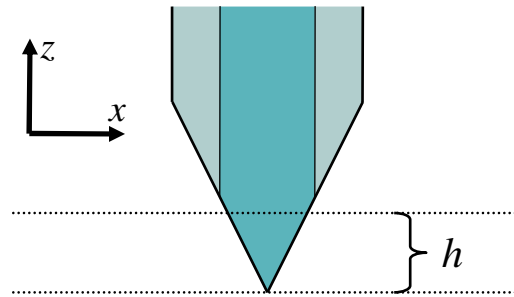
Fig. 4.4: The effective plane of detection is situated a distance *h* above the apex plane.

When the topography identification was performed in the beginning of the Ph.D. work, the transfer function calculations described in the following chapter had not yet been performed, and the exact relation between the near-field and the optical image was unclear. For this reason it was assumed that the detected signal determined by the power carried by the guided mode of the single-mode fiber is proportional to the near-field intensity in an effective plane of detection (EPD) located at the height *h* relative to the tip apex as illustrated in Fig. 4.4. When using *s* polarization, there is only one electric field component along the *y* axis, and it was assumed that the SNOM images reflect the absolute square of the E_y component. For *p*-polarized light, however, the detected field has two field components along the *x* and *z* axes, and as the relative contributions of the E_x and E_z components to the fiber mode were unknown, the relation between the measured signal and the near-field intensity was more complicated. Therefore the analysis described below is based only on measurements with *s*-polarized light.

4.2 Inversion procedure

The grating is too deep for the assumptions of perturbative methods[31-34] to solve the ISP, and we have chosen the procedure suggested in Ref. 35 instead. The method consists of a geometrical model, a direct field calculation, a comparison between modeling and experimental data and a minimization routine.

We start by defining a geometrical model with parameter vector $\mathbf{p}$. We assume that the geometry can be described by this model, and the components of the vector $\mathbf{p}$ are the parameters to be determined. Unknown non-structural information, such as the intensity and the angle of incidence of the incoming light, can also be described by these parameters.

The field above the grating is then calculated for a parameter vector. We compare the modelled field $I(\mathbf{r}\,|\,\mathbf{p})$ with the experimental data $I_{\mathrm{exp}}(\mathbf{r})$ and calculate a fit value $f(\mathbf{p})$ using:

$$f(\mathbf{p}) = \int \left[I_{\mathrm{exp}}(\mathbf{r}) - I(\mathbf{r}|\mathbf{p}) \right]^2 d\mathbf{r} \, . \tag{4.1}$$

The inversion procedure now consists of varying $\mathbf{p}$ and minimizing $f(\mathbf{p})$ with standard minimization routines. The global minimum $\mathbf{p}^{(\mathrm{min})}$ is then identified as the correct parameter vector $\mathbf{p}^{(c)}$ that reflects the geometry of the real structure. Ideally this would satisfy $f(\mathbf{p}^{(c)}) = 0$, but an inaccurate geometrical model, noise and the difficulty of measuring the exact field due to the unknown SNOM tip transfer function will result in $f(\mathbf{p}^{(c)}) > 0$.

4.3 Topography identification

We have employed a simple rectangular two-parameter geometry illustrated in Fig. 4.5 to model the grating. Here, $\mathbf{p} = (H, W)$. Fig. 4.1 hints that the real sample is better described by a more complicated model, but for reasons given in section 4.4 we have chosen to determine a simple two-parameter equivalent profile.
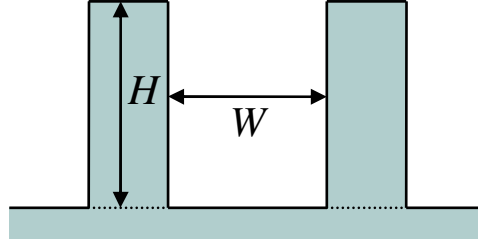
Fig. 4.5: The two-parameter geometric model of the grating is shown.

With only two parameters, the parameter landscape within reasonable limits could be calculated. The function $\log(f(\mathbf{p}))$ is illustrated in Fig. 4.6.
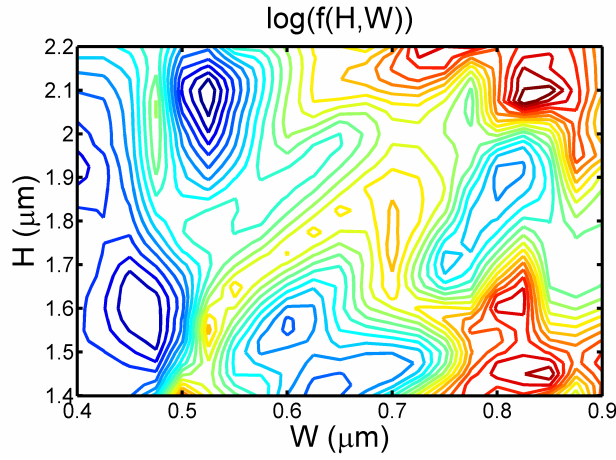
Fig. 4.6: A contour plot of the function $\log(f(\mathbf{p}))$ is shown. The minimum values are illustrated in blue.

In the comparison of the SNOM image to the modeled field intensity distribution for a parameter set $\mathbf{p}$, the relative height h of the EPD was introduced as a variable. For each set $\mathbf{p}$ the function $\log(f(\mathbf{p},h))$ was minimized with respect to h, such that

$$f(\mathbf{p}) = \min f(\mathbf{p},h)\big|_{\mathbf{p}} . \qquad (4.2)$$

In this way, the global minimum $\mathbf{p}^{(\min)}$ was unambiguously determined.

Parameter	Best fit value
H	2.093 ± 0.003 μm
W	0.516 ± 0.001 μm

Table 4.1: The $\mathbf{p}^{(\min)}$ parameters are given.

The corresponding parameters are given in Table 4.1, and the uncertainty on the parameters is due to the variation of the SNOM images along the *y* axis shown in Fig. 4.3. The modeled field intensity at $\mathbf{p}^{(min)}$ is compared to the resulting experimental *x-z* intensity profile in Fig. 4.7.
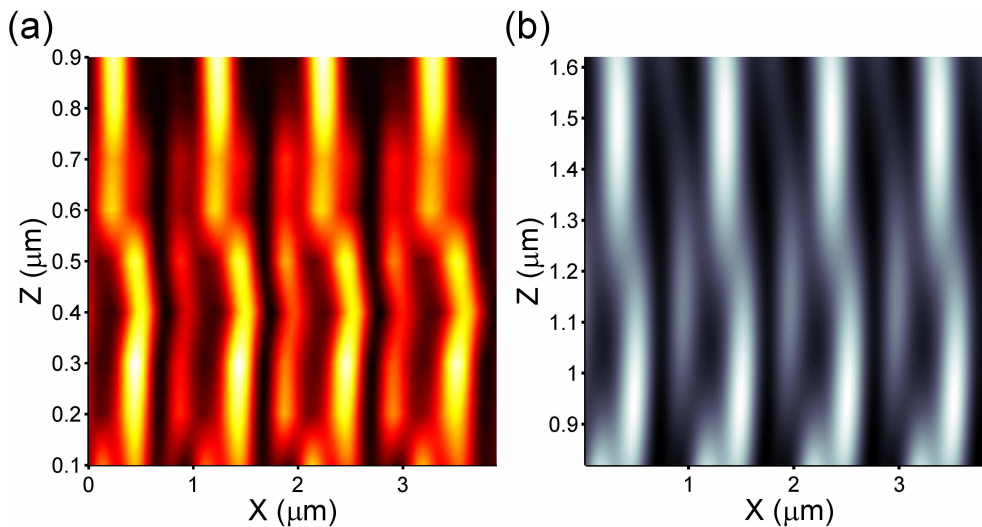
Fig. 4.7: **The experimental average *x-z* measured signal profile (a) and the modeled field intensity at the global minimum (b) are shown.**

Comparing the experimental and calculated intensity profiles one observes a displacement along the *z* axis of the latter with respect to the former by ~ 725 nm away from the surface plane. As the *z* coordinate of the experimental data refers to the height of the tip apex, this indicates that the position of the EPD is $h \approx 725$ nm above the apex. The location of an EPD a certain distance above the apex has also been noticed previously[40,41] when comparing experimental images and theoretical simulations.

4.4 Discussion

We have determined the parameters of a rectangular grating using the direct modeling of light diffraction by the grating and an optimization procedure to fit the numerically and experimentally obtained field intensity distributions. In the following, we consider the advantages and limitations of this approach.

The calculations performed are the result of direct modeling of the field scattered by a given dielectric refractive index structure. Existing well-established methods can be used and the development of dedicated inversion programs is not required.

The general problem with optimization routines is related to the difficulty of determining a global minimum rather than just a local one without searching the entire parameter space. Many methods have been proposed but a fool-proofed technique does not exist, limiting the inversion technique to very few free parameters. Furthermore, as the bijectivity of the geometry-image map is not established, one cannot automatically rule out the existence of another parameter vector $\mathbf{p}'^{(\min)}$ that minimizes the fit function equally well, i.e. that satisfies the relation $f(\mathbf{p}'^{(\min)}) = f(\mathbf{p}^{(\min)})$. If different geometries would result in an identical fit value, prior knowledge about the structure would be required to determine which parameter minimum reflects the true geometry.

In our case, Fig. 4.1 suggests that the grating profile is not rectangular and a five-parameter geometrical model with graded side walls was tested. However, with the calculation time of a few minutes per parameter vector, a complete mapping of the five-parameter space with any reasonable resolution was not feasible, and identical $f(\mathbf{p}^{(\min)})$ values were found for different local minima, indicating that the geometry-image mapping is not single-valued.

Under the questionable assumption that the two-parameter geometry model is valid, we could determine the grating geometry with small uncertainty. We should however acknowledge that more complicated geometries produce intensity field profiles very similar to that measured experimentally. For these geometries, bijectivity of the geometry-image map should be achieved by performing additional SNOM measurements, e.g., made at other wavelengths and/or for p polarization and/or for different angles of incidence. Currently, however, it is not well-understood what information about the sample geometry can be obtained with a particular measurement configuration. By increasing the number of measurements we expect to decrease the probability of having the same intensity profile for different grating geometries, but a systematic procedure that would ensure one-to-one correspondence between intensity profiles and sample geometries is yet to be developed.

Before pursuing the inversion problem of topography characterization any further, one important issue needs addressing: We have assumed that the SNOM image faithfully

represents the free space field intensity in an EPD located a certain distance above the tip apex. However, the validity of this assumption is questionable, and as the input of the topography identification routine is the intensity distribution, any deviation between the SNOM image and the field intensity can thus have the consequence that the procedure identifies the wrong geometry. This issue is the topic of the next chapter.

5 Modeling of near-field imaging

The interpretation of the measured SNOM image has been subject to debate[36,42-46]. The SNOM clearly allows detection of sub-wavelength features, but in spite of progress in the understanding of the field-image coupling mechanism, the exact relation between the optical field and the observed SNOM image remains unclear.

Near-field detection is generically complicated by the close proximity of the fiber tip, resulting in tip-sample interaction modifying the free-space optical field generated in absence of the nano-collector. The detected field thus includes a perturbation of the original free-space field, a perturbation that is generally difficult or impossible to separate. Fortunately, it has been shown that for a dielectric sample studied in the PSTM setup, the tip-sample interaction is negligible[47], and in this case the fiber tip can be approximated to a passive detector of the optical field.

Initially, the nano-collector was considered a point-like probe[42] detecting the optical field intensity at a specific position. It was argued[43] that the finite size of the probe should be taken into account by introducing a transfer function. The intensity transfer function (ITF), relating the absolute square of the electric field to the SNOM image, was proposed[43] and the conception that the image represents field intensity below a spatial frequency cut-off has been used to analyze experimental near-field data[48,49].

It was then shown that the ITF does not accurately model the field-image relation[44]. The correct general framework necessary for describing the field-image coupling was given in Ref. 45. It was pointed out[36] that when utilizing a single-mode fiber, the general TF can be described by an amplitude coupling function. General symmetry properties of this TF, correctly relating the optical field to the image, are given in Ref. 46.

Further information about the nature of the TF was obtained as comparisons[40,41] between modeling and experiment, such as the one in chapter 4, revealed correspondence between detected SNOM images and the intensities of the free space fields in planes a distance above the tip apex. This hints that the optical field may be coupled to the guided fiber mode not at the very tip apex, but in an effective detection point inside the tip.

However, to completely understand the field-image relation, exact numerical modeling of the field scattered on the fiber tip must be performed. A comparison

between free-space field and image measured using a reduced 2D tip was performed in Ref. 50 and a full 3D simulation of the field scattered by the tip was presented in Ref. 51. Unfortunately, in these works the concept of TF was not invoked.

In this chapter, we examine TFs for conical single-mode fiber tips computed using the EET. The coupling efficiency of glass fiber tips is determined in a wide range of spatial frequencies of the incident radiation for tip opening angles varying from 30° to 120°. The characteristics of the TFs describing the relation between probed optical fields and near-field images are analyzed in detail.

5.1 Near-field coupling

The general experimental SNOM setup we will model is illustrated in Fig. 5.1. A dielectric sample with subwavelength features is placed on a prism and is subject to illumination from below at the wavelength λ of 633 nm. The sample scatters the light, and a near-field is generated.
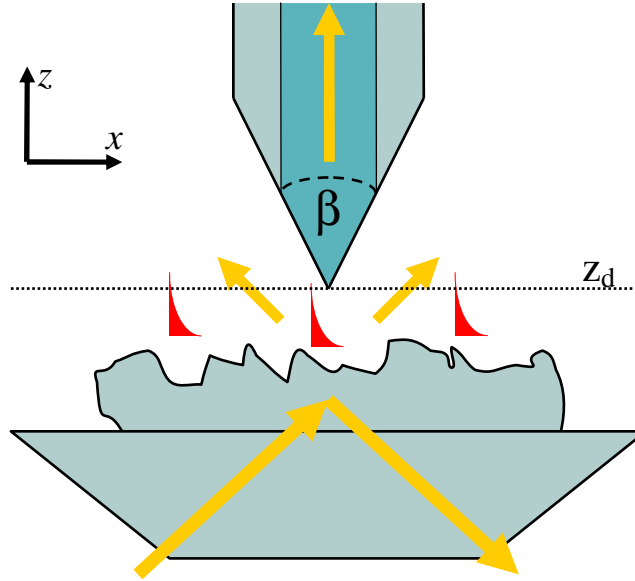

Fig. 5.1: The SNOM measurement setup is illustrated.

The SNOM fiber tip is scanned above the sample in constant height mode[9] with tip apex in a detection plane $z = z_d$. Other scanning modes, such as constant distance mode[9], are frequently used in SNOM microscopy, but the TF relates the field in a

detection plane of fixed z coordinate to the image, and the choice of constant height mode is thus necessary if field-image comparisons are to be made.

The tip couples the field to the guided fiber mode, but also scatters light back towards the sample. We restrict our analysis to SNOM measurements on weakly reflecting dielectric samples, where the tip-sample interaction can be ignored,[47] and we thus only treat the coupling of the near-field into the guided mode. Also, we assume that power levels are sufficiently low so that linear scattering can be assumed. These restrictions are made as the TF, relating the free-space field with the observed SNOM image, is only defined[45] in the linear scattering regime and in absence of multiple reflections between tip and sample. A treatment of the field-image relation in setups where the tip-sample interaction or non-linear effects are important is beyond the scope of this Ph. D. thesis.

In the detection plane $(\mathbf{r}_\perp, z_d)$ we consider the electric field $\mathbf{E} = (E_x, E_y, E_z)$. We then make the plane-wave expansion in the detection plane:

$$\mathbf{E}(\mathbf{r}_\perp, z_d) = \int \mathbf{F}(\mathbf{k}_\perp) \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) d\mathbf{k}_\perp . \tag{5.1}$$

The z dependence of the plane waves are given by $\exp(ik_z z)$ with $k_z = \sqrt{k_0^2 - |\mathbf{k}_\perp|^2}$, where $k_0 = 2\pi/\lambda$. Plane waves with $|\mathbf{k}_\perp|^2 < k_0^2$ correspond to the propagating part of the near-field, while those with $|\mathbf{k}_\perp|^2 > k_0^2$ represent its evanescent part.

The three components of the vector field $\mathbf{E} = (E_x, E_y, E_z)$ are not independent but subject to the divergence relation $\nabla \cdot \mathbf{E} = 0$, valid in free space. For this reason we will in the following only consider the transverse components $\mathbf{E}_\perp = (E_x, E_y)$ and $\mathbf{F}_\perp = (F_x, F_y)$ of the field and of the expansion coefficients respectively. The z component of the field can be determined at any time using

$$E_z(\mathbf{r}_\perp, z_d) = -\int \frac{\mathbf{F}_\perp \cdot \mathbf{k}_\perp}{k_z} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) d\mathbf{k}_\perp . \tag{5.2}$$

Now, the optical field couples to guided fiber modes n with angular momenta m with normalized mode profiles $\mathbf{G}^{n,m}(r) \exp(im\varphi)$. Following Ref. 45, we introduce vector coupling coefficients $\mathbf{H}^{n,m}(\mathbf{k}_\perp)$ describing the coupling between a plane wave with
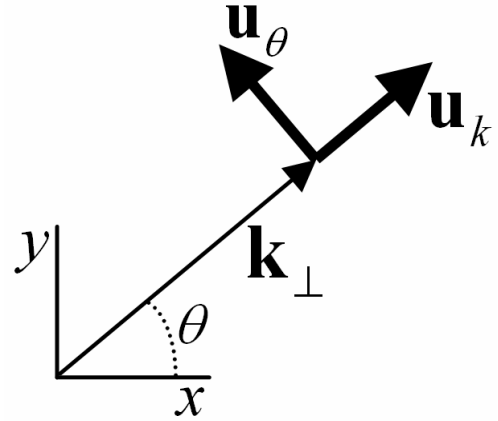
wave vector $\mathbf{k}_\perp$ and the guided fiber mode with field profile of the form $A^{n,m}\mathbf{G}^{n,m}(r)\exp(im\varphi)$. The contribution from the total field to a guided mode is then

$$A^{n,m}(\mathbf{r}_\perp) = \int \mathbf{H}^{n,m}(\mathbf{k}_\perp) \cdot \mathbf{F}_\perp(\mathbf{k}_\perp)\exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)d\mathbf{k}_\perp , \qquad (5.3)$$

where $\mathbf{r}_\perp$ refers to the position of the apex in the plane $z = z_d$. The mode profiles $\mathbf{G}^{n,m}(r)\exp(im\varphi)$ are normalized such that the total power S detected at the end of the fiber is given by $S(\mathbf{r}_\perp) = \sum_{n,m}\left|A^{n,m}(\mathbf{r}_\perp)\right|^2$.

To take advantage of the rotational symmetry of the optical fiber, the plane wave vector components of Eq. (5.1) may be expressed in cylindrical coordinates as $\mathbf{F}_\perp(k,\theta) = F_p(k,\theta)\mathbf{u}_k + F_s(k,\theta)\mathbf{u}_\theta$, corresponding to s- and p-polarized parts of the field. We obtain:

$$\mathbf{E}_\perp(\mathbf{r}_\perp, z_d) = \int \left(F_p(k,\theta)\mathbf{u}_k + F_s(k,\theta)\mathbf{u}_\theta\right)\exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)d\mathbf{k}_\perp . \qquad (5.4)$$

The cylindrical unit vectors are illustrated in Fig. 5.2. We remark in passing that when calculating the z component of the electric field using Eq. (5.2), only the p-polarized part of the field gives a contribution.
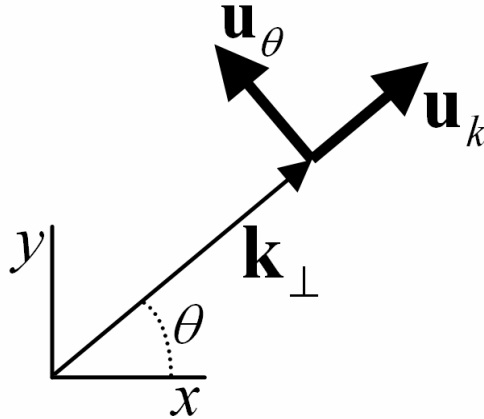
For a single-mode fiber of weak refractive index contrast we only need to include coefficients (A_x, A_y) of the two orthogonal linearly polarized modes of the form $A_{x,y}\mathbf{G}_{x,y}(\mathbf{r}_\perp)$, where the profiles $\mathbf{G}_{x,y}(\mathbf{r}_\perp)$ as previously are normalized such that the total fiber power is $S(\mathbf{r}_\perp) = |A_x(\mathbf{r}_\perp)|^2 + |A_y(\mathbf{r}_\perp)|^2$. The field contribution to the two modes is then given by:

$$\begin{bmatrix} A_x(\mathbf{r}_\perp) \\ A_y(\mathbf{r}_\perp) \end{bmatrix} = \int \begin{bmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} H_s(k)F_s(k,\theta) \\ H_p(k)F_p(k,\theta) \end{bmatrix} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)d\mathbf{k}_\perp. \qquad (5.5)$$

The detailed derivation of this equation is given in appendix 8.4. In Eq. (5.5), $H_s(k)$ and $H_p(k)$ represent the coupling of an incoming plane wave of s and p polarization respectively with in-plane wave vector of norm k to a guided fiber mode. Due to symmetry, the functions $H_s(k)$ and $H_p(k)$ are independent of the angle of incidence θ. The functions $H_s(k)$ and $H_p(k)$ are the TFs that describe the detection capability of a particular fiber tip, that we compute numerically using the procedure described in chapter 3. In the following, the TFs are presented as functions of k/k_0 for convenience.

5.2 Convergence

When performing computations of the TF using the EET, we would ideally like to include an infinite number of eigenmodes in the expansion of the field. Assuming that the eigenmodes constitute a complete basis set, an infinite number of modes would ensure that the true optical field was identical to its eigenmode expansion.

Unfortunately, we must, for numerical reasons, restrict ourselves to including a finite number N of eigenmodes. Even though this means that our computation is no longer exact, we can still in general expect that only a limited number of eigenmodes are necessary to assure that the computation reflects reality within a reasonable error margin.

Furthermore, we are generally interested in the properties of a structure placed in free space, but the EET requires that we enclose our geometry inside a metal cylinder. Parasitic reflections from the metal walls will influence the system under study, but

we can expect their influence to become negligible when the radius of the metal cylinder is sufficiently large.

To assure that the number of included modes in the computation is large enough and that parasitic reflections do not influence our results, convergence checks must be performed. The convergence issues related to the determination of the TF using the EET are non-trivial, and for this reason we devote this subsection to a convergence study of TFs of uncoated and metal-coated fiber tips. The computations for uncoated tips are all full 3D, but for reasons explained below 2D calculations are also considered for metal-coated tips.

5.2.1 Uncoated fiber tips

The uncoated fiber tip geometry modeled in this work is that of a tapered single-mode fiber with core and cladding refractive indices of 1.4622 and 1.4572, respectively, and a core diameter of 4 μm. The refractive index geometry of the computation is illustrated in Fig. 5.3. The tip opening angle β varies, but the wavelength λ is 633 nm in all calculations in this thesis.
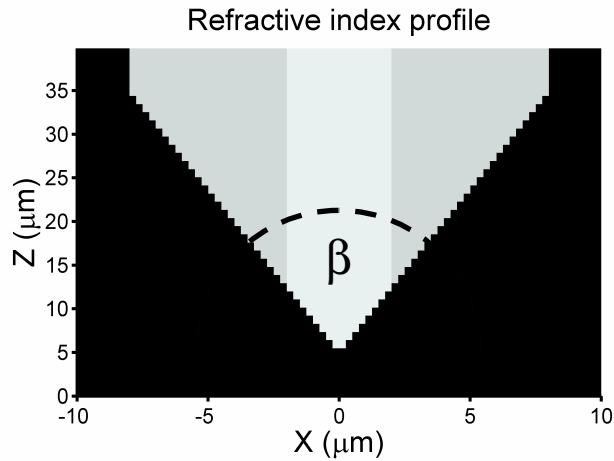

Fig. 5.3: The index geometry of a SNOM fiber tip is illustrated.

In the computation of the TF for the uncoated fiber tip, the complex modes described in section 3.3.2 did not appear, and the TF could thus be calculated using the SA approach as well as using the FD and PW methods. TFs computed for an increasing

number of eigenmodes using the three mode calculation techniques are illustrated below:
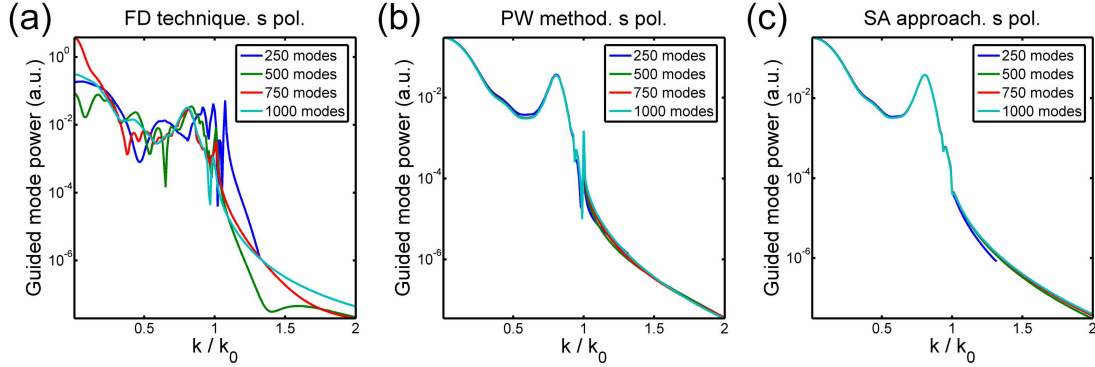

Fig. 5.4: The figure shows TFs obtained using the finite difference technique (a), plane-wave expansion (b) and the semi-analytical approach (c) using a varying number of modes. $\beta = 30°$ and $L_r = 30$ μm.

Inspecting the figure, we observe that the convergence of the TF calculated using the FD technique is very poor. The TF computed using the PW expansion converges nicely, however, a discontinuity is observed at $k/k_0 = 1$. This discontinuity is not physical, and its presence reveals a difficulty of the PW expansion in computing the correct guided mode coupling at this value of k/k_0. The best convergence is, not surprisingly, obtained using the SA approach. In Fig. 5.4(c), we do observe a small shoulder near $k/k_0 = 1$, most likely caused by parasitic reflections from the metal cylinder, however, the curves from the SA calculation are still much smoother than those of the PW computation.

From Fig. 5.4(c), it is clear that only ~ 500 eigenmodes are necessary to achieve convergence. Except for the procedure used to determine eigenmodes, the computations performed are completely identical, and we can thus conclude that, even though the FD technique handles field discontinuities well, the general quality of its mode profiles is inferior compared to that of modes of the other methods. On the other hand, the calculations using PW expansion generally yields nice performance. The PW expansion has difficulty describing field discontinuities, but we conclude that the index contrast and thus the field discontinuities are weak enough so that they may be described with sufficient precision using the number of basis modes included in the computations above. The best TFs are obtained using the SA approach, and in the

following convergence studies the SA approach was employed to determine the mode profiles.

Convergence of the TF calculation for an increasing number of eigenmodes is necessary, however, other issues need addressing as well. First of all, real optical fibers feature a cladding diameter of ~ 125 μm. However, the guided mode profile only extends a couple of microns into the cladding, and we thus expect that computation on a reduced fiber tip with a much smaller cladding diameter to be sufficient.
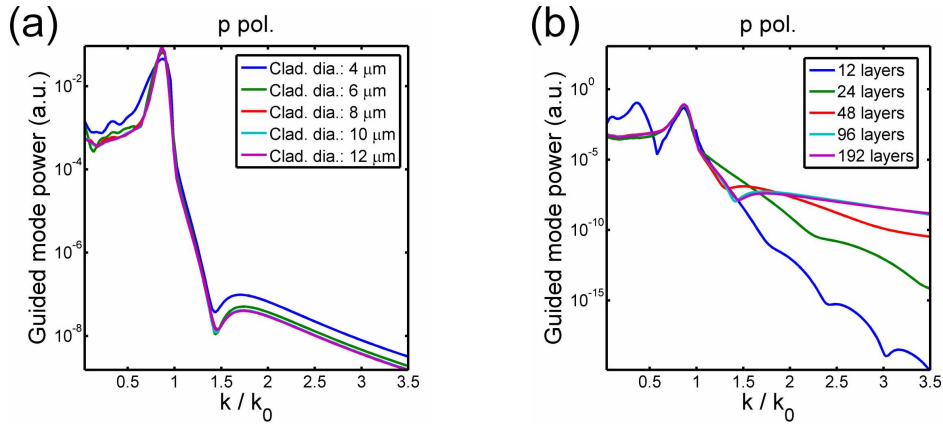
Fig. 5.5: The TF is illustrated for varying cladding diameter (a) and varying number of layers (b). $\beta = 60°$ and $L_r = 10$ μm and in (b) the cladding diameter is 12 μm.

TFs for fiber tips of varying cladding diameter are depicted in Fig. 5.5(a). We observe that the TF converges very quickly when the cladding diameter is increased, and it is thus only necessary to consider a reduced fiber tip with a cladding diameter of ~ 12 μm when calculating the transfer function.

A second issue is the number of layers used in the representation of the tip. As the EET treats structures consisting of layers of uniform index profile along the z axis, we model the conically shaped tip by using a large number of thin layers, resulting in the saw-tooth profile shown in Fig. 5.3. The number of layers required is studied in Fig. 5.5(b). For $k/k_0 < 1$, only about 48 layers are required to achieve convergence, however, to obtain a converging TF for all values of k about 192 layers must be used.

Finally, we study the effect of enclosing the fiber tip in a metal cylinder. We aim at determining the TF for a fiber tip in open space, but we are forced to perform computations on a fiber tip placed inside a metal cylinder as explained earlier.

However, when we increase the metal cylinder radius towards infinity, we expect the TF to converge towards its free space value.
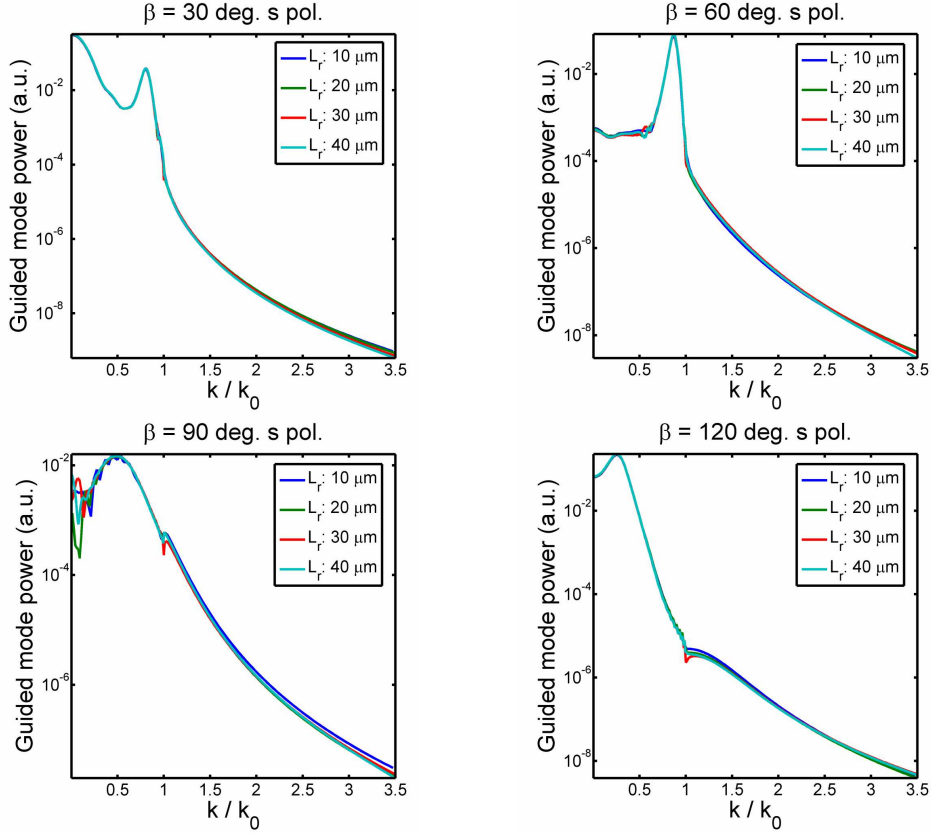
Fig. 5.6: The TFs for varying metal cylinder radius L_r are shown for $\beta = 30°$, $60°$, $90°$ and $120°$.

We study the influence of the metal cylinder radius on the TF by inspecting the figure above. For a fiber tip having an opening angle β equal to $30°$, $60°$ or $120°$ we observe virtually no dependence of the TF on the radius of the cylinder. This may be somewhat surprising as one expects the optical field distribution to highly perturbed by the metal wall in particular for the small cylinder radius of 10 µm. But whatever parasitic reflections are introduced by the metal wall, the figure shows that, when $\beta = 30°$, $60°$ or $120°$, these reflections do not influence the magnitude of the coupling to the guided fiber mode for the angles. Inspecting the TF for $\beta = 90°$, we observe an oscillation behaviour near $k/k_0 < 0.5$, and in this interval the TF does not converge when the cylinder radius is increased. We conclude that for this particular opening angle, the parasitic reflections do influence the TF for $k/k_0 < 0.5$.

5.2.2 Metal coated tips

Inspired by the success in calculating TFs for uncoated fiber tips, the possibility of determining the TF for a fiber tip covered with a thin layer of metal was examined. Optical resolution of ~ $\lambda/40$ has been achieved experimentally[52] with a metal-coated tip, and this geometry more closely resembles the original design by Synge.

We would like to calculate the TF of a fiber tip with the index profile illustrated in Fig. 5.7(a).
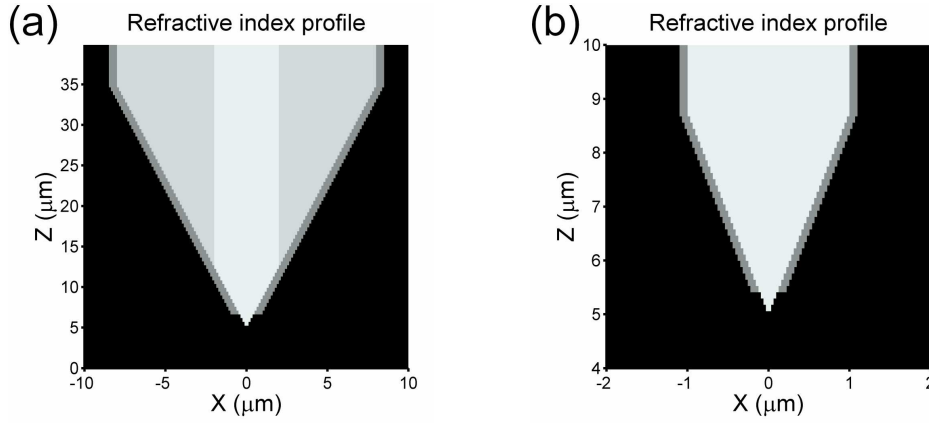
Fig. 5.7: The geometries of a full tip profile (a) and a reduced tip (b) with metal coatings are shown.

The fiber is covered with a layer of aluminum except at the very apex of the tip, and an aperture is thus be formed. The typical thickness t of a metal layer covering a SNOM fiber tip is usually ~ 25-100 nm, and the coating is thus fairly thin. However, the real challenge lies in the refractive index of aluminum, which is ~ $1.30 + 7.48i$ at the wavelength $\lambda = 633$ nm, and a highly absorbing material is thus introduced into the structure.

As calculation time for a complex refractive index profile is very long, the reduced metal-coated tip structure depicted in Fig. 5.7(b) was investigated. Calculations were performed only using the SA approach and the PW expansion as the performance of the FD technique in the previous section was found to be very poor.

When determining modes in a zone with a complex refractive index profile the perturbative method described in the end of section 3.4.3 was employed in the SA approach, whereas the PW code handles the complex refractive index directly.

TFs obtained for the reduced metal-coated tip for a varying number of modes are shown in the figure below:
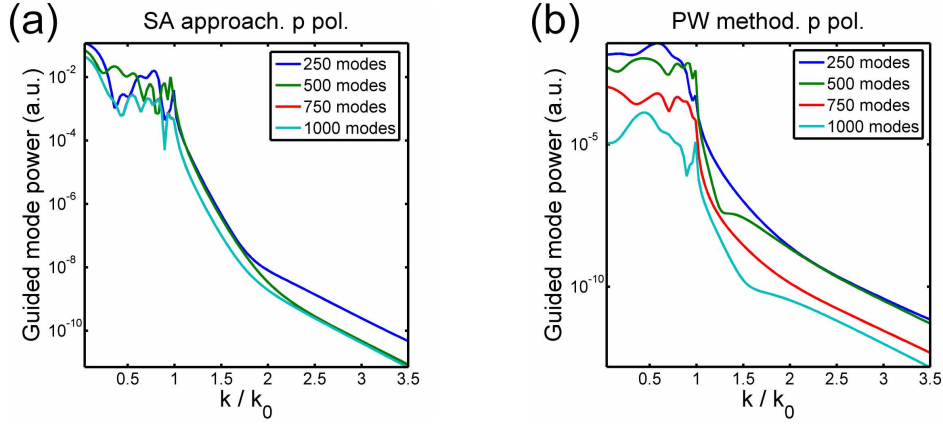
Fig. 5.8: The TFs for the reduced metal-coated tip obtained using the SA approach (a) and the PW expansion (b) for a varying number of modes are shown. $\beta = 30°$, $L_r = 10$ μm and $t = 50$ nm.

Inspecting Fig. 5.8(a) we observe a slower convergence of the TF when the number of modes is increased compared to the case of the uncoated metal tip. However, the curves for 750 and 1000 modes are identical. and this would suggest that the SA approach can handle the metal-coating and yield convergence for a reasonable number of eigenmodes. This, however, is only partly true and the problem lies in the determination of the propagation constants of eigenmodes. The perturbative method used to determine eigenmodes was optimized with respect to speed while keeping a reasonable accuracy, but even then, the typical calculation time necessary to find propagation constants for 1000 complex eigenmodes was above 2 hours and wrong propagation constants would still appear from time to time due to insufficient accuracy. Even though the perturbation method initially seems to work, it is slow and not very stable.

In the previous section, we saw that the PW expansion method performed reasonably well for uncoated tips, and the method was also tested for the metal-coated tip. The resulting TFs are depicted in Fig. 5.8(b). We observe that even though the general shape of the curves matches those obtained using the SA approach, the TF does not converge. The problem can be understood by studying the profiles of eigenmodes obtained using the two methods:

69
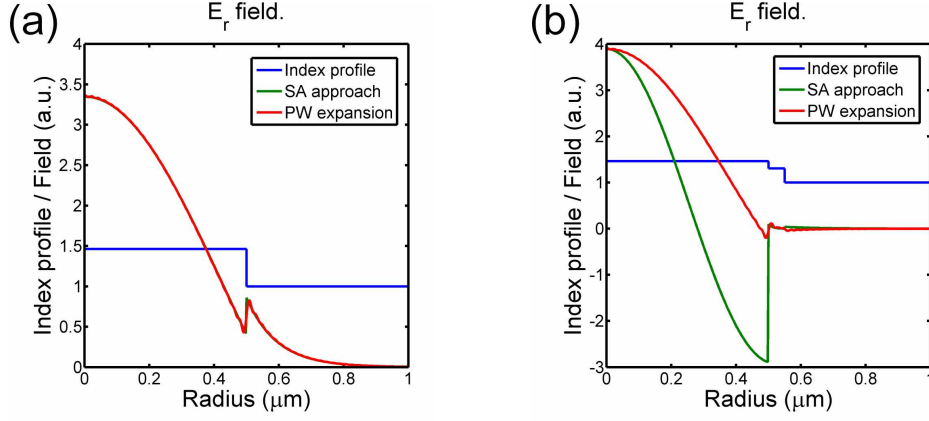
Fig. 5.9: **The radial component of the fundamental mode in a glass waveguide (a) and a metal-coated guide (b) calculated using the SA and the PW methods are shown. The real part of the field and the index profile are depicted.**

The fundamental mode profile in a glass waveguide obtained using the PW expansion method with 2000 basis modes is compared to the perfect mode profile derived using the SA approach in Fig. 5.9(a). We observe that, since the index step is so small, the PW expansion can nicely model the discontinuity at the glass-air interface, and this results in the fast convergence of the method for uncoated tips. In Fig. 5.9(b), mode profiles obtained in a metal-coated waveguide are shown. The large refractive index of aluminum causes a steep discontinuity of the radial field component at the glass-metal interface clearly visible on the profile obtained using the SA approach. Comparing the two profiles, we observe that the curve calculated using PW expansion does not resemble the perfect curve obtained using the SA approach. As discussed at the end of section 0, a discontinuity in the refractive index results in a discontinuity of the radial electric field component, which is difficult to expand on a limited set of basis modes. We can thus expect the PW expansion technique to fail in computing the correct eigenmodes in structures featuring large index steps, as it is the case for the structure illustrated in Fig. 5.9(b).

As the problem is caused by the discontinuity, we can expect the PW method to perform better in structures where the steep discontinuities have been replaced with linear gradients. A graded index profile will result in a continuous radial field component, which is more easily described using plane waves. For this reason we examine a geometry with linear gradients at the metal interfaces of the form $\varepsilon_r(r) = ar + b$.
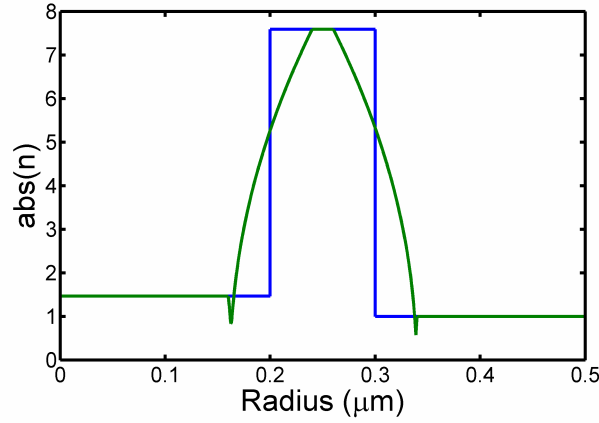
Fig. 5.10: The absolute value of the refractive index profile of a standard zone without gradients (blue) and a zone with gradients at the glass-metal and metal-air interfaces (green) is shown.

The absolute value of the radial index profiles for a standard structure and for a geometry with gradients is depicted in Fig. 5.10. The glass waveguide has a radius of 200 nm and is covered with a 100 nm thick layer of aluminum. In the graded profile, 80 nm thick gradient regions are included where the dielectric constant ε_r changes linearly. As the figure depicts the absolute value of the refractive index rather than ε_r, the linear gradients are shown in a slightly curved shape.

Unfortunately, the calculation of operator matrices $\overline{\overline{O}}_1$ in the cylindrically symmetric version of the PW expansion involves the evaluation of the indefinite integral $\int J_l(a_m r) J_l(a_n r) dr$, which has no known anti-derivative. The integrals can be evaluated numerically, but, in practice, this is far too slow. However, instead of examining a 3D geometry, we can consider a 2D fiber tip in Cartesian coordinates with gradients along the x axis of the form $\varepsilon_r(x) = ax + b$ and with uniformity along the y axis. In Cartesian coordinates, primitive functions for all the indefinite integrals exist, and the determination of the operator matrix is thus done in a reasonable amount of time. Even though a 2D calculation is a simplification of the full 3D problem, 2D computations will allow us to evaluate the effect of implementing a gradient at the metal interfaces.

71
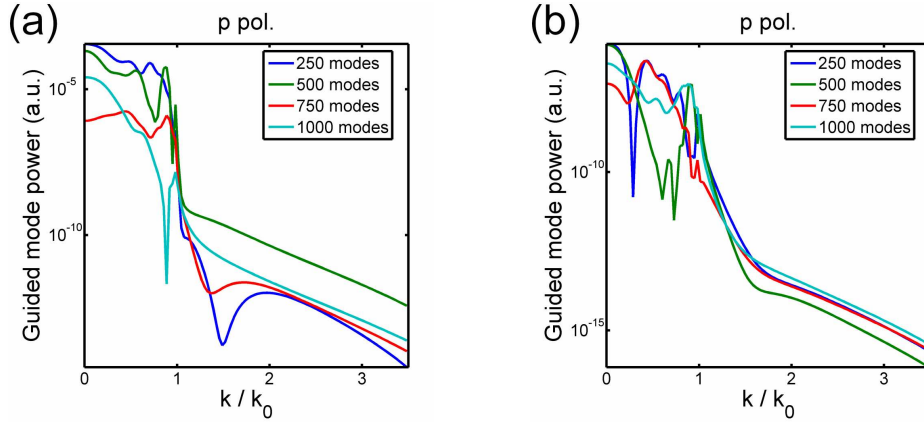
Fig. 5.11: TFs calculated using the 2D PW method are shown for the standard metal-coated reduced tip (a) and for a tip with gradients at the metal interfaces (b).
$\beta = 30°$, $L_x = 20\ \mu m$ and $t = 100\ nm$.

In Fig. 5.11(a), TFs calculated for a 2D metal-coated tip without gradients is shown, and as in the 3D case the convergence is very poor. TFs for the same tip, where the index discontinuities have been replaced with gradients, are illustrated in Fig. 5.11(b). Comparing the convergence in the case of the graded vs. the ungraded structure, a slight improvement is observed for the graded structure, however the convergence is still not acceptable. The lack of convergence is most likely due to the fact that, even though the discontinuity has been converted into a gradient, the gradient width is still less than 0.5 % of the total geometry width, and the change in the field is thus still very steep compared to the width of the geometry. Softer gradients of widths around ~ 500 nm or more can of course be considered, but a geometry with such thick gradients deviates greatly from the true physical geometry, and the extent to which a calculation on such a graded structure reflects the physical reality is questionable.

The possibility of calculating the TF of a metal-coated fiber tip was, due to time constraints, not investigated further.

5.2.3 Discussion

The EET has proven to be an efficient tool when simulating the scattering of light on 3D rotationally symmetric uncoated fiber tips. The main strength of the EET is its ability to handle geometries with uniformity or periodicity along the propagation axis. The tapered tip features neither, but the computation of its TF is still precise and reasonably fast considering the size of the geometry.

The EET simulates light scattering in geometries that are limited by artificial boundaries introducing parasitic reflections. We do not calculate the field scattered by a structure placed in free space, but we are forced to consider the influence of the metal walls. By extending the cylinder radius or by introducing advanced boundary conditions, this influence can be reduced, however, in the case of TFs calculated for uncoated fiber tips, the influence of parasitic reflections is generally negligible even for small cylinder radii.

5.3 Transfer functions

The TF computed using a full 3D simulation allows us to determine the power carried by the guided fiber mode for a given incoming light intensity as function of $k = \left| \mathbf{k}_\perp \right|$.

The following TFs are calculated for an electric field amplitude of 27.5 kV/m in the detection plane corresponding to an illumination intensity of 1 $\mu W/\mu m^2$ for a propagating wave.

TFs determined for s and p polarization for various opening angles are depicted in Fig. 5.12.
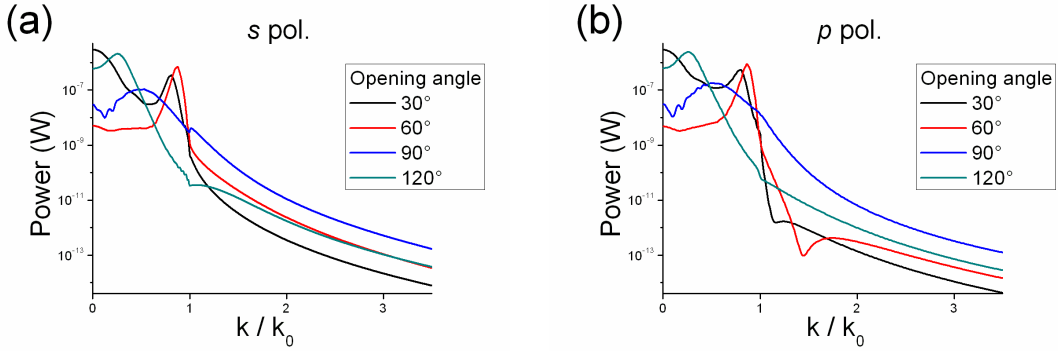
Fig. 5.12: TFs for opening angles β from 30° to 120° for s polarization (a) and p polarization (b) are shown.

In the propagating regime ($k/k_0 < 1$), we observe an oscillatory behavior for all the TFs of both polarizations. The peaks in this part of the spectrum represent coupling resonances, the positions of which depend on geometry. In the evanescent regime, we generally observe a monotonic near-exponential decrease except for p polarization in

the interval $1 < k/k_0 < 1.5$ when $\beta = 60°$ or $90°$. We also note that, for $k/k_0 > 1$, the coupling efficiency as function of opening angle seems to be at a maximum for an opening angle β around $90°$. However, it should be stressed that determination of the exact optimal opening angle requires a study of TFs for more than four opening angles.

5.4 Resolving power

One of the motivations for calculating a TF is that it allows us to determine the resolving power of a SNOM fiber tip. The definition of resolution, however, depends on whether the optical fields distributions are limited in real space or in k space.

To introduce the concept of resolution for distributions limited in real space, we consider an electric field distribution whose x or z component in the detection plane is of the shape of a delta-function. In the first case the in-plane field distribution is given by $\mathbf{E}_\perp(\mathbf{r}_\perp, z_d) = \delta(\mathbf{r}_\perp)\mathbf{u}_x$, resulting in $F_x(\mathbf{k}_\perp) = 1/(2\pi)^2$ for the plane wave expansion coefficients. $\nabla \cdot \mathbf{E} = 0$ must be satisfied however, so the z component of the electric field distribution is non-zero and can be determined using Eq. (5.2). Secondly, when the z component of the field is a delta-function we have $E_z(\mathbf{r}_\perp, z_d) = \delta(\mathbf{r}_\perp)$ and $F_z(\mathbf{k}_\perp) = 1/(2\pi)^2$, and this time the requirement $\nabla \cdot \mathbf{E} = 0$ leads to a non-zero in-plane field. As s-polarized waves do not contribute to the z component of the field, we set $F_s = 0$ and obtain the p-polarized in-plane field contributions from $F_p(\mathbf{k}_\perp) = -k_z F_z(\mathbf{k}_\perp)/|\mathbf{k}_\perp|$, a rotationally symmetric function.

The calculated SNOM images for the delta-function distribution along the x axis are shown in Fig. 5.13(a). For $\beta = 30°$ the SNOM image profile features a near-Gaussian shape, but for the larger opening angles we generally observe an oscillating decay along the x axis.

74
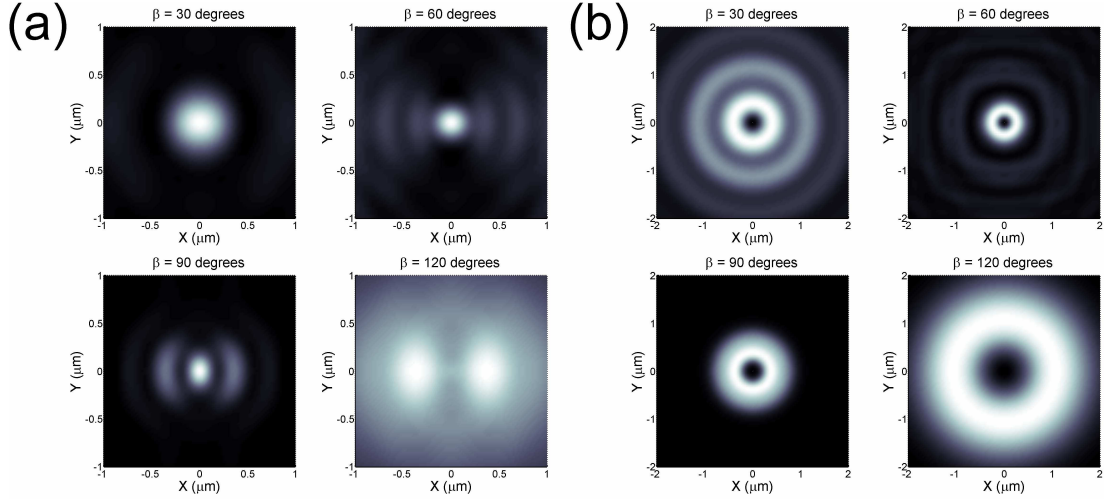
Fig. 5.13: SNOM images of near-field delta functions as function of opening angle β are shown. The *x* component of the field is a delta function in (a) and the *z* component is a delta function in (b).

This is not very surprising as the near-exponential decay of the TFs for $k/k_0 > 1$ serves as an effective cut-off, and we can thus roughly approximate the product $\mathbf{H}^{n,m}(\mathbf{k}_\perp) \cdot \mathbf{F}_\perp(\mathbf{k}_\perp)$ of Eq. (5.3) to a step function equal to zero for $k/k_0 > 1$. Making the Fourier transform we obtain a sinc function having the oscillating decay observed in Fig. 5.13(a) for opening angles of 60°, 90° and 120°.

We characterize the spatial extent of the imaged spots using the average deviation (AD) along the two axes given by $\int S(x,0)|x|dx \Big/ \int S(x,0)dx$ along the *x* axis, and the determined ADs are presented in Table 5.1.

	$\mathbf{E}_\perp = \delta(\mathbf{r}_\perp)\mathbf{u}_x$		$E_z = \delta(\mathbf{r}_\perp)$
Opening angle β	**Average deviation along *x* axis**	**Average deviation along *y* axis**	**Average deviation**
30°	262 nm	160 nm	1175 nm
60°	1029 nm	88 nm	845 nm
90°	255 nm	104 nm	506 nm
120°	698 nm	683 nm	1199 nm

Table 5.1: Average deviations of the images of delta-functions.

The smallest AD is obtained for β = 60° along the *y* axis, however, this opening angle also features the largest AD along the *x* axis due to the slow oscillating decay. The minimum AD along the *x* axis is given for the β = 90° opening angle with an AD along the *y* axis only slightly larger than that for β = 60°. This behavior reflects the

flatness of the TFs: The TF for β = 90° is rather broad and falls off slowly resulting in a wide step function giving a narrow sinc-like spot. On the other hand, the TF for β = 120° decreases much faster, the step function is of shorter width and the resulting spot is wider.

Fig. 5.13(b) presents the SNOM images obtained when the z component of the field is a delta-function. The images preserve the rotational symmetry of the $E_z(\mathbf{r}_\perp, z_d) = \delta(\mathbf{r}_\perp)$ distribution, and we generally observe a donought-shaped profile. We can understand the node at the origin by inspecting Eq. (5.5) for rotationally symmetric profiles $F_s(\mathbf{k}_\perp)$ and $F_p(\mathbf{k}_\perp)$. In this case we have $A_{x,y}(\mathbf{r}_\perp) = -A_{x,y}(-\mathbf{r}_\perp)$ and $A_{x,y}(0)$ must thus equal 0. AD values for the donought-shaped spots are given in Table 5.1, and, as previously, the narrower profiles are obtained for opening angles of 60° and 90°.

We observe that the exact shape of the SNOM image of the delta-functions depends on the nature of the TF. The shape is generally not of Gaussian form, and this complicates the interpretation of an experimentally obtained SNOM image. In the case where the x component of the field is a delta-function, we have good one-to-one correspondence between delta-function and detected spot for β = 30°, but when β = 90° the two weaker spots neighbouring the main spot can be mistaken for near-field point sources of weaker intensity. When the z component of the field is a delta-function the approximation of the images to Gaussian-shaped spots becomes even more ambiguous. We conclude that, even when the concept of resolution is introduced, the identification of the SNOM image as the local electric field intensity is generally incorrect.

If we ignore the non-Gaussian shape of the imaged spots, we can roughly define the resolution of a SNOM tip as twice the AD values given in Table 5.1. We note that the resolution limit defined here is related to the shape of the tip and is independent of noise.

The definition of resolution described above is meaningful when imaging optical field distributions limited in real space. However, when imaging plane waves that can be considered point sources in k space, the concept of resolution should be treated differently:

In previous works,[53,54] the resolution of SNOM setups was determined by imaging counter propagating waves. Two such waves of identical spatial frequency and

amplitude propagating along the x axis and decaying along the z axis result in a standing wave pattern of the form $\cos(k_x x)\exp(-|k_z|z)$ with $k_z = \sqrt{k_0^2 - k_x^2}$, where $k_x > k_0$. The intensity pattern features an image contrast $C_i = (I_{max} - I_{min})/(I_{max} + I_{min})$ of 1, which, without background noise, would be preserved in the SNOM image. However, for increasing k_x and thus increasing $|k_z|$ the calculated TFs show us that the standing wave pattern intensity in the image decays near-exponentially, and when it approaches the background noise floor, the image contrast will deteriorate, and the limit of the spatial resolution has been reached. It should be stressed that, contrary to the previous case, the resolution limit when imaging a point-source in k space is due to background noise: The coupling coefficient drops exponentially for increasing k, but it is never zero. Even though the fiber tip has a finite size, the image contrast would, without noise, be equal to 1 regardless of the value of the spatial frequency.

Keeping the illumination intensity fixed and choosing a noise floor, resolutions of the various tips for the two polarizations can be extracted directly from Fig. 5.12 as the intersection of the TF curves with the noise floor. Resolutions of the tips described in the figure are given in Table 5.2 for (arbitrarily chosen) noise floors of 1 pW and 100 fW.

Opening angle β	Noise floor			
	1 pW	100 fW	1 pW	100 fW
	Resolution (*s* pol.)		Resolution (*p* pol.)	
30°	365 nm	264 nm	437 nm	300 nm
60°	286 nm	209 nm	471 nm	251 nm
90°	219 nm	156 nm	248 nm	180 nm
120°	289 nm	205 nm	317 nm	222 nm

Table 5.2: The SNOM tip resolving power is listed.

Obviously, the values of the resolving power presented here depend directly on the illumination power and the background noise floor, but relative comparison between the resolutions of various fiber tips can still be made: First, the resolutions for *s* polarization are slightly better than those for *p*-polarized light. And second, we

observe that the resolution is generally improved with ~ 40 % when the noise floor is decreased from 1 pW to 100 fW.

Whether we consider point sources in real space or in k space, the best resolution is not obtained for $\beta = 30°$ as one might expect but rather for the large opening angle $\beta = 90°$. The data show that if one is interested in improving the resolution of a SNOM microscope, one should not aim at producing a fiber tip with smallest opening angle possible.

5.5 Effective plane of detection approximation

In the effective plane of detection (EPD) approximation, the coupling of the near field to the fiber tip with apex in the detection plane is proportional to the value of the free-space field produced in absence of tip a certain height h above the detection plane. From $k_z = \sqrt{k_0^2 - k^2}$ we have that $|k_z| \to k$ for $k >> k_0$, so an evanescent wave with spatial dependence of the form $\exp(ikr - |k_z|z)$ in the reference plane will have decayed by a factor ~ $\exp(-kh)$ at the height h. We recognize this exponential decay as the near-linear regime in Fig. 5.12 of the TF curves plotted using semi-logarithmic scale for $k/k_0 > 2$. The coupling of the near-field to the fiber tip in this regime can thus be interpreted as coupling at an effective plane of detection positioned a distance h above the apex.

However, for $k/k_0 < 2$ the TF curves are not quite linear. The EPD height h can still be defined, but it should then be a function of k: For a given k we solve the two equations $S(k) = a \exp(2ik_z(k)h)$ and $S(k + \Delta k) = a \exp(2ik_z(k + \Delta k)h)$ for the parameters a and h when $\Delta k \to 0$ (the factor 2 is included as we fit to power, not field strength). We obtain the function $h(k)$, illustrated in Fig. 5.14, representing the local EPD height, near the spatial frequency k.

The function $h(k)$ is highly non-uniform when $k/k_0 < 2$ for both polarizations, and for p-polarized light we observe negative h values for opening angles of 30° and 60° in the intervals, where the TFs have a positive slope.
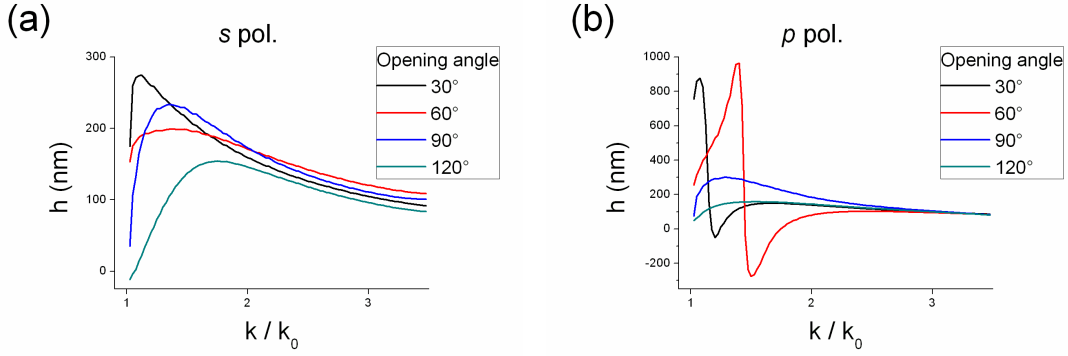
Fig. 5.14: The local effective plane of detection height is shown for *s* polarization (a) and *p* polarization (b).

When $k/k_0 > 2$, the curves are nearly independent of opening angle and, while still not constant, they vary much less than in the previous regime. The curves are generally slowly decreasing with average value of ~ 125 nm for both polarizations in the interval $2 < k/k_0 < 3.5$. We note that this value agrees well with the EPD height determined when comparing the experimental and modeled data of Ref. 40.

The EPD approximation which assumes a k independent effective plane of detection is intuitively appealing, but Fig. 5.14 reveals that the approximation is not ideal as it only holds for spatial frequency components of $k/k_0 > 2$. This means that it can be used only when prior knowledge of the field distribution is available, knowledge that allows us to rule out the existence of components with $k/k_0 < 2$. Furthermore, even when restricting the approximation to $k/k_0 > 2$, the effective detection height h is still not exactly constant but continues to decrease with k.

In many experiments however, as when imaging the propagation of surface plasmon polaritons[10] or the guided modes in photonic crystal structures[11], the spectrum of the optical field distribution is limited to a very narrow interval in k space and light of only one polarization is present. In this case the SNOM image does indeed reflect the intensity distribution in an effective detection plane and, if knowledge of the average spatial frequency is available, the z coordinate of this plane can be determined from Fig. 5.14.

5.6 Fiber tip sharpness

To examine the importance of employing a sharp fiber tip, we investigated the TFs of fiber tips of opening angle $\beta = 90°$ having their outer end cut off as illustrated in Fig. 5.15.
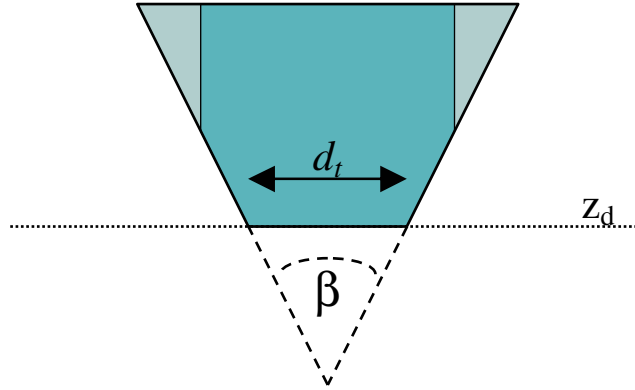
Fig. 5.15: The cut-off tip geometry is illustrated. The end of the cut-off tip has diameter d_t.

The detection plane is then raised to the end of the cut-off tip. The TFs for varying tip cut-off diameters d_t are illustrated in Fig. 5.16.
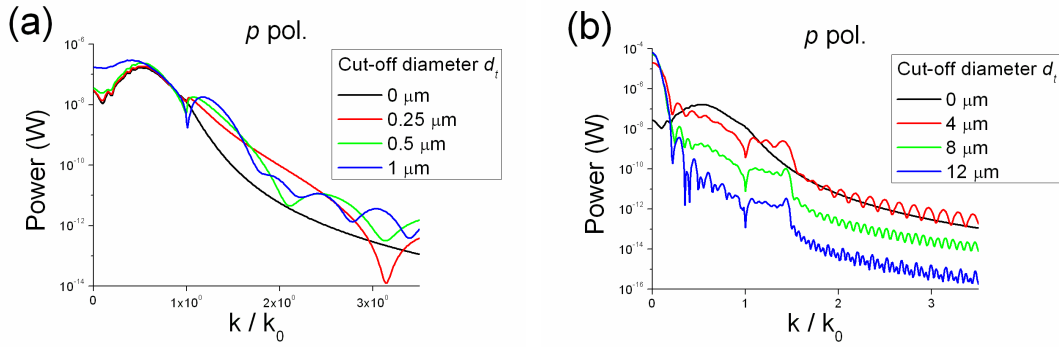
Fig. 5.16: Transfer functions for p polarization for tips of fixed opening angle β of 90° with varying cut-off diameter d_t.

For $d_t = 0$ μm, nothing has been cut off and the TF is that of Fig. 5.12(b) for $\beta = 90°$. For $d_t > 0$ μm the curves feature a complex oscillatory behavior. However, for the $d_t \leq 4$ μm curves, the average power level is identical to or slightly better than that of the uncut tip. This indicates that a perfectly sharp fiber tip results in a near-linear TF curve, where the EPD approximation may be used. But, surprisingly, a perfectly sharp tip is not necessary to achieve good coupling in the evanescent regime. For

80

example, considering the detection of evanescent waves with relatively low spatial frequencies $1 < k/k_0 < 2.5$, the tip cut with $d_t = 0.25$ μm is notably more efficient than the sharp one and preferential over other cut tips due to a monotonous decay of the transfer function.

However, we observe an overall drop in the power level when d_t is increased beyond 4 μm. Also, for cut tips with d_t above ~ 1 μm we notice an improved coupling of propagating waves with k/k_0 below ~ 0.4. The amplification of propagating modes in the SNOM image is inconvenient as they will dominate over the evanescent field of interest. We conclude that even though the tip need not be perfectly sharp, the conical part of the tapered fiber beyond $d_t = 1$ μm improves the near-field image and should not be omitted.

5.7 2D vs. 3D calculations

We have presented TFs determined using a 3D model. Full 3D vectorial computations are usually very demanding, whereas the numerical modeling of a 2D geometry in general requires only modest computing power. For this reason people often resort to simplified 2D simulations, even though results obtained from a 2D calculation do not necessarily hold for a true 3D geometry.

In the case of the TF, it is of interest to compare TFs calculated using the full 3D model with TFs obtained for a simplified 2D geometry. The 2D tip we will consider has the profile along the *x* axis depicted in Fig. 5.1, but is uniform along the *y* axis.

With uniformity along the *y*-axis the polarizations split, and we can perform computations for each polarization separately. This means that only about half the number of modes is required to obtain convergence compared to the 3D cylindrically symmetric geometry. The 2D calculations presented below were performed using the plane-wave expansion code with periodic boundary conditions.

In Fig. 5.17 the TFs calculated using the 2D code are compared to those obtained using the full 3D code.
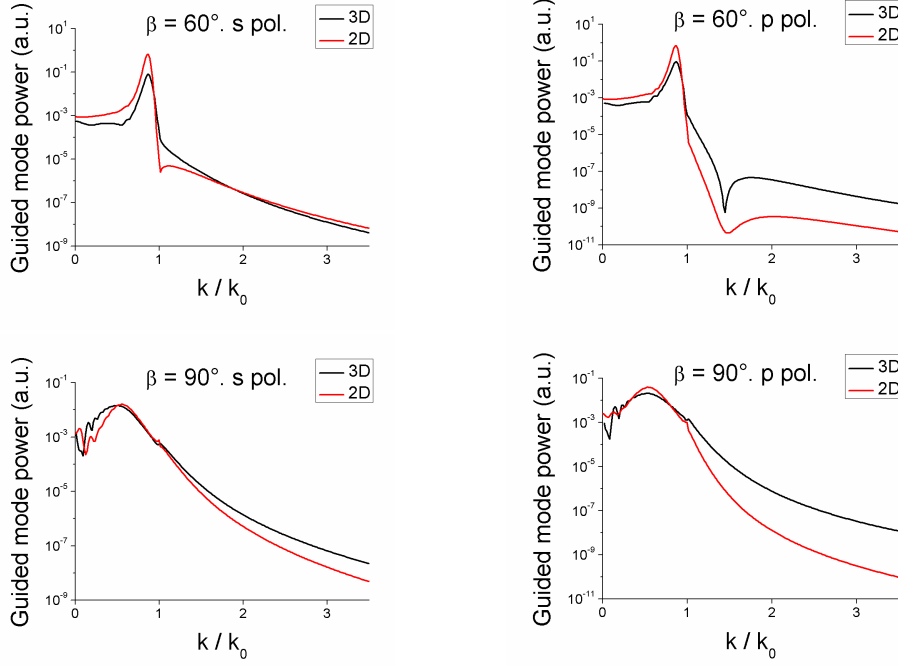
Fig. 5.17: The TFs calculated using the 2D and full 3D code are shown for opening angles β of 60° and 90°.

We observe that the shape of the 2D TFs is very similar to the 3D TF curves, when $k/k_0 < 1$. In the evanescent regime, the 2D curves tend to decay faster than the full 3D curves, in particular for *p* polarization.

However, a reasonable qualitative correspondence is observed indicating that even though 3D computations are necessary to obtain the exact TF, fast 2D calculations may be useful to reveal its general shape.

5.8 Comparison with experiment

In the topography identification described in chapter 4, the SNOM image is assumed to be reflecting the free-space field intensity distribution at the effective plane of detection. Now, we have determined the TF for the uncoated fiber tip used to obtain the experimental SNOM images and the above assumption is no longer necessary. We can introduce the TF into the inversion routine and take advantage of our knowledge of the relation between the free space optical field and the SNOM image to improve the quality of the topography identification.

Before implementing the TF into the inversion routine, we start by performing a test simulation of the experimental measurement sketched in Fig. 4.2. The geometry of

the grating was determined separately[55] using AFM and optical diffraction microscopy, and exact knowledge of the grating profile is thus available. Of course, we will in general not have prior knowledge about the topography of the sample under study, but for now we are only interested in verifying that our numerically determined TF does indeed describe the field-image relation correctly.

When the grating geometry, the angle of incidence of light, the TF of the tip and all other relevant parameters are known, we can perform a simulation of the experimental SNOM measurement.
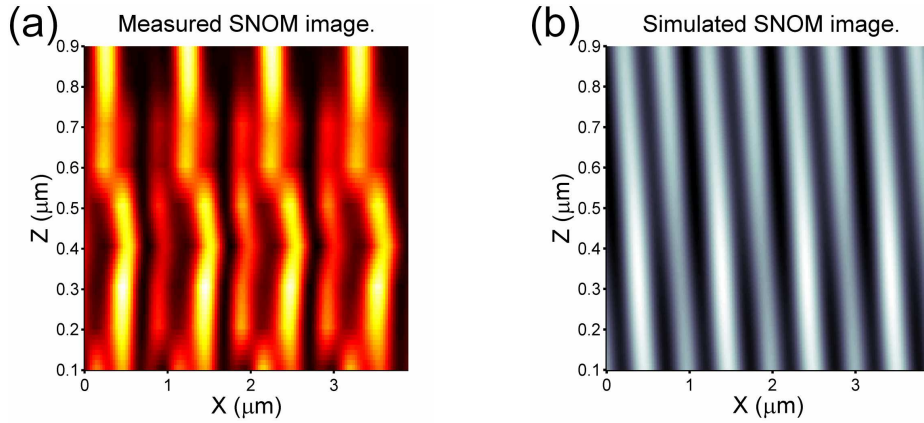
Fig. 5.18: The experimentally measured (a) and the simulated (b) SNOM images are depicted. The z coordinate is that of the tip apex. $\beta = 24°$.

The experimentally obtained SNOM image is compared with the result of the numerical modeling in Fig. 5.18. With exact knowledge of all parameters, we would expect to obtain good correspondence between theory and experiment. However, even though some agreement can be observed for the high-intensity lines in Fig. 5.18, the general deviation between the experimental and the modeled SNOM image is still somehow disappointing. As modeling of the scattering of light on a grating using the EET is well-established, we suspect that the reason for the discrepancy is related to the TF.

An experimental characterization of the TF was performed[56] by Ilya P. Radko at Aalborg University. In this work, a fiber tip was illuminated with laser light of wavelength 633 nm and total power of ~ 1.5 mW and the power of the guided fiber mode was then detected.
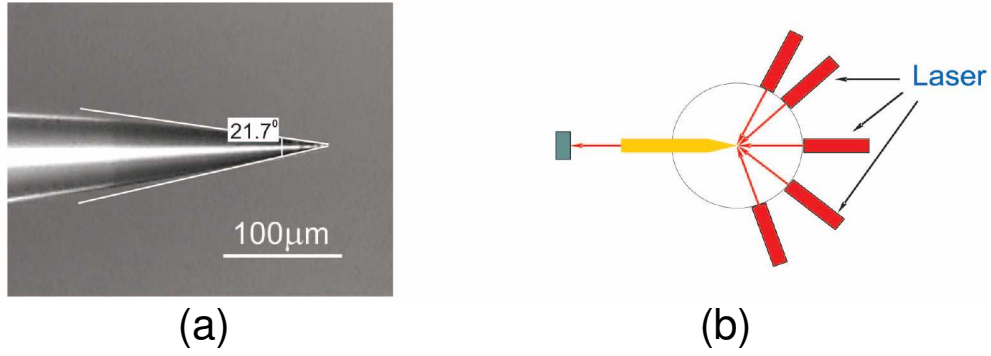
Fig. 5.19: The fiber tip used in the TF measurements is shown in (a) and the experimental setup is sketched in (b). Illustrations by Ilya P. Radko.

The fiber tip which was used in the measurement is illustrated in Fig. 5.19(a). It features a slightly smaller opening angle, but should otherwise be identical to the one used in the topography identification. On the scale of the fiber tip apex where the coupling takes place, the Gaussian laser beam can be considered a plane wave.

The experimental setup used in the far-field measurements is illustrated in Fig. 5.19(b). By changing the angle of incidence of the illumination beam the far-field measurements allow a determination of the TF in the interval $k / k_0 \leq 1$. Near-field measurements where the tip was illuminated with evanescent waves having $k / k_0 > 1$ were also performed, but since the near field above the grating is dominated by its propagating components we focus on the far field measurement.
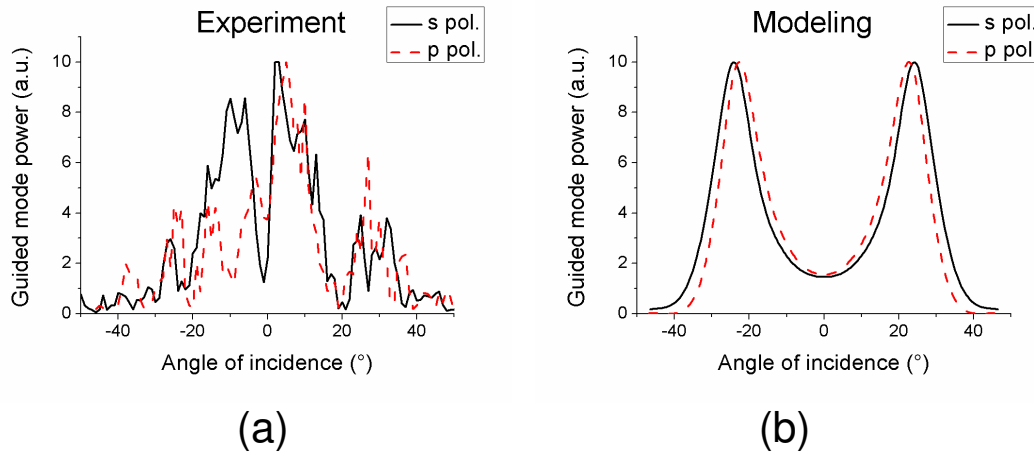
Fig. 5.20: The experimentally measured (a) and the simulated (b) TFs are depicted. $\beta = 22°$.

The experimentally measured TF in the far-field setup is given in Fig. 5.20(a). First, we notice that the curves for both polarizations are not symmetric, and this suggests that the tip shape is not perfectly conical. Second, even though we observe two major

84

peaks centered at approximately ± 10°, the spectrum is generally dominated by fast oscillations. On the other hand, the modeled TF shown in Fig. 5.20(b) is symmetric and does not feature the same rapid oscillations. One possible explanation for the lack of fast oscillations in the modeling is an error in the numerical computation, but since the asymmetry of the experimental curves already indicate that the tip shape is not ideal, it is more likely that the oscillations are due to imperfections of the tip shape near the outer end of the tip, where the coupling takes place.

In the experiment, repeated measurements on the same tip were performed giving identical results which suggests that the measurement setup was working properly. However, measurements were also performed on different tips fabricated using the same etching procedure.[38] Variations were observed in the measured TFs, and this supports our suspicion that the etching process is not perfect, but produces random imperfections in the tip shape.
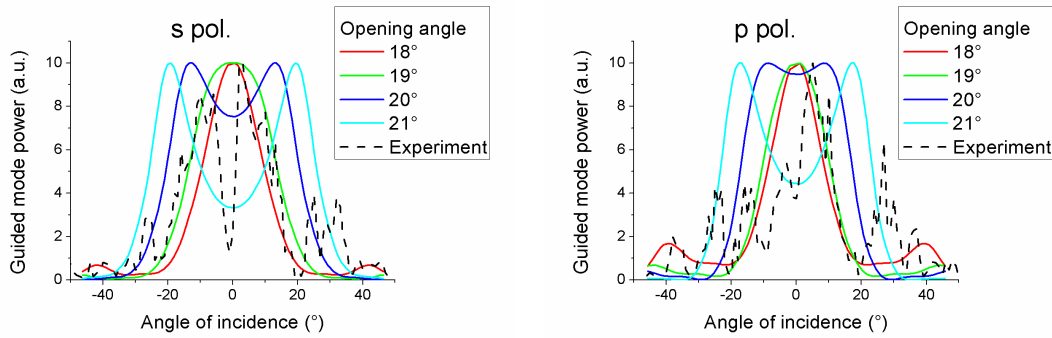
Fig. 5.21: TFs for varying opening angle β are depicted for *s* and *p* polarization.

The dependence of the TF on small variations of the tip opening angle β is shown in Fig. 5.21, and it is apparent that the TF is sensitive to a change of the order of just a few degrees. Even though it is hard to compare the shape of the experimental curves with the modeled ones due to asymmetry and fast oscillations, a rough estimate is that the tip opening angle at the very apex is ~ 18-19° rather than the 21.7° indicated in Fig. 5.19(a), which is on a scale of hundreds of microns.

It is thus possible that the discrepancy between experiment and modeling observed in Fig. 5.18 could be caused by the fiber tip used in the measurement having an opening angle near the apex different from the assumed value of 24°. To examine this possibility simulations of SNOM images obtained using tips with opening angles

between 15° and 30° were performed, but the agreement between experiment and modeling did not improve significantly for any angle in this interval.

As Fig. 5.21 reveals that the TF is very sensitive to the exact shape of the fiber tip profile, the most plausible explanation for the discrepancy between experiment and modeling seems to be the influence of imperfections in the fiber tip shape on the TF. These imperfections perturb the TF to an extent where it can no longer be properly modeled using a simple conical geometry model for the fiber tip and the experimental SNOM image thus cannot be reproduced in our simulation.

5.9 Discussion

The fact that arbitrarily small field variations are not observed in the SNOM image can be roughly understood by introducing the concept of resolution, where the TF is considered a step-function, constant for $k < 2\pi / r$, where r is the resolution limit, and zero otherwise. The SNOM images of delta-functions, however, demonstrate that the SNOM image does not faithfully depict the optical field intensity and that the concept of resolution cannot alone explain the field-image relation.

From the calculated TFs it is clear that, in the evanescent regime, the curves are better approximated to an exponential decay characteristic for an effective plane of detection above the tip apex than to the constant obtained, if the field was effectively coupled at the very apex. Unfortunately, the EPD approximation is only valid for perfectly sharp fiber tips in certain spatial frequency regimes.

The limitations of the concept of resolution and of the EPD approximation suggest that, to interpret the SNOM image correctly, a more general approach is needed: We can consider the SNOM image detected for a given optical field as the result of operating on the field with a coupling operator $\hat{H}$. If this operator is invertible, we can operate with its inverse $\hat{H}^{-1}$ on the SNOM image to obtain the optical field. From Eq. (5.5) we observe that, in the general case, the operator $\hat{H}$ is invertible if the complex functions $A_x(\mathbf{r}_\perp)$ and $A_y(\mathbf{r}_\perp)$ are known, and we thus require that our SNOM microscope can detect phase and power of the two linearly polarized guided modes (A_x, A_y) individually. If this is possible, the components of the optical field are given by the inverse Fourier transform of Eq. (5.5):

$$\begin{bmatrix} F_s(k,\theta) \\ F_p(k,\theta) \end{bmatrix} = \frac{1}{4\pi^2} \begin{bmatrix} -\sin\theta/H_s(k) & \cos\theta/H_s(k) \\ \cos\theta/H_p(k) & \sin\theta/H_p(k) \end{bmatrix} \int \begin{bmatrix} A_x(\mathbf{r}_\perp) \\ A_y(\mathbf{r}_\perp) \end{bmatrix} \exp(-i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)d\mathbf{r}_\perp . \quad (5.6)$$

Eq. (5.6) shows that all information about the vectorial electric field in the reference plane is available if we can measure $\left(A_x(\mathbf{r}_\perp), A_y(\mathbf{r}_\perp)\right)$ and we have prior knowledge of the functions $H_s(k)$ and $H_p(k)$. If the field is uniform along the y axis, the integrals in Eq. (5.6) vanish for angles θ different from 0 (mod π) and the polarizations are separated. In this particular case, we can reconstruct the field without individual measurements of $A_x(\mathbf{r}_\perp)$ and $A_y(\mathbf{r}_\perp)$.

Theoretically, the functions $H_s(k)$ and $H_p(k)$ can be completely arbitrary. In practice there will be noise present, and we cannot use the inversion of Eq. (5.6) for spatial frequencies that couple to power levels below the noise floor. To overcome the limitation to resolution due to noise, a nano-collector with a near-constant amplitude of the TF is preferred: With a near-constant amplitude, a weak propagating background field will not be amplified in the SNOM image compared to field components of high spatial frequency, and the evanescent field of interest can then be studied.

However, the use of the inversion in Eq. (5.6) requires that we can model the coupling at the tip of the optical field to the guided fiber mode correctly. We can calculate the TF of a rotationally symmetric fiber tip, but the geometry of fiber tips produced experimentally deviates from the perfect conical structure. This deviation causes the TF of the experimental tip to differ from the modeled one to an extent, where it does not make sense to implement the TF in the inversion procedure of the topography characterization. We conclude that, in order to analyze experimentally measured SNOM images correctly and perform topography characterizations of better quality, the discrepancy between the simulated and the experimentally determined TFs must be addressed. On the experimental side one can imagine the invention of better tip fabrication techniques that result in tip shapes closer to the perfect cone. As for the modeling one may consider the calculation of TFs for tip shapes without rotational symmetry. The numerical modeling of such tips is much more demanding than that of rotationally symmetric structures, but with the rapid increase of computation power of modern computers, such calculations may be possible in the near future.

6 Local density of states in micropillars

The possibility of modifying the spontaneous emission rate in a solid state material and thereby controlling light emission was first suggested by Yablonovitch[57]. He introduced the photonic crystal, a structure with a periodic refractive index profile, and pointed out that the photonic crystal can feature a photonic bandgap, which is frequency interval where no light can propagate, in the same way that a crystalline material can exhibit an electronic bandgap.

Of particular interest in quantum optics is the micro-cavity which may be created by introducing a defect into the photonic crystal. At the defect, a mode with a frequency inside the photonic bandgap may exist. The bandgap, however, prevents the light from propagating into the photonic crystal and the mode is thus localized to the defect. When a quantum dot (QD) is placed at the defect, strong coupling effects between the QD and the localized mode can be achieved.

The quantity governing the coupling between the QD and the optical field is the local density of states (LDOS). In this brief chapter, we examine the possibility of calculating the LDOS of a micropillar.

6.1 Local density of states

In the weak coupling regime, the spontaneous emission rate Γ at the position $\mathbf{r}$ and frequency ω is given by Fermi's golden rule[58]:

$$\Gamma = \frac{\pi}{2\hbar^2}|M|^2 N(\mathbf{d}, \omega, \mathbf{r}),\qquad(6.1)$$

where M is the dipole matrix element representing the atomic transition, $\mathbf{d}$ is its orientation and $N(\mathbf{d}, \omega, \mathbf{r})$ is the LDOS defined by:

$$N(\mathbf{d}, \omega, \mathbf{r}) = \sum_\alpha \delta(\omega - \omega_\alpha)|\mathbf{d} \cdot \mathbf{E}_\alpha(\mathbf{r})|^2.\qquad(6.2)$$

In this equation, the summation is made over all solutions to Eq. (3.1) for all frequencies. The spontaneous emission rate is thus proportional to the LDOS and we observe in Eq. (6.2) that, if a periodicity in the geometry results in a frequency

interval with no solutions to Eq. (3.1), the LDOS will be zero and light emission will be inhibited.

The theory of spontaneous emission and LDOS is well-developped[59]. However, the numerical computation of the LDOS involves the determination of all modes at a given frequency, and in complex structures this poses a challenge. In a photonic crystal without defects, due to periodicity, only a single unit cell needs to be considered and the calculation of the LDOS can be performed[60,61]. When the periodicity is broken by a defect, the geometry can be treated using a super-cell approximation[61], but for the approximation to be valid the computation must be performed on a large cell, and this can be very numerically demanding.

6.2 The micropillar

In the following we aim at computing the LDOS in a micropillar.[62] The geometry is a rotationally symmetric stack of alternating layers of GaAs and AlAs on top of a GaAs substrate. The stack consists of two distributed Bragg reflectors (DBRs) surrounding a micro-cavity, and the structure is very similar to a VCSEL except for the diameter of the pillar being much smaller. The geometry is illustrated below:
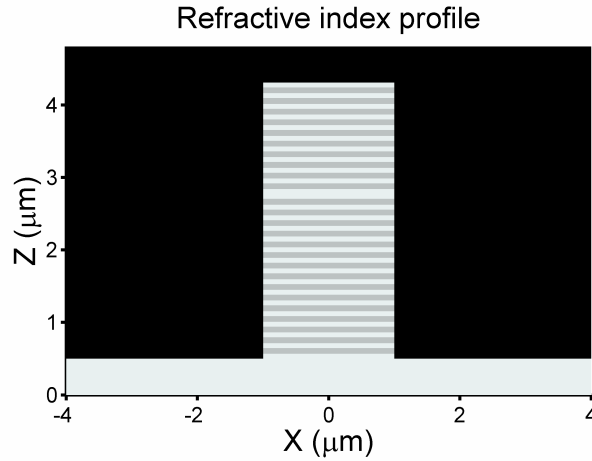

Fig. 6.1: The index profile of the micropillar.

If a QD is placed in the center of the resonator, there is a high probability that a photon generated by the de-excitation of the QD will couple to a cavity mode. And if the top DBR is thinner than the bottom reflector, the photon will be emitted through the top. One can improve the Q-factor of the stack by increasing the number of layer

pairs in the DBRs and thereby increase the probability that the photon couples to a cavity mode. In a realistic structure, however, there will be a material loss, and if the DBRs become too thick, the photon will be absorbed before leaving the micropillar. Now, if the LDOS can be determined using numerical modeling, the optimal balance between DBR reflectivity and material losses can be found such that the emission efficiency of the micropillar is maximized. And this is our motivation for calculating the LDOS in the micropillar.

6.3 Mode numeration

In the following we assume for simplicity that the structure is lossless. Now, to calculate the LDOS at a given point, we must determine all modes $\mathbf{E}_\alpha$ that are solutions to Eq. (3.1). We can define a mode profile $\mathbf{E}_\alpha$ by illuminating the micropillar from the bottom with a single eigenmode as illustrated in the figure below:
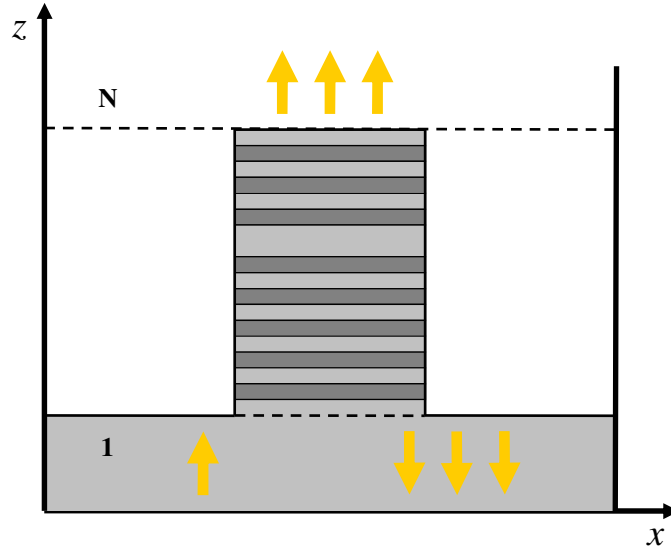
Fig. 6.2: The illumination of the micropillar from below using a single eigenmode is illustrated.

The incoming eigenmode will be reflected at the base of the micropillar, and in layer 1 the backwards propagating field will be a superposition of eigenmodes of layer 1. On the other side of the pillar, the field in layer N will be a superposition of eigenmodes of layer N. The entire mode profile, however, is completely determined by the label $\alpha = (l, h_{l,n}, \beta)$ of the illuminating eigenmode, where the labels l, $h_{l,n}$ and β refer to its angular momentum, the norm of its in-plane k vector and its propagation

constant respectively. $h_{l,n}$ takes a discrete set of values so that the boundary conditions are fulfilled, β is positive and the relation $h_{l,n}^2 + \beta^2 = \varepsilon_{r;GaAs}k_0^2$ holds. In a similar fashion, one can define a mode by illuminating the micropillar from above with the eigenmode specified by $\alpha = (l, h_{l,n}, \beta)$, where β is negative and $h_{l,n}^2 + \beta^2 = k_0^2$ is valid. In this way, the label $\alpha = (l, h_{l,n}, \beta)$ specifies all modes of the micropillar geometry and we can write the normalization condition as:

$$\int \varepsilon_r \mathbf{E}_\alpha^* \cdot \mathbf{E}_{\alpha'} d\mathbf{r} = \delta_{ll'}\delta_{nn'}\delta(\beta - \beta') . \tag{6.3}$$

We should remark that, to keep the modes physical, we only illuminate the micropillar with propagating waves and thus require that $\beta^2 > 0$.

6.4 Convergence

As it was the case for the TFs, we would ideally like to perform a computation for a micropillar in free space, and, as previously, we are forced to enclose the geometry in a metal cylinder due to the limitations of the EET. The metal cylinder will influence the LDOS, however, when the radius of the metal cylinder increases we expect the value of the LDOS to converge.

Before examining the micropillar, we study the influence of the metal cylinder on the LDOS of an empty vacuum geometry without micropillar. In the following we consider a dipole orientation **d** directed along the *r* axis and we calculate the LDOS at the point *r* = 0 so that calculations need to be performed for only one angular momentum.

The LDOS, normalized to its free-space value, is given in Fig. 6.3 as function of metal cylinder radius L_r at the wavelength λ = 950 nm. We observe that the LDOS converges towards unity, and a precision better than 1 % is obtained immediately. This indicates that we may not have to worry about the influence of the metal cylinder walls on the LDOS.
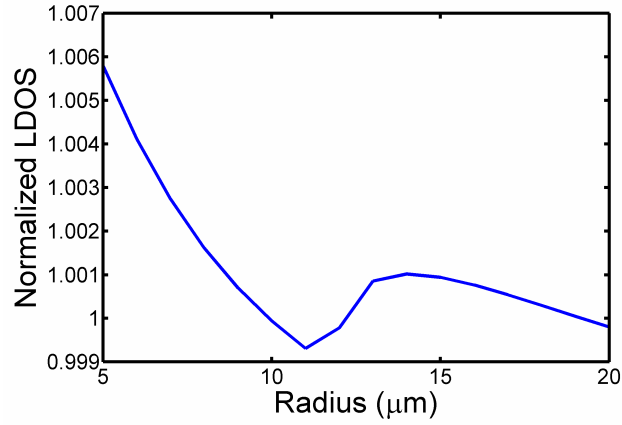
Fig. 6.3: The LDOS of the vacuum geometry as function of L_r is depicted. λ = 950 nm.

We now take a look at the micropillar. The structure we will examine has 10 and 15 layer pairs in the top and bottom DBR, respectively, and refractive indices of 3.5 and 3 are chosen for GaAs and AlAs, respectively. The structure is designed to be resonant at 950 nm and the LDOS is determined at the center of the cavity.

(a)

(b)
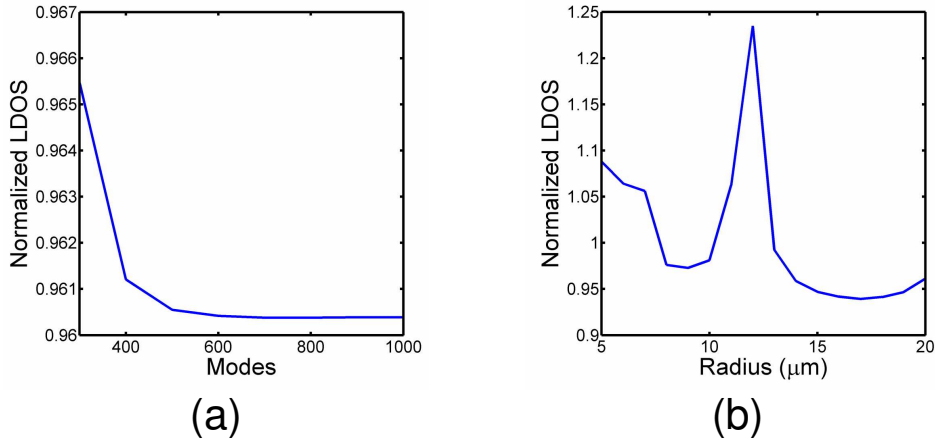
Fig. 6.4: The LDOS of the micropillar as function of included modes (a) and as function of the metal cylinder radius L_r (b). In (a) L_r = 20 μm and in (b) 400 modes are used. λ = 950 nm.

In Fig. 6.4(a), the LDOS as function of included modes in the computation is depicted and we observe that about 400 modes are required to achieve an accuracy of 1 ‰. The influence of the metal cylinder radius L_r on the LDOS can be studied by inspecting Fig. 6.4(b). The normalization here is made to the LDOS of bulk GaAs, and since the LDOS in the cavity is perturbed by the micropillar geometry, we should not expect the normalized LDOS to equal unity. We observe that the influence of the

93

metal wall causes a deviation of the LDOS of ~ 20 % from its average value and it is not clear that convergence is achieved.

6.5 Discussion

The EET can readily be used to compute the LDOS in a micropillar enclosed in a metal cylinder. We are, however, interested in the LDOS of a micropillar in free space, and convergence of the LDOS when the cylinder radius is increased is required, to be confident that the LDOS is not influenced by the metal walls.

In our initial calculation, we do not observe this convergence and an unsatisfying uncertainty of ~ 20 % is obtained. To improve the precision we can of course extend the cylinder radius beyond 20 μm and re-examine convergence. Since computations become more demanding for increased cylinder radius, this is the "brute force" method. However, another more elegant option is to introduce advanced boundary conditions. In other geometries, the implementation of PMLs has proven effective[22-24] in reducing the reflections from the boundary, and it is thus likely that the implementation of PMLs in our micropillar geometry can improve convergence without increasing the cylinder radius.

Due to time constraints, however, these possibilities will not be explored until after the submission deadline of this Ph.D. thesis.

7 Conclusion

This thesis summarizes my Ph.D. work on developing optical methods for performing characterization of structures on a subwavelength scale.

The numerical method chosen in this work to simulate the scattering of light on microscopic structures was the EET. The theory behind the method has been given, multiple techniques for solving the eigenmode problem have been presented, and the strengths and weaknesses of each technique have been discussed. The EET has proven effective in calculating the TF of an uncoated fiber tip, but the high quality simulation of the optical field possible by using the semi-analytical approach for solving the eigenmode problem relies on the correct determination of the propagations constants. It seems possible to model metal-coated fiber tips, however, the determination of propagation constants in zones with complex refractive index profiles is very cumbersome. If an efficient method for finding propagation constants is developed, reasonably fast simulations of TFs of metal coated fiber tips should be possible and the properties of the TFs of these probes can be explored.

The sample subject to subwavelength characterization in this work was a deep grating, and an optimization technique to identify the grating topography was tested. The technique becomes cumbersome when analyzing a geometry modeled using a large number of free parameters, but if the number of parameters is limited, the technique works well. However, it leaves us with the problem of establishing bijectivity of the geometry-image mapping. The illumination of samples with different topographies can result in the same optical field distribution, and the optimization technique is only efficient when we possess an understanding of the measurements necessary to establish a one-to-one correspondence between object and image. In this work, an optically equivalent topography, which is similar, but not identical to the true sample topography has been identified.

Furthermore, the calculation of the TF of SNOM fiber tips has been presented. The knowledge obtained of the TF has been used to perform an in-depth analysis of the relation between the free-space optical field and the measured SNOM image. When working with uncoated fiber tips, we have learned that the detected SNOM image only represents the optical field intensity distribution in an effective plane of detection in a few particular cases. In general, an inversion operation must be performed to

reconstruct the optical field. This reconstruction requires exact knowledge of the transfer function of the fiber tip and the capability to measure phase and power of the two orthogonal modes of the single-mode fiber individually. To extend the achievable resolution, not only good coupling, but also a flat amplitude of the transfer function is preferred. Our simulations indicate that for uncoated fiber tips at the wavelength of 633 nm, a tip having a ~ 90° degree opening angle is the optimal choice. The tip can be sharp, but an equally good coupling is obtained for a cut-off tip with diameter d_t up to 1 μm.

The TF was determined for an uncoated conical fiber tip with an opening angle identical to that of the fiber tip used in the experimental SNOM measurements. However, the tip used in the experiment deviates from the ideal conical profile, and this deviation prevented us from reproducing the experimental SNOM images in our simulation. As the reconstruction of the free-space optical field requires exact knowledge of the TF, we concluded that the imperfections in the realistic tip profile pose an obstacle to the correct interpretation of the SNOM images and that better tip fabrication techniques are required to improve the correspondence between simulation and experiment.

Finally, the possibility of computing the LDOS of micropillars was briefly touched upon. Even though parasitic reflections from metal walls clearly influence the LDOS in an initial calculation, the EET still seems to be an efficient tool in evaluating the LDOS in a micropillar.

8 Appendix

8.1 Mode profiles

In this section, the analytic expressions for the mode profiles used in the plane-wave expansion technique and in the semi-analytical approach are given. These solutions to the wave equation are valid for uniform refractive index profiles. To obtain the profiles we consider the wave equation for the E_z and H_z field components in the frequency domain:

$$\left(\nabla_\perp^2 + k^2\right)\begin{bmatrix} E_z(\mathbf{r}) \\ H_z(\mathbf{r}) \end{bmatrix} = 0 . \tag{8.1}$$

Here $k^2 = \varepsilon_r k_0^2$, and the time dependence $\exp(-i\omega t)$ is assumed. The z dependence is separated using $F_z(\mathbf{r}) = f_z(\mathbf{r}_\perp)\exp(i\beta z)$, and Eq. (8.1) becomes:

$$\left(\nabla_\perp^2 + k^2 - \beta^2\right)\begin{bmatrix} e_z(\mathbf{r}_\perp) \\ h_z(\mathbf{r}_\perp) \end{bmatrix} = 0 . \tag{8.2}$$

This equation allows two classes of solutions: The transverse electric (TE) modes having E_z equal to zero and the transverse magnetic (TM) modes with H_z equal to zero. Eqs. (8.1) and (8.2) only describe z components of the field, however, when these are known the remaining four field components can be calculated using the relations:

$$\begin{bmatrix} \mathbf{e}_\perp(\mathbf{r}_\perp) \\ \mathbf{h}_\perp(\mathbf{r}_\perp) \end{bmatrix} = \frac{i}{k^2 - \beta^2}\begin{bmatrix} \beta\nabla_\perp e_z - k(\mu_0/\varepsilon)^{\frac{1}{2}}\mathbf{u}_z \times \nabla_\perp h_z \\ \beta\nabla_\perp h_z + k\varepsilon_r(\varepsilon/\mu_0)^{\frac{1}{2}}\mathbf{u}_z \times \nabla_\perp e_z \end{bmatrix}, \tag{8.3}$$

that are obtained by manipulation of Maxwell's equations.

The analytic solutions in Cartesian and cylindrical coordinates are given below. In both coordinate sets, we consider the boundary condition of a perfectly conducting metal wall requiring that the tangential components of the electric field should be zero. In Cartesian coordinates, we also consider the periodic boundary condition.

8.1.1 Mode profiles in Cartesian coordinates

If we express the Laplace operator in Cartesian coordinates, Eq. (8.2) becomes:

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+k^2-\beta^2\right)\begin{bmatrix}e_z(\mathbf{r}_\perp)\\h_z(\mathbf{r}_\perp)\end{bmatrix}=0,\tag{8.4}$$

and its solutions are products of exponential, sine and cosine functions. The exact nature of the solutions depends on the boundary conditions.

Solutions in a geometry limited by metal walls at $x=0$, $x=L_x$, $y=0$ and $y=L_y$ are given in Table 8.1 below:

Field component	TE modes	TM modes
e_x	$-i\omega\mu_0/h^2 b\cos(ax)\sin(by)$	$i\beta/h^2 a\cos(ax)\sin(by)$
e_y	$i\omega\mu_0/h^2 a\sin(ax)\cos(by)$	$i\beta/h^2 b\sin(ax)\cos(by)$
e_z	0	$\sin(ax)\sin(by)$
h_x	$-i\beta/h^2 a\sin(ax)\cos(by)$	$-i\omega\varepsilon_0/h^2 b\sin(ax)\cos(by)$
h_y	$-i\beta/h^2 b\cos(ax)\sin(by)$	$i\omega\varepsilon_0/h^2 a\cos(ax)\sin(by)$
h_z	$\cos(ax)\cos(by)$	0

Table 8.1: TE and TM modes in a rectangular vacuum geometry with metal walls.

To respect the boundary conditions, the conditions $a=n_x\pi/L_x$ and $b=n_y\pi/L_y$ apply with n_x and n_y being positive integers. Either n_x or n_y may be zero, but not both at the same time. The relation $k^2-\beta^2=a^2+b^2\equiv h^2$ also holds.

When using the periodic boundary condition stating that $\mathbf{e}(\mathbf{r}_\perp+n_x L_x\mathbf{u}_x+n_y L_y\mathbf{u}_y)=\mathbf{e}(\mathbf{r}_\perp)$, a field component of a solution is usually chosen to be a travelling wave of the form $\exp(iax+iby)$. However, in the plane-wave expansion technique, complex basis modes lead to a complex-valued operator matrix, which is computationally much more demanding to solve than a real-valued matrix. For this reason we employ sine and cosine functions as basis set when using periodic boundary conditions, even though this leads to an increased amount of book-keeping. The in-plane electric field components of the basis modes used when the boundary conditions are periodic are listed below:

Field component	TE modes			
	Set 1	Set 2	Set 3	Set 4
e_x	$\dfrac{i\beta}{h^2}a\cos(ax)\cos(by)$	$\dfrac{i\beta}{h^2}a\cos(ax)\sin(by)$	$-\dfrac{i\beta}{h^2}a\sin(ax)\cos(by)$	$-\dfrac{i\beta}{h^2}a\sin(ax)\sin(by)$
e_y	$-\dfrac{i\beta}{h^2}b\sin(ax)\sin(by)$	$\dfrac{i\beta}{h^2}b\sin(ax)\cos(by)$	$-\dfrac{i\beta}{h^2}b\cos(ax)\sin(by)$	$\dfrac{i\beta}{h^2}b\cos(ax)\cos(by)$

	TM modes			
	Set 1	Set 2	Set 3	Set 4
e_x	$\dfrac{i\omega\mu_0}{h^2}b\cos(ax)\cos(by)$	$-\dfrac{i\omega\mu_0}{h^2}b\cos(ax)\sin(by)$	$\dfrac{i\omega\mu_0}{h^2}b\sin(ax)\cos(by)$	$-\dfrac{i\omega\mu_0}{h^2}b\sin(ax)\sin(by)$
e_y	$\dfrac{i\omega\mu_0}{h^2}a\sin(ax)\sin(by)$	$\dfrac{i\omega\mu_0}{h^2}a\sin(ax)\cos(by)$	$-\dfrac{i\omega\mu_0}{h^2}a\cos(ax)\sin(by)$	$-\dfrac{i\omega\mu_0}{h^2}a\cos(ax)\cos(by)$

Table 8.2: TE and TM modes in a rectangular vacuum geometry with periodic boundary conditions.

The conditions on a and b are $a = n_x 2\pi / L_x$ and $b = n_y 2\pi / L_y$ with n_x and n_y being positive integers. When they appear in a cosine function, n_x and n_y can also take the value of zero. As previously we have $k^2 - \beta^2 = a^2 + b^2 \equiv h^2$.

8.1.2 Mode profiles in cylindrical coordinates

Writing out the Laplace operator in cylindrical coordinates, Eq. (8.2) is given by:

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} + k^2 - \beta^2\right)\begin{bmatrix}e_z(\mathbf{r}_\perp)\\ h_z(\mathbf{r}_\perp)\end{bmatrix} = 0. \qquad (8.5)$$

We proceed by separating the angular dependence of the field using $h_z(\mathbf{r}_\perp) = h_z(r)\exp(il\theta)$ and obtain:

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{l^2}{r^2} + k^2 - \beta^2\right)\begin{bmatrix}e_z(r)\\ h_z(r)\end{bmatrix} = 0. \qquad (8.6)$$

When $k^2 - \beta^2 \equiv h^2 > 0$, the solutions to this equation are the Bessel functions $J_l(hr)$ and $Y_l(hr)$ of first and second kind. If $k^2 - \beta^2 \equiv -q^2 < 0$, the solutions are the modified Bessel functions $I_l(qr)$ and $K_l(qr)$ of first and second kind. Even though all four functions, depending on the sign of $k^2 - \beta^2$, are solutions, only the Bessel function $J_l(hr)$ fulfils the boundary conditions both at the center $r = 0$, where the field

should remain finite, and at the metal wall $r = L_r$. The following table lists the mode profiles:

Field component	TE modes	TM modes
e_r	$-\omega\mu_0/h^2(l/r)J_l(hr)$	$i\beta/h^2(\partial/\partial r)J_l(hr)$
e_θ	$-i\omega\mu_0/h^2(\partial/\partial r)J_l(hr)$	$-\beta/h^2(l/r)J_l(hr)$
e_z	0	$J_l(hr)$
h_r	$i\beta/h^2(\partial/\partial r)J_l(hr)$	$\omega\varepsilon/h^2(l/r)J_l(hr)$
h_θ	$-\beta/h^2(l/r)J_l(hr)$	$i\omega\varepsilon/h^2(\partial/\partial r)J_l(hr)$
h_z	$J_l(hr)$	0

Table 8.3: TE and TM modes in vacuum geometries in cylindrical coordinates with angular dependence given by exp($il\theta$).

The boundary condition requires that $J_l(hL_r) = 0$ and $(\partial/\partial r)J_l(hL_r) = 0$ for TM and TE modes respectively. Zero points of the functions $J_l(r)$ and $(\partial/\partial r)J_l(r)$ must thus be determined, and this is done using a standard search routine.

It is useful to also introduce the "sum and difference" coordinates defined by the transformations $f_+ \equiv (f_r + f_\theta/i)/\sqrt{2}$ and $f_- \equiv (f_r - f_\theta/i)/\sqrt{2}$. Using the Bessel recurrence relations $J_{l\pm 1} = (l/r \mp \partial/\partial r)J_l$, the solutions from Table 8.3 become:

Field component	TE modes	TM modes
e_+	$-\omega\mu/hJ_{l-1}(hr)/\sqrt{2}$	$i\beta/hJ_{l-1}(hr)/\sqrt{2}$
e_-	$-\omega\mu/hJ_{l+1}(hr)/\sqrt{2}$	$-i\beta/hJ_{l+1}(hr)/\sqrt{2}$
e_z	0	$J_l(hr)$
h_+	$i\beta/hJ_{l-1}(hr)/\sqrt{2}$	$\omega\varepsilon/hJ_{l-1}(hr)/\sqrt{2}$
h_-	$-\beta/hJ_{l+1}(hr)/\sqrt{2}$	$\omega\varepsilon/hJ_{l+1}(hr)/\sqrt{2}$
h_z	$J_l(hr)$	0

Table 8.4: TE and TM modes in vacuum geometries in "sum and difference" cylindrical coordinates with angular dependence given by exp($il\theta$).

We have now gotten rid of the derivative and the factor $1/r$, and the solutions are now easier to work with.

8.2 Derivation of reflection and transmission matrices

Referring to Fig. 3.16, we require that the tangential components of the electric and magnetic fields are continuous at the interface. From Eqs. (3.31) and (3.32) we obtain:

$$\mathbf{e}_{I;\perp;m}(\mathbf{r}_\perp) + \sum_j \mathbf{e}_{I;\perp;j}(\mathbf{r}_\perp) R_{I,II;jm} = \sum_j \mathbf{e}_{II;\perp;j}(\mathbf{r}_\perp) T_{I,II;jm} \tag{8.7}$$

$$\mathbf{h}_{I;\perp;m}(\mathbf{r}_\perp) - \sum_j \mathbf{h}_{I;\perp;j}(\mathbf{r}_\perp) R_{I,II;jm} = \sum_j \mathbf{h}_{II;\perp;j}(\mathbf{r}_\perp) T_{I,II;jm} \ . \tag{8.8}$$

We now introduce the inner product given by:

$$\langle \mathbf{e}_m | \mathbf{h}_n \rangle = \int (\mathbf{e}_m \times \mathbf{h}_n) \cdot \mathbf{u}_z d\mathbf{r}_\perp \ . \tag{8.9}$$

In the following, we assume that our mode profiles have been normalized using this inner product. If our eigenmodes are non-degenerate, we then have from the orthogonality relation of Eq. (3.18) that $\langle \mathbf{e}_{q,m} | \mathbf{h}_{q,n} \rangle = \delta_{mn}$, where q is the layer index. Multiplying the bra version of Eq. (8.7) with $|\mathbf{h}_{I,n}\rangle$ and the ket version of Eq. (8.8) with $\langle \mathbf{e}_{I,n} |$ gives:

$$\delta_{mn} + R_{I,II;nm} = \sum_j \langle \mathbf{e}_{II;j} | \mathbf{h}_{I;n} \rangle T_{I,II;jm} \tag{8.10}$$

$$\delta_{mn} - R_{I,II;nm} = \sum_j \langle \mathbf{e}_{I;n} | \mathbf{h}_{II;j} \rangle T_{I,II;jm} \ . \tag{8.11}$$

Adding and subtracting these two equations, we obtain:

$$2\delta_{mn} = \sum_j \left(\langle \mathbf{e}_{II;j} | \mathbf{h}_{I;n} \rangle + \langle \mathbf{e}_{I;n} | \mathbf{h}_{II;j} \rangle \right) T_{I,II;jm} \tag{8.12}$$

$$R_{I,II;nm} = \frac{1}{2} \sum_j \left(\langle \mathbf{e}_{II;j} | \mathbf{h}_{I;n} \rangle - \langle \mathbf{e}_{I;n} | \mathbf{h}_{II;j} \rangle \right) T_{I,II;jm} \ . \tag{8.13}$$

In matrix form this can be written as:

$$\overline{\overline{T_{I,II}}} = 2\left(\overline{\overline{\langle \mathbf{e}_{II} | \mathbf{h}_I \rangle}}^T + \overline{\overline{\langle \mathbf{e}_I | \mathbf{h}_{II} \rangle}} \right)^{-1} \tag{8.14}$$

$$\overline{\overline{R_{I,II}}} = \frac{1}{2}\left(\overline{\overline{\langle \mathbf{e}_{II} | \mathbf{h}_I \rangle}}^T - \overline{\overline{\langle \mathbf{e}_I | \mathbf{h}_{II} \rangle}} \right) \overline{\overline{T_{I,II}}} \ . \tag{8.15}$$

These equations define the reflection and transmission matrices. The expressions for $\overline{\overline{R_{II,I}}}$ and $\overline{\overline{T_{II,I}}}$ are derived in a similar manner.

8.3 Expressions for the scattering matrices

Referring to Fig. 3.17 and the subsequent discussion, the total reflection of light incident from below by the three-layer structure is given by the sum:

$$SR_{I,III} = R_{I,II} + T_{II,I}P_{II}R_{II,III}P_{II}T_{I,II} + T_{II,I}P_{II}R_{II,III}P_{II}\left(R_{II,I}P_{II}R_{II,III}P_{II}\right)T_{I,II}$$
$$+ T_{II,I}P_{II}R_{II,III}P_{II}\left(R_{II,I}P_{II}R_{II,III}P_{II}\right)^2 T_{I,II} + ... \qquad (8.16)$$

where the product $R_{II,I}P_{II}R_{II,III}P_{II}$ describes a round trip in layer II, and the double bars have been omitted for simplicity. Using the relation $1 + A + A^2 + ... = (1-A)^{-1}$, valid for matrices with absolute eigenvalues below unity, the defining recursive relation becomes:

$$SR_{I,III} = R_{I,II} + T_{II,I}P_{II}R_{II,III}P_{II}\left(1 - R_{II,I}P_{II}R_{II,III}P_{II}\right)^{-1}T_{I,II}. \qquad (8.17)$$

As explained in section 3.5.2, the corresponding scattering matrix $SR_{1,q}$ for $q > 3$ can be determined simply by replacing the matrices $R_{I,II}$, $T_{I,II}$ etc. with $SR_{1,q-1}$, $ST_{1,q-1}$ etc. The matrix $ST_{1,q}$ is determined in a similar manner.

The defining recursive relations for all the scattering matrices used in the eigenmode expansion technique are given below:

$$ST_{1,q+1} = T_{q,q+1}P_q\left(1 - SR_{q,1}P_qR_{q,q+1}P_q\right)^{-1}ST_{1,q} \qquad (8.18)$$

$$SR_{1,q+1} = SR_{1,q} + ST_{q,1}P_qR_{q,q+1}P_q\left(1 - SR_{q,1}P_qR_{q,q+1}P_q\right)^{-1}ST_{1,q} \qquad (8.19)$$

$$ST_{q+1,1} = ST_{q,1}P_qT_{q+1,q} + ST_{q,1}P_qR_{q,q+1}P_q\left(1 - SR_{q,1}P_qR_{q,q+1}P_q\right)^{-1}SR_{q,1}P_qT_{q+1,q} \qquad (8.20)$$

$$SR_{q+1,1} = R_{q+1,q} + T_{q,q+1}P_q\left(1 - SR_{q,1}P_qR_{q,q+1}P_q\right)^{-1}SR_{q,1}P_qT_{q+1,q} \qquad (8.21)$$

$$ST_{n,q-1} = T_{q,q-1}P_qST_{n,q} + T_{q,q-1}P_qSR_{q,n}P_q\left(1 - R_{q,q-1}P_qSR_{q,n}P_q\right)^{-1}R_{q,q-1}P_qST_{n,q} \qquad (8.22)$$

$$SR_{n,q-1} = SR_{n,q} + ST_{q,n}P_q\left(1 - R_{q,q-1}P_qSR_{q,n}P_q\right)^{-1}R_{q,q-1}P_qST_{n,q} \qquad (8.23)$$

$$ST_{q-1,n} = ST_{q,n}P_q\left(1 - R_{q,q-1}P_qSR_{q,n}P_q\right)^{-1}T_{q-1,q} \qquad (8.24)$$

$$SR_{q-1,n} = R_{q-1,q} + T_{q,q-1}P_qSR_{q,n}P_q\left(1 - R_{q,q-1}P_qSR_{q,n}P_q\right)^{-1}T_{q-1,q} \qquad (8.25)$$

In the above expressions it is understood that $SR_{1,2} = R_{1,2}$, $ST_{1,2} = T_{1,2}$ etc.

Finally the expansion coefficients $\overline{a_q^+}$ and $\overline{a_q^-}$ describing the forward and backward propagating parts of the field in a layer q in a structure with a total of n layers are given by the expressions:

$$a_q^+ = (1 - SR_{q,1}P_q SR_{q,n}P_q)^{-1}(ST_{1,q}a_1'^+ + SR_{q,1}P_q ST_{n,q}a_n'^-) \quad (8.26)$$

$$a_q^- = (1 - SR_{q,1}P_q SR_{q,n}P_q)^{-1}(SR_{q,n}P_q ST_{1,q}a_1'^+ + ST_{n,q}a_n'^-) . \quad (8.27)$$

From these coefficients, the electric field is then given by:

$$\mathbf{E}_{q;\perp}(\mathbf{r}) = \sum_n \left(a_{q,n}^+ e^{i\beta_{q;n}(z-z_{q-1,q})} + a_{q,n}^- e^{-i\beta_{q;n}(z-z_{q,q+1})}\right)\mathbf{e}_{q;\perp;n}(\mathbf{r}_\perp) . \quad (8.28)$$

8.4 Derivation of Eq. (5.5)

Using bra-ket notation, we define $|k,\theta,s\rangle \equiv \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)\mathbf{e}_\theta$ and $|k,\theta,p\rangle \equiv \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)\mathbf{e}_k$ and Eq. (5.4) becomes:

$$\mathbf{E}(\mathbf{r}_\perp,z_d) = \int \left(F_s(k,\theta)|k,\theta,s\rangle + F_p(k,\theta)|k,\theta,p\rangle\right)d\mathbf{k}_\perp . \quad (8.29)$$

We now introduce the coupling operator $\hat{H}$ coupling the plane wave $|k,\theta,s\rangle$ to the guided mode $\mathbf{G}^{\pm 1}(r)\exp(\pm i\varphi)$ of angular momentum ± 1 of a single-mode fiber. The coupling coefficient is then of the form $\langle \pm 1|\hat{H}|k,\theta,s\rangle$. We then exploit that, due to cylindrical symmetry, the rotation operator $e^{-iJ_z\theta}$ and the coupling operator $\hat{H}$ commute:

$$\langle \pm 1|\hat{H}|k,\theta,s\rangle = \langle \pm 1|\hat{H}e^{-iJ_z\theta}|k,0,s\rangle = e^{\mp i\theta}\langle \pm 1|\hat{H}|k,0,s\rangle . \quad (8.30)$$

Similar expressions hold for p-polarization. We now define coupling coefficients $H_s^{\pm 1}(k) \equiv i\sqrt{2}\langle \pm 1|\hat{H}|k,0,s\rangle$ and $H_p^{\pm 1}(k) \equiv \sqrt{2}\langle \pm 1|\hat{H}|k,0,p\rangle$. The coupling of the total field to the fiber mode becomes:

$$A^{\pm 1}(\mathbf{r}_\perp) = \int \left(\frac{1}{i\sqrt{2}}H_s^{\pm 1}(k)F_s(k,\theta) + \frac{1}{\sqrt{2}}H_p^{\pm 1}(k)F_p(k,\theta)\right)\exp(\mp i\theta + i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)d\mathbf{k}_\perp . \quad (8.31)$$

This expression describes the coupling to the guided modes $\mathbf{G}^{\pm 1}(r)\exp(\pm i\varphi)$. However, for a weak index contrast between core and cladding, there exist linear combinations $\mathbf{G}_{x,y}(\mathbf{r}_\perp)$ of the modes $\mathbf{G}^{\pm 1}(r)\exp(\pm i\varphi)$, which, to a good approximation, are linearly polarized. We are therefore usually more interested in coupling to the modes $\mathbf{G}_{x,y}(\mathbf{r}_\perp)$, and we thus define $|x\rangle = (|1\rangle + |-1\rangle)/\sqrt{2}$ and $|y\rangle = (|1\rangle - |-1\rangle)/(i\sqrt{2})$. Now, as incoming s- and p-polarized light of angle $\theta = 0$ is directed along the y and x axis respectively, we have $\langle x|\hat{H}|k,0,s\rangle = 0$ and $\langle y|\hat{H}|k,0,p\rangle = 0$ and obtain the properties $H_s^1 = -H_s^{-1} \equiv H_s$ and $H_p^1 = H_p^{-1} \equiv H_p$ for the coupling coefficients. The coupling of the field to the mode profiles $A_{x,y}\mathbf{G}_{x,y}(\mathbf{r}_\perp)$ is then given by the expression:

$$\begin{bmatrix} A_x(\mathbf{r}_\perp) \\ A_y(\mathbf{r}_\perp) \end{bmatrix} = \int \begin{bmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} H_s(k)F_s(k,\theta) \\ H_p(k)F_p(k,\theta) \end{bmatrix} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)d\mathbf{k}_\perp . \qquad (8.32)$$

List of publications

This thesis is based on the following publications:

Journal papers:

N. Gregersen, B. Tromborg and S. I. Bozhevolnyi, "Vectorial modeling of near-field imaging with uncoated fiber probes: Transfer function and resolving power," http://arxiv.org/physics/0608291, to be published in Appl. Opt.

I. P. Radko, S. I. Bozhevolnyi and N. Gregersen, "Transfer function and near-field detection of evanescent waves", Appl. Opt. **45**, 4054-4061 (2006).

N. Gregersen, B. Tromborg, V. S. Volkov, S. I. Bozhevolnyi and J. Holm, "Topography characterization of a deep grating using near-field imaging," Appl. Opt. **45**, 117-121 (2006).

Other contributions:

N. Gregersen, B. Tromborg, V. S. Volkov and S. I. Bozhevolnyi, "Image interpretation and topography characterization using near-field microscopy," to be published in DOPS-Nyt (popular magazine of the Danish Optical Society).

N. Gregersen, "Transfer functions for uncoated SNOM fiber tips," poster presentation at the annual meeting of the Danish Optical Society, Risø National Laboratory, Roskilde, Denmark, November 2005.

N. Gregersen, "Vectorial modeling of transfer functions of fiber tips in collection SNOM," poster presentation at the NFO-9 conference, Lausanne, Switzerland, September 2006.

References

1. http://www.intel.com/technology/silicon/.

2. T. J. Lemaire and A. Bassrei, "Three-dimensional reconstruction of dielectric objects by the coupled-dipole method," Appl. Opt. **39**, 1272-1278 (2000).

3. M. Baribaud, "Microwave imagery: Analytical method and maximum entropy method," J. Phys. D **23**, 269-288 (1990).

4. S. Caorsi, G. L. Gragnani and M. Pasorino,"Redundant electromagnetic data for microwave imaging of three dimensional dielectric objects," IEEE Trans. Antennas Propag. **42**, 581-589 (1994).

5. E. H. Synge, "Suggested method for extending microscopic resolution into the ultra-microscopic region," Philosophical Magazine **6**, 356-362 (1928).

6. E. A. Ash & G. Nicholls, "Super-resolution aperture scanning microscope," Nature **237**, 510-513 (1972).

7. D. W. Pohl, W. Denk and M. Lanz, "Optical stethoscopy: Image recording with resolution $\lambda / 20$," Appl. Phys. Lett. **44**, 651-653 (1984).

8. A. Lewis, M. Isaacson, A. Harootunian and A. Muray, "Development of a 500 Å spatial resolution light microscope. I. Light is efficiently transmitted through $\lambda / 16$ diameter apertures," Ultramicr. **13**, 227-231 (1984).

9. D. Courjon, *Near-Field Microscopy and Near-Field Optics* (Imperial College Press, 2003).

10. B. Hecht, H. Bielefeldt, L. Novotny, Y. Inouye and D. W. Pohl, "Local Excitation, Scattering and Interference of Surface Plasmons", Phys. Rev. Lett. **77**, 1889-1892 (1996).

11. S. I. Bozhevolnyi, V. S. Volkov, T. Søndergaard, A. Boltasseva, P. I. Borel and M. Kristensen, "Near-field imaging of light propagation in photonic crystal waveguides: Explicit role of Bloch harmonics," Phys. Rev. B **66**, 235204 (2002).

12. S. Hudlet, S. Aubert, A. Bruyant, R. Bachelot, P.-M. Adam, J.-L. Bijeon, G. Lérondel, P. Royer and A. A. Stashkevich, "Apertureless near field optical microscopy: a contribution to the understanding of the signal detected in the presence of a background field," Opt. Commun. **230**, 245-251 (2004).

13. P. L. Phillips, J. C. Knight, J. M. Pottage, G. Kakarantzas, P. St. J. Russell, "Direct measurement of optical phase in the near field," Appl. Phys. Lett. **76**, 541-543 (2000).

14. A. Taflove and S. C. Hagness, *Computational Electromagnetics: The Finite-Difference Time-Domain Method, 3rd Ed.* (Artech House Publishers, 2005).

15. J. N. Reddy, *An Introduction to the Finite Element Method, 3rd Ed.* (McGraw-Hill Science/Engineering/Math, 2005).

16. O. J. F. Martin, C. Girard and A. Dereux, "Generalized Field Propagator for Electromagnetic Scattering and Light Confinement," Phys. Rev. Lett **74**, 526-529 (1995).

17. C. Hafner, *The Generalized Multipole Technique for Computational Electromagnetics* (Artech House Publishers, 1990).

18. J. E. Goell, "A circular-harmonic computer analysis of rectangular dielectric waveguides," Bell Syst. Tech. J. **51**, 2133-2160 (1969).

19. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Kluwer Academic Publishers, 2000).

20. S. Peng and A. Oliner, "Guidance and leakage properties of open dielectric waveguides. I. Mathemathical formulations," IEEE Trans. Microw. Theory Tech. **29**, 843-855 (1981).

21. A. S. Sudbø, "Film mode matching: a versatile numerical method for vector mode field calculations in dielectric waveguides," J. Europ. Opt. Soc. A **2**, 211-233 (1993).

22. P. Bienstman and R. Baets, "Optical modelling of photonic crystals and VCSELs using eigenmode expansion and perfectly matched layers," Opt. Quantum Electron. **33**, 327-341 (2001).

23. P. Bienstman, H. Derudder, R. Baets, F. Olyslager, D. D. Zutter, "Analysis of cylindrical waveguide discontinuities using vectorial eigenmodes and perfectly matched layers," IEEE Trans. Microw. Theory Tech. **49**, 349-354 (2001).

24. P. Bienstman, *Rigorous and efficient modelling of wavelength scale photonic components*, Ph.D. thesis, Ghent University (2001).

25. M. Stilling, *Coupled-Cavity Bottom-Emitting VCSELs*, Master's thesis, University of Southern Denmark (2004). http://www.mortenstilling.com/.

26. CAMFR, http://camfr.sourceforge.net/.

27. A. S. Sudbø, "Why Are Accurate Computations of Mode Fields in Rectangular Dielectric Waveguides Difficult?" J. Lightwave Technol. 10, 418-419 (1992).

28. J. Arnbak, "Leaky waves on a dielectric rod," Electron. Lett. **5**, 41-42 (1969).

29. H. Sagan, *Boundary and eigenvalue problems in mathematical physics* (Dover Publications Inc., 1989).

30. L. M. Delves and J. N. Lynnes, "A numerical method for locating the zeros of an analytic function," Math. Comput. **21**, 543-560 (1967).

31. J.-J. Greffet, A. Sentenac, R. Carminati, "Surface profile reconstruction using near-field data," Opt. Commun. **116**, 20-24 (1995).

32. R. Carminati and J.-J. Greffet, "Reconstruction of the dielectric contrast profile from near-field data," Ultramicr. **61**, 11-16 (1995).

33. P. S. Carney and J. C. Schotland, "Determination of three-dimensional structure in photon scanning tunnelling microscopy," J. Opt. A **4**, 140-144 (2002).

34. P. S. Carney, R. A. Frazin, S. I. Bozhevolnyi, V. S. Volkov, A. Boltasseva and J. C. Schotland, "Computational Lens for the Near Field," Phys. Rev. Lett. **92**, 163903 (2004).

35. D. Macías, A. Vial and D. Barchiesi, "Application of evolution strategies for the solution of an inverse problem in near-field optics," J. Opt. Soc. Am. A. **21**, 1465-1471 (2004).

36. S. I. Bozhevolnyi, B. Vohnsen and E. A. Bozhevolnaya, "Transfer functions in collection scanning near-field optical microscopy," Opt. Commun. **172**, 171-179 (1999).

37. DME-DualScopeTM, Herlev, Denmark.

38. R. Stockle, C. Fokas, V. Deckert and R. Zenoby, "High-quality near-field optical probes by tube etching," Appl. Phys. Lett. **75**, 160-162 (1999).

39. A. Shchemelinin, M. Rudman, K. Lieberman and A. Lewis, "A simple lateral force sensing technique for near-field micropattern generation," Rev. Sci. Instrum. **64**, 3538-3541 (1993).

40. S. I. Bozhevolnyi, V. A. Markel, V. Coello, W. Kim, and V. M. Shalaev, "Direct observation of localized dipolar excitations on rough nanostructured surfaces," Phys. Rev. B **58**, 11441-11448 (1998).

41. A. Nesci, R. Dändliker, M. Salt, H. P. Herzig, "Measuring amplitude and phase distribution of fields generated by gratings with sub-wavelength resolution," Opt. Commun. **205**, 229-238 (2002).

42. D. Van Labeke and D. Barchiesi, "Probes for scanning tunneling optical microscopy: a theoretical comparison," J. Opt. Soc. Am. A **10**, 2193-2201 (1993).

43. R. Carminati and J.-J. Greffet, "Two-dimensional numerical simulation of the photon scanning tunneling microscope. Concept of transfer-function," Opt. Commun. **116**, 316-321 (1995).

44. S. I. Bozhevolnyi, B. Vohnsen, E. A. Bozhevolnaya and S. Berntsen, "Self-consistent model for photon scanning tunneling microscopy: implications for image formation and light scattering near a phase-conjugating mirror," J. Opt. Soc. Am. A **13**, 2381-2392 (1996).

45. J.-J. Greffet and R. Carminati, "Image formation in near-field optics," Prog. Surf. Sci. **56**, 133-237 (1997).

46. S. I. Bozhevolnyi, "Near-field mapping of surface polariton fields", J. Micr. **202**, 313-319 (2000).

47. J. C. Weeber, F. de Fornel and J. P. Goudonnet, "Numerical study of the tip-sample interaction in the photon scanning tunneling microscope," Opt. Commun. **126**, 285-292 (1996).

48. J. C. Weeber, E. Bourillot, A. Dereux, J. P. Goudonnet, Y. Chen and C. Girard, "Observation of light confinement effects with a near-field optical microscope," Phys. Rev. Lett. **77**, 5332-5335 (1996).

49. J. R. Krenn, A. Dereux, J. C. Weeber, E. Bourillot, Y. Lacroute, J. P. Goudonnet, G. Schider, W. Gotschy, A. Leitner, F. R. Aussenegg and C. Girard, "Squeezing the optical near-field zone by plasmon coupling of metallic nanoparticles," Phys. Rev. Lett. **82**, 2590-2593 (1999).

50. S. Goumri-Said, L. Salomon, J. P. Dufour, F. de Fornel and A. V. Zayats, "Numerical simulations of photon scanning tunneling microscopy: role of a probe tip geometry in image formation," Opt. Commun. **244**, 245-258 (2005)

51. O. J. F. Martin, "3D simulations of the experimental signal measured in near-field optical microscopy," J. Micr. **194**, 235-239 (1999).

52. E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner and R. L. Kostelak, "Breaking the Diffraction Barrier: Optical Microscopy on a Nanometric Scale," Science **251**, 1468-1470 (1991).

53. J. Schöfer, M. J. Gregor, P. G. Blome and R. G. Ulbrich, "Influence of aperture diameter on image contrast and resolution in scanning near-field optical microscopy," J. Appl. Phys. **81**, 5871-5877 (1997).

54. S. I. Bozhevolnyi and E. A. Bozhevolnaya, "Near-Field imaging of the interference pattern of counterpropagating evanescent waves", Opt. Lett. **24**, 747-749 (1999).

55. J. Garnaes, P. -E. Hansen, N. Agersnap, J. Holm, F. Borsetto and A. Kühle, "Profiles of a high-aspect-ratio grating determined by spectroscopic scatterometry and atomic-force microscopy," Appl. Opt. **45**, 3201-3212 (2006).

56. I. P. Radko, S. I. Bozhevolnyi, and N. Gregersen, "Transfer function and near-field detection of evanescent waves," Appl. Opt. **45**, 4054-4061 (2006).

57. E. Yablonovitch, "Inhibited spontaneous emission in solid-state physics and electronics," Phys. Rev. Lett. **58**, 2059-2062 (1987).

58. C. Gerry and P. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2004).

59. N. Vats, S. John and K. Busch, "Theory of fluorescence in photonic crystals," Phys. Rev. A **65**, 043808 (2002).

60. K. Busch and S. John, "Photonic band gap formation in certain self-organizing systems," Phys. Rev. E **58**, 3896-3908 (1998).

61. T. Søndergaard, "Spontaneous Emission in Two-Dimensional Photonic Crystal Microcavities," IEEE J. Quantum Electron. **36**, 450-457 (2000).

62. J. P. Reithmaier, G. Sęk, A. Löffler, C. Hofmann, S. Kuhn, S. Reitzenstein, L. V. Keldysh, V. D. Kulakovskii, T. L. Reinecke and A. Forchel, "Strong coupling in a single quantum dot-semiconductor microcavity system," Nature **432**, 197-200 (2004).