Dissertation

submitted to the

Combined Faculties of the Natural Sciences and Mathematics

of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

André Junker

born in Schwetzingen, Germany

Oral examination: May 30, 2018

Advances in the Performance and Applicability of Modal Electromagnetic Simulations

Referees:

Prof. Dr. Karl-Heinz Brenner

Prof. Dr. Frank Wyrowski

Advances in the Performance and Applicability of Modal Electromagnetic Simulations

The 'rigorous coupled-wave analysis' (RCWA) is advanced in performance and extended in its application range. First, the RCWA framework is adapted to the treatment of structured incident and transmitted regions in order to enable the exact modeling of axially extended objects such as integrated optical components, long waveguides or fibers without incurring unwanted back-reflections from distant interfaces to homogeneous regions. Furthermore, a method to determine the propagation direction of eigenmodes is derived. Second, the treatment of coherent bidirectional light incidence is introduced and applied to the simulation of sample-induced aberrations in 4π-microscopy. In this context, a consistent formulation to describe arbitrary polarization states of structured incident light sources is derived. Third, the 'fast rigorous iterative method' (FRIM) is developed, an algorithm based on an iterative approach, which enables the rigorous simulation of structures such as certain diffractive optical elements with a significantly higher mode count than presently possible. This is achieved by replacing the computationally complex eigenmode decomposition inherent to standard modal methods by a sequence of efficient matrix multiplications. Thereby, the numerical cost is reduced from $\mathcal{O}(\bar{N}^3)$ to $\mathcal{O}(\bar{N}\log\bar{N})$, and at the same time the memory requirement is eased from $\mathcal{O}(\bar{N}^2)$ to $\mathcal{O}(\bar{N})$, $\bar{N}$ being the number of modes in the calculation.

Fortschritte im Leistungsvermögen und Anwendungsspektrum von Modalen Elektromagnetischen Simulationen

Die 'rigorous coupled-wave analysis' (RCWA) wird sowohl im Leistungsvermögen als auch im Anwendungsspektrum erweitert. Einerseits wird der RCWA-Algorithmus zur Behandlung strukturierter Randregionen angepasst, um die exakte Modellierung axial ausgedehnter Strukturen wie integrierter optischer Komponenten, langer Wellenleiter oder Fasern zu ermöglichen, ohne unerwünschte Rückreflexionen von entfernten Grenzflächen zu homogenen Regionen betrachten zu müssen. In diesem Zusammenhang wird zudem eine Methode zur Bestimmung der Propagationsrichtung von Eigenmoden hergeleitet. Zweitens wird die kohärente beidseitige Lichteinstrahlung in die RCWA eingeführt und auf die Simulation von objekt-induzierten Aberrationen in der 4π-Mikroskopie angewendet. Dies beinhaltet außerdem eine konsistente Formulierung zur Beschreibung beliebiger Polarisationszustände von strukturierten Beleuchtungen. Drittens wird die 'fast rigorous iterative method' (FRIM) entwickelt, ein auf einem iterativen Ansatz basierender Algorithmus, der die rigorose Simulation von Strukturen wie z.B. bestimmter diffraktiver optischer Elemente mit einer signifikant höheren Modenzahl als bisher ermöglicht. Dies wird dadurch erreicht, dass die numerisch komplexe Eigenmodenzerlegung, die in Modalmethoden inhärent enthalten ist, durch eine Aneinanderreihung effizienter Matrixmultiplikationen ersetzt wird. Dadurch erfolgt eine Reduzierung der numerische Komplexität von $\mathcal{O}(\bar{N}^3)$ auf $\mathcal{O}(\bar{N}\log\bar{N})$ und eine gleichzeitige Verringerung des Speicherbedarfs von $\mathcal{O}(\bar{N}^2)$ auf $\mathcal{O}(\bar{N})$, wobei $\bar{N}$ die Gesamtanzahl der in der Rechnung verwendeten Moden beschreibt.

Contents

List of symbols

Roman symbols

Notation	Description
a	Amplitude of dielectric waveguide eigenfunctions in CMT.
$\hat{a}^{(l)}$	Matrix used in the ETMA.
a_{lm}	Coefficient in the spherical wave decomposition.
$a\left(\mathbf{r}\right) = \left\lvert t\left(\mathbf{r}\right)\right\rvert$	Amplitude of the transmission function $t\left(\mathbf{r}\right)$.
A, ∂A	Surface area, closed path around a surface area.
$\hat{A}_{\pm}^{ll'}$	Matrix in the derivation of the S-matrix approach.
$\hat{b}^{(l)}$	Matrix used in the ETMA.
b_{lm}	Coefficient in the spherical wave decomposition.
$\mathbf{B}\left(\mathbf{r}, t\right)$	Magnetic flux density as function of $\mathbf{r}$ and t.
$\boldsymbol{\mathcal{B}}\left(\mathbf{r}, \omega\right)$	Magnetic flux density as function of $\mathbf{r}$ and ω.
$c_0 = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$	Speed of light in vacuum.
$\mathbf{c}_{\pm}^{(l)}$	Forward (+) and backward (-) propagating eigenmode coefficient vector.
$\hat{C}$	Circulant matrix (without upper index).
$\hat{C}^{(\mathrm{I/III})}$	Conversion matrix from the electric to the magnetic field Fourier mode coefficient vector in region I/III.
$d = \sum_l d^{(l)}$	Thickness of the grating.
d_{ap}	Aperture diameter.
d_{core}	Core diameter of a planar dielectric waveguide.
d_{gap}	Distance between the cores of two adjacent waveguides.
$d^{(l)}$	Thickness of the l^{th} grating layer.
$\mathrm{d}A$	Scalar surface element.
$\mathrm{d}\mathbf{A} = \mathbf{n}\,\mathrm{d}A$	Vectorial surface element.
$\mathrm{d}\mathbf{s}$	Unit vector in the direction of the vectorial path $\mathbf{s}$.
$\mathrm{d}x = \frac{P_x}{M}$	Position space sampling in x-direction.
$\mathrm{d}y = \frac{P_y}{N}$	Position space sampling in y-direction.
$\hat{D}$	Some diagonal matrix.
$\mathbf{D}\left(\mathbf{r}, t\right)$	Electric flux density as function of $\mathbf{r}$ and t.
$\boldsymbol{\mathcal{D}}\left(\mathbf{r}, \omega\right)$	Electric flux density as function of $\mathbf{r}$ and ω.

$\hat{\mathbf{e}}^{\mathrm{TE/TM}}$	Unit vector in TE/TM-direction.
$\hat{\mathbf{e}}_{x/y}$	Unit vector in x/y-direction.
$\mathrm{err}_{\mathrm{it}}\,(n_{\mathrm{it}})$	Error estimate of the FRIM iteration by comparison of two successive iterations.
$\mathrm{err}_{\mathrm{RCWA}}\,(n_{\mathrm{it}})$	Error estimate of the FRIM iteration by comparison to the RCWA result.
$\mathbf{E}\,(\mathbf{r},t)$	Electric field as function of $\mathbf{r}$ and t.
$\boldsymbol{\mathcal{E}}\,(\mathbf{r},\omega)$	Electric field as function of $\mathbf{r}$ and ω.
$\widetilde{\boldsymbol{\mathcal{E}}}_{mn}\,(z)$	Polarization vector of the electric field Fourier mode coefficient (m,n) inside the grating.
f	Focal length.
$f\,(x),\ f\,(x,y)$	Piecewise continuous, piecewise smooth, bounded periodic function of x (and y) with periods P_x (and P_y).
$\hat{f}^{(l)}$	Matrix used in the ETMA.
$\mathbf{f}$	Direction of a waveplate's fast axis.
$\hat{F}$	Fourier matrix.
$\mathbf{F}\,(\mathbf{r},t)$	General field as function of $\mathbf{r}$ and t.
$\boldsymbol{\mathcal{F}}\,(\mathbf{r},\omega)$	General field as function of $\mathbf{r}$ and ω.
g	Real coefficient.
$g\,(x),\ g\,(x,y)$	Piecewise continuous, piecewise smooth, bounded periodic function of x (and y) with periods P_x (and P_y).
$\hat{g}^{(l)}$	Matrix used in the ETMA.
h	Small axial propagation step in the RK method.
H	Height.
$\mathbf{H}\,(\mathbf{r},t)$	Magnetic field as function of $\mathbf{r}$ and t.
$\boldsymbol{\mathcal{H}}\,(\mathbf{r},\omega)$	Magnetic field as function of $\mathbf{r}$ and ω.
$\widetilde{\boldsymbol{\mathcal{H}}}_{mn}\,(z)$	Polarization vector of the magnetic field Fourier mode coefficient (m,n) inside the grating.
$i=\sqrt{-1}$	Imaginary unit.
$I_{\mathrm{unperturbed}}$	Intensity without the SNOM fiber tip.
j_l	Spherical Bessel functions.
$\mathbf{j}_{\mathrm{free}}\,(\mathbf{r},t)$	Free electric current density as function of $\mathbf{r}$ and t.
$k=\frac{2\pi n}{\lambda}$	Length of the wave vector in a medium.
$k_0=\frac{2\pi}{\lambda}$	Length of the wave vector in vacuum.
k_{x0}	Offset of the wave vector's x-component.
$k_{x;mn}$	x-component of the wave vector for the mode (m,n).
k_{y0}	Offset of the wave vector's y-component.
$k_{y;mn}$	y-component of the wave vector for the mode (m,n).
$k_{z;mn}$	z-component of the wave vector for the mode (m,n).
$\mathbf{k}$	Wave vector in the medium.
$\hat{K}_x=\mathrm{diag}\left(\frac{k_{x;mn}}{k_0}\right)$	Diagonal matrix with $\frac{k_{x;mn}}{k_0}$ on the diagonal.
$\hat{K}_y=\mathrm{diag}\left(\frac{k_{y;mn}}{k_0}\right)$	Diagonal matrix with $\frac{k_{y;mn}}{k_0}$ on the diagonal.

$\mathbf{K}_{1/2/3/4}$	Vectors in the RK4 method.
$l = 0 \,..\, L+1$	Layer index. $l = 0$: incident region, $l = L+1$: transmitted region.
L	Number of layers in the grating.
$\widetilde{\boldsymbol{\mathcal{L}}}_{mn}^{(\mathrm{I/III})}$	Polarization vector of the incident Fourier mode coefficient (m,n) in region I/III.
m	As a lower index: Fourier coefficient index with respect to x.
$M = 2M_0 + 1$	Total number of modes in x for a mode truncated system.
M_0	One-sided number of modes in x for a mode truncated system.
$\hat{M}^{(l)}$	Coupling matrix of the l^{th} layer.
$\mathbf{M}\,(\mathbf{r}, t)$	Magnetization as function of $\mathbf{r}$ and t.
n	As an elevated attribute: component normal to the interface.
n	Refractive index of a medium, $n = \sqrt{\epsilon\mu}$. As a lower index: Fourier coefficient index with respect to y.
n_{cladding}	Refractive index of the waveguide cladding.
n_{core}	Refractive index of the waveguide core.
n_{fast}	Refractive index of a birefringent material's fast axis.
n_{it}	Iteration variable.
n_{min}	Minimum of the refractive index of a structured medium.
n_{peak}	Peak refractive index at the center of a GRIN waveguide.
n_{slow}	Refractive index of a birefringent material's slow axis.
$\bar{n} = \langle n \rangle_{x,y}$	Average refractive index of a medium.
$\mathbf{n}$	Unit vector normal to an interface.
$N = 2M_0 + 1$	Total number of modes in y for a mode truncated system.
N_0	One-sided number of modes in y for a mode truncated system.
N_{it}	Number of iterations in the FRIM.
$N^{(l)}$	Number of propagation steps inside one layer for the RK method.
N_{o}	Expansion order of the exponential function in the FRIM.
$\bar{N} = MN$	Total number of modes in x and y for a mode truncated system.
$\mathbf{p}^{(l)}$	Vector used in the ETMA.
P_{coupled}	Amount of power coupled into a SNOM fiber.
P_{diss}	Dissipated power.
P_{fl}	Fluorescence power.
P_x	Period in x-direction.
P_y	Period in y-direction.
$\hat{P}_{\pm}^{(l)}\,(z)$	Diagonal forward (+) and backward (-) propagation matrix.

$\mathbf{P}(\mathbf{r},t)$	Polarization as function of $\mathbf{r}$ and t.
$\mathbf{q}^{(l)}$	Vector used in the ETMA.
$r=\lvert\mathbf{r}\rvert$	Absolute value of the spatial coordinate vector.
r_{hwp}	Radius of the half-wave plate.
$\mathbf{r}=(x,y,z)^T$	Spatial coordinate vector.
$\mathbf{r}_{\mathrm{I}}$	Point on an interface.
$\mathbf{r}_{\mathrm{I}}^{\pm}$	Location close to $\mathbf{r}_{\mathrm{I}}$, shifted infinitesimally in $\mathbf{n}$-direction to either side of the interface.
$\mathbf{r}^{(l)}$	Vector used in the ETMA.
R	Radius.
$\hat{R}$	R-matrix.
$\widetilde{\boldsymbol{\mathcal{R}}}_{mn}$	Polarization vector of the reflected Fourier mode coefficient (m,n) in region I.
$\widetilde{\boldsymbol{\mathcal{R}}}^{\mathrm{aprx}}$	Some approximation to the reflected Fourier mode coefficient vector.
$\mathbf{s}$	Vectorial path, or direction of a waveplate's fast axis.
$\mathbf{s}^{(l)}$	Vector used in the ETMA.
$\mathbf{S}(\mathbf{r},t)$	Poynting vector as function of $\mathbf{r}$ and t.
$\hat{S}$	S-matrix.
t	As an elevated attribute: component tangential to the interface.
t	Time in [s].
$t(\mathbf{r})$	Transmission function as function of $\mathbf{r}$.
$\mathbf{t}$	Unit vector tangential to an interface.
$\hat{T}$	T-matrix.
$\widetilde{\boldsymbol{\mathcal{T}}}_{mn}$	Polarization vector of the transmitted Fourier mode coefficient (m,n) in region III.
u	Eigenmode of a planar dielectric waveguide.
$\widetilde{\mathbf{v}}$	Some Fourier space vector.
V, ∂V	Volume, closed surface around a volume.
$\hat{V}^{(l)}$	Coupling and conversion matrix of the l^{th} layer from the eigenmode coefficient vector to the magnetic field Fourier mode coefficient vector.
w	Mixing parameter in the FRIM.
W	Width.
$\hat{W}^{(l)}$	Coupling matrix of the l^{th} layer from the eigenmode coefficient vector to the electric field Fourier mode coefficient vector.
x	x-component of the spatial coordinate vector.
$\hat{X}^{(l)}$	Diagonal eigenmode vector propagation matrix of the l^{th} layer, $X^{(l)}_{qq'}=e^{ik_0\sqrt{\lambda_q^{(l)}}d^{(l)}}\,\delta_{qq'}$.
y	y-component of the spatial coordinate vector.

y_l	Spherical Bessel functions.
$Y_l^m(\theta, \varphi)$	Spherical harmonics.
z	z-component of the spatial coordinate vector.
z_{crit}	Critical propagation distance (of AS or Sommerfeld propagation).

Greek symbols

Notation	Description
α	Apex angle of a SNOM fiber tip, or coefficient of $\hat{\mathbf{e}}^{\text{TE}}$-vector.
β	Coefficient of $\hat{\mathbf{e}}^{\text{TM}}$-vector.
β_q	Axial propagation constant in a planar dielectric waveguide.
γ	Axial propagation constant in CMT.
γ_q	Lateral propagation constant in a planar dielectric waveguide.
$\hat{\Gamma}$	Transition matrix unsed in the FRIM.
δn	Residual refractive index variation ($\delta n = n - \bar{n}$) in the BPM, or refractive index difference of a birefringent material ($\delta n = n_{\text{slow}} - n_{\text{fast}}$).
δx	Shift in x-direction.
$\Delta\phi$	Phase difference.
Δx	Lateral offset.
Δz	Axial offset.
ϵ_0	Electric permittivity of the vacuum.
$\hat{\epsilon}^{(l)}(x, y)$	Electric permittivity (3×3-tensor) in the l^{th} layer as function of x and y, constant in z.
$\epsilon(\mathbf{r}, \omega)$	Electric permittivity (scalar) as function of $\mathbf{r}$ and ω.
$\hat{\epsilon}(\mathbf{r}, \omega)$	Electric permittivity (3×3-tensor) as function of $\mathbf{r}$ and ω.
θ	Polar angle in spherical coordinates
θ_c	Critical angle for total internal reflection.
θ_q	Propagation angle of a planar dielectric waveguide.
κ	Coupling strength between waveguides in CMT.
λ	Vacuum wavelength.
$\lambda_q^{(l)}$	Diagonal entries of the eigenvalue matrix $\hat{\Lambda}^{(l)}$.
$\lambda_{\text{max}}^{(l)}$	Eigenvalue of the matrix $ik_0\hat{M}^{(l)}d^{(l)}$ with the largest absolute value.
$\lambda_{\text{max},\hat{\Xi}}$	Eigenvalue of the matrix $\hat{\Xi}$ with the largest absolute value.
$\hat{\Lambda}^{(l)}$	Diagonal eigenvalue matrix of the l^{th} layer with entries $\lambda_q^{(l)}$.
μ_0	Magnetic permeability of the vacuum.
$\hat{\mu}^{(l)}(x, y)$	Magnetic permeability (3×3-tensor) in the l^{th} layer as function of x and y, constant in z.
$\mu(\mathbf{r}, \omega)$	Magnetic permeability (scalar) as function of $\mathbf{r}$ and ω.
$\hat{\mu}(\mathbf{r}, \omega)$	Magnetic permeability (3×3-tensor) as function of $\mathbf{r}$ and ω.

$\boldsymbol{\nu} = \frac{\mathbf{k}}{2\pi}$	Spatial frequency.
$\xi \in \mathbb{R}$	Small real number that is sent to zero.
$\hat{\Xi}$	Contraction operator in the FRIM.
$\rho_{\text{free}}(\mathbf{r}, t)$	Free electric charge density as function of $\mathbf{r}$ and t.
σ	Width of a Gaussian function.
ϕ	Phase of a complex number.
$\phi(\mathbf{r}) = \arg[t(\mathbf{r})]$	Phase angle of the transmission function $t(\mathbf{r})$.
φ	Azimuthal angle in spherical coordinates.
ψ	Polarization angle.
ω	Angular frequency.

Mathematical symbols and operators

Notation	Description
$(\cdots)_\perp$	Combines the x and y components of a quantity in one vector, for instance $\boldsymbol{\mathcal{F}}_\perp = (\boldsymbol{\mathcal{F}}_x, \boldsymbol{\mathcal{F}}_y)^T$.
$(\cdots)^T$	Transpose.
$(\cdots)^*$	Complex conjugate.
$\langle\cdots\rangle_t$	Time averaging operator.
$[\![\cdots]\!]_{m,m'}$, $[\![\cdots]\!]_{mn,m'n'}$	Toeplitz matrix (2d or 4d) for the application of the direct rule in x and y.
$\lceil\cdots\rceil_{m,m'}(y)$	Toeplitz matrix (2d) for the application of the direct rule only in x.
$\lfloor\cdots\rfloor_{n,n'}(x)$	Toeplitz matrix (2d) for the application of the direct rule only in y.
$\lfloor\lceil\cdots\rceil\rfloor_{mn,m'n'}$	Toeplitz matrix (2d) for the application of the inverse rule only in x and the direct rule in y.
$\lceil\lfloor\cdots\rfloor\rceil_{mn,m'n'}$	Toeplitz matrix (2d) for the application of the direct rule only in x and the inverse rule in y.
$\widetilde{(\cdots)}_m(y)$, $\widetilde{(\cdots)}_n(x)$, $\widetilde{(\cdots)}_{mn}$	Fourier transfored quantity with respect to x, y or both. Transformed quantities are identified by the indexing. If there is no indexing, then the transformation is always in both coordinates.
$(\cdots)^{(\mathrm{I/II/III}/l)}$	As an elevated attribute: specifies the region. I: incident, II: grating, III: transmitted, l: layer number.
$\mathbb{0}$	Zero matrix.
$\mathbb{1}$	Identity matrix.
$\delta_{nn'} = \begin{cases} 1 & n = n' \\ 0 & n \neq n' \end{cases}$	Kronecker delta.
$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)^T$	Three dimensional spatial differentiation operator.
arg	Returns the phase of a complex number.

diag	Creates a diagonal matrix from a vector.
$\mathscr{F}$	Fourier transformation operator.
$\Im$	Imaginary part of a complex number.
$\mathcal{O}(\ldots)$	Order of accuracy of a method ('big-O-notation').
$\Re$	Real part of a complex number.
sgn	Sign of a real number.

Acronyms

Acronym	Expansion
Al	Aluminum
arb.u.	Arbitrary units
AS	Angular spectrum
BPM	Beam propagation method
BTTB	Block-Toeplitz Toeplitz-block
c.c.	Complex conjugate
CFL	Courant-Friedrichs-Lewy condition
CMT	Coupled mode theory
DOE	Diffractive optical element
ETMA	Enhanced transmittance matrix approach
FDTD	Finite-difference time-domain
FFT	Fast Fourier transform
FMM	Fourier modal method
FRIM	Fast rigorous iterative method
FZP	Fresnel zone plate
GRIN	Graded index
HWP	Half-wave plate
LHS	Left hand side
LIF-RCWA	Localized input field RCWA
NA	Numerical aperture
PML	Perfectly matched layers
RAM	Random access memory
RCWA	Rigorous coupled-wave analysis
RHS	Right hand side
RK	Runge-Kutta
RK4	Fourth order accurate RK
S-matrix	Scattering matrix
SNOM	Near-field scanning optical microscopy
TE	Transversal electric
TEA	Thin element approximation
TM	Transversal magnetic
T-matrix	Transfer matrix
VBPM	Vectorial BPM

VTEA	Vectorial TEA
VWPM	Vectorial WPM
WPM	Wave propagation method

1 | Introduction

1.1 Scientific context

The 'rigorous coupled-wave analysis' (RCWA) was initially developed in the year 1981 by Moharam and Gaylord as an exact solver of Maxwell's equations for the long wavelength electromagnetic domain[1]. The application of the algorithm in the optical wavelength regime was initially not intended and was also impeded by the lack of sufficient processing power and memory space, which would have been necessary to simulate realistic physical settings. Nevertheless, the interest to apply the RCWA also to optical wavelengths existed, but was, in consequence, limited to the simulation of structures in the nanooptical scale with dimensions in the order of a few tens of wavelengths or to periodic problems with periods in the size of a few tens of wavelengths.

Within the last two decades, the progress in lithographic fabrication was one of the driving forces for the advancement of rigorous electromagnetic simulations. While the dimensions of the smallest fabricable structures were still well above 1 µm in the beginning of the 1980s, presently the possible resolution comes down to some 14 nm[2], which is significantly less than one tenth of optical wavelengths. This development lead to several enhancements of the initial RCWA. In 1993, the algorithmic framework was extended by Li to handle multiple grating layers[3], and in the same year, the treatment of two-dimensional gratings was introduced by Bräuer and Bryngdahl[4]. In 1995, conical light incidence was included[5]. While the RCWA seemed to work well for transversal electric (TE) polarization, it was noted that the convergence properties of the algorithm were not satisfactory for transversal magnetic (TM) polarization, especially when one dealt with metallic gratings. In 1996, first an empiric solution to this problem was found independently by Lalanne and Morris[6] and Granet and Guizal[7]. The latter was then formalized by Li, who found out that the bad convergence behavior was caused by an insufficient Fourier reconstruction of position space products of functions with pairwise-concurrent jump discontinuities. These findings subsequently lead to the so-called 'Li's factorization rules'[8, 9]. Moreover, the inherent stability of the RCWA was also greatly improved by new methods to enforce the boundary conditions, namely the introduction of the 'R-matrix approach' and the 'scattering matrix approach' (S-matrix approach)[10], as well as the 'enhanced transmittance matrix approach' (ETMA)[11, 12]. In 2001,

the concept of 'perfectly matched layers' (PMLs) was introduced, which facilitated the simulation of non-periodic structures[13]. Furthermore, Popov and Nevière[14] and later Li[15] extended the RCWA to the treatment of linear anisotropic media. Based on an idea formulated in [14], Schuster introduced normal vector fields for the description of refractive index structures, this way being able to generalize Li's factorization rules to material interfaces with arbitrary orientation[16]. In 2014, Auer and Brenner published how structured light incidence could be efficiently implemented in the existing framework without changing the complexity of the RCWA algorithm[17].

Despite all the advances in this field, and despite the huge progress made in computer technology during the last couple of decades, I would like to emphasize that the RCWA is still far from being fully developed. In fact, one of its main problems - the limitation to very small problem sizes - remains, and, so far, there exists no satisfactory solution to reduce the extreme memory consumption and algorithmic complexity of the RCWA down to an acceptable level. The only attempt known to me is the approach of Semenikhin and Zanuccoli[18], who recently claimed to have reduced the numerical complexity of modal methods in certain cases from $\mathcal{O}\left(\bar{N}^3\right)$ to $\mathcal{O}\left(\bar{N}^2\right)$ by calculating the eigenmode expansion coefficients via an iterative procedure, where in the case of a modal description $\bar{N}$ represents the number of modes in the calculation. Indeed, scalar wave-optical propagation algorithms are still mostly preferred to exact Maxwell solvers like the RCWA or the 'finite-difference time-domain' (FDTD) method[19], which are only used if truly necessary due to their severe time and memory constraints. Furthermore, it should be noted that, still, not all problem geometries can be represented in the standard RCWA framework.

1.2 Structure of this dissertation

This dissertation advances the performance of the RCWA and extends its application range to a much wider set of problems than presently possible. To this end, several modifications and generalizations are made at multiple levels of the currently existing framework. In the following, the contributions of this dissertation to the field are introduced. The structure is as follows.

Chapter 2
The basic principles of optics, as far as required within the scope of this dissertation, are explained and derived based on Maxwell's equations. The latter include the derivation of the continuity conditions across material interfaces, the modeling of light matter interactions for general and linear media including absorption, the derivation of the Poynting theorem, and an overview and the categorization of several established propagation algorithms. These include, amongst others, the 'Sommerfeld propagation' and the 'angular spectrum propagation' in homogeneous media, and the

'thin element approximation' (TEA), the 'beam propagation method' (BPM), the 'wave propagation method' (WPM) and the FDTD for structured media. Finally, it is shown how the electromagnetic fields can be normalized in order to change to a symmetric and dimensionless description of Maxwell's equations.

Chapter 3
The framework of the RCWA is derived. The latter includes a definition of the RCWA problem geometry, the Fourier analysis of Maxwell's equations for linear isotropic and linear anisotropic media using Li's factorization rules, the derivation of the most general solution of the resulting system of coupled differential equations, and the introduction of several standard methods for enforcing boundary conditions like the S-matrix approach and the ETMA. Furthermore, the consistency requirements for tilted structured light incidence as proposed by Auer an Brenner[17] are derived. The numerical complexity and the memory requirement of the RCWA is motivated. The framework is subsequently applied to the simulation of a 'near-field scanning optical microscopy' (SNOM)[20] experiment, which was is published in [A1].

Chapter 4
It is demonstrated how the RCWA is extended to structured incident and transmitted regions. To this end, the boundary value problem of the RCWA is reformulated and the energy flow properties of eigenmodes in structured media are investigated. The aim is the fully vectorial modeling of axially extended objects such as infinitely long waveguides or fibers. The numerical results are compared to the analytically known eigenmodes of planar dielectric waveguides. Furthermore, comparative simulations are conducted with approximative algorithms like the 'coupled mode theory' (CMT)[21, 22] (evanescent coupling) and the BPM. This chapter is based on the publication [A2].

Chapter 5
The RCWA framework is extended to the treatment of coherent bidirectional light incidence and the simulation of 4π-microscopy. To this end, three different concepts are combined. First, the idea of structured illumination presented in [17] is applied in order to model focused beams in a fast and consistent manner. Second, it is derived how structured light sources of different polarization types, such as linear, radial, azimuthal or elliptic polarization, can be realized in a general form. Third, it is shown how coherent bidirectional structured light incidence is integrated into the existing RCWA framework. The mentioned concepts are combined to investigate the near field of light propagating in the vicinity of a specimen in 4π-microscopy. This chapter is based on the publications [A3, A4].

Chapter 6
As mentioned in the introductory part, one of the main drawbacks of rigorous Maxwell solvers is the limitation to very small calculation domains due to the high algorithmic complexity and memory consumption. In this chapter, the 'fast rigor-

ous iterative method' (FRIM) is developed, an exact algorithm based on an iterative approach, which, under certain conditions, allows solving also large size problems approximation free. The latter is achieved by replacing the computationally complex eigenmode decomposition inherent to standard modal methods, which is responsible for the algorithmic complexity of $\mathcal{O}\big(\bar{N}^3\big)$, by a sequence of efficient matrix multiplications. It is shown that, thereby, the numerical cost is reduced from $\mathcal{O}\big(\bar{N}^3\big)$ to $\mathcal{O}\big(\bar{N}\log\bar{N}\big)$, and the memory requirement is eased from $\mathcal{O}\big(\bar{N}^2\big)$ to $\mathcal{O}\big(\bar{N}\big)$, where $\bar{N}$ represent the number of modes in the calculation. Apart from the derivation of the FRIM iteration scheme, the mathematical and the empiric convergence properties are analyzed in detail. The validity of the FRIM is verified by comparison to other simulation methods, amongst others the standard RCWA for small-scale problems, and the TEA and the WPM for larger problem sizes. The decrease of the numerical complexity achieved with the FRIM enables to raise the lateral dimensions of rigorously analyzable structures like certain large sized diffractive optical elements (DOEs) to a much higher level, with mode counts in the order of 2000×2000 and more being possible. In this context, the FRIM becomes especially useful in cases, where a small scale periodicity, such as in gratings, is not present and, therefore, large calculation domains are inevitable. Furthermore, it is demonstrated that the FRIM can be directly extended to the treatment of linear anisotropic media. This chapter is based on the publications [A5, A6].

2 | Theoretical basis

In this chapter, the basic principles of optics that are needed throughout this theses are derived. In Sect. 2.1, a brief historical background about the field of optics is given, starting from the early beginnings and closing with the discovery of Maxwell's equations. In Sect. 2.2, Maxwell's equations are presented and the transition to the temporal frequency domain is performed. In Sect. 2.3, the continuity conditions of the electromagnetic fields across material interfaces are derived using Gauss's and Stokes' theorem. In Sect. 2.4, general and linear media are introduced. Sect. 2.5 derives the Poynting theorem and establishes a connection between the electric permittivity, the magnetic permeability and the absorption properties of a material. In Sect. 2.6, the wave equation and the Helmholtz equation are derived. Furthermore, the plane wave decomposition and the spherical wave decomposition are introduced. Sect. 2.7 gives an overview over the spectrum of numerical propagation algorithms in general. Furthermore, several selected methods are presented in more detail. This includes the angular apectrum and Sommerfeld propagation in homogeneous media, and the BPM, the WPM and the FDTD for structured media. Finally, in Sect. 2.8, it is shown how the electromagnetic fields are normalized in order to obtain a symmetric and dimensionless description of Maxwell's equations.

Unless indicated otherwise, the derivations shown in this chapter can be found in [23, 24].

2.1 A brief history of optics

The properties and characteristics of light have been subject of scientific research for thousands of years. Already the Egyptians and the ancient Greek investigated the properties of crystal lenses and developed theories about the propagation of light[25]. In the early beginnings of classical optics, the field was mainly dominated by mathematicians like Euclid, who viewed optics from a very geometrical perspective. At this early stage, it was thought that light was spreading along straight lines, which they called 'rays'[26]. Optics experienced its first renaissance with the advent of early modern astronomy in the beginning of the 17$^{\text{th}}$ century. At that time, Johannes Kepler and Galileo Galilei studied the movement of the stars and planets in the sky with their at that time very sophisticated refractive telescopes. They discovered

and characterized new effects, such as the atmospheric refraction[27, 28]. Only little later, in 1621, the Dutch astronomer Willebrord Snellius succeeded in mathematically formalizing the refraction of light at material interfaces in his famous law of refraction[29]. In the late 17th century, the phenomenon of dispersion was discovered by the English scientist Isaac Newton, when he found out that white light could be decomposed into its spectral colors by a prism. From this, he concluded that every refractive telescope would suffer from chromatic aberrations[30, 31]. He eliminated this problem by the design of a reflecting telescope, the quality of whose mirrors he judged very accurately by means of 'Newton's rings'[31]. The foundation of diffractive optics and the wave theory of light was subsequently laid by the Dutch mathematician Christiaan Huygens in 1690, whose main proposition was that the speed of light was finite[32]. In 1801, his theories were picked up by the English physician Thomas Young in the explanation of his famous double slit interferometer experiment[33]. In the early 19th century (1815-1818), the young french physicist Augustin-Jean Fresnel backed Huygens's and Young's results with a series of detailed analyses of the interference and diffraction properties of differently polarized light sources[34, 35].

At about the same time as Fresnel did his research on the properties of light, the scientists Michael Faraday and André-Marie Ampère worked on problems in a seemingly completely different field. In 1822, Ampère found out that an electric current passing through two wires causes these wires to attract or repel each other depending on the direction of the current[36]. In 1831, Faraday discovered the law of electromagnetic induction when experimenting with two insulated coils wrapped around an iron core[37]. Their work was studied little later by the Scottish scientist James Clerk Maxwell, who, in 1861, published a work, in which he proposed how electricity and magnetism could be described in just one single linked set of differential equations[38]. At that time, he also noticed that one possible solution of his new system of equations represented an oscillating electric and magnetic field traveling through empty space at approximately the same speed that one had before measured for the propagation of light[39]. It was, indeed, not until this point in time in the late 19th century that the scientific community realized that all these phenomena, light, electricity and magnetism, were attributed to one and the same underlying physical effect. The famous Maxwell's equations in the vectorial notation as we know them today were, however, developed at a later point by the English physicist Oliver Heaviside[40], the American scientist Josiah Willard Gibbs[41] and the German physicist Heinrich Hertz[42, 43] as a simplification of the initially 20 equations published by Maxwell.

2.2 Maxwell's equations

The macroscopic Maxwell's equations are a set of partial differential equations that describe the interplay of electric and magnetic fields with electric charges, currents

and matter. In their differential form they read

$$\nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t} \qquad \textbf{Maxwell-Faraday equation} \tag{2.1a}$$

$$\nabla \times \mathbf{H}(\mathbf{r},t) = \mathbf{j}_{\text{free}}(\mathbf{r},t) + \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} \qquad \textbf{Ampère's circuital law} \tag{2.1b}$$

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = \rho_{\text{free}}(\mathbf{r},t) \qquad \textbf{Gauss's law} \tag{2.1c}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r},t) = 0 \qquad \textbf{Gauss's law for magnetism} \tag{2.1d}$$

The Maxwell-Faraday equation states that a time-varying magnetic flux density $\mathbf{B}(\mathbf{r},t)$ induces an electric field $\mathbf{E}(\mathbf{r},t)$. Following Ampère's circuital law, a magnetic field $\mathbf{H}(\mathbf{r},t)$ is induced by a current $\mathbf{j}_{\text{free}}(\mathbf{r},t)$ of free electric charges and the so-called 'displacement current', which is the derivative of the electric flux density $\mathbf{D}(\mathbf{r},t)$ with respect to time. Gauss's law states that the free electric charge density $\rho_{\text{free}}(\mathbf{r},t)$ is the source of the electric flux density $\mathbf{D}(\mathbf{r},t)$. Similarly, Gauss's law for magnetism states that the magnetic flux density $\mathbf{B}(\mathbf{r},t)$ is always source-free, which is equivalent to the statement that there exist no free magnetic monopoles.

Within the scope of this thesis, no free currents or free charges are involved. Therefore, we set $\mathbf{j}_{\text{free}}(\mathbf{r},t) = \mathbf{0}$ and $\rho_{\text{free}}(\mathbf{r},t) = 0$. With these simplifications, Maxwell's equations read

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

$$\nabla \times \mathbf{H}(\mathbf{r},t) = +\frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} \tag{2.2b}$$

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = 0 \tag{2.2c}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r},t) = 0\,. \tag{2.2d}$$

The latter are Maxwell's equations with the fields being functions of position $\mathbf{r} = (x,y,z)^T$ and time t. Maxwell solvers like, for instance, the FDTD directly operate on Eqs. (2.1) or Eqs. (2.2). Other solvers, like for instance the RCWA, are Fourier domain methods, i.e., only monochromatic fields are considered. For these methods, Eqs. (2.2) are transformed into the temporal frequency domain. This is done by defining

$$\mathbf{F}(\mathbf{r},t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \boldsymbol{\mathcal{F}}(\mathbf{r},\omega)\, e^{-i\omega t}\, d\omega\,, \tag{2.3}$$

where $\mathbf{F} \in \{\mathbf{E}, \mathbf{B}, \mathbf{H}, \mathbf{D}\}$ and $\boldsymbol{\mathcal{F}} \in \{\boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{H}}, \boldsymbol{\mathcal{D}}\}$. Eqs. (2.2) can then be rewritten in terms of the transformed fields $\boldsymbol{\mathcal{F}}(\mathbf{r}, \omega)$,

$$\nabla \times \boldsymbol{\mathcal{E}}(\mathbf{r}, \omega) = +i\omega \boldsymbol{\mathcal{B}}(\mathbf{r}, \omega) \tag{2.4a}$$
$$\nabla \times \boldsymbol{\mathcal{H}}(\mathbf{r}, \omega) = -i\omega \boldsymbol{\mathcal{D}}(\mathbf{r}, \omega) \tag{2.4b}$$
$$\nabla \cdot \boldsymbol{\mathcal{D}}(\mathbf{r}, \omega) = 0 \tag{2.4c}$$
$$\nabla \cdot \boldsymbol{\mathcal{B}}(\mathbf{r}, \omega) = 0\,. \tag{2.4d}$$

Throughout this thesis, the calligraphic font always indicates a field, which is Fourier transformed with respect to the time coordinate. On the other hand, the normal font always indicates a field in the time-domain. The arguments ω or t may be dropped in order to facilitate the readability.

2.3 Continuity conditions

It is commonly known that the tangential electric and magnetic field components are continuous across a discontinuous material interface. Moreover, also the normal components of the electric and magnetic flux densities are continuous. In the following, these continuity conditions are derived directly from Maxwell's equations via Gauss's and Stokes' theorem.

2.3.1 Electric and magnetic flux density

Assume a material interface as depicted in Fig. 2.1 and a Gaussian box, which is arranged symmetrically around a point $\mathbf{r}_\mathrm{I}$ on the interface. Following Gauss's theorem, from Eq. (2.2c) one obtains

$$\int_V \nabla \cdot \mathbf{D}(\mathbf{r}, t)\, \mathrm{d}V = \int_{\partial V} \mathbf{D}(\mathbf{r}, t) \cdot \mathrm{d}\mathbf{A} = 0\,. \tag{2.5}$$

Now let the height H of the box go to zero symmetrically around the interface, i.e., such that the upper/lower circular boundary of the box remains on the upper/lower side of the interface. In the limit $H \to 0$, the area of the cylinder's side walls goes to zero and, therefore, does not contribute to the surface integral in Eq. (2.5). Only the circular areas on the top/bottom contribute. Then, let the radius R of the box go to zero. In the limit $R \to 0$ one obtains

$$\lim_{\substack{R\to 0\\ H\to 0}} \int_{\substack{\text{circular}\\ \text{areas}}} \mathbf{D}(\mathbf{r}, t) \cdot \frac{\mathrm{d}\mathbf{A}}{\pi R^2} = \mathbf{D}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^+, t\right) - \mathbf{D}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^-, t\right) = 0\,, \tag{2.6}$$

where $\mathbf{r}_\mathrm{I}^\pm = \lim\limits_{\xi\to 0} \mathbf{r}_\mathrm{I} \pm \xi\, \mathbf{n}$, and $\mathbf{D}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^\pm, t\right) = \left[\mathbf{D}\left(\mathbf{r}_\mathrm{I}^\pm, t\right) \cdot \mathbf{n}\right] \mathbf{n}$ represent the normal components of the electric flux density on the two sides of the interface. From Eq.

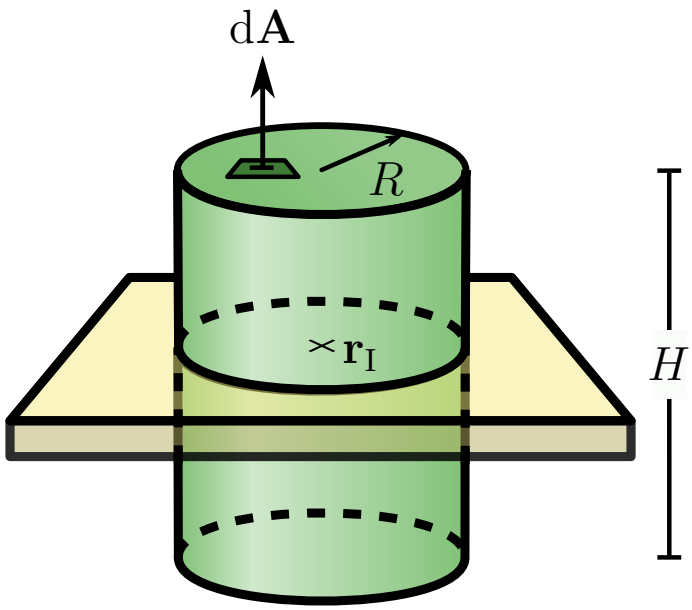


Figure 2.1: Gaussian box (green cylinder) with radius R and height H, symmetrically arranged around the material interface (yellow). The vectorial surface element is $\mathrm{d}\mathbf{A} = \mathbf{n}\,\mathrm{d}A$, $\mathbf{n}$ being the normal vector to the interface.

(2.6) follows the continuity of the normal component of the electric flux density,

$$\mathbf{D}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^+, t\right) = \mathbf{D}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^-, t\right) . \tag{2.7}$$

The same line of arguments can be followed for the magnetic flux density. One obtains

$$\mathbf{B}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^+, t\right) = \mathbf{B}^\mathrm{n}\left(\mathbf{r}_\mathrm{I}^-, t\right) . \tag{2.8}$$

2.3.2 Electric and magnetic fields

Assume a material interface as depicted in Fig. 2.2 and a Stokes path, which is arranged symmetrically around a point $\mathbf{r}_\mathrm{I}$ on the interface. Following Stokes' theorem, from Eq. (2.2a) one obtains

$$-\int_A \frac{\partial \mathbf{B}\left(\mathbf{r}, t\right)}{\partial t} \cdot \mathrm{d}\mathbf{A} = \int_A \left[\nabla \times \mathbf{E}\left(\mathbf{r}, t\right)\right] \cdot \mathrm{d}\mathbf{A} = \int_{\partial A} \mathbf{E}\left(\mathbf{r}, t\right) \cdot \mathrm{d}\mathbf{s} . \tag{2.9}$$

Now let the height H of the rectangle go to zero symmetrically around the interface, i.e., such that the two paths parallel to the surface remain on the upper and lower side of the interface, respectively. In the limit $H \to 0$, the area of the rectangle goes to zero, i.e.

$$\lim_{H \to 0} \int_A \frac{\partial \mathbf{B}\left(\mathbf{r}, t\right)}{\partial t} \cdot \mathrm{d}\mathbf{A} = 0 . \tag{2.10}$$

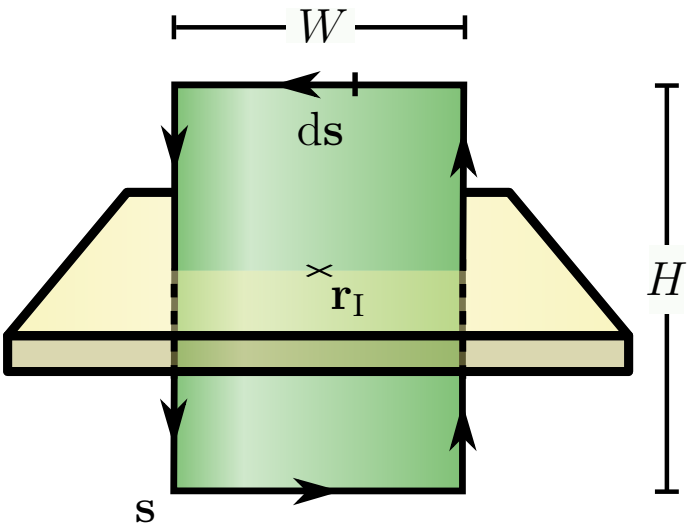


Figure 2.2: Stokes path **s** (border of the green rectangle) with width W and height H, symmetrically arranged around a material interface (yellow). The unit vector parallel to the Stokes path is denoted by d**s**.

Furthermore, the paths perpendicular to the interface vanish and, therefore, do not contribute to the line integral on the right side of Eq. (2.9). Only the paths tangential to the interface contribute. Then, let also the width W of the rectangle go to zero. One obtains

$$\lim_{\substack{H\to 0\\ W\to 0}} \int_{\substack{\text{tangential}\\ \text{paths}}} \mathbf{E}\left(\mathbf{r},t\right)\cdot\frac{\mathrm{d}\mathbf{s}}{W} \quad = \quad \mathbf{E}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^+,t\right)-\mathbf{E}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^-,t\right) \quad = \quad 0\,, \tag{2.11}$$

with $\mathbf{r}_\mathrm{I}^\pm$ as before. Here, $\mathbf{E}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^\pm,t\right)=\left[\mathbf{E}\left(\mathbf{r}_\mathrm{I}^\pm,t\right)\cdot\mathbf{t}\right]\mathbf{t}$ are the tangential components of the electric field on the two sides of the interface, where $\mathbf{t}$ is a unit vector tangential to the interface at $\mathbf{r}_\mathrm{I}$. From Eq. (2.11) follows the continuity of the tangential component of the electric field,

$$\mathbf{E}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^+,t\right)=\mathbf{E}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^-,t\right)\,. \tag{2.12}$$

Since the orientation of the unit vector $\mathbf{t}$ is arbitrary within the tangential plane of the interface, *all* tangential components of the electric field are continuous across the interface. The same line of arguments can be followed for the magnetic field. One obtains

$$\mathbf{H}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^+,t\right)=\mathbf{H}^\mathrm{t}\left(\mathbf{r}_\mathrm{I}^-,t\right)\,. \tag{2.13}$$

2.4 Material properties

2.4.1 General media

The quantities $\mathbf{E}(\mathbf{r},t)$, $\mathbf{B}(\mathbf{r},t)$, and $\mathbf{D}(\mathbf{r},t)$, $\mathbf{H}(\mathbf{r},t)$ denote the electric and magnetic fields on a microscopic and macroscopic level. Generally, the material response is described by the electric polarization $\mathbf{P}(\mathbf{r},t)$ and the magnetization $\mathbf{M}(\mathbf{r},t)$. The dependence is

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t) \tag{2.14a}$$
$$\mathbf{B}(\mathbf{r},t) = \mu_0 \mathbf{H}(\mathbf{r},t) + \mu_0 \mathbf{M}(\mathbf{r},t)\ . \tag{2.14b}$$

This relation is, by definition, exact, but in many applications impractical when the exact dependence of $\mathbf{P}(\mathbf{r},t)$ on $\mathbf{E}(\mathbf{r},t)$, or $\mathbf{M}(\mathbf{r},t)$ on $\mathbf{H}(\mathbf{r},t)$ is either unknown or very complicated. Therefore, to characterize a material, the medium response is usually expanded in powers of the external field. In first order, the relation is just linear, which is a good approximation to a large number of real materials. The latter are called 'linear dispersive media'. Nevertheless, there exist also materials that exhibit also a higher order response to the external fields. In literature, these media are then referred to as 'non-linear materials'.

2.4.2 Linear media

Linear dispersive media

For linear dispersive media, the fields in the temporal Fourier domain obey the following linear relation,

$$\boldsymbol{\mathcal{D}}(\mathbf{r},\omega) = \epsilon_0\, \hat{\epsilon}(\mathbf{r},\omega)\ \boldsymbol{\mathcal{E}}(\mathbf{r},\omega) \tag{2.15a}$$
$$\boldsymbol{\mathcal{B}}(\mathbf{r},\omega) = \mu_0 \hat{\mu}(\mathbf{r},\omega)\, \boldsymbol{\mathcal{H}}(\mathbf{r},\omega)\ , \tag{2.15b}$$

where $\epsilon_0 \hat{\epsilon}(\mathbf{r},\omega)$ and $\mu_0 \hat{\mu}(\mathbf{r},\omega)$ are the 3×3-tensors of the electric permittivity and the magnetic permeability. Dispersion is defined as a frequency dependence of the electric permittivity $\hat{\epsilon}(\mathbf{r},\omega)$ or the magnetic permeability $\hat{\mu}(\mathbf{r},\omega)$.

Linear dispersionless media

If the electric permittivity and magnetic permeability do not depend on ω, a material is called 'dispersionless', i.e.

$$\boldsymbol{\mathcal{D}}(\mathbf{r},\omega) = \epsilon_0\, \hat{\epsilon}(\mathbf{r})\ \boldsymbol{\mathcal{E}}(\mathbf{r},\omega) \tag{2.16a}$$
$$\boldsymbol{\mathcal{B}}(\mathbf{r},\omega) = \mu_0 \hat{\mu}(\mathbf{r})\, \boldsymbol{\mathcal{H}}(\mathbf{r},\omega)\ . \tag{2.16b}$$

In this case, this relation also directly transfers to the fields in the temporal domain,

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0\,\hat{\epsilon}(\mathbf{r})\;\mathbf{E}(\mathbf{r},t) \tag{2.17a}$$
$$\mathbf{B}(\mathbf{r},t) = \mu_0\hat{\mu}(\mathbf{r})\,\mathbf{H}(\mathbf{r},t)\;. \tag{2.17b}$$

Linear isotropic media

For isotropic media, the material response is independent of direction, i.e., the material tensors become scalar quantities,

$$\boldsymbol{\mathcal{D}}(\mathbf{r},\omega) = \epsilon_0\,\epsilon(\mathbf{r},\omega)\;\boldsymbol{\mathcal{E}}(\mathbf{r},\omega) \tag{2.18a}$$
$$\boldsymbol{\mathcal{B}}(\mathbf{r},\omega) = \mu_0\mu(\mathbf{r},\omega)\,\boldsymbol{\mathcal{H}}(\mathbf{r},\omega)\;. \tag{2.18b}$$

Linear dispersionless isotropic media

Linear dispersionless isotropic materials combine the properties of dispersionless and isotropic media. One obtains the relations

$$\boldsymbol{\mathcal{D}}(\mathbf{r},\omega) = \epsilon_0\,\epsilon(\mathbf{r})\;\boldsymbol{\mathcal{E}}(\mathbf{r},\omega) \tag{2.19a}$$
$$\boldsymbol{\mathcal{B}}(\mathbf{r},\omega) = \mu_0\mu(\mathbf{r})\,\boldsymbol{\mathcal{H}}(\mathbf{r},\omega) \tag{2.19b}$$

and

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0\,\epsilon(\mathbf{r})\;\mathbf{E}(\mathbf{r},t) \tag{2.20a}$$
$$\mathbf{B}(\mathbf{r},t) = \mu_0\mu(\mathbf{r})\,\mathbf{H}(\mathbf{r},t)\;. \tag{2.20b}$$

2.5 Energy conservation

2.5.1 Poynting theorem

Multiply Eq. (2.2b) with $\mathbf{E}(\mathbf{r},t)$ from the left,

$$\mathbf{E}(\mathbf{r},t)\cdot[\nabla\times\mathbf{H}(\mathbf{r},t)] = \mathbf{E}(\mathbf{r},t)\cdot\frac{\partial\mathbf{D}(\mathbf{r},t)}{\partial t}\,, \tag{2.21}$$

and rewrite the left side using a vectorial identity,

$$\mathbf{H}(\mathbf{r},t)\cdot[\nabla\times\mathbf{E}(\mathbf{r},t)] - \nabla\cdot[\mathbf{E}(\mathbf{r},t)\times\mathbf{H}(\mathbf{r},t)] = \mathbf{E}(\mathbf{r},t)\cdot\frac{\partial\mathbf{D}(\mathbf{r},t)}{\partial t}\,. \tag{2.22}$$

By replacing $\nabla\times\mathbf{E}(\mathbf{r},t)$ using Eq. (2.2a) one obtains

$$\nabla\cdot\mathbf{S}(\mathbf{r},t) = -\mathbf{E}(\mathbf{r},t)\cdot\frac{\partial\mathbf{D}(\mathbf{r},t)}{\partial t} - \mathbf{H}(\mathbf{r},t)\cdot\frac{\partial\mathbf{B}(\mathbf{r},t)}{\partial t}\,, \tag{2.23}$$

where $\mathbf{S}(\mathbf{r},t) = \mathbf{E}(\mathbf{r},t) \times \mathbf{H}(\mathbf{r},t)$ is identified as the Poynting vector, which has units of an energy flux density $[\mathbf{S}] = Jm^{-2}s^{-1}$. By integrating Eq. (2.23) over an arbitrary connected volume V and applying Gauss's theorem on the left side, one obtains

$$\int_{\partial V} \mathbf{S}(\mathbf{r},t)\,\mathrm{d}\mathbf{A} = -\int_V \left[\mathbf{E}(\mathbf{r},t)\cdot\frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} + \mathbf{H}(\mathbf{r},t)\cdot\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}\right]\mathrm{d}V\,. \tag{2.24}$$

For the special case of linear dispersionless isotropic media this simplifies to

$$\int_{\partial V} \mathbf{S}(\mathbf{r},t)\,\mathrm{d}\mathbf{A} = -\frac{\partial}{\partial t}\int_V u(\mathbf{r},t)\;\mathrm{d}V \tag{2.25}$$

with the local energy density

$$u(\mathbf{r},t) = \frac{1}{2}\epsilon_0\epsilon(\mathbf{r})\;\mathbf{E}^2(\mathbf{r},t) + \frac{1}{2}\mu_0\mu(\mathbf{r})\;\mathbf{H}^2(\mathbf{r},t)\;. \tag{2.26}$$

Eq. (2.25) is known as the Poynting theorem. It states that the total energy flux per unit time across a closed surface must be equal to the negative rate of change of the internal energy inside the enclosed volume. In other words, if there is a total net energy flux per unit time across a closed surface, there must be a source or sink inside the volume (corresponding to gain or absorption).

2.5.2 Time averaging

The oscillation frequency of visible light is in the order of several 10^{14} Hz, which is significantly faster than the time scale of the processes that are usually of interest in optics. Therefore, the time average of the physical fields or dependent quantities are generally far more important. In the following, it is shown how the time averaged quantities can be calculated from the time dependent quantities. To this end, monochromatic fields of the form $\mathbf{F}(\mathbf{r},t) = \Re\left[\boldsymbol{\mathcal{F}}(\mathbf{r},\omega)\;e^{-i\omega t}\right]$ are assumed. Only the real part of the complex field is considered because, as a function of position and time, the physical fields must be real-valued.

Time average of an isolated monochromatic field

$$\begin{aligned}\langle\mathbf{F}(\mathbf{r},t)\rangle_t &= \langle\Re\left[\boldsymbol{\mathcal{F}}(\mathbf{r},\omega)\;e^{-i\omega t}\right]\rangle_t \\ &= \frac{1}{2}\langle\boldsymbol{\mathcal{F}}(\mathbf{r},\omega)\;e^{-i\omega t}\rangle_t + \text{c.c.} \\ &= \frac{1}{2}\boldsymbol{\mathcal{F}}(\mathbf{r},\omega)\;\langle e^{-i\omega t}\rangle_t + \text{c.c.} \\ &= 0\end{aligned} \tag{2.27}$$

The time average of an isolated monochromatic field is always zero.

Time average of the product of two monochromatic fields

$$
\begin{aligned}
&\langle \mathbf{F}_1(\mathbf{r},t)\cdot\mathbf{F}_2(\mathbf{r},t)\rangle_t \\
&= \langle \Re\left[\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\ e^{-i\omega t}\right]\cdot\Re\left[\boldsymbol{\mathcal{F}}_2(\mathbf{r},\omega)\ e^{-i\omega t}\right]\rangle_t \\
&= \frac{1}{4}\langle\left[\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\ e^{-i\omega t}+\text{c.c.}\right]\cdot\left[\boldsymbol{\mathcal{F}}_2(\mathbf{r},\omega)\ e^{-i\omega t}+\text{c.c.}\right]\rangle_t \\
&= \frac{1}{4}\langle\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2(\mathbf{r},\omega)\ e^{-2i\omega t}+\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\rangle_t\ +\ \text{c.c.} \\
&= \frac{1}{4}\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\ +\ \text{c.c.} \\
&= \frac{1}{2}\Re\left[\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\right]
\end{aligned}
\tag{2.28}
$$

From this result, for the absolute value squared of a field follows

$$
\langle|\mathbf{F}(\mathbf{r},t)|^2\rangle_t = \frac{1}{2}\left|\boldsymbol{\mathcal{F}}(\mathbf{r},\omega)\right|^2 . \tag{2.29}
$$

The same calculation leads to

$$
\langle\mathbf{F}_1(\mathbf{r},t)\times\mathbf{F}_2(\mathbf{r},t)\rangle_t = \frac{1}{2}\Re\left[\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\times\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\right] . \tag{2.30}
$$

Time average of the product of a monochromatic field and the time derivative of a monochromatic field

$$
\begin{aligned}
&\langle\mathbf{F}_1(\mathbf{r},t)\cdot\frac{\partial\mathbf{F}_2(\mathbf{r},t)}{\partial t}\rangle_t \\
&= \langle\Re\left[\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\ e^{-i\omega t}\right]\cdot\frac{\partial}{\partial t}\Re\left[\boldsymbol{\mathcal{F}}_2(\mathbf{r},\omega)\ e^{-i\omega t}\right]\rangle_t \\
&= \frac{1}{4}\langle\left[\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\ e^{-i\omega t}+\text{c.c.}\right]\cdot\left[-i\omega\boldsymbol{\mathcal{F}}_2(\mathbf{r},\omega)\ e^{-i\omega t}+\text{c.c.}\right]\rangle_t \\
&= \frac{1}{4}\langle-i\omega\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2(\mathbf{r},\omega)\ e^{-2i\omega t}+i\omega\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\rangle_t\ +\ \text{c.c.} \\
&= \frac{1}{4}i\omega\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\ +\ \text{c.c.} \\
&= \frac{1}{2}\Re\left[i\omega\boldsymbol{\mathcal{F}}_1(\mathbf{r},\omega)\cdot\boldsymbol{\mathcal{F}}_2^*(\mathbf{r},\omega)\right]
\end{aligned}
\tag{2.31}
$$

From this result, it follows

$$
\langle\mathbf{F}(\mathbf{r},t)\cdot\frac{\partial}{\partial t}\mathbf{F}(\mathbf{r},t)\rangle_t = 0 . \tag{2.32}
$$

2.5.3 Local absorption and gain

By time averaging Eq. (2.24) and applying Eq. (2.31), one obtains

$$\begin{aligned}&\langle \int_{\partial V} \mathbf{S}\left(\mathbf{r},t\right) \mathrm{d}\mathbf{A} \rangle_t \\ &= -\int_V \frac{1}{2} \Re\left[i\omega \boldsymbol{\mathcal{E}}\left(\mathbf{r},\omega\right) \cdot \boldsymbol{\mathcal{D}}^*\left(\mathbf{r},\omega\right) + i\omega \boldsymbol{\mathcal{H}}\left(\mathbf{r},\omega\right) \cdot \boldsymbol{\mathcal{B}}^*\left(\mathbf{r},\omega\right) \right] \mathrm{d}V \,. \end{aligned} \tag{2.33}$$

The integrand on the right side denotes the local absorption density. In the special case of linear isotropic media one obtains

$$\begin{aligned}&\langle \int_{\partial V} \mathbf{S}\left(\mathbf{r},t\right) \mathrm{d}\mathbf{A} \rangle_t \\ &= -\int_V \frac{\omega}{2} \left\{ \epsilon_0 \Im\left[\epsilon\left(\mathbf{r},\omega\right)\right] \left|\boldsymbol{\mathcal{E}}\left(\mathbf{r},\omega\right)\right|^2 + \mu_0 \Im\left[\mu\left(\mathbf{r},\omega\right)\right] \left|\boldsymbol{\mathcal{H}}\left(\mathbf{r},\omega\right)\right|^2 \right\} \mathrm{d}V \,. \end{aligned} \tag{2.34}$$

One observes that the amount of absorption of a medium is linearly dependent on the imaginary part of the electric permittivity and magnetic permeability. A positive imaginary part describes absorption, whereas a negative imaginary part corresponds to gain.

2.6 Wave equation and Helmholtz equation

Consider a volume, which is homogeneously filled with a linear dispersionless isotropic medium. Inside this volume, the electric permittivity and the magnetic permeability are constant in space. Therefore, from Eqs. (2.20) one obtains the relations

$$\mathbf{D}\left(\mathbf{r},t\right) = \epsilon_0 \, \epsilon \, \mathbf{E}\left(\mathbf{r},t\right) \tag{2.35a}$$

$$\mathbf{B}\left(\mathbf{r},t\right) = \mu_0 \mu \, \mathbf{H}\left(\mathbf{r},t\right) \,. \tag{2.35b}$$

The wave equation can be directly derived from Maxwell's equations. To this end, calculate the rotation of Eq. (2.2a),

$$\nabla \times \left[\nabla \times \mathbf{E}\left(\mathbf{r},t\right)\right] \quad = \quad -\nabla \times \frac{\partial}{\partial t} \left[\mathbf{B}\left(\mathbf{r},t\right)\right] \,. \tag{2.36}$$

Rewrite the left side and, on the right side replace $\mathbf{B}\left(\mathbf{r},t\right)$ using Eq. (2.35b),

$$\nabla \left[\nabla \cdot \mathbf{E}\left(\mathbf{r},t\right)\right] - \nabla^2 \mathbf{E}\left(\mathbf{r},t\right) \quad = \quad -\mu_0 \mu \frac{\partial}{\partial t} \left[\nabla \times \mathbf{H}\left(\mathbf{r},t\right)\right] \,. \tag{2.37}$$

On the left side, the divergence of the electric field is zero, since $\nabla \cdot \mathbf{D}\left(\mathbf{r},t\right) = \nabla \cdot \left[\epsilon_0 \epsilon \, \mathbf{E}\left(\mathbf{r},t\right)\right] = \epsilon_0 \epsilon \, \nabla \cdot \mathbf{E}\left(\mathbf{r},t\right) = 0$. On the right side, substitute $\nabla \times \mathbf{H}\left(\mathbf{r},t\right)$ using

Eq. (2.2b),

$$\nabla^2 \mathbf{E}(\mathbf{r},t) = \mu_0\mu \frac{\partial}{\partial t}\left[\frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t}\right]. \tag{2.38}$$

By replacing $\mathbf{D}(\mathbf{r},t)$ using Eq. (2.35a), one obtains the wave equation,

$$\nabla^2 \mathbf{E}(\mathbf{r},t) - \left(\frac{n}{c_0}\right)^2 \frac{\partial^2}{\partial t^2}\mathbf{E}(\mathbf{r},t) = 0, \tag{2.39}$$

where $n = \sqrt{\epsilon\mu}$ is the refractive index and $c_0 = \sqrt{\epsilon_0\mu_0}^{-1}$ denotes the speed of light in vacuum. By transforming the electromagnetic fields into the temporal Fourier domain, one obtains the vectorial Helmholtz equation,

$$\nabla^2 \boldsymbol{\mathcal{E}}(\mathbf{r},\omega) + \left(\frac{n\omega}{c_0}\right)^2 \boldsymbol{\mathcal{E}}(\mathbf{r},\omega) = 0. \tag{2.40}$$

As one can see, the three components of the electromagnetic field entirely decouple in this equation. That is why also the scalar Helmholtz equation is of high relevance in optics. It has the identical form as Eq. (2.40), whereas the electric field is assumed to be a scalar quantity.

2.6.1 Plane wave decomposition

Plane waves are a complete and orthogonal set of solutions of the wave equation, Eq. (2.39). They take the form[23]

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E_0}\, e^{i(\mathbf{kr}-\omega t)}, \tag{2.41}$$

where $\mathbf{k} = (k_x, k_y, k_z)^T$ is the wave vector, which defines the propagation direction. It can be easily checked that a valid solution fulfills the dispersion relation,

$$|\mathbf{k}| = nk_0 = \frac{n\omega}{c_0}, \tag{2.42}$$

where

$$k_0 = \frac{\omega}{c_0} = \frac{2\pi}{\lambda} \tag{2.43}$$

is the wave vector in vacuum, and λ is the vacuum wavelength[1]. Consequently, the spatial components of the wave vector are not independent of each other. One often defines the axial component $k_z = \pm\sqrt{n^2k_0^2 - \mathbf{k}_\perp^2}$ as a function of the lateral components. Since the differential equation is only solved for this particular choice of k_z, in optics one usually Fourier transforms the electromagnetic fields only in the

[1] $\mathbf{k}$ is related to the spatial frequencies $\boldsymbol{\nu}$ via $\mathbf{k} = 2\pi\boldsymbol{\nu}$.

two coordinates x and y,

$$\widetilde{\boldsymbol{\mathcal{E}}}\left(\mathbf{k}_{\perp}, z\right) = \int_{-\infty}^{+\infty} \boldsymbol{\mathcal{E}}\left(\mathbf{r}_{\perp}, z\right) \, e^{-i\mathbf{k}_{\perp}\cdot\mathbf{r}_{\perp}} \mathrm{d}^2 r_{\perp} \,, \tag{2.44}$$

where the ω-dependence of the quantities is dropped. Eq. (2.44) is called the 'plane wave decomposition' or 'angular spectrum decomposition' of the electric field. The latter is illustrated in Fig. 2.3.

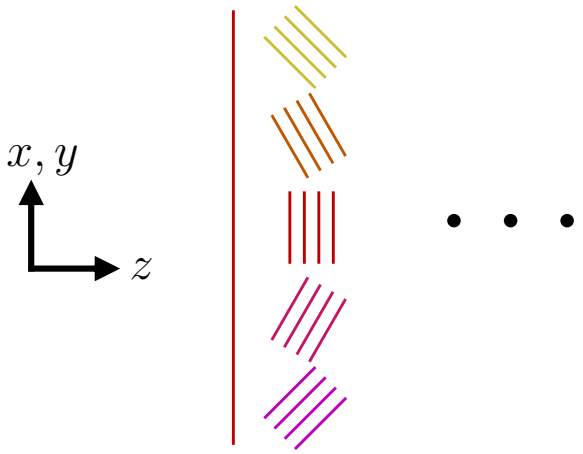


Figure 2.3: The angular spectrum decomposition of the electromagnetic field corresponds to a superposition of plane waves with different amplitudes (colors) and propagation directions.

If $k_z \in \mathbb{R}$, the corresponding plane wave is called 'propagating', since $e^{ik_z z}$ is a pure phase factor. If $k_z \in i\mathbb{R}$, the corresponding plane wave is called 'evanescent', since $e^{ik_z z}$ is an exponentially increasing or decreasing amplitude factor. One can think of evanescent waves as the waves with propagation angles larger than 90°. If k_z is some other complex number, there must be either absorption or gain in the medium of propagation - in this case, the distinction between 'propagating' and 'evanescent' waves is more intricate.

2.6.2 Spherical wave decomposition

In spherical coordinates, (r, θ, φ), the most general solution of the scalar Helmholtz equation is[44]

$$\mathcal{E}\left(\mathbf{r}, \omega\right) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \left[a_{lm} j_l\left(kr\right) + b_{lm} y_l\left(kr\right)\right] \, Y_l^m\left(\theta, \varphi\right) \,, \tag{2.45}$$

where $j_l\left(kr\right)$ and $y_l\left(kr\right)$ are the spherical Bessel functions, and $Y_l^m\left(\theta, \varphi\right)$ are the spherical harmonics, which represent a complete set of orthogonal functions on the

sphere[45]. The lowest mode of this expansion, i.e., $l = m = 0$, is well-known as 'the' spherical wave

$$\mathcal{E}(\mathbf{r}, \omega) \propto \frac{e^{ikr}}{r}, \tag{2.46}$$

which is rotationally symmetric.

2.7 Optical propagation algorithms

2.7.1 Development of numerical optics

As more and more computer processing power became available in the second half of the 20$^{\text{th}}$ century, scientific research increasingly dealt with the numerical analysis of physical problems. In optics, the numerical modeling of propagating electromagnetic fields was of particular interest. In the 1950s, several basic approaches already existed as, for instance, the propagation via the Kirchhoff diffraction integral[46], the Sommerfeld diffraction integral[47], or the angular spectrum (AS) propagation method[48, 49]. The mentioned algorithms, which were by the way all proven mathematically equivalent by Weyl[50], are representatives of what is commonly known as 'scalar propagation algorithms', because all components of the vector-valued electromagnetic fields are treated equally, or as it is sometimes formulated: the vectorial character of the electromagnetic fields is 'neglected'. However, since these propagation algorithms lacked the ability to describe both polarization and reflection, the interest in fully vectorial solvers was quite high. In 1966, for the first time such an algorithm was presented by Yee, the FDTD method[19], which operates in the time and spatial domain. As its direct counterpart in the spatial and temporal frequency domain, the RCWA was published by Moharam and Gaylord in 1981[1]. Both the FDTD and the RCWA could, indeed, correctly predict reflections and treat polarization, but unfortunately, their efficiency in terms of memory requirements and numerical complexity was so poor, that only extremely small problem sizes with dimensions in the order of a few wavelengths could be analyzed. Therefore, also new scalar propagation algorithms were developed as, for example, the BPM by Feit and Fleck in 1978[51] or the WPM in 1993 by Brenner and Singer[52]. Many variations and combinations of the different algorithms, and also extensions to vector-valued fields exist[53, 54]. Generally, it can be said that, even today, scalar propagation methods are still extensively applied if polarization and reflection effects play a minor role. Due to their severe time and memory constraints, exact Maxwell solvers like the FDTD or the RCWA are only used if truly necessary. The latter is usually given when the characteristic length scales of the refractive index structures become significantly smaller than the wavelength of light, which nowadays is most often the case in structures written with state-of-the-art lithographic devices.

Optical propagation algorithms can be subdivided into several groups of methods.

There are at least five different categories.

1. Homogeneous and inhomogeneous medium propagation
2. Exact and approximate methods
3. Scalar and vectorial methods
4. Paraxial and high numerical aperture (NA) methods
5. Position space and frequency space methods

Fig. 2.4 shows a sorted list of several selected established methods. Of course, this list is not complete. It is only meant for the reader to obtain a quick overview over the categorization of the most important methods in the field. In the following, several representatives are introduced in more detail.

2.7.2 Homogeneous medium propagation

Angular spectrum propagation

The AS propagation was first developed by [48, 49] as a scalar propagation method. Even so, it can be directly extended to vectorial electromagnetic fields[56]. In temporal and spatial Fourier space, the propagation of each individual plane wave is represented by the multiplication with the corresponding propagator given in Eq. (2.41). Therefore, the electromagnetic field is decomposed into its angular spectrum, as illustrated in Fig. 2.3. The propagation of a superposition of plane waves along the z-axis is then carried out via

$$\boldsymbol{\mathcal{E}}\left(\mathbf{r}_\perp, z\right) = \int_{-\infty}^{+\infty} \widetilde{\boldsymbol{\mathcal{E}}}\left(\mathbf{k}_\perp, z=0\right)\, e^{\pm i\sqrt{n^2k_0^2-\mathbf{k}_\perp^2}\,z}\, e^{i\mathbf{k}_\perp\cdot\mathbf{r}_\perp} \frac{\mathrm{d}^2k_\perp}{\left(2\pi\right)^2}\,, \tag{2.47}$$

where the sign $\pm$ determines the propagation direction[23]. In the actual implementation on a computer, the assignment in Eq. (2.47) is normally carried out via two 'fast Fourier transformations' (FFTs) and a multiplication,

$$\boldsymbol{\mathcal{E}}\left(\mathbf{r}_\perp, z\right) = \mathscr{F}^{-1}\left\{\mathscr{F}\left[\boldsymbol{\mathcal{E}}\right]\left(\mathbf{k}_\perp, z=0\right)\, e^{\pm i\sqrt{n^2k_0^2-\mathbf{k}_\perp^2}\,z}\right\}\left(\mathbf{r}_\perp\right)\,. \tag{2.48}$$

The numerical complexity of the AS propagation is dominated by the numerical complexity for carrying out the FFTs, which is of the order $\mathcal{O}\left(\bar{N}\log\bar{N}\right)$, $\bar{N}$ representing the number of Fourier modes in the calculation. The memory requirement is of the order $\mathcal{O}\left(\bar{N}\right)$. It can be shown that, due to the finite sampling of the electromagnetic field and the propagator, the propagation distance is limited to[58],

$$z \leq z_{\mathrm{crit}} = \frac{P_x\,\mathrm{d}x}{\lambda}\,, \tag{2.49}$$

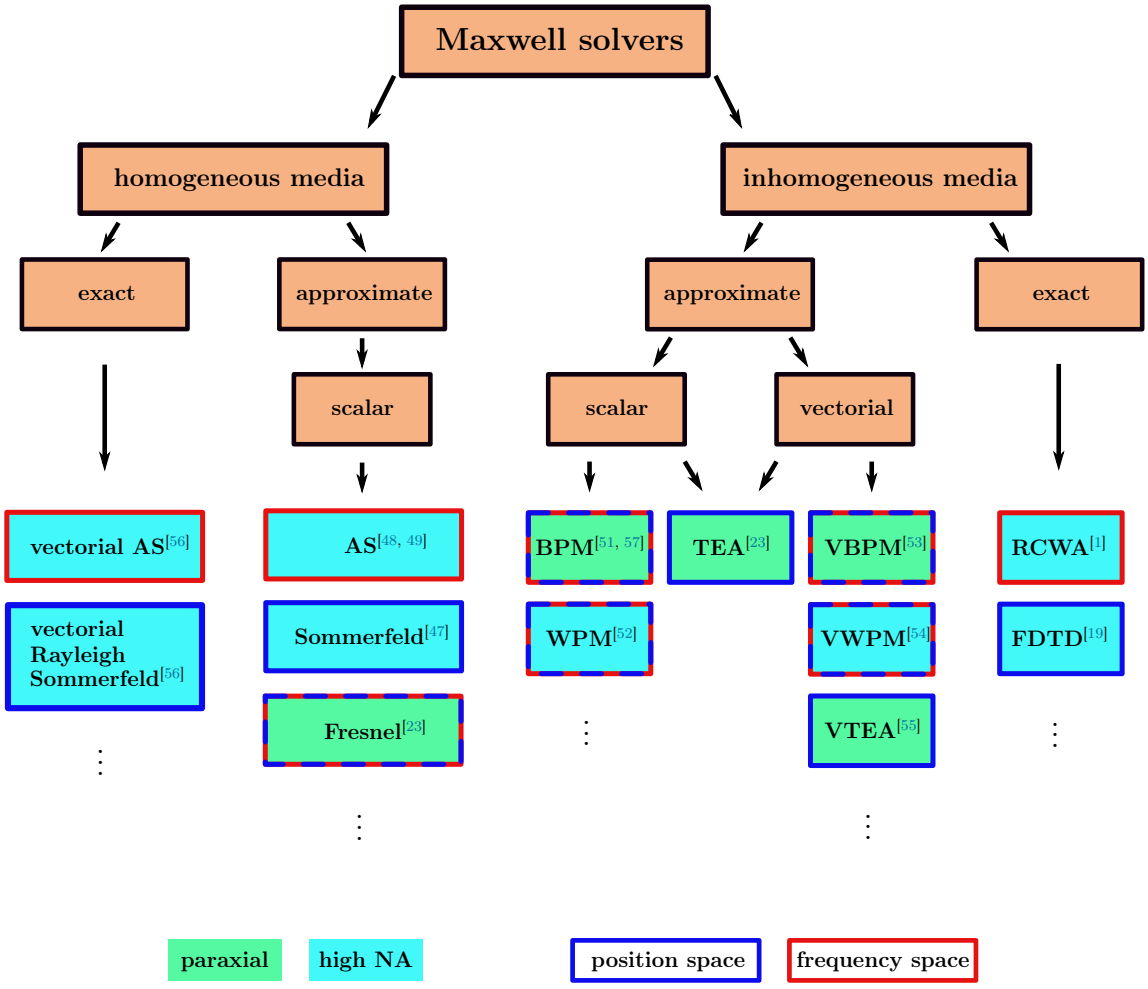


Figure 2.4: Categorization of the most important numerical propagation algorithms. The fill color of the boxes indicates whether the propagation algorithm is paraxial (green) or high NA (blue). The border color of the boxes indicates whether the method operates in position space (blue) or frequency space (red). For some methods, both is possible or operations are performed in both spaces. In these cases, the border color is a blue-red dashed line.

where P_x is the diameter of the computational grid area and $\mathrm{d}x$ is the position space sampling. Hence, the AS propagation is a near field propagation method.

Sommerfeld propagation

The propagation according to Sommerfeld[47] is a scalar propagation method, which is realized by a convolution in position space,

$$\boldsymbol{\mathcal{E}}\left(\mathbf{r}_\perp, z\right) = -\frac{1}{2\pi}\int_{-\infty}^{+\infty} \boldsymbol{\mathcal{E}}\left(\mathbf{r}'_\perp, z'=0\right)\ \frac{\partial}{\partial z}\left[\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\right]\Bigg|_{z'=0} \mathrm{d}^2 r'_\perp\,, \tag{2.50}$$

in contrast to the multiplication in Fourier space just seen in the AS approach. The mathematical equivalence of both methods was proven by Weyl[50]. As can be seen from Eq. (2.50), each point of the field at $z = 0$ can be interpreted as the source not of spherical waves as supposed by Huygens, but of the z-derivative of spherical waves as illustrated in Fig. 2.5. An alternative and mathematically equivalent form

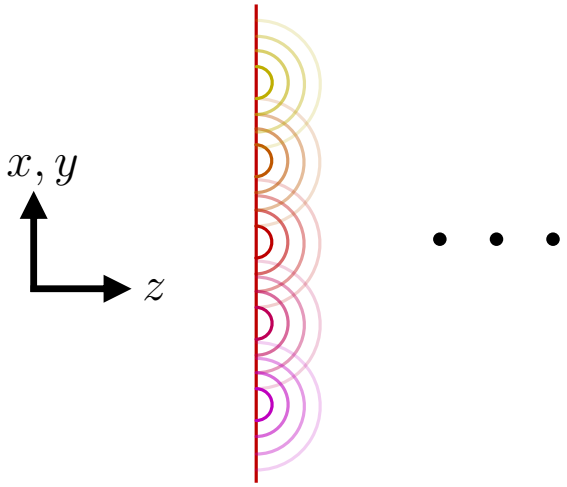


Figure 2.5: Sommerfeld diffraction: each point of the field at $z = 0$ can be interpreted as the source of the z-derivative of a spherical wave.

of Eq. (2.50) is given by

$$\boldsymbol{\mathcal{E}}\left(\mathbf{r}_\perp, z\right) = -\frac{1}{2\pi}\int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial z}\boldsymbol{\mathcal{E}}\right]\left(\mathbf{r}'_\perp, z'=0\right)\ \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\Bigg|_{z'=0} \mathrm{d}^2 r'_\perp\,. \tag{2.51}$$

One way of implementing these equations on a computer is carrying out the convolution as a matrix multiplication. The numerical complexity of this procedure is then, however, of the order $\mathcal{O}\left(\bar{N}^2\right)$. An alternative way is a similar procedure as for the AS propagation, but where the propagator is determined by the FFT of the sampled position space propagator. In this case, the numerical complexity comes back

down to $\mathcal{O}(\bar{N}\log\bar{N})$. The propagation according to Sommerfeld then is a far field propagation method, i.e., $z \geq z_{\text{crit}}$[58]. Vectorial extensions like the one presented in [56] also exist.

2.7.3 Structured medium propagation

Thin element approximation

The TEA is probably the most simplified way to model the transition of light through a structured object. A good description can be found in [23]. The principle is illustrated in Fig. 2.6. The following two assumptions are made. 1) The incident

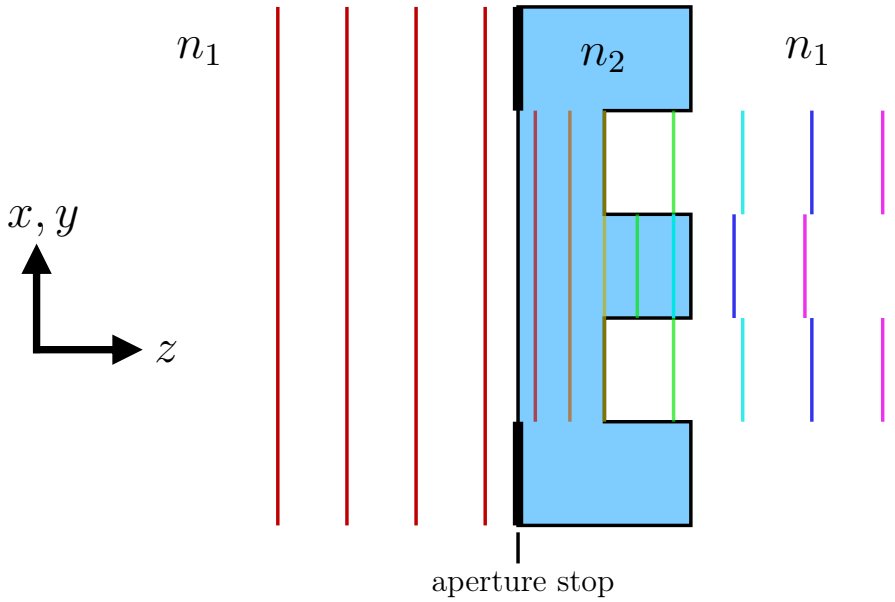


Figure 2.6: TEA: an object is approximated by a transmission function $t(\mathbf{r}_\perp) = a(\mathbf{r}_\perp)\exp[i\phi(\mathbf{r}_\perp)]$, whose amplitude corresponds to absorption (or gain), and whose phase describes the optical path differences accumulated by the traversing light.

electromagnetic field contains only small propagation angles (paraxial approximation). 2) The diffraction object is infinitely thin, i.e., all changes of the field in terms of amplitude and phase occur in only one single definite location in the object plane. Consequently, the electromagnetic field in front and behind the object are connected via the complex transmission function $t(\mathbf{r}_\perp) = a(\mathbf{r}_\perp)\exp[i\phi(\mathbf{r}_\perp)]$,

$$\boldsymbol{\mathcal{E}}'(\mathbf{r}_\perp, z) = a(\mathbf{r}_\perp)\, e^{i\phi(\mathbf{r}_\perp)}\, \boldsymbol{\mathcal{E}}(\mathbf{r}_\perp, z)\ . \tag{2.52}$$

The changes in the amplitude, $a(\mathbf{r}_\perp)$, model absorption (or gain), for instance at an aperture stop. The changes in the phase, $\phi(\mathbf{r}_\perp)$, correspond to optical path differences, for instance a DOE or a thin lens. The treatment of a diffraction object in TEA becomes problematic if the object becomes too thick, or if the lateral coupling within the object plays an important role.

Beam propagation method

The BPM method was initially developed by Feit and Fleck in 1978 for the scalar simulation of microoptical components and waveguides[51]. Both spatial and frequency domain BPMs exist. Both cases have in common the assumption that the refractive index variations and the propagation angles are and remain small throughout the propagation process. In the spatial domain solution, the scalar Helmholtz equation is solved with a separation ansatz in slowly varying envelope approximation[51]. However, also a very elegant split step solution exist[57], which is illustrated in Fig. 2.7 and shortly outlined here.

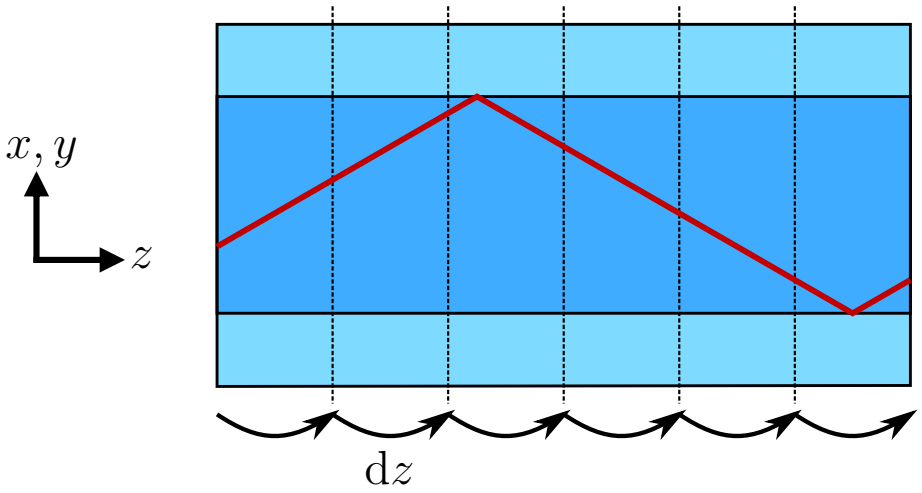


Figure 2.7: The BPM propagates the electromagnetic field in small axial steps $\mathrm{d}z$ through the structured medium. In each step, the propagation operator is split into an AS propagation and a spatial transmission function.

For the mentioned assumptions, operator splitting can be applied to the exact propagator, which then separates into a purely position-dependent and a purely frequency-dependent part,

$$\begin{aligned}\mathcal{E}\left(\mathbf{r}_\perp, z+\mathrm{d}z\right) &= \int_{-\infty}^{+\infty} \widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right)\, e^{i\sqrt{n^2(\mathbf{r}_\perp)k_0^2-\mathbf{k}_\perp^2}\,\mathrm{d}z}\, e^{i\mathbf{k}_\perp\cdot\mathbf{r}_\perp}\frac{\mathrm{d}^2k_\perp}{(2\pi)^2} \\ &\approx e^{ik_0\,\mathrm{d}z\,\delta n(\mathbf{r}_\perp)} \int_{-\infty}^{+\infty} \widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right)\, e^{i\sqrt{\bar{n}^2k_0^2-\mathbf{k}_\perp^2}\,\mathrm{d}z}\, e^{i\mathbf{k}_\perp\cdot\mathbf{r}_\perp}\frac{\mathrm{d}^2k_\perp}{(2\pi)^2}\,, \end{aligned} \tag{2.53}$$

where $\bar{n} = \langle n\left(\mathbf{r}_\perp\right)\rangle_{x,y}$ is the averaged refractive index in x and y and $\delta n\left(\mathbf{r}_\perp\right) = n\left(\mathbf{r}_\perp\right) - \bar{n}$. The propagation through the structured medium is done in small axial steps $\mathrm{d}z$. To this end, the electric field is alternately transformed to the frequency

space and back,

$$\begin{aligned}
\mathcal{E}\left(\mathbf{r}_\perp, z\right) &\xrightarrow{\mathscr{F}} \widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right) \\
\widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right) &\longrightarrow \widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right) e^{i\sqrt{\bar{n}^2 k_0^2 - \mathbf{k}_\perp^2}\,\mathrm{d}z} =: \widetilde{\mathcal{E}}'\left(\mathbf{k}_\perp, z+\mathrm{d}z\right) \\
\widetilde{\mathcal{E}}'\left(\mathbf{k}_\perp, z+\mathrm{d}z\right) &\xrightarrow{\mathscr{F}^{-1}} \mathcal{E}'\left(\mathbf{r}_\perp, z+\mathrm{d}z\right) \\
\mathcal{E}'\left(\mathbf{r}_\perp, z+\mathrm{d}z\right) &\longrightarrow e^{ik_0\,\mathrm{d}z\,\delta n(\mathbf{r}_\perp)}\,\mathcal{E}'\left(\mathbf{r}_\perp, z+\mathrm{d}z\right) =: \mathcal{E}\left(\mathbf{r}_\perp, z+\mathrm{d}z\right) .
\end{aligned} \tag{2.54}$$

The Fourier transformations can be implemented using the FFT. Therefore, the numerical complexity of the BPM is of the order $\mathcal{O}\left(\bar{N}\log\bar{N}\right)$. Upon each propagation step, a small numerical error is accumulated. Therefore, the reliable propagation distance is limited.

Wave propagation method

In contrast to the BPM, the WPM[52] keeps the exact propagator and no operator splitting is applied. As in the BPM, the propagation through the structured medium is done in small axial steps $\mathrm{d}z$. To this end, the electric field is again alternately transformed to the frequency space and back,

$$\begin{aligned}
\mathcal{E}\left(\mathbf{r}_\perp, z\right) &\xrightarrow{\mathscr{F}} \widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right) \\
\mathcal{E}\left(\mathbf{r}_\perp, z+\mathrm{d}z\right) &\xrightarrow{'\mathscr{F}^{-1}{}'} \int_{-\infty}^{+\infty} \widetilde{\mathcal{E}}\left(\mathbf{k}_\perp, z\right) e^{i\sqrt{n^2(\mathbf{r}_\perp)k_0^2 - \mathbf{k}_\perp^2}\,\mathrm{d}z}\, e^{i\mathbf{k}_\perp\cdot\mathbf{r}_\perp} \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} .
\end{aligned} \tag{2.55}$$

Since the propagator does not separate into purely spatial and frequency dependent components, the inverse Fourier transform cannot be calculated via the FFT, but must be implemented as a matrix multiplication. Therefore, the WPM has the numerical complexity $\mathcal{O}\left(\bar{N}^2\right)$. For small numbers of different refractive indices in lateral direction, the procedure may be accelerated up to, in the best case, $\mathcal{O}\left(\bar{N}\log\bar{N}\right)$[59]. With the WPM, also high NA components can be calculated. The advantages over the BPM lie, for instance, in a much more reliable simulation of strongly focused fields. However, as the BPM, a similar small numerical error is accumulated upon every propagation step. Therefore, the reliable propagation distance is also limited.

Finite-difference time-domain method

The FDTD was first published by Yee in 1966[19]. It is a fully vectorial and exact optical propagation method, i.e., it does not assume approximations prior to solving

Maxwell's equations. As the name already suggests, the system of equations is discretized in both the time and all spatial coordinates by finite difference quotients. Thereby, the FDTD implements a second order accurate leapfrog scheme[60] for the advancement of the electromagnetic fields in time. In his publication, Yee furthermore introduced the so-called 'Yee grid', which samples the electric and the magnetic fields not at the same spatial positions, but always half a grid cell from each other in each direction. This way, Yee achieved inherently divergence-free fields. The FDTD has the following advantages.

+ There is no limitation to narrow bands of frequencies. The impulse response of the materials to all frequencies contained in the electromagnetic spectrum can be modeled in one single step. This is particularly interesting for the simulation of incident short (and therefore broadband) pulses.

+ Nonlinear materials can be handled in a very uncomplicated manner.

+ The algorithm is easily implementable on a parallel architecture.

On the other hand, the disadvantages of the FDTD are the following.

- Only very small computation volumes can be handled. The reason is that any calculation area must be sampled in the order of the smallest electromagnetic wavelength, which quickly leads to very high numbers of sampling points.

- The response function of the materials to an external electromagnetic field must be explicitly modeled in the time domain. Depending on the material, this calculation is computationally expensive.

- Upon the discretization, the Courant-Friedrichs-Lewy (CFL) condition[61] must be satisfied. Otherwise, the Leapfrog scheme may become unstable.

- Boundary conditions need to be introduced at the boundaries of the calculation domain.

- The electromagnetic field at a distant point outside the computation domain cannot be determined in a straightforward manner.

The FDTD can be referred to as the direct counterpart of the RCWA, which solves Maxwell's equations in the temporal and spatial frequency domain (cf. Chapter 3).

2.8 Normalized fields

The electric and magnetic fields occurring in Maxwell's equations are dimensionful numbers. In many cases, it is, however, favorable to rescale these quantities such,

that the equations for the scaled fields contain no or as little as possible dimensionful constants. As suggested in [62], one defines

$$\begin{aligned} \boldsymbol{\mathcal{E}}(\mathbf{r},\omega) &= E_0\,\boldsymbol{\mathcal{E}}'(\mathbf{r},\omega) \\ \boldsymbol{\mathcal{H}}(\mathbf{r},\omega) &= \sqrt{\frac{\epsilon_0}{\mu_0}}E_0\,\boldsymbol{\mathcal{H}}'(\mathbf{r},\omega)\ , \end{aligned} \tag{2.56}$$

where E_0 carries the dimension of the electric field and the primed functions are now dimensionless or 'normalized' fields. For linear dispersive media, Maxwell's equations in the temporal frequency domain, Eqs. (2.4), then transform to

$$\nabla\times\boldsymbol{\mathcal{E}}'(\mathbf{r},\omega) = +ik_0\hat{\mu}(\mathbf{r})\,\boldsymbol{\mathcal{H}}'(\mathbf{r},\omega) \tag{2.57a}$$

$$\nabla\times\boldsymbol{\mathcal{H}}'(\mathbf{r},\omega) = -ik_0\hat{\epsilon}(\mathbf{r})\,\boldsymbol{\mathcal{E}}'(\mathbf{r},\omega) \tag{2.57b}$$

$$\nabla\cdot[\hat{\epsilon}(\mathbf{r})\,\boldsymbol{\mathcal{E}}'(\mathbf{r},\omega)] = 0 \tag{2.57c}$$

$$\nabla\cdot[\hat{\mu}(\mathbf{r})\,\boldsymbol{\mathcal{H}}'(\mathbf{r},\omega)] = 0\,. \tag{2.57d}$$

Starting from Chapter 3, all the occurring electromagnetic fields are defined as normalized fields, unless explicitly stated otherwise.

3 | Rigorous coupled-wave analysis

In this chapter, the framework of the RCWA is derived. A detailed outline of the scientific background and an overview of the advancements in the field from the early 1980s until today have already been given in the introductory part of this thesis. The organization is as follows. In Sect. 3.1, the typical problem geometry treated by the RCWA is defined. Furthermore, the mathematical formalism used to describe the grating and the electromagnetic fields is introduced. Sect. 3.2 presents the function arithmetic in Fourier truncated systems, including the product-convolution relation and Li's factorization rules in one and two dimensional systems. In Sect. 3.3, Maxwell's equations are expressed in terms of the frequency domain quantities for the case of linear isotropic and linear anisotropic media. The general solution of this system of equations is derived in Sect. 3.4. In Sect. 3.5, several methods for enforcing the boundary conditions are presented and compared. The numerical complexity and memory requirement of the RCWA is considered in Sect. 3.7. Sect. 3.6 discusses the consistency requirements of the RCWA algorithm for tilted structured light incidence. A sample RCWA simulation of a SNOM measurement is shown in Sect. 3.8.

3.1 Problem definition

If not otherwise stated, the reader may also refer to [5, 63] for the definitions and derivations shown in this section.

3.1.1 Geometry

The geometry of a typical RCWA problem is illustrated in Fig. 3.1. A superposition of plane waves, which is characterized by the Fourier coefficients $\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{mn}$, is incident on a lamellar grating structure of thickness d, which is periodic in x and y with periods P_x and P_y. The incident light is diffracted into different reflected ($\widetilde{\boldsymbol{\mathcal{R}}}_{mn}$) and transmitted ($\widetilde{\boldsymbol{\mathcal{T}}}_{mn}$) mode channels. The grating properties and the incident light are given. From this, the reflected and transmitted Fourier mode coefficients can be determined with the RCWA.

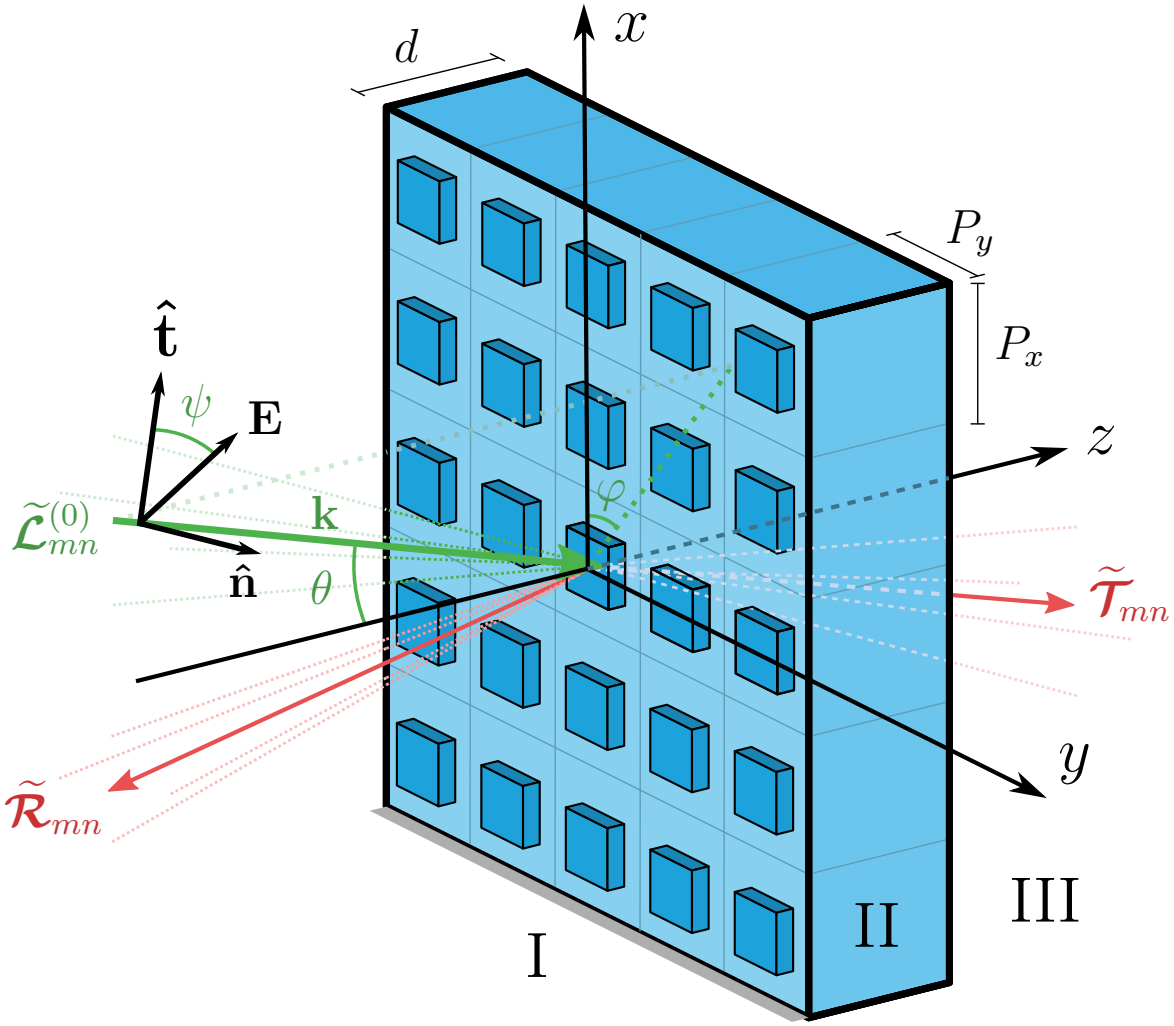


Figure 3.1: RCWA grating problem. A superposition of plane waves $\widetilde{\mathcal{L}}^{(0)}_{mn}$ (green), whose central mode is parameterized by the angles θ and φ, is incident on an x-y-periodic grating structure (blue) of thickness d. The polarization of the wave is given by the angle ψ. The mode coefficients of the transmitted and reflected light, $\widetilde{\mathcal{T}}_{mn}$) and $\widetilde{\mathcal{R}}_{mn}$ (red), are unknown and shall be calculated.

3.1.2 Refractive index expansion

Along the z-axis, space is divided into three different regions. Region I ($z < 0$) and region III ($z > d$) are assumed to be linear isotropic homogeneous media, i.e. the electric permittivity $\epsilon^{(\mathrm{I/III})}$ and the magnetic permeability $\mu^{(\mathrm{I/III})}$ are constant in space. Region II ($0 \leq z \leq d$) is a lamellar diffraction grating with the following properties.

1. The electric permittivity $\hat{\epsilon}(\mathbf{r})$ and the magnetic permeability $\hat{\mu}(\mathbf{r})$ are periodic functions of x and y with the fix periods P_x and P_y.

2. 'Staircase approximation': in z-direction, the grating is subdivided into L layers. In each layer $l = 1..L$, the electric permittivity $\hat{\epsilon}^{(l)}(x, y)$ and the magnetic permeability $\hat{\mu}^{(l)}(x, y)$ do not depend on z - cf. Fig. 3.2b.

3. 'Zigzag approximation': The functions $\hat{\epsilon}^{(l)}(x, y)$ and $\hat{\mu}^{(l)}(x, y)$ are approximated by a zigzag contour, whose segments are parallel to the x and y coordinate axes. This way, material discontinuities are either only in x direction or only in y direction[2] - cf. Fig. 3.2c.

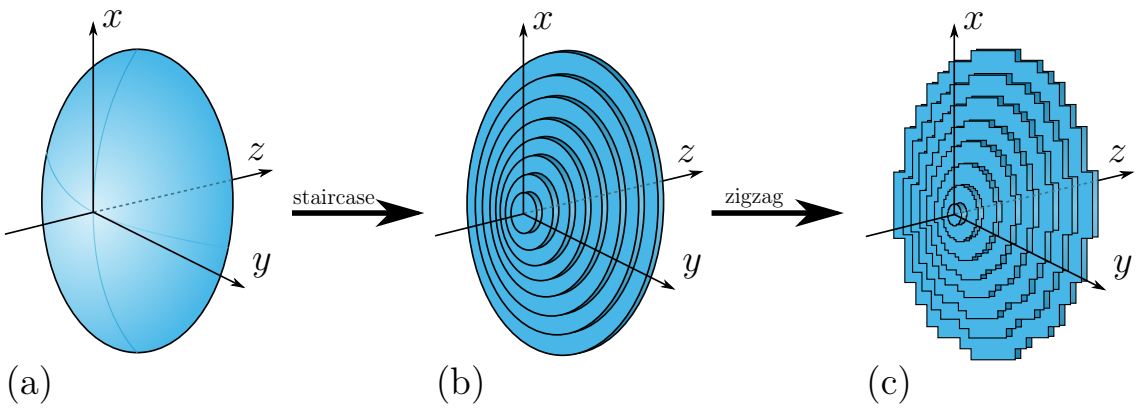


Figure 3.2: Continuous object with z-dependent permittivity and permeability distribution (a), which is approximated by the staircase approximation (b) and the zigzag approximation (c). In (b), the structure is subdivided into layers with piecewise constant material properties in z. In (c), the contour is additionally approximated by segments parallel to the x and y coordinate axes.

In the following, an upper index in parentheses always indicates a layer dependence. In this context, the quantity l represents the layer index. For the sake of consistent notation, $l = 0$ and $l = L + 1$ denote the incident and transmitted regions I and III.

[2] This approximation is needed when transforming Maxwell's equations with Li's factorization rules. A more general approach was suggested by [14, 16, 64], which works for arbitrary contour orientations. The latter goes, however, beyond the scope of this thesis and is, therefore, not shown here.

If the structure under consideration has a z-dependent permittivity or permeability distribution, it has to be approximated according to the staircase and zigzag approximation. Clearly, the higher the number of layers, the better the approximation of the original structure. The computational complexity then grows linearly with the number of layers in the grating.

The RCWA is a special case of the so-called group of 'modal methods'. The characteristic of this group of methods is that both the electromagnetic fields as well as the electric permittivity and magnetic permeability are expanded in a complete and orthogonal set of functions. In the RCWA, the latter are chosen to be the Fourier modes, which is why the RCWA is also called the 'Fourier modal method' (FMM). Since the grating structure is assumed to be periodic in x and y with periods P_x and P_y, the spectra of the electric permittivity and the magnetic permeability become discrete. Hence, the functions $\hat{\epsilon}^{(l)}(x,y)$ and $\hat{\mu}^{(l)}(x,y)$ can be expanded into a Fourier series,

$$\hat{\epsilon}^{(l)}(x,y) = \sum_{mn} \hat{\epsilon}^{(l)}_{mn} \exp\left\{2\pi i \left(\frac{m}{P_x}x + \frac{n}{P_y}y\right)\right\} \tag{3.1a}$$

$$\hat{\mu}^{(l)}(x,y) = \sum_{mn} \hat{\mu}^{(l)}_{mn} \exp\left\{2\pi i \left(\frac{m}{P_x}x + \frac{n}{P_y}y\right)\right\}, \tag{3.1b}$$

where the Fourier coefficients are given by

$$\hat{\epsilon}^{(l)}_{mn} = \frac{1}{P_x P_y} \int_{P_x}\int_{P_y} \hat{\epsilon}^{(l)}(x,y) \exp\left\{-2\pi i \left(\frac{m}{P_x}x + \frac{n}{P_y}y\right)\right\} \mathrm{d}x\,\mathrm{d}y \tag{3.2a}$$

$$\hat{\mu}^{(l)}_{mn} = \frac{1}{P_x P_y} \int_{P_x}\int_{P_y} \hat{\mu}^{(l)}(x,y) \exp\left\{-2\pi i \left(\frac{m}{P_x}x + \frac{n}{P_y}y\right)\right\} \mathrm{d}x\,\mathrm{d}y\,. \tag{3.2b}$$

In the standard RCWA by Moharam and Gaylord [1, 5], these coefficients are calculated analytically or numerically with a high precision. However, also other possibilities, like a sampling with the mode count number, exist. The advantages and disadvantages of these choices have been discussed in [17, 63].

3.1.3 Electromagnetic field expansion

Grating region

In the grating region, the electromagnetic field propagates within the lamellar diffraction grating, which is x-y-periodic with periods P_x and P_y. According to the Floquet-Bloch theorem, the most general solution of such a system are pseudo-periodic fields

of the form

$$\boldsymbol{\mathcal{E}}^{(\mathrm{II})}(\mathbf{r}) = \sum_{m,n} \widetilde{\boldsymbol{\mathcal{E}}}_{mn}(z)\ \exp\{ik_{x;mn}x + ik_{y;mn}y\} \tag{3.3a}$$

$$\boldsymbol{\mathcal{H}}^{(\mathrm{II})}(\mathbf{r}) = \sum_{m,n} \widetilde{\boldsymbol{\mathcal{H}}}_{mn}(z) \exp\{ik_{x;mn}x + ik_{y;mn}y\}\ , \tag{3.3b}$$

which is simply a Fourier series in the lateral coordinates. The wave vector components are given by

$$k_{x;mn} = k_{x0} + 2\pi\frac{m}{P_x} \qquad k_{x0} = k_0\sqrt{\epsilon^{(0)}\mu^{(0)}}\sin\theta\cos\varphi \tag{3.4a}$$

$$k_{y;mn} = k_{y0} + 2\pi\frac{n}{P_y} \qquad k_{y0} = k_0\sqrt{\epsilon^{(0)}\mu^{(0)}}\sin\theta\sin\varphi \tag{3.4b}$$

The offsets k_{x0} and k_{y0} correspond to a three-dimensional tilt of the entire mode set by the angles θ and φ in spherical coordinates.

Throughout this thesis, the tilde $\widetilde{(\cdots)}$ always indicates a field, which is Fourier transformed with respect to the x and/or y coordinate. In this context, the index m is always used for the Fourier coefficients with respect to x, n for the Fourier coefficients with respect to y. Also partly transformed fields are marked with the tilde. In this case, the transformed coordinate is recognizable by the indexing. In some situations, the arguments x and y, or the indexing m and n may be dropped in order to facilitate the readability of formulas. If a quantity lacks both indices, m and n, it is *always* Fourier transformed with respect to both coordinates x and y.

Incident and transmitted region

In the standard RCWA, the incident and transmitted regions are assumed to be isotropic and homogeneous media. Following Sect. 2.6.1, the most general solution of Maxwell's equations in such media is a superposition of forward and backward propagating plane waves,

$$\begin{aligned}\boldsymbol{\mathcal{E}}^{(0)}(\mathbf{r}) \quad = \quad & \sum_{mn} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{mn} \quad \exp\left\{i\left(k_{x;mn}x + k_{y;mn}y + k^{(0)}_{z;mn}z\right)\right\} \\ & + \sum_{mn} \widetilde{\boldsymbol{\mathcal{R}}}_{mn} \quad \exp\left\{i\left(k_{x;mn}x + k_{y;mn}y - k^{(0)}_{z;mn}z\right)\right\}\end{aligned} \tag{3.5a}$$

$$\begin{aligned}\boldsymbol{\mathcal{E}}^{(L+1)}(\mathbf{r}) = \quad & \sum_{mn} \widetilde{\boldsymbol{\mathcal{T}}}_{mn} \quad \exp\left\{i\left(k_{x;mn}x + k_{y;mn}y + k^{(L+1)}_{z;mn}z\right)\right\} \\ & + \sum_{mn} \widetilde{\boldsymbol{\mathcal{L}}}^{(L+1)}_{mn} \exp\left\{i\left(k_{x;mn}x + k_{y;mn}y - k^{(L+1)}_{z;mn}z\right)\right\}\ ,\end{aligned} \tag{3.5b}$$

where $\widetilde{\boldsymbol{\mathcal{L}}}_{mn}^{(0/L+1)}$, $\widetilde{\boldsymbol{\mathcal{R}}}_{mn}$ and $\widetilde{\boldsymbol{\mathcal{T}}}_{mn}$ are the Fourier coefficients of the incident, reflected and transmitted light, and the z-components of the wave vectors are calculated via

$$k_{z;mn}^{(0/L+1)} = \sqrt{k_0^2 \epsilon^{(0/L+1)} \mu^{(0/L+1)} - k_{x;mn}^2 - k_{y;mn}^2}\,. \tag{3.6}$$

It should be noted that, so far, RCWA calculations have only been conducted for one-sided light incidence, i.e. $\widetilde{\boldsymbol{\mathcal{L}}}_{mn}^{(L+1)} = \mathbf{0}$. In Chapter 5, the framework is extended for actual two-sided light incidence.

Structured illumination and mode set tilt

Part of this section about structured illumination is taken from the publication [A3].

In the first publications related to the RCWA, the illuminating light was always assumed to be an isolated plane wave at the origin of the frequency space, i.e. $\widetilde{\boldsymbol{\mathcal{L}}}_{mn}^{(0)} \propto \delta_{m0}\delta_{n0}$ [1, 5]. A tilt of the incident wave vector was then usually realized via a k-space shift by k_{x0} and k_{y0} according to Eq. (3.4). In position space, the incident wave is then

$$\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}(\mathbf{r}) = \begin{bmatrix} \cos\psi\cos\theta\cos\varphi - \sin\psi\sin\varphi \\ \cos\psi\cos\theta\sin\varphi + \sin\psi\cos\varphi \\ -\cos\psi\sin\varphi \end{bmatrix} e^{ik_0 n_\mathrm{I}(\sin\theta\cos\varphi\, x + \sin\theta\sin\varphi\, y + \cos\theta\, z)}\,, \tag{3.7}$$

where ψ defines the linear polarization angle of the plane wave. This situation is illustrated in Fig. 3.3a. However, realizing a tilt like this is normally sensible only for single mode light incidence as it is the case in the RCWA initially developed by [1, 5]. In many cases the incident field should be, however, not a plane wave, but some arbitrary structured illumination like for instance a focused or Gaussian beam. These types of illuminations can generally be represented as a superposition of multiple plane waves. In the past, simulations with structured illumination were often realized by carrying out as many RCWA calculations as there are incident Fourier modes [65]. However, this method is not only computationally very costly, but also not consistent as shown in [17] and Sect. 3.6. Instead, [17] suggests to realize tilted light incidence by directly exciting other than the central mode in the incident light Fourier mode vector (cf. Fig. 3.3b). In this case, the offsets defined in Eqs. (3.4a, 3.4b) are set to zero. This approach also allows the consistent description of structured illumination by simply changing the Kronecker-type input vector into a fully occupied Fourier mode vector. Sect. 3.6 investigates under which circumstances these two approaches are consistent.

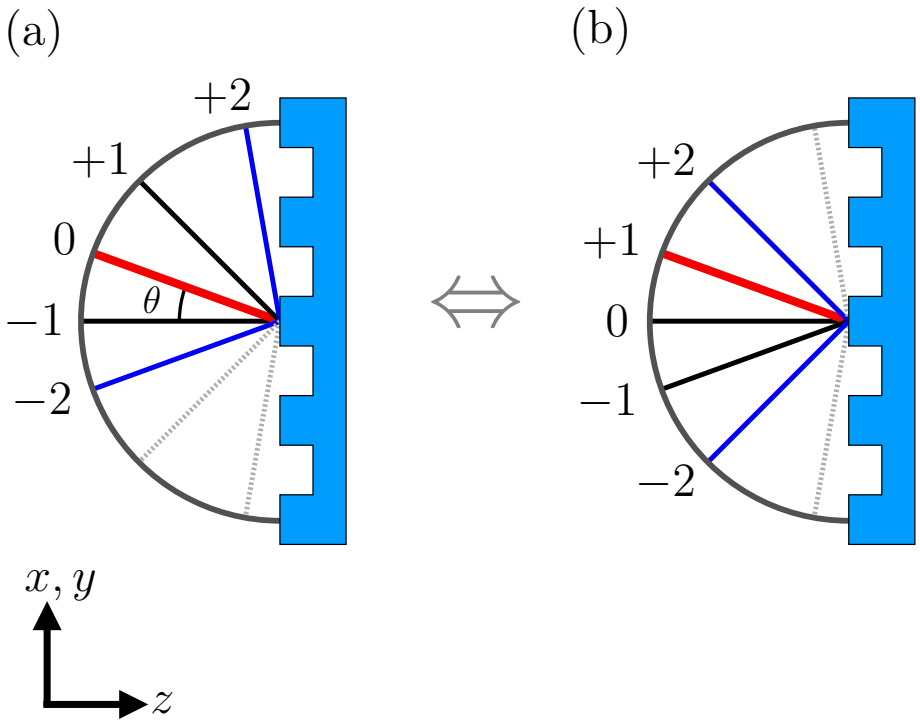


Figure 3.3: Two ways to model a tilted plane wave. (a) Excitation of the zero order harmonic with a shift in k-space, (b) excitation of mn^{th} order harmonic with a constant factor. The bold red line represents the excited mode.

3.1.4 Mode truncation

The sums that occur in the refractive index expansion and in the electromagnetic field expansion generally run over all integers $m, n = -\infty..\infty$. In the RCWA, these sums are truncated symmetrically around the zero mode, i.e. the indices run from $m = -M_0 \,..\, M_0$ and $n = -N_0 \,..\, N_0$, i.e., one does the transition

$$\sum_{mn} \to \sum_{m=-M_0}^{+M_0} \sum_{n=-N_0}^{+N_0} . \tag{3.8}$$

The one-dimensional number of modes in x and y are denoted by $M = 2M_0 + 1$ and $N = 2N_0 + 1$, respectively. In the following, the symmetric mode truncation is *always* implicitly assumed and not written out. In case that no mode truncation is assumed, i.e. the infinite system is regarded, this is explicitly indicated.

3.2 Function arithmetic in Fourier mode truncated systems

Due to the assumed periodicity of the grating and the electromagnetic fields, it is advantageous to rewrite Maxwell's equations in terms of the Fourier transformed

quantities. This transformation is often called the 'Fourier analysis of Maxwell's equations'. On the one hand, Maxwell's equations contain spatial derivative operators, which transform to a simple multiplication in Fourier space. On the other hand, Maxwell's equations also contain direct products of position space functions, namely $\hat{\epsilon}(\mathbf{r})\boldsymbol{\mathcal{E}}(\mathbf{r})$ and $\hat{\mu}(\mathbf{r})\boldsymbol{\mathcal{H}}(\mathbf{r})$, which need to be transformed accordingly. In this section, the Fourier analysis of such position space products is considered, and 'Li's factorization rules' are introduced.

3.2.1 Product-convolution relation

Given two periodic functions $f(x)$ and $g(x)$ with period P_x. The Fourier coefficients of these functions are

$$\widetilde{f}_m = \frac{1}{P_x}\int_{P_x} f(x)\exp\left\{-2\pi i\frac{m}{P_x}x\right\}\mathrm{d}x \tag{3.9a}$$

$$\widetilde{g}_m = \frac{1}{P_x}\int_{P_x} g(x)\exp\left\{-2\pi i\frac{m}{P_x}x\right\}\mathrm{d}x\,. \tag{3.9b}$$

The direct product in position space,

$$[fg](x) = f(x)\,g(x)\,, \tag{3.10}$$

is described by a convolution in frequency space. For the non-truncated system, the relation is

$$\widetilde{(fg)}_m = \frac{1}{P_x}\int_{P_x} f(x)\,g(x)\exp\left\{-2\pi i\frac{m}{P_x}x\right\}\mathrm{d}x = \sum_{m'}\widetilde{f}_{m-m'}\,\widetilde{g}_{m'}\,. \tag{3.11}$$

On the right side of Eq. (3.11), the quantity $\widetilde{f}_{m-m'}$ can be understood as a matrix with entries

$$[\![f]\!]_{m,m'} = \widetilde{f}_{m-m'}\,, \tag{3.12}$$

which is by definition a Toeplitz matrix. The latter is a matrix, whose diagonal and all secondary diagonals are constant. The concept is illustrated in Fig. 3.4. Hence, the product of two periodic functions $f(x)$ and $g(x)$ in position space is equivalent to a convolution in Fourier space, which can be represented as the matrix multiplication

$$\widetilde{(fg)}_m = \sum_{m'}[\![f]\!]_{m,m'}\,\widetilde{g}_{m'}\,, \tag{3.13}$$

where $[\![f]\!]$ is a Toeplitz matrix as defined in Eq. (3.12).

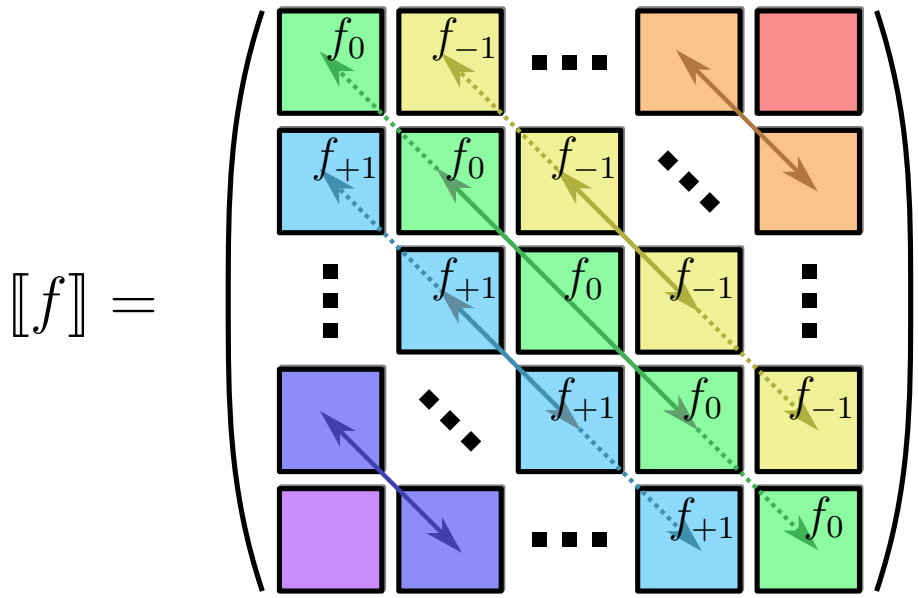


Figure 3.4: A Toeplitz matrix is a matrix, whose diagonal and secondary diagonals are constant.

3.2.2 Li's factorization rules

Direct rule

Consider Eq. (3.13), which is exact for non-truncated systems. Assume that the system is now truncated symmetrically around the central mode. In this case, both the set of coefficients $\widetilde{g}_{m'}$ and the Toeplitz matrix are truncated symmetrically. After the truncation, the vector has the dimension $M \times 1$, and matrix has the dimensions $M \times M$. The product is then approximated by the finite Matrix multiplication

$$(fg)_m^{(\text{direct rule})} = \sum_{m'=-M_0}^{+M_0} [\![f]\!]_{m,m'} \, \widetilde{g}_{m'} \, . \tag{3.14}$$

If the mode truncation is conducted according to Eq. (3.14), it is said that the 'direct rule' or 'Laurent's rule' is applied[9].

Inverse rule

In the early beginnings of the RCWA, all direct position space products were Fourier analyzed by the direct rule. However, it turned out that the resulting algorithm converges well for TE-polarized light, but is characterized by a rather slow convergence behavior for TM-polarization and especially metallic gratings. It was demonstrated that the reason for the slow convergence is a bad position space reconstruction of the products $\hat{\epsilon}(\mathbf{r})\,\boldsymbol{\mathcal{E}}(\mathbf{r})$ and $\hat{\mu}(\mathbf{r})\,\boldsymbol{\mathcal{H}}(\mathbf{r})$[8]. Across a material interface, the normal component of these products is continuous, i.e., both $\hat{\epsilon}(\mathbf{r})$ and $\boldsymbol{\mathcal{E}}(\mathbf{r})$, and $\hat{\mu}(\mathbf{r})$ and

$\boldsymbol{\mathcal{H}}(\mathbf{r})$ possess concurrent and pairwise complementary jump discontinuities. However, when reconstructing the position space product from the mode truncated functions using the direct rule, a strong Gibbs phenomenon may be present near the interface. This problem and Li's solution is discussed in the following.

As an example, assume the functions

$$f(x) = \begin{cases} \frac{1}{2} & -5 \leq x \leq 0 \\ \frac{3}{2} & 0 < x < 5 \end{cases} \tag{3.15a}$$

$$g(x) = \begin{cases} 2 & -5 \leq x \leq 0 \\ \frac{2}{3} & 0 < x < 5 \end{cases}, \tag{3.15b}$$

which are illustrated in Fig. 3.5. Assume that these function are periodically continued outside the interval $-5 \leq x < 5$. Both functions have discontinuities at $x = 0$ and at the period boundaries $x = \pm P_x/2$, whereas the product $f(x)\, g(x) = 1$ is continuous for all x. Fig. 3.6 (solid blue line) shows the reconstructed product

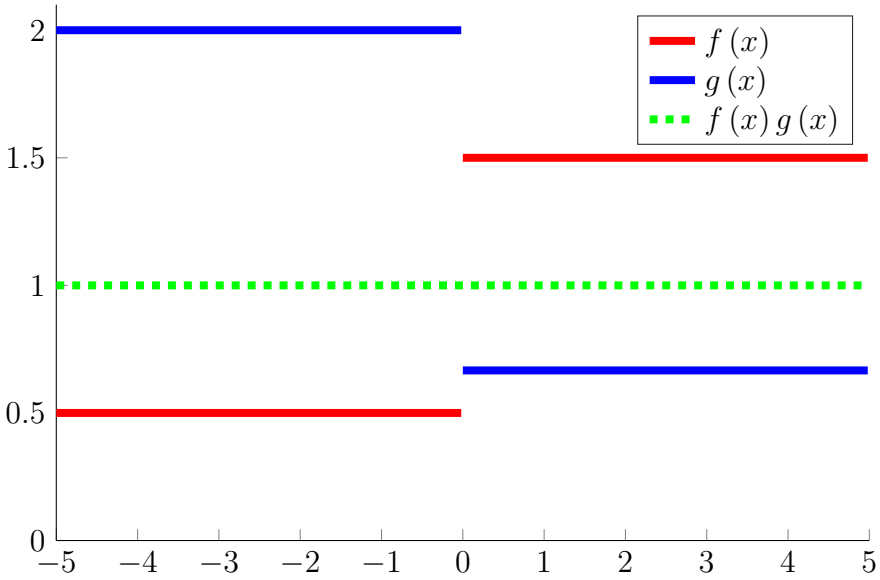


Figure 3.5: Two functions $f(x)$ (solid blue) and $g(x)$ (solid red) with a simultaneous discontinuity at $x = 0$. The product $f(x)\, g(x) = 1$ (dashed green) is continuous for all x.

function,

$$(fg)^{(\text{direct rule})}(x) = \sum_{m,m'=-M_0}^{+M_0} [\![f]\!]_{m,m'}\, \widetilde{g}_{m'} \exp\left\{2\pi i\, \frac{m}{P_x} x\right\}. \tag{3.16}$$

It possesses strong Gibbs oscillations near the discontinuites $x = 0$ and $x = \pm P_x/2$, whereas the analytic value of the product function is $(fg)^{(\text{analytic})}(x) = 1$. According

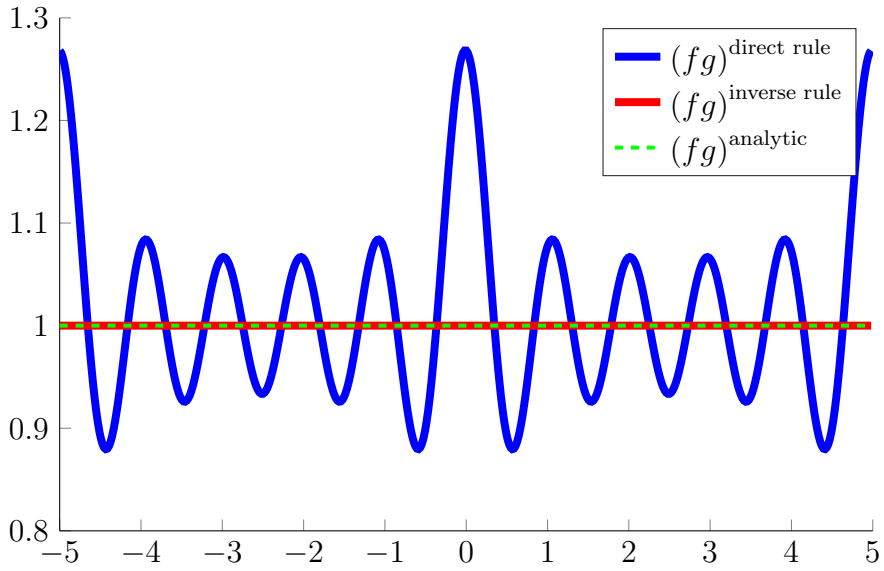


Figure 3.6: Product function $(fg)^{(\text{direct rule})}(x)$ (solid blue) reconstructed from the truncated mode coefficient vector using the direct rule. Strong Gibbs oscillations are present near the discontinuities $x = 0$ and $x = \pm P_x/2$. The reconstruction using the inverse rule, $(fg)^{(\text{inverse rule})}(x)$ (solid red), and the analytic reconstruction, $(fg)^{(\text{analytic})}(x)$ (dashed green), have the constant value 1.

to [8], the problem described above can be resolved by applying the so-called 'inverse rule'. In essence, the mode truncation is simply not done on Eq. (3.10), but on the transformed equation

$$g(x) = \frac{1}{f(x)}\,(fg)(x)\ . \tag{3.17}$$

The functions $\frac{1}{f(x)}$ and $(fg)(x)$ do not have simultaneous jump discontinuities, since $(fg)(x)$ is continuous in x. Therefore, the direct rule applies for this product. The mode truncated version of Eq. (3.17) then is

$$\widetilde{g}_m = \sum_{m'=-M_0}^{+M_0} \left[\!\left[\frac{1}{f}\right]\!\right]_{m,m'} \widetilde{(fg)}_{m'}^{(\text{inverse rule})}\ , \tag{3.18}$$

which can be transformed to

$$\widetilde{(fg)}_m^{(\text{inverse rule})} = \sum_{m'=-M_0}^{+M_0} \left[\!\left[\frac{1}{f}\right]\!\right]^{-1}_{m,m'} \widetilde{g}_{m'} \tag{3.19}$$

by matrix multiplication. This procedure is called the 'inverse rule'. Fig. 3.6 (solid

red line) shows the corresponding reconstructed product function,

$$(fg)^{(\text{inverse rule})}(x) = \sum_{m,m'=-M_0}^{+M_0} \left[\!\left[\frac{1}{f}\right]\!\right]^{-1}_{m,m'} \widetilde{g}_{m'} \exp\left\{2\pi i \, \frac{m}{P_x} x\right\} . \tag{3.20}$$

In contrast to the result of the direct rule, no strong Gibbs oscillations are present near the discontinuities $x = 0$ and $x = \pm P_x/2$. Moreover, the reconstructed function is nearly equal to the analytic product function $(fg)^{(\text{analytic})}(x) = 1$ at the sampling points.

Summary - Li's factorization rules

Li's factorization rules for Fourier factorization are a set of three rules, which summarize the above observations[9]. Be $f(x)$ and $g(x)$ two piecewise-continuous, piecewise smooth, bounded, periodic functions of x.

1. If $f(x)$ and $g(x)$ have no concurrent jump discontinuities (product type 1), the direct rule must be applied for Fourier factorization.

2. If $f(x)$ and $g(x)$ have pairwise complementary jump discontinuities (product type 2), the inverse rule must be applied for Fourier factorization.

3. If $f(x)$ and $g(x)$ have pairwise concurrent but not complementary jump discontinuities (product type 3), neither the direct rule nor the inverse rule can be applied for Fourier factorization. Hence these types of products must be avoided.

It should be noted that, for the non-truncated infinite system, the direct rule and the inverse rule, Eq. (3.14) and Eq. (3.19), are mathematically identical.

Extension to two dimensional functions

So far, Li's factorization rules were presented for product functions of one variable x only. In the following, this framework is extended to functions of two variables x and y as presented in [9], which is necessary for the Fourier analysis of Maxwell's equations in two coordinates. To this end, assume two x-y-periodic functions $f(x, y)$ and $g(x, y)$ with periods P_x and P_y. The Fourier transformation can now be calculated

for the variable x, y, or both. The notation is

$$\widetilde{f}_m(y) = \frac{1}{P_x} \int\limits_{P_x} f(x,y) \exp\left\{-2\pi i \frac{m}{P_x} x\right\} \mathrm{d}x \tag{3.21a}$$

$$\widetilde{f}_n(x) = \frac{1}{P_y} \int\limits_{P_y} f(x,y) \exp\left\{-2\pi i \frac{n}{P_y} y\right\} \mathrm{d}y \tag{3.21b}$$

$$\widetilde{f}_{mn}(x) = \frac{1}{P_x P_y} \int\limits_{P_x P_y} f(x,y) \exp\left\{-2\pi i \left(\frac{m}{P_x} x + \frac{n}{P_y} y\right)\right\} \mathrm{d}x \, \mathrm{d}y \,. \tag{3.21c}$$

In order to describe the position space product of f and g as a convolution in Fourier space, one has to differentiate between three different cases.

1. The direct rule applies to the Fourier transformation in both x and y, i.e.,

$$\widetilde{(fg)}_{mn}^{\binom{\text{direct rule in } x}{\text{direct rule in } y}} = \sum_{m'=-M_0}^{+M_0} \sum_{n'=-N_0}^{+N_0} [\![f]\!]_{mn,m'n'} \, \widetilde{g}_{m'n'} \,, \tag{3.22a}$$

   where

$$[\![f]\!]_{mn,m'n'} = \widetilde{f}_{m-m',n-n'} \,. \tag{3.22b}$$

2. The inverse rule applies to the Fourier transformation in x, and the direct rule applies to the Fourier transformation in and y, i.e.,

$$\widetilde{(fg)}_{mn}^{\binom{\text{inverse rule in } x}{\text{direct rule in } y}} = \sum_{m'=-M_0}^{+M_0} \sum_{n'=-N_0}^{+N_0} \lfloor\!\lceil f \rceil\!\rfloor_{mn,m'n'} \, \widetilde{g}_{m'n'} \,, \tag{3.23a}$$

   where

$$\lfloor\!\lceil f \rceil\!\rfloor_{mn,m'n'} = \frac{1}{P_y} \int\limits_{P_y} \left[\frac{1}{f}\right]^{-1}_{m,m'} (y) \exp\left\{-2\pi i \frac{n-n'}{P_y} y\right\} \mathrm{d}y \tag{3.23b}$$

$$\left[\frac{1}{f}\right]_{m,m'} (y) = \frac{1}{P_x} \int\limits_{P_x} \frac{1}{f(x,y)} \exp\left\{-2\pi i \frac{m-m'}{P_x} x\right\} \mathrm{d}x \,. \tag{3.23c}$$

3. The direct rule applies to the Fourier transformation in x, and the inverse rule applies to the Fourier transformation in y, i.e.,

$$\widetilde{(fg)}_{mn}^{\binom{\text{direct rule in } x}{\text{inverse rule in } y}} = \sum_{m'=-M_0}^{+M_0} \sum_{n'=-N_0}^{+N_0} \lceil\!\lfloor f \rfloor\!\rceil_{mn,m'n'} \, \widetilde{g}_{m'n'} \,, \tag{3.24a}$$

where

$$\lceil\!\lfloor f \rfloor\!\rceil_{mn,m'n'} = \frac{1}{P_x} \int\limits_{P_x} \left\lfloor \frac{1}{f} \right\rfloor_{n,n'}^{-1} (x) \exp\left\{ -2\pi i \frac{m-m'}{P_x} x \right\} \mathrm{d}x \tag{3.24b}$$

$$\left\lfloor \frac{1}{f} \right\rfloor_{n,n'} (x) = \frac{1}{P_y} \int\limits_{P_y} \frac{1}{f\,(x,y)} \exp\left\{ -2\pi i \frac{n-n'}{P_y} y \right\} \mathrm{d}y \,. \tag{3.24c}$$

3.2.3 Reduction of dimensions - mapping 4d $\rightarrow$ 2d

Consider Eqs. (3.22-3.24), which contain the two-dimensional quantitiy $\widetilde{g}_{mn}$ and three different forms of four-dimensional Toeplitz matrices, $[\![f]\!]_{mn,m'n'}$, $\lfloor\lceil f \rceil\rfloor_{mn,m'n'}$ and $\lceil\!\lfloor f \rfloor\!\rceil_{mn,m'n'}$. In order to work with these types of quantities on a computer, it is necessary to reduce the dimensionality of the system by a factor of two. The latter is achieved by combining two and two dimensions in one, i.e. one implements the mapping

$$\underbrace{\widetilde{g}_{mn}}_{\text{2d}} \rightarrow \underbrace{\widetilde{g}_{mN+n}}_{\text{1d}} \qquad \text{and} \qquad \underbrace{[\![f]\!]_{mn,m'n'}}_{\text{4d}} \rightarrow \underbrace{[\![f]\!]_{mN+n,m'N+n'}}_{\text{2d}} \,. \tag{3.25}$$

While the mapping of $\widetilde{g}_{mn}$ from 2d $\rightarrow$ 1d is fairly easy to imagine, an intuitive picture for the mapping of the Toeplitz matrix $[\![f]\!]_{mn,m'n'}$ from 4d $\rightarrow$ 2d is given in Fig. 3.7. In the literature [63], the form of this matrix is often described as 'block-Toeplitz Toeplitz-block' (BTTB), because it represents a Toeplitz block matrix, whose blocks again consist out of Toeplitz matrices. The Toeplitz matrices $\lfloor\lceil f \rceil\rfloor_{mn,m'n'}$ and $\lceil\!\lfloor f \rfloor\!\rceil_{mn,m'n'}$ have a very similar structure, but with inverted blocks.

3.3 Fourier analysis of Maxwell's equations

In this section, Maxwell's equations, as shown in Eq. (2.57), are Fourier analyzed in the variables x and y following Li's factorization rules. In the first part, this transformation is done for linear isotropic materials, which is the most frequent case. In the second part, the derivation is conducted for the more general case of linear anisotropic media.

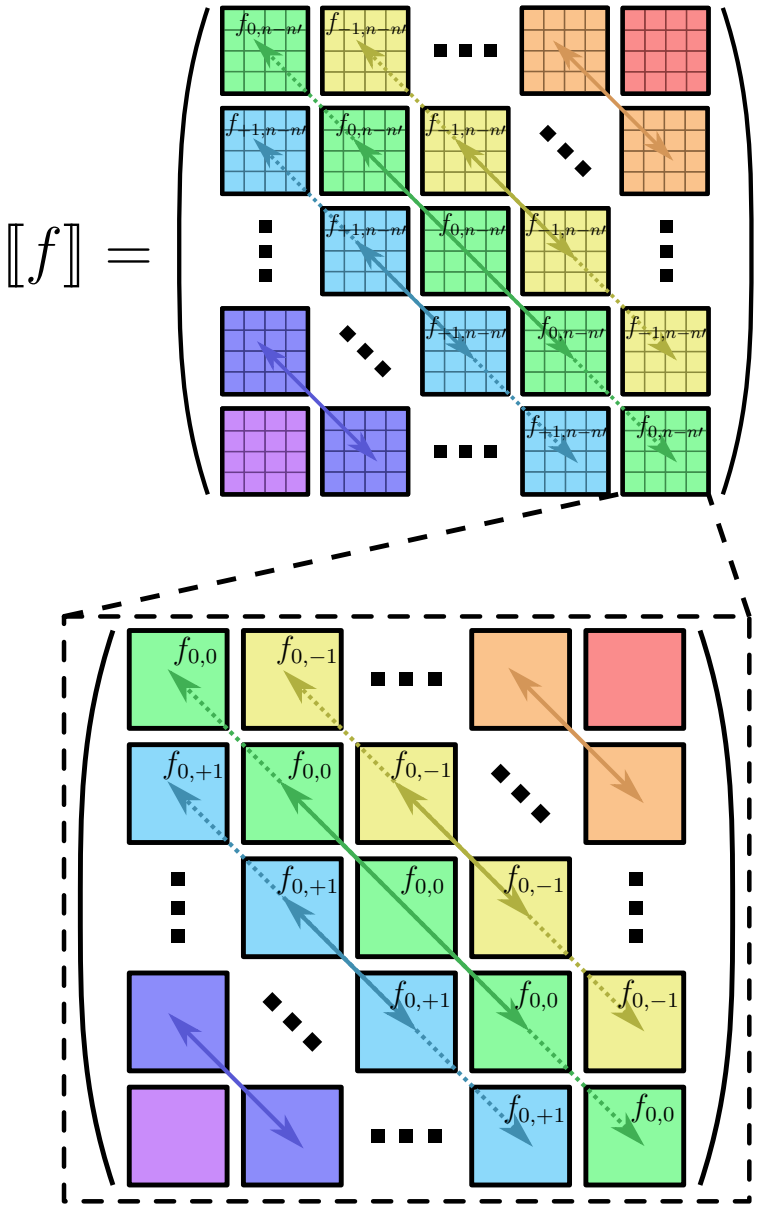


Figure 3.7: BTTB matrix: a Toeplitz block matrix, whose blocks again consist out of Toeplitz matrices.

3.3.1 Linear isotropic media

The following approach was initially presented by Li in 1997[9]. Maxwell's equations are Fourier factorized in x and y according to Li's factorization rules as presented in Sect. 3.2.2. In the l^{th} layer of the grating, the first two Maxwell equations for isotropic materials and normalized fields read

$$\partial_y \mathcal{E}_z(\mathbf{r}) - \partial_z \mathcal{E}_y(\mathbf{r}) = +ik_0 \mu^{(l)}(x,y) \mathcal{H}_x(\mathbf{r}) \tag{3.26a}$$
$$\partial_z \mathcal{E}_x(\mathbf{r}) - \partial_x \mathcal{E}_z(\mathbf{r}) = +ik_0 \mu^{(l)}(x,y) \mathcal{H}_y(\mathbf{r}) \tag{3.26b}$$
$$\partial_x \mathcal{E}_y(\mathbf{r}) - \partial_y \mathcal{E}_x(\mathbf{r}) = +ik_0 \mu^{(l)}(x,y) \mathcal{H}_z(\mathbf{r}) \tag{3.26c}$$

$$\partial_y \mathcal{H}_z(\mathbf{r}) - \partial_z \mathcal{H}_y(\mathbf{r}) = -ik_0 \epsilon^{(l)}(x,y) \mathcal{E}_x(\mathbf{r}) \tag{3.26d}$$
$$\partial_z \mathcal{H}_x(\mathbf{r}) - \partial_x \mathcal{H}_z(\mathbf{r}) = -ik_0 \epsilon^{(l)}(x,y) \mathcal{E}_y(\mathbf{r}) \tag{3.26e}$$
$$\partial_x \mathcal{H}_y(\mathbf{r}) - \partial_y \mathcal{H}_x(\mathbf{r}) = -ik_0 \epsilon^{(l)}(x,y) \mathcal{E}_z(\mathbf{r}) \tag{3.26f}$$

In the following, the upper index $^{(l)}$, which indicates the layer number, is dropped to ensure readability.

Fourier transformation of Eqs. (3.26c, 3.26f)

First, Eqs. (3.26c, 3.26f) are Fourier transformed with respect to the variables x and y. The quantities $\mathcal{H}_z(\mathbf{r})$ and $\mathcal{E}_z(\mathbf{r})$ on the right side are both continuous in x and y. Therefore, the direct rule applies,

$$ik_{x;mn}\widetilde{\mathcal{E}}_{y;mn}(z) - ik_{y;mn}\widetilde{\mathcal{E}}_{x;mn}(z) = +ik_0 \sum_{m'n'} [\![\mu]\!]_{mn,m'n'} \widetilde{\mathcal{H}}_{z;m'n'}(z) \tag{3.27a}$$
$$ik_{x;mn}\widetilde{\mathcal{H}}_{y;mn}(z) - ik_{y;mn}\widetilde{\mathcal{H}}_{x;mn}(z) = -ik_0 \sum_{m'n'} [\![\epsilon]\!]_{mn,m'n'} \widetilde{\mathcal{E}}_{z;m'n'}(z)\,. \tag{3.27b}$$

Fourier transformation of Eqs. (3.26a, 3.26d)

Eqs. (3.26a, 3.26d) are Fourier transformed only with respect to the variable x. Consider first Eq. (3.26a). The right side is equal to the magnetic flux density $\mathcal{B}_x(\mathbf{r})$, whose normal component is continuous across material interfaces. Therefore, the product $\mu(x,y)\,\mathcal{H}_x(\mathbf{r})$ is a continuous function of x, whereas the individual factors $\mu(x,y)$ and $\mathcal{H}_x(\mathbf{r})$ possess concurrent and pairwise complementary jump discontinuities. With the same line of arguments it is concluded that, in Eq. (3.26d), the product $\epsilon(x,y)\,\mathcal{E}_x(\mathbf{r})$ is a continuous function of x, and that the individual factors $\epsilon(x,y)$ and $\mathcal{E}_x(\mathbf{r})$ possess concurrent and complementary jump discontinuities.

Therefore, in both cases the inverse rule applies for the Fourier transformation in y,

$$\partial_y \widetilde{\mathcal{E}}_{z;m}(y,z) - \partial_z \widetilde{\mathcal{E}}_{y;m}(y,z) = +ik_0 \sum_{m'} \lceil 1/\mu \rceil^{-1}_{m,m'} \widetilde{\mathcal{H}}_{x;m'}(y,z) \tag{3.28a}$$

$$\partial_y \widetilde{\mathcal{H}}_{z;m}(y,z) - \partial_z \widetilde{\mathcal{H}}_{y;m}(y,z) = -ik_0 \sum_{m'} \lceil 1/\epsilon \rceil^{-1}_{m,m'} \widetilde{\mathcal{E}}_{x;m'}(y,z) \,. \tag{3.28b}$$

In Eq. (3.28a), the right side is a continuous function of y. This directly follows from the continuity of $\widetilde{\mathcal{H}}_{x;m'}(y,z)$ in y. The same is true for the right side of Eq. (3.28b). Therefore, in both cases the direct rule applies,

$$ik_{y;mn} \widetilde{\mathcal{E}}_{z;mn}(z) - \partial_z \widetilde{\mathcal{E}}_{y;mn}(z) = +ik_0 \sum_{m'n'} \lfloor \lceil \mu \rfloor \rfloor_{mn,m'n'} \widetilde{\mathcal{H}}_{x;m'n'}(z) \tag{3.29a}$$

$$ik_{y;mn} \widetilde{\mathcal{H}}_{z;mn}(z) - \partial_z \widetilde{\mathcal{H}}_{y;mn}(z) = -ik_0 \sum_{m'n'} \lfloor \lceil \epsilon \rceil \rfloor_{mn,m'n'} \widetilde{\mathcal{E}}_{x;m'n'}(z) \,. \tag{3.29b}$$

Fourier transformation of Eqs. (3.26b, 3.26e)

Eqs. (3.26b, 3.26e) are Fourier transformed only with respect to the variable y. Following the same line of arguments as above, it is concluded that the products $\mu(x,y)\,\mathcal{H}_y(\mathbf{r})$ and $\epsilon(x,y)\,\mathcal{E}_y(\mathbf{r})$ are both continuous functions of y, whereas the constituent functions possess concurrent and complementary discontinuities. Therefore, the inverse rule applies,

$$\partial_z \widetilde{\mathcal{E}}_{x;n}(x,z) - \partial_x \widetilde{\mathcal{E}}_{z;n}(x,z) = +ik_0 \sum_{n'} \lfloor 1/\mu \rfloor^{-1}_{n,n'} \widetilde{\mathcal{H}}_{y;n'}(x,z) \tag{3.30a}$$

$$\partial_z \widetilde{\mathcal{H}}_{x;n}(x,z) - \partial_x \widetilde{\mathcal{H}}_{z;n}(x,z) = -ik_0 \sum_{n'} \lfloor 1/\epsilon \rfloor^{-1}_{n,n'} \widetilde{\mathcal{E}}_{y;n'}(x,z) \,. \tag{3.30b}$$

In Eq. (3.30a), the right side is a continuous function of x. This directly follows from the continuity of $\widetilde{\mathcal{H}}_{y;n'}(x,z)$ in x. The same argument is true for the right side of Eq. (3.30b). Therefore, in both cases the direct rule applies for the Fourier transformation in x,

$$\partial_z \widetilde{\mathcal{E}}_{x;mn}(z) - ik_{x;mn} \widetilde{\mathcal{E}}_{z;mn}(z) = +ik_0 \sum_{m'n'} \lceil \lfloor \mu \rfloor \rceil_{mn,m'n'} \widetilde{\mathcal{H}}_{y;m'n'}(z) \tag{3.31a}$$

$$\partial_z \widetilde{\mathcal{H}}_{x;mn}(z) - ik_{x;mn} \widetilde{\mathcal{H}}_{z;mn}(z) = -ik_0 \sum_{m'n'} \lceil \lfloor \epsilon \rfloor \rceil_{mn,m'n'} \widetilde{\mathcal{E}}_{y;m'n'}(z) \,. \tag{3.31b}$$

Matrix equation

By solving Eqs. (3.29, 3.31) for the ∂_z derivatives and eliminating $\widetilde{\mathcal{H}}_{z;mn}(z)$ and $\widetilde{\mathcal{E}}_{z;mn}(z)$ using Eqs. (3.27), one obtains the matrix equation

$$\frac{\partial}{\partial z}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix} = ik_0\begin{bmatrix}\mathbb{0} & \hat{M}_{12}\\ \hat{M}_{21} & \mathbb{0}\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}, \tag{3.32a}$$

where

$$\hat{M}_{12} = \begin{bmatrix}\hat{K}_x [\![\epsilon]\!]^{-1} \hat{K}_y & -\hat{K}_x [\![\epsilon]\!]^{-1} \hat{K}_x + \lceil\lfloor\mu\rfloor\rceil\\ \hat{K}_y [\![\epsilon]\!]^{-1} \hat{K}_y - \lfloor\lceil\mu\rfloor\rfloor & -\hat{K}_y [\![\epsilon]\!]^{-1} \hat{K}_x\end{bmatrix} \tag{3.32b}$$

$$\hat{M}_{21} = \begin{bmatrix}-\hat{K}_x [\![\mu]\!]^{-1} \hat{K}_y & \hat{K}_x [\![\mu]\!]^{-1} \hat{K}_x - \lceil\lfloor\epsilon\rfloor\rceil\\ -\hat{K}_y [\![\mu]\!]^{-1} \hat{K}_y + \lfloor\lceil\epsilon\rfloor\rfloor & \hat{K}_y [\![\mu]\!]^{-1} \hat{K}_x\end{bmatrix} \tag{3.32c}$$

$$\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z) = \begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_x(z)\\ \widetilde{\boldsymbol{\mathcal{E}}}_y(z)\end{bmatrix} \tag{3.32d}$$

$$\widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z) = \begin{bmatrix}\widetilde{\boldsymbol{\mathcal{H}}}_x(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_y(z)\end{bmatrix} \tag{3.32e}$$

$$\hat{K}_{\rho;mn,m'n'} = \frac{k_{\rho;mn}}{k_0}\,\delta_{m,m'}\,\delta_{n,n'}\,. \tag{3.32f}$$

Here, note that the Fourier indices are omitted. Furthermore, one has to keep in mind that the layer index was dropped in the beginning of the derivation and that, therefore, the coupling matrix in Eq. (3.32a) is layer dependent, i.e. $\hat{M}_{12} \equiv \hat{M}_{12}^{(l)}$ and $\hat{M}_{21} \equiv \hat{M}_{21}^{(l)}$.

3.3.2 Linear anisotropic media

The subsequent part follows the procedure from [15]. However, it must be noted that the result in [15] does not simplify to the result of Sect. 3.2.2 in the linear isotropic case. The subsequent approach is, therefore, modified such, that both approaches are consistent. In the following, the arguments of the functions are dropped in order to facilitate the readability[3]. Start from Maxwell's equations, Eq. (2.57),

$$\partial_y \mathcal{E}_z - \partial_z \mathcal{E}_y = +ik_0\left[\mu_{xx}\mathcal{H}_x + \mu_{xy}\mathcal{H}_y + \mu_{xz}\mathcal{H}_z\right] \tag{3.33a}$$

$$\partial_z \mathcal{E}_x - \partial_x \mathcal{E}_z = +ik_0\left[\mu_{yx}\mathcal{H}_x + \mu_{yy}\mathcal{H}_y + \mu_{yz}\mathcal{H}_z\right] \tag{3.33b}$$

$$\partial_x \mathcal{E}_y - \partial_y \mathcal{E}_x = +ik_0\left[\mu_{zx}\mathcal{H}_x + \mu_{zy}\mathcal{H}_y + \mu_{zz}\mathcal{H}_z\right] \tag{3.33c}$$

[3] The arguments can be reconstructed from the normal/calligraphic style as explained earlier.

$$\partial_y \mathcal{H}_z - \partial_z \mathcal{H}_y = -ik_0 \left[\epsilon_{xx}\mathcal{E}_x + \epsilon_{xy}\mathcal{E}_y + \epsilon_{xz}\mathcal{E}_z\right] \tag{3.33d}$$
$$\partial_z \mathcal{H}_x - \partial_x \mathcal{H}_z = -ik_0 \left[\epsilon_{yx}\mathcal{E}_x + \epsilon_{yy}\mathcal{E}_y + \epsilon_{yz}\mathcal{E}_z\right] \tag{3.33e}$$
$$\partial_x \mathcal{H}_y - \partial_y \mathcal{H}_x = -ik_0 \left[\epsilon_{zx}\mathcal{E}_x + \epsilon_{zy}\mathcal{E}_y + \epsilon_{zz}\mathcal{E}_z\right] . \tag{3.33f}$$

Partial Fourier transforms of Eqs. (3.33a, 3.33d) in x or in y

First, Fourier transform Eqs. (3.33a, 3.33d) in x. To this end, rewrite,

$$\partial_y \mathcal{E}_z - \partial_z \mathcal{E}_y = +ik_0 \mu_{xx} \left[\mathcal{H}_x + \mu_{xx}^{-1}\mu_{xy}\ \mathcal{H}_y + \mu_{xx}^{-1}\mu_{xz}\ \mathcal{H}_z\right] \tag{3.34a}$$
$$\partial_y \mathcal{H}_z - \partial_z \mathcal{H}_y = -ik_0 \epsilon_{xx} \left[\mathcal{E}_x + \epsilon_{xx}^{-1}\epsilon_{xy}\ \mathcal{E}_y + \epsilon_{xx}^{-1}\epsilon_{xz}\ \mathcal{E}_z\right] . \tag{3.34b}$$

Consider Eq. (3.34a). The left hand side (LHS) is a continuous function of x. The reason is that LHS $= \mathcal{B}_x(\mathbf{r})$, whose normal component is continuous across material interfaces. Next, consider the bracket on the right hand side (RHS). The functions $\mathcal{H}_y$,$\mathcal{H}_z$ are continuous in x, i.e. the direct rule applies for their prefactors. On the other hand, μ_{xx} possesses jump discontinuities in x. From the continuity of the LHS, it is concluded that, on the RHS, μ_{xx} and the bracket must possess concurrent and complementary jump discontinuities, i.e. the inverse rule applies. The same line of arguments can be followed to transform Eq. (3.34b) in x,

$$\partial_y \widetilde{\mathcal{E}}_{z;m} - \partial_z \widetilde{\mathcal{E}}_{y;m} = +ik_0 \lceil \mu_{xx}^{-1} \rceil^{-1}_{m,m'} \Big[\widetilde{\mathcal{H}}_{x;m'} + \lceil \mu_{xx}^{-1}\mu_{xy} \rceil_{m',m''} \widetilde{\mathcal{H}}_{y;m''} + \lceil \mu_{xx}^{-1}\mu_{xz} \rceil_{m',m''} \widetilde{\mathcal{H}}_{z;m''}\Big] \tag{3.35a}$$

$$\partial_y \widetilde{\mathcal{H}}_{z;m} - \partial_z \widetilde{\mathcal{H}}_{y;m} = -ik_0 \lceil \epsilon_{xx}^{-1} \rceil^{-1}_{m,m'} \Big[\widetilde{\mathcal{E}}_{x;m'} + \lceil \epsilon_{xx}^{-1}\epsilon_{xy} \rceil_{m',m''} \widetilde{\mathcal{E}}_{y;m''} + \lceil \epsilon_{xx}^{-1}\epsilon_{xz} \rceil_{m',m''} \widetilde{\mathcal{E}}_{z;m''}\Big] , \tag{3.35b}$$

where one sums over repetitive indices (sum convention).

Next, one Fourier transforms Eqs. (3.33a, 3.33d) in y. To this end, one solves Eqs. (3.33b, 3.33e) for $\mathcal{E}_y, \mathcal{H}_y$ and inserts the result into the RHS of Eqs. (3.33a, 3.33d),

$$\partial_y \mathcal{E}_z - \partial_z \mathcal{E}_y = +ik_0 \Big[-i/k_0\, \mu_{xy}\mu_{yy}^{-1} (\partial_z \mathcal{E}_x - \partial_x \mathcal{E}_z) + \left\{\mu_{xx} - \mu_{xy}\mu_{yy}^{-1}\mu_{yx}\right\} \mathcal{H}_x + \left\{\mu_{xz} - \mu_{xy}\mu_{yy}^{-1}\mu_{yz}\right\} \mathcal{H}_z\Big] \tag{3.36a}$$

$$\partial_y \mathcal{H}_z - \partial_z \mathcal{H}_y = -ik_0 \Big[+i/k_0\, \epsilon_{xy}\epsilon_{yy}^{-1} (\partial_z \mathcal{H}_x - \partial_x \mathcal{H}_z) + \left\{\epsilon_{xx} - \epsilon_{xy}\epsilon_{yy}^{-1}\epsilon_{yx}\right\} \mathcal{E}_x + \left\{\epsilon_{xz} - \epsilon_{xy}\epsilon_{yy}^{-1}\epsilon_{yz}\right\} \mathcal{E}_z\Big] . \tag{3.36b}$$

Consider the bracket on the RHS of Eq. (3.36a). The quantities $\mathcal{B}_y = \partial_z \mathcal{E}_x - \partial_x \mathcal{E}_z$, $\mathcal{H}_x$ and $\mathcal{H}_z$ are continuous in y. Therefore, the direct rule applies to the prefac-

tors. The same argument applies to the transformation of Eq. (3.36b). By Fourier transformation in y one obtains

$$\begin{aligned} ik_y\mathcal{E}_{z;n} - \partial_z\mathcal{E}_{y;n} &= +ik_0\Big[-i/k_0\,\lfloor\mu_{xy}\mu_{yy}^{-1}\rfloor_{n,n'}\,(\partial_z\mathcal{E}_{x;n'} - \partial_x\mathcal{E}_{z;n'}) \\ &\quad +\lfloor\mu_{xx} - \mu_{xy}\mu_{yy}^{-1}\mu_{yx}\rfloor_{n,n'}\mathcal{H}_{x;n'} \\ &\quad +\lfloor\mu_{xz} - \mu_{xy}\mu_{yy}^{-1}\mu_{yz}\rfloor_{n,n'}\mathcal{H}_{z;n'}\Big] \end{aligned} \tag{3.37a}$$

$$\begin{aligned} ik_y\mathcal{H}_{z;n} - \partial_z\mathcal{H}_{y;n} = -ik_0\Big[+i/k_0\,\lfloor\epsilon_{xy}\epsilon_{yy}^{-1}\rfloor_{n,n'}\,(\partial_z\mathcal{H}_{x;n'} - \partial_x\mathcal{H}_{z;n'}) \\ +\lfloor\epsilon_{xx} - \epsilon_{xy}\epsilon_{yy}^{-1}\epsilon_{yx}\rfloor_{n,n'}\mathcal{E}_{x;n'} \\ +\lfloor\epsilon_{xz} - \epsilon_{xy}\epsilon_{yy}^{-1}\epsilon_{yz}\rfloor_{n,n'}\mathcal{E}_{z;n'}\Big] \end{aligned} \tag{3.37b}$$

Fourier transform Eqs. (3.33b, 3.33e) in x or in y

In the same manner as before, one partially Fourier transforms Eqs. (3.33b, 3.33e) in x and y. The result is

$$\begin{aligned} \partial_z\mathcal{E}_{x;m} - ik_x\mathcal{E}_{z;m} &= +ik_0\Big[-i/k_0\,\lceil\mu_{yx}\hat{\mu}_{xx}^{-1}\rceil_{m,m'}\,(\partial_y\mathcal{E}_{z;m'} - \partial_z\mathcal{E}_{y;m'}) \\ &\quad +\lceil\mu_{yy} - \mu_{yx}\hat{\mu}_{xx}^{-1}\mu_{xy}\rceil_{m,m'}\mathcal{H}_{y;m'} \\ &\quad +\lceil\mu_{yz} - \mu_{yx}\hat{\mu}_{xx}^{-1}\mu_{xz}\rceil_{m,m'}\mathcal{H}_{z;m'}\Big] \end{aligned} \tag{3.38a}$$

$$\begin{aligned} \partial_z\mathcal{H}_{x;m} - ik_x\mathcal{H}_{z;m} = -ik_0\Big[+i/k_0\,\lceil\epsilon_{yx}\epsilon_{xx}^{-1}\,(\partial_y\mathcal{H}_{z;m'} - \partial_z\mathcal{H}_{y;m'}) \\ +\lceil\epsilon_{yy} - \epsilon_{yx}\epsilon_{xx}^{-1}\epsilon_{xy}\rceil_{m,m'}\mathcal{E}_{y;m'} \\ +\lceil\epsilon_{yz} - \epsilon_{yx}\epsilon_{xx}^{-1}\epsilon_{xz}\rceil_{m,m'}\mathcal{E}_{z;m'}\Big] \end{aligned} \tag{3.38b}$$

and

$$\begin{aligned} \partial_z\mathcal{E}_{x;n} - \partial_x\mathcal{E}_{z;n} &= +ik_0\lfloor\mu_{yy}^{-1}\rfloor_{n,n'}^{-1}\,\Big[\lfloor\mu_{yy}^{-1}\mu_{yx}\rfloor_{n',n''}\mathcal{H}_{x;n''} + \mathcal{H}_{y;n'} \\ &\quad +\lfloor\mu_{yy}^{-1}\mu_{yz}\rfloor_{n',n''}\mathcal{H}_{z;n''}\Big] \end{aligned} \tag{3.39a}$$

$$\begin{aligned} \partial_z\mathcal{H}_{x;n} - \partial_x\mathcal{H}_{z;n} = -ik_0\lfloor\epsilon_{yy}^{-1}\rfloor_{n,n'}^{-1}\,\Big[\lfloor\epsilon_{yy}^{-1}\epsilon_{yx}\rfloor_{n',n''}\mathcal{E}_{x;n''} + \mathcal{E}_{y;n'} \\ +\lfloor\epsilon_{yy}^{-1}\epsilon_{yz}\rfloor_{n',n''}\mathcal{E}_{z;n''}\Big]\,. \end{aligned} \tag{3.39b}$$

Fourier transform Eqs. (3.33c, 3.33f) in x

Solve Eqs. (3.34a, 3.34b) for $\mathcal{H}_x, \mathcal{E}_x$ and insert into the RHS of Eqs. (3.33c, 3.33f),

$$\begin{aligned} \partial_x\mathcal{E}_y - \partial_y\mathcal{E}_x &= +ik_0\Big[-i/k_0\,\mu_{zx}\mu_{xx}^{-1}\,(\partial_y\mathcal{E}_z - \partial_z\mathcal{E}_y) \\ &\quad +\left(\mu_{zy} - \mu_{zx}\mu_{xx}^{-1}\mu_{xy}\right)\mathcal{H}_y \\ &\quad +\left(\mu_{zz} - \mu_{zx}\mu_{xx}^{-1}\mu_{xz}\right)\mathcal{H}_z\Big] \end{aligned} \tag{3.40a}$$

$$\begin{aligned}\partial_x \mathcal{H}_y - \partial_y \mathcal{H}_x = -ik_0 \big[&+i/k_0\, \epsilon_{zx}\epsilon_{xx}^{-1} \left(\partial_y \mathcal{H}_z - \partial_z \mathcal{H}_y\right) \\ &+ \left(\epsilon_{zy} - \epsilon_{zx}\epsilon_{xx}^{-1}\epsilon_{xy}\right) \mathcal{E}_y \\ &+ \left(\epsilon_{zz} - \epsilon_{zx}\epsilon_{xx}^{-1}\epsilon_{xz}\right) \mathcal{E}_z \big] \end{aligned} \tag{3.40b}$$

Consider the bracket on the RHS of Eq. (3.40a). The quantities $\mathcal{B}_x = \partial_y \mathcal{E}_z - \partial_z \mathcal{E}_y$, $\mathcal{H}_y$ and $\mathcal{H}_z$ are continuous in x. Therefore, the direct rule applies to the prefactors. The same argument applies to the transformation of Eq. (3.40b). One obtains

$$\begin{aligned}ik_x \mathcal{E}_{y;m} - \partial_y \mathcal{E}_{x;m} \quad = +ik_0 \big[&-i/k_0 \left\lceil \mu_{zx}\mu_{xx}^{-1} \right\rceil_{m,m'} \left(\partial_y \mathcal{E}_{z;m'} - \partial_z \mathcal{E}_{y;m'}\right) \\ &+ \left\lceil \left(\mu_{zy} - \mu_{zx}\mu_{xx}^{-1}\mu_{xy}\right) \right\rceil_{m,m'} \mathcal{H}_{y;m'} \\ &+ \left\lceil \left(\mu_{zz} - \mu_{zx}\mu_{xx}^{-1}\mu_{xz}\right) \right\rceil_{m,m'} \mathcal{H}_{z;m'} \big] \end{aligned} \tag{3.41a}$$

$$\begin{aligned}ik_x \mathcal{H}_{y;m} - \partial_y \mathcal{H}_{x;m} = -ik_0 \big[&+i/k_0 \left\lceil \epsilon_{zx}\epsilon_{xx}^{-1} \right\rceil_{m,m'} \left(\partial_y \mathcal{H}_{z;m'} - \partial_z \mathcal{H}_{y;m'}\right) \\ &+ \left\lceil \left(\epsilon_{zy} - \epsilon_{zx}\epsilon_{xx}^{-1}\epsilon_{xy}\right) \right\rceil_{m,m'} \mathcal{E}_{y;m'} \\ &+ \left\lceil \left(\epsilon_{zz} - \epsilon_{zx}\epsilon_{xx}^{-1}\epsilon_{xz}\right) \right\rceil_{m,m'} \mathcal{E}_{z;m'} \big] \end{aligned} \tag{3.41b}$$

Partially transformed Maxwell's equations

Now insert

1. Eqs. (3.35) in Eqs. (3.41),
2. Eqs. (3.39) in Eqs. (3.37),
3. Eqs. (3.35) in Eqs. (3.38).

One obtains two sets of partially transformed Maxwell's equations - one that is Fourier transformed only in x, and another one that is Fourier transformed only in y,

$$\partial_y\, \mathcal{E}_{z;m} - \partial_z\, \mathcal{E}_{y;m} = +ik_0 \big[\hat{M}^{(x)}_{xx;m,m'} \mathcal{H}_{x;m'} + \hat{M}^{(x)}_{xy;m,m'} \mathcal{H}_{y;m'} + \hat{M}^{(x)}_{xz;m,m'} \mathcal{H}_{z;m'} \big] \tag{3.42a}$$

$$\partial_z\, \mathcal{E}_{x;m} - ik_x \mathcal{E}_{z;m} = +ik_0 \big[\hat{M}^{(x)}_{yx;m,m'} \mathcal{H}_{x;m'} + \hat{M}^{(x)}_{yy;m,m'} \mathcal{H}_{y;m'} + \hat{M}^{(x)}_{yz;m,m'} \mathcal{H}_{z;m'} \big] \tag{3.42b}$$

$$ik_x \mathcal{E}_{y;m} - \partial_y\, \mathcal{E}_{x;m} = +ik_0 \big[\hat{M}^{(x)}_{zx;m,m'} \mathcal{H}_{x;m'} + \hat{M}^{(x)}_{zy;m,m'} \mathcal{H}_{y;m'} + \hat{M}^{(x)}_{zz;m,m'} \mathcal{H}_{z;m'} \big] \tag{3.42c}$$

$$\partial_y\, \mathcal{H}_{z;m} - \partial_z\, \mathcal{H}_{y;m} = -ik_0 \big[\hat{Q}^{(x)}_{xx;m,m'} \mathcal{E}_{x;m'} + \hat{Q}^{(x)}_{xy;m,m'} \mathcal{E}_{y;m'} + \hat{Q}^{(x)}_{xz;m,m'} \mathcal{E}_{z;m'} \big] \tag{3.42d}$$

$$\partial_z\, \mathcal{H}_{x;m} - ik_x \mathcal{H}_{z;m} = -ik_0 \big[\hat{Q}^{(x)}_{yx;m,m'} \mathcal{E}_{x;m'} + \hat{Q}^{(x)}_{yy;m,m'} \mathcal{E}_{y;m'} + \hat{Q}^{(x)}_{yz;m,m'} \mathcal{E}_{z;m'} \big] \tag{3.42e}$$

$$ik_x \mathcal{H}_{y;m} - \partial_y\, \mathcal{H}_{x;m} = -ik_0 \big[\hat{Q}^{(x)}_{zx;m,m'} \mathcal{E}_{x;m'} + \hat{Q}^{(x)}_{zy;m,m'} \mathcal{E}_{y;m'} + \hat{Q}^{(x)}_{zz;m,m'} \mathcal{E}_{z;m'} \big] , \tag{3.42f}$$

where

$$\hat{M}^{(x)} = \begin{bmatrix} \lceil \frac{1}{\mu_{xx}} \rceil^{-1} & \lceil \frac{1}{\mu_{xx}} \rceil^{-1} \lceil \frac{\mu_{xy}}{\mu_{xx}} \rceil & \lceil \frac{1}{\mu_{xx}} \rceil^{-1} \lceil \frac{\mu_{xz}}{\mu_{xx}} \rceil \\ \lceil \frac{\mu_{yx}}{\mu_{xx}} \rceil \lceil \frac{1}{\mu_{xx}} \rceil^{-1} & \lceil \frac{\mu_{yx}}{\mu_{xx}} \rceil \lceil \frac{1}{\mu_{xx}} \rceil^{-1} \lceil \frac{\mu_{xy}}{\mu_{xx}} \rceil + \lceil \mu_{yy} - \frac{\mu_{yx}\mu_{xy}}{\mu_{xx}} \rceil & \lceil \frac{\mu_{yx}}{\mu_{xx}} \rceil \lceil \frac{1}{\mu_{xx}} \rceil^{-1} \lceil \frac{\mu_{xz}}{\mu_{xx}} \rceil + \lceil \mu_{yz} - \frac{\mu_{yx}\mu_{xz}}{\mu_{xx}} \rceil \\ \lceil \frac{\mu_{zx}}{\mu_{xx}} \rceil \lceil \frac{1}{\mu_{xx}} \rceil^{-1} & \lceil \frac{\mu_{zx}}{\mu_{xx}} \rceil \lceil \frac{1}{\mu_{xx}} \rceil^{-1} \lceil \frac{\mu_{xy}}{\mu_{xx}} \rceil + \lceil \mu_{zy} - \frac{\mu_{zx}\mu_{xy}}{\mu_{xx}} \rceil & \lceil \frac{\mu_{zx}}{\mu_{xx}} \rceil \lceil \frac{1}{\mu_{xx}} \rceil^{-1} \lceil \frac{\mu_{xz}}{\mu_{xx}} \rceil + \lceil \mu_{zz} - \frac{\mu_{zx}\mu_{xz}}{\mu_{xx}} \rceil \end{bmatrix} \tag{3.42g}$$

$$\hat{Q}^{(x)} = \begin{bmatrix} \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} & \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} \lceil \frac{\epsilon_{xy}}{\epsilon_{xx}} \rceil & \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} \lceil \frac{\epsilon_{xz}}{\epsilon_{xx}} \rceil \\ \lceil \frac{\epsilon_{yx}}{\epsilon_{xx}} \rceil \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} & \lceil \frac{\epsilon_{yx}}{\epsilon_{xx}} \rceil \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} \lceil \frac{\epsilon_{xy}}{\epsilon_{xx}} \rceil + \lceil \epsilon_{yy} - \frac{\epsilon_{yx}\epsilon_{xy}}{\epsilon_{xx}} \rceil & \lceil \frac{\epsilon_{yx}}{\epsilon_{xx}} \rceil \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} \lceil \frac{\epsilon_{xz}}{\epsilon_{xx}} \rceil + \lceil \epsilon_{yz} - \frac{\epsilon_{yx}\epsilon_{xz}}{\epsilon_{xx}} \rceil \\ \lceil \frac{\epsilon_{zx}}{\epsilon_{xx}} \rceil \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} & \lceil \frac{\epsilon_{zx}}{\epsilon_{xx}} \rceil \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} \lceil \frac{\epsilon_{xy}}{\epsilon_{xx}} \rceil + \lceil \epsilon_{zy} - \frac{\epsilon_{zx}\epsilon_{xy}}{\epsilon_{xx}} \rceil & \lceil \frac{\epsilon_{zx}}{\epsilon_{xx}} \rceil \lceil \frac{1}{\epsilon_{xx}} \rceil^{-1} \lceil \frac{\epsilon_{xz}}{\epsilon_{xx}} \rceil + \lceil \epsilon_{zz} - \frac{\epsilon_{zx}\epsilon_{xz}}{\epsilon_{xx}} \rceil \end{bmatrix}, \tag{3.42h}$$

and

$$ik_y \mathcal{E}_{z;n} - \partial_z \mathcal{E}_{y;n} = +ik_0 \big[\hat{M}^{(y)}_{xx;n,n'} \mathcal{H}_{x;n'} + \hat{M}^{(y)}_{xy;n,n'} \mathcal{H}_{y;n'} + \hat{M}^{(y)}_{xz;n,n'} \mathcal{H}_{z;n'} \big] \tag{3.43a}$$

$$\partial_z \mathcal{E}_{x;n} - \partial_x \mathcal{E}_{z;n} = +ik_0 \big[\hat{M}^{(y)}_{yx;n,n'} \mathcal{H}_{x;n'} + \hat{M}^{(y)}_{yy;n,n'} \mathcal{H}_{y;n'} + \hat{M}^{(y)}_{yz;n,n'} \mathcal{H}_{z;n'} \big] \tag{3.43b}$$

$$\partial_x \mathcal{E}_{y;n} - ik_y \mathcal{E}_{x;n} = +ik_0 \big[\hat{M}^{(y)}_{zx;n,n'} \mathcal{H}_{x;n'} + \hat{M}^{(y)}_{zy;n,n'} \mathcal{H}_{y;n'} + \hat{M}^{(y)}_{zz;n,n'} \mathcal{H}_{z;n'} \big] \tag{3.43c}$$

$$ik_y \mathcal{H}_{z;n} - \partial_z \mathcal{H}_{y;n} = -ik_0 \big[\hat{Q}^{(y)}_{xx;n,n'} \mathcal{E}_{x;n'} + \hat{Q}^{(y)}_{xy;n,n'} \mathcal{E}_{y;n'} + \hat{Q}^{(y)}_{xz;n,n'} \mathcal{E}_{z;n'} \big] \tag{3.43d}$$

$$\partial_z \mathcal{H}_{x;n} - \partial_x \mathcal{H}_{z;n} = -ik_0 \big[\hat{Q}^{(y)}_{yx;n,n'} \mathcal{E}_{x;n'} + \hat{Q}^{(y)}_{yy;n,n'} \mathcal{E}_{y;n'} + \hat{Q}^{(y)}_{yz;n,n'} \mathcal{E}_{z;n'} \big] \tag{3.43e}$$

$$\partial_x \mathcal{H}_{y;n} - ik_y \mathcal{H}_{x;n} = -ik_0 \big[\hat{Q}^{(y)}_{zx;n,n'} \mathcal{E}_{x;n'} + \hat{Q}^{(y)}_{zy;n,n'} \mathcal{E}_{y;n'} + \hat{Q}^{(y)}_{zz;n,n'} \mathcal{E}_{z;n'} \big], \tag{3.43f}$$

where

$$\hat{M}^{(y)} = \begin{bmatrix} \lfloor \frac{\mu_{xy}}{\mu_{yy}} \rfloor \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} \lfloor \frac{\mu_{yx}}{\mu_{yy}} \rfloor + \lfloor \mu_{xx} - \frac{\mu_{xy}\mu_{yx}}{\mu_{yy}} \rfloor & \lfloor \frac{\mu_{xy}}{\mu_{yy}} \rfloor \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} & \lfloor \frac{\mu_{xy}}{\mu_{yy}} \rfloor \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} \lfloor \frac{\mu_{yz}}{\mu_{yy}} \rfloor + \lfloor \mu_{xz} - \frac{\mu_{xy}\mu_{yz}}{\mu_{yy}} \rfloor \\ \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} \lfloor \frac{\mu_{yx}}{\mu_{yy}} \rfloor & \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} & \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} \lfloor \frac{\mu_{yz}}{\mu_{yy}} \rfloor \\ \lfloor \frac{\mu_{zy}}{\mu_{yy}} \rfloor \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} \lfloor \frac{\mu_{yx}}{\mu_{yy}} \rfloor + \lfloor \mu_{zx} - \frac{\mu_{zy}\mu_{yx}}{\mu_{yy}} \rfloor & \lfloor \frac{\mu_{zy}}{\mu_{yy}} \rfloor \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} & \lfloor \frac{\mu_{zy}}{\mu_{yy}} \rfloor \lfloor \frac{1}{\mu_{yy}} \rfloor^{-1} \lfloor \frac{\mu_{yz}}{\mu_{yy}} \rfloor + \lfloor \mu_{zz} - \frac{\mu_{zy}\mu_{yz}}{\mu_{yy}} \rfloor \end{bmatrix} \tag{3.43g}$$

$$\hat{Q}^{(y)} = \begin{bmatrix} \lfloor \frac{\epsilon_{xy}}{\epsilon_{yy}} \rfloor \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} \lfloor \frac{\epsilon_{yx}}{\epsilon_{yy}} \rfloor + \lfloor \epsilon_{xx} - \frac{\epsilon_{xy}\epsilon_{yx}}{\epsilon_{yy}} \rfloor & \lfloor \frac{\epsilon_{xy}}{\epsilon_{yy}} \rfloor \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} & \lfloor \frac{\epsilon_{xy}}{\epsilon_{yy}} \rfloor \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} \lfloor \frac{\epsilon_{yz}}{\epsilon_{yy}} \rfloor + \lfloor \epsilon_{xz} - \frac{\epsilon_{xy}\epsilon_{yz}}{\epsilon_{yy}} \rfloor \\ \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} \lfloor \frac{\epsilon_{yx}}{\epsilon_{yy}} \rfloor & \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} & \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} \lfloor \frac{\epsilon_{yz}}{\epsilon_{yy}} \rfloor \\ \lfloor \frac{\epsilon_{zy}}{\epsilon_{yy}} \rfloor \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} \lfloor \frac{\epsilon_{yx}}{\epsilon_{yy}} \rfloor + \lfloor \epsilon_{zx} - \frac{\epsilon_{zy}\epsilon_{yx}}{\epsilon_{yy}} \rfloor & \lfloor \frac{\epsilon_{zy}}{\epsilon_{yy}} \rfloor \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} & \lfloor \frac{\epsilon_{zy}}{\epsilon_{yy}} \rfloor \lfloor \frac{1}{\epsilon_{yy}} \rfloor^{-1} \lfloor \frac{\epsilon_{yz}}{\epsilon_{yy}} \rfloor + \lfloor \epsilon_{zz} - \frac{\epsilon_{zy}\epsilon_{yz}}{\epsilon_{yy}} \rfloor \end{bmatrix}. \tag{3.43h}$$

The two equations marked in gray are not needed in the subsequent derivation. In the following, these partially transformed Maxwell equations are transformed also in the second coordinate.

Fourier transform Eqs. (3.42a, 3.42d) in y

Solve Eqs. (3.42b, 3.42e) for $\mathcal{H}_{y;m}, \mathcal{E}_{y;m}$ and insert the result into the RHS of Eqs. (3.42a, 3.42d),

$$\begin{aligned}\partial_y \mathcal{E}_{z;m} - \partial_z \mathcal{E}_{y;m} = +ik_0 \Big[& -i/k_0 \left\{ \hat{M}^{(x)}_{xy} \hat{M}^{(x)\,-1}_{yy} \right\}_{m,m'} \left(\partial_z \mathcal{E}_{x;m'} - ik_x \mathcal{E}_{z;m'} \right) \\ & + \left\{ \hat{M}^{(x)}_{xx} - \hat{M}^{(x)}_{xy} \hat{M}^{(x)\,-1}_{yy} \hat{M}^{(x)}_{yx} \right\}_{m,m'} \mathcal{H}_{x;m'} \\ & + \left\{ \hat{M}^{(x)}_{xz} - \hat{M}^{(x)}_{xy} \hat{M}^{(x)\,-1}_{yy} \hat{M}^{(x)}_{yz} \right\}_{m,m'} \mathcal{H}_{z;m'} \Big] \end{aligned} \tag{3.44a}$$

$$\begin{aligned}\partial_y \mathcal{H}_{z;m} - \partial_z \mathcal{H}_{y;m} = -ik_0 \Big[& +i/k_0 \left\{ \hat{Q}^{(x)}_{xy} \hat{Q}^{(x)\,-1}_{yy} \right\}_{m,m'} \left(\partial_z \mathcal{H}_{x;m'} - ik_x \mathcal{H}_{z;m'} \right) \\ & + \left\{ \hat{Q}^{(x)}_{xx} - \hat{Q}^{(x)}_{xy} \hat{Q}^{(x)\,-1}_{yy} \hat{Q}^{(x)}_{yx} \right\}_{m,m'} \mathcal{E}_{x;m'} \\ & + \left\{ \hat{Q}^{(x)}_{xz} - \hat{Q}^{(x)}_{xy} \hat{Q}^{(x)\,-1}_{yy} \hat{Q}^{(x)}_{yz} \right\}_{m,m'} \mathcal{E}_{z;m'} \Big] . \end{aligned} \tag{3.44b}$$

Consider the bracket on the RHS of Eq. (3.44a). The quantities $\mathcal{B}_{y,m'} = \partial_z \mathcal{E}_{x;m'} - \partial_x \mathcal{E}_{z;m'}$, $\mathcal{H}_{x;m'}$ and $\mathcal{H}_{z;m'}$ are continuous in y. Therefore, the direct rule applies to the prefactors. The same argument applies to the transformation of Eq. (3.44b). By Fourier transformation in y one obtains

$$\begin{aligned} ik_y \mathcal{E}_{z;mn} - \partial_z \mathcal{E}_{y;mn} = +ik_0 \Big[& -i/k_0 \lfloor \hat{M}^{(x)}_{xy} \hat{M}^{(x)\,-1}_{yy} \rfloor_{mn,m'n'} \left(\partial_z \mathcal{E}_{x;m',n'} - ik_x \mathcal{E}_{z;m',n'} \right) \\ & + \lfloor \hat{M}^{(x)}_{xx} - \hat{M}^{(x)}_{xy} \hat{M}^{(x)\,-1}_{yy} \hat{M}^{(x)}_{yx} \rfloor_{mn,m'n'} \mathcal{H}_{x;m',n'} \\ & + \lfloor \hat{M}^{(x)}_{xz} - \hat{M}^{(x)}_{xy} \hat{M}^{(x)\,-1}_{yy} \hat{M}^{(x)}_{yz} \rfloor_{mn,m'n'} \mathcal{H}_{z;m',n'} \Big] \end{aligned} \tag{3.45a}$$

$$\begin{aligned} ik_y \mathcal{H}_{z;mn} - \partial_z \mathcal{H}_{y;mn} = -ik_0 \Big[& +i/k_0 \lfloor \hat{Q}^{(x)}_{xy} \hat{Q}^{(x)\,-1}_{yy} \rfloor_{mn,m'n'} \left(\partial_z \mathcal{H}_{x;m',n'} - ik_x \mathcal{H}_{z;m',n'} \right) \\ & + \lfloor \hat{Q}^{(x)}_{xx} - \hat{Q}^{(x)}_{xy} \hat{Q}^{(x)\,-1}_{yy} \hat{Q}^{(x)}_{yx} \rfloor_{mn,m'n'} \mathcal{E}_{x;m',n'} \\ & + \lfloor \hat{Q}^{(x)}_{xz} - \hat{Q}^{(x)}_{xy} \hat{Q}^{(x)\,-1}_{yy} \hat{Q}^{(x)}_{yz} \rfloor_{mn,m'n'} \mathcal{E}_{z;m',n'} \Big] . \end{aligned} \tag{3.45b}$$

Fourier transform Eqs. (3.43b, 3.43e) in x

In the same way, one Fourier transforms Eqs. (3.43b, 3.43e) in x. The result is

$$\begin{aligned} \partial_z \mathcal{E}_{x;mn} - ik_x \mathcal{E}_{z;mn} = +ik_0 \Big[& -i/k_0 \lceil \hat{M}^{(y)}_{yx} \hat{M}^{(y)\,-1}_{xx} \rceil_{mn,m'n'} \left(ik_y \mathcal{E}_{z;m',n'} - \partial_z \mathcal{E}_{y;m',n'} \right) \\ & + \lceil \hat{M}^{(y)}_{yy} - \hat{M}^{(y)}_{yx} \hat{M}^{(y)\,-1}_{xx} \hat{M}^{(y)}_{xy} \rceil_{mn,m'n'} \mathcal{H}_{y;m',n'} \\ & + \lceil \hat{M}^{(y)}_{yz;n,n'} - \hat{M}^{(y)}_{yx} \hat{M}^{(y)\,-1}_{xx} \hat{M}^{(y)}_{xz} \rceil_{mn,m'n'} \mathcal{H}_{z;m',n'} \Big] \end{aligned} \tag{3.46a}$$

$$\begin{aligned} \partial_z \mathcal{H}_{x;mn} - ik_x \mathcal{H}_{z;mn} = -ik_0 \Big[& +i/k_0 \lceil \hat{Q}^{(y)}_{yx} \hat{Q}^{(y)\,-1}_{xx} \rceil_{mn,m'n'} \left(ik_y \mathcal{H}_{z;m',n'} - \partial_z \mathcal{H}_{y;m',n'} \right) \\ & + \lceil \hat{Q}^{(y)}_{yy} - \hat{Q}^{(y)}_{yx} \hat{Q}^{(y)\,-1}_{xx} \hat{Q}^{(y)}_{xy} \rceil_{mn,m'n'} \mathcal{E}_{y;m',n'} \\ & + \lceil \hat{Q}^{(y)}_{yz} - \hat{Q}^{(y)}_{yx} \hat{Q}^{(y)\,-1}_{xx} \hat{Q}^{(y)}_{xz} \rceil_{mn,m'n'} \mathcal{E}_{z;m',n'} \Big] . \end{aligned} \tag{3.46b}$$

Fourier transform Eqs. (3.42c, 3.42f) in y

In the same way, one Fourier transforms Eqs. (3.42c, 3.42f) in y. The result is

$$\begin{aligned} ik_x\mathcal{E}_{y;mn}-ik_y\mathcal{E}_{x;mn}=+ik_0\Big[&-i/k_0\,\lfloor\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor_{mn,m'n'}\big(\partial_z\mathcal{E}_{x;m',n'}-ik_x\mathcal{E}_{z;m',n'}\big)\\ &+\lfloor\hat{M}_{zx}^{(x)}-\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yx}^{(x)}\rfloor_{mn,m'n'}\mathcal{H}_{x;m',n'}\\ &+\lfloor\hat{M}_{zz}^{(x)}-\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yz}^{(x)}\rfloor_{mn,m'n'}\mathcal{H}_{z;m',n'}\Big] \end{aligned} \tag{3.47a}$$

$$\begin{aligned} ik_x\mathcal{H}_{y;mn}-ik_y\mathcal{H}_{x;mn}=-ik_0\Big[&+i/k_0\,\lfloor\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor_{mn,m'n'}\big(\partial_z\mathcal{H}_{x;m',n'}-ik_x\mathcal{H}_{z;m',n'}\big)\\ &+\lfloor\hat{Q}_{zx}^{(x)}-\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yx}^{(x)}\rfloor_{mn,m'n'}\mathcal{E}_{x;m',n'}\\ &+\lfloor\hat{Q}_{zz}^{(x)}-\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yz}^{(x)}\rfloor_{mn,m'n'}\mathcal{E}_{z;m',n'}\Big], \end{aligned} \tag{3.47b}$$

Fully transformed Maxwell's equations

Now insert

1. Eqs. (3.46) in Eqs. (3.45) [obtain Eqs. (3.48a, 3.48d)],
2. Eqs. (3.45) in Eqs. (3.46) [obtain Eqs. (3.48b, 3.48e)], and then
3. Eqs. (3.48b, 3.48e) in Eqs. (3.47) [obtain Eqs. (3.48c, 3.48f)].

Now Maxwell's equations are Fourier transformed in both x and y,

$$ik_{y;mn}\mathcal{E}_{z;mn}-\partial_z\mathcal{E}_{y;mn}=+ik_0\Big[\big\{\hat{M}_{xx}\mathcal{H}_x\big\}_{mn}+\big\{\hat{M}_{xy}\mathcal{H}_y\big\}_{mn}+\big\{\hat{M}_{xz}\mathcal{H}_z\big\}_{mn}\Big] \tag{3.48a}$$

$$\partial_z\mathcal{E}_{x;mn}-ik_{x;mn}\mathcal{E}_{z;mn}=+ik_0\Big[\big\{\hat{M}_{yx}\mathcal{H}_x\big\}_{mn}+\big\{\hat{M}_{yy}\mathcal{H}_y\big\}_{mn}+\big\{\hat{M}_{yz}\mathcal{H}_z\big\}_{mn}\Big] \tag{3.48b}$$

$$ik_{x;mn}\mathcal{E}_{y;mn}-ik_{y;mn}\mathcal{E}_{x;mn}=+ik_0\Big[\big\{\hat{M}_{zx}\mathcal{H}_x\big\}_{mn}+\big\{\hat{M}_{zy}\mathcal{H}_y\big\}_{mn}+\big\{\hat{M}_{zz}\mathcal{H}_z\big\}_{mn}\Big] \tag{3.48c}$$

$$ik_{y;mn}\mathcal{H}_{z;mn}-\partial_z\mathcal{H}_{y;mn}=-ik_0\Big[\big\{\hat{Q}_{xx}\mathcal{E}_x\big\}_{mn}+\big\{\hat{Q}_{xy}\mathcal{E}_y\big\}_{mn}+\big\{\hat{Q}_{xz}\mathcal{E}_z\big\}_{mn}\Big] \tag{3.48d}$$

$$\partial_z\mathcal{H}_{x;mn}-ik_{x;mn}\mathcal{H}_{z;mn}=-ik_0\Big[\big\{\hat{Q}_{yx}\mathcal{E}_x\big\}_{mn}+\big\{\hat{Q}_{yy}\mathcal{E}_y\big\}_{mn}+\big\{\hat{Q}_{yz}\mathcal{E}_z\big\}_{mn}\Big] \tag{3.48e}$$

$$ik_{x;mn}\mathcal{H}_{y;mn}-ik_{y;mn}\mathcal{H}_{x;mn}=-ik_0\Big[\big\{\hat{Q}_{zx}\mathcal{E}_x\big\}_{mn}+\big\{\hat{Q}_{zy}\mathcal{E}_y\big\}_{mn}+\big\{\hat{Q}_{zz}\mathcal{E}_z\big\}_{mn}\Big], \tag{3.48f}$$

where

$$\hat{Q}_{xx}=\Big\{\mathbb{1}-\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\Big\}^{-1}\lfloor\hat{Q}_{xx}^{(x)}-\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yx}^{(x)}\rfloor \tag{3.49a}$$

$$\hat{Q}_{xy}=\Big\{\mathbb{1}-\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\Big\}^{-1}\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\lceil\hat{Q}_{yy}^{(y)}-\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\hat{Q}_{xy}^{(y)}\rceil \tag{3.49b}$$

$$\begin{aligned}\hat{Q}_{xz}=&\Big\{\mathbb{1}-\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\Big\}^{-1}\\ &\times\Big\{\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\lceil\hat{Q}_{yz}^{(y)}-\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\hat{Q}_{xz}^{(y)}\rceil+\lfloor\hat{Q}_{xz}^{(x)}-\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yz}^{(x)}\rfloor\Big\}\end{aligned} \tag{3.49c}$$

$$\hat{Q}_{yx}=\Big\{\mathbb{1}-\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xx}^{(x)}-\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yx}^{(x)}\rfloor \tag{3.49d}$$

$$\hat{Q}_{yy}=\Big\{\mathbb{1}-\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}\lceil\hat{Q}_{yy}^{(y)}-\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\hat{Q}_{xy}^{(y)}\rceil \tag{3.49e}$$

$$\hat{Q}_{yz}=\Big\{\mathbb{1}-\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}$$
$$\times\Big\{\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xz}^{(x)}-\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yz}^{(x)}\rfloor+\lceil\hat{Q}_{yz}^{(y)}-\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\hat{Q}_{xz}^{(y)}\rceil\Big\} \tag{3.49f}$$

$$\hat{Q}_{zx}=\lfloor\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\{\mathbb{1}-\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}$$
$$\times\Big\{\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xx}^{(x)}-\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yx}^{(x)}\rfloor\Big\}$$
$$+\lfloor\hat{Q}_{zx}^{(x)}-\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yx}^{(x)}\rfloor \tag{3.49g}$$

$$\hat{Q}_{zy}=\lfloor\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\{\mathbb{1}-\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}\lceil\hat{Q}_{yy}^{(y)}-\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\hat{Q}_{xy}^{(y)}\rceil \tag{3.49h}$$

$$\hat{Q}_{zz}=\lfloor\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\{\mathbb{1}-\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}$$
$$\times\Big\{\lceil\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\rceil\lfloor\hat{Q}_{xz}^{(x)}-\hat{Q}_{xy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yz}^{(x)}\rfloor+\lceil\hat{Q}_{yz}^{(y)}-\hat{Q}_{yx}^{(y)}\,\hat{Q}_{xx}^{(y)\,-1}\hat{Q}_{xz}^{(y)}\rceil\Big\}$$
$$+\lfloor\hat{Q}_{zz}^{(x)}-\hat{Q}_{zy}^{(x)}\,\hat{Q}_{yy}^{(x)\,-1}\hat{Q}_{yz}^{(x)}\rfloor \tag{3.49i}$$

$$\hat{M}_{xx}=\Big\{\mathbb{1}-\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\Big\}^{-1}\lfloor\hat{M}_{xx}^{(x)}-\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yx}^{(x)}\rfloor \tag{3.49j}$$

$$\hat{M}_{xy}=\Big\{\mathbb{1}-\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\Big\}^{-1}\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\lceil\hat{M}_{yy}^{(y)}-\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\hat{M}_{xy}^{(y)}\rceil \tag{3.49k}$$

$$\hat{M}_{xz}=\Big\{\mathbb{1}-\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\Big\}^{-1}$$
$$\times\Big\{\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\lceil\hat{M}_{yz}^{(y)}-\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\hat{M}_{xz}^{(y)}\rceil+\lfloor\hat{M}_{xz}^{(x)}-\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yz}^{(x)}\rfloor\Big\} \tag{3.49l}$$

$$\hat{M}_{yx}=\Big\{\mathbb{1}-\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xx}^{(x)}-\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yx}^{(x)}\rfloor \tag{3.49m}$$

$$\hat{M}_{yy}=\Big\{\mathbb{1}-\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}\lceil\hat{M}_{yy}^{(y)}-\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\hat{M}_{xy}^{(y)}\rceil \tag{3.49n}$$

$$\hat{M}_{yz}=\Big\{\mathbb{1}-\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}$$
$$\times\Big\{\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xz}^{(x)}-\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yz}^{(x)}\rfloor+\lceil\hat{M}_{yz}^{(y)}-\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\hat{M}_{xz}^{(y)}\rceil\Big\} \tag{3.49o}$$

$$\hat{M}_{zx}=\lfloor\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\{\mathbb{1}-\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}$$
$$\times\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xx}^{(x)}-\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yx}^{(x)}\rfloor$$
$$+\lfloor\hat{M}_{zx}^{(x)}-\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yx}^{(x)}\rfloor \tag{3.49p}$$

$$\hat{M}_{zy}=\lfloor\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\{\mathbb{1}-\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}\lceil\hat{M}_{yy}^{(y)}-\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\hat{M}_{xy}^{(y)}\rceil \tag{3.49q}$$

$$\hat{M}_{zz}=\lfloor\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\{\mathbb{1}-\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\rfloor\Big\}^{-1}$$
$$\times\Big\{\lceil\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\rceil\lfloor\hat{M}_{xz}^{(x)}-\hat{M}_{xy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yz}^{(x)}\rfloor+\lceil\hat{M}_{yz}^{(y)}-\hat{M}_{yx}^{(y)}\,\hat{M}_{xx}^{(y)\,-1}\hat{M}_{xz}^{(y)}\rceil\Big\}$$
$$+\lfloor\hat{M}_{zz}^{(x)}-\hat{M}_{zy}^{(x)}\,\hat{M}_{yy}^{(x)\,-1}\hat{M}_{yz}^{(x)}\rfloor\,. \tag{3.49r}$$

Matrix equation

In Eqs. (3.48), eliminate $\mathcal{E}_{z;mn}$ and $\mathcal{H}_{z;mn}$ using Eqs. (3.48c, 3.48f) and solve for $\partial_z \mathcal{E}_{x/y;mn}$ and $\partial_z \mathcal{H}_{x/y;mn}$. One obtains the first order differential equation

$$\frac{\partial}{\partial z}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp(z)\end{bmatrix} = ik_0\begin{bmatrix}\hat{M}_{11} & \hat{M}_{12}\\ \hat{M}_{21} & \hat{M}_{22}\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp(z)\end{bmatrix}, \tag{3.50a}$$

where

$$\hat{M}_{11} = \begin{bmatrix}+\hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{K}_y - \hat{K}_x\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & -\hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{K}_x - \hat{K}_x\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\\ +\hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{K}_y - \hat{K}_y\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & -\hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{K}_x - \hat{K}_y\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\end{bmatrix} \tag{3.50b}$$

$$\hat{M}_{12} = \begin{bmatrix}+\hat{M}_{yx}+\hat{K}_x\hat{Q}_{zz}^{-1}\hat{K}_y - \hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{M}_{zx} & +\hat{M}_{yy}-\hat{K}_x\hat{Q}_{zz}^{-1}\hat{K}_x - \hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{M}_{zy}\\ -\hat{M}_{xx}+\hat{K}_y\hat{Q}_{zz}^{-1}\hat{K}_y + \hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{M}_{zx} & -\hat{M}_{xy}-\hat{K}_y\hat{Q}_{zz}^{-1}\hat{K}_x + \hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{M}_{zy}\end{bmatrix} \tag{3.50c}$$

$$\hat{M}_{21} = \begin{bmatrix}-\hat{Q}_{yx}-\hat{K}_x\hat{M}_{zz}^{-1}\hat{K}_y + \hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & -\hat{Q}_{yy}+\hat{K}_x\hat{M}_{zz}^{-1}\hat{K}_x + \hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\\ +\hat{Q}_{xx}-\hat{K}_y\hat{M}_{zz}^{-1}\hat{K}_y - \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & +\hat{Q}_{xy}+\hat{K}_y\hat{M}_{zz}^{-1}\hat{K}_x - \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\end{bmatrix} \tag{3.50d}$$

$$\hat{M}_{22} = \begin{bmatrix}-\hat{K}_x\hat{M}_{zz}^{-1}\hat{M}_{zx} - \hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{K}_y & -\hat{K}_x\hat{M}_{zz}^{-1}\hat{M}_{zy} + \hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{K}_x\\ -\hat{K}_y\hat{M}_{zz}^{-1}\hat{M}_{zx} + \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{K}_y & -\hat{K}_y\hat{M}_{zz}^{-1}\hat{M}_{zy} - \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{K}_x\end{bmatrix}, \tag{3.50e}$$

with $\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)$, $\widetilde{\boldsymbol{\mathcal{H}}}_\perp(z)$ and $\hat{K}_{x/y}$ as before. Just like for the linear isotropic case, also here the layer index was dropped in the beginning of the derivation, i.e. the coupling matrix in Eq. (6.25a) is layer dependent, i.e. $\hat{M}_{11} \equiv \hat{M}_{11}^{(l)}$, $\hat{M}_{12} \equiv \hat{M}_{12}^{(l)}$, $\hat{M}_{21} \equiv \hat{M}_{21}^{(l)}$, and $\hat{M}_{22} \equiv \hat{M}_{22}^{(l)}$. In fact, it can be easily checked that the coupling matrix simplifies to the same result as derived in Sect. 3.3.1 in the linear isotropic case.

3.4 Eigenspace solution and eigenmodes

3.4.1 Linear isotropic media

Solution for the electric field

The matrix equation for the l^{th} layer, Eq. (3.32a), can be transformed into the second order differential equation for the electric field,

$$\frac{\partial^2}{\partial z^2}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) = -k_0^2\hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z), \tag{3.51}$$

$$\hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)} = \begin{bmatrix} -\hat{K}_x \left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1} \hat{K}_x \lfloor\lceil\epsilon^{(l)}\rceil\rfloor - \lceil\lfloor\mu^{(l)}\rfloor\rceil \hat{K}_y \left[\!\left[\mu^{(l)}\right]\!\right]^{-1} \hat{K}_y + \lceil\lfloor\mu^{(l)}\rfloor\rceil \lfloor\lceil\epsilon^{(l)}\rceil\rfloor & \cdots \\ \lfloor\lceil\mu^{(l)}\rceil\rfloor \hat{K}_x \left[\!\left[\mu^{(l)}\right]\!\right]^{-1} \hat{K}_y - \hat{K}_y \left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1} \hat{K}_x \lfloor\lceil\epsilon^{(l)}\rceil\rfloor & \end{bmatrix}$$

$$\begin{bmatrix} & \cdots \\ -\hat{K}_x \left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1} \hat{K}_y \lceil\lfloor\epsilon^{(l)}\rfloor\rceil + \lceil\lfloor\mu^{(l)}\rfloor\rceil \hat{K}_y \left[\!\left[\mu^{(l)}\right]\!\right]^{-1} \hat{K}_x & \\ -\hat{K}_y \left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1} \hat{K}_y \lceil\lfloor\epsilon^{(l)}\rfloor\rceil - \lfloor\lceil\mu^{(l)}\rceil\rfloor \hat{K}_x \left[\!\left[\mu^{(l)}\right]\!\right]^{-1} \hat{K}_x + \lfloor\lceil\mu^{(l)}\rceil\rfloor \lceil\lfloor\epsilon^{(l)}\rfloor\rceil & \end{bmatrix}. \tag{3.52}$$

Be $\hat{W}^{(l)}$ the right eigenvector matrix of $\hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)}$. Then Eq. (3.51) can be decoupled via

$$\frac{\partial^2}{\partial z^2}\left[\hat{W}^{(l)-1}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)\right] = -k_0^2 \underbrace{\hat{W}^{(l)-1}\hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)}\hat{W}^{(l)}}_{\hat{\Lambda}^{(l)}} \hat{W}^{(l)-1}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) \tag{3.53}$$

where $\hat{\Lambda}^{(l)} = \operatorname{diag}\left(\lambda_1^{(l)}, \cdots, \lambda_{2\bar{N}}^{(l)}\right)$ is the corresponding diagonal eigenvalue matrix, and $\bar{N} = MN$ is the total number of modes in the calculation. The most general solution of this differential equation is

$$\left[\hat{W}^{(l)-1}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)\right]_q = c_{+;q}^{(l)} e^{+ik_0\sqrt{\lambda_q^{(l)}}z} + c_{-;q}^{(l)} e^{-ik_0\sqrt{\lambda_q^{(l)}}\left(z-d^{(l)}\right)}. \tag{3.54}$$

where $d^{(l)}$ is the thickness and the position $z = 0$ is defined as the left interface of the l^{th} layer[4]. They are typically interpreted as the forward and backward propagating eigenmode coefficients inside a layer. It is shown in Chapter 4 that this assignment is correct for most physical materials, but not generally true. Eq. (3.54) can be solved for $\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)$,

$$\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) = \begin{bmatrix}\hat{W}^{(l)} & \hat{W}^{(l)}\end{bmatrix} \begin{bmatrix} \hat{P}_+^{(l)}(z) & \mathbb{0} \\ \mathbb{0} & \hat{P}_-^{(l)}(z) \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(l)} \\ \mathbf{c}_-^{(l)} \end{bmatrix}, \tag{3.55}$$

where $P_{+;qq'}^{(l)}(z) = e^{+ik_0\sqrt{\lambda_q^{(l)}}z}\,\delta_{qq'}$ and $P_{-;qq'}^{(l)}(z) = e^{-ik_0\sqrt{\lambda_q^{(l)}}\left(z-d^{(l)}\right)}\,\delta_{qq'}$.

Solution for the magnetic field

From the matrix equation, Eq. (3.32a), one knows the relation

$$\begin{aligned}\widetilde{\boldsymbol{\mathcal{H}}}_\perp(z) &= -\frac{i}{k_0}\hat{M}_{12}^{(l)-1}\frac{\partial}{\partial z}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) \\ &= -\frac{i}{k_0}\hat{M}_{21}^{(l)}\hat{W}^{(l)}\hat{\Lambda}^{(l)-1}\hat{W}^{(l)-1}\frac{\partial}{\partial z}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z).\end{aligned} \tag{3.56}$$

[4]Mathematically, the z-shift by $d^{(l)}$ in the second term is not required, since the effect is only a multiplication of the coefficient vector $\mathbf{c}_{-;m}^{(l)}$ by a constant. When enforcing the boundary conditions in Sect. 3.5, this shift is *only* assumed for the ETMA.

By inserting $\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)$ from Eq. (3.55) and carrying out the z-derivative one obtains

$$\widetilde{\boldsymbol{\mathcal{H}}}_{\perp;m'}(z) = \sum_{q'} \underbrace{\left[\hat{M}_{21}^{(l)} \hat{W}^{(l)} \hat{\Lambda}^{(l)-\frac{1}{2}}\right]_{qq'}}_{V_{qq'}} \left(c_{+;q'}^{(l)}\, e^{+ik_0\sqrt{\lambda_{q'}^{(l)}}z} - c_{-;q'}^{(l)}\, e^{-ik_0\sqrt{\lambda_{q'}^{(l)}}\left(z-d^{(l)}\right)} \right) . \tag{3.57}$$

In matrix form, this is

$$\widetilde{\boldsymbol{\mathcal{H}}}_\perp(z) = \begin{bmatrix} \hat{V}^{(l)} & -\hat{V}^{(l)} \end{bmatrix} \begin{bmatrix} \hat{P}_+^{(l)}(z) & \mathbb{0} \\ \mathbb{0} & \hat{P}_-^{(l)}(z) \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(l)} \\ \mathbf{c}_-^{(l)} \end{bmatrix} . \tag{3.58}$$

Combined solution

Eqs. (3.55, 3.58) can be combined in one equation,

$$\underbrace{\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp(z) \end{bmatrix}}_{\text{Fourier modes}} = \underbrace{\begin{bmatrix} +\hat{W}^{(l)} & +\hat{W}^{(l)} \\ +\hat{V}^{(l)} & -\hat{V}^{(l)} \end{bmatrix}}_{\substack{\text{conversion and} \\ \text{coupling matrix}}} \underbrace{\begin{bmatrix} \hat{P}_+^{(l)}(z) & \mathbb{0} \\ \mathbb{0} & \hat{P}_-^{(l)}(z) \end{bmatrix}}_{\text{propagation}} \underbrace{\begin{bmatrix} \mathbf{c}_+^{(l)} \\ \mathbf{c}_-^{(l)} \end{bmatrix}}_{\substack{\text{eigenmode} \\ \text{coefficients}}} , \tag{3.59}$$

which describes the light propagation in axial direction inside the l^{th} layer. It decomposes into three parts.

1. The eigenmode coefficient vector $\begin{bmatrix} \mathbf{c}_+^{(l)} & \mathbf{c}_-^{(l)} \end{bmatrix}^T$: at this point the coefficients are unknown and need to be determined (cf. Sect. 3.5).
2. The propagation matrix: consists of the z-dependent forward $\hat{P}_+^{(l)}(z)$ and backward $\hat{P}_-^{(l)}(z)$ propagator.
3. The conversion and coupling operator: the matrix $\hat{W}^{(l)}$ converts from eigenmode space to the electric field's Fourier space. The matrix $\hat{V}^{(l)}$ does the same, but subsequently conducts an additional conversion from the electric to the magnetic field.

Interpretation of $\sqrt{\lambda_q^{(l)}}$

Almost all real passive materials have an at least slight positive absorption rate, i.e. their electric permittivity and/or magnetic permeability have a positive imaginary part. The latter transfers directly onto the eigenvalues $\lambda_q^{(l)}$. Therefore, for these types of materials, $\sqrt{\lambda_q^{(l)}}$ lies in the first quadrant, i.e. has a positive real and imaginary part. As a result, as one can see from the form of $\hat{P}_\pm^{(l)}(z)$, the forward/backward propagating mode is exponentially decreasing/increasing in $+z$-direction. If, however, the initial assumption (the medium has a positive absorption rate) is violated,

for instance for active materials, then the normalization of the eigenmode coefficients has to be reconsidered (cf. Chapter 4).

3.4.2 Linear anisotropic media

The coupling matrix of the l^{th} layer from Eq. (6.25),

$$\frac{\partial}{\partial z}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix} = ik_0\begin{bmatrix}\hat{M}_{11}^{(l)} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \hat{M}_{22}^{(l)}\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}, \tag{3.60}$$

is diagonalized, so that

$$\frac{\partial}{\partial z}\left\{\hat{W}^{(l)-1}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}\right\} = ik_0\underbrace{\hat{W}^{(l)-1}\begin{bmatrix}\hat{M}_{11}^{(l)} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \hat{M}_{22}^{(l)}\end{bmatrix}\hat{W}^{(l)}}_{\hat{\Lambda}^{(l)}}\hat{W}^{(l)-1}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}, \tag{3.61}$$

where $\hat{\Lambda}^{(l)} = \operatorname{diag}\left(\lambda_1^{(l)},\cdots,\lambda_{4\bar{N}}^{(l)}\right)$ is the corresponding diagonal eigenvalue matrix[5]. The most general solution of this first order differential equation is

$$\left\{\hat{W}^{(l)-1}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}\right\}_q = c_q^{(l)}\, e^{+ik_0\lambda_q^{(l)}z}. \tag{3.62}$$

Eq. (3.62) can be solved for $\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}$,

$$\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix}_q = \sum_{q'} W_{qq'}^{(l)} c_{q'}^{(l)}\, e^{+ik_0\lambda_{q'}^{(l)}z}. \tag{3.63}$$

$$\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}(z)\end{bmatrix} = \hat{W}^{(l)}\hat{P}^{(l)}(z)\,\mathbf{c}^{(l)}, \tag{3.64}$$

where $P_{qq'}^{(l)}(z) = e^{+ik_0\lambda_q^{(l)}z}\,\delta_{qq'}$.

[5] It is fairly easy to show that, in the isotropic case, the eigenvalues found with this approach are indeed equal to plus or minus the square root of the eigenvalues found in Sect. 3.4.1. The reason is that, in isotropic structures, there exist always two identical eigenmodes, that differ only by the sign of their axial propagation constant.

3.5 Enforcing boundary conditions

The boundary conditions are defined by the light, which is incident from $z = \pm\infty$ on the grating. In this section, several methods for enforcing the boundary conditions are presented, which allow determining the reflected and transmitted Fourier mode coefficients and the eigenmode coefficients inside each layer. In literature, the different approaches to solve the boundary value problem can be classified in groups, depending on the quantities that are related to each other. Fig. 3.8 gives a first overview of the situation.

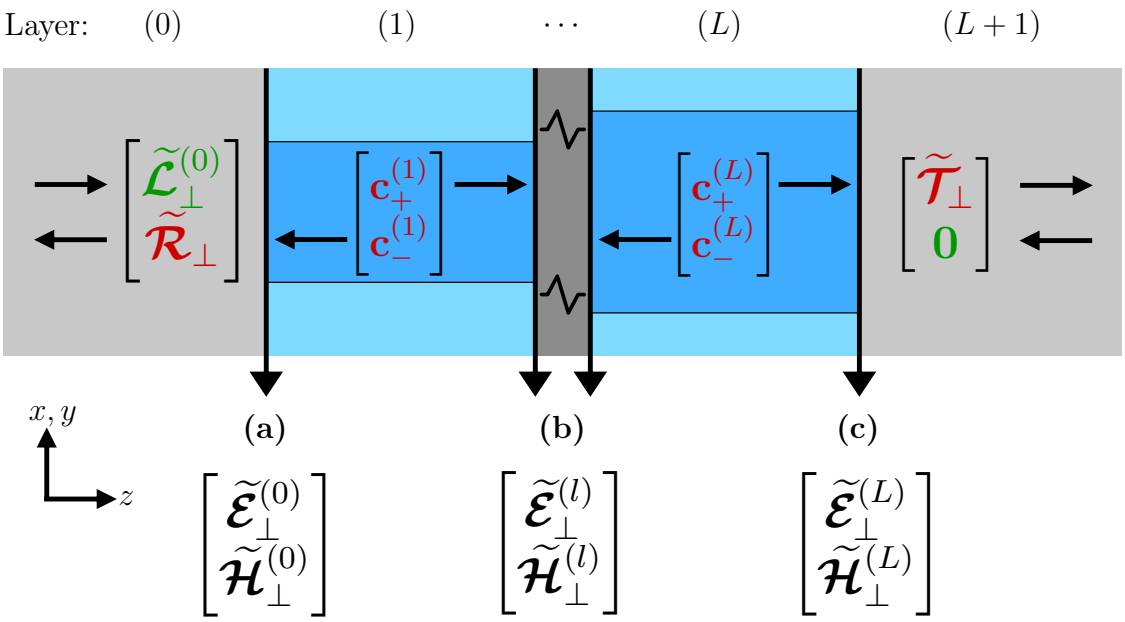


Figure 3.8: ***Given***: incident light Fourier mode coefficients from the left ($\widetilde{\boldsymbol{\mathcal{L}}}_\perp^{(0)}$) and right ($\mathbf{0}$). ***Unknown***: reflected ($\widetilde{\boldsymbol{\mathcal{R}}}_\perp$) and transmitted ($\widetilde{\boldsymbol{\mathcal{T}}}_\perp$) Fourier mode coefficients, eigenmode coefficients of each layer ($\mathbf{c}_\pm^{(l)}$). At the layer interfaces (a-c), the tangential electric and magnetic field components ($\widetilde{\boldsymbol{\mathcal{E}}}_\perp$, $\widetilde{\boldsymbol{\mathcal{H}}}_\perp$) must be continuous.

3.5.1 T-matrix approach

The 'transfer matrix approach' (T-matrix approach)[10] relates the Fourier components of the electric and magnetic fields on the two sides of a diffracting structure via the T-matrix. By evaluating Eq. (3.59), one obtains

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp^{(l)} \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp^{(l)} \end{bmatrix} = \underbrace{\begin{bmatrix} +\hat{W}^{(l)} & +\hat{W}^{(l)} \\ +\hat{V}^{(l)} & -\hat{V}^{(l)} \end{bmatrix} \begin{bmatrix} \hat{X}^{(l)} & \mathbb{0} \\ \mathbb{0} & \hat{X}^{(l)-1} \end{bmatrix} \begin{bmatrix} +\hat{W}^{(l)} & +\hat{W}^{(l)} \\ +\hat{V}^{(l)} & -\hat{V}^{(l)} \end{bmatrix}^{-1}}_{\hat{T}^{(l)}} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp^{(l-1)} \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp^{(l-1)} \end{bmatrix}, \quad (3.65)$$

where $\hat{X}^{(l)}$ are the diagonal eigenmode propagation matrices, with $X^{(l)}_{qq'} = e^{ik_0\sqrt{\lambda_q^{(l)}}d^{(l)}}\,\delta_{qq'}$, and $\hat{T}^{(l)}$ is the T-matrix of the l^{th} layer. By combining the T-matrices of all layers, one obtains

$$\underbrace{\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(L+1)} & -\hat{C}^{(L+1)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{T}}}_\perp \\ \mathbf{0} \end{bmatrix}}_{\left[\widetilde{\boldsymbol{\mathcal{E}}}_\perp^{(L)} \quad \widetilde{\boldsymbol{\mathcal{H}}}_\perp^{(L)}\right]^T} = \underbrace{\prod_{l=L}^{1} \hat{T}^{(l)}}_{\hat{T}} \underbrace{\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_\perp^{(0)} \\ \widetilde{\boldsymbol{\mathcal{R}}}_\perp \end{bmatrix}}_{\left[\widetilde{\boldsymbol{\mathcal{E}}}_\perp^{(0)} \quad \widetilde{\boldsymbol{\mathcal{H}}}_\perp^{(0)}\right]^T}, \tag{3.66}$$

where $\hat{T}$ is the T-matrix of the whole grating and $\hat{C}^{(l)}$ is defined as the matrix that converts the electric- into the magnetic-field Fourier mode coefficient vector,

$$\widetilde{\boldsymbol{\mathcal{H}}}_\perp = \hat{C}^{(l)} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \tag{3.67a}$$

$$\hat{C}^{(l)} = \begin{bmatrix} -\frac{\hat{K}_x \hat{K}_y}{\mu^{(l)} \hat{K}_z^{(l)}} & -\frac{\hat{K}_y^2 + \hat{K}_z^{(l)\,2}}{\mu^{(l)} \hat{K}_z^{(l)}} \\ \frac{\hat{K}_x^2 + \hat{K}_z^{(l)\,2}}{\mu^{(l)} \hat{K}_z^{(l)}} & \frac{\hat{K}_x \hat{K}_y}{\mu^{(l)} \hat{K}_z^{(l)}} \end{bmatrix}. \tag{3.67b}$$

This form of the conversion matrix can be obtained directly from Eqs. (2.57a, 2.57b) and is only valid for homogeneous layers. Mathematically, the T-matrix approach seems straightforward and very elegant. However, it is numerically unstable[10]. The reason is that in each of the layer T-matrices, $\hat{T}^{(l)}$, a propagation of the eigenmodes is done in both positive and negative direction, which is represented by the matrices $\hat{X}^{(l)}$ and $\hat{X}^{(l)-1}$. The inverse of the matrix $\hat{X}^{(l)}$ contains, however, potentially very large terms, which is always the case when evanescent waves or absorption are involved. In this case, the matrix multiplication in Eq. (3.65) is not properly conducted in a numerical sense due to arithmetic overflow.

3.5.2 S-matrix approach

In contrast to the T-matrix approach, the S-matrix relates the *eigenmode* coefficients in front and behind a diffracting structure to each other. More precisely, the S-matrix of the l^{th} layer relates the forward and backward propagating eigenmode coefficients $\mathbf{c}_\pm^{(l+1)}$ and $\mathbf{c}_\pm^{(l-1)}$ via[66]

$$\begin{bmatrix} \mathbf{c}_-^{(l-1)} \\ \mathbf{c}_+^{(l+1)} \end{bmatrix} = \hat{S}^{(l)} \begin{bmatrix} \mathbf{c}_+^{(l-1)} \\ \mathbf{c}_-^{(l+1)} \end{bmatrix}. \tag{3.68}$$

The first publication, in which this approach was applied to modal methods, was by Li in 1996[10]. More modern publications involve, for instance, [66] from 2011 with an updated notation. In order to better understand the term 'scattering matrix', consider Fig. 3.9. The 2×2-blocks of $\hat{S}^{(l)}$ have the following meaning.

- $\hat{S}_{21}^{(l)}$: forward scattering of the incident light from the left.

- $\hat{S}_{12}^{(l)}$: backward scattering of the incident light from the right.
- $\hat{S}_{11}^{(l)}$: backward scattering of the incident light from the left.
- $\hat{S}_{22}^{(l)}$: forward scattering of the incident light from the right.

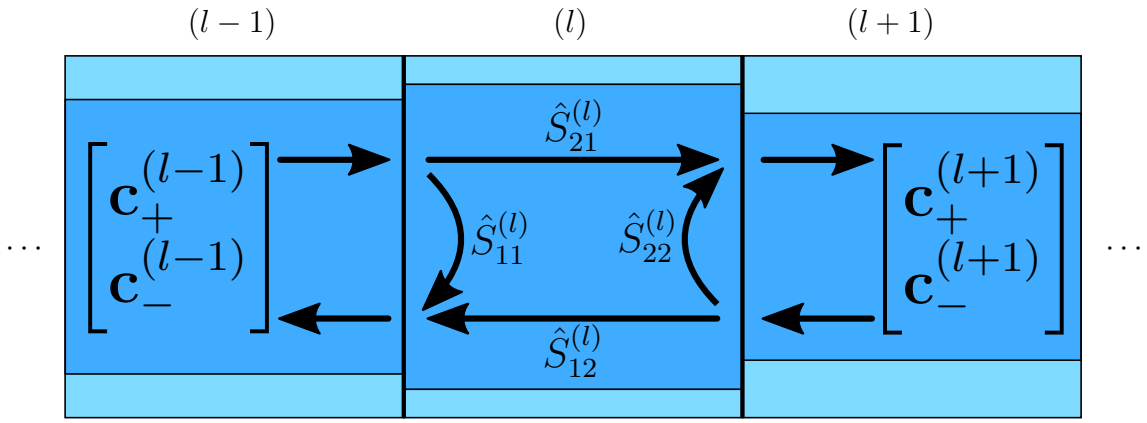


Figure 3.9: Meaning of the S-matrix elements. The four 2×2-blocks of $\hat{S}^{(l)}$ describe the scattering of each of the two input/incident channels into each of the two output channels.

The advantage of this approach becomes apparent when the S-matrix elements are written out explicitly. By evaluating Eq. (3.59) at both interfaces of the l^{th} layer, one obtains

$$\hat{S}_{11}^{(l)} = \left(\hat{A}_{+}^{(l,l-1)} - \hat{X}^{(l)}\hat{A}_{-}^{(l,l+1)}\hat{A}_{+}^{(l,l+1)-1}\hat{X}^{(l)}\hat{A}_{-}^{(l,l-1)}\right)^{-1} \times \left(\hat{X}^{(l)}\hat{A}_{-}^{(l,l+1)}\hat{A}_{+}^{(l,l+1)-1}\hat{X}^{(l)}\hat{A}_{+}^{(l,l-1)} - \hat{A}_{-}^{(l,l-1)}\right) \tag{3.69a}$$

$$\hat{S}_{12}^{(l)} = \left(\hat{A}_{+}^{(l,l-1)} - \hat{X}^{(l)}\hat{A}_{-}^{(l,l+1)}\hat{A}_{+}^{(l,l+1)-1}\hat{X}^{(l)}\hat{A}_{-}^{(l,l-1)}\right)^{-1}\hat{X}^{(l)} \times \left(\hat{A}_{+}^{(l,l+1)} - \hat{A}_{-}^{(l,l+1)}\hat{A}_{+}^{(l,l+1)-1}\hat{A}_{-}^{(l,l+1)}\right) \tag{3.69b}$$

$$\hat{S}_{21}^{(l)} = \left(\hat{A}_{+}^{(l,l+1)} - \hat{X}^{(l)}\hat{A}_{-}^{(l,l-1)}\hat{A}_{+}^{(l,l-1)-1}\hat{X}^{(l)}\hat{A}_{-}^{(l,l+1)}\right)^{-1}\hat{X}^{(l)} \times \left(\hat{A}_{+}^{(l,l-1)} - \hat{A}_{-}^{(l,l-1)}\hat{A}_{+}^{(l,l-1)-1}\hat{A}_{-}^{(l,l-1)}\right) \tag{3.69c}$$

$$\hat{S}_{22}^{(l)} = \left(\hat{A}_{+}^{(l,l+1)} - \hat{X}^{(l)}\hat{A}_{-}^{(l,l-1)}\hat{A}_{+}^{(l,l-1)-1}\hat{X}^{(l)}\hat{A}_{-}^{(l,l+1)}\right)^{-1} \times \left(\hat{X}^{(l)}\hat{A}_{-}^{(l,l-1)}\hat{A}_{+}^{(l,l-1)-1}\hat{X}^{(l)}\hat{A}_{+}^{(l,l+1)} - \hat{A}_{-}^{(l,l+1)}\right) \tag{3.69d}$$

$$\hat{A}_{\pm}^{ll'} = \hat{W}^{(l)-1}\hat{W}^{(l')} \pm \hat{V}^{(l)-1}\hat{V}^{(l')}\,. \tag{3.69e}$$

A more detailed derivation is presented in [66]. In the S-matrix elements, Eqs. (3.69), the matrix $\hat{X}^{(l)}$ is not directly inverted at any position, i.e. the arithmetic overflow

is avoided. In order to determine the S-matrix of a layer stack, it is necessary to combine the S-matrices of two or more successive layers or structures to one single composite S-matrix. Following [66, 67], this is done via the so-called Redheffer Star Product,

$$\begin{bmatrix} \hat{S}_{11}^{(AB)} & \hat{S}_{12}^{(AB)} \\ \hat{S}_{21}^{(AB)} & \hat{S}_{22}^{(AB)} \end{bmatrix} = \begin{bmatrix} \hat{S}_{11}^{(A)} & \hat{S}_{12}^{(A)} \\ \hat{S}_{21}^{(A)} & \hat{S}_{22}^{(A)} \end{bmatrix} \otimes \begin{bmatrix} \hat{S}_{11}^{(B)} & \hat{S}_{12}^{(B)} \\ \hat{S}_{21}^{(B)} & \hat{S}_{22}^{(B)} \end{bmatrix} \tag{3.70a}$$

$$\hat{S}_{11}^{(AB)} = \hat{S}_{11}^{(A)} + \hat{S}_{12}^{(A)} \left[\mathbb{1} - \hat{S}_{11}^{(B)} \hat{S}_{22}^{(A)} \right]^{-1} \hat{S}_{11}^{(B)} \hat{S}_{21}^{(A)} \tag{3.70b}$$

$$\hat{S}_{12}^{(AB)} = \hat{S}_{12}^{(A)} \left[\mathbb{1} - \hat{S}_{11}^{(B)} \hat{S}_{22}^{(A)} \right]^{-1} \hat{S}_{12}^{(B)} \tag{3.70c}$$

$$\hat{S}_{21}^{(AB)} = \hat{S}_{21}^{(B)} \left[\mathbb{1} - \hat{S}_{22}^{(A)} \hat{S}_{11}^{(B)} \right]^{-1} \hat{S}_{21}^{(A)} \tag{3.70d}$$

$$\hat{S}_{22}^{(AB)} = \hat{S}_{22}^{(B)} + \hat{S}_{21}^{(B)} \left[\mathbb{1} - \hat{S}_{22}^{(A)} \hat{S}_{11}^{(B)} \right]^{-1} \hat{S}_{22}^{(A)} \hat{S}_{12}^{(B)} \,. \tag{3.70e}$$

Eventually, one is left with the global S-matrix equation,

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} \end{bmatrix} = \hat{S} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \mathbf{0} \end{bmatrix} . \tag{3.71}$$

This equation relates the set of known quantities (right side) to the unknown quantities (left side) via the global S-matrix.

3.5.3 R-matrix approach

The R-matrix approach is very similar to the S-matrix approach. However, instead of the eigenmode coefficients, the R-matrix of the l^{th} layer relates the Fourier coefficient vectors of the electric and magnetic fields at the two layer interfaces to each other. The electric and magnetic fields appear on opposite sides of the equation sign. For completeness, the definition of the R-matrix is given at this point[10],

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_{\perp}^{(l)} \\ \widetilde{\boldsymbol{\mathcal{E}}}_{\perp}^{(l-1)} \end{bmatrix} = \hat{R}^{(l)} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}^{(l)} \\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp}^{(l-1)} \end{bmatrix} . \tag{3.72}$$

The solution principle of the R-matrix approach is very similar to the S-matrix approach and is not shown here, since is not needed in the context of this thesis. The interested reader may refer to [3, 10, 68].

3.5.4 Enhanced transmittance matrix approach

In the following, the so-called 'partial solution approach' of the ETMA as first proposed by [11] is presented, which means that it only applies for one-sided light incidence. The notation and derivation sticks very closely to that used in Chapter 5,

where the method is extended to two-sided light incidence based on the publication [A3]. Even though the name suggests otherwise, the algorithm provides the full solution. In the following, the basic principle is explained by means of Fig. 3.8. First, the tangential electric and magnetic field components, which are continuous in axial direction across the layer interfaces, are set equal. To this end, one calculates the electric and magnetic field mode coefficient vectors at the layer boundaries using Eq. (3.59), including the propagator shift by $d^{(l)}$, as mentioned in the derivation. At the leftmost interface, Fig. 3.8a, the continuity condition gives

$$\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(1)} & +\hat{W}^{(1)}\hat{X}^{(1)} \\ +\hat{V}^{(1)} & -\hat{V}^{(1)}\hat{X}^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(1)} \\ \mathbf{c}_{-}^{(1)} \end{bmatrix} . \tag{3.73}$$

At an intermediate interface between two layers $l-1$ and l, Fig. 3.8b, one gets

$$\begin{bmatrix} +\hat{W}^{(l-1)}\hat{X}^{(l-1)} & +\hat{W}^{(l-1)} \\ +\hat{V}^{(l-1)}\hat{X}^{(l-1)} & -\hat{V}^{(l-1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(l-1)} \\ \mathbf{c}_{-}^{(l-1)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(l)} & +\hat{W}^{(l)}\hat{X}^{(l)} \\ +\hat{V}^{(l)} & -\hat{V}^{(l)}\hat{X}^{(l)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(l)} \\ \mathbf{c}_{-}^{(l)} \end{bmatrix} . \tag{3.74}$$

At the rightmost interface, Fig. 3.8c, one obtains

$$\begin{bmatrix} +\hat{W}^{(L)}\hat{X}^{(L)} & +\hat{W}^{(L)} \\ +\hat{V}^{(L)}\hat{X}^{(L)} & -\hat{V}^{(L)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(L)} \\ \mathbf{c}_{-}^{(L)} \end{bmatrix} = \begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(L+1)} & -\hat{C}^{(L+1)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} \\ \mathbf{0} \end{bmatrix} . \tag{3.75}$$

Starting from the rightmost interface, Eq. (3.75) can be written as

$$\begin{bmatrix} +\hat{W}^{(L)}\hat{X}^{(L)} & +\hat{W}^{(L)} \\ +\hat{V}^{(L)}\hat{X}^{(L)} & -\hat{V}^{(L)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(L)} \\ \mathbf{c}_{-}^{(L)} \end{bmatrix} = \begin{bmatrix} \hat{f}^{(L+1)} \\ \hat{g}^{(L+1)} \end{bmatrix} \mathbf{c}_{+}^{(L+1)} , \tag{3.76}$$

where one identifies $\mathbf{c}_{+}^{(L+1)} = \widetilde{\boldsymbol{\mathcal{T}}}_{\perp}$ and $\hat{f}^{(L+1)} = \mathbb{1}$, $\hat{g}^{(L+1)} = \hat{C}^{(L+1)}$. By rearranging Eq. (3.76), one obtains

$$\begin{bmatrix} -\hat{W}^{(L)} & \hat{f}^{(L+1)} \\ \hat{V}^{(L)} & \hat{g}^{(L+1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{-}^{(L)} \\ \mathbf{c}_{+}^{(L+1)} \end{bmatrix} = \begin{bmatrix} \hat{W}^{(L)}\hat{X}^{(L)} \\ \hat{V}^{(L)}\hat{X}^{(L)} \end{bmatrix} \mathbf{c}_{+}^{(L)} , \tag{3.77}$$

which can be transformed to

$$\mathbf{c}_{-}^{(L)} = \hat{a}^{(L)}\mathbf{c}_{+}^{(L)} \tag{3.78a}$$

$$\mathbf{c}_{+}^{(L+1)} = \hat{b}^{(L)}\mathbf{c}_{+}^{(L)} \tag{3.78b}$$

with

$$\begin{bmatrix} \hat{a}^{(L)} \\ \hat{b}^{(L)} \end{bmatrix} = \begin{bmatrix} -\hat{W}^{(L)} & \hat{f}^{(L+1)} \\ \hat{V}^{(L)} & \hat{g}^{(L+1)} \end{bmatrix}^{-1} \begin{bmatrix} \hat{W}^{(L)}\hat{X}^{(L)} \\ \hat{V}^{(L)}\hat{X}^{(L)} \end{bmatrix} . \tag{3.79}$$

Eq. (3.78) is now being substituted back into Eq. (3.74) for $l = L$, which expresses the continuity conditions between the second last and last layer,

$$\begin{bmatrix} +\hat{W}^{(L-1)}\hat{X}^{(L)} & +\hat{W}^{(L-1)} \\ +\hat{V}^{(L-1)}\hat{X}^{(L)} & -\hat{V}^{(L-1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(L-1)} \\ \mathbf{c}_{-}^{(L-1)} \end{bmatrix} = \begin{bmatrix} \hat{f}^{(L)} \\ \hat{g}^{(L)} \end{bmatrix} \mathbf{c}_{+}^{(L)} , \tag{3.80}$$

where

$$\begin{bmatrix} \hat{f}^{(L)} \\ \hat{g}^{(L)} \end{bmatrix} = \begin{bmatrix} \hat{W}^{(L)} \left(\mathbb{1} + \hat{X}^{(L)}\hat{a}^{(L)} \right) \\ \hat{V}^{(L)} \left(\mathbb{1} - \hat{X}^{(L)}\hat{a}^{(L)} \right) \end{bmatrix} . \tag{3.81}$$

One observes that Eq. (3.80) is the same as Eq. (3.76), except that the index is decremented. The procedure [Eqs. (3.76-3.81)] is therefore repeated for the remaining layer interfaces in decreasing order until the leftmost interface, where one is left with

$$\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix} = \begin{bmatrix} \hat{f}^{(1)} \\ \hat{g}^{(1)} \end{bmatrix} \mathbf{c}_{\mathbf{1}}^{+} . \tag{3.82}$$

This can finally be solved for $\widetilde{\boldsymbol{\mathcal{R}}}_{\perp}$ and $\mathbf{c}_{\mathbf{1}}^{+}$ via

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \\ \mathbf{c}_{\mathbf{1}}^{+} \end{bmatrix} = \begin{bmatrix} -\mathbb{1} & \hat{f}^{(1)} \\ \hat{C}^{(0)} & \hat{g}^{(1)} \end{bmatrix}^{-1} \begin{bmatrix} \mathbb{1} \\ \hat{C}^{(0)} \end{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} . \tag{3.83}$$

Eq. (3.78) and its equivalents for the other layers can subsequently be used to determine the remaining unknowns $\mathbf{c}_{+}^{(l)}$, $\mathbf{c}_{-}^{(l)}$ and $\widetilde{\boldsymbol{\mathcal{T}}}_{\perp}$ iteratively.

3.6 Consistency requirements

Part of this section is based on [A3].

In Sect. 3.1.3, two different methods for realizing a tilt of the incident light modes are presented. In method 1, the amplitude of the central mode is set to one and a shift of $k_{x0} = nk_0 \sin\theta \cos\varphi$ and $k_{y0} = nk_0 \sin\theta \sin\varphi$ is applied to the lateral components of the k-vectors. In method 2, a higher order Fourier mode is excited in the incident mode coefficient vector, while the k-vector shift is omitted. If the angles corresponding to the k-vector shift in method 1 are equal to the angles of the excited higher order grating mode in method 2,

$$\begin{bmatrix} \sqrt{\epsilon^{(0)}\mu^{(0)}}\, sin\theta\, cos\varphi \\ \sqrt{\epsilon^{(0)}\mu^{(0)}}\, sin\theta\, sin\varphi \end{bmatrix} = \begin{bmatrix} m\lambda/P_x \\ n\lambda/P_y \end{bmatrix} , \tag{3.84}$$

the identical physical situation is described[6]. One naturally expects that the result is independent of the applied method. The authors of [17] demonstrated that this is not automatically the case. The explanation is that the mode truncation in the two scenarios is different. In method 1, the mode truncation is asymmetric, whereas method 2 results in a symmetric mode truncation. In order to avoid inconsistencies between the two methods, the suggestion in [17] was requiring that the resulting set of differential equations must be identical in both cases (*'angular shift-invariance'*[17]). In this very publication, it was shown that, for mode-truncated systems, this requirement is fulfilled under the following prerequisites (cf. also Fig. 3.10).

- The electromagnetic field's Fourier coefficient vectors, $\widetilde{\boldsymbol{\mathcal{E}}}_{\perp}$, must be truncated symmetrically around zero. Outside the truncation boundaries, the Fourier mode coefficients are assumed to be continued in a periodic manner.
- Outside the truncation boundaries, the lateral wave vector components, $k_{x/y;mn}$, are assumed to be continued periodically. This way, it is ensured that the highest mode $+M/N$ is always followed by the lowest mode $-M/N$ (*'centered bandwidth limitation'*[17]).
- The electric permittivity $\hat{\epsilon}\,(x,y)$ and the magnetic permeability $\hat{\mu}\,(x,y)$ are assumed to be band limited to the chosen mode truncation boundaries. The corresponding Toeplitz matrices, $[\![\epsilon]\!]$ and $[\![\mu]\!]$, are continued in a circulant fashion in the upper right and lower left corners.

It should be noted that, for the circulant definition of the Toeplitz matrices, the application of the direct rule and the inverse rule as described in Sect. 3.2 result in mathematically identical matrices, i.e., $[\![\epsilon]\!] = [\![1/\epsilon]\!]^{-1}$. Consequently, Li's factorization rules are automatically fulfilled.

Recently, there has been some debate about Li's factorization rules. It was shown that, due to the non-circulant definition of the Toeplitz matrices, the energy balance between the external fields and the local absorption is not consistent[69]. In the mentioned publication, it is also demonstrated that this problem is solved for the circulant definition of the Toeplitz matrices. In combination with [17, 63], this result suggests that Li's factorization rules merely counteract the symptoms of mode truncation rather than solving the true underlying problem, i.e., the inconsistent occupation of the Toeplitz matrices.

3.7 Numerical complexity and memory requirement

The numerical complexity of the standard RCWA algorithm is dominated by

[6] If the tilt according to method 1 is not equal to a grating mode angle, no corresponding mode can be found with method 2. The reason is that, in this case, the incident electromagnetic field does not feature the same lateral periodicity as the grating.

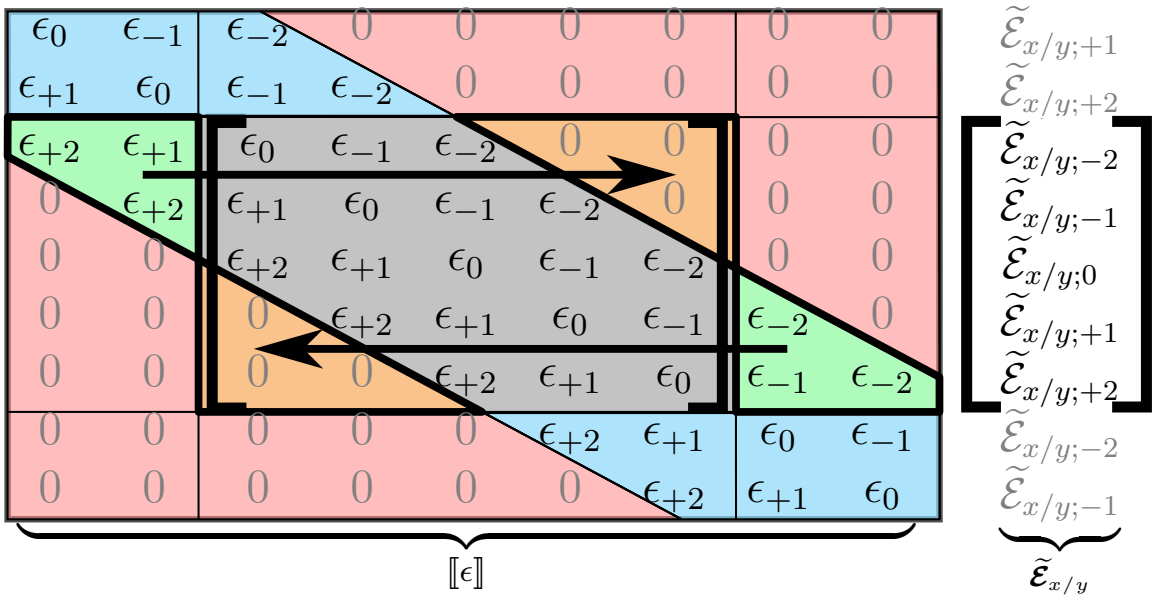


Figure 3.10: The electromagnetic field Fourier coefficient vector is continued periodically outside the symmetric truncation limits. Furthermore, the mode truncation corresponds to an effective band limitation of $\epsilon(x, y)$ or $\mu(x, y)$, respectively. Therefore, the corners of the truncated Toeplitz matrix are completed in a cyclic fashion - it then becomes a circulant matrix.

1. the diagonalization of the coupling matrix in the process of solving the coupled differential equation, Eqs. (3.53, 3.61), and

2. the matrix inversions in the process of enforcing the boundary conditions, Eqs. (3.69, 3.79).

These matrices have the dimensions $2\bar{N} \times 2\bar{N}$ or $4\bar{N} \times 4\bar{N}$, respectively. The numerical complexity of both the diagonalization and the inversion are of the order $\mathcal{O}(\bar{N}^3)$. The memory requirement is given as the number of elements in the matrix, which is of the order $\mathcal{O}(\bar{N}^2)$.

In consequence, only very small problem sizes can actually be solved with the RCWA. To make an example, consider a grating, which should be sampled with $M_0 = N_0 = 50$, i.e. $\bar{N} = 101^2 \approx 10^4$ modes. Consequently, the coupling matrix has $(4\bar{N})^2 \approx 1.6 \cdot 10^9$ complex elements. Assuming a double-precision arithmetic, the memory requirement is already 24 GB. On an off-the-shelf hexacore server with 3.70 GHz clock rate and 64 GB RAM used for this work, the calculation time of this problem is approximately 8 h.

3.8 Simulation of a SNOM measurement

In this section, a simulation from the early beginnings of my PhD studies is shown, which was published in [A1] and nicely illustrates a number of problems inherent to the current RCWA. In the following chapters, the majority of these problems will be solved.

SNOM is a method, which acquires evanescent field information, thus overcoming the Abbe resolution limit. It is a technique used for sub-λ imaging at optical frequencies in experimental physics[20]. SNOM is generally thought to be understood well, since it works in practice and, experimentally, one finds that the resolution is in the order of a few tens of nm[70]. In the past, there were several attempts to simulate the coupling of light into a SNOM fiber tip, using for example the BPM[71] or the FDTD[72]. Here, the recently developed 'localized input field RCWA' (LIF-RCWA)[17] is applied to simulate SNOM measurements. Two different simulations are shown. First, in Sect. 3.8.1, the scanning of a narrow focal spot is considered and the influence of different parameters on the resolution of SNOM images is investigated. Second, Sect. 3.8.2 studies how the presence of the SNOM-probe changes the spatial electromagnetic field during a measurement. To this end, the undisturbed electric field intensity is compared to the field distribution altered by the presence of the SNOM probe.

3.8.1 Scanning a narrow focal spot

A two-dinensional SNOM-measurement is simulated, where a narrow focal spot serves as a sample. Fig. 3.11 illustrates the electric permittivity distribution of the fiber tip, where d_{ap} denotes the aperture diameter, α corresponds to the apex angle of the fiber tip, Δz denotes the vertical distance between sample and aperture and Δx is the lateral offset of the sample with respect to aperture. In the following, the focus is moved laterally across the aperture and an RCWA calculation is conducted at each Δx. One obtains an intensity profile by plotting the amount of power coupled into the fiber core versus Δx. In all simulations, a wavelength of $\lambda = 850\,\mathrm{nm}$, a calculation period of $P = 2\,\mu\mathrm{m}$ and a total mode count of $M = 101$ are assumed. In order to obtain a narrow focus in front of the SNOM-tip, the LIF-RCWA[17] is used to excite the 81 central modes of the incident field in TM-polarization. This results in a peak width of approximately 40 nm. The power density distribution of one simulation is shown in Fig. 3.12. It can be clearly seen how light couples through the tip into the fiber. Also, one can see that the amount of energy coupled into the fiber is small compared to the incident power. In the following, the influence of the parameters d_{ap}, α, and Δz on the width of the resulting intensity profile is summarized. First, one observes that the width of the recorded peak is approximately proportional to the width of the aperture diameter d_{ap}. The measured Gaussian width ranges from $\sigma = 24\,\mathrm{nm}$ at $d_{\mathrm{ap}} = 40\,\mathrm{nm}$ to $\sigma = 84\,\mathrm{nm}$ at $d_{\mathrm{ap}} = 300\,\mathrm{nm}$. As

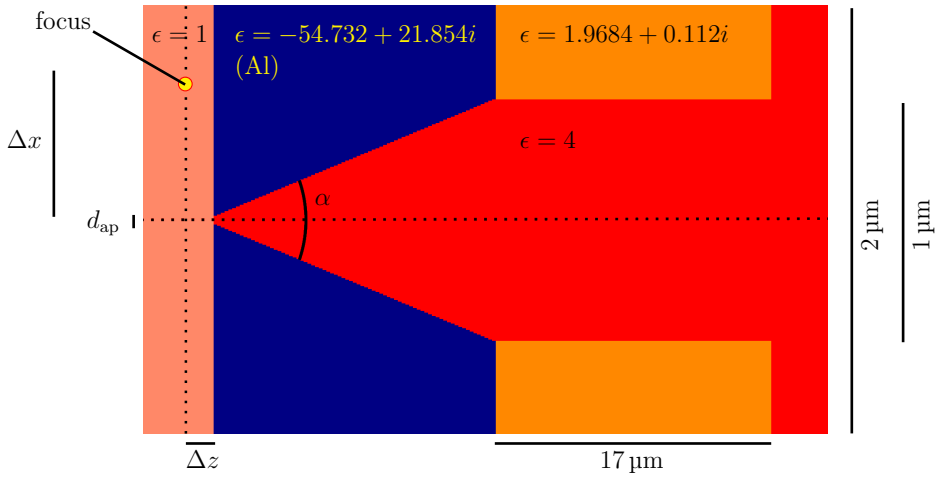


Figure 3.11: Permittivity distribution of the fiber tip. The transmitted region (on the right side) is homogeneous and index-matched with the fiber core to reduce unwanted reflections. The fiber cladding is slightly absorbing to remove non-guided modes from the fiber.

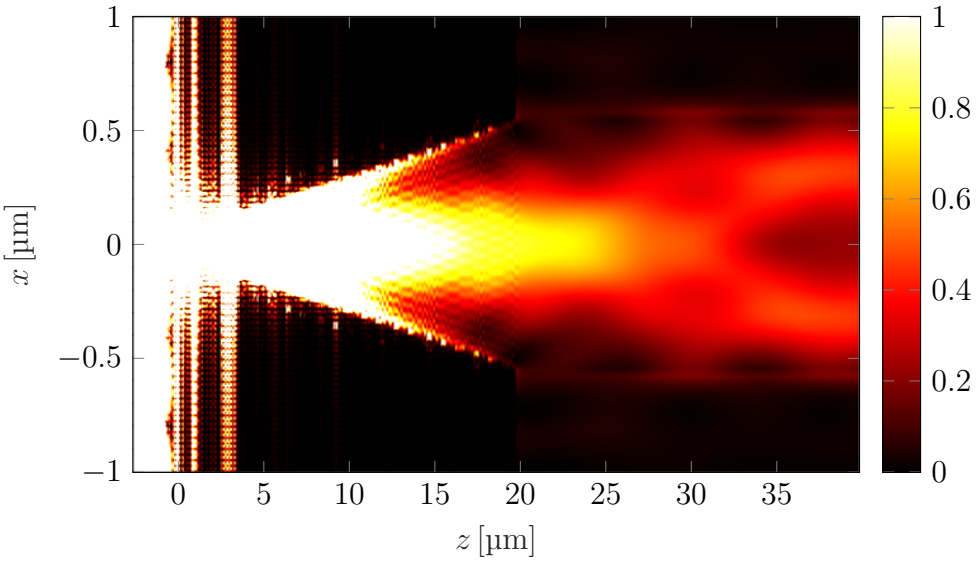


Figure 3.12: Clipped power density distribution inside a SNOM-probe for $d_{\text{ap}} = 40\,\text{nm}$, $\alpha = 25°$, $\Delta x = 0\,\text{nm}$, and $\Delta z = 30\,\text{nm}$.

expected, the amount of intensity coupled into the fiber core increases significantly for larger aperture diameters. Second, varying the apex angle α does not significantly change the width of the recorded peak. However, the amount of intensity coupled into the fiber increases significantly for larger α. Third, one observes that the measured peak width grows with increasing Δz. The Gaussian width ranges from $\sigma = 24\,\text{nm}$ at $\Delta z = 30\,\text{nm}$ to $\sigma = 143\,\text{nm}$ at $\Delta z = 200\,\text{nm}$. This observation can be explained by the characteristic exponential decay of the evanescent modes, which is where the sub-λ spatial characteristics are encoded. Due to the exponential decay of their amplitudes, the propagation distance from the focus to the fiber acts as a low-pass filter.

3.8.2 Scanning a grating structure

In the following, a SNOM-measurement of a back-illuminated grating structure is simulated in order to investigate the influence the SNOM fiber tip on the measurement. Fig. 3.13 illustrates the electric permittivity distribution. Furthermore, a fixed $d_{\text{ap}} = 40\,\text{nm}$ and $\Delta z = 30\,\text{nm}$ are assumed. All remaining parameters are chosen in the same way as in Sect. 3.8.1. In Fig. 3.14, the power density distribution

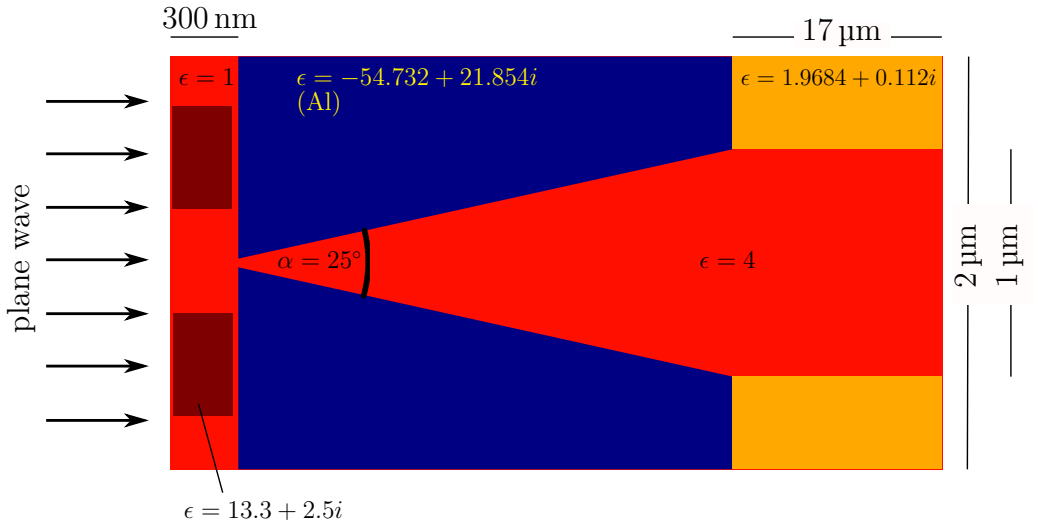


Figure 3.13: Permittivity distribution of the grating structure and the fiber tip. In lateral direction, the grating blocks on the left are 500 nm wide each.

is shown without (left) and with (right) the fiber tip inserted for the tip position $\Delta x = 325\,\text{nm}$. It can be seen that the presence of the tip severely disturbs the power density distribution.

One might think that the amount of light coupled into the fiber is proportional to the integrated disturbed power density distribution in front of the SNOM aperture. Fig. 3.15 (left) shows a lateral line scan of the integrated undisturbed power density distribution. In contrast, Fig. 3.15 (right) shows the amount of power, which is

actually coupled into the fiber core as a function of Δx. It can be clearly seen that both curves are not in agreement. The width and position of the 'step' in Fig. 3.15 (right) are determined by a least squares fit of a smoothened step function. One finds that the width of the step, $\sigma = 43\,\text{nm} \approx \lambda/20$, is small, but its position is shifted by approximately $\delta x \approx 82\,\text{nm}$ with respect to the original step position in the grating.

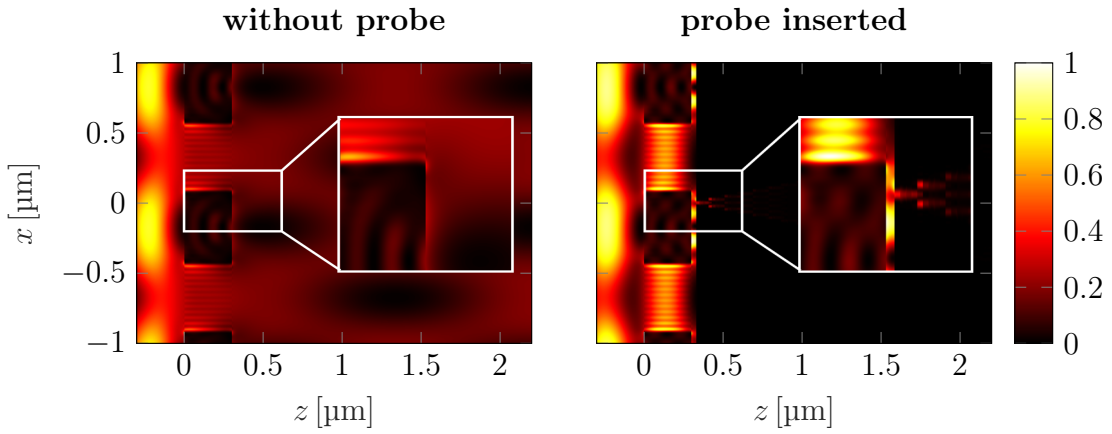


Figure 3.14: Undisturbed power density distribution with only the grating (left, arbitrary intensity scale) and power density distribution with fiber probe inserted (right, same intensity scale as in left image).

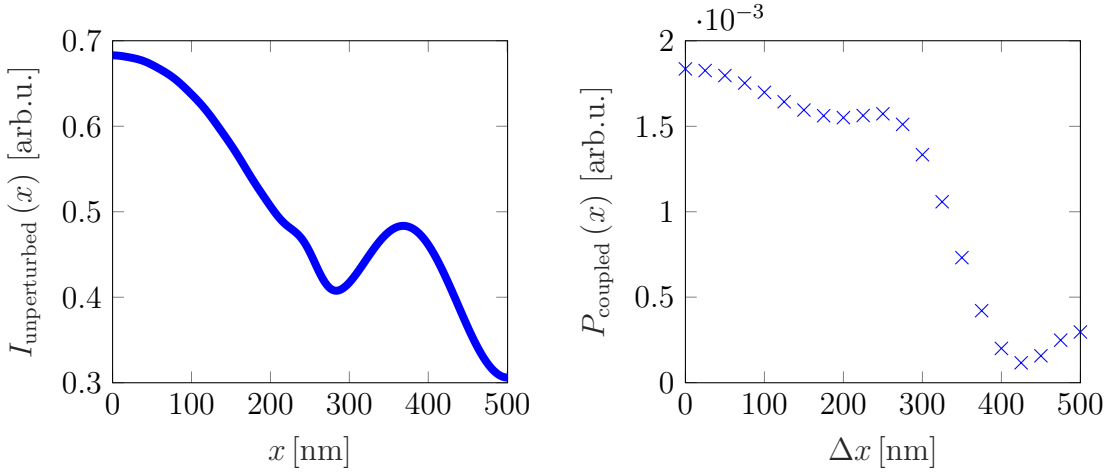


Figure 3.15: Line scan of the unperturbed power density distribution at $\Delta z = 30\,\text{nm}$ as a function of the lateral position x (left, arbitrary intensity scale), and amount of power coupled into the fiber as a function of the SNOM probe position Δx (right, same scaling as in left image).

In summary, one finds that the aperture size, the apex angle of the fiber tip, and the distance between sample and tip play a critical role when it comes to the

resolving power of SNOM measurements. Furthermore, the simulations show that retroactive effects of the fiber tip on the power density distribution are important and, consequently, the intensity profile of the measurement may deviate from the true underlying structure of the sample.

3.8.3 Analysis and discussion

Since the first publication in the 1980s, many useful extensions, generalizations and stabilization mechanisms were introduced to enhance the original RCWA. However, situations and parameter ranges still exist, which are currently inaccessible, be it due to time or memory constraints, or for lack of algorithmic development. Consider the example simulation of the SNOM fiber tip from the previous section.

1. The incident and transmitted regions have to be homogeneous in the RCWA. Therefore, the light coupled into the SNOM fiber eventually radiates into the homogeneous transmitted region, which is index-matched to the fiber core in order to minimize reflections. Clearly, it would be advantageous if the transmitted region could be defined as a structured region, so that a very long fiber can be modeled.

2. Light is only allowed to be incident from one direction in the RCWA. For an inversion of the light path, a backside illumination would be the easiest and most direct solution. However, in the standard framework, this possibility does not exist. The only way to emulate a backside illumination consists in mirroring the entire refractive index distribution. This procedure is, however, prone to implementation errors. Furthermore, it does not allow a simultaneous two-sided illumination.

3. The resolution of RCWA simulations is restricted to very small problem sizes due to time and memory constraints. In most cases, only 2d-calculations are possible in order to retain a sensible sampling.

In the following chapters, the three mentioned problems are considered. In Chapter 4, the RCWA is generalized as to allowing also structured incident and transmitted regions, where the incident light is described as a superposition of eigenmodes. In Chapter 5, the RCWA framework is extended for coherent bidirectional light incidence. In Chapter 6, the limitation to very small problem sizes is removed for the case of certain thin optical elements, which is achieved by an iterative solution approach.

4 | Structured incident and transmitted regions

This chapter is partly based on the publication [A2][7].

4.1 Motivation and structure

The standard RCWA framework is limited to the simulation of grating layer stacks, which are embedded between two homogeneous regions. However, especially in cases where the refractive index structures are constant over a wide axial range, the influence of distant interfaces usually needs to be neglected. This is, for instance, the case for the simulation of long waveguides, fibers or integrated optical components, i.e., when the spatial coherence length is smaller than the axial extend of the object. Therefore, in this chapter, the common RCWA is extended to the treatment of structured incident and transmitted regions. To my knowledge, the very basic idea was actually first brought up by Lalanne[73], but an adapted procedure was indeed never formulated. Furthermore, at this point it was not recognized that, in order to be able to set sensible boundary conditions at the outer interfaces, it is essential to unambiguously determine the propagation direction of the eigenmodes in the incident and transmitted regions. While this is fairly simple for homogeneous regions and Fourier modes, it becomes nontrivial for structured layers. Therefore, it is investigated how the propagation direction of individual eigenmodes is determined unambiguously. It is shown that the standard RCWA algorithm with its homogeneous incident and transmitted regions is a special case of this new generalized description. Throughout this chapter, linear isotropic materials are assumed.

This chapter is organized as follows. In Sect. 4.2, the modified boundary value problem for structured incident and transmitted regions is formulated. Sect. 4.3 compares the two cases, with homogeneous and with structured incident and transmitted regions, and shows that the standard RCWA is just a special case of the generalized structured treatment. Furthermore, it is demonstrated how the transition from structured to homogeneous regions is expressed in the boundary conditions. In Sect. 4.4, the energy flow properties of structured layer eigenmodes are

[7] © 2016 IEEE.

derived. Furthermore, it is shown how the propagation direction of an eigenmodes is unambiguously determined from the Ponynting vector. The latter is necessary to set the boundary conditions in a sensible manner. In Sect. 4.5, several simulation examples are shown in order to verify the findings. First, the numerical eigenmodes are compared to the analytic eigenmodes of a planar dielectric waveguide. Second, the evanescent coupling between two closely adjacent waveguides is simulated and compared to the results of CMT in the case of weak and strong coupling. Third, the mode coupling from vacuum into a long graded index (GRIN) waveguide with a sech-profile is simulated and compared to the analytic solution.

4.2 Reformulation of the boundary value problem

In the standard implementation of the RCWA, a periodic grating structure is situated between the incident and transmitted regions with known and constant indices of refraction. A superposition of Fourier modes is assumed to be incident from $z = -\infty$, no light is incident from $z = +\infty$ (cf. Chapter 3). In the following, the restriction to homogeneous incident and transmitted regions is removed. As illustrated in Fig. 4.1, these regions should now be treated as structured media instead. In order to

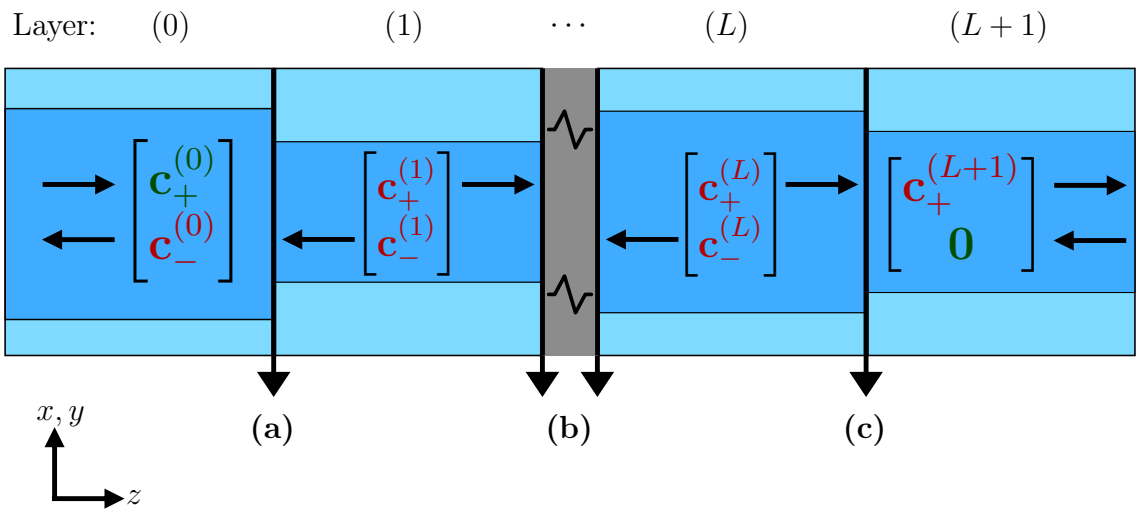


Figure 4.1: Boundary conditions for the RCWA with structured incident and transmitted regions. The eigenmode coefficients of the incident field, $\mathbf{c}_+^{(0)}$, are given. No light is assumed to be incident from the right half space. The reflected and transmitted eigenmode coefficients, $\mathbf{c}_-^{(0)}$ and $\mathbf{c}_+^{(L+1)}$, are unknown, so are the eigenmode coefficients $\mathbf{c}_\pm^{(l)}$ within the layers.

do so, it is necessary to conduct an eigenmode decomposition as it is usually done for the grating layers (cf. Sect. 3.4.1). Furthermore, now, a linear superposition of *eigen*modes is assumed to be incident from $z = -\infty$. At this point, no light is

assumed to be incident from $z = +\infty$. Nevertheless, as shown in Chapter 5, bidirectional light incidence is fully compatible with this approach. So far, no definition for a 'forward' and 'backward' propagating eigenmode was given. At this point, it is simply assumed that this distinction can be made and that $\mathbf{c}_+$ ($\mathbf{c}_-$) correspond to the 'forward' ('backward') propagating eigenmodes. A full definition is given later in Sect. 4.4.4. In order to set the boundary conditions, the tangential electric and magnetic field components must be matched at the layer interfaces similar to Sect. 3.5. For the interface between the incident region and the first layer, Fig. 4.1a, this reads

$$\begin{bmatrix} +\hat{W}^{(0)} & +\hat{W}^{(0)} \\ +\hat{V}^{(0)} & -\hat{V}^{(0)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(0)} \\ \mathbf{c}_-^{(0)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(1)} & +\hat{W}^{(1)}\hat{X}^{(1)} \\ +\hat{V}^{(1)} & -\hat{V}^{(1)}\hat{X}^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(1)} \\ \mathbf{c}_-^{(1)} \end{bmatrix} . \tag{4.1}$$

At an intermediate interface between two layers $l-1$ and l, Fig. 4.1b, one gets

$$\begin{bmatrix} +\hat{W}^{(l-1)}\hat{X}^{(l-1)} & +\hat{W}^{(l-1)} \\ +\hat{V}^{(l-1)}\hat{X}^{(l-1)} & -\hat{V}^{(l-1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(l-1)} \\ \mathbf{c}_-^{(l-1)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(l)} & +\hat{W}^{(l)}\hat{X}^{(l)} \\ +\hat{V}^{(l)} & -\hat{V}^{(l)}\hat{X}^{(l)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(l)} \\ \mathbf{c}_-^{(l)} \end{bmatrix} . \tag{4.2}$$

At the rightmost interface, Fig. 4.1c, one obtains

$$\begin{bmatrix} +\hat{W}^{(L)}\hat{X}^{(L)} & +\hat{W}^{(L)} \\ +\hat{V}^{(L)}\hat{X}^{(L)} & -\hat{V}^{(L)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(L)} \\ \mathbf{c}_-^{(L)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(L+1)} & +\hat{W}^{(L+1)} \\ +\hat{V}^{(L+1)} & -\hat{V}^{(L+1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(L+1)} \\ \mathbf{c}_-^{(L+1)} \end{bmatrix} . \tag{4.3}$$

These equations are the direct equivalent to the continuity conditions given in the derivation of the ETMA, Eqs. (3.73-3.75). Based on Eqs. (4.1-4.3), the ETMA is, indeed, derived in just the same way as done in Sect. 3.5.4, so that the 'new' algorithm now solves for the eigenmode coefficients of a structured layer. The interested reader finds a complete reformulation of the ETMA *including* the implementation of bidirectional illumination in Chapter 5. There, it is also shown that the S-matrix approach is also applicable.

4.3 Structured vs. homogeneous regions

In the derivation of the RCWA, the general solution of Maxwell's equations for x-y-periodic structured layers of linear isotropic materials was derived. On the other hand, it is known that the solutions of Maxwell's equations in homogeneous linear isotropic materials are plane waves. Consequently, one expects that, in the limit of homogeneous layers, there is a smooth transition of the very general RCWA solution into a superposition of plane waves. In order to investigate the connection, consider again Eq. (3.51),

$$\frac{\partial^2}{\partial z^2}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) = -k_0^2 \hat{M}_{12}^{(l)} \hat{M}_{21}^{(l)} \widetilde{\boldsymbol{\mathcal{E}}}_\perp(z) . \tag{4.4}$$

In the special case of a homogeneous layer l, the Toeplitz matrices built from the electric permittivity and magnetic permeability become diagonal. It is straightforward to show that, in consequence, also the coupling matrix in Eq. (4.4) becomes diagonal,

$$\hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)} \to \begin{bmatrix} \epsilon^{(l)}\mu^{(l)}\mathbb{1} - \hat{K}_x^2 - \hat{K}_y^2 & 0 \\ 0 & \epsilon^{(l)}\mu^{(l)}\mathbb{1} - \hat{K}_x^2 - \hat{K}_y^2 \end{bmatrix}, \tag{4.5}$$

i.e., the system of differential equations is already fully decoupled. Therefore, a diagonalization as conducted in Sect. 3.4.1 is not necessary. Instead, one directly onbtains $\hat{W}^{(l)} \to \mathbb{1}$ as the eigenvector matrix, and $\hat{\Lambda}^{(l)} \to \hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)}$ as the diagonal eigenvalue matrix. The diagonal entries of $\hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)}$ can, indeed, be identified as the z-components of the wave vectors in homogeneous media. The matrix $\hat{V}^{(l)}$ turns into the matrix $\hat{C}^{(l)}$, which is the homogeneous space conversion matrix from the electric field Fourier mode vector into the magnetic field Fourier mode vector,

$$\widetilde{\boldsymbol{\mathcal{H}}}_\perp = \hat{C}^{(l)}\widetilde{\boldsymbol{\mathcal{E}}}_\perp\,, \tag{4.6}$$

as previously defined in Eq. (3.67). In summary, one obtains the relations

$$\hat{W}^{(l)} \to \mathbb{1} \tag{4.7a}$$

$$\hat{\Lambda}^{(l)} \to \hat{M}_{12}^{(l)}\hat{M}_{21}^{(l)} = \hat{K}_z^{(l)} \tag{4.7b}$$

$$\hat{V}^{(l)} \to \hat{C}^{(l)}\,. \tag{4.7c}$$

Hence, in the limit of homogeneous layers, the RCWA eigenmode solution turns into a simple superposition of plane waves. From this follows that, indeed, the eigenmodes of homogeneous regions are the Fourier modes, i.e., plane waves!

Now, consider again the boundary conditions at the layer interfaces, Eqs. (4.1-4.3). Note that, by making the transition to homogeneous layers, as shown above, one directly retrieves Eqs. (3.73-3.75) from the standard ETMA. Hence, the generalization presented in Sect. 4.2 is consistent with the existing framework. Therefore, also homogeneous layers can be treated with this very general approach.

4.4 Energy flow direction and Poynting vector

As presented in the previous section, it is assumed that a given superposition of eigenmodes of the incident region illuminates the grating from $z = -\infty$. On the other hand, an unknown 'transmitted' and 'reflected' superposition of eigenmodes propagates away from the grating in the structured incident and transmitted regions towards $z = \pm\infty$. The above statements contains the implicit assumption that the propagation direction of the eigenmodes is already known. In fact, this is a fairly simple task for Fourier modes, i.e., in the case of the standard RCWA with homo-

geneous incident and transmitted regions, where the z-component of the Poynting vector is usually taken as a measure. However, the definition of the propagation direction of a *general* eigenmode of a structured layer is not anymore straightforward. In the following, it is derived how the propagation direction of a general eigenmode can be unambiguously determined.

4.4.1 Plane waves in homogeneous media

First, consider a homogeneous region, for which the solution of Maxwell's equations is known to be a superposition of plane waves,

$$\boldsymbol{\mathcal{E}}(\mathbf{r},\omega) = \sum_q \widetilde{\boldsymbol{\mathcal{E}}}_q \, e^{+ik_{z;q}z} \, e^{ik_{x;q}x+ik_{y;q}y} \tag{4.8a}$$

$$\begin{aligned}\boldsymbol{\mathcal{H}}(\mathbf{r},\omega) &= -\frac{i}{\mu k_0}\nabla\times\boldsymbol{\mathcal{E}}(\mathbf{r},\omega)\\ &= \frac{1}{\mu k_0}\sum_q \mathbf{k}_q\times\widetilde{\boldsymbol{\mathcal{E}}}_q \, e^{+ik_{z;q}z} \, e^{ik_{x;q}x+ik_{y;q}y}\,.\end{aligned} \tag{4.8b}$$

In general, the time-averaged Poynting vector is calculated from the electric and magnetic fields via

$$\begin{aligned}\langle\mathbf{S}\rangle_t(\mathbf{r}) &= \frac{1}{2}\Re\left[\boldsymbol{\mathcal{E}}(\mathbf{r},\omega)\times\boldsymbol{\mathcal{H}}^*(\mathbf{r},\omega)\right]\\ &= \frac{1}{2}\sum_{qq'}\Re\left[\frac{1}{\mu^* k_0}\widetilde{\boldsymbol{\mathcal{E}}}_q\times\left(\mathbf{k}^*_{q'}\times\boldsymbol{\mathcal{E}}^*_{q'}\right) \, e^{i\left(k_{x;q}-k_{x;q'}\right)x+i\left(k_{y;q}-k_{y;q'}\right)y+i\left(k_{z;q}-k^*_{z;q'}\right)z}\right]\end{aligned} \tag{4.9}$$

For an isolated propagating plane wave in a non-absorbing medium, i.e., $\widetilde{\boldsymbol{\mathcal{E}}}_q = \widetilde{\boldsymbol{\mathcal{E}}}_Q\,\delta_{qQ}$ and $\Im[\epsilon] = \Im[\mu] = \Im[k_z] = 0$, the Poynting vector simplifies,

$$\langle\mathbf{S}\rangle_t^{[Q]}(\mathbf{r}) = \langle\mathbf{S}\rangle_t^{[Q]} = \frac{1}{2}\left|\boldsymbol{\mathcal{E}}_Q\right|^2\frac{\mathbf{k}_Q}{\mu k_0}\,. \tag{4.10}$$

In this case, the energy flow is in the direction of $\mathbf{k}_Q$ and independent of $\mathbf{r}$. On the other hand, it can be shown that for evanescent waves, energy flows only in x-y-direction, and not in axial direction. Another interesting property of plane waves is also that

$$\int\langle\mathbf{S}\rangle_t(\mathbf{r})\,\mathrm{d}x\,\mathrm{d}y = \sum_q\langle\mathbf{S}\rangle_t^{[q]}\,, \tag{4.11}$$

i.e., the x-y-integrated energy flow of the superposition of plane waves is equal to the sum of the plane waves' individual energy flows. Not that, therefore, it is sensible to define the diffraction efficiencies of the Fourier modes as $\mathrm{DE}_q \propto \langle\mathbf{S}\rangle_t^{[q]}$[5]. For

a linear superposition of forward *and* backward propagating Fourier modes, mixed terms appear in the expression for the Poynting vector. These terms are, for instance, responsible for the finite energy flow in the tunneling region of an evanescent coupler. Therefore, even though an isolated evanescent wave does not carry energy in axial direction, for a superposition of a forward and backward propagating evanescent waves this is very well possible.

4.4.2 Eigenmodes in structured media

Now consider the more general case of a structured layer. From Eqs. (3.54, 3.57) one obtains

$$\widetilde{\mathcal{E}}_{\perp;q}(z) = \sum_{q'} W_{qq'} c_{+;q'} \, e^{+ik_0\sqrt{\lambda_{q'}}z} \tag{4.12a}$$

$$\widetilde{\mathcal{H}}_{\perp;q}(z) = \sum_{q'} V_{qq'} \; c_{+;q'} \, e^{+ik_0\sqrt{\lambda_{q'}}z} \,, \tag{4.12b}$$

where only the forward propagating eigenmodes are considered, i.e., $c_{-;q} = 0$. In order to calculate the energy flow of an isolated eigenmode, one assigns $c_{+;q} = c_{+;Q}\,\delta_{qQ}$, where an arbitrary eigenmode can be chosen by setting Q to a constant index. By separating the matrices $\hat{W}$ and $\hat{V}$ into an upper and lower half, $\hat{W} = \begin{bmatrix}\hat{W}_x & \hat{W}_y\end{bmatrix}^T$ and $\hat{V} = \begin{bmatrix}\hat{V}_x & \hat{V}_y\end{bmatrix}^T$, one obtains

$$\mathcal{E}^{[Q]}_{x;q}(\mathbf{r},\omega) = \sum_{q} W_{x;qQ} \; c_{+;Q} \; e^{+ik_0\sqrt{\lambda_Q}z} \; e^{ik_{x;q}x + ik_{y;q}y} \tag{4.13a}$$

$$\mathcal{E}^{[Q]}_{y;q}(\mathbf{r},\omega) = \sum_{q} W_{y;qQ} \; c_{+;Q} \; e^{+ik_0\sqrt{\lambda_Q}z} \; e^{ik_{x;q}x + ik_{y;q}y} \tag{4.13b}$$

$$\mathcal{H}^{[Q]}_{x;q}(\mathbf{r},\omega) = \sum_{q} V_{x;qQ} \; c_{+;Q} \; e^{+ik_0\sqrt{\lambda_Q}z} \; e^{ik_{x;q}x + ik_{y;q}y} \tag{4.13c}$$

$$\mathcal{H}^{[Q]}_{y;q}(\mathbf{r},\omega) = \sum_{q} V_{y;qQ} \; c_{+;Q} \; e^{+ik_0\sqrt{\lambda_Q}z} \; e^{ik_{x;q}x + ik_{y;q}y} \,, \tag{4.13d}$$

where $q = mN + n$ is the concatenated index, which runs simultaneously over both Fourier dimensions k_x and k_y as described in Sect. 3.2.3. The time-averaged Poynting vector is calculated from the electric and magnetic fields via

$$\langle \mathbf{S} \rangle_t (\mathbf{r}) = \frac{1}{2} \Re \left[\boldsymbol{\mathcal{E}}(\mathbf{r},\omega) \times \boldsymbol{\mathcal{H}}^*(\mathbf{r},\omega) \right] . \tag{4.14}$$

By inserting Eqs. (4.13) into Eq. (4.14), one obtains

$$\langle S_z \rangle_t^{[Q]}(\mathbf{r}) = \frac{1}{2}\, e^{-2k_0 \Im\left[\sqrt{\lambda_Q}\right] z} \sum_{qq'} \Re\left[\left(W_{x;qQ} V^*_{y;q'Q} - W_{y;qQ} V^*_{x;q'Q}\right) \times e^{i\left(k_{x;q}-k_{x;q'}\right)x + i\left(k_{y;q}-k_{y;q'}\right)y}\right]. \tag{4.15}$$

The z-component of the Poynting vector is only z-dependent via the imaginary part of $\sqrt{\lambda_Q}$, which characterizes the absorption strength of the medium. For non-absorbing media, the energy flow $\langle S_z \rangle_t^{[Q]}(\mathbf{r})$ is constant in z. If absorption (amplification) is present, the z-dependence is an exponential decrease (increase), depending on the sign of $\Im\left[\sqrt{\lambda_Q}\right]$. Note that, in the lateral coordinates x and y, the energy flow of an eigenmode is not necessarily constant as it is the case for plane waves. As an example, Fig. 4.2 shows the numerically calculated Poynting vector of a planar slab waveguide eigenmode.

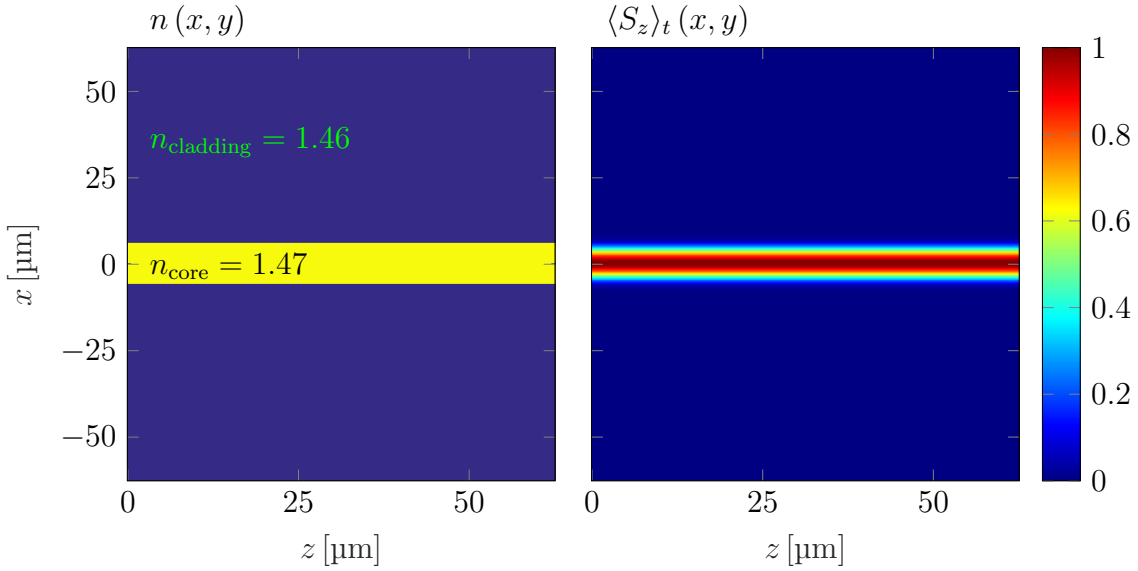


Figure 4.2: Numerically calculated z-component of the Poynting vector (right image, normalized to a maximum amplitude of 1) for the lowest order eigenmode of the slab waveguide shown in the left image (core diameter $d_{\text{core}} = 12\,\mu\text{m}$). RCWA parameters: $P_x = 125\,\mu\text{m}$, $M_0 = 200$, $\lambda = 1.55\,\mu\text{m}$.

4.4.3 Analytic Poynting vector for the 2d-case

In the two-dimensional case, i.e., if both the grating and the fields are constant in y, the analytic Poynting vector can be determined. For $k_{y;mn} = 0$, the eigenvalue

decomposition splits into two parts,

$$\begin{aligned}\hat{M}_{12}\hat{M}_{21} &= \begin{bmatrix} -\hat{K}_x\lceil\epsilon\rceil^{-1}\hat{K}_x\lceil\epsilon^{-1}\rceil + \lceil\mu\rceil\lceil\epsilon^{-1}\rceil & \mathbb{0} \\ \mathbb{0} & -\lceil\mu^{-1}\rceil\hat{K}_x\lceil\mu\rceil^{-1}\hat{K}_x + \lceil\mu^{-1}\rceil\lceil\epsilon\rceil \end{bmatrix} \\ &= \begin{bmatrix} \hat{W}_{11} & \mathbb{0} \\ \mathbb{0} & \hat{W}_{22} \end{bmatrix} \begin{bmatrix} \hat{\Lambda}_1 & \mathbb{0} \\ \mathbb{0} & \hat{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \hat{W}_{11}^{-1} & \mathbb{0} \\ \mathbb{0} & \hat{W}_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \hat{W}_{11}\hat{\Lambda}_1\hat{W}_{11}^{-1} & \mathbb{0} \\ \mathbb{0} & \hat{W}_{22}\hat{\Lambda}_2\hat{W}_{22}^{-1} \end{bmatrix}. \end{aligned} \tag{4.16}$$

Furthermore, one obtains

$$\begin{aligned}\hat{V} &= \hat{M}_{21}\hat{W}\hat{\Lambda}^{-1/2} \\ &= \begin{bmatrix} \mathbb{0} & -\lceil\mu^{-1}\rceil\hat{W}_{22}\hat{\Lambda}_2^{1/2} \\ \lceil\epsilon^{-1}\rceil^{-1}\hat{W}_{11}\hat{\Lambda}_1^{-1/2} & \mathbb{0} \end{bmatrix}. \end{aligned} \tag{4.17}$$

By inserting Eqs. (4.16, 4.17) into Eq. (4.15), one obtains

$$\begin{aligned}\langle S_z\rangle_t^{[Q]}(x,z) = \frac{1}{2}\, e^{-2k_0\Im\left[\sqrt{\lambda_Q}\right]z} \\ \times \sum_{qq'} \Re\Big\{ \Big(\left[\hat{W}_{11}^* \quad \mathbb{0}\right]_{qQ} \left[\lceil\epsilon^{-1}\rceil^{-1}\hat{W}_{11}\hat{\Lambda}_1^{-1/2} \quad \mathbb{0}\right]_{q'Q} \\ + \left[\mathbb{0} \quad \hat{W}_{22}^*\right]_{qQ} \left[\mathbb{0} \quad \lceil\mu^{-1}\rceil\hat{W}_{22}\hat{\Lambda}_2^{1/2}\right]_{q'Q} \Big)\, e^{-i\left(k_{x;q}-k_{x;q'}\right)x} \Big\}. \end{aligned} \tag{4.18}$$

In the two summands, due to the Fourier back transformation, the Toeplitz matrices $\lceil\epsilon^{-1}\rceil^{-1}$ and $\lceil\mu^{-1}\rceil$ can be replaced by the product with the functions $\epsilon_{\mathrm{rec}}(x)$ and $\mu_{\mathrm{rec}}^{-1}(x)$, i.e., the mode-truncated reconstructions of $\epsilon(x)$ and $\mu^{-1}(x)$. Furthermore, the matrices $\hat{\Lambda}_1^{-1/2}$ and $\hat{\Lambda}_2^{1/2}$ are diagonal and can be separated. One obtains

$$\begin{aligned}\langle S_z\rangle_t^{[Q]}(x,z) = \frac{1}{2}\, e^{-2k_0\Im\left[\sqrt{\lambda_Q}\right]z} \Bigg(\Re\left\{\epsilon_{\mathrm{rec}}(x)\,\hat{\Lambda}_{1;Q}^{-1/2}\right\} \sum_q \left| \left[\hat{W}_{11} \quad \mathbb{0}\right]_{qQ} e^{ik_{x;q}x} \right|^2 \\ + \Re\left\{\mu_{\mathrm{rec}}^{-1}(x)\,\hat{\Lambda}_{2;Q}^{+1/2}\right\} \sum_q \left| \left[\mathbb{0} \quad \hat{W}_{22}\right]_{qQ} e^{ik_{x;q}x} \right|^2 \Bigg). \end{aligned} \tag{4.19}$$

As in Eq. (4.15), the sign of $\langle S_z\rangle_t^{[Q]}(x,z)$ does not depend on z. Furthermore, one can see that the expression separates into two parts. Depending on the choice of Q, the contribution either comes from the first (x-polarized modes) or second (y-polarized modes) summand. No mixing of the two terms occurs. For the x-polarized modes, the sign of the energy flow is completely determined by $\Re\left\{\epsilon_{\mathrm{rec}}(x)\,\hat{\Lambda}_{1;Q}^{-1/2}\right\}$. For absorption-free materials, this means that the energy flow changes its sign simul-

taneously with $\epsilon_{\mathrm{rec}}(x)$. In analogy, the sign of the energy flow for y-polarized modes is completely determined by $\Re\left\{\mu_{\mathrm{rec}}^{-1}(x)\,\hat{\Lambda}_{2;Q}^{+1/2}\right\}$, i.e., for absorption-free materials the energy flow changes its sign simultaneously with $\mu_{\mathrm{rec}}(x)$.

4.4.4 Propagation direction of eigenmodes

In Eqs. (4.15, 4.19), the z-component of the time averaged Poynting vector of an eigenmode, $\langle S_z\rangle_t^{[Q]}(\mathbf{r})$, is derived. Furthermore, it is shown that the quantity generally is a function of x, y and z. However, since the z-dependence is only an exponential decay or growth, the sign does explicitly *not* depend on the variable z. Therefore, $\langle S_z\rangle_t^{[Q]}(\mathbf{r})$ can be regarded as a measure for the propagation direction of the Q^{th} eigenmode. However, as mentioned before, the energy flow in z-direction may be subject to sign changes within a period in x and y, i.e., in some lateral intervals energy flows in positive direction, whereas in others it flows in negative direction. Therefore, one must consider the direction of the net energy flow in z-direction, which is obtained by integrating Eq. (4.15) over one period in both x and y,

$$\iint_{P_x\,P_y}\langle\mathbf{S}(\mathbf{r},t)\rangle_t^{[Q]}\,\mathrm{d}x\,\mathrm{d}y=\frac{1}{2}\,e^{-2k_0\Im\left[\sqrt{\lambda_Q}\right]z}\,\sum_q\Re\left[\left(W_{x;qQ}V_{y;qQ}^{*}-W_{y;qQ}V_{x;qQ}^{*}\right)\right]\,. \tag{4.20}$$

The propagation direction of the Q^{th} eigenmode is then defined as the sign of this quantity,

$$\sigma^{[Q]}=\mathrm{sgn}\left[\iint_{P_x\,P_y}\langle\mathbf{S}(\mathbf{r},t)\rangle_t^{[Q]}\,\mathrm{d}x\,\mathrm{d}y\right]\,. \tag{4.21}$$

It is straightforward to show that changing the sign of $\sqrt{\lambda_Q}$ in Eq. (4.12) causes also a sign change in the energy flow, i.e., in Eq. (4.21). Therefore, by choosing the correct sign for each $\sqrt{\lambda_q}$ separately, one can achieve a proper assignment of the coefficients $c_{+;q}$ / $c_{-;q}$ as the forward / backward propagating eigenmodes. In principle, the freedom in this assignment also allows providing the transmitted field and subsequently using the RCWA to calculate the necessary back-side illumination (cf. Chapter 5).

4.5 Simulation examples and verification

In this section, several RCWA simulation examples are presented, which make use of the possibility for structured incident and/or transmitted regions. In Sect. 4.5.1, a comparison between the analytic eigenmodes of a planar dielectric waveguide and the

numerically determined RCWA eigenmodes is made. In Sect. 4.5.2, the evanescent coupling of the electric field between two adjacent waveguides is demonstrated and compared to CMT. In Sect. 4.5.3, the mode coupling across a fiber interface into a GRIN fiber with sech profile is investigated and compared to a BPM simulation.

4.5.1 Planar dielectric waveguide: comparison to analytic eigenmodes

The two-dimensional planar dielectric waveguide as depicted in Fig. 4.3 is one of the few examples of structured media, for which analytic eigenmodes are known. A detailed derivation can be found in [24] (Chapter 8.2) and [74] (Chapter 2.4). In the following, the most important results of this derivation is shortly outlined. The

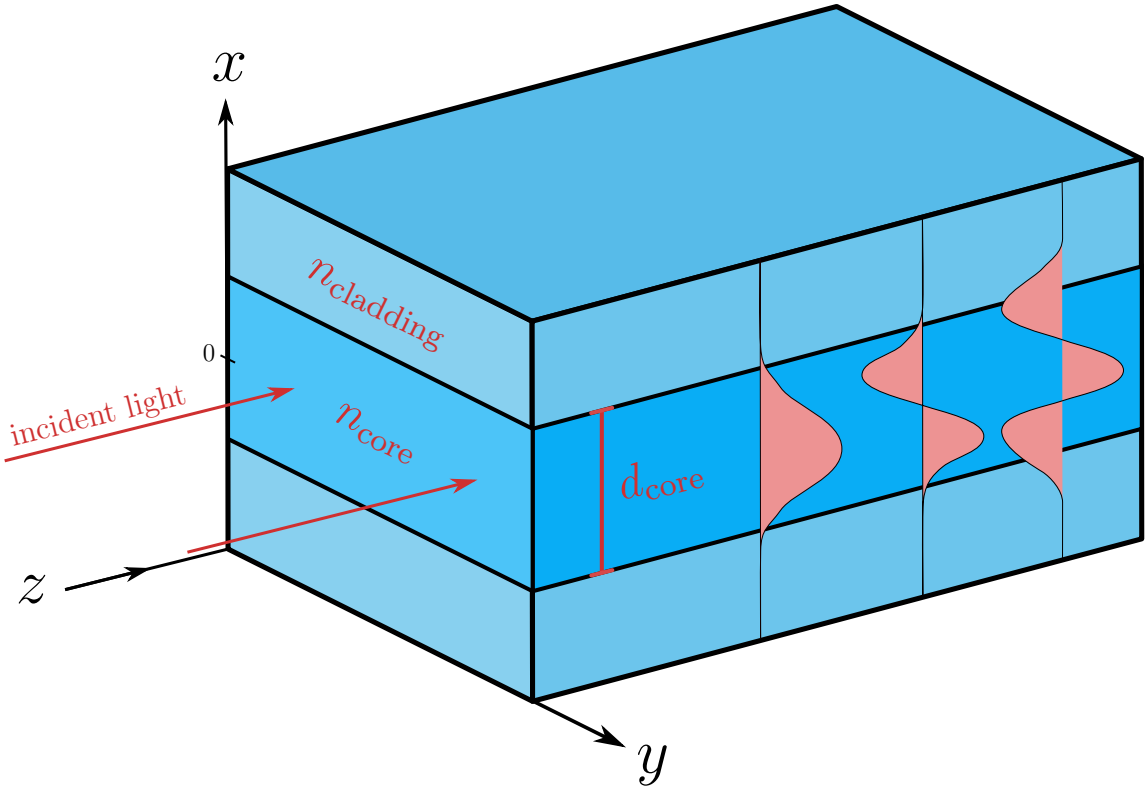


Figure 4.3: Planar dielectric waveguide of diameter d_{core} with refractive indices n_{core} and n_{cladding}. The refractive index distribution is constant in y.

eigenmodes of a planar dielectric waveguide are characterized by their individual propagation angle θ_m, which is determined by a consistency equation. Here, the consideration is restricted to TE-polarization, i.e., the electric field polarization vector oscillates in y-direction. Similar results are obtained for TM-polarization. In the

TE case, the consistency equation reads

$$\tan\left(\pi\frac{d\,n_{\text{core}}}{\lambda}\sin\theta_q - q\frac{\pi}{2}\right) = \sqrt{\frac{\cos^2\theta_{\text{c}}}{\sin^2\theta_q} - 1}\,, \tag{4.22}$$

where $\theta_{\text{c}} = \operatorname{asin}\left(n_{\text{cladding}}/n_{\text{core}}\right)$ is the critical angle for total internal reflection. This equation only has a solution for a finite number of m, i.e., the number of propagative eigenmodes is limited. From the highest propagative eigenmode angle, the NA of such a waveguide can be determined,

$$\text{NA}_{\text{waveguide}} = \sqrt{n_{\text{core}}^2 - n_{\text{cladding}}^2}\,. \tag{4.23}$$

The electric field of the q^{th} eigenmode is of the form

$$\mathcal{E}_{x;q}^{\text{TE}}(\mathbf{r}) \propto \begin{cases} \cos\left(2\pi\frac{n_{\text{core}}\sin\theta_q}{\lambda}x\right)e^{i\beta_q z} & \text{for } q = 0, 2, 4\ldots \\ \sin\left(2\pi\frac{n_{\text{core}}\sin\theta_q}{\lambda}x\right)e^{i\beta_q z} & \text{for } q = 1, 3, 5\ldots \end{cases} \tag{4.24}$$

for the propagation inside the waveguide core, $-d_{\text{core}}/2 \leq x \leq +d_{\text{core}}/2$, with $\beta_q = n_{\text{core}}\,k_0\,\cos\theta_q$. In the cladding, $|x| > d_{\text{core}}/2$, one obtains an exponential decrease in the direction away from the core,

$$\mathcal{E}_{x;q}^{\text{TE}}(\mathbf{r}) \propto e^{-\gamma_q|x|}\,e^{i\beta_q z}\,, \tag{4.25}$$

where $\gamma_q = \sqrt{\beta_q^2 - n_{\text{cladding}}^2 k_0^2}$. At the interface between the cladding and the core, the electric field is continuous. In Fig. 4.4, the normalized electric field amplitudes $\mathcal{E}_{x;q}^{\text{TE}}(\mathbf{r})$ are plotted for the parameters $n_{\text{core}} = 1.47$, $n_{\text{cladding}} = 1.46$ and $d_{\text{core}} = 12\,\mu\text{m}$. For this parameter set, only three propagative eigenmodes exist. In Fig. 4.5, these analytic eigenmodes are compared to the TE eigenmodes that are determined numerically by means of the RCWA. It is observed that the eigenmode decomposition of the RCWA reproduces the analytic eigenmodes very well. In the following, assume a homogeneous medium with $n^{(0)} = 1.47$ for $z < 0$, and a Gaussian beam light incidence with waist $\sigma_0 = 3.5\,\mu\text{m}$. In the half space $z \geq 0$, assume the planar dielectric multi-mode waveguide with the parameters as given before. Fig. 4.6 shows the resulting normalized electric field amplitude calculated numerically with the RCWA. The stated simulation parameters are chosen to be the same as in [63] (Sect. 3.10.4), who simulated a $1400\,\mu\text{m}$ long waveguide slab sandwiched between two index-matched homogeneous regions. However, in contrast to [63], in the simulation conducted here, the waveguide explicitly has *no* right interface. Indeed, the simulation results for the field inside the fiber look similar, but visible differences exist. Furthermore, the treatment of the waveguide as structured transmitted region without a rear interface has the following advantages.

- Unwanted back-reflections originating from the rear waveguide interface are

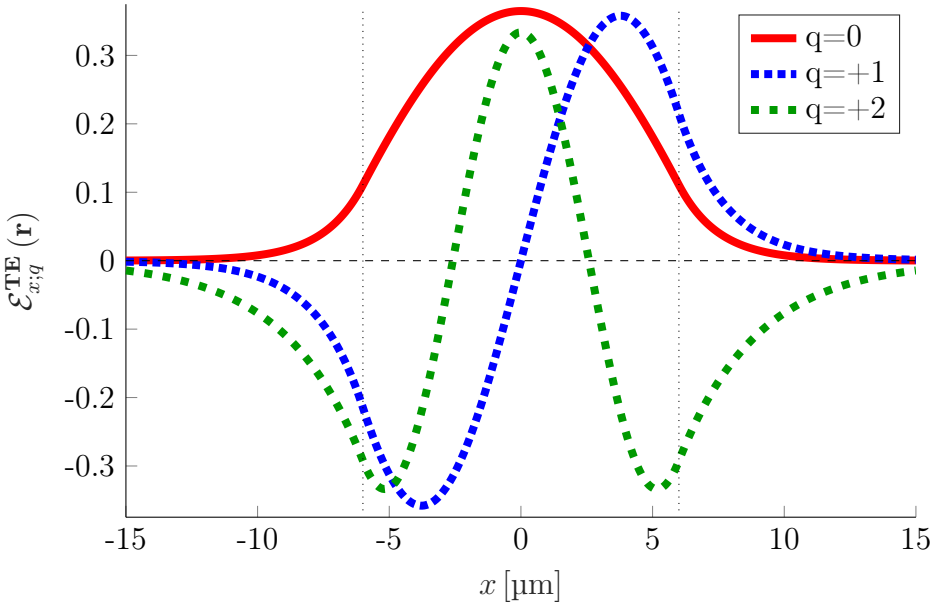


Figure 4.4: Normalized electric field amplitudes $\left|\mathcal{E}^{\mathrm{TE}}_{x;q}(\mathbf{r})\right|$ for the planar dielectric waveguide eigenmodes. For the given parameters, only three propagative solutions ($q = 0, 1, 2$) exist. The vertical dashed black lines indicate the core boundaries.

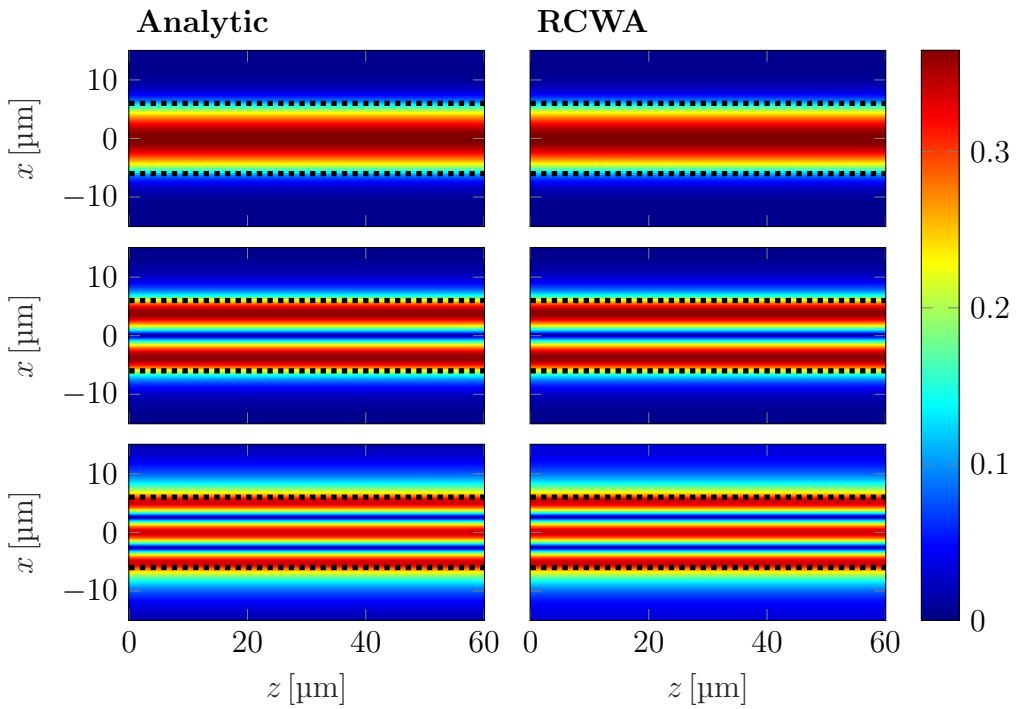


Figure 4.5: Normalized electric field amplitudes $\left|\mathcal{E}^{\mathrm{TE}}_{x;q}(\mathbf{r})\right|$ of the waveguide TE-eigenmodes. Analytic eigenmodes (left) and numerically determined eigenmodes (right). The dashed black lines indicate the border of the core. RCWA parameters: $P_x = 30\,\mu\mathrm{m}$, $M_0 = 200$, $\lambda = 1.55\,\mu\mathrm{m}$.

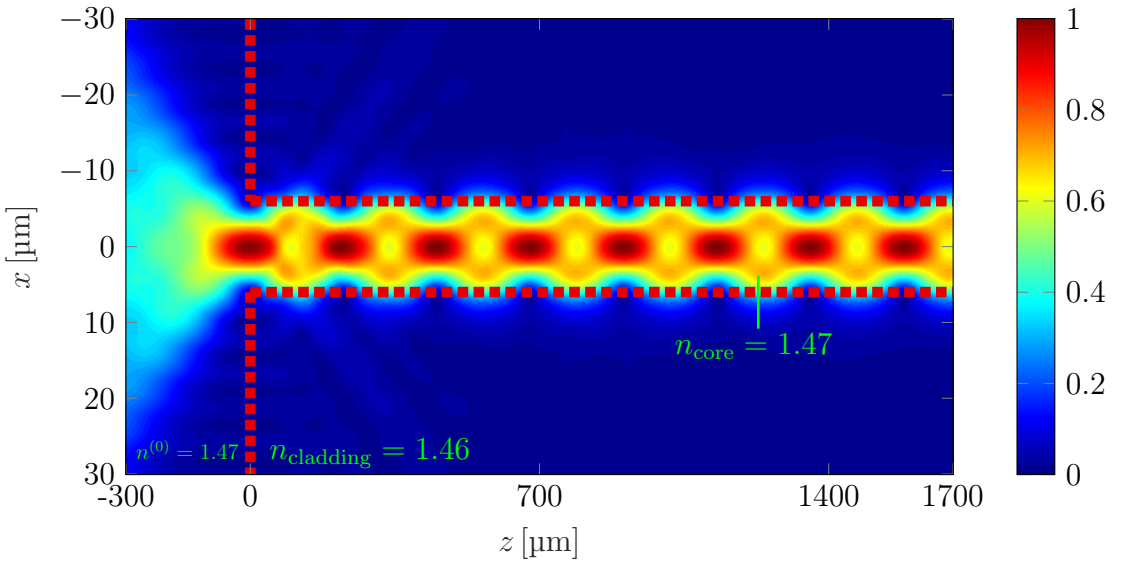


Figure 4.6: Gaussian beam incident on the left interface of a planar dielectric waveguide (normalized electric field amplitude, parameters given in the text). No right waveguide interface is assumed. The dashed red lines indicate the core boundaries. RCWA parameters: $P_x = 125\,\mu\text{m}$, $M_0 = 200$, $\lambda = 1.55\,\mu\text{m}$.

inherently avoided. In contrast, previous publications still had to sandwich the waveguide between two index-matched homogeneous regions[63].

- The transmitted waveguide eigenmode coefficients can be directly identified as the coupling efficiencies of the illuminating light into the different mode channels across the waveguide interface.

- Arbitrary propagative or non-propagative waveguide eigenmodes can be directly excited as incident light. This may, for instance, be useful in the calculation of the mode coupling *out* of a fiber. Before, the simulation of such a setting was only possible in very restricted cases. For instance, [63] motivated how an isolated fundamental mode in a single mode waveguide could be produced in the standard RCWA framework.

4.5.2 Evanescent coupling between waveguides

Consider the setup shown in Fig. 4.7. Two planar dielectric waveguides of diameter d_{core} are aligned parallel to each other with the separation d_{gap}. The refractive indices of the core and cladding are n_{core} and n_{cladding}. From the analytic eigenmodes of the isolated waveguides, Eqs. (4.24, 4.25), one can see that the electric field is not strictly zero outside the waveguide core, but that there is an exponential decay of the amplitude inside the cladding. If two waveguides are sufficiently close to each other, the eigenfunctions of the waveguides may to a certain degree overlap with the close-by other waveguide via the exponentially decaying branch, as illustrated in

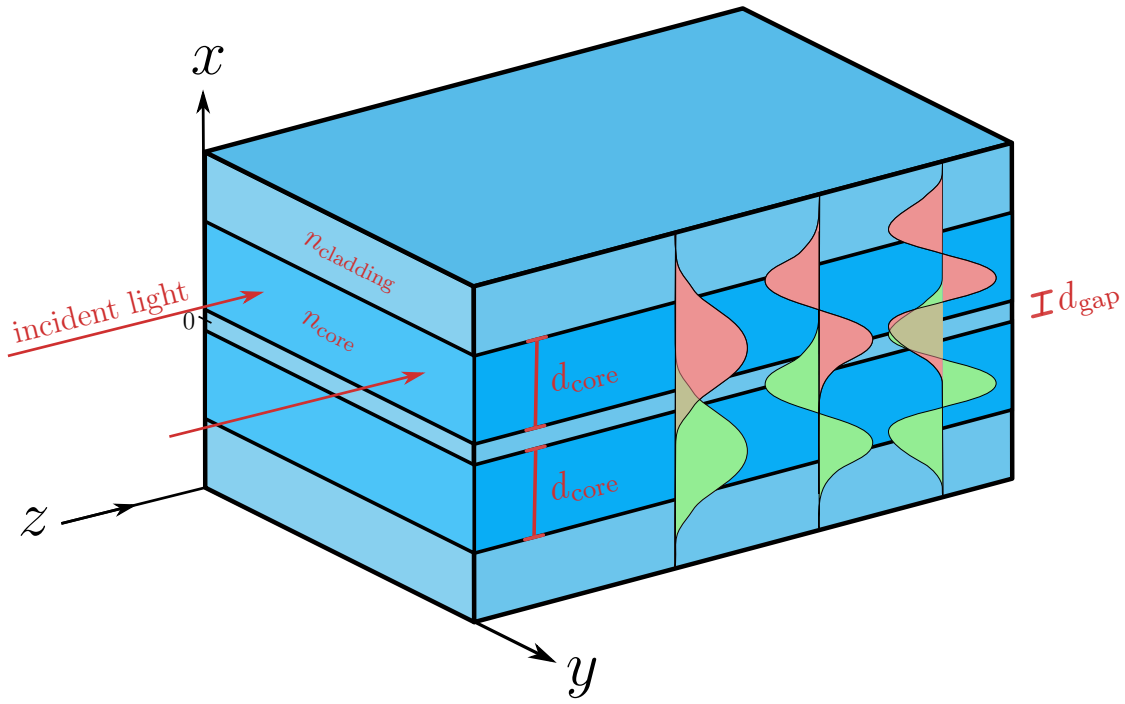


Figure 4.7: Two planar dielectric waveguides of diameter d_{core} and refractive index n_{core} at a separation d_{gap}. The refractive index of the cladding is n_{cladding}. Assume that the waveguide has no interface in axial direction.

Fig. 4.8. If an electric field is excited in one of the waveguides, due to this overlap, it is possible to transfer this excitation across the barrier. This effect is called 'evanescent coupling', because the coupling is solely communicated via the exponentially decaying or 'evanescent' branch of the eigenfunctions. A good description of this effect can be found in [24]. In order to find the exact solution of this problem, one must find the eigenmodes of the global system. For strong coupling, $d_{\text{gap}} \ll \lambda$, the resulting eigenmodes naturally deviate from those of the isolated waveguides. For weak coupling, $d_{\text{gap}} \gg \lambda$, one expects the eigenfunction of the two waveguides to separate and look similar to the isolated waveguide eigenmodes.

First, consider the case of weak coupling (parameters given in Fig. 4.9) and that the eigenmode of an isolated waveguide is illuminated in the upper waveguide. In Fig. 4.9 (top), the axial propagation of the field is calculated with the RCWA. The exact solution of these types of systems are often very complex and tedious to calculate due to the limited amount of computation power. For larger sized three-dimensional systems, this problem still persists today. In the case of weakly coupled single-mode waveguides, however, CMT is a good approximative theory. A good derivation can be found in [21, 22, 24]. In the following, the results obtained with the RCWA are compared to the result of CMT. To this end, the derivation of [24] is sketched at this point. Be $u_1(x) \exp(i\beta_1 z)$ and $u_2(x) \exp(i\beta_2 z)$ the eigenfunctions of the isolated single-mode waveguides. The major assumption of CMT is that the mode coupling is

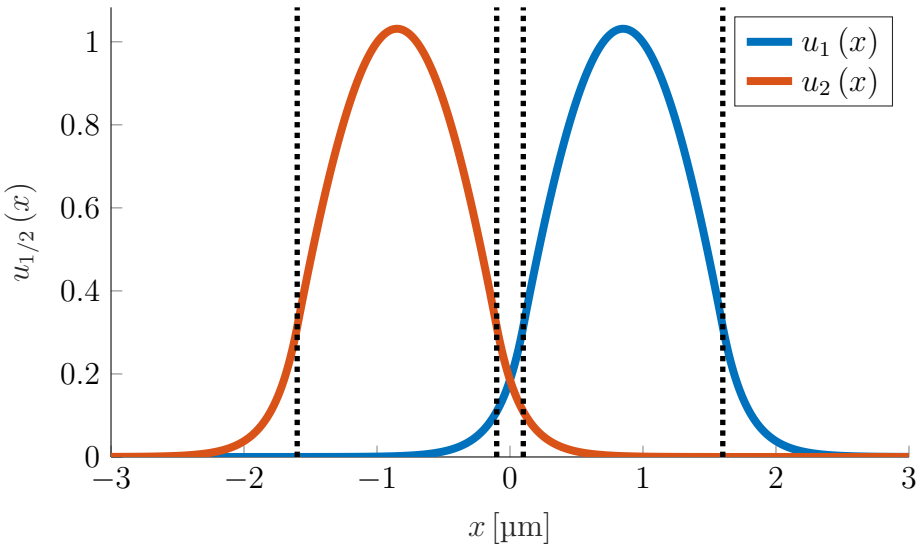


Figure 4.8: Eigenfunctions of two close-by adjacent waveguides (fundamental modes). The exponentially decaying branch of the eigenmode in the left waveguide overlaps with the right waveguide and vice versa. The dashed black lines indicate the core boundaries.

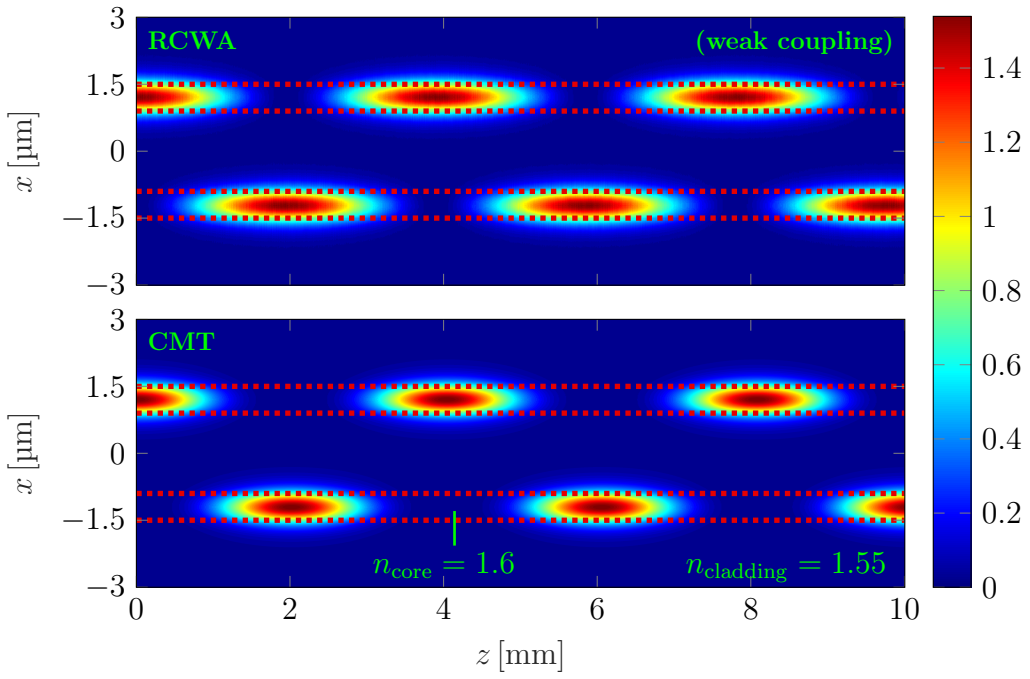


Figure 4.9: Evanescent coupling between two fibers for weak coupling ($d_{\text{core}} = 0.6\,\mu\text{m}$, $d_{\text{gap}} = 1.8\,\mu\text{m}$). Electric field intensity calculated with the RCWA (top) and with CMT (bottom). The dashed red lines indicate the core boundaries. RCWA parameters: $P_x = 10\,\mu\text{m}$, $M_0 = 200$, $\lambda = 633\,\text{nm}$.

communicated solely via the field amplitudes. The part of the electric field amplitude in waveguide one, which overlaps with the solution of waveguide two, is regarded as a source term for waveguide two and vice versa. This way, one obtains two coupled differential equations for the amplitudes a_1 and a_2 of the waveguide eigenmodes,

$$\frac{\mathrm{d}a_1}{\mathrm{d}z} = -i\kappa_{21} e^{+i\Delta\beta z} a_2(z) \tag{4.26a}$$

$$\frac{\mathrm{d}a_2}{\mathrm{d}z} = -i\kappa_{12} e^{-i\Delta\beta z} a_1(z) \ , \tag{4.26b}$$

where $\Delta\beta = \beta_1 - \beta_2$. The coupling strengths κ_{12} and κ_{21} are determined via the overlap integrals as defined in [24]. For $a_2(0) = 0$, this system of differential equations is solved by

$$|a_1(z)|^2 = |a_1(0)|^2 \left[\cos^2(\gamma z) + \left(\frac{\Delta\beta}{2\gamma}\right)^2 \sin^2(\gamma z)\right] \tag{4.27a}$$

$$|a_2(z)|^2 = |a_1(0)|^2 \frac{|\kappa_{21}|}{\gamma^2} \sin^2(\gamma z) \ , \tag{4.27b}$$

with $\gamma^2 = \left(\frac{\Delta\beta}{2}\right)^2 + \kappa_{12}\kappa_{21}$. From this equation, one can see that, upon propagation in z, power is transferred from one waveguide to the other and back. For identical waveguides one even obtains a full transfer of power,

$$|a_1(z)|^2 = |a_1(0)|^2 \cos^2(\kappa z) \tag{4.28a}$$

$$|a_2(z)|^2 = |a_1(0)|^2 \sin^2(\kappa z) \ , \tag{4.28b}$$

where $\kappa = \kappa_{12} = \kappa_{21}$. Fig. 4.9 (bottom) shows the electric field intensity inside the fiber simulated with the RCWA. For the given parameters, CMT predicts a full transfer of power after $\pi/(2\kappa) \approx 2019\,\mu\mathrm{m}$, which is just slightly more than the $1950\,\mu\mathrm{m}$ predicted by the RCWA calculation (deviation of 3.5 %). Still, the intensity maxima in the RCWA simulation are significantly wider in axial direction.

Now consider the case of strong coupling. For the parameters given in Fig. 4.10, the waveguides are not single-mode waveguides, but support three propagative eigenmodes. In the following, the lowest eigenmode of an isolated waveguide is excited in the upper waveguide. The electric field intensity inside the fiber calculated with the RCWA and CMT is shown in Fig. 4.10. Also in this case, the signal oscillates back and forth between the two waveguides due to evanescent coupling. However, the intensity peaks in the RCWA calculation are far more perturbed and look different from one peak to another. The reason is that light is also coupled into the higher eigenmodes of the waveguides and not only into the fundamental eigenmode. For these parameters, CMT predicts a full energy transfer after $\pi/(2\kappa) \approx 134\,\mu\mathrm{m}$, which significantly deviates from the RCWA result by 18% ($\sim 158\,\mu\mathrm{m}$).

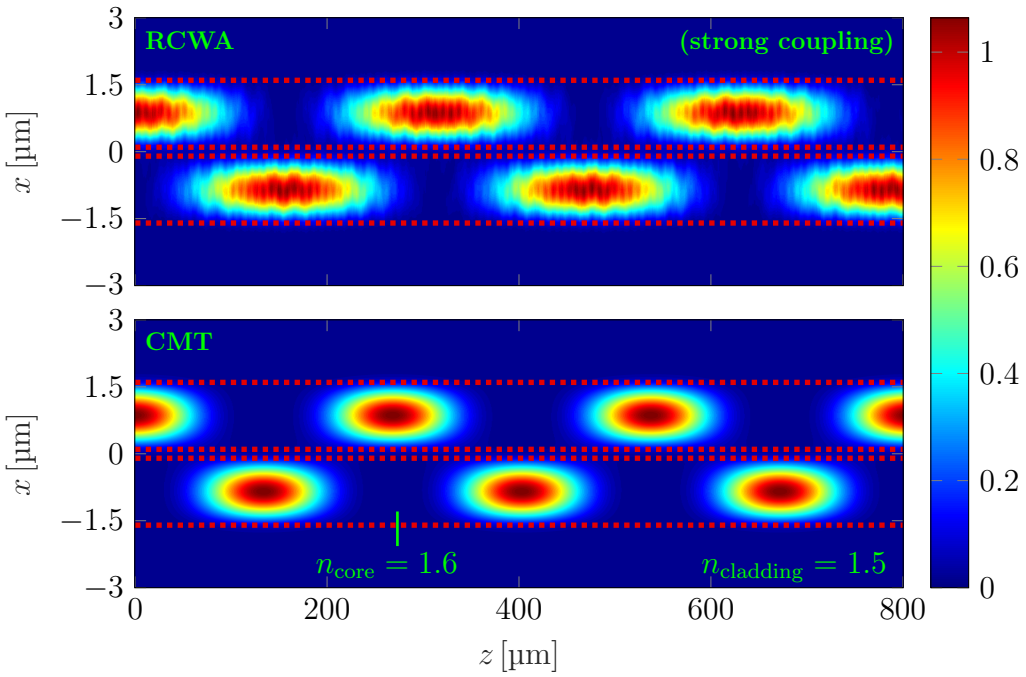


Figure 4.10: Evanescent coupling between two fibers for strong coupling ($d_{\text{core}} = 1.5\,\mu\text{m}$, $d_{\text{gap}} = 0.2\,\mu\text{m}$). Electric field intensity calculated with the RCWA (top) and with CMT (bottom). The dashed red lines indicate the core boundaries. RCWA parameters: $P_x = 10\,\mu\text{m}$, $M_0 = 200$, $\lambda = 633\,\text{nm}$.

In summary, CMT makes very strong assumptions about the system and predictions become unreliable for either multi-mode waveguides and/or strong coupling. On the other hand, in the RCWA calculation, the coupling coefficients are inherently determined in the correct manner, and the mode coupling into all the higher waveguide eigenmodes is automatically taken into account.

4.5.3 Mode coupling into a long graded index fiber

In the following, assume a two-dimensional GRIN waveguide, where the refractive index distribution of the core possesses a sech- or $\cosh^{-1}$-profile as illustrated in Fig. 4.11. The light propagation in sech-profile fibers can be solved analytically, and

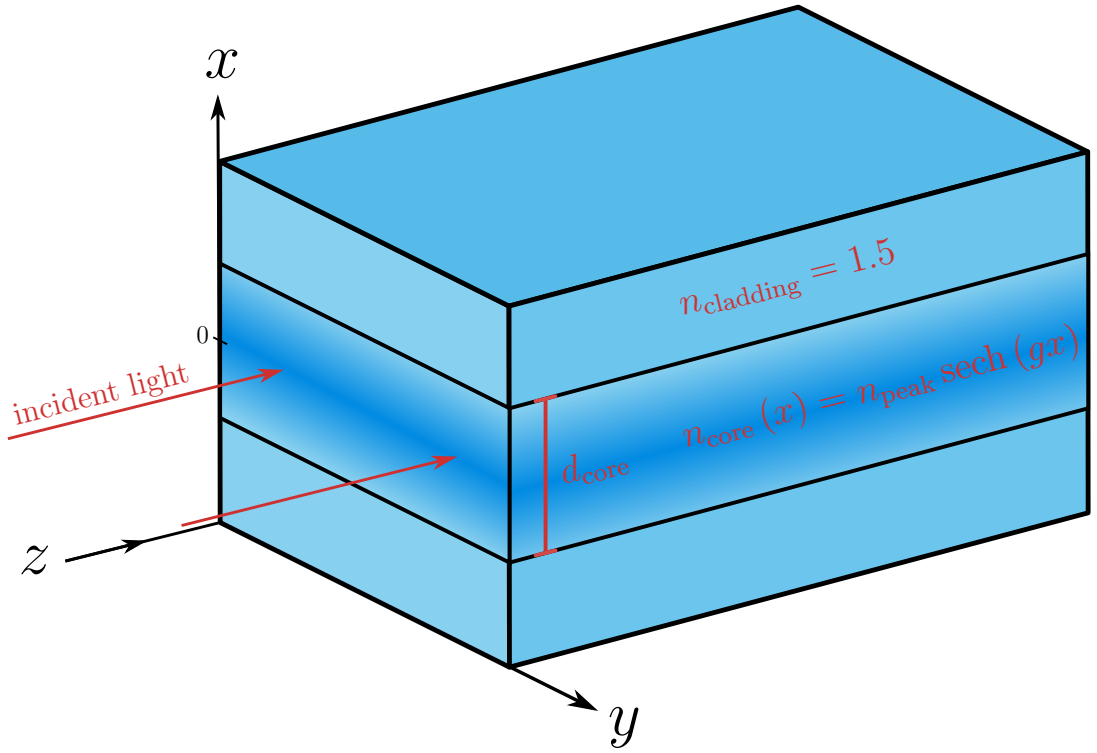


Figure 4.11: GRIN waveguide (2d) with sech-profile core of diameter d_{core}, and constant refractive index in the cladding. Sample parameters: $d_{\text{core}} = 60\,\mu\text{m}$, $n_{\text{cladding}} = 1.5$, $n_{\text{peak}} = 1.6$, $g = \frac{2}{d}\operatorname{sech}^{-1}\left(\frac{n_{\text{cladding}}}{n_{\text{peak}}}\right)$. The NA of the waveguide for these parameters is $\text{NA}_{\text{waveguide}} = 0.56$. The refractive index is constant in y-direction.

it is known that this type of GRIN fiber has imaging characteristics, i.e., the field distribution at any axial position is repeatedly imaged across some characteristic length in axial direction[75]. As an example, consider an interface between vacuum ($n^{(0)} = 1$) and a sech-profile waveguide with parameters as given in Fig. 4.11. Assume that the waveguide is very long, i.e., its rear interface is far away and does not play a role. At $z = 0$, the waveguide's left interface is illuminated with a slightly off-centered diffraction limited focal spot. Fig. 4.12 shows the RCWA and BPM simulations of this setting. Indeed, the imaging characteristics is correctly predicted by the RCWA calculation. In comparison to the BPM, the following differences are observed.

1. The BPM predicts a slightly shorter axial imaging length than the RCWA.
2. The peak height of the focus at $z = 0$ is significantly elevated in the RCWA due to the reflection at the fiber interface.

3. The the overall field amplitude within the fiber is lower in the RCWA calculation due to reflection losses.

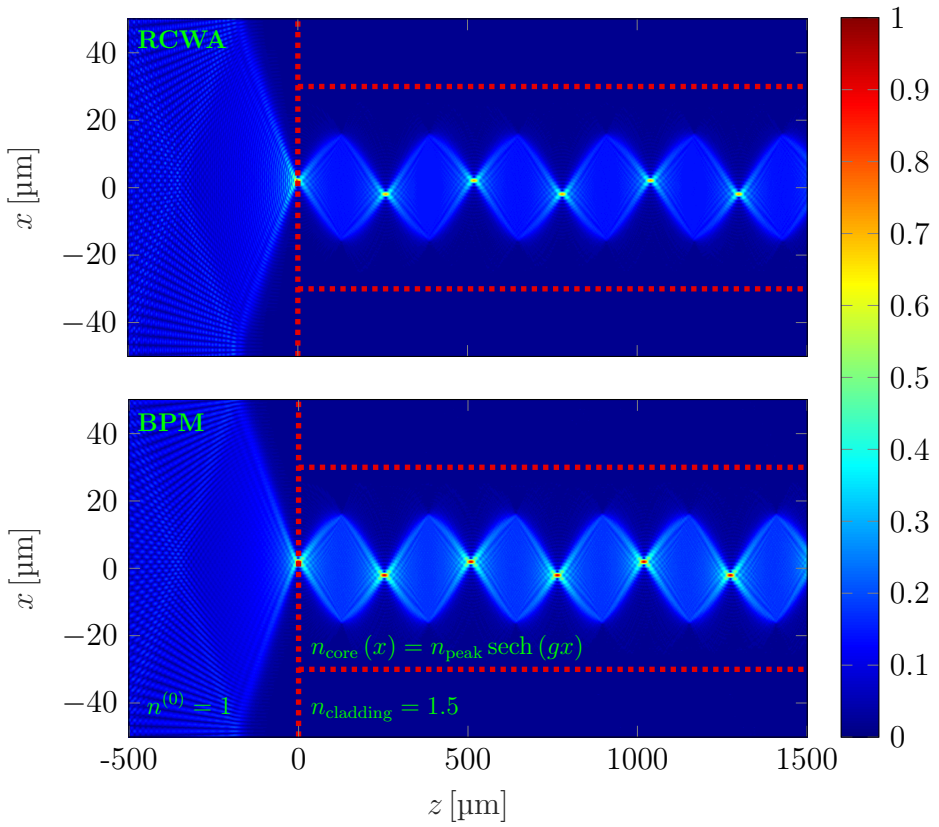


Figure 4.12: Interface of a sech-index fiber illuminated by a diffraction limited focus (NA = 0.3), which is shifted 2 µm off-center. Electric field amplitude calculated with the RCWA (top) and BPM (bottom). The dashed red lines indicate the core boundaries. RCWA/BPM parameters: $P_x = 100\,\mu\text{m}$, $M_0 = 200$, $\lambda = 633\,\text{nm}$.

Furthermore, the RCWA also allows a valid and fast field calculation for quite large propagation distances, i.e., far away from the waveguide interface. This is explicitly *not* the case for other established algorithms that are often applied for light propagation in fibers, also for the BPM. As mentioned in Sect. 2.7, the BPM accumulates errors upon every propagation step. The resulting effect is shown in Fig. 4.13, which shows the electric field amplitude 18-20 mm inside the waveguide. Clearly, the BPM result is already highly perturbed after a few millimeters of propagation, while the RCWA still gives reasonable results. The propagation distance possible with the RCWA is, in principle, only limited by the numerical errors in the eigenvalue decomposition.

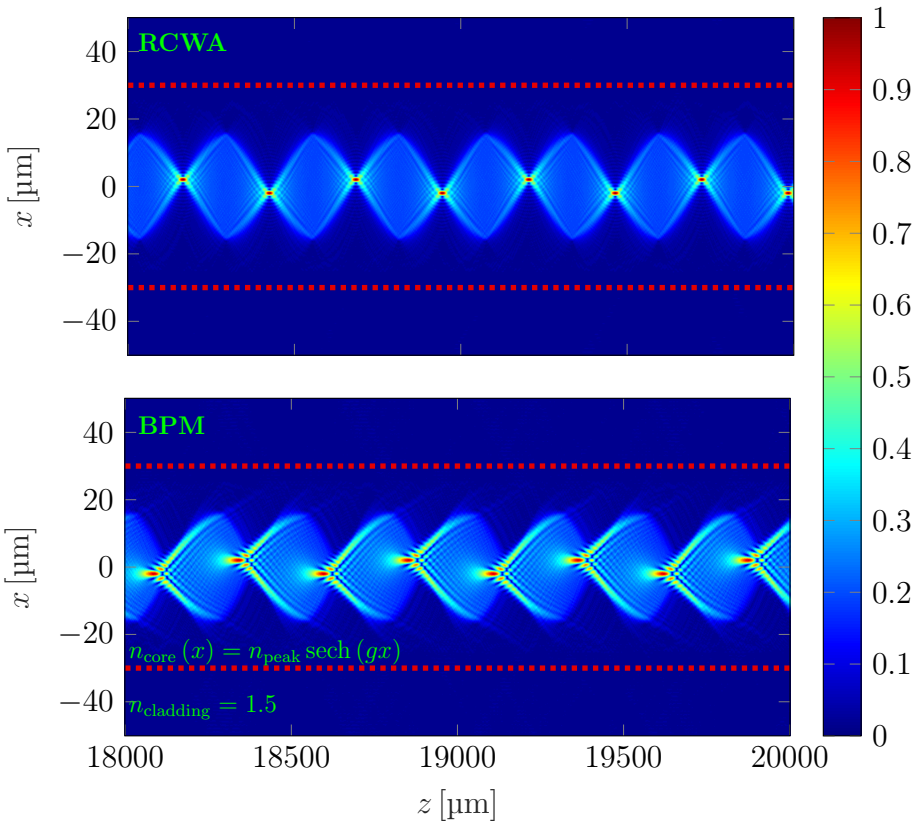


Figure 4.13: Electric field amplitude calculated with the RCWA (top) and BPM (bottom) 18 mm inside the fiber. The BPM simulation features severe perturbations, whereas the RCWA simulation still gives reliable results. The dashed red lines indicate the core boundaries. RCWA/BPM parameters: $P_x = 100\,\mu\text{m}$, $M_0 = 200$, $\lambda = 633\,\text{nm}$.

In the given setup, an interesting interpretation also concerns the transmitted eigenmode coefficients, which represent the coupling efficiencies about the interface. The absolute value of the eigenmode coefficients in the waveguide region corresponds to the degree of excitation of the individual eigenmodes.

4.6 Conclusions

In this chapter, the RCWA framework is extended to allow for the normally homogeneous incident and transmitted regions to be structured as the grating layers. Possible applications constitute the light propagation in structured layers of large axial extent, for instance the mode coupling into or out of long waveguides, or generally all cases where possible reflections from an unwanted distant interface of a structured region should be neglected. This way, it is possible to treat waveguides with only one or even with no interface to a homogeneous region at all. In order to set sensible boundary conditions, the propagation direction of the individual eigenmodes must be known in these regions. The RCWA extension derived in this

chapter is verified by a comparison to the analytic eigenmodes of a planar dielectric waveguide. Furthermore, the evanescent coupling between two adjacent waveguides is successfully reproduced, including a positive comparison to CMT for a weakly coupled system. On the other hand, it is shown that, as expected, the predictions of CMT significantly deviate from the RCWA result for a strongly coupled system. Moreover, the mode coupling from vacuum into a long GRIN fiber is simulated. The typical imaging characteristic of a sech-profile waveguide is demonstrated. Furthermore, it is demonstrated that the RCWA also allows a valid and fast field calculation for quite large propagation distances, i.e., far away from interfaces. This is explicitly *not* the case for other propagation algorithms that are often applied to waveguides or fibers, such as for instance the BPM[51] or the WPM[52], which accumulate errors upon every propagation step. The propagation distance possible with the RCWA is, in principle, only limited by the numerical error in the eigenvalue decomposition.

5 | Bidirectional light incidence in RCWA

This chapter is partly based on the publications [A3] and [A4][8].

5.1 Motivation and structure

The standard RCWA framework is limited to unidirectional plane wave light incidence. However, in applications like, for instance, 4π-microscopy, coherent wave fronts are actually incident from both sides onto the object[76, 77] (cf. Fig. 5.1). Therefore, in this chapter, the framework of the RCWA is extended to bidirectional

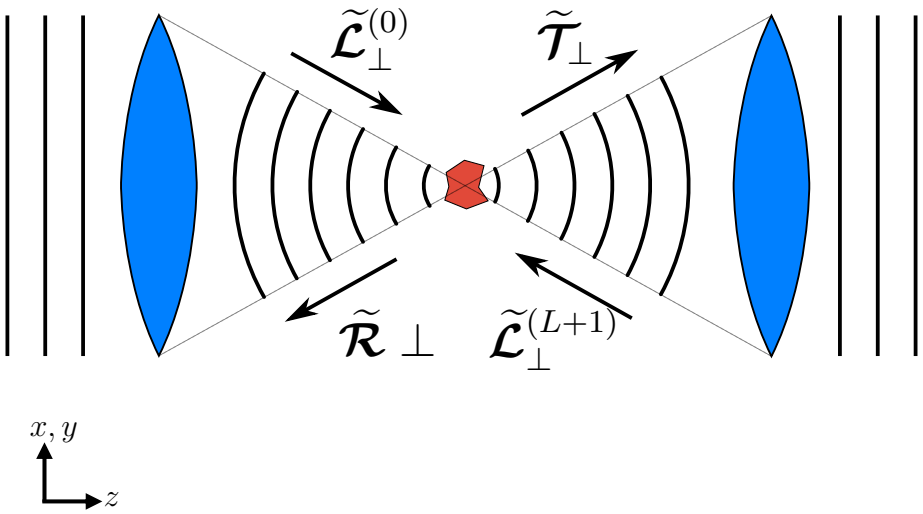


Figure 5.1: Principle of a 4π-microscope. A specimen is illuminated with two coherent converging wave fronts from opposite directions.

coherent light incidence. In order to model the setting of the 4π-microscope in the RCWA, three distinct concepts are combined. First, the previously introduced idea of structured illumination is applied to model focused beams in a fast and consistent manner (cf. [17] and Sect. 3.1.3). Second, in Sect. 5.2, it is shown how structured

[8] © 2017 IEEE.

light sources of different polarization types, such as linear, radial, azimuthal or elliptical polarization, can be realized in a general form. Eventually, in Sect. 5.3, the concept of coherent bidirectional light incidence is integrated into the RCWA framework. In Sect. 5.4, these techniques are combined to simulate bidirectional focused illumination and sample-induced wave front aberrations in 4π-microscopy.

5.2 Polarization of structured incident fields

In this section it is shown how polarization is introduced into a multi-mode treatment of rigorous diffraction. This problem has been addressed much earlier before in the context of high NA imaging[78, 79, 80]. These authors, unfortunately, did not provide a clear cut derivation. In [78, 79], only a plausibility consideration is given, whereas in [80] the final result is presented as given. The basic idea of the derivation shown here goes back to an idea of Brenner[81]. The full derivation is provided here. Interestingly, the result that is obtained can be considered as the very general case of the results presented in [78, 79, 80].

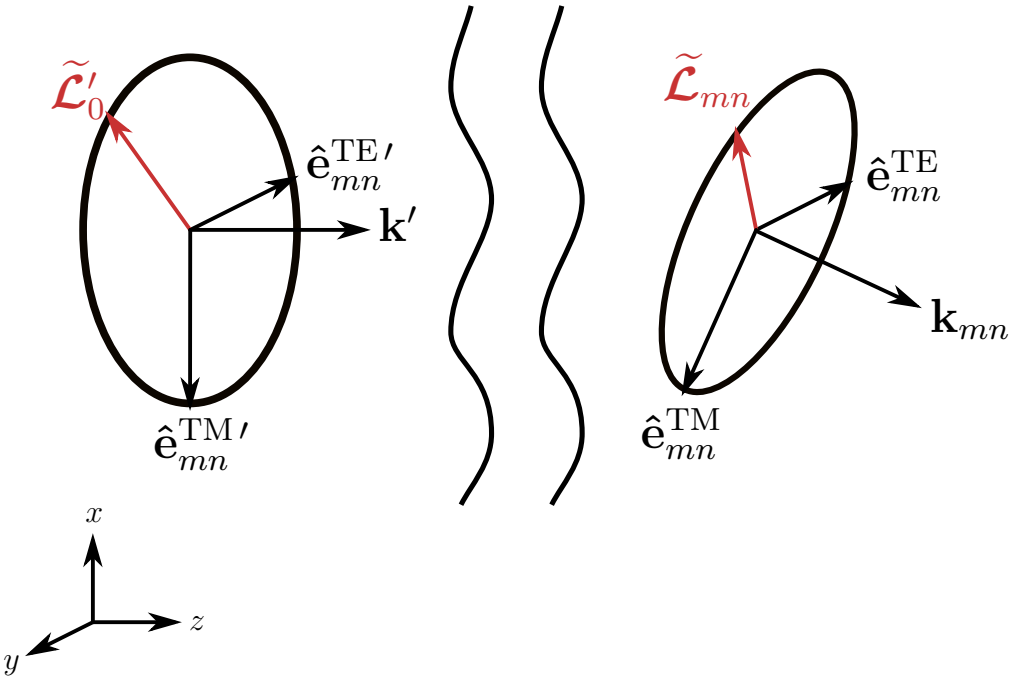


Figure 5.2: A plane wave with amplitude $\widetilde{\mathcal{L}}'_0$ and wave vector $\mathbf{k}' = k\hat{\mathbf{e}}_z$ passes an optical system, which transforms it into a superposition of plane waves with amplitudes $\widetilde{\mathcal{L}}_{mn}$ and wave vectors $\mathbf{k}_{mn}$.

In an imaging experiment, as well as in microscopy, a structured field distribution is usually formed by an initially polarized wave front, which is structured by some kind of optical system. In imaging, the optical system consists of a mask illuminated by a plane wave, which is imaged by some kind of imaging optics. In confocal microscopy,

the optical system reshapes a planar wave front into a diffraction limited focus. Consider the situation depicted in Fig. 5.2. A plane wave with given amplitude and polarization state $\widetilde{\boldsymbol{\mathcal{L}}}_0'$, and wave vector $\mathbf{k}' = k\hat{\mathbf{e}}_z$ is incident upon an optical system. The latter transforms the incident plane wave into a spatially structured wave front, which consists of a superposition of multiple plane waves with wave vectors $\mathbf{k}_{mn}$. Now consider one specific $\mathbf{k}_{mn}$. The plane wave's particular decomposition into TE and TM components depends only on $\mathbf{k}_{mn}$ itself. The TE-direction is specified as perpendicular to the plane spanned by the vectors $\mathbf{k}_{mn}$ and $\mathbf{k}'$. Thus, $\hat{\mathbf{e}}_{mn}^{\mathrm{TE}}$ is given by

$$\hat{\mathbf{e}}_{mn}^{\mathrm{TE}} = \frac{1}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \begin{bmatrix} -k_{y;mn} \\ +k_{x;mn} \\ 0 \end{bmatrix} . \tag{5.1}$$

The TM polarization vector then follows from orthogonality, $\mathbf{k}_{mn} \perp \hat{\mathbf{e}}_{mn}^{\mathrm{TM}} \perp \hat{\mathbf{e}}_{mn}^{\mathrm{TE}}$. Thus, the TM direction is given by

$$\hat{\mathbf{e}}_{mn}^{\mathrm{TM}} = \frac{1}{|\mathbf{k}_{mn}|} \frac{1}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \begin{bmatrix} k_{x;mn}k_{z;mn} \\ k_{y;mn}k_{z;mn} \\ -k_{x;mn}^2 - k_{y;mn}^2 \end{bmatrix} . \tag{5.2}$$

Now, it is assumed that the initial polarization state $\hat{\mathbf{e}}_{mn}^{\mathrm{TE}\prime}$, $\hat{\mathbf{e}}_{mn}^{\mathrm{TM}\prime}$ of the plane wave is converted into the new set of polarization components $\hat{\mathbf{e}}_{mn}^{\mathrm{TE}}$, $\hat{\mathbf{e}}_{mn}^{\mathrm{TM}}$ by a rotation of the coordinate system around the $\hat{\mathbf{e}}_{mn}^{\mathrm{TE}}$ direction. Thus,

$$\begin{aligned} \hat{\mathbf{e}}_{mn}^{\mathrm{TE}\prime} &= \frac{1}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \begin{bmatrix} -k_{y;mn} \\ +k_{x;mn} \\ 0 \end{bmatrix} \\ \hat{\mathbf{e}}_{mn}^{\mathrm{TM}\prime} &= \frac{1}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \begin{bmatrix} +k_{x;mn} \\ +k_{y;mn} \\ 0 \end{bmatrix} . \end{aligned} \tag{5.3}$$

From the transversality of the electromagnetic fields, it can be concluded that any given polarization vector $\widetilde{\boldsymbol{\mathcal{L}}}_0'$ can be decomposed into

$$\frac{\widetilde{\boldsymbol{\mathcal{L}}}_0'}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_0'\right|} = \alpha\hat{\mathbf{e}}_{mn}^{\mathrm{TE}\prime} + \beta\hat{\mathbf{e}}_{mn}^{\mathrm{TM}\prime} , \tag{5.4}$$

where α and β are complex coefficients. Since the TE and TM polarization vectors are obtained by rotation, the coefficients α and β are unchanged in the new

coordinate system,

$$\frac{\widetilde{\boldsymbol{\mathcal{L}}}_{mn}}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_{mn}\right|} = \alpha\hat{\mathbf{e}}_{mn}^{\mathrm{TE}} + \beta\hat{\mathbf{e}}_{mn}^{\mathrm{TM}}\,. \tag{5.5}$$

Eq. (5.5) specifies how the Fourier components of a structured light distribution are polarized for a given incident polarization. The description in Eq. (5.5) allows to describe arbitrary polarization states, including linear, azimuthal, radial and elliptic polarization. In the following, these cases are outlined.

5.2.1 Azimuthal polarization

A pure TE-polarization is achieved for $\alpha = 1$ and $\beta = 0$. It can be seen directly from the form of $\hat{\mathbf{e}}_{mn}^{\mathrm{TE}\prime}$ in Eq. (5.3) that this is equivalent to an azimuthal polarization ($\mathbf{e}_{mn}^{\mathrm{TE}\prime} \mathrel{\hat{=}} \hat{\mathbf{e}}_{\varphi}$). In this case, from Eq. (5.5) one obtains

$$\frac{\widetilde{\boldsymbol{\mathcal{L}}}_{mn}}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_{mn}\right|} = \frac{1}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \begin{bmatrix} -k_{y;mn} \\ k_{x;mn} \\ 0 \end{bmatrix}\,. \tag{5.6}$$

Fig. 5.3 (left) shows an azimuthally polarized Gaussian profile.

5.2.2 Radial polarization

A pure TM-polarization is achieved for $\alpha = 0$ and $\beta = 1$. It can be seen directly from the form of $\hat{\mathbf{e}}_{mn}^{\mathrm{TM}\prime}$ in Eq. (5.3) that this is equivalent to a radial polarization ($\mathbf{e}_{mn}^{\mathrm{TM}\prime} \mathrel{\hat{=}} \hat{\mathbf{e}}_{r}$). In this case, from Eq. (5.5) one obtains

$$\frac{\widetilde{\boldsymbol{\mathcal{L}}}_{mn}}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_{mn}\right|} = \frac{1}{|\mathbf{k}_{mn}|} \frac{1}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \begin{bmatrix} k_{x;mn}k_{z;mn} \\ k_{y;mn}k_{z;mn} \\ -k_{x;mn}^2 - k_{y;mn}^2 \end{bmatrix}\,. \tag{5.7}$$

Fig. 5.3 (right) shows a radially polarized Gaussian profile.

5.2.3 Linear polarization

In order to achieve linear polarization, one requires

$$\frac{\widetilde{\boldsymbol{\mathcal{L}}}_{0}'}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_{0}'\right|} = \cos\psi\,\hat{\mathbf{e}}_x + \sin\psi\,\hat{\mathbf{e}}_y\,, \tag{5.8}$$

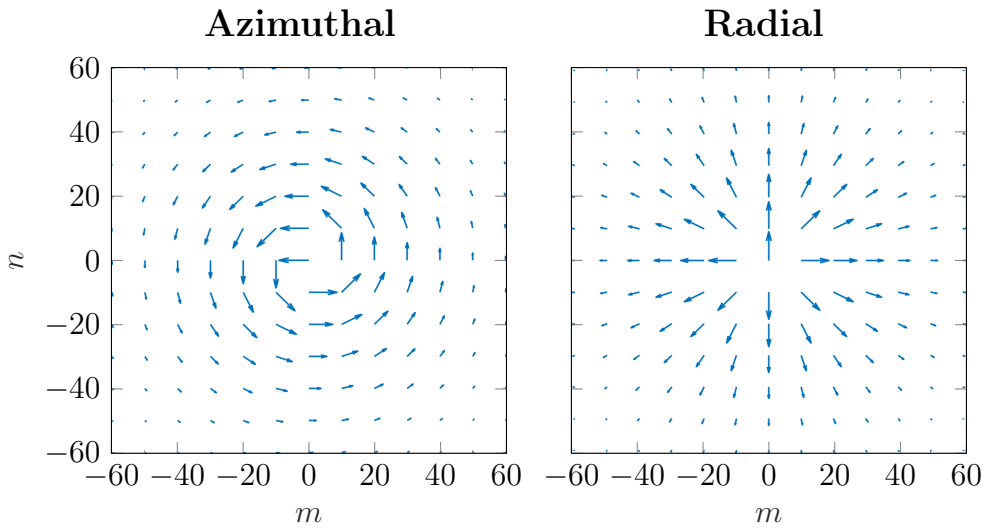


Figure 5.3: x-y-polarization vectors of a sample Gaussian illumination for azimuthal (left) and radial (right) polarization.

where ψ denotes the angle between the x-axis and the polarization direction. The coefficients α and β are then calculated via

$$\begin{aligned}\alpha &= \frac{\widetilde{\boldsymbol{\mathcal{L}}}_0'}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_0'\right|} \cdot \hat{\mathbf{e}}_{mn}^{\mathrm{TE}\prime} = \frac{-k_{y;mn}\cos\psi + k_{x;mn}\sin\psi}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} \\ \beta &= \frac{\widetilde{\boldsymbol{\mathcal{L}}}_0'}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_0'\right|} \cdot \hat{\mathbf{e}}_{mn}^{\mathrm{TM}\prime} = \frac{+k_{x;mn}\cos\psi + k_{y;mn}\sin\psi}{\sqrt{k_{x;mn}^2 + k_{y;mn}^2}} .\end{aligned} \tag{5.9}$$

With Eqs. (5.5, 5.9), one then obtains

$$\begin{aligned}\frac{\widetilde{\boldsymbol{\mathcal{L}}}_{mn}}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_{mn}\right|} =& \frac{-k_{y;mn}\cos\psi + k_{x;mn}\sin\psi}{k_{x;mn}^2 + k_{y;mn}^2} \begin{bmatrix} -k_{y;mn} \\ k_{x;mn} \\ 0 \end{bmatrix} \\ &+ \frac{1}{|\mathbf{k}_{mn}|} \frac{k_{x;mn}\cos\psi + k_{y;mn}\sin\psi}{k_{x;mn}^2 + k_{y;mn}^2} \begin{bmatrix} k_{x;mn}k_{z;mn} \\ k_{y;mn}k_{z;mn} \\ -k_{x;mn}^2 - k_{y;mn}^2 \end{bmatrix} .\end{aligned} \tag{5.10}$$

This result can be directly compared to Eq. (5) in [78]. Eq. (5.10) can be implemented for, for instance, a linearly polarized focused beam. Fig. 5.4 shows a linearly polarized Gaussian profile for $\psi = 0°$ (left) and $\psi = 45°$ (right).

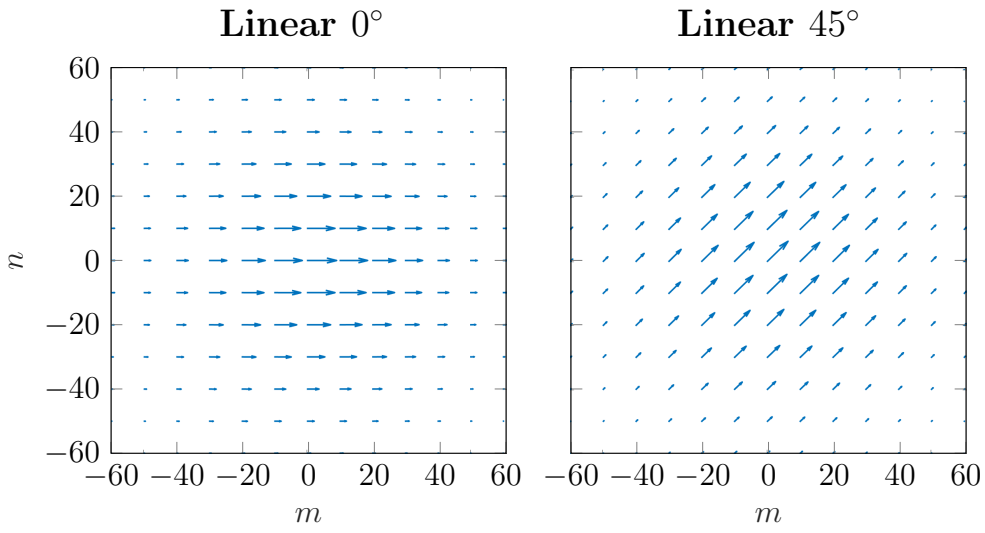


Figure 5.4: x-y-polarization vectors of a sample Gaussian illumination for 0°-linear (left) and 45°-linear (right) polarization. Only every 10^{th} vector is plotted in each dimension.

5.2.4 Elliptical / circular polarization

In the case of elliptical polarization, the incident polarization vector is

$$\frac{\widetilde{\boldsymbol{\mathcal{L}}}'_0}{\left|\widetilde{\boldsymbol{\mathcal{L}}}'_0\right|} = \cos\psi\,\hat{\mathbf{e}}_x + i\sin\psi\,\hat{\mathbf{e}}_y\,, \tag{5.11}$$

where the only difference to linear polarization is an additional imaginary unit in the y component. Accordingly, the result is very similar to linear polarization,

$$\begin{aligned}\frac{\widetilde{\boldsymbol{\mathcal{L}}}_{mn}}{\left|\widetilde{\boldsymbol{\mathcal{L}}}_{mn}\right|} =& \frac{-k_{y;mn}\cos\psi + ik_{x;mn}\sin\psi}{k_{x;mn}^2 + k_{y;mn}^2}\begin{bmatrix}-k_{y;mn}\\ k_{x;mn}\\ 0\end{bmatrix}\\ &+ \frac{1}{|\mathbf{k}_{mn}|}\frac{k_{x;mn}\cos\psi + ik_{y;mn}\sin\psi}{k_{x;mn}^2 + k_{y;mn}^2}\begin{bmatrix}k_{x;mn}k_{z;mn}\\ k_{y;mn}k_{z;mn}\\ -k_{x;mn}^2 - k_{y;mn}^2\end{bmatrix}.\end{aligned} \tag{5.12}$$

5.3 Bidirectional light incidence in RCWA

In this section it is described how bidirectional light incidence can be incorporated in the existing RCWA framework. Fig. 5.5 shows an L-layer structure with light incident from both sides, left *and* right. To be most general, a possible structuring of the incident and transmitted regions as described in Chapter 4 is included in the following derivation. The incident light is given in the form of the eigenmode coefficient vectors $\mathbf{c}_+^{(0)}$ and $\mathbf{c}_-^{(L+1)}$ of the incident and transmitted region. In the

case of homogeneous incident and transmitted regions, these coefficients are equal to the Fourier mode coefficients $\widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)}$ and $\widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)}$, respectively. It is clear that a

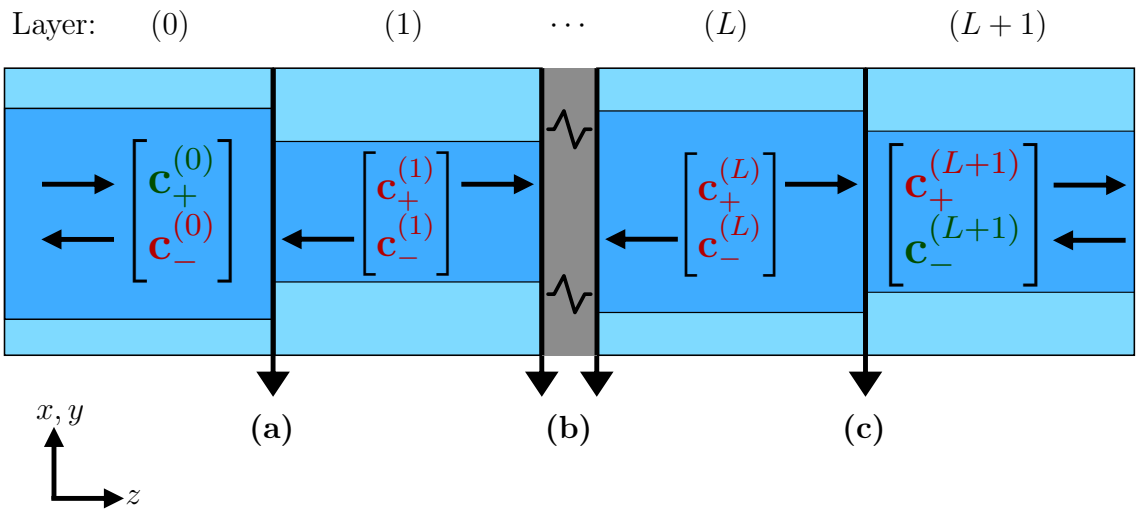


Figure 5.5: RCWA layout with a stack of L layers and light incident from both sides. The mode coefficients of the incident light, $\mathbf{c}_+^{(0)}$ and $\mathbf{c}_-^{(L+1)}$, are known. The reflected and transmitted mode coefficients, $\mathbf{c}_-^{(0)}$ and $\mathbf{c}_+^{(L+1)}$, as well as the eigenmode coefficients inside the grating layers, $\mathbf{c}_+^{(l)}$ and $\mathbf{c}_-^{(l)}$, are unknown.

change in the illumination only has an effect on the boundary conditions. Therefore, the modifications do not concern the eigenmode decomposition within the layers. The changes rather relate to the method, by which the boundary conditions are enforced. In the literature one can find various ways to solve the boundary value problem (cf. Sect. 3.5). The most prominent methods are the ETMA[11] and the S-matrix approach. The latter was first described in the context of RCWA by [10], a more modern description can be found in [66]. In principle, two-sided light incidence can be incorporated in all the described methods.

5.3.1 Adapting the S-matrix approach

In the S-matrix approach, the transition of light through the l^{th} layer is modeled by the so-called S-Matrix. As described in Sect. 3.5.2, the latter connects the amplitudes of the forward and backward propagating eigenmodes of the previous and successive layer via[66]

$$\begin{bmatrix} \mathbf{c}_-^{(l-1)} \\ \mathbf{c}_+^{(l+1)} \end{bmatrix} = \begin{bmatrix} \hat{S}_{11}^{(l)} & \hat{S}_{12}^{(l)} \\ \hat{S}_{21}^{(l)} & \hat{S}_{22}^{(l)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(l-1)} \\ \mathbf{c}_-^{(l+1)} \end{bmatrix} . \tag{5.13}$$

Furthermore, the global S-matrix of a multi-layer stack can be obtained from the individual S-matrices of the layers via the Redheffer star product, cf. Sect. 3.5.2

and [66]. The global S-matrix then connects the vector of unknowns to the vector of knowns, i.e., the injected light at the entry and exit of the multi-layer stack. The amplitude of the back-propagating wave at the stack exit is commonly set to zero. In order to account for bidirectional illumination, this vector must be changed from zero to the amplitudes of the injected light modes. In Eq. (5.14) this change is highlighted in dark blue,

$$\begin{bmatrix} \mathbf{c}_{-}^{(0)} \\ \mathbf{c}_{+}^{(L+1)} \end{bmatrix} = \begin{bmatrix} \hat{S}_{11}^{\text{global}} & \hat{S}_{12}^{\text{global}} \\ \hat{S}_{21}^{\text{global}} & \hat{S}_{22}^{\text{global}} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(0)} \\ \mathbf{c}_{-}^{(L+1)} \end{bmatrix} . \tag{5.14}$$

In the case of homogeneous incident and transmitted layers, the relation simplifies to

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} \end{bmatrix} = \begin{bmatrix} \hat{S}_{11}^{\text{global}} & \hat{S}_{12}^{\text{global}} \\ \hat{S}_{21}^{\text{global}} & \hat{S}_{22}^{\text{global}} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)} \end{bmatrix} . \tag{5.15}$$

Mathematically, this change only means that the zero vector is replaced by a non-zero vector. Thus, the boundary value problem is equally solved for coherent bidirectional illumination. Moreover, the numerical complexity and the memory requirement are the same as before, which can seen from the comparison with Sect. 3.7 and [66]. The same S-matrix is used for both illumination situations. Therefore, the computation time for bidirectional illumination in the S-matrix approach is practically the same as for one-sided illumination.

5.3.2 Adapting the ETMA approach

In the following, the so-called 'partial solution approach' of the ETMA as first proposed by [11] and presented in Sect. 3.5.4 is modified, which in its original version only applies for one-sided light incidence. Furthermore, it is extended to structured incident and transmitted regions as suggested in Chapter 4. Even though the name of the approach suggests otherwise, the algorithm provides the full solution of the boundary value problem. First, the lateral electric- and magnetic field components, which are continuous in z-direction across the layer interfaces, are set equal. To this end, one calculates the electric and magnetic field mode coefficient vectors at the layer boundaries using Eq. (3.59). At the leftmost interface, Fig. 5.5a, the continuity condition reads

$$\begin{bmatrix} +\hat{W}^{(0)} & +\hat{W}^{(0)} \\ +\hat{V}^{(0)} & -\hat{V}^{(0)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(0)} \\ \mathbf{c}_{-}^{(0)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(1)} & +\hat{W}^{(1)}\hat{X}^{(1)} \\ +\hat{V}^{(1)} & -\hat{V}^{(1)}\hat{X}^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(1)} \\ \mathbf{c}_{-}^{(1)} \end{bmatrix} . \tag{5.16}$$

At an intermediate interface between two layers $l-1$ and l, Fig. 5.5b, one gets

$$\begin{bmatrix} +\hat{W}^{(l-1)}\hat{X}^{(l-1)} & +\hat{W}^{(l-1)} \\ +\hat{V}^{(l-1)}\hat{X}^{(l-1)} & -\hat{V}^{(l-1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(l-1)} \\ \mathbf{c}_{-}^{(l-1)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(l)} & +\hat{W}^{(l)}\hat{X}^{(l)} \\ +\hat{V}^{(l)} & -\hat{V}^{(l)}\hat{X}^{(l)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(l)} \\ \mathbf{c}_{-}^{(l)} \end{bmatrix} . \tag{5.17}$$

At the rightmost interface, Fig. 5.5c, one obtains

$$\begin{bmatrix} +\hat{W}^{(L)}\hat{X}^{(L)} & +\hat{W}^{(L)} \\ +\hat{V}^{(L)}\hat{X}^{(L)} & -\hat{V}^{(L)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(L)} \\ \mathbf{c}_-^{(L)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(L+1)} & +\hat{W}^{(L+1)} \\ +\hat{V}^{(L+1)} & -\hat{V}^{(L+1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(L+1)} \\ \mathbf{c}_-^{(L+1)} \end{bmatrix} . \tag{5.18}$$

Starting from the rightmost interface, Eq. (5.18) can be written as

$$\begin{bmatrix} +\hat{W}^{(L)}\hat{X}^{(L)} & +\hat{W}^{(L)} \\ +\hat{V}^{(L)}\hat{X}^{(L)} & -\hat{V}^{(L)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(L)} \\ \mathbf{c}_-^{(L)} \end{bmatrix} = \begin{bmatrix} \hat{f}^{(L+1)} \\ \hat{g}^{(L+1)} \end{bmatrix} \mathbf{c}_+^{(L+1)} + \begin{bmatrix} \mathbf{p}^{(L+1)} \\ \mathbf{q}^{(L+1)} \end{bmatrix} , \tag{5.19}$$

where one identifies $\hat{f}^{(L+1)} = \hat{W}^{(L+1)}$, $\hat{g}^{(L+1)} = \hat{V}^{(L+1)}$, $\mathbf{p}^{(L+1)} = \hat{W}^{(L+1)}\mathbf{c}_-^{(L+1)}$ and $\mathbf{q}^{(L+1)} = -\hat{V}^{(L+1)}\mathbf{c}_-^{(L+1)}$. Rearrange Eq. (5.19),

$$\begin{bmatrix} -\hat{W}^{(L)} & \hat{f}^{(L+1)} \\ \hat{V}^{(L)} & \hat{g}^{(L+1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_-^{(L)} \\ \mathbf{c}_+^{(L+1)} \end{bmatrix} = \begin{bmatrix} \hat{W}^{(L)}\hat{X}^{(L)} \\ \hat{V}^{(L)}\hat{X}^{(L)} \end{bmatrix} \mathbf{c}_+^{(L)} - \begin{bmatrix} \mathbf{p}^{(L+1)} \\ \mathbf{q}^{(L+1)} \end{bmatrix} . \tag{5.20}$$

From Eq. (5.20) one obtains

$$\mathbf{c}_-^{(L)} = \hat{a}^{(L)}\mathbf{c}_+^{(L)} + \mathbf{r}^{(L)} , \tag{5.21a}$$

$$\mathbf{c}_+^{(L+1)} = \hat{b}^{(L)}\mathbf{c}_+^{(L)} + \mathbf{s}^{(L)} , \tag{5.21b}$$

where

$$\begin{bmatrix} \hat{a}^{(L)} \\ \hat{b}^{(L)} \end{bmatrix} = + \begin{bmatrix} -\hat{W}^{(L)} & \hat{f}^{(L+1)} \\ \hat{V}^{(L)} & \hat{g}^{(L+1)} \end{bmatrix}^{-1} \begin{bmatrix} \hat{W}^{(L)}\hat{X}^{(L)} \\ \hat{V}^{(L)}\hat{X}^{(L)} \end{bmatrix} , \tag{5.21c}$$

$$\begin{bmatrix} \mathbf{r}^{(L)} \\ \mathbf{s}^{(L)} \end{bmatrix} = - \begin{bmatrix} -\hat{W}^{(L)} & \hat{f}^{(L+1)} \\ \hat{V}^{(L)} & \hat{g}^{(L+1)} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{p}^{(L+1)} \\ \mathbf{q}^{(L+1)} \end{bmatrix} . \tag{5.21d}$$

Eq. (5.21a) is now being substituted back into Eq. (5.17) for $l = L$, which describes the boundary conditions between the second last and last layer,

$$\begin{bmatrix} +\hat{W}^{(L-1)}\hat{X}^{(L)} & +\hat{W}^{(L-1)} \\ +\hat{V}^{(L-1)}\hat{X}^{(L)} & -\hat{V}^{(L-1)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_+^{(L-1)} \\ \mathbf{c}_-^{(L-1)} \end{bmatrix} = \begin{bmatrix} \hat{f}^{(L)} \\ \hat{g}^{(L)} \end{bmatrix} \mathbf{c}_+^{(L)} + \begin{bmatrix} \mathbf{p}^{(L)} \\ \mathbf{q}^{(L)} \end{bmatrix} , \tag{5.22a}$$

where

$$\begin{bmatrix} \hat{f}^{(L)} \\ \hat{g}^{(L)} \end{bmatrix} = \begin{bmatrix} \hat{W}^{(L)} \left(\mathbb{1} + \hat{X}^{(L)}\hat{a}^{(L)} \right) \\ \hat{V}^{(L)} \left(\mathbb{1} - \hat{X}^{(L)}\hat{a}^{(L)} \right) \end{bmatrix} , \tag{5.22b}$$

$$\begin{bmatrix} \mathbf{p}^{(L)} \\ \mathbf{q}^{(L)} \end{bmatrix} = \begin{bmatrix} +\hat{W}^{(L)}\hat{X}^{(L)}\mathbf{r}^{(L)} \\ -\hat{V}^{(L)}\hat{X}^{(L)}\mathbf{r}^{(L)} \end{bmatrix} . \tag{5.22c}$$

One observes that Eq. (5.22a) is virtually the same as Eq. (5.19), but with decremented index. The procedure [Eqs. (5.19-5.22)] is, therefore, repeated for the remaining layer interfaces in decreasing order until the leftmost interface, where one is left with

$$\begin{bmatrix} +\hat{W}^{(0)} & +\hat{W}^{(0)} \\ +\hat{V}^{(0)} & -\hat{V}^{(0)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{+}^{(0)} \\ \mathbf{c}_{-}^{(0)} \end{bmatrix} = \begin{bmatrix} \hat{f}^{(1)} \\ \hat{g}^{(1)} \end{bmatrix} \mathbf{c}_{+}^{(1)} + \begin{bmatrix} \mathbf{p}^{(1)} \\ \mathbf{q}^{(1)} \end{bmatrix}. \tag{5.23}$$

This can finally be solved for $\mathbf{c}_{-}^{(0)}$ and $\mathbf{c}_{+}^{(1)}$ via

$$\begin{bmatrix} \mathbf{c}_{-}^{(0)} \\ \mathbf{c}_{+}^{(1)} \end{bmatrix} = \begin{bmatrix} -\hat{W}^{(0)} & \hat{f}^{(1)} \\ \hat{V}^{(0)} & \hat{g}^{(1)} \end{bmatrix}^{-1} \left(\begin{bmatrix} \hat{W}^{(0)} \\ \hat{V}^{(0)} \end{bmatrix} \mathbf{c}_{+}^{(0)} - \begin{bmatrix} \mathbf{p}^{(1)} \\ \mathbf{q}^{(1)} \end{bmatrix} \right). \tag{5.24}$$

Eq. (5.21) and its equivalents for the other layers can subsequently be used to determine the remaining unknowns $\mathbf{c}_{+}^{(l)}$, $\mathbf{c}_{-}^{(l)}$ iteratively.

In terms of the numerical complexity, it can be seen from the comparison with Sect. 3.5.4 and [11] that $\mathbf{p}^{(l)}$, $\mathbf{q}^{(l)}$, $\mathbf{r}^{(l)}$ and $\mathbf{s}^{(l)}$ are the only newly introduced quantities. These quantities are vectors with a length determined by the number of modes, $2\bar{N}$, considered in the calculation. On the other hand, the matrices $\hat{a}^{(l)}$, $\hat{b}^{(l)}$,$\hat{f}^{(l)}$, $\hat{g}^{(l)}$, which are already needed in the original one-sided algorithm, are $2\bar{N} \times 2\bar{N}$-matrices. Thus, the treatment of bidirectional illumination does not add any substantial memory or time complexity to the case of single sided illumination. Furthermore, it can be seen from Eqs. (5.21c, 5.21d) that the matrix inversion can be reused, thus the computational effort is basically unchanged.

5.4 Application to the 4π-microscope

In the following, the algorithm presented above is applied to 4π-microscopy. The principle was already illustrated in the motivation of this chapter, Fig. 5.1. The latter is a variation of confocal microscopy that was introduced in the 1990s with the intent to increase the axial resolution for thick samples. The basics are well explained in [76, 77]. The general idea is to employ two coherent counter-propagating converging waves as a light source instead of applying focused light only from one side. Since in most cases the objects of interest are in the sub-wavelength regime, a scalar treatment is not sufficiently accurate. Therefore, exact optical methods are considered as relevant for the simulation of 4π-microscopy. Fig. 5.6 illustrates an x-z slice through the undisturbed foci produced by both one-sided (2π) and coherent two-sided (4π) illumination. In this example, the 2π illumination generates one single diffraction limited focal spot with the well-known lateral and axial extents proportional to λ/NA and λ/NA^2[82]. Bidirectional 4π-illumination results in a standing wave interference pattern, which exhibits an axial intensity modulation with a sharp $\lambda/4$ sized central peak and two side lobes. The position of the excitation focus can

be controlled by shifting the relative phase of the two illumination beams, cf. [76, 77].

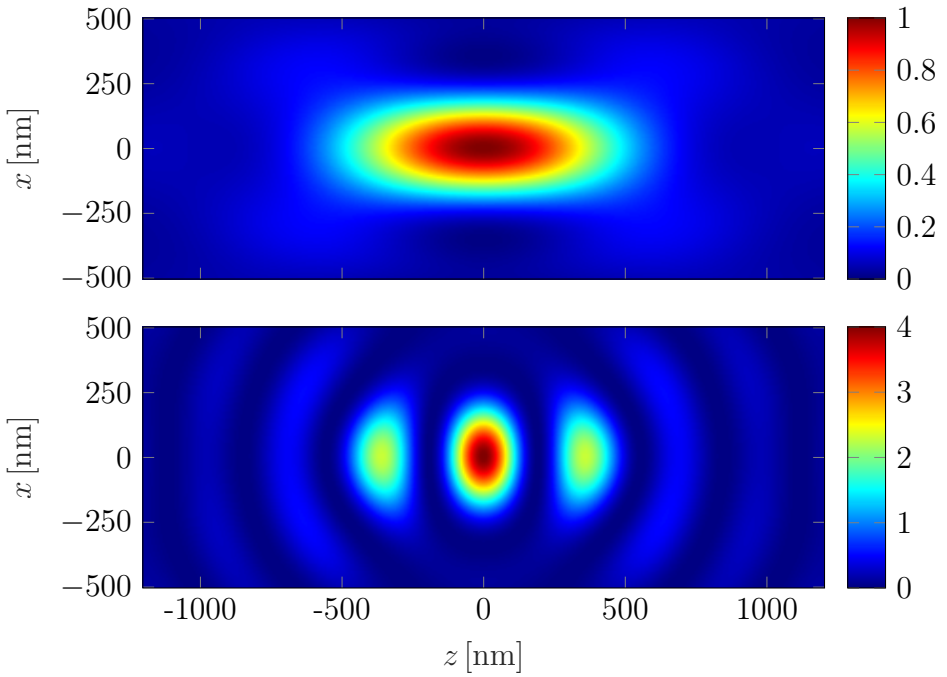


Figure 5.6: Intensity of the undisturbed foci (central x-z slice) with one-sided 2π- (top) and coherent two-sided 4π-illumination (bottom). The incident electric field is linearly polarized in y-direction. Number of modes $M = N = 41$, grid period $P_x = P_y = 10\,\mu\text{m}$, wavelength $\lambda = 800\,\text{nm}$, surrounding material water $n = 1.33$, NA $= 1.28$, linear polarization angle $\psi = 90°$.

Consider the situation illustrated in Fig. 5.7. In 4π-microscopy, certain regions of the sample are prepared with fluorescence markers that are subsequently excited by the illumination. For detection, only the global fluorescent light is measured[83]. For testing the position and resolution accuracy, one assumes a microscopic ovoid protein sample suspended in water, which contains a small sub-λ region prepared with fluorescence markers at its right boundary. Since the fluorescence molecule produces a local absorption of the excitation light, the fluorescence markers are modeled by a positive imaginary part in the local permittivity $\epsilon(\vec{r})$.

First, it is investigated how the form of the focus changes in the presence of the sample. To this end, the region prepared with the fluorescence markers is positioned at the center of the undisturbed 4π-focus. An x-z slice through the aberrated 4π-focus is shown in Fig. 5.8. It can be observed that the aberrations introduced by the sample significantly change the form of the focus. On the one hand, the amplitude of the central peak is decreased and intensity is spilled into the side lobes, which are no longer arranged symmetrically around the central focus. On the other hand, the

focus position is shifted in z-direction by more than 100 nm out of its original central position.

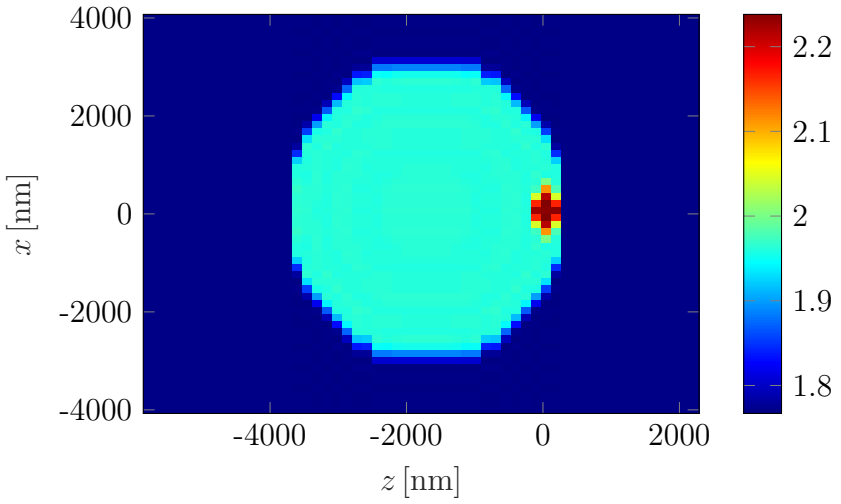


Figure 5.7: Spatial permittivity distribution $\epsilon(\vec{r})$ (central x-z slice), which consists of an ovoid protein envelope with axis diameters $d_{p,x} \times d_{p,y} \times d_{p,z} = 4\,\mu\text{m} \times 6\,\mu\text{m} \times 4\,\mu\text{m}$ ($\epsilon_p = 1.96$) in water ($\epsilon_w = 1.77$), containing another small ovoid structure with axis diameters $d_{fl,x} \times d_{fl,y} \times d_{fl,z} = 0.8\,\mu\text{m} \times 0.8\,\mu\text{m} \times 0.4\,\mu\text{m}$ prepared with fluorescence markers ($\epsilon_p = 2.24 + 0.3i$). For the calculation, the sample is divided into 28 layers and Gaussian apodization is applied to the Fourier components of $\epsilon(\vec{r})$.

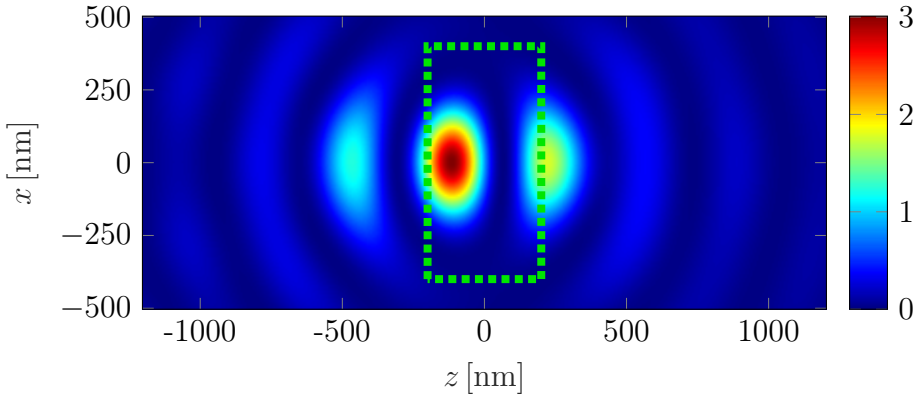


Figure 5.8: Intensity of the aberrated focus generated by 4π-illumination of the sample shown in Fig. 5.7 (central x-z slice, parameters as before). The focus form and peak height are significantly changed (cf. Fig. 5.6 bottom). The dashed green box indicates the position and size of the fluorescent structure. Clearly, the focus is not located centrally inside the sample.

In a second step, the sample is shifted along the z-axis, and the amount of fluorescence light at several positions is determined. To this end, one assumes that the

fluorescence light power is proportional to the light power that is absorbed in the sample[84]. Fig. 5.9 shows the resulting curve (solid red line) and for comparison the same curve for the sample consisting of only the fluorescent structure without the protein envelope (blue dotted line). Clearly, the peak is shifted by approximately 100 nm compared to the case without protein envelope and, additionally, the form of the peak is changed. This means that the small protein structure introduces significant aberrations, which affect the measurement. The authors of [85] address the problems generated by sample-induced aberrations in thick biological samples and propose to implement an adaptive optics scheme in the experimental setup. With the combination of the above concepts, these kinds of aberrations can now be consistently simulated and, therefore, be better understood and avoided.

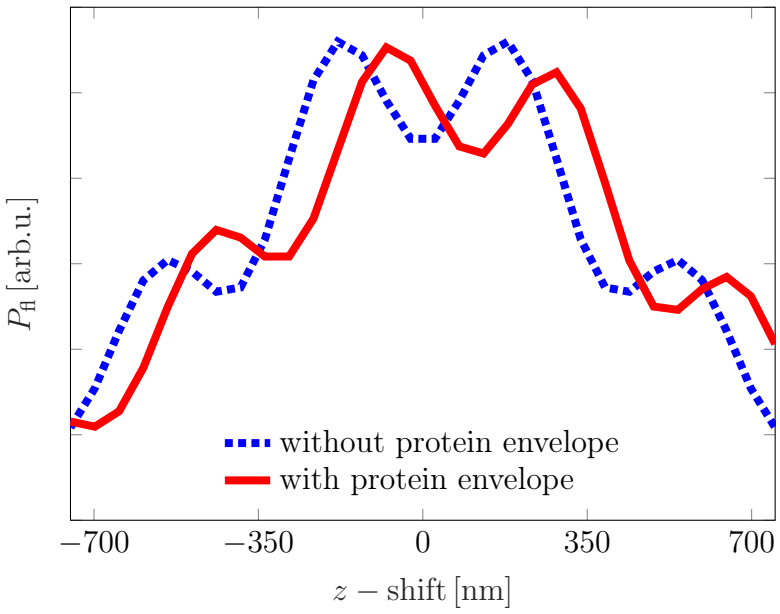


Figure 5.9: Fluorescence light power P_{fl} vs. z-shift of the sample with (solid red) and without (blue dotted) protein envelope.

5.5 Conclusions

In this chapter, the RCWA is applied to the simulation of 4π-microscopy. To this end, various modifications are introduced. For the focused illumination, the localized field approach by Auer and Brenner[17] is adopted for a fast and consistent handling of arbitrarily structured illumination. Furthermore, a derivation for describing the polarization state of an arbitrary structured illumination is provided. As an example, the expressions for radial, azimuthal, linear and elliptic polarization are derived. Finally, the introduction of an additional backside illumination into the existing framework of the RCWA is presented. As an application, it is investigated how the focus in 4π-microscopy is aberrated due to the presence of clustered protein

structures. One observes that, if not corrected, this may result in a position shift of the fluorescence peak in the range of several hundred nanometers. Hereby, it is shown that sample-induced aberrations can be simulated with an exact method for realistic 4π-illumination scenarios. Therefore, it is expected that this enables to better understand and correct the aberrations occurring in these types of systems in the future.

6 | Fast rigorous iterative method

This chapter is partly based on the publications [A5, A6].

6.1 Motivation and structure

The application of rigorous optical simulation algorithms is limited to the nanooptical scale due to severe computing time and memory constraints. This is true even for today's high performance computers. To address this problem, the FRIM is presented in this chapter, an algorithm that reduces the numerical complexity of modal methods from $\mathcal{O}(\bar{N}^3)$ to $\mathcal{O}(\bar{N}\log\bar{N})$ under certain conditions. At the same time, the memory requirement decreases from $\mathcal{O}(\bar{N}^2)$ to $\mathcal{O}(\bar{N})$. The FRIM is based on circumventing the numerically complex eigenmode decomposition, which is responsible for the $\mathcal{O}(\bar{N}^3)$ dependence, by replacing this operation by a sequence of efficient matrix multiplications. Apart from speed, another major advantage of the iterative FRIM is the possibility to trade accuracy versus calculation time at the desired mode count, which is not possible in standard modal methods. In the latter, the only possible trade-off is a reduction of the number of modes considered in the calculation. In the following, it is shown that, with the FRIM, structures like for example certain large sized DOEs can be simulated exactly with a high mode count in the order of 2000×2000 as shown in Table 6.1. This table illustrates the reduction in calculation time and memory space achieved with the FRIM in direct comparison to the RCWA.

Also the limitations of the FRIM have to be considered, which concern the layer thickness, the NA, and the amount of absorption. It is shown that only objects with thicknesses up to a few wavelengths can be simulated. Second, the NA of the calculation should be moderate, i.e., $\mathrm{NA}_{\mathrm{calc}} \leq n_{\mathrm{min}}$, where n_{min} is the smallest refractive index in the grating structure. Third, only weak absorption should be present in the object medium. Within the scope of these limitations, the FRIM can be considered as an ideal method for the exact electromagnetic simulation of large sized DOEs, which in most cases are presently modeled by treating the DOE in the TEA with a subsequent scalar AS propagation. Especially at higher incident angles θ, the TEA is known to loose its validity, producing unreliable results with no satisfactory alter-

Table 6.1: Calculation time and memory demand of the RCWA and the FRIM for different mode counts.

	Calculation time		Memory demand (RAM)	
$\bar{N}$	**RCWA**	**FRIM**	**RCWA**	**FRIM**
$\mathbf{20^2}$	5 s	6 s	47 MB	27 kB
$\mathbf{100^2}$	8 h	46 s	25 GB	638 kB
$\mathbf{200^2}$	22 d	279 s	389 GB	2.5 MB
$\mathbf{2000^2}$	60 000 yr	11 h	3822 TB	244 MB

For the RCWA, the numbers at $\bar{N} = 200^2$ and $\bar{N} = 2000^2$ are extrapolated values, where the calculation time and the memory demand are assumed to grow with $\mathcal{O}(\bar{N}^3)$ and $\mathcal{O}(\bar{N}^2)$, respectively. All the other values originate from actual calculations conducted with an off-the-shelf hexacore server with 3.70 GHz clock rate and 64 GB RAM.

native simulation method being available. In this chapter, the differences between simulations conducted with the FRIM and the TEA are investigated. In the cases, for which the TEA is known to fail, the FRIM is compared to the WPM[52] in order to validate the results. Finally, it is illustrated that the FRIM becomes especially useful in cases, where a small scale periodicity, such as in gratings, is not present and, therefore, large calculation domains are inevitable. As an example, a numerical simulation of the Zernike phase contrast method is performed.

This chapter is organized as follows. In Sect. 6.2, the iteration scheme of the FRIM is derived. In Sect. 6.3, it is shown how the proposed algorithm can be efficiently implemented at a numerical cost of $\mathcal{O}\big(\bar{N}\log\bar{N}\big)$ and with a memory requirement of only $\mathcal{O}\big(\bar{N}\big)$. Sect. 6.4 investigates the convergence properties of the FRIM. On the one hand, a strict mathematical convergence criterion is developed. It is derived that the convergence behavior is exponential and shown how the number of necessary iterations can be predicted before starting the calculation. On the other hand, an empiric convergence condition is given in terms of the grating thickness, the object's absorption properties and the NA of the calculation. In Sect. 6.5, several simulation results are shown in order to validate the FRIM. As a first test, the FRIM is directly compared to the RCWA for a small sized DOE in order to verify that equal results are obtained. In the second part, a large size Fresnel zone plate (FZP) out of fused silica and with a phase height difference of π is simulated with the FRIM. The latter cannot be simulated with the RCWA due to time and memory constraints. The result is subsequently validated by both testing the energy conservation properties and by a comparison to a scalar simulation treating the grating in TEA and a subsequent AS propagation. In the third part, the FRIM is compared to a WPM simulation for the case of a FZP with a phase height difference of 2π. This is especially interesting because the intensity pattern behind the element originates only from diffraction at the lateral interfaces, whereas the TEA would act as if no DOE were present at all. In the fourth part, the Zernike phase contrast method is rigorously simulated with

the FRIM. Finally, in Sect. 6.6, the FRIM is extended to the treatment of linear anisotropic media.

6.2 Iteration scheme

In this section, an iterative algorithm for the full rigorous optical simulation of spatially structured objects is developed. In the derivation of the RCWA, Sect. 3.3, the propagation of the electromagnetic field within one grating layer slab was found to be described by the first order differential equation[9]

$$\frac{\partial}{\partial z}\left\{\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp\end{bmatrix}(z)\right\} = ik_0\begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp\end{bmatrix}(z)\ , \tag{6.1}$$

where $\widetilde{\boldsymbol{\mathcal{E}}}_\perp$ and $\widetilde{\boldsymbol{\mathcal{H}}}_\perp$ denote the Fourier coefficients of the lateral electric and magnetic field components, and $\hat{M}_{12}^{(l)}$ and $\hat{M}_{21}^{(l)}$ are the mode coupling matrices of the l^{th} grating layer. The general solution of Eq. (6.1) can be expressed as a matrix exponential,

$$\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp\end{bmatrix}(z) = \exp\left\{ik_0\begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}z\right\}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp\end{bmatrix}(0)\ . \tag{6.2}$$

Hence, Eq. (6.2) can be used to back-propagate the electromagnetic field across the respective layer via

$$\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp;\text{left}}^{(l)}\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp;\text{left}}^{(l)}\end{bmatrix} = \exp\left\{-ik_0\begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}d^{(l)}\right\}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp;\text{right}}^{(l)}\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp;\text{right}}^{(l)}\end{bmatrix}\ , \tag{6.3}$$

$d^{(l)}$ being the thickness of the l^{th} layer. Consequently, the electromagnetic fields at the left and right interface of the entire L-layer grating are connected via

$$\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp;\text{left}}^{(1)}\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp;\text{left}}^{(1)}\end{bmatrix} = \prod_{l=1}^{L}\exp\left\{-ik_0\begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}d^{(l)}\right\}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_{\perp;\text{right}}^{(l)}\\ \widetilde{\boldsymbol{\mathcal{H}}}_{\perp;\text{right}}^{(l)}\end{bmatrix}\ . \tag{6.4}$$

To obtain the most generality, as introduced in Chapter 5, also here it is assumed that light may be incident from both the left and the right onto the grating. In terms of the incident ($\widetilde{\boldsymbol{\mathcal{L}}}_\perp^{(0/L+1)}$), reflected ($\widetilde{\boldsymbol{\mathcal{R}}}_\perp$) and transmitted ($\widetilde{\boldsymbol{\mathcal{T}}}_\perp$) Fourier mode

[9] The algorithm also works with anisotropic permittivities / permeabilities, for which all blocks of the coupling matrix are fully occupied (cf. Sect. 3.3.2). This extension is shown in Sect. 6.6 of this chapter.

coefficients, Eq. (6.4) reads

$$\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix} = \prod_{l=1}^{L} \exp\left\{ -ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} d^{(l)} \right\} \times \begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(L+1)} & -\hat{C}^{(L+1)} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)} \end{bmatrix}, \tag{6.5}$$

where $\hat{C}^{(0/L+1)}$ are the conversion matrices from the electric- to the magnetic field Fourier mode coefficients in the homogeneous incident and transmitted regions as given in Sect. 3.5. Eq. (6.5) can now be solved for either $\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} & \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix}^T$ or $\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} & \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)} \end{bmatrix}^T$, whichever is needed,

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix} = \hat{\Gamma} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)} \end{bmatrix} \tag{6.6a}$$

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{T}}}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)} \end{bmatrix} = \hat{\Gamma}^{-1} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix}, \tag{6.6b}$$

where

$$\hat{\Gamma} = \underbrace{\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix}^{-1}}_{*\mathrm{c}} \underbrace{\prod_{l=1}^{L} \exp\left\{ -ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} d^{(l)} \right\}}_{*\mathrm{b}} \times \underbrace{\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(L+1)} & -\hat{C}^{(L+1)} \end{bmatrix}}_{*\mathrm{a}} \tag{6.7a}$$

$$\hat{\Gamma}^{-1} = \begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(L+1)} & -\hat{C}^{(L+1)} \end{bmatrix}^{-1} \prod_{l=L}^{1} \exp\left\{ +ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} d^{(l)} \right\} \times \begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix}. \tag{6.7b}$$

The matrices $\hat{\Gamma}$ and $\hat{\Gamma}^{-1}$ are each a product of three different kinds of matrices denoted by (*a), (*b) and (*c). Subsequently, Eqs. (6.6) are used to construct the FRIM iteration scheme, which is illustrated in Fig. 6.1.

1. On the left side of the grating, initialize the iteration with the stacked vector $\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp;0}^{(0)} & \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;0} \end{bmatrix}^T = \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)} & \widetilde{\boldsymbol{\mathcal{R}}}_{\perp}^{\mathrm{aprx}} \end{bmatrix}^T$, where the lower indices denote the iteration index. $\widetilde{\boldsymbol{\mathcal{R}}}_{\perp}^{\mathrm{aprx}}$ is a first approximation to the reflected mode coefficient vector[10].

[10] One good choice for the initial $\widetilde{\boldsymbol{\mathcal{R}}}_{\perp}^{\mathrm{aprx}}$ can, for instance, be obtained from the back propagation of $\widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)}$ through the grating in TEA.

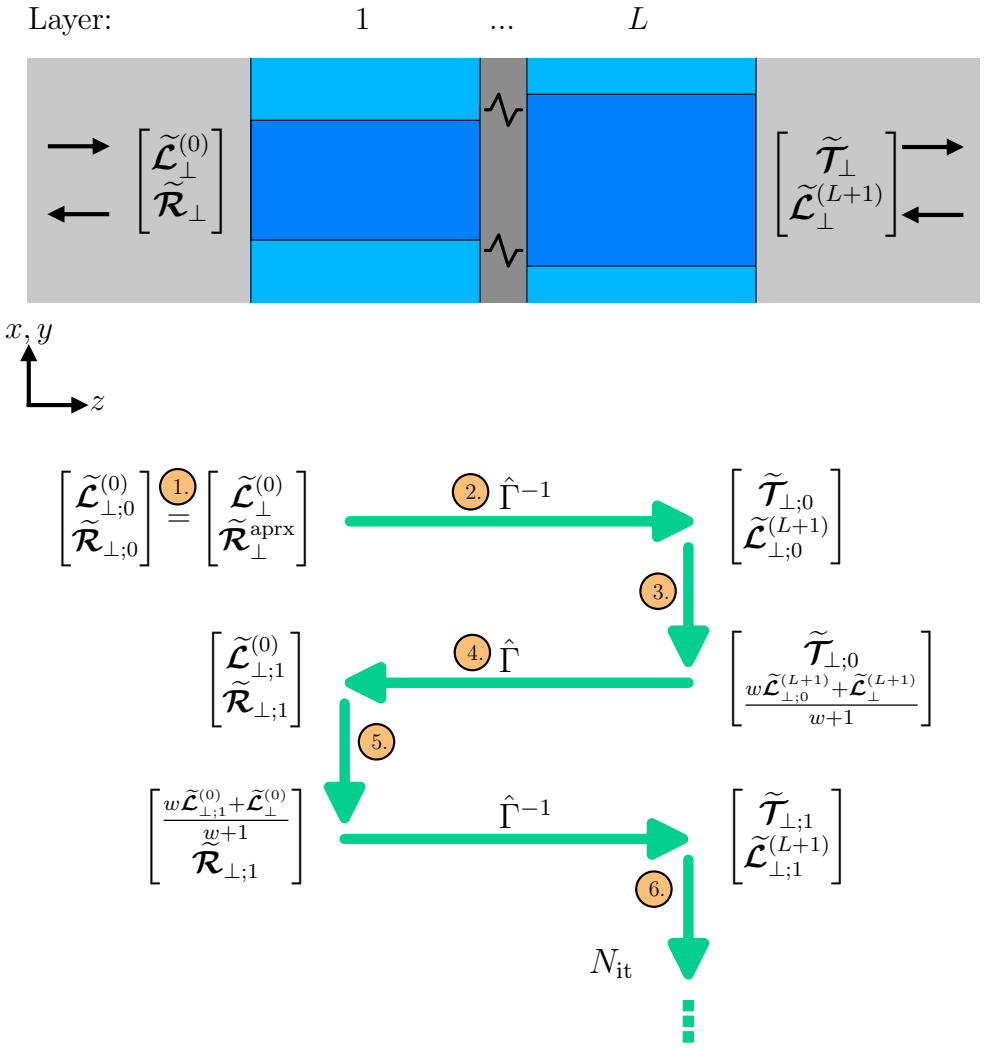


Figure 6.1: The FRIM iteration scheme alternately applies $\hat{\Gamma}$ and $\hat{\Gamma}^{-1}$ to propagate the mode coefficient vectors back and forth across the grating while partially mixing in the true incident light coefficients at each side given some mixing parameter $w \in \mathbb{R}_0^+$.

2. Apply $\hat{\Gamma}^{-1}$ to propagate the mode coefficient vector through the structure to the right grating interface.

3. Mix the lower half of the resulting mode coefficient vector with the true incident light coefficients $\widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(L+1)}$.

4. Apply $\hat{\Gamma}$ to propagate the modified mode coefficient vector back through the structure to the left grating interface.

5. Mix the upper half of the resulting mode coefficient vector with the true incident light coefficients $\widetilde{\boldsymbol{\mathcal{L}}}_{\perp}^{(0)}$.

6. Repeat steps 2-5 until reaching convergence. The total number of iterations is denoted by N_{it}.

6.3 Fast and memory efficient implementation

In the following, it is shown how the iteration scheme derived in Sect. 6.2 can be implemented fast and at the same time memory efficient by conducting efficient matrix multiplications. It is proven that the proposed method has a numerical complexity of the order $\mathcal{O}\big(\bar{N}\log\bar{N}\big)$ and a memory requirement of the order $\mathcal{O}\big(\bar{N}\big)$, $\bar{N}$ being the total number of modes in the calculation. These complexities are, in principle, the same as for conducting a scalar AS propagation.

The numerical performance of the iteration scheme's implementation in terms of memory efficiency and calculation time mainly relies on the ability to efficiently multiply an arbitrary known vector to both $\hat{\Gamma}$ and $\hat{\Gamma}^{-1}$. In the following, it is shown that these multiplications can be conducted such that the calculation time is of the order $\mathcal{O}\big(\bar{N}\log\bar{N}\big)$, and that the memory requirement is of the order $\mathcal{O}\big(\bar{N}\big)$. In the following, only the efficient multiplication of the matrix $\hat{\Gamma}$ is explained, since the same procedure also applies to the multiplication of $\hat{\Gamma}^{-1}$. As shown in Eq. (6.7), the matrix $\hat{\Gamma}$ is made up of a multiplication of three separate matrices. In order to multiply the entire matrix to a known vector, one has to conduct a consecutive multiplication of all three constituent matrices.

6.3.1 First constituent matrix

The efficient multiplication of the first constituent matrix (*a),

$$\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(L+1)} & -\hat{C}^{(L+1)} \end{bmatrix}, \tag{6.8}$$

to a known vector is straightforward, since $\mathbb{1}$ is the identity matrix, and the matrix $\hat{C}^{(L+1)}$ is by definition a 2×2 block diagonal matrix. Clearly, this multiplication can be implemented such that both the memory requirement and the numerical complexity of the involved operations are of the order $\mathcal{O}\big(\bar{N}\big)$.

6.3.2 Second constituent matrix

The second constituent matrix consists of a product of matrix exponentials (*b), which need to be multiplied to a known vector. Ref. [86] summarizes several different possibilities to compute the result numerically. In the following, two different ways are presented - the first via a Taylor series expansion and the second via the 'Runge-Kutta' (RK) method. However, mind that there still exist numerous alternative numerical integration schemes that could be used here. In principle, the applied integration scheme only plays a minor role, if run at a sufficient accuracy level. Though, differences may be observed in the runtime of the algorithm.

Taylor series expansion

The most direct and therefore easiest way is to expand the exponential function into a Taylor series via

$$\exp\left\{-ik_0\begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}d^{(l)}\right\} \approx \sum_{n=0}^{N_\mathrm{o}} \frac{\left(-ik_0 d^{(l)}\right)^n}{n!}\begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}^n . \tag{6.9}$$

The expansion is truncated at some adequate order N_o. Given the desired accuracy of the calculation, the necessary order N_o can be obtained as follows. First, calculate the eigenvalue of the matrix

$$ik_0\hat{M}d^{(l)} = \begin{bmatrix}\mathbb{0} & \hat{M}_{12}^{(l)}\\ \hat{M}_{21}^{(l)} & \mathbb{0}\end{bmatrix}d^{(l)} \tag{6.10}$$

with the largest absolute value, $\lambda_{\max}^{(l)}$. This can be done iteratively, for instance, via the Arnoldi iteration[87] using the same efficient multiplication techniques that are described later in this section. Subsequently, determine the order N_o up to which

$$\exp\left(\left|\lambda_{\max}^{(l)}\right|\right) = \sum_{n=0}^{N_\mathrm{o}} \frac{\left|\lambda_{\max}^{(l)}\right|^n}{n!} \tag{6.11}$$

needs to be expanded to reach the desired accuracy by estimating the size of the remainder[88]. For layer thicknesses in the order of the wavelength, one typically needs to expand up to the order $N_\mathrm{o} \approx 40$ to reach machine precision. In order to multiply the entire matrix shown in Eq. (6.9) to a known vector, it is sufficient to be able to multiply the sub-matrices[5, 14, 86]

$$\hat{M}_{12}^{(l)} = \begin{bmatrix}\hat{K}_x\left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1}\hat{K}_y & -\hat{K}_x\left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1}\hat{K}_x + \lceil\!\lfloor\mu^{(l)}\rfloor\!\rceil\\ \hat{K}_y\left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1}\hat{K}_y - \lfloor\!\lceil\mu^{(l)}\rceil\!\rfloor & -\hat{K}_y\left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1}\hat{K}_x\end{bmatrix} \tag{6.12a}$$

$$\hat{M}_{21}^{(l)} = \begin{bmatrix}-\hat{K}_x\left[\!\left[\mu^{(l)}\right]\!\right]^{-1}\hat{K}_y & \hat{K}_x\left[\!\left[\mu^{(l)}\right]\!\right]^{-1}\hat{K}_x - \lceil\!\lfloor\epsilon^{(l)}\rfloor\!\rceil\\ -\hat{K}_y\left[\!\left[\mu^{(l)}\right]\!\right]^{-1}\hat{K}_y + \lfloor\!\lceil\epsilon^{(l)}\rceil\!\rfloor & \hat{K}_y\left[\!\left[\mu^{(l)}\right]\!\right]^{-1}\hat{K}_x\end{bmatrix} \tag{6.12b}$$

efficiently. The matrices that appear in the blocks of Eq. (6.12) are, however, not entirely arbitrary, but have a specific structure. On the one hand, $\hat{K}_{x/y}$ are the diagonal matrices containing the normalized lateral wave vector components. On the other hand, $\left[\!\left[\epsilon^{(l)}\right]\!\right]$ and $\left[\!\left[\mu^{(l)}\right]\!\right]$ denote the Toeplitz matrices made up of the Fourier components of the electric permittivity $\epsilon^{(l)}(x,y)$ and magnetic permeability $\mu^{(l)}(x,y)$, respectively. The multiplication of the diagonal matrices $\hat{K}_{x/y}$ to a known vector is again straightforward. The multiplication of the Toeplitz matrices, for instance $\left[\!\left[\epsilon^{(l)}\right]\!\right]$, is equivalent to a convolution. This convolution can be carried out

as a point-wise multiplication in position space, i.e.,

$$\left[\!\left[\epsilon^{(l)}\right]\!\right] \widetilde{\mathbf{v}} = \mathscr{F}\left\{\epsilon^{(l)}\left(x,y\right)\mathscr{F}^{-1}\left[\widetilde{\mathbf{v}}\right]\left(x,y\right)\right\}\,, \tag{6.13}$$

where $\mathscr{F}$ is the two dimensional Fourier operator in the lateral coordinates x and y. For continuous functions and an infinite number of modes this relation is exact. However, for discrete systems and when using the FFT for Fourier transformation it only remains exact, if the matrix $\left[\!\left[\epsilon^{(l)}\right]\!\right]$ is defined in a circulant fashion[11,12]. The reason is that, independent of their specific entries, all circulant matrices share the same eigenvector matrix, which is known analytically and corresponds to the $M \times M$ Fourier matrix

$$F_{mm'} = \frac{1}{\sqrt{M}}\, e^{2\pi i \frac{mm'}{M}}\,, \tag{6.14}$$

familiar from the discrete Fourier transform (DFT)[89]. In the two-dimensional case, the Fourier transformation operator reads

$$F_{mn,m'n'} = \frac{1}{\sqrt{MN}}\, e^{2\pi i \frac{mm'}{M}}\, e^{2\pi i \frac{nn'}{N}}\,. \tag{6.15}$$

Therefore, discrete counterpart of Eq. (6.13) is

$$\begin{aligned}\sum_{m'n'} \left[\!\left[\epsilon^{(l)}\right]\!\right]_{mn,m'n'} \widetilde{v}_{m'n'} &= \sum_{m'n'} \left[\hat{F}\,\mathrm{diag}\left(\epsilon^{(l)}\right)\hat{F}^{-1}\right]_{mn,m'n'} \widetilde{v}_{m'n'} \\ &= \sum_{m'n'} \hat{F}_{mn,m'n'}\; \epsilon^{(l)}\left(x_{m'},y_{n'}\right)\left[\hat{F}^{-1}\widetilde{\mathbf{v}}\right]_{m'n'}\,. \end{aligned} \tag{6.16}$$

Identically, the multiplication of the inverse of the Toeplitz matrix, $\left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1}$, can be conducted via

$$\left[\!\left[\epsilon^{(l)}\right]\!\right]^{-1} \mathbf{v} = \mathscr{F}\left\{\epsilon^{(l)-1}\left(x,y\right)\mathscr{F}^{-1}\left[\mathbf{v}\right]\left(x,y\right)\right\}\,. \tag{6.17}$$

The multiplication of the matrices $\left[\!\left[\mu^{(l)}\right]\!\right]$ and $\left[\!\left[\mu^{(l)}\right]\!\right]^{-1}$ is done in the same way. Altogether, the numerical complexity of multiplying the entire second constituent matrix is dominated by the Fourier transformation, and is, therefore, of the order $\mathcal{O}\big(\bar{N}\log\bar{N}\big)$. The memory requirement is of the order $\mathcal{O}\big(\bar{N}\big)$.

[11] Additionally, an implicit band limitation of the permittivities and permeabilities is assumed and, at the same time, also a periodic continuation of the truncated mode coefficient vectors of the fields as explained in [17].

[12] In this case, Li's factorization rules are not necessary, since the direct rule and the inverse rule produce identical results[17]. Therefore, in Eq. (6.12), all Toeplitz matrices involving the inverse rule can be identically replaced and treated like the Toeplitz matrices involving the direct rule.

Runge-Kutta Method

Another possibility to compute the exponential matrix-vector product is via the RK method[90], which solves first order differential equations of the form of Eq. (6.1). Here, the 'fourth-order RK' method, commonly referred to as 'RK4', is applied. To this end, each layer is subdivided into $N^{(l)}$ small slices of thickness $h^{(l)} = d^{(l)}/N^{(l)}$. From one layer slice to the next, the field then changes according to

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp \end{bmatrix} \left(z + h^{(l)}\right) = \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp \end{bmatrix} (z) + \frac{1}{6} \left(\mathbf{K}_1 + 2\mathbf{K}_2 + 2\mathbf{K}_3 + \mathbf{K}_4\right) + \mathcal{O}\left(h^{(l)\,5}\right) , \tag{6.18a}$$

where

$$\mathbf{K}_1 = h^{(l)}\ ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp \end{bmatrix} (z) \tag{6.18b}$$

$$\mathbf{K}_2 = h^{(l)}\ ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} \left\{ \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp \end{bmatrix} (z) + \frac{\mathbf{K}_1}{2} \right\} \tag{6.18c}$$

$$\mathbf{K}_3 = h^{(l)}\ ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} \left\{ \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp \end{bmatrix} (z) + \frac{\mathbf{K}_2}{2} \right\} \tag{6.18d}$$

$$\mathbf{K}_4 = h^{(l)}\ ik_0 \begin{bmatrix} \mathbb{0} & \hat{M}_{12}^{(l)} \\ \hat{M}_{21}^{(l)} & \mathbb{0} \end{bmatrix} \left\{ \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{E}}}_\perp \\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp \end{bmatrix} (z) + \mathbf{K}_3 \right\} \tag{6.18e}$$

This way, the electric and magnetic fields can be traced through the grating step by step. This procedure can be applied for both the forward- as well as for the backward-propagating direction, the latter with negative step size. Furthermore, if this procedure is applied, it is not necessary to first approximate a continuous and z-dependent refractive index distribution by the staircase approximation as in the RCWA - the RK4 integrator can be directly applied to the z-dependent structure.

6.3.3 Third constituent matrix

The third constituent matrix (*c) is the matrix inverse of a 2×2 block matrix, where each block by itself is invertible. Therefore, one can rewrite

$$\begin{bmatrix} \mathbb{1} & \mathbb{1} \\ \hat{C}^{(0)} & -\hat{C}^{(0)} \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} \mathbb{1} & -\hat{C}^{(0)-1} \\ \mathbb{1} & +\hat{C}^{(0)-1} \end{bmatrix} . \tag{6.19}$$

Moreover, since the matrix $\hat{C}^{(0)}$ is a 2×2 block diagonal matrix, the inversion can be conducted analytically. Therefore, as for the first constituent matrix, both the memory requirement and the numerical complexity are of the order $\mathcal{O}\left(\bar{N}\right)$.

6.4 Convergence and limitations

In this section, the convergence conditions of the algorithm proposed in Sect. 6.2 is investigated. In the first part, a mathematical convergence criterion based on matrix eigenvalues is developed. In the second part, it is shown that the empiric convergence criteria for the iterative method are as follows.

1. The grating thickness is in the order of few wavelengths or lower.

2. The NA of the calculation is approximately equal or smaller than the lowest refractive index in the grating.

3. The grating structure is weakly or non-absorbing.

6.4.1 Mathematical convergence criterion

Consider the iteration scheme described in Sect. 6.2. In one full iteration step the following procedure is conducted,

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}+1} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}+1} \end{bmatrix} = \hat{\Gamma} \left\{ \begin{bmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & \frac{w}{w+1}\mathbb{1} \end{bmatrix} \hat{\Gamma}^{-1} \left(\begin{bmatrix} \frac{w}{w+1}\mathbb{1} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} + \begin{bmatrix} \frac{\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp}}{w+1} \\ \mathbf{0} \end{bmatrix} \right) + \begin{bmatrix} \mathbf{0} \\ \frac{\widetilde{\boldsymbol{\mathcal{L}}}^{(L+1)}_{\perp}}{w+1} \end{bmatrix} \right\} . \tag{6.20}$$

If one inserts

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} = \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix} + \begin{bmatrix} \Delta\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \Delta\widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} \tag{6.21}$$

into Eq. (6.20) and simplifies, one directly obtains the relations

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}+1} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}+1} \end{bmatrix} = \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp} \end{bmatrix} + \underbrace{\hat{\Gamma} \begin{bmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & \frac{w}{w+1}\mathbb{1} \end{bmatrix} \hat{\Gamma}^{-1} \begin{bmatrix} \frac{w}{w+1}\mathbb{1} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} \end{bmatrix}}_{\hat{\Xi}} \begin{bmatrix} \Delta\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \Delta\widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} \tag{6.22}$$

and

$$\begin{bmatrix} \Delta\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \Delta\widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} = \hat{\Xi}^{n_{\mathrm{it}}} \begin{bmatrix} \Delta\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;0} \\ \Delta\widetilde{\boldsymbol{\mathcal{R}}}_{\perp;0} \end{bmatrix} , \tag{6.23}$$

where $\hat{\Xi}$ is the contraction operator. Be $\lambda_{\mathrm{max};\hat{\Xi}}$ the eigenvalue of $\hat{\Xi}$ with the largest absolute value, which can be computed iteratively for instance via the Arnodi iteration[87].

From Eqs. (6.22, 6.23), a convergence behavior of the form

$$\left\| \begin{bmatrix} \Delta\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \Delta\widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} \right\|, \left\| \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}+1} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}+1} \end{bmatrix} - \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \\ \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}} \end{bmatrix} \right\| \propto \left| \lambda_{\max;\hat{\Xi}} \right|^{n_{\mathrm{it}}}, \tag{6.24}$$

is expected, i.e., the iteration series converges if and only if $\left|\lambda_{\max;\hat{\Xi}}\right| \leq 1$. From this it is also clear that the convergence rate increases for smaller $\left|\lambda_{\max;\hat{\Xi}}\right|$. In order to test the convergence behavior, a FRIM simulation of a small sized focusing DOE as shown in Fig. 6.2 is conducted, which is illuminated with the plane wave $\hat{\mathbf{e}}_x \exp(ik_0 z)$. The DOE is obtained by the scalar design method described in [91].

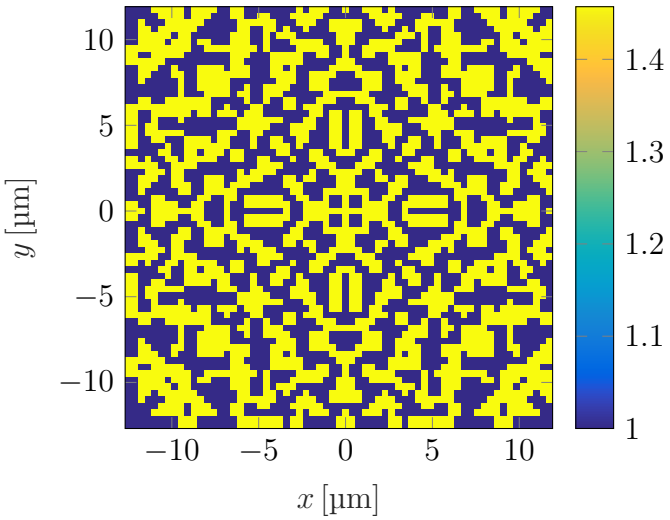


Figure 6.2: Refractive index distribution of a small sized DOE (scalar design[91]) made of fused silica $n_{\mathrm{FS}} = 1.4607$[92], wavelength $\lambda = 532\,\mathrm{nm}$, thickness $d = 577\,\mathrm{nm}$. The element is designed to produce a focus with NA $= 0.5$ at $z = f = 1.2\,\mathrm{mm}$. FRIM parameters: $P_x = P_y = 25\,\mu\mathrm{m}$, $M = N = 65$.

Fig. 6.3 shows the convergence curve of the FRIM for this particular simulation. It is in agreement with the analysis involving Eq. (6.24), i.e., the algorithm converges exponentially until reaching the numerical limit. This includes another major advantage of the FRIM. Unlike standard methods, the iterative method allows trading accuracy for speed at a given mode count by terminating the iteration at any desired accuracy before reaching full convergence. It is observed that the convergence curve flattens when the average error of a single mode coefficient is approaching machine precision. It is remarkable that the convergence behavior does not depend on the form of the incident light at all, but depends only on the wavelength, the grating properties, the number of sampling points, and the choice of the mixing parameter w. The exact influence of these quantities on $\lambda_{\max;\hat{\Xi}}$ is, however, unknown. For the mixing parameter, it was empirically found that a value of $w = 0.5\,..\,1$ works quite well in most cases.

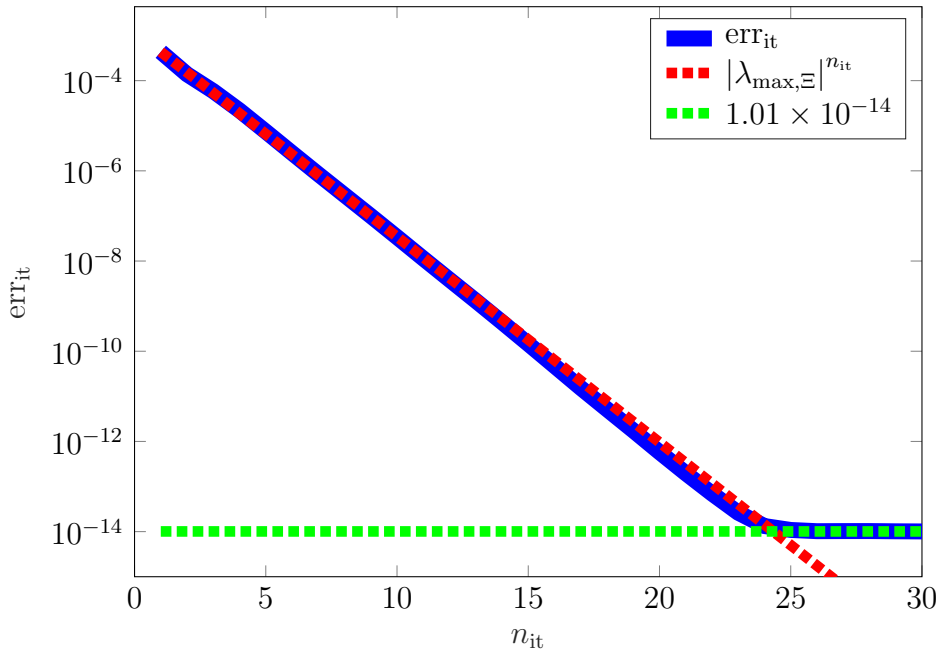


Figure 6.3: Convergence curve of the FRIM (solid blue line). The error estimate $\mathrm{err}_{\mathrm{it}}\left(n_{\mathrm{it}}\right) = \frac{1}{N}\left|\left[\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}+1} \quad \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}+1}\right]^T - \left[\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}} \quad \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}}\right]^T\right|$ is plotted versus the iteration number n_{it}. An exponential convergence curve according to $\left|\lambda_{\max;\hat{\Xi}}\right|^{n_{\mathrm{it}}}$ (dotted red line) is observed, as predicted. The quantity $\left|\lambda_{\max;\hat{\Xi}}\right| \approx 0.3501$ has been calculated via the Arnoldi iteration[87]. The convergence curve flattens upon reaching almost machine precision.

6.4.2 Empiric convergence criteria

In Sect. 6.4.1 a mathematical convergence criterion was developed for the iteration series presented in Sect. 6.2. However, no analytic relation between the physical setting, the simulation parameters and the eigenvalues of $\hat{\Gamma}$ and $\hat{\Xi}$ is known. Therefore, at this point, the empirically determined limits, within which the iterative algorithm typically converges, are presented and motivated. Note that the three conditions stated below are not strict limits. They should rather be seen as guidance values, i.e., it is possible to violate one to a certain degree at the expense of another.

Grating thickness $d \approx \lambda$

The grating thickness should be in the order of a few wavelengths or smaller. The reason is that for too thick layers the Taylor expansion of the matrix exponential in Eq. (6.9) has to be performed to very high orders N_{o}.

Moderate NA

The NA of the calculation, $\mathrm{NA}_{\mathrm{calc}} = \lambda\sqrt{\left(\frac{M_0}{P_x}\right)^2 + \left(\frac{N_0}{P_y}\right)^2}$, should be approximately $\mathrm{NA}_{\mathrm{calc}} < n_{\mathrm{min}}$, where n_{min} is the smallest refractive index in the grating, which is typically air. As before, $P_{x/y}$ are defined as the diameter of the computation field and M_0/N_0 as the one dimensional one-sided number of modes in the calculation. It is observed that upon passing the evanescent limit in the corners of the calculation area, i.e., $\mathrm{NA}_{\mathrm{calc}} > n_{\mathrm{min}}$, the quantity $\left|\lambda_{\mathrm{max};\hat{\Xi}}\right|$ quickly becomes larger than one, which leads to a diverging iteration. By going to even higher NAs, i.e., $\mathrm{NA}_{\mathrm{calc}} \gg n_{\mathrm{min}}$, the highest evanescent mode coefficients may become excessively dominant in the matrix exponential, Eq. (6.9). In consequence, the double precision arithmetic may at some point numerically cut off relevant contributions from the remaining modes.

Weakly absorbing structures

The exact relation between the absorption strength within the grating and the quantity $\left|\lambda_{\mathrm{max};\hat{\Xi}}\right|$, which determines the convergence behavior of the FRIM, is not straightforward. However, empirically one observes the following. Compared to the non-absorbing case, the quantity $\left|\lambda_{\mathrm{max};\hat{\Xi}}\right|$ is just slightly affected by weak absorption. When gradually increasing the absorption strength there is, in first approximation, also a gradual increase in $\left|\lambda_{\mathrm{max};\hat{\Xi}}\right|$. At some intermediate absorption level one approaches $\left|\lambda_{\mathrm{max};\hat{\Xi}}\right| = 1$, where the FRIM would start to diverge. This transition point varies, however, depending on the grating thickness and the particular structure.

As stated above, it is observed that a trade-off between the three mentioned limits is always possible up to a certain degree. One may violate one limit to a certain degree at the expense of another. Example: a grating of thickness $d = 4\lambda$ may be simulated with the FRIM if no absorption is present in the grating at all. At the other extreme, one could also simulate a grating of thickness $d = \lambda/20$ with a higher NA or a comparably high absorption.

6.5 Simulation results

In the first part of this section, the FRIM is validated by a direct comparison to the RCWA. To this end, a small sized focusing DOE is simulated with both methods and the results are compared. In the second part, the FRIM is used to rigorously simulate a large sized focusing FZP (phase height π) out of fused silica, which cannot be simulated with the RCWA due to time and memory constraints. The result is validated by testing the conservation of energy, and it is compared to a scalar simulation treating the grating in TEA and a subsequent AS propagation (TEA-AS). In the third part, the FRIM is compared to the WPM for a FZP with a phase

height of 2π. In the fourth part, the Zernike phase contrast method is rigorously simulated with the FRIM.

6.5.1 Comparison: FRIM vs. RCWA (small sized DOE)

As a first test, the FRIM is validated by comparing its results to an actual RCWA simulation. To this end, one considers again the small sized focusing DOE shown in Fig. 6.2, which can still be simulated with the RCWA. The DOE, which is illuminated with the plane wave $\hat{\mathbf{e}}_x \exp(ik_0 z)$ at $\lambda = 532\,\mathrm{nm}$, is simulated with both the RCWA and the FRIM. In the RCWA simulation, the circulant definition of the Toeplitz matrices as proposed in [17] is applied. The reason is that this circulant definition is inherently contained in the FRIM algorithm due to the representation of the Toeplitz matrices as a multiplication in position space. Fig. 6.4 shows an x-z slice through the simulated focus produced by the DOE for both the FRIM (top) and the RCWA (bottom). The differences between the RCWA and the FRIM simulations are in the order of machine precision.

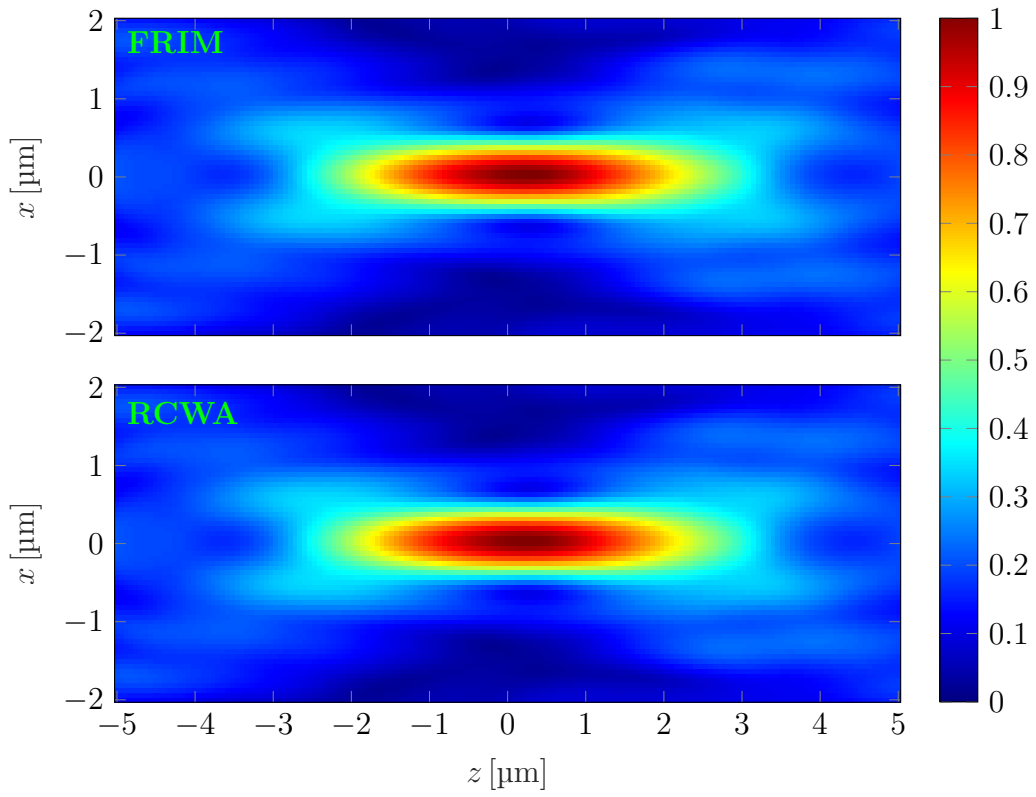


Figure 6.4: FRIM (top) and RCWA simulations (bottom). Electric field amplitude (x-z slice through the focus) produced by the DOE shown in Fig. 6.2 for straight plane wave incidence. The RCWA and FRIM results are identical up to machine precision. FRIM/RCWA parameters: $M_0 = N_0 = 32$, $N_\mathrm{o} = 46$, $N_\mathrm{it} = 30$, $\mathrm{NA}_\mathrm{calc} = 0.96$.

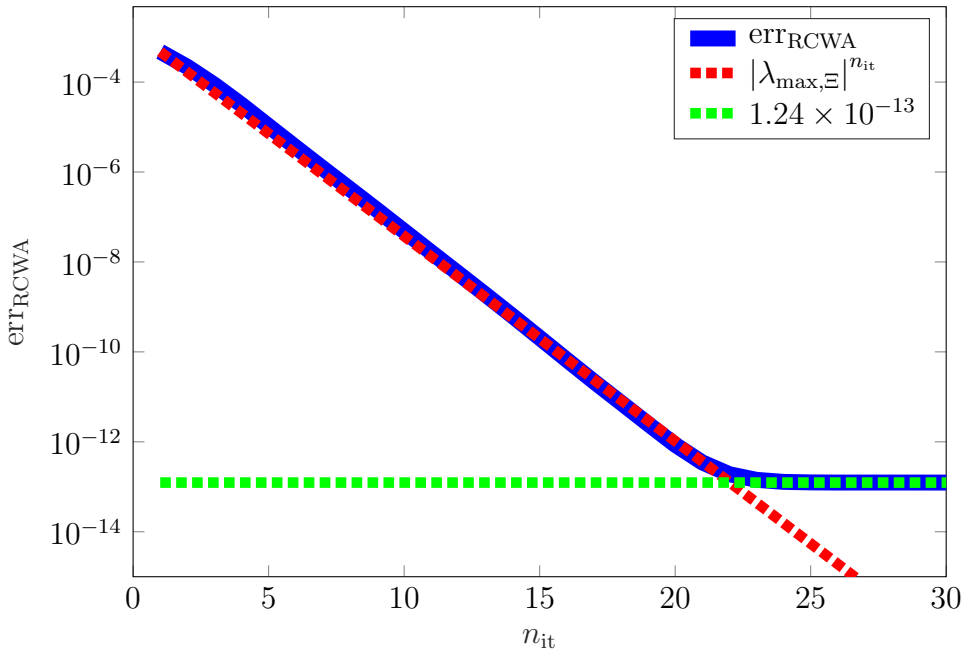


Figure 6.5: Convergence curve of the FRIM (solid blue line). The error estimate $\mathrm{err}_{\mathrm{RCWA}}(n_{\mathrm{it}}) = \frac{1}{N}\left|\left[\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;n_{\mathrm{it}}+1} \;\; \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;n_{\mathrm{it}}+1}\right]^T - \left[\widetilde{\boldsymbol{\mathcal{L}}}^{(0)}_{\perp;\mathrm{RCWA}} \;\; \widetilde{\boldsymbol{\mathcal{R}}}_{\perp;\mathrm{RCWA}}\right]^T\right|$, which denotes the error with respect to the RCWA result, is plotted versus the iteration number n_{it}. The same exponential convergence behavior as shown in Fig. 6.3 is observed.

In Fig. 6.3, convergence was defined as the difference between successive iterations. In Fig. 6.5, convergence is considered as the difference to the RCWA result. For each iteration step, n_{it}, one defines the deviation of the iterative result from the RCWA result via the error measure $\mathrm{err}_{\mathrm{RCWA}}(n_{\mathrm{it}})$. One clearly sees that the algorithm converges exponentially to the RCWA result until reaching a limit of approximately 10^{-13}. Also, here the $\left|\lambda_{\mathrm{max};\hat{\Xi}}\right|^{n_{\mathrm{it}}}$ convergence behavior predicted in Sect. 6.4.1 is confirmed.

6.5.2 Comparison: FRIM vs. TEA-AS (large sized FZP)

Now, a problem is regarded, which is not solvable with the RCWA due to time and memory constraints. Consider a large sized FZP with a phase height of π and the following parameters. Material fused silica, $n_{\mathrm{FS}} = 1.4607$[92], wavelength $\lambda = 532\,\mathrm{nm}$, thickness $d = 577\,\mathrm{nm}$, number of radii $n_{\mathrm{radii}} = 160$, focus at $z = f = 200\,\mu\mathrm{m}$, computational grid period $P_x = P_y = 300\,\mu\mathrm{m}$, number of pixels $M = N = 783$. This FZP is illuminated with the plane wave $\hat{\mathbf{e}}_x \exp(ik_0 z)$. Fig. 6.6 shows an x-y slice through the focus simulated with the FRIM (left) and the TEA-AS (right). One observes that the TEA-AS focus is perfectly rotationally symmetric. In contrast, this is not what one expects from a vectorial calculation. At this size scale, one anticipates an asymmetry in x-y-direction due to polarization effects, which the FRIM result

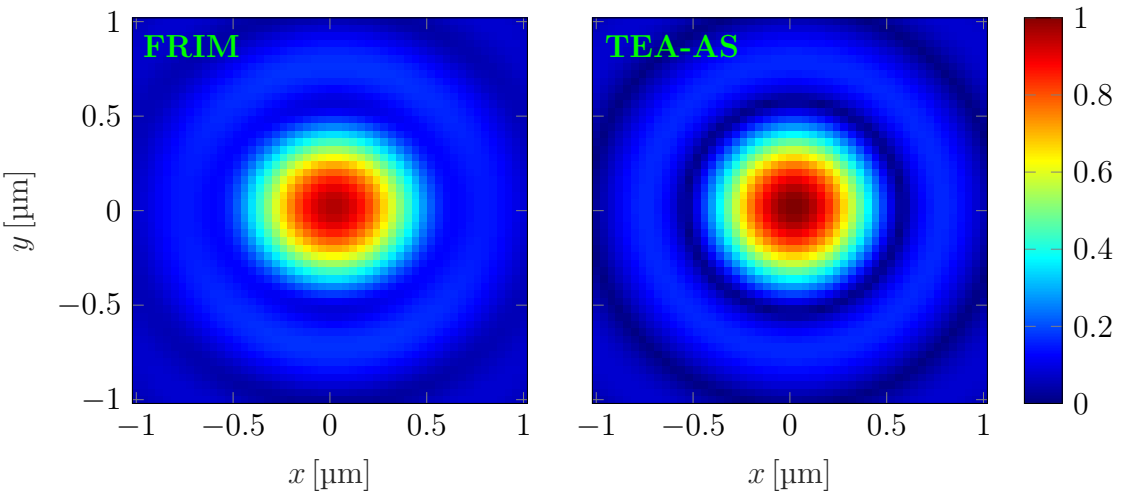


Figure 6.6: Exact FRIM simulation (left) and scalar TEA simulation (right) of the focus. Electric field amplitude (x-y slice) generated by the large size π-FZP for straight plane wave incidence. Due to polarization effects, the FRIM focus is clearly smeared out in x-direction, whereas the TEA focus is perfectly rotationally symmetric. FRIM parameters: $M_0 = N_0 = 391$, $N_\mathrm{o} = 30$, $N_\mathrm{it} = 20$, $\mathrm{NA}_\mathrm{calc} = 0.98$.

clearly shows. This characteristic asymmetric oval focus form is typically known from a lens focus with linear incident polarization. Unfortunately, it is not possible to compare the FRIM result in Fig. 6.6 to the RCWA. The reason is that the RCWA calculation would take approximately 216 years on the same computer, and it would require at least 90 TB of RAM (cf. Table 6.1). Therefore, one has to rely on consistency requirements to assess whether the FRIM result is correct. In this case, the conservation of energy is considered. Since the DOE is non-absorbing, the amount of energy contained in the incident wave must equal the amount of energy contained in the reflected and transmitted waves combined. In this particular example, the power dissipation is $P_\mathrm{diss} < 10^{-13}$. Hence, the energy is conserved almost up to machine precision.

6.5.3 Comparison: FRIM vs. WPM (large sized FZP)

In the following, consider the same illumination and the same FZP structure as in Sect. 6.5.2, but now with a phase height of 2π, i.e., a thickness of $d = 1155\,\mathrm{nm}$. In this case, the TEA treats the structure as a constant transmission object and one can, therefore, expect that the incident plane wave is not altered by the TEA at all. One of the few methods for treating inhomogeneous media at high NA is the WPM[52], which is scalar and unidirectional. It was recently demonstrated[93] as very comparable to rigorous COMSOL simulations[94]. Therefore, the FRIM is compared to the result of the WPM. Fig. 6.7 shows a central x-z slice (top) and y-z slice (center) of the electric field amplitude behind the 2π-FZP for the FRIM, and a central x-z slice for the WPM (bottom). First, one observes that, for both simulation

methods, the wave form differs significantly from that of a plane wave, which is what the TEA would predict in the transmitted region. Furthermore, in all three images two foci are seen at $z \approx 95\,\mu\text{m}$ and $z \approx 200\,\mu\text{m}$, respectively. In principle, also other features, such as the tails of the foci, look quite similar. Note, however, that, in the scalar WPM simulation, the peak amplitude of the second focus is predicted higher than in the exact simulation. Moreover, in the scalar WPM simulation the second focus is predicted to be nearly symmetric in axial direction, whereas this is clearly not the case in the exact FRIM simulation. The asymmetry is a polarization artifact, since one observes significant differences in the x- and y-polarization slices for the FRIM.

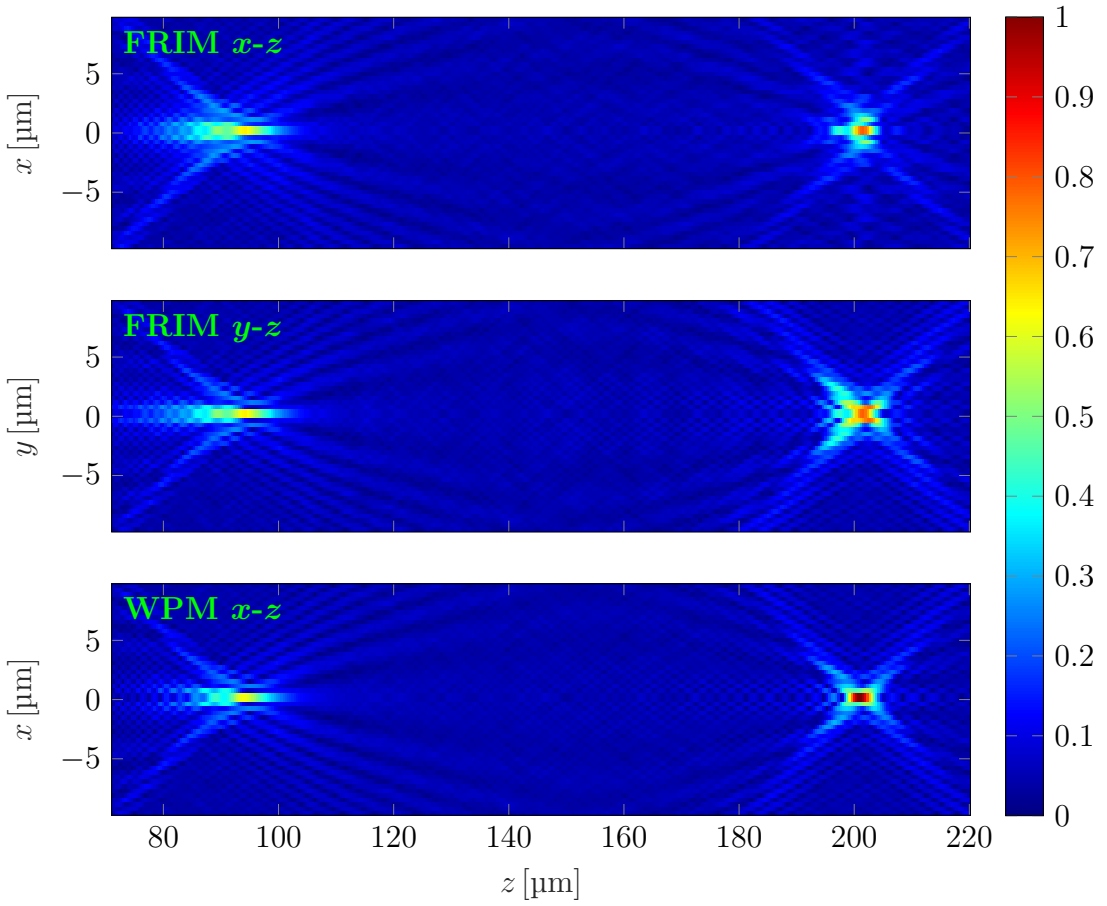


Figure 6.7: Electric field amplitude behind a 2π-FZP for plane wave light incidence. Rigorous FRIM simulation (top: x-z slice; center: y-z slice) and WPM simulation (bottom: x-z slice). Both simulations exhibit similar features and differ significantly from the form of a plane wave. Differences in the focus form and amplitude are observed. FRIM parameters: $M_0 = N_0 = 391$, $N_\text{o} = 60$, $N_\text{it} = 18$, $\text{NA}_\text{calc} = 0.98$

6.5.4 FRIM simulation of the Zernike phase contrast method

To give an example for a large scale problem, the FRIM is applied to simulate the Zernike phase contrast method[95]. With this method, pure phase objects can be made visible on a detector by filtering with a phase plate in the Fourier domain. The corresponding optical setup is illustrated in Fig. 6.8. Here, the electric field behind the phase object, which is shown in Fig. 6.9, is considered as given. It is assumed that the phase object is very thin, i.e., the TEA can be applied here and the polarization of the initial wave is not affected upon the transition. The combination of the lens and the propagation to the phase plate is treated by a Fourier transformation, here realized by a FFT. The result is used as the input for the rigorous FRIM calculation, where the polarization direction of the electric field modes in front of the phase plate, $\widetilde{\boldsymbol{\mathcal{L}}}(\mathbf{k}) \,/\, \left|\widetilde{\boldsymbol{\mathcal{L}}}(\mathbf{k})\right|$, is obtained as described for linear polarization in Sect. 5.2. The propagation from the back side of the phase plate to the detector is again treated by a FFT. The FRIM is applied only for the transmission through the phase plate, because there the illumination is a strongly converging wave with high NA, whereas the thin phase object is only illuminated by a plane wave.

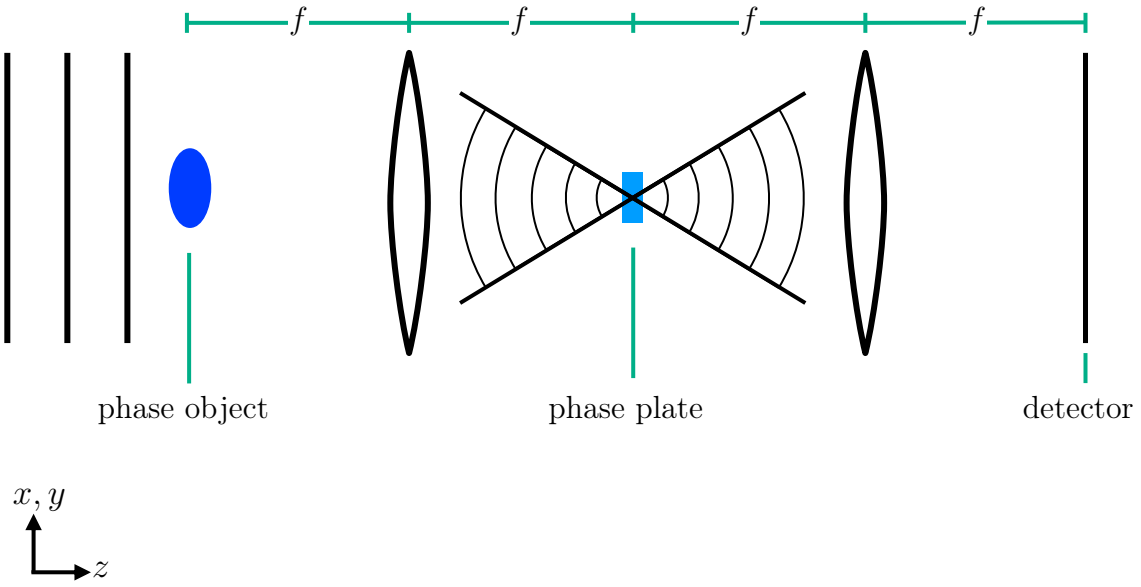


Figure 6.8: Optical setup of the Zernike phase contrast method. A phase object is illuminated by a plane wave, followed by an optical Fourier transformation. In the Fourier plane, the light passes through a circular $\lambda/4$ phase plate with a diameter of 2.5 µm. Another optical Fourier transformation represents the propagation to the detector.

In the following, it is assumed that $\lambda = 633\,\mathrm{nm}$ and that the initial incident plane wave is y-polarized, i.e., $\hat{\mathbf{e}}_{\mathbf{y}} \exp(ik_0 z)$. Fig. 6.10 shows the simulated detector images for the case that the transition through the phase plate is, on the one hand, calculated with the FRIM (left) and, on the other hand, with the TEA (right). Generally, both

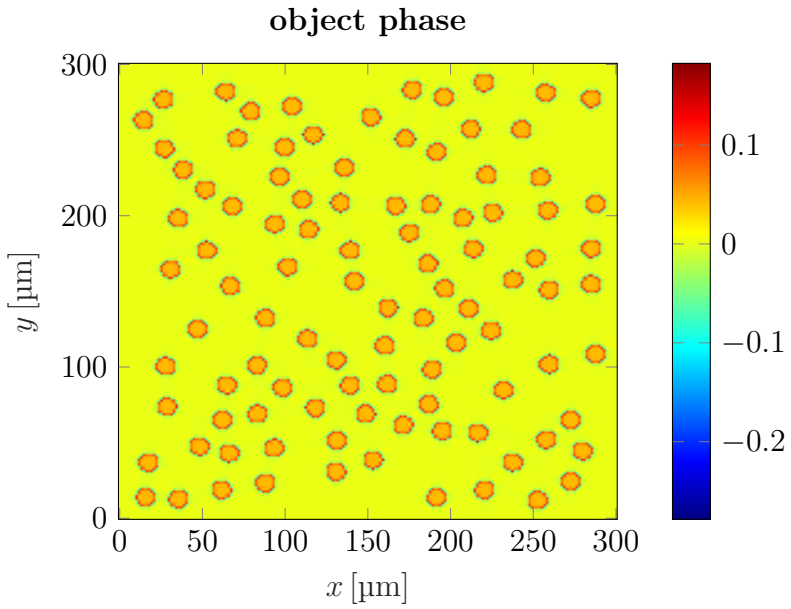


Figure 6.9: Sample phase object consisting of 100 randomly distributed circles with radius $r = 2\,\mu\text{m}$ incident at the optical setup shown in Fig. 6.8. The phase jump within a circle is $\Delta\phi = 0.1\,\pi$. Grid period $P_x = P_y = 100\,\mu\text{m}$, number of pixels $M = N = 201$.

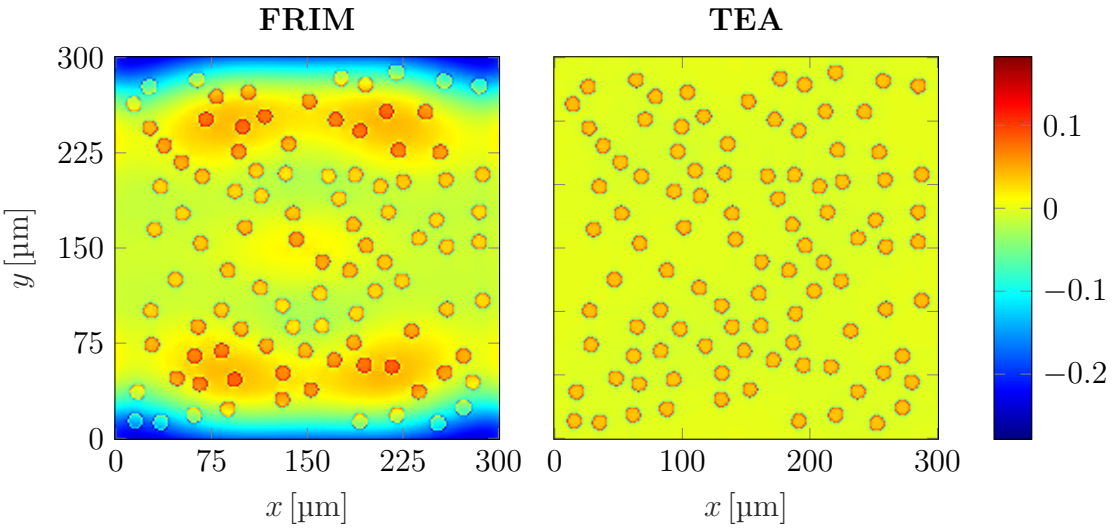


Figure 6.10: Rigorous FRIM simulation (left) and scalar TEA simulation (right) of the Zernike phase contrast method (electric field amplitude at the detector) for the incident phase object shown in Fig. 6.9. FRIM parameters: $M_0 = N_0 = 100$, $N_\text{o} = 30$, $N_\text{it} = 30$, $\text{NA}_\text{calc} = 0.633$

detector images reproduce the initial phase object. However, the background in the FRIM image is much more strongly modulated than in the TEA image. The reason is that, in contrast to the TEA, in the exact FRIM simulation, reflections and mode coupling across the lateral phase plate interfaces are taken into account. Therefore, a different fraction of the actual object information crosses the phase plate.

6.6 Extension to anisotropic media

The FRIM can be directly extended to anisotropic media by replacing the isotropic coupling matrix in Eq. (6.1) by its anisotropic counterpart derived in Sect. 3.3.2,

$$\frac{\partial}{\partial z}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp(z)\end{bmatrix} = ik_0\begin{bmatrix}\hat{M}_{11} & \hat{M}_{12}\\ \hat{M}_{21} & \hat{M}_{22}\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{\mathcal{E}}}_\perp(z)\\ \widetilde{\boldsymbol{\mathcal{H}}}_\perp(z)\end{bmatrix}, \tag{6.25a}$$

with

$$\hat{M}_{11} = \begin{bmatrix}+\hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{K}_y-\hat{K}_x\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & -\hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{K}_x-\hat{K}_x\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\\ +\hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{K}_y-\hat{K}_y\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & -\hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{K}_x-\hat{K}_y\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\end{bmatrix} \tag{6.25b}$$

$$\hat{M}_{12} = \begin{bmatrix}+\hat{M}_{yx}+\hat{K}_x\hat{Q}_{zz}^{-1}\hat{K}_y-\hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{M}_{zx} & +\hat{M}_{yy}-\hat{K}_x\hat{Q}_{zz}^{-1}\hat{K}_x-\hat{M}_{yz}\hat{M}_{zz}^{-1}\hat{M}_{zy}\\ -\hat{M}_{xx}+\hat{K}_y\hat{Q}_{zz}^{-1}\hat{K}_y+\hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{M}_{zx} & -\hat{M}_{xy}-\hat{K}_y\hat{Q}_{zz}^{-1}\hat{K}_x+\hat{M}_{xz}\hat{M}_{zz}^{-1}\hat{M}_{zy}\end{bmatrix} \tag{6.25c}$$

$$\hat{M}_{21} = \begin{bmatrix}-\hat{Q}_{yx}-\hat{K}_x\hat{M}_{zz}^{-1}\hat{K}_y+\hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & -\hat{Q}_{yy}+\hat{K}_x\hat{M}_{zz}^{-1}\hat{K}_x+\hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\\ +\hat{Q}_{xx}-\hat{K}_y\hat{M}_{zz}^{-1}\hat{K}_y-\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx} & +\hat{Q}_{xy}+\hat{K}_y\hat{M}_{zz}^{-1}\hat{K}_x-\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\end{bmatrix} \tag{6.25d}$$

$$\hat{M}_{22} = \begin{bmatrix}-\hat{K}_x\hat{M}_{zz}^{-1}\hat{M}_{zx}-\hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{K}_y & -\hat{K}_x\hat{M}_{zz}^{-1}\hat{M}_{zy}+\hat{Q}_{yz}\hat{Q}_{zz}^{-1}\hat{K}_x\\ -\hat{K}_y\hat{M}_{zz}^{-1}\hat{M}_{zx}+\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{K}_y & -\hat{K}_y\hat{M}_{zz}^{-1}\hat{M}_{zy}-\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{K}_x\end{bmatrix}. \tag{6.25e}$$

As shown in the derivation of the fast and efficient matrix multiplication scheme in Sect. 6.3, all the constituent matrices are required to be either diagonal or circulant matrices for the algorithm to work. In fact, also in the anisotropic case, the matrices $\hat{K}_{x/y}$ are diagonal and the matrices $\hat{M}_{\alpha\beta}$ and $\hat{Q}_{\alpha\beta}$ retain a circulant form. The latter may be not directly obvious, but can be derived from some basic properties of circulant matrices.

6.6.1 Properties of circulant matrices

As mentioned before, all circulant matrices have the same eigenvector matrix (the Fourier matrix)[89]. From this property, the following can be concluded.

1. All circulant matrices commute,

$$\hat{C}_1\hat{C}_2 = \hat{C}_2\hat{C}_1\,. \tag{6.26}$$

2. The sum of two circulant matrices is a circulant matrix. The eigenvalues of

the sum equals the sum of the eigenvalues,

$$\hat{C}_1 + \hat{C}_2 = \hat{F}\hat{D}_1\hat{F}^{-1} + \hat{F}\hat{D}_2\hat{F}^{-1} = \hat{F}\left(\hat{D}_1 + \hat{D}_2\right)\hat{F}^{-1}\,, \tag{6.27}$$

where $\hat{D}_{1/2}$ are the diagonal eigenvalue matrices.

3. The product of two circulant matrices is a circulant matrix. The eigenvalues of the product equals the product of the eigenvalues,

$$\hat{C}_1\hat{C}_2 = \hat{F}\hat{D}_1\hat{F}^{-1}\hat{F}\hat{D}_2\hat{F}^{-1} = \hat{F}\hat{D}_1\hat{D}_2\hat{F}^{-1}\,. \tag{6.28}$$

Obviously, it is possible to operate with circulant matrices just as one would with complex numbers. Since the matrices $\hat{M}_{\alpha\beta}$ and $\hat{Q}_{\alpha\beta}$ only consist of sums and products of circulant matrices, they themselves must also have this form. Consequently, the fast multiplication scheme shown in Sect. 6.3 can be applied to the anisotropic coupling matrix as is.

6.6.2 Simulation of a half-wave plate

As an example, consider a circular-shaped half-wave plate (HWP) with radius $r_{\text{hwp}} = 30\,\mu\text{m}$ as shown in Fig. 6.11, which consists out of the material calcite - a material that exhibits birefringence. At the wavelength $\lambda = 590\,\text{nm}$, the refractive indices of the ordinary and extraordinary axes are $n_{\text{slow}} = 1.658$ and $n_{\text{fast}} = 1.486$, respectively[96]. Assume that this HWP is illuminated centrally and head-on by a

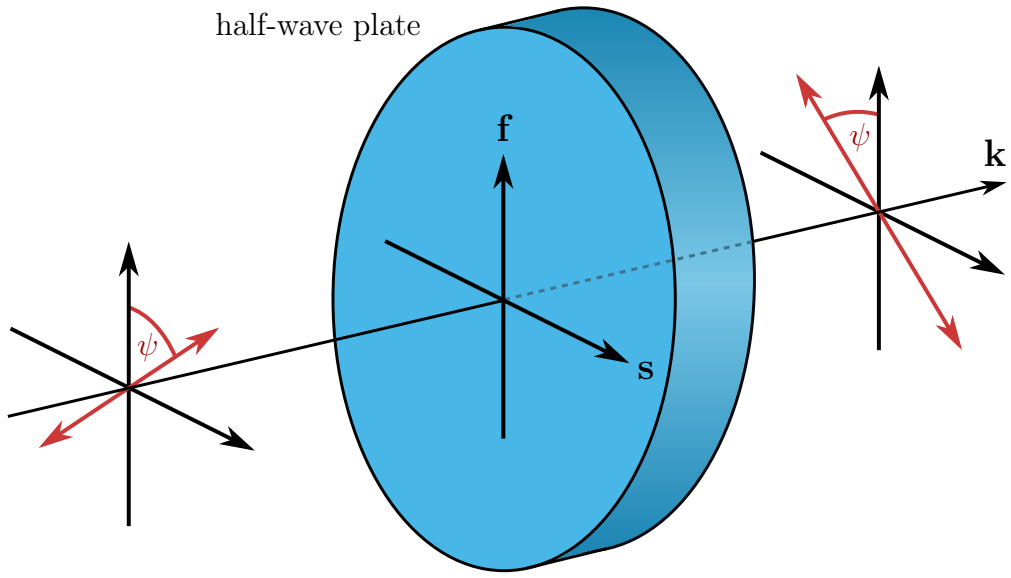


Figure 6.11: A HWP is illuminated with a linearly polarized Gaussian beam at the angle $\psi = 45°$ to the fast axis. The polarization angle of the incident light is turned by $2\psi = 90°$.

linearly polarized Gaussian beam with $\sigma = 40\,\mu\text{m}$, where the electric field polariza-

tion vector and the fast axis form the angle $\psi = 45°$. Due to the refractive index differences along the fast and slow axis, a relative phase difference $\delta\phi = k_0\,\delta n\,d$ is created between the polarization components upon the propagation through the waveplate, where $\delta n = n_{\text{slow}} - n_{\text{fast}}$ and d is the thickness of the waveplate. The thickness of a HWP is chosen such that $\delta\phi = \pi$, i.e., the component along the slow axis is retarded by half an oscillation cycle, which corresponds to a sign change. In consequence, the HWP is expected to turn the plane of polarization by $2\psi = 90°$ relative to the incident polarization. Fig. 6.12 shows the orientation of the polarization vectors before (left) and behind (right) the waveplate calculated with a FRIM simulation. It is clearly seen that, as expected, the HWP rotates the polarization

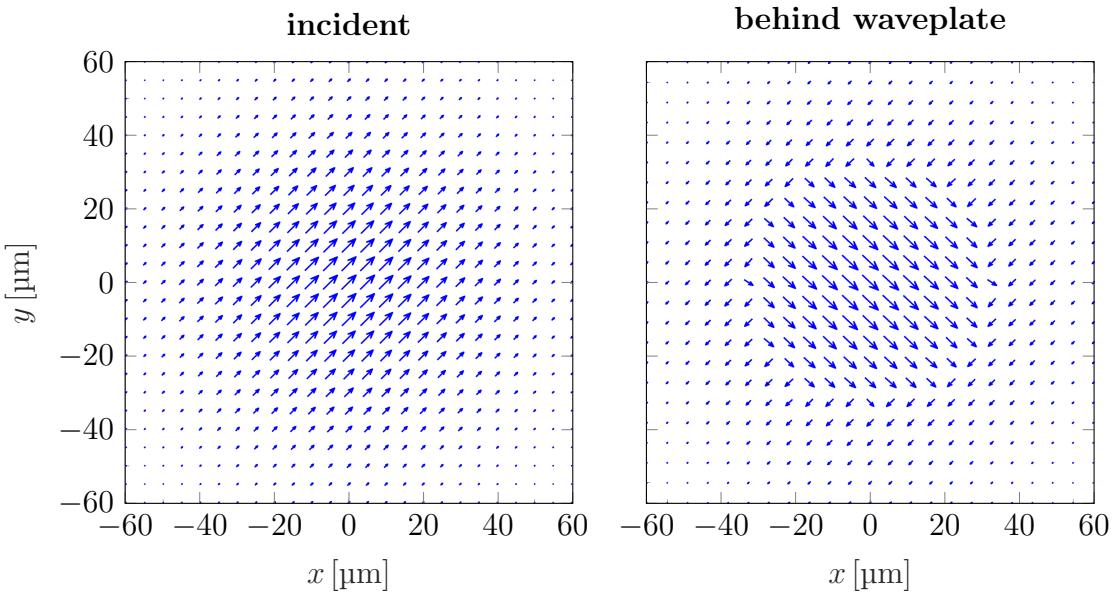


Figure 6.12: Anisotropic FRIM simulation: fraction of linearly polarized light at the front (left) and back (right) of the HWP. Every 10^{th} vector is plotted in each dimension. FRIM parameters: $P_x = P_y = 200\,\mu\text{m}$, $M = N = 401$.

vector of the incident light by $2\psi = 90°$, whereas the original polarization state remains unchanged outside the waveplate area.

6.7 Conclusions

In this chapter, the FRIM is introduced, a new algorithm for the exact simulation of large scale and high resolution optical problems. It is shown that, under certain conditions, both an eigenvalue decomposition and large numeric matrix inversions as inherent to modal methods can be circumvented by replacing these operations by a sequence of efficient matrix multiplications. This approach reduces the numerical complexity from $\mathcal{O}(\bar{N}^3)$ in modal methods to $\mathcal{O}(\bar{N}\log\bar{N})$. At the same time, the memory requirement decreases from $\mathcal{O}(\bar{N}^2)$ to $\mathcal{O}(\bar{N})$. This makes the complexity of

the FRIM comparable to standard scalar propagation methods like, for instance, the AS propagation. A major advantage of the iterative nature of the FRIM is the possibility to trade accuracy versus calculation time at any desired mode count, which is not possible in standard modal methods, for which the only possible trade-off is a reduction of the number of modes considered in the calculation. It is demonstrated that, with the FRIM, large scale structures, such as certain DOEs, can be simulated rigorously at very high mode counts, with 2000×2000 sampling points and more being possible. Due to this dense sampling of the angular spectrum, the FRIM enables to raise the lateral dimensions of rigorously analyzable structures to a much higher level. Furthermore, it is proven that the results of the FRIM coincide with the RCWA, provided that an RCWA simulation can still be conducted. When no such comparison is possible, it is shown that, on the one hand, the FRIM result is indeed fully energy conserving, and, on the other hand, gives similar results as established scalar propagation methods like the TEA and the WPM. Furthermore, it is demonstrated that the FRIM is also especially useful for the simulation of structures that do *not* exhibit a small scale periodicity as in gratings, i.e., in cases where large calculation domains are inevitable. As an example, a numerical simulation of the Zernike phase contrast method is performed. Eventually, it is shown how the FRIM can be extended to the treatment of anisotropic media as well. To this end, the transition of a linearly polarized Gaussian beam through a HWP is simulated.

Furthermore, also the convergence properties of the FRIM are investigated. First, a strictly mathematical convergence criterion is derived, which permits estimating the number of necessary iterations as well as the calculation time before actually starting with the iteration. Second, the empiric convergence behavior of the FRIM is motivated. It is found that the physical limitations concern the layer thickness (up to a few wavelengths), the NA of the calculation ($\mathrm{NA}_{\mathrm{calc}} \leq n_{\mathrm{min}}$), and the amount of absorption in the grating (weak). Nevertheless, it is also shown that one may violate one limit to a certain degree at the expense of another, i.e., the mentioned limits are not strict and a trade-off is always possible.

In the section about the convergence properties it is concluded that, when any of the iterations diverges, this does not necessarily mean that the limited accuracy of the double precision arithmetic is the cause. Therefore, it is expected that convergence improvements of the FRIM are still possible and that, potentially, also other iteration schemes with a better convergence behavior may still be found.

7 | Conclusions and outlook

In this dissertation, the application range of the RCWA is extended. The modifications of the framework are presented in three parts.

In Chapter 4, the RCWA is extended to allow for the simulation of structured incident and transmitted regions. This way, axially extended objects, such as long waveguides, fibers or integrated optical components can be simulated in an exact and fully vectorial manner, while reflections from unwanted distant interfaces with homogeneous regions are completely avoided. Furthermore, it enables treating structured layers of large axial extent with only one or even no interface to a homogeneous region at all. Possible examples include the mode coupling into or out of long waveguides, perturbations and small defects in long waveguides and their effect on eigenmode propagation, or the evanescent coupling between waveguides. In order to extend the framework, the boundary value problem of the RCWA is reformulated and established solution algorithms like the S-matrix approach or the ETMA are adapted. Furthermore, the energy flow properties and the propagation direction of eigenmodes in structured media are derived, so that sensible boundary conditions can be set at the outer layer interfaces.

In order to verify the approach, first, it is demonstrated that the numerically calculated eigenmodes of a planar dielectric waveguide are in good agreement with the analytic eigenmodes. Second, it is shown that the evanescent coupling between two adjacent waveguides can be reproduced. A comparison to CMT is done for both a weakly and a strongly coupled system. A good agreement is, indeed, observed in the weakly coupled regime. Third, the mode coupling of a focused Gaussian beam from vacuum into a long GRIN fiber without a rear interface is simulated. On the one hand, the analytically known imaging characteristic of sech-profile waveguides is demonstrated. On the other hand, it is observed that, for medium to large propagation distances, the BPM develops strong artifacts, which are not present in the RCWA simulation. The reason is that the BPM accumulates errors upon every propagation step $\mathrm{d}z$. This is clearly not the case for the RCWA, which allows a valid and fast field calculation even for large propagation distances. The latter is, in principle, only limited by the numerical error from the eigenvalue decomposition.

In Chapter 5, the framework of the RCWA is extended to the treatment of coherent bidirectional structured light incidence. To this end, the following concepts are com-

bined. First, the localized field approach by Auer and Brenner[17] is adopted for a fast and consistent modeling of arbitrary structured illumination. Second, a framework is introduced, which allows modeling an arbitrary given polarization state on structured wave fronts. As an example, expressions for different polarization types such as linear, radial, azimuthal or elliptic polarization are stated. Third, coherent bidirectional structured light incidence is integrated into the existing RCWA framework. The additional introduction of a backside illumination symmetrizes and, therefore, completes the description of the RCWA. Finally, the mentioned concepts are combined to investigate how the light focus in 4π-microscopy is aberrated due to the presence of clustered protein structures. One observes that, if not corrected, the aberrations may result in a fluorescence peak position shift in the range of several hundred nanometers in the vicinity of a specimen. In summary, it is shown that, in principle, sample-induced wave front aberrations can be simulated with an exact method for realistic 4πillumination scenarios. Therefore, it is expected that this enables a better understanding and correction of the aberrations occurring in these types of systems.

In Chapter 6, the FRIM is developed, which represents a new algorithm for the exact simulation of large scale and high resolution optical problems based on an iterative approach. It is shown that, under certain conditions, the FRIM can reduce the algorithmic complexity from $\mathcal{O}\left(\bar{N}^3\right)$ of modal methods to $\mathcal{O}\left(\bar{N}\log\bar{N}\right)$, where at the same time the memory requirement decreases from $\mathcal{O}\left(\bar{N}^2\right)$ to $\mathcal{O}\left(\bar{N}\right)$. The latter is achieved by circumventing the computationally complex eigenmode decomposition inherent to standard modal methods, which is responsible for the algorithmic complexity of $\mathcal{O}\left(\bar{N}^3\right)$, by replacing this operation by a sequence of efficient matrix multiplications. Indeed, this makes the numerical complexity of the FRIM comparable to standard scalar propagation methods like, for instance, the AS propagation. This way, one of the main drawbacks of rigorous Maxwell solvers - the limitation to very small calculation domains - may be eased. It is demonstrated that the FRIM is applicable to certain large scale structures, such as certain DOEs, which can be simulated rigorously at very high mode counts, with 2000×2000 sampling points and more being possible. Due to this dense sampling of the angular spectrum, the lateral dimensions of rigorously analyzable structures is raised to a much higher level. Furthermore, a major advantage of the FRIM is the possibility to terminate the iteration process at any desired accuracy, i.e., it is possible to trade accuracy versus calculation time at any desired fix mode count. The latter is not possible in standard modal methods, for which the only possible trade-off is a reduction of the number of modes considered in the calculation.

To verify the validity of the FRIM, several comparisons to established simulation methods are conducted in different application scenarios. First, it is shown that the FRIM and the RCWA coincide for small problem sizes - here a focusing DOE with small-scale periodicity - provided that an RCWA simulation can still be conducted. Second, in the case of a large sized FZP with a phase height of π, it is demonstrated that the FRIM result is comparable to the TEA, whereas the typical asymmetric

focus shape is, indeed, reproduced by the rigorous FRIM. Since in this case no comparison with the RCWA is possible, it is pointed out that the FRIM result indeed conserves the energy in the order of 10^{-13}. Third, for a large sized FZP with a phase height of 2π, the FRIM is compared to the high-NA WPM, since the intensity pattern behind the phase object originates only from diffraction at the lateral interfaces and the TEA acts as if no phase object were present at all. It is shown that both the FRIM and the WPM results have similar characteristics. However, especially in focused regions the WPM deviates significantly from the FRIM. It is shown that the latter likely is a polarization effect, since, in these regions, one observes differences in the FRIM results for TE- and TM-polarization. Fourth, it is shown that the FRIM becomes especially useful in cases, where a small scale periodicity, such as in gratings, is not present and, therefore, large calculation domains are inevitable. As an example, a numerical simulation of the Zernike phase contrast method is performed. Finally, it is shown that the FRIM can be directly extended to the treatment of anisotropic media. To this end, the transition of a linearly polarized Gaussian beam through a HWP is simulated. As expected, a rotation of the polarization direction is observed behind the waveplate.

In addition to the mere results, the mathematical and the empiric convergence properties of the FRIM are analyzed in detail. On the one hand, a strictly mathematical convergence criterion is derived on the basis of a contraction operator. Since the mathematical convergence is nearly exponential, the number of necessary iterations as well as the calculation time may be estimated before actually starting with the calculation. On the other hand, the empiric convergence behavior of the FRIM is motivated. It is found that the physical limitations concern the layer thickness (up to a few wavelengths), the NA of the calculation ($\mathrm{NA}_{\mathrm{calc}} \leq n_{\mathrm{min}}$) and the amount of absorption in the grating (weak). Furthermore, it is shown that it is possible to violate one limit to a certain degree at the expense of another, i.e., the mentioned limits are not strict and a trade-off is always possible.

Outlook

With the extensions presented in this dissertation, the RCWA becomes applicable to a much wider set of realistic problems. First, the extension to structured incident and transmitted regions is expected to facilitate the exact simulation of waveguides, optical fibers and integrated optical components, since eigenmode illumination is inherently treated and no artificial interfaces to homogeneous regions must be enforced at the boundaries. The latter should be especially useful in the modeling of small-diameter optical fibers, where polarization effects play a critical role. Furthermore, in the context of waveguides and fibers, for the first time the error-free numerical calculation of fiber coupling efficiencies is possible within the RCWA framework, since unwanted reflections from unphysical interfaces can be entirely eliminated. Second, the RCWA extension to coherent bidirectional light incidence allows an exact simulation of multi-directional illumination scenarios, which may be applied, for instance,

to the simulation of 4π-microscopy experiments. Another advantage of this extension is the possibility to invert the light path in a simulation without having to rotate or mirror the electric permittivity and magnetic permeability distribution. This way, implementation errors can easily be detected and avoided. Third, the FRIM allows simulating certain problem geometries with an unprecedented resolution. It is expected that the rigorous design of, for instance, DOEs is greatly facilitated or even made possible at all, since the inverse grating problem is known to require a larger number of iteration steps. Furthermore, in the section about the FRIM convergence properties, it is motivated that a diverging iteration must not necessarily be caused by running into the limits of the double-precision floating point arithmetic. The decision about convergence or divergence should rather be regarded as an intrinsic property of the iteration scheme with the grating and the mode count as parameters. Therefore, convergence improvements may still be found, for instance by applying alternative iteration schemes. This way, the present limitations on the FRIM may be eased and more general problem geometries may be simulated with a high resolution in the near future.

List of figures

List of tables

A | Bibliography

[1] Moharam, M. G. and Gaylord, T. K. "Rigorous coupled-wave analysis of planar-grating diffraction". In: *J. Opt. Soc. Am.* 71.7 (July 1981), pp. 811–818 (cit. on pp. 1, 18, 20, 30, 32).

[2] Thompson, S. E. and Parthasarathy, S. "Moore's law: the future of Si microelectronics". In: *Mater. Today* 9.6 (2006), pp. 20–25 (cit. on p. 1).

[3] Li, L. "Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity". In: *J. Opt. Soc. Am. A* 10.12 (Dec. 1993), pp. 2581–2591 (cit. on pp. 1, 59).

[4] Bräuer, R. and Bryngdahl, O. "Electromagnetic diffraction analysis of two-dimensional gratings". In: *Opt. Commun.* 100 (July 1993), pp. 1–5 (cit. on p. 1).

[5] Moharam, M. G., Gaylord, T. K., Grann, E. B., and Pommet, D. A. "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings". In: *J. Opt. Soc. Am. A* 12.5 (May 1995), pp. 1068–1076 (cit. on pp. 1, 27, 30, 32, 73, 111).

[6] Lalanne, P. and Morris, G. M. "Highly improved convergence of the coupled-wave method for TM polarization". In: *J. Opt. Soc. Am. A* 13.4 (Apr. 1996), pp. 779–784 (cit. on p. 1).

[7] Granet, G. and Guizal, B. "Efficient implementation of the coupled-wave method for metallic lamellar gratings in TM polarization". In: *J. Opt. Soc. Am. A* 13.5 (May 1996), pp. 1019–1023 (cit. on p. 1).

[8] Li, L. "Use of Fourier series in the analysis of discontinuous periodic structures". In: *J. Opt. Soc. Am. A* 13.9 (Sept. 1996), pp. 1870–1876 (cit. on pp. 1, 35, 37).

[9] Li, L. "New formulation of the Fourier modal method for crossed surface-relief gratings". In: *J. Opt. Soc. Am. A* 14.10 (Oct. 1997), pp. 2758–2767 (cit. on pp. 1, 35, 38, 42).

[10] Li, L. "Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings". In: *J. Opt. Soc. Am. A* 13.5 (May 1996), pp. 1024–1035 (cit. on pp. 1, 56, 57, 59, 97).

[11] Moharam, M. G., Pommet, D. A., and Grann, E. B. "Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach". In: *J. Opt. Soc. Am. A* 12.5 (1995), pp. 1077–1086 (cit. on pp. 1, 59, 97, 98, 100).

[12] Tan, E. L. "Note on formulation of the enhanced scattering- (transmittance-) matrix approach". In: *J. Opt. Soc. Am. A* 19.6 (June 2002), pp. 1157–1161 (cit. on p. 1).

[13] Silberstein, E., Lalanne, P., Hugonin, J.-P., and Cao, Q. "Use of grating theories in integrated optics". In: *J. Opt. Soc. Am. A* 18.11 (Nov. 2001), pp. 2865–2875 (cit. on p. 2).

[14] Popov, E. and Nevière, M. "Maxwell equations in Fourier space: fast-converging formulation for diffraction by arbitrary shaped, periodic, anisotropic media". In: *J. Opt. Soc. Am. A* 18.11 (Nov. 2001), pp. 2886–2894 (cit. on pp. 2, 29, 111).

[15] Li, L. "Fourier modal method for crossed anisotropic gratings with arbitrary permittivity and permeability tensors". In: *J. Opt. A: Pure Appl. Opt.* 5.4 (2003), p. 345 (cit. on pp. 2, 44).

[16] Schuster, T., Ruoff, J., Kerwien, N., Rafler, S., and Osten, W. "Normal vector method for convergence improvement using the RCWA for crossed gratings". In: *J. Opt. Soc. Am. A* 24.9 (Sept. 2007), pp. 2880–2890 (cit. on pp. 2, 29).

[17] Auer, M. and Brenner, K.-H. "Localized input fields in rigorous coupled-wave analysis". In: *J. Opt. Soc. Am. A* 31.11 (2014), pp. 2385–2393 (cit. on pp. 2, 3, 30, 32, 62, 64, 91, 103, 112, 118, 130).

[18] Semenikhin, I. and Zanuccoli, M. "Application of the iterative approach to modal methods for the solution of Maxwell's equations". In: *J. Comput. Phys.* 300.Supplement C (2015), pp. 438–454 (cit. on p. 2).

[19] Yee, K. "Numerical solution of inital boundary value problems involving maxwell's equations in isotropic media". In: *IEEE Trans. Antennas Propag.* 14 (May 1966), pp. 302–307 (cit. on pp. 2, 18, 20, 24).

[20] Pohl, D. W., Denk, W., and Lanz, M. "Optical stethoscopy: Image recording with resolution $\lambda/20$". In: *Appl. Phys. Lett.* 44.7 (1984), pp. 651–653 (cit. on pp. 3, 64).

[21] Kogelnik, H. "Coupled wave theory for thick hologram gratings". In: *Bell Syst. Tech. J.* Vol. 48. 9. IEEE, 1969, pp. 2909–2947 (cit. on pp. 3, 82).

[22] Huang, W.-P. "Coupled-mode theory for optical waveguides: an overview". In: *J. Opt. Soc. Am. A* 11.3 (Mar. 1994), pp. 963–983 (cit. on pp. 3, 82).

[23] Goodman, J. W. *Introduction to Fourier Optics*. McGraw-Hill Series in Electrical and Computer Engineering: Communications and Signal Processing. McGraw-Hill, 1996 (cit. on pp. 5, 16, 19, 20, 22).

[24] Saleh, B. E. A. and Teich, M. C. *Fundamentals of Photonics*. Hoboken, New Jersey, USA: Wiley, 2007 (cit. on pp. 5, 78, 82, 84).

[25] Bradbury, S. *The Evolution of the Microscope*. London: Pergamon Press, 1965 (cit. on p. 5).

[26] Lindberg, D. C. *Theories of Vision from Al-Kindi to Kepler*. Chicago: University of Chicago Press, 1976 (cit. on p. 5).

[27] Kepler, J., Dyck, W. von, Caspar, M., Hammer, F., and Wissenschaften, B. A. der. *Johannes Keplers gesammelte Werke: Astronomiae pars optica*. Gesammelte Werke. C. H. Beck, 1938 (cit. on p. 6).

[28] Drake, S. *Galileo: Pioneer Scientist*. Toronto: The University of Toronto Press, 1990 (cit. on p. 6).

[29] Kwan, A., Dudley, J., and Lantz, E. "Who really discovered Snell's law?" In: *Phys. World* 15.4 (2002), p. 64 (cit. on p. 6).

[30] Darrigol, O. *A History of Optics: From Greek Antiquity to the Nineteenth Century*. Oxford: Oxford University Press, 2012 (cit. on p. 6).

[31] Newton, I. *Opticks: or, a treatise of the reflexions, refractions, inflexions and colours of light. Also two treatises of the species and magnitude of curvilinear figures*. London: Sam Smith and Benjamin Walford, Royal Society of London, 1704 (cit. on p. 6).

[32] Huygens, C. *Traité de la lumière*. Leiden, Netherlands: Pierre van der Aa, 1690 (cit. on p. 6).

[33] Robinson, A. *The Last Man Who Knew Everything*. New York, USA: Pi Press, 2006, pp. 123–124 (cit. on p. 6).

[34] Fresnel, A.-J. "Mémoire sur la diffraction de la lumière". In: *Ann. Chim. Phys.* 1 (1816), pp. 239–81 (cit. on p. 6).

[35] Senarmont, H. de, Verdet, E., and Fresnel, L. *Mémoire sur la diffraction de la lumière*. Impr. impériale, 1866-1870 (cit. on p. 6).

[36] Œrsted, Arago, Ampère, Davy, H., Biot, Erman, Schweiger, and Rive, D. L. *Exposé des nouvelles découvertes sur l'électricité et le magneétisme*. Paris, France: Méquignon-Marvis, 1822 (cit. on p. 6).

[37] Faraday, M. and Day, P. *The philosopher's tree: a selection of Michael Faraday's writings*. London: CRC Press, 1999 (cit. on p. 6).

[38] Maxwell, J. C. "On physical lines of force". In: *Philos. Mag.* 90 (Feb. 2010), pp. 11–23 (cit. on p. 6).

[39] Maxwell, J. C. "A Dynamical Theory of the Electromagnetic Field". In: *Philos. Trans. Royal Soc. Lond. Series I* 155 (1865), pp. 459–512 (cit. on p. 6).

[40] Heaviside, O. "The Induction of Currents in Cores". In: *The Electrician* 13 (1884), pp. 133–4 (cit. on p. 6).

[41] Gibbs, J. W. "On multiple algebra". In: *Proceedings of the American Association for the Advancement of Science*. Vol. 35. 1886, pp. 37–66 (cit. on p. 6).

[42] Hertz, H. "Über die Grundgleichungen der Elektrodynamik für ruhende Körper". In: *Ann. Phys.* 276.8 (1890), p. 577 (cit. on p. 6).

[43] Hertz, H. "Über die Grundgleichungen der Elektrodynamik für bewegte Körper". In: *Ann. Phys.* 277.11 (1890), pp. 369–399 (cit. on p. 6).

[44] Abramowitz, M. and Stegun, I. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications, 1964 (cit. on p. 17).

[45] Courant, R. and Hilbert, D. *Methods of Mathematical Physics*. Bd. 1. Wiley, 2008 (cit. on p. 18).

[46] Kirchhoff, G. "Zur Theorie der Lichtstrahlen". In: *Ann. Phys.* 254 (1882), pp. 663–695 (cit. on p. 18).

[47] Sommerfeld, A. (cit. on pp. 18, 20, 21).

[48] Scheffers, H. "Vereinfachte Ableitung der Formeln für die Fraunhoferschen Beugungserscheinungen". In: *Ann. Phys.* 434 (1942), pp. 211–215 (cit. on pp. 18–20).

[49] Bouwkamp, C. J. "Diffraction Theory". In: *Rep. Prog. Phys.* 17 (1954), pp. 35–100 (cit. on pp. 18–20).

[50] Weyl, H. "Ausbreitung elektromagnetischer Wellen über einem ebenen Leiter". In: *Ann. Phys.* 365.21 (1919), pp. 481–500 (cit. on pp. 18, 21).

[51] Feit, M. D. and Fleck, J. A. "Light propagation in graded-index optical fibers". In: *Appl. Opt.* 17.24 (Dec. 1978), pp. 3990–3998 (cit. on pp. 18, 20, 23, 89).

[52] Brenner, K.-H. and Singer, W. "Light propagation through microlenses: a new simulation method". In: *Appl. Opt.* 32.26 (Sept. 1993), pp. 4984–4988 (cit. on pp. 18, 20, 24, 89, 106, 120).

[53] Liu, J. M. and Gomelsky, L. "Vectorial beam propagation method". In: *J. Opt. Soc. Am. A* 9.9 (Sept. 1992), pp. 1574–1585 (cit. on pp. 18, 20).

[54] Fertig, M. and Brenner, K.-H. "Vector wave propagation method". In: *J. Opt. Soc. Am. A* 27.4 (Apr. 2010), pp. 709–717 (cit. on pp. 18, 20).

[55] Kerwien, N., Schuster, T., Rafler, S., Osten, W., and Totzeck, M. "Vectorial thin-element approximation: a semirigorous determination of Kirchhoff's boundary conditions". In: *J. Opt. Soc. Am. A* 24.4 (Apr. 2007), pp. 1074–1084 (cit. on p. 20).

[56] Liu, P. and Lü, B. "The vectorial angular-spectrum representation and Rayleigh-Sommerfeld diffraction formulae". In: *Opt. Laser Technol.* 39.4 (2007), pp. 741–744 (cit. on pp. 19, 20, 22).

[57] Korpel, A., Lin, H. H., and Mehrl, D. J. "Convenient operator formalism for Fourier optics and inhomogeneous and nonlinear wave propagation". In: *J. Opt. Soc. Am. A* 6.5 (May 1989), pp. 630–635 (cit. on pp. 20, 23).

[58] Lawrence, G. N. "Optical Modeling". In: *Applied Optics and Optical Engineering, Volume XI*. Ed. by Shannon, R. R. and Wyant, J. C. Vol. 11. 1992, p. 125 (cit. on pp. 19, 22).

[59] Schmidt, S., Tiess, T., Schröter, S., Hambach, R., Jäger, M., Bartelt, H., Tünnermann, A., and Gross, H. "Wave-optical modeling beyond the thin-element-approximation". In: *Opt. Express* 24.26 (Dec. 2016), pp. 30188–30200 (cit. on p. 24).

[60] Birdsall, C. and Langdon, A. *Plasma Physics via Computer Simulation*. Series in Plasma Physics and Fluid Dynamics. Taylor & Francis, 2004 (cit. on p. 25).

[61] Courant, R., Friedrichs, K., and Lewy, H. "Über die partiellen Differenzengleichungen der mathematischen Physik". In: *Math. Ann.* 100 (1928), pp. 32–74 (cit. on p. 25).

[62] Brenner, K.-H. "Normalized electromagnetic fields". In: *Annual Report 2015-2016 of the Chair of Optoelectronics (Institute of Computer Engineering, Universität Heidelberg)* (2017), p. 8 (cit. on p. 26).

[63] Auer, M. "Numerical treatment of localized fields in rigorous diffraction theory and its application to light absorption in structured layers". PhD thesis. Universität Heidelberg, 2016 (cit. on pp. 27, 30, 40, 62, 79, 81).

[64] Schuster, T. "Simulation von Lichtbeugung an Kreuzgitter-Strukturen und deren Anwendung in der Scatterometrie". PhD thesis. Universität Stuttgart, Institut für Technische Optik, 2010 (cit. on p. 29).

[65] Wu, S.-D., Gaylord, T. K., Glytsis, E. N., and Wu, Y.-M. "Three-dimensional converging–diverging Gaussian beam diffraction by a volume grating". In: *J. Opt. Soc. Am. A* 22.7 (July 2005), pp. 1293–1303 (cit. on p. 32).

[66] Rumpf, R. C. "Improved formulation of scattering matrices for semi-analytical methods that is consistent with convention". In: *Prog. Electromagn. Res. B* 13.35 (2011), pp. 241–261 (cit. on pp. 57–59, 97, 98).

[67] Redheffer, R. "Difference equations and functional equations in transmission-line theory". In: *Modern Mathematics for the Engineer*. Ed. by Beckenbach, E. F. McGraw-Hill, New York, 1961. Chap. 12, pp. 282–337 (cit. on p. 59).

[68] Li, L. "Bremmer series, R-matrix propagation algorithm, and numerical modeling of diffraction gratings". In: *J. Opt. Soc. Am. A* 11.11 (Nov. 1994), pp. 2829–2836 (cit. on p. 59).

[69] Brenner, K.-H. "Verification of Near-field Calcuations by Conservation Laws". In: *15th Workshop on Information Optics (WIO)*. IEEE, July 2016 (cit. on p. 62).

[70] Dürig, U., Pohl, D. W., and Rohner, F. "Near field optical scanning microscopy". In: *J. Appl. Phys.* 59.10 (1986), pp. 3318–3327 (cit. on p. 64).

[71] Lamela, J., Jaque, F., Cantelar, E., Jaque, D., Kaminskii, A. A., and Lifante, G. "BPM simulation of SNOM measurements of waveguide arrays induced by periodically poled BNN crystals". In: *Opt. Quantum Electron.* 39.10 (Aug. 2007), pp. 805–811 (cit. on p. 64).

[72] Sawada, K., Nakamura, H., Maruoka, T., Tamura, Y., Imura, K., Saiki, T., and Okamoto, H. "FDTD Simulated Observation of a Gold Nanorod by Scanning Near-Field Optical Microscopy". In: *J. Plasma Fusion Res.* 5 (Jan. 2010) (cit. on p. 64).

[73] Lalanne, P. and Silberstein, E. "Fourier-modal methods applied to waveguide computational problems". In: *Opt. Lett.* 25.15 (Aug. 2000), pp. 1092–1094 (cit. on p. 69).

[74] Grau, G. and Freude, W. *Optische Nachrichtentechnik*. Berlin: Springer, 1991 (cit. on p. 78).

[75] Bao, C., Srivastava, R., Gómez-Reino, C., and Pérez, M. V. "Gaussian beam propagation and imaging in hyperbolic secant profile media". In: *J. Opt. A: Pure Appl. Opt.* 5.1 (1996), p. 15 (cit. on p. 86).

[76] Hell, S. and Stelzer, E. H. "Properties of a 4Pi confocal fluorescence microscope". In: *J. Opt. Soc. Am. A* 9.12 (1992), pp. 2159–2166 (cit. on pp. 91, 100, 101).

[77] Hell, S., Stelzer, E. H., Lindek, S., and Cremer, C. "Confocal microscopy with an increased detection aperture: type-B 4Pi confocal microscopy". In: *Opt. Lett.* 19.3 (1994), pp. 222–224 (cit. on pp. 91, 100, 101).

[78] Mansuripur, M. "Distribution of light at and near the focus of high-numerical-aperture objectives". In: *J. Opt. Soc. Am. A* 3.12 (Dec. 1986), pp. 2086–2093 (cit. on pp. 92, 95).

[79] Mansuripur, M. "Certain computational aspects of vector diffraction problems". In: *J. Opt. Soc. Am. A* 6.6 (June 1989), pp. 786–805 (cit. on p. 92).

[80] Flagello, D. G., Milster, T., and Rosenbluth, A. E. "Theory of high-NA imaging in homogeneous thin films". In: *J. Opt. Soc. Am. A* 13.1 (Jan. 1996), pp. 53–64 (cit. on p. 92).

[81] Brenner, K.-H. "Localized light sources in the exact optical simulation of multilayer stacks". In: *Annual Report 2013 of the Chair of Optoelectronics (Institute of Computer Engineering, Universität Heidelberg)* (2014), p. 6 (cit. on p. 92).

[82] Born, M., Wolf, E., and Bhatia, A. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 2000 (cit. on p. 100).

[83] Schrader, M., Hell, S. W., and Voort, H. T. M. van der. "Three-dimensional super-resolution with a 4Pi-confocal microscope using image restoration". In: *J. Appl. Phys.* 84.8 (1998), pp. 4033–4042 (cit. on p. 101).

[84] Herten, D.-P. Institute for Physical Chemistry, Universität Heidelberg, Im Neuenheimer Feld 267, D-69120 Heidelberg, Germany (personal communication, 2017). 2017 (cit. on p. 103).

[85] Gould, T. J., Burke, D., Bewersdorf, J., and Booth, M. J. "Adaptive optics enables 3D STED microscopy in aberrating specimens". In: *Opt. Express* 20.19 (2012), pp. 20998–21009 (cit. on p. 103).

[86] Popov, E. *Gratings: Theory and Numeric Applications*. Chapter 7: Differential Methods for Periodic Structures. Marseille, France: Presses Universitaires de Provence, 2014, p. 249 (cit. on pp. 110, 111).

[87] Arnoldi, W. E. "The principle of minimized iterations in the solution of the matrix eigenvalue problem". In: *Quart. Appl. Math.* 9 (1951), pp. 17–29 (cit. on pp. 111, 114, 116).

[88] Apostol, T. *Calculus*. Chapter 7.6. New York: Wiley, 1967 (cit. on p. 111).

[89] Davis, P. *Circulant Matrices*. AMS Chelsea Publishing Series. Chelsea, 1994 (cit. on pp. 112, 124).

[90] Süli, E. and Mayers, D. *An Introduction to Numerical Analysis*. Cambridge: Cambridge University Press, 2003 (cit. on p. 113).

[91] Stenau, T. and Brenner, K.-H. "Diffractive Lenses with Overlapping Aperture - A New Tool in Scanning Microscopy". In: *Imaging and Applied Optics - Imaging Systems and Applications*. Optical Society of America, July 2016 (cit. on p. 115).

[92] *Refractive index database*. https://refractiveindex.info/. 2018 (cit. on pp. 115, 119).

[93] Schmidt, S., Tiess, T., Schröter, S., Hambach, R., Jäger, M., Bartelt, H., Tünnermann, A., and Gross, H. "Wave-optical modeling beyond the thin-element-approximation". In: *Opt. Express* 24.26 (Dec. 2016), pp. 30188–30200 (cit. on p. 120).

[94] *COMSOL Multiphysics*. https://www.comsol.com/. 2016 (cit. on p. 120).

[95] Zernike, F. "How I Discovered Phase Contrast". In: *Science* 121.3141 (1955), pp. 345–349 (cit. on p. 122).

[96] Ghosh, G. "Dispersion-equation coefficients for the refractive index and birefringence of calcite and quartz crystals". In: *Opt. Commun.* 163.1 (1999), pp. 95–102 (cit. on p. 125).

B | Own publications

[A1] Junker, A. and Brenner, K.-H. "Simulation and Analysis of SNOM Measurements using Rigorous Coupled-Wave Analysis". In: *DGaO Proceedings 2015*. DGaO, July 2015 (cit. on pp. 3, 64).

[A2] Junker, A. and Brenner, K.-H. "Structuring the incident and transmitted regions in rigorous coupled-wave analysis". In: *15th Workshop on Information Optics (WIO)*. IEEE, July 2016 (cit. on pp. 3, 69).

[A3] Junker, A. and Brenner, K.-H. "Two-sided illumination in rigorous coupled-wave analysis applied to the 4π-microscope". In: *J. Opt. Soc. Am. A* 34.10 (Oct. 2017), pp. 1769–1775 (cit. on pp. 3, 32, 60, 61, 91).

[A4] Junker, A. and Brenner, K.-H. "Bidirectional light incidence in rigorous coupled-wave analysis applied to 4π-microscopy". In: *16th Workshop on Information Optics (WIO)*. IEEE, Sept. 2017 (cit. on pp. 3, 91).

[A5] Junker, A. and Brenner, K.-H. "Achieving a high mode count in the exact electromagnetic simulation of diffractive optical elements". In: *J. Opt. Soc. Am. A* 35.3 (Mar. 2018), pp. 377–385 (cit. on pp. 4, 105).

[A6] Junker, A. and Brenner, K.-H. "High mode count rigorous simulation of diffractive optical elements by an iterative solution approach". In: *EOS Topical Meeting on Diffractive Optics 2017*. EOS, Sept. 2017, pp. 23–24 (cit. on pp. 4, 105).

[A7] Junker, A., Stenau, T., and Brenner, K.-H. "Scalar wave-optical reconstruction of plenoptic camera images". In: *Appl. Opt.* 53.25 (Sept. 2014), pp. 5784–5790.

C | Acknowledgments

The time as a PhD student at the Institute of Computer Engineering (ZITI) of the Universität Heidelberg provided numerous possibilities for my personal and professional development. I would like to take the opportunity to express my deep gratitude to all the people who, directly or indirectly, contributed to this work, and to those who accompanied and supported me throughout the past years.

First of all, I am deeply indebted to Professor Dr. Karl-Heinz Brenner for giving me the opportunity to join his group at the Chair of Optoelectronics as a PhD student. Thank you for providing such an interesting topic, which turned out to offer numerous prospects for new developments. Thank you also for leaving me so much freedom in choosing the focus of my work, and for being passionate and excited about all new findings no matter what. I also greatly appreciated the open-door policy, which enabled many of our fruitful discussions in the first place. Thank you for the support you gave me in every aspect of my work.

I also want to express my gratitude to Professor Dr. Frank Wyrowski from the Friedrich-Schiller-Universität Jena, who I consider myself lucky to have met on several conferences, and who I always experienced as a passionate and fair discussion partner. Thank you for agreeing to be the second reviewer of my PhD thesis.

Furthermore, I am grateful to all the wonderful group members at the Chair of Optoelectronics, especially Tim Stenau, Max Auer, Lisa Tuschner, Xiyuan Liu and Torsten Paul, for their friendship, and, of course, also for their time and expertise when needed. Many thanks also go to Wolfgang Stumpfs, who possesses an incredible knowledge and experience as a precision mechanic. Thank you for showing me so many tips and tricks in the context of metal processing, instructing me in the operation of the lathe, the CNC milling machine, etc. The days in the workshop were definitely among the most enjoyable! Beyond that, Adriana and I had a lot of fun at the Weinheimer See, where we went diving for the first time in our lives. I hope that, in turn, you had a lot of fun flying the touring motor glider as a copilot at the airfield Herrenteich! Furthermore, I would like to especially thank Sabine Volk, who organized about everything from our conference travels to joint group events. Thank you also for your advice regarding all university related questions. Thank you to all the chair members for the always relaxed atmosphere and the numerous decent

conversations, which often went beyond the scope of work. Without you, especially lunchtime would have been only half as entertaining!

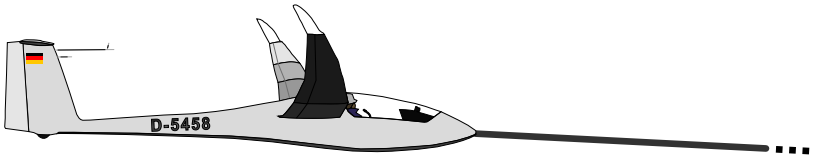
I would also like to thank my former high school teachers, Gisela Roquette, Alfred Seßler, Norbert Roßwag, Carl-Julian Pardall and Winfried Krause from Hebel-Gymnasium Schwetzingen, and Janet Wardall and Greg Eichelkraut from Merrill High School for their tremendous commitment and fantastic classes. Without your dedication, I would probably not have started to study a natural science in the first place. Therefore, this work is also yours!

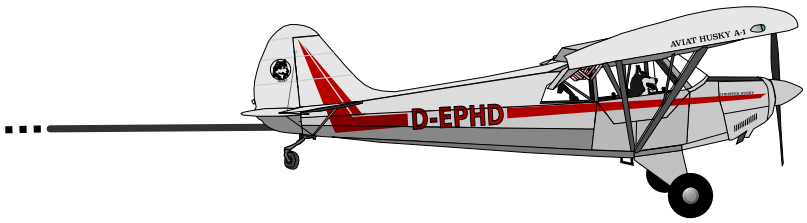
Furthermore, my PhD studies would have been by far harder - if not impossible - without my best friends, who always provided plenty of distraction to divert my attention away from work towards the things that are of real importance in life: flying, skiing, and having fun. At this point, I would like to especially mention Andreas, Patrick and all members of the 'Herrenteich skiing [and flying] club'.

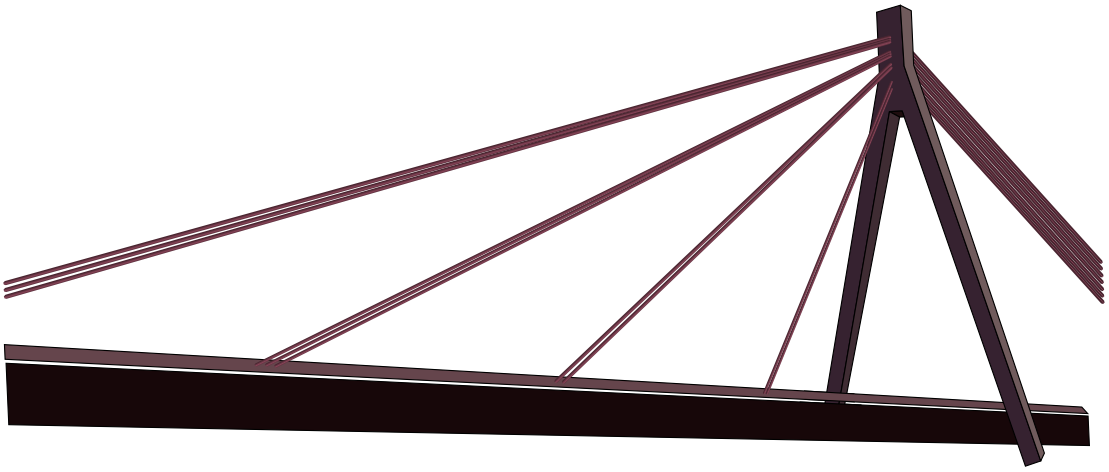
Last but not least, I would like to say thank you to my German and American families for all the love and support throughout the years. Thank you for everything you have done for me, for caring, for all the good advice, and for always being interested in what I do and what I think. Particularly, I would like to thank Adriana for sharing so many lasting experiences with me, be it in our free time, at home, on vacation, ... I consider myself incredibly fortunate to know you and to be the one you spend your life with!

Heidelberg, February 2018 André Junker

Wenn du weißt, wo du bist,
kannst du sein, wo du willst.







Wenn du nicht weißt, wo du bist,
musst du sehen, wo du bleibst!