

Forward diffraction modelling : analysis and application to grating reconstruction

Citation for published version (APA):

Kraaij, van, M. G. M. M. (2011). *Forward diffraction modelling : analysis and application to grating reconstruction*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Mathematics and Computer Science]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR702579>

DOI:

[10.6100/IR702579](https://doi.org/10.6100/IR702579)

Document status and date:

Published: 01/01/2011

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Forward Diffraction Modelling:
Analysis and Application to
Grating Reconstruction

Copyright ©2011 by M.G.M.M. van Kraaij, Eindhoven, The Netherlands.
All rights are reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of the author.

Cover shows a computer rendering of 2 slit optical interference pattern (Copyright by J.K.N. Murphy, Auckland, New Zealand)

A catalogue record is available from the Eindhoven University of Technology Library
Proefschrift. - ISBN: 978-90-386-2447-1

NUR 919

Subject headings: boundary value problems; differential operators; eigenvalue problems / electromagnetic waves; diffraction / electromagnetic scattering; numerical methods

The work described in this thesis has been carried out under the auspices of

 **ASML** - Veldhoven, The Netherlands.

Forward Diffraction Modelling: Analysis and Application to Grating Reconstruction

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College
voor Promoties in het openbaar te verdedigen
op dinsdag 22 maart 2011 om 16.00 uur

door

Markus Gerardus Martinus Maria van Kraaij

geboren te Mierlo

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. R.M.M. Mattheij

Copromotor:

dr. J.M.L. Maubach

Nomenclature

General

a	physical vector	$\mathbf{a} \cdot \mathbf{b}$	dot product of vectors
\mathbf{a}	mathematical vector	$\mathbf{a} \times \mathbf{b}$	cross product of vectors
\mathbf{a}	mathematical block vector	$ a $	absolute value of scalar
\mathbf{A}	mathematical matrix	$\llbracket \mathbf{a} \rrbracket$	jump of vector
\mathcal{A}	mathematical block matrix		
$\text{Re}[\]$	real part	x, y, z	Cartesian coordinates
$\text{Im}[\]$	imaginary part	j	imaginary unit
$\langle \cdot, \cdot \rangle_w$	weighed inner product of functions		

Operators

∇	gradient operator	$\nabla \times$	curl operator
$\nabla \cdot$	divergence operator	∇^2	Laplace operator
$\nabla_S \cdot$	surface divergence operator	$\frac{\partial}{\partial a}$	partial derivative to a

Subscripts

I	superstrate material index	i	layer index
II	substrate material index	m	harmonic/diffraction order index
B	Bloch method	s, p	s -, p -polarised part
R	RCWA method	x, y, z	x -, y -, z -component
S	surface		

Superscripts

\mathbf{a}^m	m^{th} block row of vector	e	electric
$\mathcal{A}^{m:}$	m^{th} block row of matrix	h	magnetic
$\mathcal{A}^{:n}$	n^{th} block column of matrix	i	incident
\mathcal{A}^{mn}	m^{th} block row and n^{th} block column of matrix	r	reflected
hom	homogeneous	t	transmitted
part	particular	\perp	perpendicular
r	relative	+	positive z-direction
T	transposed	-	negative z-direction
H	complex conjugate transposed	'	derivative

Overscripts

$\bar{}$	conjugate	$\hat{}$	periodic eigenvalue problem
$\tilde{}$	conical diffraction	$\check{}$	semi-periodic eigenvalue problem

Greek symbols

$\Delta\kappa^2$	difference Bloch wave vector components	μ	permeability
ε	permittivity	μ_0	permeability of vacuum
ε_0	permittivity of vacuum	ν	auxiliary refraction index variable
ζ	asymptote increasing transcendental eigenvalue function	ξ	asymptote decreasing transcendental eigenvalue function
η	diffraction efficiency	ρ	charge density phasor
θ	polar angle of incidence	ϱ	charge density
κ	Bloch wave vector component	σ	electric conductivity
λ_0	vacuum wavelength	ϕ	azimuthal angle of incidence
Λ	grating pitch	ψ	polarisation angle
		ω	angular frequency

Roman symbols

a	even basis solution Bloch mode	M	total number of harmonics
A	expansion coefficient of even basis solution Bloch mode	n	refraction index
b	odd basis solution Bloch mode	n'	real part refraction index
B	expansion coefficient of odd basis solution Bloch mode	n''	imaginary part refraction index
B	magnetic flux phasor	\mathbf{n}	normal unit vector
\mathcal{B}	magnetic flux	\mathbf{p}	auxiliary unit vector in plane of incidence
D	total grating height	P	time-averaged energy flow
D	electric flux phasor	r	distance to the origin
\mathcal{D}	electric flux	\mathbf{r}	radial unit vector
e_x, e_y, e_z	Cartesian unit vectors	R	expansion coefficient of reflected Rayleigh mode
E	electric field phasor	s	auxiliary unit vector normal to plane of incidence
\mathcal{E}	electric field	S	Poynting vector
f	scalar field	t	time variable
F	electromagnetic field phasor	T	expansion coefficient of transmitted Rayleigh mode
\mathcal{F}	electromagnetic field	u	eigenfunction x -direction
h	layer thickness	v	eigenfunction z -direction
H	magnetic field phasor	w	weight function
\mathcal{H}	magnetic field	W	Wronskian
J	electric current phasor	\mathbf{x}	position vector
\mathcal{J}	electric current	X	layer offset
k_0	wavenumber of vacuum	Y	admittance
k	wave vector	Z	layer height
K	total number of layers		
L	total number of offsets		

Mathematical vectors and matrices

\mathbf{A}	matrix to be diagonalised TE
\mathcal{A}	auxiliary matching matrix grating layers
\mathbf{B}	part of matrix to be diagonalised TM
\mathbf{C}	\mathbf{c} expansion coefficient basis solution z -direction
\mathbf{C}	part of matrix to be diagonalised TM
δ	\mathbf{d} \mathbf{d} expansion coefficient incident field
\mathbf{D}	auxiliary matrix TM polarisation
ϵ	\mathbf{E} fourier coefficient relative permittivity
\mathbf{F}	\mathcal{F} auxiliary matrix Riccati recursion or fundamental solution
\mathbf{G}	\mathcal{G} auxiliary matrix Riccati recursion or fundamental solution
\mathbf{H}	auxiliary matrix fundamental solution at bottom interface
\mathbf{I}	\mathbf{I} identity matrix

		\mathcal{J}	auxiliary anti-diagonal unit matrix
k	\mathbf{K}	\mathcal{K}	wave vector component
λ	\mathbf{L}	\mathcal{L}	auxiliary eigenvalue variable
μ	\mathbf{M}	\mathcal{M}	square root of eigenvalue
π	\mathbf{P}		fourier coefficient reciprocal relative permittivity
q	\mathbf{Q}	\mathcal{Q}	RCWA eigenvector component or Bloch coupling coefficient
R	\mathbf{r}		expansion coefficient reflected field
	\mathbf{R}	\mathcal{R}	Riccati transformation matrix one- and two-stage approach
	\mathbf{S}	\mathcal{S}	Riccati transformation matrix two-stage approach
T	\mathbf{t}		expansion coefficient transmitted field
	\mathbf{U}	\mathcal{U}	upper triangular matrix one- and two-stage approach
	\mathbf{v}		solution component z-direction
	\mathbf{V}	\mathcal{V}	upper triangular matrix two-stage approach
ω	\mathbf{W}	\mathcal{W}	fourier coefficient weight function
x	\mathbf{X}	\mathcal{X}	auxiliary coefficient exponentially decaying wave

Table 1: The first column gives the notation for the coefficient, the second column for the vector or matrix and the third column for the block vector or block matrix.

Abbreviations

a-diag	anti-diagonal	IVP	Initial Value Problem
ASR	Adaptive Spatial Resolution	NA	Numerical Aperture
BARC	Backward Anti-Reflective Coating	ODE	Ordinary Differential Equation
BVP	Boundary Value Problem	PDE	Partial Differential Equation
CD	Critical Dimension	RCWA	Rigorous Coupled-Wave Analysis
diag	diagonal	SEM	Scanning Electron Microscope
FD	Finite Differences	TE	Transverse Electric
FDTD	Finite Difference Time Domain	TM	Transverse Magnetic
FFT	Fast Fourier Transform		

Contents

Nomenclature	viii
1 Introduction	1
1.1 The electronics industry	1
1.2 Fabrication of a chip using lithography	2
1.3 Problem description	4
1.4 Outline of the thesis	6
2 Model for a diffraction problem	7
2.1 Short historic overview of optics and electromagnetism	7
2.2 Mathematical model	9
2.2.1 Maxwell's equations	9
2.2.2 Constitutive relations and boundary conditions	10
2.3 Basic assumptions and reduced model	14
2.3.1 Planar diffraction TE polarisation	17
2.3.2 Planar diffraction TM polarisation	18
2.3.3 Conical diffraction	19
2.4 Derivation Rayleigh modes outside grating layers	20
2.4.1 Energy flow and the Poynting vector	25
3 Discretisation with Bloch modes	27
3.1 Derivation Bloch modes inside grating layers	27
3.1.1 Planar diffraction	28
3.1.2 Conical diffraction	31
3.2 Properties transcendental equation	34
3.3 Matching at the interfaces and the coupling matrix	39
3.3.1 Planar diffraction	39
3.3.2 Conical diffraction	43
3.4 Generalisation of the Bloch mode method	49
4 Discretisation with RCWA	53
4.1 Derivation RCWA modes inside grating structure	53

4.1.1	Planar diffraction	54
4.1.2	Conical diffraction	57
4.2	Matching at the interfaces	59
4.2.1	Planar diffraction	59
4.2.2	Conical diffraction	63
5	Solution strategies for the truncated linear system	69
5.1	Unstable transfer matrix algorithm	69
5.2	Stable recursion with Riccati transformations	72
5.2.1	One-stage Riccati approach	72
5.2.2	Two-stage Riccati approach	75
5.2.3	Riccati versus enhanced transmittance matrix approach	78
5.3	Connection with multiple shooting and stabilised march	79
6	Modifications and improvements for RCWA	85
6.1	Adaptive spatial resolution	86
6.1.1	Overview of necessary changes to RCWA	86
6.2	Finite difference discretisation	91
6.2.1	Partitioned domain approach	92
6.2.2	Single domain approach	101
7	Numerical results	103
7.1	Comparison of the forward diffraction models and their improvements	103
7.2	Application of the forward diffraction model to grating reconstruction	111
8	Conclusions and future work	115
A	Quantities related to the Rayleigh expansion	119
A.1	Derivation of the perpendicular s -polarised and p -polarised parts	119
A.2	Derivation of the energy flow	120
B	ASR coupling matrix	123
B.1	Numerical quadrature	123
B.2	Fast Fourier Transform	124
B.3	Bessel related special functions	124
C	Alternative ASR coordinate transformations	127
	Bibliography	131
	Index	135
	Summary	137
	Samenvatting	139

Contents

xi

Acknowledgements

141

Curriculum vitae

143

Chapter 1

Introduction

First a brief overview of the electronic industry is given in Section 1.1. Then in Section 1.2 the fabrication process of a chip is discussed. This process faces many challenges of which one in particular is explained in Section 1.3. A more detailed outline of this thesis is given in Section 1.4.

1.1 The electronics industry

Today's world of electronics is so large that over 1000 billion US dollars a year is spent on electronic applications. Electronic devices like computers, tv's, mobile phones and cameras are all examples of this multi-billion dollar business. Almost all electronic equipment these days contains chips and its market is valued in the order of 200 to 300 billion dollars. For the production of these chips many different production systems are used. One of these systems are so-called lithography systems which are sold by companies like ASML. These systems which are indicated at the bottom of Figure 1.1 perform the most critical step in the fabrication process of chips.

The advances in chip technology over the past years has been captured by the famous *Moore's law*. This law states that approximately every 1.5 to 2 years the number of transistors on a chip doubles. This law is named after Intel's co-founder Gordon E. Moore, who described this in a paper back in 1965. Moore's law also implies that every 1.5 to 2 years the computing power per chip doubles at roughly equivalent power consumptions for half the price. In order to keep following this trend, the lithography systems have become very complex systems over the years. Figure 1.2 shows a close-up of a chip revealing its complex layered structure. The next section briefly explains how these lithography systems are used in the fabricating process of today's chips.



Figure 1.1: Overview of some electronic devices containing chips fabricated with lithography systems

1.2 Fabrication of a chip using lithography

Building *integrated circuits (ICs)* to form a chip is a very complex process containing many different steps that require a high level of precision. The whole process starts with *silicon*, the basic constituent of sand. Silicon is what is called a *semiconductor*, under certain conditions it conducts electricity while under other conditions it does not. This allows the material to act as a switch and the basic component that makes use of this property is a *transistor*. By combining millions of these transistors using *interconnects* complex ICs can be fabricated.

The first step in this fabrication process is purifying the silicon since impurities at this stage could render the final chips useless. Then the purified silicon is melted and grown into cylinders typically 300mm in diameter. By slicing the cylinder into circular disks and polishing them to create an ultra-smooth surface, a silicon *wafer* is created which acts as a substrate. Then a photolithographic printing process is used to build several chips layer by layer onto this wafer. Figure 1.3 shows the process of adding one such layer, in this case a patterned oxide layer. The starting point (1) is the silicon substrate (or partially processed chip). Then a thin oxide layer is deposited (2) which acts as an insulator. This oxide layer is typically grown in a furnace at high temperature in the

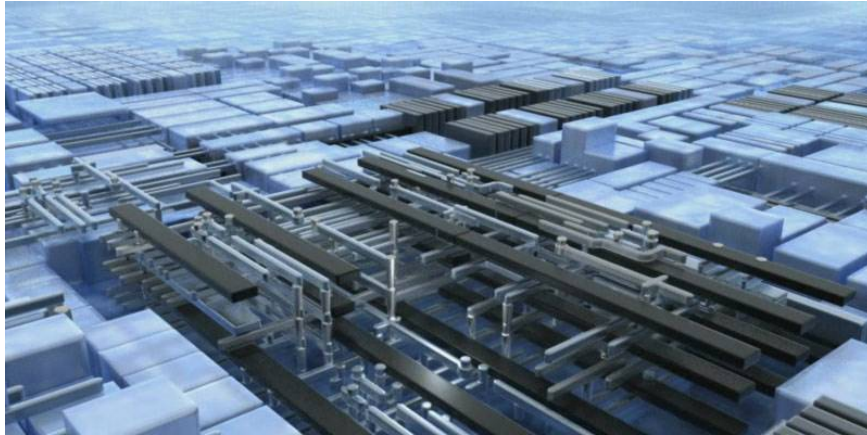


Figure 1.2: Close-up of a chip revealing its complex layered structure

presence of oxygen. In the next step (3) the oxide film is coated with a light-sensitive material called *photoresist*. This coating will be used in subsequent steps to remove certain sections of the underlying oxide thereby creating a specific oxide pattern. Photoresist is sensitive to ultraviolet light, yet resistant to certain etching chemicals required at a later stage. The important printing process (4) uses mask-pieces made of glass with both transparent and opaque areas. This mask essentially contains part of the circuit design of that specific layer. By shining ultraviolet light onto the mask which only passes through the transparent areas, the circuit design is transferred onto the photoresist. The parts of the photoresist that are exposed to the ultraviolet light become soluble and are removed using a solvent (5) revealing part of the oxide layer underneath. A chemical etching process (6) removes the exposed oxide while the remaining photoresist protects the unexposed areas and underlying oxide pattern. Finally the remaining protective photoresist is removed (7), leaving the desired oxide pattern on the silicon layer.

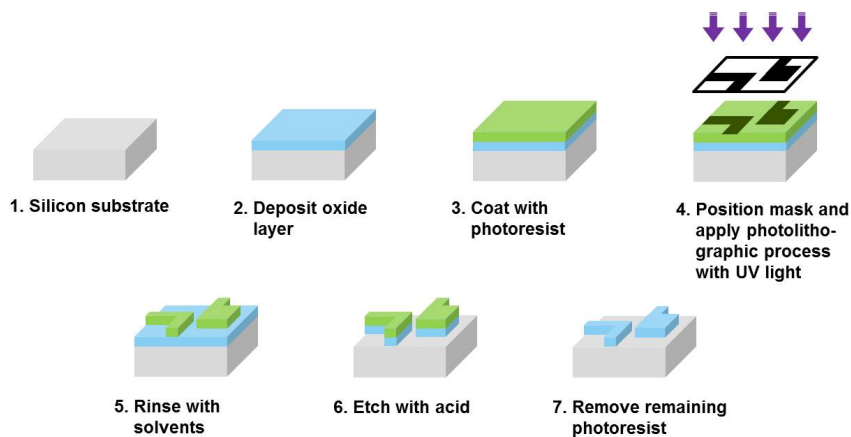


Figure 1.3: Process of adding a single patterned oxide layer

The finished oxide layer is just one part of the fabrication process. In order to arrive at a working chip many other steps are required, including

- **Adding more layers.** Other materials like for example polysilicon, which unlike oxide actually conducts electricity, are deposited on the wafer. Also now the steps of further film depositions, printing using masks and etching are required to transfer part of the circuit to these other layers. The total number of layers depends on the component being manufactured but typically lies in the order of 20-40.
- **Doping.** The *doping* process bombards the exposed areas of the silicon wafer with various chemical impurities, altering the way silicon conducts electricity in these areas. Doping is what turns silicon into silicon transistors, enabling the switching between two states, 1 for on and 0 for off.
- **Metallisation.** Layers of metal are applied to form the connections between the transistors. Typically copper is used because of its low resistance and cost effective integration into the fabrication process. The specific patterns in these metal layers can be formed using the photolithographic printing process described earlier. Also the bonding points to connect the chip to the outside world are made of metals.

Finally the whole wafer (if necessary) is planarised by chemical mechanical polishing, given a protective layer and tested to ensure the circuits are working as intended. If successful, the wafer containing the working ICs is cut and the individual ICs are mounted on supports before being packaged.

1.3 Problem description

In the previous section the fabrication process of a chip revealed that a chip consist of around 20-40 different layers containing all kinds of complex patterns. Moreover these layers which are stacked on top of each other are not processed at the same time but one after the other. This means that every time the photolithographic process prints a pattern in a new layer, this needs to be done very precisely in order to obtain a working IC in the end. Not only should the new pattern itself be printed within tight specifications, it should also be aligned properly with the patterns in the underlying layer. Typically a lot of information on the lithography process can be obtained by measuring test structures or gratings which are scattered over the wafer. These *gratings* are tiny periodic structures much smaller than ICs. With today's tight requirements a dedicated metrology tool is used for measuring these extremely small features. First the gratings are illuminated and its response (a scattered intensity) is measured. For certain applications like *overlay metrology* the asymmetry in this measured signal (due to an offset between two gratings in different layers) can be used to align the lithographic process. For other applications like *critical dimension (CD) metrology* one is interested in the shape

of the grating lines that produced the measured signal. Since this information is not directly available but encrypted in the measurement, a reconstruction algorithm is used to extract it. The reconstructed values like height, width and sidewall angle can then be related to machine settings like dose and focus which control the lithographic process.

Figure 1.4 shows an example of a grating produced at different focus levels of the machine. These images were taken with a scanning electron microscope (SEM) which is able to visualise the shape of these (trapezoidal looking) grating lines directly. The im-

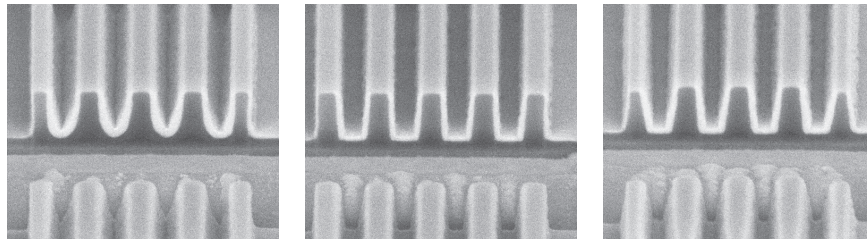


Figure 1.4: Scanning electron microscope images of a grating through focus. The image in the middle corresponds to a grating that was produced while in focus (Source: ASML).

age in the middle corresponds to a grating that was produced while in focus and shows straight lines without too much rounding at the top and bottom. The other out of focus images typically correspond to a process that falls outside of the specifications. It is therefore crucial to know the shape of these test structures so that proper action can be taken like for example rework or scraping of the current wafer, adjusting the machine settings for the next batch of wafers, etc...Note that the images in Figure 1.4 are here for illustrative purposes only since a SEM typically is a destructive, expensive and time-consuming measurement. However the metrology tool described in this thesis is non-destructive and fast but does not have direct access to this shape information. In particular this CD metrology application therefore requires rigorous mathematical models that solve optical diffraction problems for periodic gratings in combination with advanced reconstruction algorithms.

The research described in this thesis mainly focusses on the forward modelling part. A mathematical model describing the optical diffraction problem is derived and solved using several mode expansion techniques. Also the integration of this forward model into the CD metrology application is discussed. The second step of combining the forward model with a reconstruction algorithm has been addressed in [1]. There the inverse modelling part and corresponding sensitivity analysis revealed that the main computational burden is still taken up by solving many forward problems. Since the CD metrology application has very strict throughput requirements, the main goal of this thesis is therefore to develop algorithms that solve the forward problem as fast as possible while remaining both accurate and stable.

1.4 Outline of the thesis

In Chapter 2 the basic framework of (electromagnetic) diffraction theory is introduced. After explaining some general concepts of diffraction theory, the governing equations are established together with the constitutive relations and boundary conditions. For the specific diffraction problem that is considered in this thesis a reduced model is derived which is partly solved in closed form in the last subsection. The following two chapters discuss two mode expansion methods used to further discretise the reduced model. First the Bloch mode method is studied in Chapter 3 which is based on computing the exact eigenfunctions of the underlying reduced problem. This involves solving a complicated transcendental equation which for certain geometries can be solved efficiently. Finally a large linear system is derived by applying the matching boundary conditions. In Chapter 4 an alternative mode expansion method is studied, the Rigorous Coupled-Wave Analysis, which uses Fourier based expansion functions. Although these expansion functions only approximate the exact eigenfunctions of Bloch, RCWA is much more flexible and easier to generalise to more complicated geometries. A similar linear system is obtained from the matching boundary conditions.

Chapter 5 focusses on solving the large linear system, derived with both mode expansion methods, stably and efficiently. It is shown that standard techniques that do not take special care of the exponentially growing and decaying solution components are unstable. Therefore a stable algorithm is derived that not only decouples these solution components but also uses a two-stage approach for maximum efficiency. The link to other published algorithms and other standard techniques used for solving boundary value problems is discussed in the latter part of this chapter. Two modifications for the RCWA mode expansion method are the topic of Chapter 6. These modifications are aimed at improving the convergence of RCWA while maintaining its flexibility and relatively simple implementation. The first modification applies a coordinate transformation before using the regular Fourier discretisation whereas the second modification completely replaces the Fourier discretisation by a finite difference approximation. Both modifications try to get closer to the exact eigenfunctions of Bloch by properly taking care of material transitions in the underlying geometry (which is typically not done in standard RCWA).

The accuracy of both mode expansion methods and their modifications is evaluated in Chapter 7. For several representative diffraction configurations numerical results are presented. The integration of these forward diffraction models into a CD reconstruction application is addressed in the second part of this chapter. Finally Chapter 8 summarises the main results and discusses some future research topics.

Chapter 2

Model for a diffraction problem

To model the effect of an electromagnetic field on objects, we need a mathematical framework that describes the scattering of this field. In this thesis we consider objects with a characteristic length scale similar to that of the incident electromagnetic field. In that case the behaviour of the electromagnetic field is described by the theory of diffraction, which may be considered a special case of scattering. In Section 2.1 some general concepts of diffraction theory are explained in a short historical overview. Section 2.2 summarises the governing equations as well as the constitutive relations and boundary conditions. In Section 2.3 we introduce some simplifications to the model and derive a reduced model. This reduced model will be the basis for the different discretisations in the subsequent chapters. A part of the solution of the reduced model that is common to these different discretisations is derived in the final Section 2.4. Here also some quantities related to energy are introduced. These quantities are typically measured in a real life application and can also be used to check the quality and performance of an algorithm.

2.1 Short historic overview of optics and electromagnetism

An overview of important discoveries related to the development of our understanding of optical phenomena and electromagnetism can be found in [6]. Here we will briefly mention a few names and their important contributions to the field of optics and electromagnetism that form the basis of diffraction theory. This also gives us the opportunity to introduce some concepts frequently used in these fields and in this thesis.

The history of optics can be traced back to the Greek philosopher and mathematician

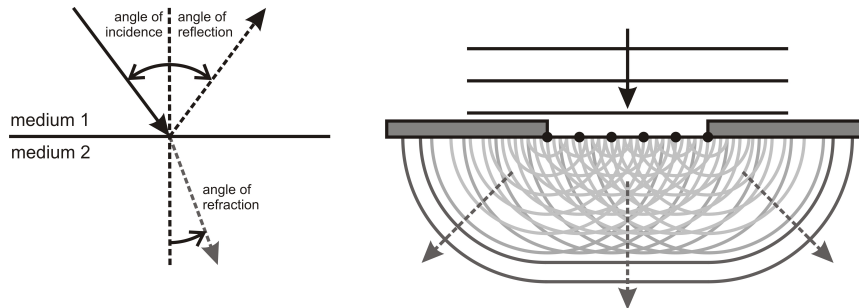


Figure 2.1: Law of reflection and refraction and the concept of interference

Euclid (c. 325-265 BC). He was one of the first to write down systematically his ideas on the propagation of light. The *law of (specular) reflection* was already known to the ancient Greeks. This law states that the direction of incoming light and the direction of reflected outgoing light make the same angle with respect to the surface normal. It was not until 1621 that Willebrord Snell (1580-1626) experimentally found the *law of refraction*. This law describes the relationship between the angles of incidence and refraction when light passes through a boundary between two different media. The first phenomenon of *interference*, the colours exhibited by thin films (Newton's rings), was discovered independently by Robert Boyle (1627-1691) and Robert Hooke (1635-1703). Around that time the wave theory of light was greatly improved by Christiaan Huygens (1629-1695) who was able to derive the laws of reflection and refraction with the *Huygens' principle*. This principle states that the wavefront of a propagating wave of light at any instant conforms to the envelope of spherical wavelets emanating from every point on the wavefront at the prior instant. This principle is also closely related to another aspect of interference, the addition (superposition) of two or more waves resulting in a new wave pattern. The wave theory by then was not able to describe the concept of *polarisation*, a property of transverse waves which describes the orientation of the oscillations in the plane perpendicular to the wave's direction of travel. Therefore the wave theory was rejected by Isaac Newton (1642-1727), which brought research on this topic to a stand still for nearly a century.

It was not until the beginning of the nineteenth century that important discoveries led to the acceptance of the revived wave theory. Although not generally accepted by the community, Thomas Young (1773-1829) supported and contributed to the wave theory explaining the principle of interference and the colours of thin films by new experiments. The breakthrough came in 1818 from Jean Fresnel (1788-1827) who combined the work of Huygens and Young into a general wave theory that could explain not only the rectilinear propagation of light but also the minute deviations from it, i.e. *diffraction* phenomena. Around that same time the phenomenon of polarisation, which had been observed by others before but not well understood, was put in the same general framework. Meanwhile the field of electricity and magnetism was developing into a leading science almost independently of optics, one of the exponents being Michael Faraday

(1791-1867). It was James Clerk Maxwell (1831-1879) who managed to summarise all previous experiences in this field in a system of equations, the famous Maxwell's equations. These equations describe the behaviour of electromagnetic waves propagating with a velocity which could be calculated from electrical measurements only. This velocity turned out to be equal to the speed of light verified by experiments in 1888 done by Heinrich Hertz (1857-1894). This led Maxwell to believe that light waves are electromagnetic waves, and thereby creating the link between the fields of optics and electromagnetism.

2.2 Mathematical model

2.2.1 Maxwell's equations

The behaviour of electromagnetic fields in diffraction theory is governed by Maxwell's equations. These equations in differential form give rise to a system of first-order partial differential equations (PDEs) that hold at every point in whose neighbourhood the physical properties of the medium are continuous

$$\nabla \times \mathcal{E}(\mathbf{x}, t) = -\frac{\partial}{\partial t} \mathcal{B}(\mathbf{x}, t), \quad (2.1a)$$

$$\nabla \times \mathcal{H}(\mathbf{x}, t) = \frac{\partial}{\partial t} \mathcal{D}(\mathbf{x}, t) + \mathcal{J}(\mathbf{x}, t), \quad (2.1b)$$

where \mathcal{E} is the *electric field*, \mathcal{B} is the *magnetic induction*, \mathcal{H} is the *magnetic field*, \mathcal{D} is the *electric displacement*, \mathcal{J} is the *electric current density*, \mathbf{x} contains the space variables and t is the time variable. The terminology is taken from [6] and will be used throughout this thesis. However, \mathcal{B} is also called the *magnetic flux density* and \mathcal{D} is also called the *electric flux density*. Equation (2.1a), also known as *Faraday's law*, shows how a time change in the magnetic induction gives a contribution to the electric field. Similarly, equation (2.1b) also known as *Ampère's law* shows how an electric current and a time change in the electric displacement give a contribution to the magnetic field. Maxwell's equations are supplemented with the law of conservation of charge or *continuity equation*

$$\nabla \cdot \mathcal{J}(\mathbf{x}, t) + \frac{\partial}{\partial t} \varrho(\mathbf{x}, t) = 0, \quad (2.2)$$

where ϱ is the *electric charge density*. Equation (2.2) shows that a time change in the electric charge density contributes to an electric current density. These three equations are supplemented with two scalar relations

$$\nabla \cdot \mathcal{B}(\mathbf{x}, t) = 0, \quad (2.3a)$$

$$\nabla \cdot \mathcal{D}(\mathbf{x}, t) = \varrho(\mathbf{x}, t). \quad (2.3b)$$

Equation (2.3a) also known as *Gauss' law for the magnetic field* implies that no free magnetic poles exist. Equation (2.3b) also known as *Gauss' law for the electric field* shows the relation between the electric displacement and charge density.

The electromagnetic quantities are functions of the spatial variable \mathbf{x} and time variable t . In the application described in Section 7.2 the light source is typically a laser or a white light source in combination with a colour filter. Both sources can be adequately modelled by a *time-harmonic field* or *monochromatic field* with a fixed *angular frequency* ω

$$\mathcal{F}(\mathbf{x}, t) = \text{Re} \left[\mathbf{F}(\mathbf{x}) e^{j\omega t} \right], \quad (2.4)$$

where \mathcal{F} is any of the previously introduced electromagnetic quantities and Re means the real part. This thesis will only consider time-harmonic fields and under this assumption Maxwell's equations become

$$\nabla \times \mathbf{E} = -j\omega \mathbf{B}, \quad (2.5a)$$

$$\nabla \times \mathbf{H} = j\omega \mathbf{D} + \mathbf{J}, \quad (2.5b)$$

supplemented with the continuity equation

$$\nabla \cdot \mathbf{J} + j\omega \rho = 0, \quad (2.6)$$

and the two scalar relations

$$\nabla \cdot \mathbf{B} = 0, \quad (2.7a)$$

$$\nabla \cdot \mathbf{D} = \rho. \quad (2.7b)$$

For time-harmonic fields the two scalar relations (2.7) are just auxiliary relations and can be derived from Maxwell's equations and the continuity equation. Taking the divergence of (2.5a) and using the vector identity $\nabla \cdot (\nabla \times \mathbf{F}) = 0$ gives (2.7a). Similarly taking the divergence of (2.5b), using the same vector identity and combining with the continuity equation gives (2.7b). The electromagnetic quantities are summarised in Table 2.1 with their units.

2.2.2 Constitutive relations and boundary conditions

Maxwell's equations and the continuity equation do not form a complete set of equations for the electromagnetic quantities. We need a set of *constitutive relations* that complement equations (2.5) and (2.6). The constitutive relations incorporate the influence of matter on the electromagnetic fields and typically have the dependencies

$$\mathbf{D} = \mathbf{D}(\mathbf{E}), \quad \mathbf{B} = \mathbf{B}(\mathbf{H}) \quad \text{and} \quad \mathbf{J} = \mathbf{J}(\mathbf{E}). \quad (2.8)$$

Symbol	Name	SI units	
E	electric field	Volt per meter	$\frac{V}{m} = \frac{kg \cdot m}{A \cdot s^3}$
H	magnetic field	Ampere per meter	$\frac{A}{m}$
D	electric displacement	Coulomb per square meter	$\frac{C}{m^2} = \frac{A \cdot s}{m^2}$
B	magnetic induction	Tesla	$T = \frac{kg}{A \cdot s^2}$
J	electric current density	Ampere per square meter	$\frac{A}{m^2}$
ρ	electric charge density	Coulomb per cubic meter	$\frac{C}{m^3} = \frac{A \cdot s}{m^3}$

Table 2.1: The electromagnetic quantities from Maxwell's equations

We will restrict ourselves to *linearly reacting* media and *time-invariant* reactions. This means that (2.8) is linear in the electric and magnetic field and does not depend on time. Moreover all media are assumed *dispersion free* and *isotropic*. This means that the material responds instantaneously and has physical properties that at each point are independent of direction. Finally only *source free* media are considered so no external sources are present. Under these assumptions the constitutive relations have the well-known form

$$\mathbf{D} = \varepsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H} \quad \text{and} \quad \mathbf{J} = \sigma \mathbf{E}. \quad (2.9)$$

Here ε is called the *permittivity* (or *dielectric constant*), μ is known as the *permeability* and σ is called the *conductivity*. Materials for which σ is negligibly small (e.g. air, glass) are called *insulators* or *dielectrics*. Their electric and magnetic properties are then completely determined by ε and μ . In this thesis we will only consider non-magnetic media resulting in a permeability for all media which is equal to the *free space permeability* μ_0 . Contrary to the permeability, the permittivity can change from material to material and is different from the *free space permittivity* ε_0 . Materials for which $\sigma \neq 0$ or is not negligibly small (e.g. metals) are called *conductors*. We will encounter some examples later on for which the conductivity is indeed non-zero. However, we will exclude the *perfectly electric conductor* from our analysis for which the conductivity goes to infinity. The material parameters are summarised in Table 2.2 with their units. Substituting the constitutive relations (2.9) into Maxwell's equations (2.5) results in

$$\nabla \times \mathbf{E} = -j\omega\mu_0\mathbf{H}, \quad (2.10a)$$

$$\nabla \times \mathbf{H} = j\omega\varepsilon_0\varepsilon^r\mathbf{E}, \quad (2.10b)$$

where we have introduced the complex-valued relative permittivity

$$\varepsilon^r = \frac{\varepsilon}{\varepsilon_0} - j\frac{\sigma}{\omega\varepsilon_0}. \quad (2.11)$$

Symbol	Name	SI units
ε	permittivity	Farad per meter $\frac{\text{F}}{\text{m}} = \frac{\text{A}^2 \cdot \text{s}^4}{\text{kg} \cdot \text{m}^3}$
μ	permeability	Henry per meter $\frac{\text{H}}{\text{m}} = \frac{\text{kg} \cdot \text{m}}{\text{A}^2 \cdot \text{s}^2}$
σ	conductivity	Siemens per meter $\frac{\text{S}}{\text{m}} = \frac{\text{kg} \cdot \text{m}^3}{\text{A}^2 \cdot \text{s}^3}$

Table 2.2: The material parameters from the constitutive equations

The *refraction index* n is related to the relative permittivity through the relation

$$n = \sqrt{\varepsilon^r} = n' - jn'' \quad (2.12)$$

where n' and n'' are both non-negative and real-valued. In optics it is usually this refraction index and not the permittivity that is used to characterise a material. If the relative permittivity is constant and does not depend on the spatial coordinates, the *Helmholtz equation* can be derived for the electromagnetic fields

$$\nabla^2 \mathbf{E} + k_0^2 \varepsilon^r \mathbf{E} = \mathbf{0}, \quad (2.13a)$$

$$\nabla^2 \mathbf{H} + k_0^2 \varepsilon^r \mathbf{H} = \mathbf{0}, \quad (2.13b)$$

where we have used the vector identity $\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}$ and where $k_0 = \omega \sqrt{\varepsilon_0 \mu_0}$ is the vacuum *wavenumber*.

While there are many functions that satisfy the differential equations (2.10), only one of them is the real solution to the problem. To determine this solution, one must know the *boundary conditions* associated with the domain. We will introduce two types of boundary conditions that describe the behaviour of the electromagnetic quantities in a domain. *Interface boundary conditions* describe the behaviour at the interface between two media with different material properties, *radiation boundary conditions* describe the behaviour at infinity. The interface boundary conditions can formally be derived from the integral representations of Maxwell's equations, see for example [20]. For a smooth interface between two media, say a surface S , the time-harmonic electromagnetic quantities must satisfy the following four equations

$$\llbracket \mathbf{n} \times \mathbf{E} \rrbracket_S = \mathbf{0}, \quad (2.14a)$$

$$\llbracket \mathbf{n} \times \mathbf{H} \rrbracket_S = \mathbf{J}_S, \quad (2.14b)$$

$$\llbracket \mathbf{n} \cdot \mathbf{B} \rrbracket_S = 0, \quad (2.14c)$$

$$\llbracket \mathbf{n} \cdot \mathbf{D} \rrbracket_S = \rho_S, \quad (2.14d)$$

where \mathbf{n} is the unit vector normal to the interface pointing from medium 1 into medium 2, \mathbf{J}_S is the *electric surface current density* and ρ_S is the *electric surface charge density*. Addi-

tionally we have the law of conservation of surface charge derived in [3]

$$\nabla_S \cdot \mathbf{J}_S + \llbracket \mathbf{n} \cdot \mathbf{J} \rrbracket_S = -j\omega\rho_S, \quad (2.15)$$

where $\nabla_S \cdot$ is the surface divergence. Because we excluded the perfectly electric conductor from our analysis, actually no surface currents can exist. Substituting the constitutive relations (2.9) into the interface boundary conditions and combining (2.14d) with (2.15) results in

$$\llbracket \mathbf{n} \times \mathbf{E} \rrbracket_S = \mathbf{0}, \quad (2.16a)$$

$$\llbracket \mathbf{n} \times \mathbf{H} \rrbracket_S = \mathbf{0}, \quad (2.16b)$$

$$\llbracket \mathbf{n} \cdot \mathbf{H} \rrbracket_S = 0, \quad (2.16c)$$

$$\llbracket \mathbf{n} \cdot (\epsilon^r \mathbf{E}) \rrbracket_S = 0. \quad (2.16d)$$

Among these four equations only two are independent [19,21]: (2.16c) and (2.16d) can be derived from (2.16a) and (2.16b) respectively. Therefore the two independent interface boundary conditions we will use are given by the continuity of the tangential fields

$$\llbracket \mathbf{n} \times \mathbf{E} \rrbracket_S = \mathbf{0}, \quad (2.17a)$$

$$\llbracket \mathbf{n} \times \mathbf{H} \rrbracket_S = \mathbf{0}. \quad (2.17b)$$

The radiation boundary condition describes the behaviour of the electromagnetic quantities at infinity. The standard *Sommerfeld radiation condition* typically looks like

$$\lim_{r \rightarrow \infty} r^{\frac{n-1}{2}} \left(\frac{\partial f}{\partial r} + jk_0 f \right) = 0. \quad (2.18)$$

Here f is a scalar field, r is the distance to the origin and n is the spatial dimension of the problem. For electromagnetic problems the radiation boundary condition is given by [13]

$$\lim_{r \rightarrow \infty} r \left(\sqrt{\mu_0/\epsilon_0} \mathbf{r} \times \mathbf{H} + \mathbf{E} \right) = \mathbf{0}, \quad (2.19a)$$

$$\lim_{r \rightarrow \infty} r \left(\mathbf{r} \times \mathbf{E} - \sqrt{\mu_0/\epsilon_0} \mathbf{H} \right) = \mathbf{0}, \quad (2.19b)$$

where \mathbf{r} is the unit vector in the radial direction. The physical interpretation of this boundary condition states that the scattered quantities are not incoming at infinity. Moreover the radial component of the scattered quantities show a decay faster than r^{-1} for an increasing distance to the origin. In this thesis we will look at unbounded (periodic) scatterers where certain components of the scattered quantities do not show any decay. Thus the electromagnetic counterpart of the Sommerfeld radiation condition does not apply here and is replaced by the *Rayleigh radiation condition*. The Rayleigh radiation condition states that the reflected field has, at some height above the surface of the scatterer, an expansion in plane waves propagating upwards and evanescent waves decaying exponentially with distance from the surface. A similar statement can be made for the transmitted field below the surface of the scatterer.

2.3 Basic assumptions and reduced model

This thesis will focus on unbounded scatterers; more specifically we will only look at infinitely periodic gratings. Although in real life gratings are never perfectly periodic nor are they infinite, often they can be approximated and modelled by an infinitely periodic grating. Before writing down the reduced equations and boundary conditions some notation is introduced. Figure 2.2 gives a schematic overview of an infinitely periodic

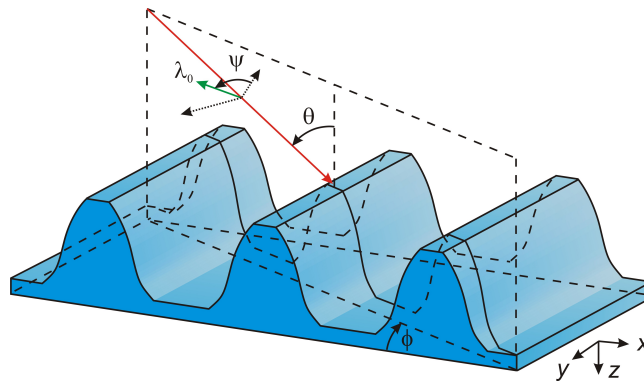


Figure 2.2: *Infinitely periodic grating with a linearly polarised incident plane wave*

grating. The periodicity of the grating is along the x -direction, the lines of the grating are along the y -direction and the z -direction is pointing downwards to complete the orthogonal coordinate system (e_x, e_y, e_z) of the 3-dimensional Euclidean space. Because the grating lines are infinitely long, the diffraction problem is invariant in the y -direction. This also means that the relative permittivity in (2.20) does not depend on the y -coordinate. The period or *pitch* of the grating is equal to Λ and the total height of the grating is equal to D . The origin of the 3-dimensional Euclidean space is chosen at

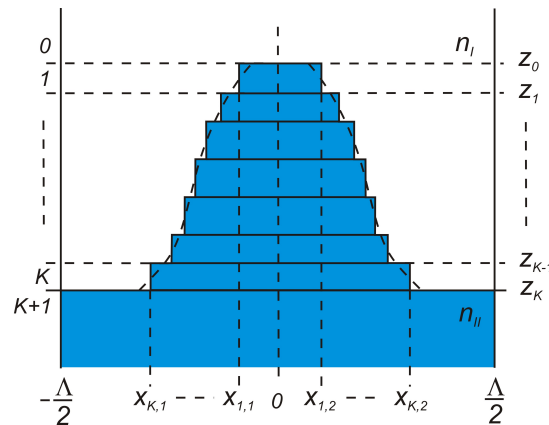


Figure 2.3: *Unit cell of an infinitely periodic grating with its layered approximation*

the top of the grating. Above the grating for $z < 0$ and below the grating for $z > D$ there are two infinite half-spaces called the *superstrate* and *substrate* respectively. For the moment assume a two medium problem where the grating consists of the same material as the substrate with refraction index n_{II} and a superstrate with refraction index n_{I} . The superstrate is assumed to be a dielectric and therefore has a purely real-valued refraction index $n_{\text{I}} = n'_{\text{I}}$ according to (2.12). On the other hand the substrate and grating can be either a dielectric or a conductor and therefore in general have a complex-valued refraction index $n_{\text{II}} = n'_{\text{II}} - jn''_{\text{II}}$.

Because of the assumption of an infinitely periodic grating we can restrict the analysis and computations to one unit cell. More importantly the grating in this unit cell is approximated by layers in which the relative permittivity no longer depends on the vertical coordinate z but only on the horizontal periodic coordinate x . In Figure 2.3 such a layered approximation of the grating is given where the layers are numbered from 0 to $K + 1$. Layer 0 corresponds to the superstrate, the layers 1 to K make up the actual grating and will be called the *grating layers* and layer $K + 1$ corresponds to the substrate. Throughout this thesis the subscript i and its values from 0 to $K + 1$ is used in variables to indicate that they belong to layer i only. The offsets in each grating layer are denoted by $X_{i,l}$ for $l = 0, \dots, L_i$ where $X_{i,0} = -\Lambda/2$ and $X_{i,L_i} = \Lambda/2$. The height of each grating layer is given by Z_i and moreover $Z_0 = 0$. The thickness of each grating layer is then $h_i = Z_i - Z_{i-1}$. For the layered grating in Figure 2.3 the relative permittivity in the grating layers is given by

$$\varepsilon_i^r(x) = \begin{cases} n_{\text{II}}^2, & X_{i,1} \leq x \leq X_{i,2}, \\ n_{\text{I}}^2, & \text{otherwise.} \end{cases} \quad (2.20)$$

A monochromatic linearly polarised *plane wave* with unit amplitude and vacuum *wavelength* $\lambda_0 = 2\pi/k_0$ is incident on the grating in layer 0 with *polar angle* θ , *azimuthal angle* ϕ and *polarisation angle* ψ . The positive direction of all three angles is depicted in Figure 2.2. The direction of the incident electric field with *wave vector* \mathbf{k} is defined by the first two angles, while the linear polarisation state or orientation of the incident electric field is defined by the third angle. The *plane of incidence* is the plane spanned by the incident wave vector \mathbf{k} and \mathbf{e}_z . In the case of normal incidence where $\theta = 0$ the plane of incidence is defined by the xz -plane. We will distinguish between three different diffraction cases

- The *planar diffraction* case corresponds to an azimuthal angle $\phi = 0$ so that the plane of incidence coincides with the xz -plane. In this case we can consider two basic linear polarisations from which all other polarisations can be derived through the superposition principle.

- *TE polarisation* or transverse electric polarisation corresponds to $\psi = \frac{\pi}{2}$ which means that the incident electric field is perpendicular to the plane of incidence and parallel to the grating lines in the y -direction.
- *TM polarisation* or transverse magnetic polarisation corresponds to $\psi = 0$ which means that the incident electric field lies in the plane of incidence. In this case the corresponding incident magnetic field is perpendicular to the plane of incidence and parallel to the grating lines in the y -direction.
- The *conical diffraction* case corresponds to an azimuthal angle $\phi \neq 0$ and can be considered the generalisation of the planar diffraction case. Contrary to the planar diffraction case we will see that now the electric and magnetic field components remain coupled through the boundary condition and cannot be separated into two basic polarisations. For this reason the planar diffraction case is still considered separately.

With this notation the incident electric field and corresponding magnetic field in layer 0 are given by

$$\mathbf{E}_0^i = e^{-jk \cdot \mathbf{x}} (\sin \psi \mathbf{s} + \cos \psi \mathbf{p}), \quad (2.21a)$$

$$\mathbf{H}_0^i = Y_1 e^{-jk \cdot \mathbf{x}} (\cos \psi \mathbf{s} - \sin \psi \mathbf{p}), \quad (2.21b)$$

$$\mathbf{k} = k_0 n_1 (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)^T, \quad (2.21c)$$

$$\mathbf{s} = (-\sin \phi, \cos \phi, 0)^T, \quad (2.21d)$$

$$\mathbf{p} = (\cos \theta \cos \phi, \cos \theta \sin \phi, -\sin \theta)^T, \quad (2.21e)$$

where $Y = \sqrt{\epsilon_0/\mu_0}$ is simply the free space *admittance* and $Y_1 = n_1 Y$ the admittance corresponding to medium 1, \mathbf{s} is an auxiliary unit vector normal to the plane of incidence and \mathbf{p} is an auxiliary unit vector in the plane of incidence so that $\mathbf{p} = \mathbf{s} \times \mathbf{k}/(k_0 n_1)$. Now that the direction of the incident field is known, we can properly define the invariance and the periodic boundary conditions of the electromagnetic quantities in a unit cell. Because Maxwell's equations are linear and the incident field is a plane wave it can be shown that the electromagnetic quantities are actually *pseudo-periodic* [31]

$$\mathbf{F}(\mathbf{x}) = e^{-jk \cdot \mathbf{e}_y y} \tilde{\mathbf{F}}(x, z) = e^{-jk_y y} \tilde{\mathbf{F}}(x, z), \quad (2.22a)$$

$$\mathbf{F}(\mathbf{x} + \Lambda \mathbf{e}_x) = e^{-jk \cdot \mathbf{e}_x \Lambda} \mathbf{F}(\mathbf{x}) = e^{-jk_x \Lambda} \mathbf{F}(\mathbf{x}). \quad (2.22b)$$

In the subsequent section we will drop the tilde in (2.22a) when substituting the electric or magnetic fields. From the context it will become clear whether we need to add the y -dependence explicitly. All these assumptions also greatly simplify Maxwell's equations (2.10) and the interface boundary conditions (2.16) for the three different diffraction cases.

2.3.1 Planar diffraction TE polarisation

For planar diffraction with TE polarisation ($\phi = 0, \psi = \frac{\pi}{2}$) the incident electric field in (2.21a) simplifies to

$$E_0^i = E_{0,y}^i(x, z) \mathbf{e}_y = e^{-jk_0 n_1 (x \sin \theta + z \cos \theta)} \mathbf{e}_y. \quad (2.23)$$

The incident electric field only has a non-zero y -component and it only depends on the x - and z -coordinate. Because of this the electric field in all layers also have this property, in accordance with (2.22a) where $k_y = 0$. Maxwell's equations (2.10) with the layered approximation for $i = 0, \dots, K + 1$ simplify to

$$\frac{\partial}{\partial z} E_{i,y} = j\omega\mu_0 H_{i,x}, \quad (2.24a)$$

$$\frac{\partial}{\partial x} E_{i,y} = -j\omega\mu_0 H_{i,z}, \quad (2.24b)$$

$$\frac{\partial}{\partial z} H_{i,x} = j\omega\varepsilon_0 \varepsilon_i^r E_{i,y} + \frac{\partial}{\partial x} H_{i,z}. \quad (2.24c)$$

The equations above can be rewritten into one second-order differential equation for the electric field component $E_{i,y}$. Eliminating the magnetic field components by substituting (2.24a) and (2.24b) into (2.24c) and dividing by k_0^2 gives

$$\frac{1}{k_0^2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + k_0^2 \varepsilon_i^r \right) E_{i,y} = 0. \quad (2.25)$$

For the layers 0 and $K + 1$, where the relative permittivity is constant, (2.25) reduces to the standard Helmholtz equation as was already derived in (2.13a). For the interface boundary condition (2.17) between two adjacent layers for $i = 0, \dots, K$ where $\mathbf{n} = \mathbf{e}_z$ and $z = Z_i$ we get

$$E_{i,y} = E_{i+1,y}, \quad (2.26a)$$

$$H_{i,x} = H_{i+1,x}. \quad (2.26b)$$

Also here the magnetic field component can be eliminated by using (2.24a) which gives

$$E_{i,y} = E_{i+1,y}, \quad (2.27a)$$

$$\frac{1}{k_0} \frac{\partial}{\partial z} E_{i,y} = \frac{1}{k_0} \frac{\partial}{\partial z} E_{i+1,y}. \quad (2.27b)$$

From this interface boundary condition we can see that the electric field is continuous across a layer interface and that also its partial derivative with respect to z is continuous. Of course the electromagnetic fields also have to satisfy the pseudo-periodic boundary condition (2.22b) in all layers and the Rayleigh radiation condition in the superstrate and substrate.

2.3.2 Planar diffraction TM polarisation

For planar diffraction with TM polarisation ($\phi = 0, \psi = 0$) the incident magnetic field in (2.21b) simplifies to

$$\mathbf{H}_0^i = H_{0,y}^i(x, z) \mathbf{e}_y = Y_1 e^{-jk_0 n_1 (x \sin \theta + z \cos \theta)} \mathbf{e}_y. \quad (2.28)$$

So now it is the incident magnetic field that only has a non-zero y -component and only depends on the x - and z -coordinate. Because of this the magnetic field in all layers also have this property, in accordance with (2.22a) where $k_y = 0$. Maxwell's equations (2.10) with the layered approximation for $i = 0, \dots, K + 1$ simplify to

$$\frac{\partial}{\partial z} E_{i,x} = -j\omega\mu_0 H_{i,y} + \frac{\partial}{\partial x} E_{i,z}, \quad (2.29a)$$

$$\frac{\partial}{\partial z} H_{i,y} = -j\omega\varepsilon_0 \varepsilon_i^r E_{i,x}, \quad (2.29b)$$

$$\frac{\partial}{\partial x} H_{i,y} = j\omega\varepsilon_0 \varepsilon_i^r E_{i,z}. \quad (2.29c)$$

The equations above can be rewritten into one second-order differential equation for the magnetic field component $H_{i,y}$. Eliminating the electric field components by substituting (2.29b) and (2.29c) into (2.29a) and dividing by k_0^2 results in

$$\frac{1}{k_0^2} \left(\varepsilon_i^r \frac{\partial}{\partial x} \frac{1}{\varepsilon_i^r} \frac{\partial}{\partial x} + \frac{\partial^2}{\partial z^2} + k_0^2 \varepsilon_i^r \right) H_{i,y} = 0. \quad (2.30)$$

Also now for the layers 0 and $K + 1$, where the relative permittivity is constant, (2.30) reduces to the standard Helmholtz equation as was already derived in (2.13b). For the interface boundary condition (2.17) between two adjacent layers for $i = 0, \dots, K$ where $\mathbf{n} = \mathbf{e}_z$ and $z = Z_i$ we get

$$H_{i,y} = H_{i+1,y}, \quad (2.31a)$$

$$E_{i,x} = E_{i+1,x}. \quad (2.31b)$$

Again we can eliminate the electric field component by using (2.29b)

$$H_{i,y} = H_{i+1,y}, \quad (2.32a)$$

$$\frac{1}{k_0} \frac{1}{\varepsilon_i^r} \frac{\partial}{\partial z} H_{i,y} = \frac{1}{k_0} \frac{1}{\varepsilon_{i+1}^r} \frac{\partial}{\partial z} H_{i+1,y}. \quad (2.32b)$$

From this interface boundary condition we can see that the magnetic field is continuous across a layer interface but contrary to TE polarisation the partial derivative with respect to z of the magnetic field is not. Also here the electromagnetic fields still need to satisfy the pseudo-periodic boundary condition (2.22b) in all layers and the Rayleigh radiation condition in the superstrate and substrate.

2.3.3 Conical diffraction

For conical diffraction ($\phi \neq 0$) the incident field in (2.21a) does not simplify any further. In general the electric and magnetic field both have three non-zero components that depend on all three spatial coordinates. However, because the grating is still invariant in the y -direction and is approximated by layers, Maxwell's equations (2.10) for $i = 0, \dots, K + 1$ simplify to

$$-jk_y E_{i,z} - \frac{\partial}{\partial z} E_{i,y} = -j\omega\mu_0 H_{i,x}, \quad (2.33a)$$

$$\frac{\partial}{\partial z} E_{i,x} - \frac{\partial}{\partial x} E_{i,z} = -j\omega\mu_0 H_{i,y}, \quad (2.33b)$$

$$\frac{\partial}{\partial x} E_{i,y} + jk_y E_{i,x} = -j\omega\mu_0 H_{i,z}, \quad (2.33c)$$

$$-jk_y H_{i,z} - \frac{\partial}{\partial z} H_{i,y} = j\omega\varepsilon_0 \varepsilon_i^r E_{i,x}, \quad (2.33d)$$

$$\frac{\partial}{\partial z} H_{i,x} - \frac{\partial}{\partial x} H_{i,z} = j\omega\varepsilon_0 \varepsilon_i^r E_{i,y}, \quad (2.33e)$$

$$\frac{\partial}{\partial x} H_{i,y} + jk_y H_{i,x} = j\omega\varepsilon_0 \varepsilon_i^r E_{i,z}, \quad (2.33f)$$

where we have dropped the term $e^{-jk_y y}$ from all equations coming from (2.22a) now that $k_y \neq 0$. The z -component of the electric and magnetic field can be eliminated with help of equations (2.33c) and (2.33f)

$$\frac{\partial}{\partial z} E_{i,y} = j\omega\mu_0 H_{i,x} + \frac{1}{j\omega\varepsilon_0} \frac{1}{\varepsilon_i^r} \left(-jk_y \frac{\partial}{\partial x} H_{i,y} + k_y^2 H_{i,x} \right), \quad (2.34a)$$

$$\frac{\partial}{\partial z} E_{i,x} = -j\omega\mu_0 H_{i,y} + \frac{1}{j\omega\varepsilon_0} \frac{\partial}{\partial x} \left(\frac{1}{\varepsilon_i^r} \frac{\partial}{\partial x} H_{i,y} + jk_y \frac{1}{\varepsilon_i^r} H_{i,x} \right), \quad (2.34b)$$

$$\frac{\partial}{\partial z} H_{i,y} = -j\omega\varepsilon_0 \varepsilon_i^r E_{i,x} + \frac{1}{j\omega\mu_0} \left(-k_y^2 E_{i,x} + jk_y \frac{\partial}{\partial x} E_{i,y} \right), \quad (2.34c)$$

$$\frac{\partial}{\partial z} H_{i,x} = j\omega\varepsilon_0 \varepsilon_i^r E_{i,y} + \frac{1}{j\omega\mu_0} \left(-jk_y \frac{\partial}{\partial x} E_{i,x} - \frac{\partial^2}{\partial x^2} E_{i,y} \right). \quad (2.34d)$$

After some straightforward algebra it is possible to derive two uncoupled second-order differential equation for the electric and magnetic field components $E_{i,x}$ and $H_{i,x}$

$$\frac{1}{k_0^2} \left(\frac{\partial}{\partial x} \frac{1}{\varepsilon_i^r} \frac{\partial}{\partial x} \varepsilon_i^r + \frac{\partial^2}{\partial z^2} + k_0^2 \varepsilon_i^r - k_y^2 \right) E_{i,x} = 0, \quad (2.35a)$$

$$\frac{1}{k_0^2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + k_0^2 \varepsilon_i^r - k_y^2 \right) H_{i,x} = 0. \quad (2.35b)$$

Similar to the planar diffraction case for the layers 0 and $K + 1$, where the relative permittivity is constant, (2.35) reduces to the standard Helmholtz equation as was already derived in (2.13). For the interface boundary conditions (2.17) between two adjacent layers for $i = 0, \dots, K$ where $\mathbf{n} = \mathbf{e}_z$ and $z = Z_i$ we get

$$E_{i,x} = E_{i+1,x}, \quad H_{i,x} = H_{i+1,x}, \quad (2.36a)$$

$$E_{i,y} = E_{i+1,y}, \quad H_{i,y} = H_{i+1,y}. \quad (2.36b)$$

Contrary to the planar diffraction case it is not possible to eliminate the y -components of the electric and magnetic field and subsequently simplify the interface boundary conditions. Therefore after solving for the x -components of the electric and magnetic field through (2.35), one still needs to compute the y -components of the electric and magnetic field with the help of (2.34b) and (2.34d) when applying the interface boundary conditions (2.36b). The pseudo-periodic boundary condition (2.22b) still applies to the electromagnetic fields in all layers together with the Rayleigh radiation condition in the superstrate and substrate.

2.4 Derivation Rayleigh modes outside grating layers

In the superstrate and substrate the relative permittivity is constant and thus Maxwell's equations reduce to the standard Helmholtz equation. The eigenfunctions of this equation with pseudo-periodic boundary conditions can be computed exactly and turn out to be plane waves. The fields in the superstrate and substrate are linear combinations of these eigenfunctions while taking care of the Rayleigh radiation condition. This expansion is also known as the *Rayleigh expansion*. We provide a derivation of the Rayleigh expansion in the superstrate and substrate for the general conical diffraction case. We show its various straightforward steps which reoccur in the next chapter when we discuss one of the discretisations.

Using the invariance in the y -direction the Helmholtz equation (2.13a) for the electric field for $i = 0, K + 1$ after dividing by k_0^2 reduces to

$$\frac{1}{k_0^2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + k_0^2 \varepsilon_i^r - k_y^2 \right) E_i(x, z) e^{-jk_y y} = 0, \quad (2.37)$$

where ε_i^r is now constant and equal to either n_I^2 in the superstrate or n_{II}^2 in the substrate. Recall that for planar diffraction with TE polarisation only the y -component of the electric field is non-zero and $k_y = 0$. For planar diffraction with TM polarisation the x - and z -components of the electric field are non-zero and again $k_y = 0$. For the conical diffraction case all components of the electric field are non-zero in general and $k_y \neq 0$. Focussing on one component $\alpha = x, y, z$ in the superstrate, this partial differential equa-

tion is easily solved by the method of *separation of variables*

$$E_{0,\alpha}(x, z) = u_\alpha(x)v_\alpha(z). \quad (2.38)$$

Substituting in (2.37), rearranging the terms and adding the proper boundary conditions gives

$$\begin{cases} \frac{1}{k_0^2}u_\alpha'' = -\mu^2 u_\alpha, & |x| < \frac{\Lambda}{2}, \\ u_\alpha(\frac{\Lambda}{2}) = e^{-jk_x\Lambda}u_\alpha(-\frac{\Lambda}{2}), \\ u_\alpha'(\frac{\Lambda}{2}) = e^{-jk_x\Lambda}u_\alpha'(-\frac{\Lambda}{2}), \end{cases} \quad (2.39a)$$

and

$$\begin{cases} \frac{1}{k_0^2}v_\alpha'' + (n_1^2 - \frac{k_y^2}{k_0^2})v_\alpha = \mu^2 v_\alpha, & z < 0, \\ \text{Rayleigh radiation condition,} \\ \text{Interface boundary condition,} \end{cases} \quad (2.39b)$$

where μ^2 is the separation constant or eigenvalue of the problem. Because the pseudo-periodic boundary conditions are identical in both the superstrate, substrate and for all values of α only one constant μ is needed. It is easy to see that the differential operator corresponding to the problem for u_α is self-adjoint. This means that the eigenvalues are real-valued and the eigenfunctions corresponding to distinct eigenvalues are orthogonal. Naturally we first need to define an inner product for the previous statement to make sense. Therefore we define the *weighed inner product* between two functions g and h as

$$\langle g, h \rangle_w := \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} w(x)g(x)\overline{h(x)} dx, \quad (2.40)$$

where the bar is used to denote complex conjugation. The eigenfunctions are then orthogonal with respect to the inner product $\langle g, h \rangle := \langle g, h \rangle_1$. The solution for u_α is now readily obtained and is given by

$$u_\alpha = A_\alpha \cos(k_0\mu x) + B_\alpha \sin(k_0\mu x). \quad (2.41)$$

In order to get a relation for the unknown expansion coefficients A_α and B_α we apply the pseudo-periodic boundary conditions. Thus,

$$A_\alpha \cos(\frac{k_0\mu\Lambda}{2}) + B_\alpha \sin(\frac{k_0\mu\Lambda}{2}) = e^{-jk_x\Lambda} (A_\alpha \cos(-\frac{k_0\mu\Lambda}{2}) + B_\alpha \sin(-\frac{k_0\mu\Lambda}{2})), \quad (2.42a)$$

$$\begin{aligned} -k_0\mu (A_\alpha \sin(\frac{k_0\mu\Lambda}{2}) - B_\alpha \cos(\frac{k_0\mu\Lambda}{2})) = \\ -k_0\mu e^{-jk_x\Lambda} (A_\alpha \sin(-\frac{k_0\mu\Lambda}{2}) - B_\alpha \cos(-\frac{k_0\mu\Lambda}{2})). \end{aligned} \quad (2.42b)$$

For this homogeneous set of linear equations to have a non-trivial solution we require

that

$$\begin{aligned} k_0\mu \left((1 - e^{-jk_x\Lambda})^2 \cos^2\left(\frac{k_0\mu\Lambda}{2}\right) + (1 + e^{-jk_x\Lambda})^2 \sin^2\left(\frac{k_0\mu\Lambda}{2}\right) \right) = \\ k_0\mu \left(1 - 2e^{-jk_x\Lambda} \cos(k_0\mu\Lambda) + e^{-2jk_x\Lambda} \right) = 0. \end{aligned} \quad (2.43)$$

The term within the outer pair of brackets simplifies to the relation $\cos(k_0\mu\Lambda) = \cos(k_x\Lambda)$ which has a countably infinite set of solutions for $m \in \mathbb{Z}$

$$k_{xm} := k_0\mu_m = k_x - \frac{2\pi m}{\Lambda}. \quad (2.44)$$

We can neglect the solution corresponding to $\mu = 0$ since this gives either the trivial or the constant solution in the case of real periodic boundary conditions. The former is not interesting at all while the latter is already captured by (2.44). Now we still need to determine the expansion coefficients A_α and B_α . One relation directly follows from the linear dependence of (2.42) while the other relation follows from normalising the eigenfunctions

$$B_{\alpha m} = -\frac{(1 - e^{-jk_x\Lambda}) \cos(k_{xm} \frac{\Lambda}{2})}{(1 + e^{-jk_x\Lambda}) \sin(k_{xm} \frac{\Lambda}{2})} A_{\alpha m} = -jA_{\alpha m}, \quad (2.45a)$$

$$1 = \langle u_{\alpha m}, u_{\alpha m} \rangle. \quad (2.45b)$$

Combining these results gives a complete set of orthonormal eigenfunctions

$$u_{\alpha m} = e^{-jk_{xm}x}. \quad (2.46)$$

This set of eigenfunctions is sometimes called the *pseudo-periodic Fourier series* which is easily explained by looking at (2.44). There we see that because k_x is just a constant it introduces an extra phase-factor in front of the standard Fourier series. Finally, looking at the equation for v_α we have infinitely many solutions of the form

$$v_{\alpha m} = R_{\alpha m} e^{jk_{1,zm}z} + T_{\alpha m} e^{-jk_{1,zm}z}, \quad (2.47a)$$

$$k_{1,zm} = \sqrt{k_0^2 n_1^2 - k_{xm}^2 - k_y^2}, \quad \text{Im}[k_{1,zm}] \leq 0. \quad (2.47b)$$

With the Rayleigh radiation condition we can now determine half of the unknown expansion coefficients. This can be seen when writing down the total solution of the electric field including the y -dependency of this field

$$\begin{aligned} E_{0,\alpha} e^{-jk_y y} &= \sum_{m=-\infty}^{\infty} u_{\alpha m} v_{\alpha m} e^{-jk_y y} \\ &= \sum_{m=-\infty}^{\infty} R_{\alpha m} e^{-j(k_{xm}x + k_y y - k_{1,zm}z)} + \sum_{m=-\infty}^{\infty} T_{\alpha m} e^{-j(k_{xm}x + k_y y + k_{1,zm}z)}. \end{aligned} \quad (2.48)$$

From this expression we see that the solution consists of two parts: plane waves moving

towards the grating with expansion coefficients $T_{\alpha m}$ and plane waves moving away from the grating with expansion coefficients $R_{\alpha m}$. The Rayleigh radiation condition now requires the former to match with the known incident field of (2.21). This means that $T_{\alpha 0} = \sin \psi s_\alpha + \cos \psi p_\alpha$ and $T_{\alpha m} = 0$ for $m \neq 0$. The second part describes the scattered field in the superstrate and these plane waves indeed satisfy the Rayleigh radiation condition. This can be seen from the fact that when $k_{I,zm}$ is real we have a *propagating wave* and when $k_{I,zm}$ is imaginary we have an *evanescent wave* decaying exponentially with distance from the grating surface. The m^{th} plane wave in the Rayleigh expansion is also called the m^{th} *diffraction order* which can be either propagating or evanescent. The interface boundary condition will be used later to determine the unknown expansion coefficients $R_{\alpha m}$ for which we also need the solution in the other layers. Analogously the Rayleigh expansion in the substrate can be derived where the only difference is the relative permittivity and the absence of an incident field. Note that because of a different relative permittivity we define $k_{II,zm}$ like in (2.47b) but with n_I replaced by n_{II} .

Summarising, the Rayleigh expansions for the electric field in the superstrate and substrate are

$$E_0 e^{-jk_y y} = E_0^r + E_0^i = \sum_{m=-\infty}^{\infty} R_m^e e^{-jk_m^r \cdot x} + E_0^i, \quad (2.49a)$$

$$E_{K+1} e^{-jk_y y} = E_{K+1}^t = \sum_{m=-\infty}^{\infty} T_m^e e^{-jk_m^t \cdot (x - Z_K e_z)}, \quad (2.49b)$$

$$\mathbf{k}_m^r = (k_{xm}, k_y, -k_{I,zm})^T, \quad (2.49c)$$

$$\mathbf{k}_m^t = (k_{xm}, k_y, k_{II,zm})^T. \quad (2.49d)$$

Note that in the substrate we have normalised the solution in the vertical direction by subtracting the total height of the grating. This is done so that the expansion coefficients of the evanescent transmitted waves remain of the same order of magnitude as the propagating transmitted waves. Because the magnetic field satisfies exactly the same Helmholtz equation in the superstrate and substrate, these Rayleigh expansions are given by

$$H_0 e^{-jk_y y} = H_0^r + H_0^i = Y_I \sum_{m=-\infty}^{\infty} R_m^h e^{-jk_m^r \cdot x} + H_0^i, \quad (2.50a)$$

$$H_{K+1} e^{-jk_y y} = H_{K+1}^t = Y_{II} \sum_{m=-\infty}^{\infty} T_m^h e^{-jk_m^t \cdot (x - Z_K e_z)}. \quad (2.50b)$$

Here we have scaled the reflected magnetic field like the incident magnetic field with the admittance of medium 1 so that again all expansion coefficients are of the same order of magnitude. The transmitted magnetic field is scaled with the admittance of medium 2, i.e. $Y_{II} = n_{II} \sqrt{\epsilon_0 / \mu_0}$. Contrary to what we stated about the electric field, now for planar diffraction with TE polarisation the x - and z -components of the magnetic field are non-zero. Similarly for planar diffraction with TM polarisation only the y -component

of the magnetic field is non-zero. For the conical diffraction case all components of the magnetic field are non-zero in general. Finally note that an extra superscript is used in the expansion coefficients to distinguish between the electric and magnetic field. From Maxwell's equations it can be seen that the following relations hold between these expansion coefficients

$$\mathbf{k}_m^r \times \mathbf{R}_m^e = k_0 n_I \mathbf{R}_m^h, \quad \mathbf{k}_m^r \times \mathbf{R}_m^h = -k_0 n_I \mathbf{R}_m^e, \quad (2.51a)$$

$$\mathbf{k}_m^t \times \mathbf{T}_m^e = k_0 n_{II} \mathbf{T}_m^h, \quad \mathbf{k}_m^t \times \mathbf{T}_m^h = -k_0 n_{II} \mathbf{T}_m^e. \quad (2.51b)$$

Although for the general conical diffraction case the electric and magnetic field components remain coupled through the boundary condition it is still common to split up the field into two basic parts. The *s-polarised* part represents that part of the electric field normal to the diffraction plane and the *p-polarised* part represents that part of the electric field in the diffraction plane. Here the *diffraction plane* is defined for each diffraction order separately and is given by the *xz*-plane rotated by an angle ϕ_m about the *z*-axis so that

$$\phi_m = \arg(k_{xm} + jk_y). \quad (2.52)$$

For the sake of uniqueness we shall define $\arg(0) = 0$ so that we can also speak about a diffraction plane when a diffraction order moves in the positive or negative *z*-direction (for example in the case of normal incidence we have now also defined a diffraction plane for the 0th diffraction order). Now the *s*-polarised and *p*-polarised parts are given by

$$R_{sm} = \mathbf{R}_m^e \cdot \mathbf{s}_m = -\mathbf{R}_m^h \cdot \mathbf{p}_m^r, \quad R_{pm} = \mathbf{R}_m^e \cdot \mathbf{p}_m^r = \mathbf{R}_m^h \cdot \mathbf{s}_m, \quad (2.53a)$$

$$T_{sm} = \mathbf{T}_m^e \cdot \mathbf{s}_m = -\mathbf{T}_m^h \cdot \mathbf{p}_m^t, \quad T_{pm} = \mathbf{T}_m^e \cdot \mathbf{p}_m^t = \mathbf{T}_m^h \cdot \mathbf{s}_m. \quad (2.53b)$$

Here $\mathbf{s}_m = (-\sin \phi_m, \cos \phi_m, 0)^T$ is an auxiliary unit vector normal to the diffraction plane and is defined for each diffraction order separately. The auxiliary unit vectors in the diffraction plane are then given by $\mathbf{p}_m^r = \mathbf{s}_m \times \mathbf{k}_m^r / (k_0 n_I)$ and $\mathbf{p}_m^t = \mathbf{s}_m \times \mathbf{k}_m^t / (k_0 n_{II})$ in the superstrate and substrate respectively. Using (2.51) we also related the magnetic field to these *s*-polarised and *p*-polarised parts. When applying the interface boundary conditions later on we are typically interested in the tangential *x*- and *y*-components of the fields. This means that in (2.53) the relations including \mathbf{s}_m are very useful since they relate these tangential components of the fields to the *s*-polarised and *p*-polarised parts. Additional equations for these tangential components can be derived by looking in the direction perpendicular to \mathbf{s}_m (a derivation can be found in Appendix A.1)

$$\mathbf{R}_m^e \cdot \mathbf{s}_m^\perp = -\frac{k_{I,zm}}{k_0 n_I} R_{pm}, \quad \mathbf{R}_m^h \cdot \mathbf{s}_m^\perp = \frac{k_{I,zm}}{k_0 n_I} R_{sm}, \quad (2.54a)$$

$$\mathbf{T}_m^e \cdot \mathbf{s}_m^\perp = \frac{k_{II,zm}}{k_0 n_{II}} T_{pm}, \quad \mathbf{T}_m^h \cdot \mathbf{s}_m^\perp = -\frac{k_{II,zm}}{k_0 n_{II}} T_{sm}, \quad (2.54b)$$

where $\mathbf{s}_m^\perp = \mathbf{s}_m \times \mathbf{e}_z = (\cos \phi_m, \sin \phi_m, 0)^T$.

2.4.1 Energy flow and the Poynting vector

In our analysis of time-harmonic fields the quantity of interest related to energy is the complex-valued time-averaged *Poynting vector*

$$\mathbf{S} = \frac{1}{2} \mathbf{E} \times \overline{\mathbf{H}}. \quad (2.55)$$

The Poynting vector represents the flow of energy or power flux per unit of area. Because we introduced the phasor notation in equation (2.4), taking the real part of the Poynting vector gives the real *energy flow density* or *power flux density*. For the incident field the Poynting vector becomes

$$\begin{aligned} \mathbf{S}_0^i &= \frac{1}{2} \mathbf{E}_0^i \times \overline{\mathbf{H}}_0^i \\ &= \frac{1}{2} Y_I \frac{1}{k_0 n_1} \mathbf{k}, \end{aligned} \quad (2.56)$$

where we have used (2.21) and the fact that by definition \mathbf{s} , \mathbf{p} and \mathbf{k} are orthogonal to each other. The time-averaged energy flow through an area $A(x, y)$ parallel to the plane of the grating (the xy -plane) is then given by

$$P = \int_{A(x,y)} \mathbf{S} \cdot \mathbf{e}_z \, dA. \quad (2.57)$$

For an infinitely periodic grating that is invariant in the y -direction, this area is defined by $0 \leq x \leq \Lambda, 0 \leq y \leq 1$. For the incident field the corresponding incoming energy in the direction of \mathbf{e}_z then becomes

$$\begin{aligned} P_0^i &= \int_0^1 \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} \mathbf{S}_0^i \cdot \mathbf{e}_z \, dx dy \\ &= \frac{1}{2} Y_I \Lambda \cos \theta. \end{aligned} \quad (2.58)$$

Similarly for the scattered field we can look at the energy for each diffraction order. The Poynting vector corresponding to a reflected or transmitted diffraction order is given by

$$\begin{aligned} \mathbf{S}_{0,m}^r &= \frac{1}{2} \mathbf{E}_{0,m}^r \times \overline{\mathbf{H}}_{0,m}^r \\ &= \frac{1}{2} Y_I e^{-2\text{Im}[k_{1,zm}]z} \mathbf{R}_m^e \times \overline{\mathbf{R}}_m^h, \end{aligned} \quad (2.59a)$$

$$\begin{aligned} \mathbf{S}_{K+1,m}^t &= \frac{1}{2} \mathbf{E}_{K+1,m}^t \times \overline{\mathbf{H}}_{K+1,m}^t \\ &= \frac{1}{2} \overline{Y_{II}} e^{2\text{Im}[k_{II,zm]}(z-Z_K)} \mathbf{T}_m^e \times \overline{\mathbf{T}}_m^h. \end{aligned} \quad (2.59b)$$

From these expressions we can see that the propagating orders in the superstrate have a Poynting vector that remains constant while the evanescent orders have a Poynting vector that decays exponentially fast with increasing distance to the grating. A similar remark can be made for the substrate if it is lossless and has a purely real-valued

refraction index. If on the other hand the substrate is lossy and has a complex-valued refraction index then all diffraction orders have a Poynting vector that decays exponentially fast with increasing distance to the grating. From the Poynting vector we can compute the time-averaged energy flow for a reflected or transmitted diffraction order (a derivation can be found in Appendix A.2)

$$P_{0,m}^r = \frac{1}{2} Y_I \Lambda e^{-2\text{Im}[k_{I,zm}]z} \left(\frac{\overline{k_{I,zm}}}{k_0 n_I} |R_{sm}|^2 + \frac{k_{I,zm}}{k_0 n_I} |R_{pm}|^2 \right), \quad (2.60a)$$

$$P_{K+1,m}^t = \frac{1}{2} \overline{Y_{II}} \Lambda e^{2\text{Im}[k_{II,zm}](z-Z_K)} \left(\frac{\overline{k_{II,zm}}}{k_0 n_{II}} |T_{sm}|^2 + \frac{k_{II,zm}}{k_0 n_{II}} |T_{pm}|^2 \right). \quad (2.60b)$$

The *diffraction efficiency* is a quantity derived from the power density in the superstrate and substrate and it is this quantity that is frequently used in numerical tests. The diffraction efficiency corresponding to a diffraction order is the real part of the ratio of the reflected or transmitted power density and the incoming power density. Moreover this ratio is evaluated at the top and bottom of the grating structure. For the general conical diffraction case this means that for a lossless medium

$$\eta_m^r = \text{Re} \left[\frac{P_{0,m}^r(0)}{P_0^i} \right] = |R_{sm}|^2 \text{Re} \left[\frac{k_{I,zm}}{k_0 n_I \cos \theta} \right] + |R_{pm}|^2 \text{Re} \left[\frac{k_{I,zm}}{k_0 n_I \cos \theta} \right], \quad (2.61a)$$

$$\eta_m^t = \text{Re} \left[\frac{P_{K+1,m}^t(Z_K)}{P_0^i} \right] = |T_{sm}|^2 \text{Re} \left[\frac{k_{II,zm}}{k_0 n_{II} \cos \theta} \right] + |T_{pm}|^2 \text{Re} \left[\frac{k_{II,zm}}{k_0 n_{II} \cos \theta} \right]. \quad (2.61b)$$

Now that the energy and, more specifically, the diffraction efficiency are known for the general conical diffraction case, we can easily derive these quantities for the planar diffraction case. This is because for planar diffraction with TE polarisation we have the relations

$$R_{sm} = \text{sign}(k_{xm}) R_{ym}^e, \quad R_{pm} = 0, \quad (2.62a)$$

$$T_{sm} = \text{sign}(k_{xm}) T_{ym}^e, \quad T_{pm} = 0, \quad (2.62b)$$

and for TM polarisation

$$R_{sm} = 0, \quad R_{pm} = \text{sign}(k_{xm}) R_{ym}^h, \quad (2.62c)$$

$$T_{sm} = 0, \quad T_{pm} = \text{sign}(k_{xm}) T_{ym}^h. \quad (2.62d)$$

Here we used the standard sign-function with the exception that $\text{sign}(0) = 1$.

Chapter 3

Discretisation with Bloch modes

Maxwell's equations with the interface, pseudo-periodic and radiation boundary conditions for the three diffraction cases can be discretised in many ways. In this chapter we will focus on a discretisation strategy based on a mode expansion method. The mode expansion method that we discuss here is also known as the Bloch mode method. Essentially the electromagnetic fields in each layer are expanded into a set of eigenfunctions of Maxwell's equations in that layer. In section 2.4 we already derived the eigenfunctions in the superstrate and substrate where the solution was given by the Rayleigh expansion. By following the same steps in Section 3.1 the eigenfunctions of Maxwell's equations inside the grating layers are derived. These eigenfunctions are actually special linear combinations of sines and cosines and the corresponding expansion is known as the *Bloch mode expansion*. Although, in principle, these Bloch modes are known in closed form, the arguments of the sines and cosines are related to the roots of a transcendental equation. This transcendental equation and its properties will be the topic of Section 3.2. Finally the layers are connected through the interface boundary condition. In Section 3.3 we will discuss how this results in a linear system of equations where the layer specific expansions introduce a coupling matrix. Solving this system stably will be the topic of Chapter 5. We conclude this chapter with some generalisations and limitations of the Bloch mode method in Section 3.4.

3.1 Derivation Bloch modes inside grating layers

In the grating layers the relative permittivity is no longer constant but a periodic piecewise constant function of x . This means that Maxwell's equations no longer reduce to the Helmholtz equation and that the plane waves are no longer the eigenfunctions. We will derive these eigenfunctions for the structure in Figure 2.3 with a permittivity

function given by (2.20). We will restrict ourselves to a symmetric lossless dielectric grating with $X_{i,2} = -X_{i,1} =: X_i$ and real-valued refraction indices. The eigenfunctions of Maxwell's equations in the grating layers are now determined using the method from Section 2.4, that is using *separation of variables*.

3.1.1 Planar diffraction

We start by applying the method of separation of variables to the planar diffraction case. Moreover, we will solve the problem in the layers $i = 1, \dots, K$ for both polarisations simultaneously by introducing a field F_{Bi} which represents the y -component of either the electric or magnetic field for TE and TM polarisation respectively

$$F_{Bi}(x, z) = f u_{Bi}(x) v_{Bi}(z), \quad (3.1)$$

where $f = 1$ for TE polarisation and $f = -jY$ for TM polarisation. Substituting into (2.25) or (2.30), rearranging the terms and adding the proper boundary conditions gives

$$\begin{cases} \frac{1}{k_0^2} u_{Bi}'' + n_{II}^2 u_{Bi} = -\mu_{Bi}^2 u_{Bi}, & |x| < X_i, \\ \frac{1}{k_0^2} u_{Bi}'' + n_I^2 u_{Bi} = -\mu_{Bi}^2 u_{Bi}, & X_i < |x| < \frac{\Lambda}{2}, \\ u_{Bi}(\pm X_i^-) = u_{Bi}(\pm X_i^+), & u_{Bi}(\frac{\Lambda}{2}) = e^{-jk_x \Lambda} u_{Bi}(-\frac{\Lambda}{2}), \\ w_{II} u'_{Bi}(\pm X_i^\mp) = w_I u'_{Bi}(\pm X_i^\pm), & u'_{Bi}(\frac{\Lambda}{2}) = e^{-jk_x \Lambda} u'_{Bi}(-\frac{\Lambda}{2}), \end{cases} \quad (3.2a)$$

and

$$\begin{cases} \frac{1}{k_0^2} v_{Bi}'' = \mu_{Bi}^2 v_{Bi}, & Z_{i-1} < z < Z_i, \\ \text{Interface boundary conditions,} \end{cases} \quad (3.2b)$$

where μ_{Bi}^2 is the separation constant or eigenvalue of the problem and where $w_I = w_{II} = 1$ for TE polarisation and $w_I = 1/n_I^2, w_{II} = 1/n_{II}^2$ for TM polarisation. It is important to note that we have added *inner layer interface boundary conditions* at $x = \pm X_i$ for the problem in the periodic x -direction. These can be derived in a similar way as the interface boundary conditions between two adjacent layers for the problem in the vertical z -direction. Both require the continuity of the tangential electromagnetic fields. For the inner layer interface boundary conditions this means that from (2.17) and by eliminating the tangential z -component of the electromagnetic field using (2.24b) and (2.29c) we obtain

$$\llbracket F_{Bi} \rrbracket_{x=\pm X_i} = 0, \quad (3.3a)$$

$$\llbracket w_i \frac{\partial}{\partial x} F_{Bi} \rrbracket_{x=\pm X_i} = 0, \quad (3.3b)$$

where $w_i = 1$ for TE polarisation and $w_i = 1/\varepsilon_i^r$ for TM polarisation. It can be shown that the differential operator corresponding to the problem for u_{Bi} is self-adjoint [9, 10]. This means that the eigenvalues are real-valued and the eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the weighed inner product $\langle g, h \rangle_{w_i}$ as defined in (2.40). Let a_i, b_i be two linearly independent basis solutions of equation (3.2a) which satisfy the inner layer interface boundary conditions at $x = \pm X_i$ (not necessarily satisfying the pseudo-periodic boundary condition) and

$$\begin{cases} a_i(0) = 1, & b_i(0) = 0, \\ a_i'(0) = 0, & b_i'(0) = 1. \end{cases} \quad (3.4)$$

Then

$$a_i(x) = \begin{cases} \cos(\kappa_{\text{II}i}x), & |x| \leq X_i, \\ \cos(\kappa_{\text{II}i}X_i) \cos(\kappa_{\text{I}i}(|x| - X_i)) - \frac{w_{\text{II}} \kappa_{\text{III}i}}{w_{\text{I}} \kappa_{\text{I}i}} \sin(\kappa_{\text{II}i}X_i) \sin(\kappa_{\text{I}i}(|x| - X_i)), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.5a)$$

and

$$b_i(x) = \begin{cases} \frac{1}{\kappa_{\text{II}i}} \sin(\kappa_{\text{II}i}x), & |x| \leq X_i, \\ \frac{1}{\kappa_{\text{II}i}} \text{sign}(x) \left(\sin(\kappa_{\text{II}i}X_i) \cos(\kappa_{\text{I}i}(|x| - X_i)) + \frac{w_{\text{II}} \kappa_{\text{III}i}}{w_{\text{I}} \kappa_{\text{I}i}} \cos(\kappa_{\text{II}i}X_i) \sin(\kappa_{\text{I}i}(|x| - X_i)) \right), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.5b)$$

where

$$\kappa_{\text{II}i}^2 = k_0^2 n_{\text{II}}^2 + k_0^2 \mu_{\text{B}i}^2, \quad \kappa_{\text{I}i}^2 = \kappa_{\text{II}i}^2 + \Delta\kappa^2, \quad \Delta\kappa^2 = k_0^2 (n_{\text{I}}^2 - n_{\text{II}}^2). \quad (3.5c)$$

Indeed, one can easily show that a_i and b_i are two linearly independent basis solutions by looking at the scaled Wronskian $w_i W(a_i, b_i)$ which is constant and equal to w_{II} according to (3.4). Moreover we see that a_i is an even function and b_i is an odd function which is very similar to the cosine and sine in the derivation of the Rayleigh modes. We now set

$$u_{Bi} = A_i a_i + B_i b_i, \quad (3.6)$$

and apply the pseudo-periodic boundary condition in order to get a relation for the unknown expansion coefficients A_i and B_i . Thus,

$$A_i a_i\left(\frac{\Lambda}{2}\right) + B_i b_i\left(\frac{\Lambda}{2}\right) = e^{-jk_x \Lambda} (A_i a_i\left(-\frac{\Lambda}{2}\right) + B_i b_i\left(-\frac{\Lambda}{2}\right)), \quad (3.7a)$$

$$A_i a_i'\left(\frac{\Lambda}{2}\right) + B_i b_i'\left(\frac{\Lambda}{2}\right) = e^{-jk_x \Lambda} (A_i a_i'\left(-\frac{\Lambda}{2}\right) + B_i b_i'\left(-\frac{\Lambda}{2}\right)). \quad (3.7b)$$

For this homogeneous set of linear equations to have a non-trivial solution we require that

$$(1 - e^{-jk_x\Lambda})^2 a_i(\frac{\Lambda}{2}) b_i'(\frac{\Lambda}{2}) - (1 + e^{-jk_x\Lambda})^2 a_i'(\frac{\Lambda}{2}) b_i(\frac{\Lambda}{2}) = 0, \quad (3.8)$$

where we have used the fact that both a_i and b_i' are even and a_i' and b_i are odd. Expanding this equation and using that $a_i(\frac{\Lambda}{2}) b_i'(\frac{\Lambda}{2}) - a_i'(\frac{\Lambda}{2}) b_i(\frac{\Lambda}{2}) = w_{\text{II}}/w_{\text{I}}$ results in the following *transcendental eigenvalue equation*

$$\begin{aligned} T_i(\kappa_{\text{II}i}^2) &:= \cos(2\kappa_{\text{II}i} X_i) \cos(\kappa_{\text{I}i}(\Lambda - 2X_i)) - \\ &\quad \frac{1}{2} \left(\frac{w_{\text{I}} \kappa_{\text{II}i}}{w_{\text{II}} \kappa_{\text{I}i}} + \frac{w_{\text{II}} \kappa_{\text{I}i}}{w_{\text{I}} \kappa_{\text{II}i}} \right) \sin(2\kappa_{\text{II}i} X_i) \sin(\kappa_{\text{I}i}(\Lambda - 2X_i)) \\ &= \cos(k_x \Lambda), \end{aligned} \quad (3.9)$$

which has a countably infinite set of real-valued solutions $\kappa_{\text{II}i,m}^2$ for $m \in \mathbb{N}_0$. Now we still need to determine the expansion coefficients A_i and B_i . One relation directly follows from the linear dependence of (3.7) while the other relation follows from normalising the eigenfunctions

$$B_{i,m} = -\frac{(1 - e^{-jk_x\Lambda}) a_{i,m}(\frac{\Lambda}{2})}{(1 + e^{-jk_x\Lambda}) b_{i,m}(\frac{\Lambda}{2})} A_{i,m} = -\frac{(1 + e^{-jk_x\Lambda}) a_{i,m}'(\frac{\Lambda}{2})}{(1 - e^{-jk_x\Lambda}) b_{i,m}'(\frac{\Lambda}{2})} A_{i,m}, \quad (3.10a)$$

$$1 = \langle u_{\text{B}i,m}, u_{\text{B}i,m} \rangle_{w_i}. \quad (3.10b)$$

Finally, looking at the equation for $v_{\text{B}i}$ we have infinitely many solutions of the form

$$v_{\text{B}i,m} = C_{\text{B}i,m}^- e^{k_0 \mu_{\text{B}i,m} (z - Z_i)} + C_{\text{B}i,m}^+ e^{-k_0 \mu_{\text{B}i,m} (z - Z_{i-1})}. \quad (3.11)$$

It can be shown that the countably infinite set of eigenfunctions $\{u_{\text{B}i,m}\}_{m=0}^{\infty}$ forms a complete set [12]. Therefore the total solution of the field in a grating layer is given by

$$F_{\text{B}i} = f \sum_{m=0}^{\infty} u_{\text{B}i,m} v_{\text{B}i,m}. \quad (3.12)$$

By looking at the z -dependent part in this expression we see that again the solution consists of two parts: waves moving in the positive z -direction with expansion coefficients $C_{\text{B}i,m}^+$ and waves moving in the negative z -direction with expansion coefficients $C_{\text{B}i,m}^-$. Because of (3.5c) and the properties of the transcendental equation which are explained in the next section, these waves consist of a finite number of propagating waves and an infinite number of evanescent waves. The interface boundary condition will be used later to determine the unknown expansion coefficients. In [9] another derivation of the Bloch modes can be found for a slightly different unit cell. However with the derivation in this section it is much easier to explain some important properties of the transcendental equation.

3.1.2 Conical diffraction

Since the conical diffraction case looks very similar to the planar diffraction case we will pay special attention to the differences. Although the conical diffraction case does not have two different polarisations anymore, it does consist of two separate problems resembling the two polarisations in the planar diffraction case. Again the method of separation of variables is used. In the layers $i = 1, \dots, K$ a field \tilde{F}_{Bi} is introduced which represents the x -component of either the magnetic or electric field respectively

$$\tilde{F}_{Bi}(x, z) = \tilde{f}\tilde{u}_{Bi}(x)\tilde{v}_{Bi}(z), \quad (3.13)$$

where $\tilde{f} = -jY$ for the magnetic field and $\tilde{f} = 1$ for the electric field. We purposely swapped the order of the magnetic and electric field in order to retain some symmetry later on. Substituting into (2.35), rearranging the terms and adding the proper boundary conditions gives

$$\left\{ \begin{array}{ll} \frac{1}{k_0^2}\tilde{u}_{Bi}'' + (n_{II}^2 - \frac{k_y^2}{k_0^2})\tilde{u}_{Bi} = -\tilde{\mu}_{Bi}^2\tilde{u}_{Bi}, & |x| < X_i, \\ \frac{1}{k_0^2}\tilde{u}_{Bi}'' + (n_I^2 - \frac{k_y^2}{k_0^2})\tilde{u}_{Bi} = -\tilde{\mu}_{Bi}^2\tilde{u}_{Bi}, & X_i < |x| < \frac{\Lambda}{2}, \\ \frac{1}{\tilde{w}_I}\tilde{u}_{Bi}(\pm X_i^-) = \frac{1}{\tilde{w}_I}\tilde{u}_{Bi}(\pm X_i^+), & \tilde{u}_{Bi}(\frac{\Lambda}{2}) = e^{-jk_x\Lambda}\tilde{u}_{Bi}(-\frac{\Lambda}{2}), \\ \tilde{u}_{Bi}'(\pm X_i^\mp) = \tilde{u}_{Bi}'(\pm X_i^\pm), & \tilde{u}_{Bi}'(\frac{\Lambda}{2}) = e^{-jk_x\Lambda}\tilde{u}_{Bi}'(-\frac{\Lambda}{2}), \end{array} \right. \quad (3.14a)$$

and

$$\left\{ \begin{array}{l} \frac{1}{k_0^2}\tilde{v}_{Bi}'' = \tilde{\mu}_{Bi}^2\tilde{v}_{Bi}, \quad Z_{i-1} < z < Z_i, \\ \text{Interface boundary conditions,} \end{array} \right. \quad (3.14b)$$

where $\tilde{\mu}_{Bi}^2$ is the separation constant or eigenvalue of the problem and where $\tilde{w}_I = \tilde{w}_{II} = 1$ for the magnetic field and $\tilde{w}_I = 1/n_I^2, \tilde{w}_{II} = 1/n_{II}^2$ for the electric field. Also here the *inner layer interface boundary conditions* at $x = \pm X_i$ can be derived from (2.33) by using the continuity of the tangential electromagnetic fields in (2.17)

$$\left[\left[\frac{1}{\tilde{w}_i}\tilde{F}_{Bi} \right] \right]_{x=\pm X_i} = 0, \quad (3.15a)$$

$$\left[\left[\tilde{w}_i \frac{\partial}{\partial x} \left(\frac{1}{\tilde{w}_i}\tilde{F}_{Bi} \right) \right] \right]_{x=\pm X_i} = 0, \quad (3.15b)$$

where $\tilde{w}_i = 1$ for the magnetic field and $\tilde{w}_i = 1/\epsilon_i'$ for the electric field. In the conical diffraction case the equation for the x -component of the magnetic (electric) field resembles the equation for the y -component of the electric (magnetic) field in the planar diffraction case with TE (TM) polarisation. Moreover note that there is a small difference in the inner layer interface boundary conditions where the reciprocal weight function

now also appears in front of the field. The differential operator corresponding to the problem for \tilde{u}_{Bi} is again self-adjoint and therefore all eigenvalues are real-valued and the eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the inner product $\langle g, h \rangle_{1/\tilde{w}_i}$. From the remainder of this section it will become clear why we have the reciprocal weight function in this inner product. Let \tilde{a}_i, \tilde{b}_i be two linearly independent basis solutions of equation (3.14a) which satisfy the inner layer interface boundary conditions at $x = \pm X_i$ (not necessarily satisfying the pseudo-periodic boundary condition) and

$$\begin{cases} \tilde{a}_i(0) = \tilde{w}_{\Pi}, & \tilde{b}_i(0) = 0, \\ \tilde{a}'_i(0) = 0, & \tilde{b}'_i(0) = \tilde{w}_{\Pi}. \end{cases} \quad (3.16)$$

Then

$$\tilde{a}_i(x) = \begin{cases} \tilde{w}_{\Pi} \cos(\tilde{\kappa}_{\Pi i} x), & |x| \leq X_i, \\ \tilde{w}_I \left(\cos(\tilde{\kappa}_{\Pi i} X_i) \cos(\tilde{\kappa}_{I i}(|x| - X_i)) - \frac{\tilde{w}_{\Pi} \tilde{\kappa}_{\Pi i}}{\tilde{w}_I \tilde{\kappa}_{I i}} \sin(\tilde{\kappa}_{\Pi i} X_i) \sin(\tilde{\kappa}_{I i}(|x| - X_i)) \right), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.17a)$$

and

$$\tilde{b}_i(x) = \begin{cases} \frac{\tilde{w}_{\Pi}}{\tilde{\kappa}_{\Pi i}} \sin(\tilde{\kappa}_{\Pi i} x), & |x| \leq X_i, \\ \frac{\tilde{w}_I}{\tilde{\kappa}_{\Pi i}} \text{sign}(x) \left(\sin(\tilde{\kappa}_{\Pi i} X_i) \cos(\tilde{\kappa}_{I i}(|x| - X_i)) + \frac{\tilde{w}_{\Pi} \tilde{\kappa}_{\Pi i}}{\tilde{w}_I \tilde{\kappa}_{I i}} \cos(\tilde{\kappa}_{\Pi i} X_i) \sin(\tilde{\kappa}_{I i}(|x| - X_i)) \right), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.17b)$$

where

$$\tilde{\kappa}_{\Pi i}^2 = k_0^2 n_{\Pi}^2 + k_0^2 \mu_{Bi}^2 - k_y^2, \quad \tilde{\kappa}_{I i}^2 = \tilde{\kappa}_{\Pi i}^2 + \Delta \kappa^2. \quad (3.17c)$$

Comparing the basis solutions of the conical diffraction case (3.17) with the basis solutions of the planar diffraction case (3.5), it is clear that when $\kappa_{\Pi i} = \tilde{\kappa}_{\Pi i}$ these solutions are related through $\tilde{a}_i = w_i a_i$ and $\tilde{b}_i = w_i b_i$. In this relation one should realise that the weight function w_i in the planar TE (or TM) case is actually identical to the weight function \tilde{w}_i in the conical case for the magnetic (or electric) field. Since the solutions in the planar diffraction case were linearly independent, so are the solutions in the conical diffraction case. Here the scaled Wronskian $(1/\tilde{w}_i)W(\tilde{a}_i, \tilde{b}_i)$ is constant and equal to \tilde{w}_{Π} according to (3.16). Moreover \tilde{a}_i is again an even function and \tilde{b}_i is an odd function. We now set

$$\tilde{u}_{Bi} = \tilde{A}_i \tilde{a}_i + \tilde{B}_i \tilde{b}_i, \quad (3.18)$$

and apply the pseudo-periodic boundary condition in order to get a relation for the unknown expansion coefficients \tilde{A}_i and \tilde{B}_i . Exactly the same relations as in (3.7) hold but now with a tilde above the expansion coefficients and basis solutions. Consequently the same transcendental eigenvalue equation as in (3.9) can be derived except that we have to replace $\kappa_{\text{II}i}$ with $\tilde{\kappa}_{\text{II}i}$. This means that for all conical angles of incidence with the same k_x we only need to solve one transcendental eigenvalue equation: the corresponding planar diffraction case! Because the factor k_y is still different for each conical angle of incidence with fixed k_x we do get different eigenvalues $\tilde{\mu}_{\text{B}i,m}$ although $\kappa_{\text{II},m} = \tilde{\kappa}_{\text{II},m}$ for $m \in \mathbb{N}_0$. Now we still need to determine the expansion coefficients \tilde{A}_i and \tilde{B}_i . One relation directly follows from the linear dependence of (3.7) while the other relation follows from normalising the eigenfunctions

$$\tilde{B}_{i,m} = -\frac{(1 - e^{-jk_x\Lambda})a_{i,m}(\frac{\Lambda}{2})}{(1 + e^{-jk_x\Lambda})b_{i,m}(\frac{\Lambda}{2})}\tilde{A}_{i,m} = -\frac{(1 + e^{-jk_x\Lambda})a'_{i,m}(\frac{\Lambda}{2})}{(1 - e^{-jk_x\Lambda})b'_{i,m}(\frac{\Lambda}{2})}\tilde{A}_{i,m}, \quad (3.19a)$$

$$1 = \langle \tilde{u}_{\text{B}i,m}, \tilde{u}_{\text{B}i,m} \rangle_{1/\tilde{w}_i}. \quad (3.19b)$$

Note that in (3.19a) we already replaced the conical basis solutions with planar basis solutions. These planar basis solution obviously correspond to the planar diffraction case with the same k_x . Then the conical basis solutions are related to these planar basis solutions through the weight function as mentioned before which in turn drops out of the fraction. Moreover, because the solutions of the planar diffraction case were orthogonal with respect to the inner product in (3.10b), now we have orthogonality and normalise with respect to the inner product in (3.19b). Combining these relations gives $\tilde{A}_{i,m} = A_{i,m}$ and therefore also $\tilde{B}_{i,m} = B_{i,m}$. Summarising, for a conical diffraction case with a certain value of k_x the solution for the problem in the periodic x -direction is given in terms of the solution of the corresponding planar diffraction case by $\tilde{u}_{\text{B}i,m} = w_i u_{\text{B}i,m}$. Finally, looking at the equation for $\tilde{v}_{\text{B}i}$ we have infinitely many solutions of the form

$$\tilde{v}_{\text{B}i,m} = \tilde{C}_{\text{B}i,m}^- e^{k_0 \tilde{\mu}_{\text{B}i,m} (z - Z_i)} + \tilde{C}_{\text{B}i,m}^+ e^{-k_0 \tilde{\mu}_{\text{B}i,m} (z - Z_{i-1})}. \quad (3.20)$$

The total solution of the field in a grating layer is then given by

$$\tilde{F}_{\text{B}i} = \tilde{f} \sum_{m=0}^{\infty} \tilde{u}_{\text{B}i,m} \tilde{v}_{\text{B}i,m} = \tilde{f} w_i \sum_{m=0}^{\infty} u_{\text{B}i,m} v_{\text{B}i,m}. \quad (3.21)$$

Naturally the same remarks hold about the propagating and evanescent waves as in the planar diffraction case. Also now the interface boundary conditions will be used later to determine the unknown expansion coefficients. In [24] a different derivation of the conical diffraction case can be found. There a special decomposition of the field is introduced whereas here we solve directly for the uncoupled field components that come out of Maxwell's equations. This seems more natural and allows us to compare the results better with the second discretisation strategy that we will discuss in the next chapter.

3.2 Properties transcendental equation

One of the main challenges with the Bloch mode method is finding the eigenvalues $\mu_{\text{Bi},m}^2$, which are related to the roots $\kappa_{\text{II},m}^2$ of the transcendental equation (3.9) through (3.5c) for the planar diffraction case. Since the root-finding in the conical diffraction case reduces to a planar diffraction case problem (and a similar remark can be made about the eigenfunctions) we only need to focus on the planar diffraction case in this section. Before zooming in to the numerical computation of these roots, we first discuss some general properties of the transcendental equation.

The first property deals with special angles of incidence and the corresponding simplification of the transcendental equation. As can be seen from (3.8) we can consider two special cases where the term between one of the pair of brackets enclosing the exponential becomes equal to zero. This happens when either $e^{-jk_x\Lambda} = 1$, the *periodic eigenvalue problem*, or when $e^{-jk_x\Lambda} = -1$, the *semi-periodic eigenvalue problem*. Both cases correspond to a so called *Littrow mount* where the angle of incidence is such that the diffracted light of one of the diffraction orders follows the same path as the incident one. The former corresponds to a Littrow mount where an even diffraction order goes back in the direction of the incident light whereas the latter corresponds to a Littrow mount where an odd diffraction order goes back in the direction of the incident light. For the periodic eigenvalue problem the roots of the transcendental equation can be found by solving for those κ_{II}^2 that satisfy the two separate problems

$$\begin{cases} a_i'(\frac{\Lambda}{2}) = 0, \\ u_i = A_i a_i, \end{cases} \quad \text{and} \quad \begin{cases} b_i(\frac{\Lambda}{2}) = 0, \\ u_i = B_i b_i, \end{cases} \quad (3.22)$$

where the second line gives the corresponding eigenfunction. The coefficients in the eigenfunctions are again found by normalising the solution according to (3.10b). Moreover the functions a_i and b_i defined in (3.5) simplify dramatically and are given by

$$a_i(x) = \begin{cases} \cos(\kappa_{\text{II}}x), & |x| \leq X_i, \\ \frac{\cos(\kappa_{\text{II}}X_i)}{\cos(\kappa_{\text{II}}(\frac{\Lambda}{2} - X_i))} \cos(\kappa_{\text{II}}(\frac{\Lambda}{2} - |x|)), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.23a)$$

and

$$b_i(x) = \begin{cases} \frac{1}{\kappa_{\text{II}}} \sin(\kappa_{\text{II}}x), & |x| \leq X_i, \\ \frac{1}{\kappa_{\text{II}}} \frac{\sin(\kappa_{\text{II}}X_i)}{\sin(\kappa_{\text{II}}(\frac{\Lambda}{2} - X_i))} \text{sign}(x) \sin(\kappa_{\text{II}}(\frac{\Lambda}{2} - |x|)), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.23b)$$

Here we used the relations $a_i'(\frac{\Lambda}{2}) = 0$ and $b_i(\frac{\Lambda}{2}) = 0$ to rewrite the constant in front of the $\sin(\kappa_{\text{II}}(|x| - X_i))$ appearing in (3.5) and applied the angle sum and difference

identities for the sine and cosine. Similarly for the semi-periodic eigenvalue problem we can write the two separate problems

$$\begin{cases} a_i(\frac{\Lambda}{2}) = 0, \\ u_i = A_i a_i, \end{cases} \quad \text{and} \quad \begin{cases} b_i'(\frac{\Lambda}{2}) = 0, \\ u_i = B_i b_i. \end{cases} \quad (3.24)$$

Also now the functions a_i and b_i simplify dramatically and by following the same steps as for the periodic eigenvalue problem we arrive at

$$a_i(x) = \begin{cases} \cos(\kappa_{\text{II}i} x), & |x| \leq X_i, \\ \frac{\cos(\kappa_{\text{II}i} X_i)}{\sin(\kappa_{\text{II}i}(\frac{\Lambda}{2} - X_i))} \sin(\kappa_{\text{II}i}(\frac{\Lambda}{2} - |x|)), & X_i \leq |x| \leq \frac{\Lambda}{2}, \end{cases} \quad (3.25a)$$

and

$$b_i(x) = \begin{cases} \frac{1}{\kappa_{\text{II}i}} \sin(\kappa_{\text{II}i} x), & |x| \leq X_i, \\ \frac{1}{\kappa_{\text{II}i} \cos(\kappa_{\text{II}i}(\frac{\Lambda}{2} - X_i))} \text{sign}(x) \cos(\kappa_{\text{II}i}(\frac{\Lambda}{2} - |x|)), & X_i \leq |x| \leq \frac{\Lambda}{2}. \end{cases} \quad (3.25b)$$

The second property is formulated in a theorem that describes the general behaviour of the function T_i in the transcendental equation. The proof of this theorem can be found in [15]. If we denote the roots of the periodic eigenvalue problem by $\widehat{\kappa}_{\text{II}i,m}^2$ and of the semi-periodic eigenvalue problem by $\widetilde{\kappa}_{\text{II}i,m}^2$ then

Theorem 3.1 (i). *The roots of the periodic and semi-periodic eigenvalue problem occur in the order*

$$\widehat{\kappa}_{\text{II}i,0}^2 < \widetilde{\kappa}_{\text{II}i,0}^2 \leq \widetilde{\kappa}_{\text{II}i,1}^2 < \widehat{\kappa}_{\text{II}i,1}^2 \leq \widehat{\kappa}_{\text{II}i,2}^2 < \widetilde{\kappa}_{\text{II}i,2}^2 \leq \widetilde{\kappa}_{\text{II}i,3}^2 < \dots$$

(ii). *In the intervals $[\widehat{\kappa}_{\text{II}i,2m}^2, \widetilde{\kappa}_{\text{II}i,2m}^2]$, $T_i(\kappa_{\text{II}i}^2)$ decreases from 1 to -1 .*

(iii). *In the intervals $[\widetilde{\kappa}_{\text{II}i,2m+1}^2, \widehat{\kappa}_{\text{II}i,2m+1}^2]$, $T_i(\kappa_{\text{II}i}^2)$ increases from -1 to 1.*

(iv). *In the intervals $(-\infty, \widehat{\kappa}_{\text{II}i,0}^2)$ and $(\widehat{\kappa}_{\text{II}i,2m+1}^2, \widehat{\kappa}_{\text{II}i,2m+2}^2)$, $T_i(\kappa_{\text{II}i}^2) > 1$.*

(v). *In the intervals $(\widetilde{\kappa}_{\text{II}i,2m}^2, \widetilde{\kappa}_{\text{II}i,2m+1}^2)$, $T_i(\kappa_{\text{II}i}^2) < -1$.*

In Figure 3.1 a plot of a typical function T_i is shown where $n_{\text{I}} = 1$, $n_{\text{II}} = 3.77$, $X_i = 0.25$, $\Lambda = 1$, $\lambda_0 = 0.6328$. The left picture shows the behaviour on the interval $[0, 5000]$ while the right picture zooms in on that same interval. Here the outermost dashed lines correspond to the periodic and semi-periodic eigenvalue problem whereas the middle dashed line at height $\frac{1}{2}$ corresponds to a general eigenvalue problem. The intersections of the horizontal dashed line at height 1 with the oscillatory solid line are the roots $\widehat{\kappa}_{\text{II}i,m}^2$ of the periodic eigenvalue problem. Similarly the roots $\widetilde{\kappa}_{\text{II}i,m}^2$ of the semi-periodic eigenvalue problem are given by the intersections of the horizontal dashed line at height

-1 with the oscillatory solid line. For a general angle of incidence not corresponding to a Littrow mount the roots of the transcendental equation are given by the intersections of the oscillatory solid line and a horizontal line at some height $-1 < \cos(k_x \Lambda) < 1$ of which the dashed line at height $\frac{1}{2}$ is an example. The behaviour of the function

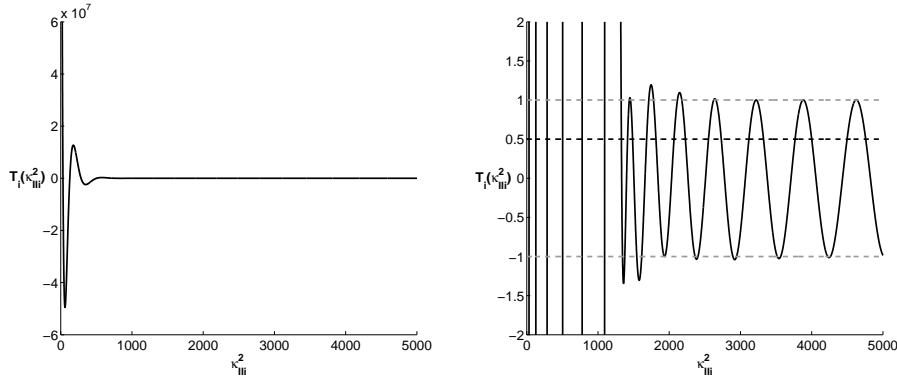


Figure 3.1: Typical plot of the function T_i where $n_I = 1$, $n_{II} = 3.77$, $X_i = 0.25$, $\Lambda = 1$, $\lambda_0 = 0.6328$. The left picture shows the behaviour on the interval $[0, 5000]$ while the right picture zooms in on that same interval. The outermost dashed lines correspond to the periodic and semi-periodic eigenvalue problem whereas the middle dashed line at height $\frac{1}{2}$ corresponds to a general eigenvalue problem.

T_i is in agreement with Theorem 3.1 but is clearly different in the interval $[0, -\Delta\kappa^2]$ and $[-\Delta\kappa^2, 5000]$ with $-\Delta\kappa^2 = 1303$. The latter corresponds to the case where κ_{IIi} and κ_{Ii} are both real-valued and the function T_i is the sum of the product of two sines and cosines. The former corresponds to the case where κ_{IIi} is still real-valued but κ_{Ii} is imaginary-valued resulting in a function T_i , which is the sum of the product of a sine and hyperbolic sine and of a cosine and hyperbolic cosine. For negative values of κ_{IIi}^2 the function T_i is the sum of the product of two hyperbolic sines and hyperbolic cosines and is therefore always greater than 1. In this case there is no intersection point with any of the dashed lines for negative values of κ_{IIi}^2 . When $n_I > n_{II}$ the role of κ_{IIi} and κ_{Ii} is somewhat reversed. Then the crossover point located at zero is defined for κ_{IIi} and T_i is always greater than one for $\kappa_{IIi}^2 = -\Delta\kappa^2 < 0$, effectively translating the graph from Figure 3.1 to the left over a distance $\Delta\kappa^2$.

Now that we know some general properties of the transcendental equation we can start discussing the strategy of actually computing the roots of this transcendental equation [32]. It is important to note that one typically searches for the first M roots of the transcendental equation and not just any M roots. This is because the first few roots are related to the dominant behaviour of the electromagnetic field, while the other roots are related to the higher order effects. We will first discuss the case of a Littrow mounting with periodic boundary conditions. From Figure 3.1 one might think that finding these roots is very complicated since they are so close together. It turns out that this is not a real problem because of the first property that we described at the beginning of this sec-

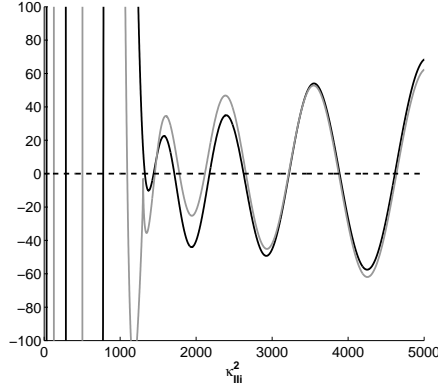


Figure 3.2: Typical plot of $a_i'(\frac{\Lambda}{2})$ (solid black) and $b_i(\frac{\Lambda}{2})$ (solid grey) but now as a function of κ_{II}^2 where $n_I = 1, n_{II} = 3.77, X_i = 0.25, \Lambda = 1, \lambda_0 = 0.6328$ for the periodic eigenvalue problem.

tion. There we explained that the roots of the transcendental equation can be found by solving two separate problems (3.22). The two separate problems also separate the roots as can be seen in Figure 3.2. This figure shows $a_i'(\frac{\Lambda}{2})$ and $b_i(\frac{\Lambda}{2})$ but now as a function of κ_{II}^2 where we look for the intersections with the dashed line or x -axis. The equation for $a_i'(\frac{\Lambda}{2}) = 0$ reads

$$-\kappa_{II} \cos(\kappa_{II} X_i) \sin(\kappa_{II} (\frac{\Lambda}{2} - X_i)) - \frac{w_{II}}{w_I} \kappa_{II} \sin(\kappa_{II} X_i) \cos(\kappa_{II} (\frac{\Lambda}{2} - X_i)) = 0. \quad (3.26)$$

Dividing by both cosines and multiplying by -1 gives

$$\kappa_{II} \tan(\kappa_{II} (\frac{\Lambda}{2} - X_i)) + \frac{w_{II}}{w_I} \kappa_{II} \tan(\kappa_{II} X_i) = 0. \quad (3.27)$$

This expression is very useful because both parts are monotonically increasing functions. The asymptotes of both tangents can be used to define intervals in which there is exactly one solution to the original equation (3.26). Note that for the interval $[0, -\Delta\kappa^2]$ the tangent with κ_{II} becomes a hyperbolic tangent and therefore has no asymptotes anymore. The collection of asymptotes is given by

$$\tilde{\zeta}_i^2 = \begin{cases} \left(\frac{\frac{\pi}{2} + p\pi}{X_i}\right)^2, & p = 0, 1, \dots, \\ \left(\frac{\frac{\pi}{2} + q\pi}{\frac{\Lambda}{2} - X_i}\right)^2 - \Delta\kappa^2, & q = 0, 1, \dots \end{cases} \quad (3.28)$$

After sorting these asymptotes in increasing order the interval in which (3.26) has exactly one solution is simply given by two subsequent values in the sorted list. Note that we also have to add 0 to this list since the expression on the left of the equal sign in (3.27) is negative and therefore has its first root in the interval from 0 to the first asymp-

tote. For the root-finding algorithm we used the standard routine `fzero` in Matlab V7.3 (R2006b) which uses a combination of bisection, secant, and inverse quadratic interpolation methods [16]. To speed up computations one can also use a standard Newton scheme and use `fzero` as an alternative if Newton fails to find a root in the desired interval. Note that although the original formulation (3.26) did not consist of monotonic functions, it still has the same roots as (3.27) and therefore the intervals derived for the tangent formulation can also be used here to guarantee exactly one root. Similarly the equation for $b_i(\frac{\Lambda}{2}) = 0$ reads

$$\frac{1}{\kappa_{\text{II}}} \sin(\kappa_{\text{II}} X_i) \cos(\kappa_{\text{I}}(\frac{\Lambda}{2} - X_i)) + \frac{w_{\text{II}}}{w_{\text{I}}} \frac{1}{\kappa_{\text{I}}} \cos(\kappa_{\text{II}} X_i) \sin(\kappa_{\text{I}}(\frac{\Lambda}{2} - X_i)) = 0. \quad (3.29)$$

Dividing by both sines and multiplying by $\kappa_{\text{I}} \kappa_{\text{II}}$ gives

$$\kappa_{\text{I}} \cot(\kappa_{\text{I}}(\frac{\Lambda}{2} - X_i)) + \frac{w_{\text{II}}}{w_{\text{I}}} \kappa_{\text{II}} \cot(\kappa_{\text{II}} X_i) = 0. \quad (3.30)$$

The same procedure can be repeated as before except that now we have monotonically decreasing functions. The collection of asymptotes is given by

$$\widehat{\xi}_i^2 = \begin{cases} \left(\frac{p\pi}{X_i}\right)^2, & p = 1, 2, \dots, \\ \left(\frac{q\pi}{\frac{\Lambda}{2} - X_i}\right)^2 - \Delta\kappa^2, & q = 0, 1, \dots \end{cases} \quad (3.31)$$

Note that $p = 0$ is not an asymptote since the limit for $\kappa_{\text{II}} \rightarrow 0$ exists for the expression on the left of the equal sign in (3.30). If this limit is positive we still need to add zero to the sorted list of asymptotes because then the first root lies in the interval from 0 to the first asymptote. The root-finding algorithm remains the same which concludes the strategy for the periodic eigenvalue problem.

For a Littrow mounting with semi-periodic boundary conditions exactly the same strategy can be used. Also here the roots lie very close together but the transcendental equation can again be separated. Now we look for the solutions of the problems $a_i(\frac{\Lambda}{2}) = 0$ and $b_i'(\frac{\Lambda}{2}) = 0$. By using the monotonicity and asymptotes of the corresponding tangent and cotangent formulations we can derive the intervals in which there lies exactly one root. This root is found with the same root-finding algorithm and thereby completing the strategy for the semi-periodic eigenvalue problem.

For a general angle of incidence not corresponding to a Littrow mount the roots of the transcendental are no longer the roots of two separate problems. The strategy we suggest here is to find the roots of the periodic and semi-periodic eigenvalue problem first and then use the properties of Theorem 3.1. This means that there is exactly one root in the interval given by (ii) and (iii) of the theorem which can be found by using once more the root-finding algorithm mentioned before.

3.3 Matching at the interfaces and the coupling matrix

The electromagnetic fields in the superstrate and substrate are described by the Rayleigh expansion, whereas the fields in the grating layers are given by the Bloch mode expansion. In this section the interface boundary conditions are applied in order to get the necessary relations between the unknown expansion coefficients that still appear in these expansions. Since these expansions are based on the eigenfunctions of each layer, they also differ from layer to layer and a coupling operator needs to be computed.

3.3.1 Planar diffraction

For the planar diffraction case the necessary equations for the unknown expansion coefficients follow from the interface boundary conditions (2.27) or (2.32) depending on the polarisation. From these equations it can be seen that only the y -component of the electromagnetic fields is required. Recall that the Rayleigh expansions in the superstrate and substrate were derived for a general conical angle of incidence. This means that here we only use the y -component of the expansions in (2.49) and (2.50) for TE and TM polarisation respectively. Additionally the solution inside the grating layers is given by the Bloch mode expansion in (3.12) for both polarisations. Summarising we get for $z = Z_0$

$$\nu_I \sum_{m=-\infty}^{\infty} (C_{B0,m}^- + \delta_{0,m}) e^{-jk_{xm}x} = \sum_{m=0}^{\infty} u_{B1,m} v_{B1,m}(0), \quad (3.32a)$$

$$j\omega_0 \nu_I \sum_{m=-\infty}^{\infty} \frac{k_{1,zm}}{k_0} (C_{B0,m}^- - \delta_{0,m}) e^{-jk_{xm}x} = \frac{1}{k_0} \sum_{m=0}^{\infty} w_1 u_{B1,m} v'_{B1,m}(0), \quad (3.32b)$$

for $z = Z_i$ and $i = 1, \dots, K-1$

$$\sum_{m=0}^{\infty} u_{B i,m} v_{B i,m}(Z_i) = \sum_{m=0}^{\infty} u_{B i+1,m} v_{B i+1,m}(Z_i), \quad (3.32c)$$

$$\frac{1}{k_0} \sum_{m=0}^{\infty} w_i u_{B i,m} v'_{B i,m}(Z_i) = \frac{1}{k_0} \sum_{m=0}^{\infty} w_{i+1} u_{B i+1,m} v'_{B i+1,m}(Z_i), \quad (3.32d)$$

and for $z = Z_K$

$$\sum_{m=0}^{\infty} u_{BK,m} v_{BK,m}(Z_K) = \nu_{II} \sum_{m=-\infty}^{\infty} C_{BK+1,m}^+ e^{-jk_{xm}x}, \quad (3.32e)$$

$$\frac{1}{k_0} \sum_{m=0}^{\infty} w_K u_{BK,m} v'_{BK,m}(Z_K) = -j\omega_{K+1} \nu_{II} \sum_{m=-\infty}^{\infty} \frac{k_{II,zm}}{k_0} C_{BK+1,m}^+ e^{-jk_{xm}x}, \quad (3.32f)$$

with $\nu_I = \nu_{II} = 1$ for TE polarisation and $\nu_I = jn_I$, $\nu_{II} = jn_{II}$ for TM polarisation where the latter comes from dividing out the term $-jY$. Moreover we extended the defi-

inition of the weight function w_i to the superstrate and substrate. The coefficients $C_{B0,m}^-$ represent either the expansion coefficients R_{ym}^e or R_{ym}^h depending on the polarisation. A similar remark can be made for the coefficients $C_{BK+1,m}^+$ in the substrate. Finally the known expansion coefficients of the incident field are given by $\delta_{0,m}$, the discrete Dirac delta function which is equal to one for $m = 0$ and zero otherwise. The system of equations (3.32) in known expansions functions (the eigenfunctions of Maxwell's equations in each layer) and with unknown expansion coefficients can be solved by the *method of moments* [18]. This method first defines a set of linearly independent *test functions*. Then the inner product of (3.32) with each of these test functions is taken, thereby eliminating the x -dependency of the equations. Finally the resulting algebraic linear system of equations of infinite dimension is truncated and solved for the unknown expansion coefficients. Choosing these test functions still allows for some freedom and in this thesis we will mention two approaches frequently used for the grating diffraction problem. The first approach is known as the *homogeneous method* where for each interface the same set of test functions is used for both interface boundary conditions. Moreover this set of test functions is chosen equal to the set of eigenfunctions in one of the layers adjacent to the interface. In this thesis we write down the second approach known as the *hybrid method* where for each interface first two different sets of test functions are used for the two interface boundary conditions. Moreover these two sets of test functions are chosen equal to the eigenfunctions in the two layers adjacent to the interface. The advantage of this approach is that the so called energy balance and reciprocity criteria are always satisfied, independent of the truncation order [7, 32]. Then we multiply (3.32a,c,e) with the conjugates of $e^{-jk_{xp}x}$, $w_i u_{B i,q}$ and $w_K u_{B K,q}$ respectively and after dividing by the pitch we integrate over the pitch. Similarly we multiply (3.32b,d,f) with the conjugates of $u_{B 1,q}$, $u_{B i+1,q}$ and $e^{-jk_{xp}x}$ respectively and after dividing by the pitch we again integrate over the pitch. Finally using the orthogonality of the eigenfunctions and substituting the solution for $v_{B i,m}$, the following set of equations can be derived for $p \in \mathbb{Z}$ and $q \in \mathbb{N}_0$

$$\nu_1 (C_{B0,p}^- + \delta_{0,p}) = \sum_{m=0}^{\infty} (C_{B1,m}^- x_{B1,m} + C_{B1,m}^+) \overline{q_{B0,mp}}, \quad (3.33a)$$

$$j\omega_0 \nu_1 \sum_{m=-\infty}^{\infty} \frac{k_{1,zm}}{k_0} (C_{B0,m}^- - \delta_{0,m}) q_{B0,qm} = \mu_{B1,q} (C_{B1,q}^- x_{B1,q} - C_{B1,q}^+), \quad (3.33b)$$

for $q \in \mathbb{N}_0$ and $i = 1, \dots, K-1$

$$C_{B i,q}^- + C_{B i,q}^+ x_{B i,q} = \sum_{m=0}^{\infty} (C_{B i+1,m}^- x_{B i+1,m} + C_{B i+1,m}^+) \overline{q_{B i,mq}}, \quad (3.33c)$$

$$\sum_{m=0}^{\infty} \mu_{B i,m} (C_{B i,m}^- - C_{B i,m}^+ x_{B i,m}) q_{B i,qm} = \mu_{B i+1,q} (C_{B i+1,q}^- x_{B i+1,q} - C_{B i+1,q}^+), \quad (3.33d)$$

and for $p \in \mathbb{Z}$ and $q \in \mathbb{N}_0$

$$C_{\text{BK},q}^- + C_{\text{BK},q}^+ x_{\text{BK},q} = \nu_{\text{II}} \sum_{m=-\infty}^{\infty} C_{\text{BK}+1,m}^+ \overline{q_{\text{BK},mq}}, \quad (3.33e)$$

$$\sum_{m=0}^{\infty} \mu_{\text{BK},m} (C_{\text{BK},m}^- - C_{\text{BK},m}^+ x_{\text{BK},m}) q_{\text{BK},pm} = -j\omega_{K+1} \nu_{\text{II}} \frac{k_{\text{II},zp}}{k_0} C_{\text{BK}+1,p}^+, \quad (3.33f)$$

with the coupling coefficients between two adjacent layers for $i = 1, \dots, K-1$

$$q_{\text{B}0,qm} = \langle e^{-jk_{xm}x}, u_{\text{B}1,q} \rangle, \quad q_{\text{B}i,qm} = \langle u_{\text{B}i,m}, u_{\text{B}i+1,q} \rangle_{w_i}, \quad q_{\text{BK},pm} = \langle u_{\text{BK},m}, e^{-jk_{xp}x} \rangle_{w_K}, \quad (3.34)$$

and where $x_{\text{B}i,m} = e^{-k_0 \mu_{\text{B}i,m} h_i}$. In order to solve this algebraic linear system of equations for the unknown expansion coefficients we truncate the series. This means that the index p and all sums that run from $-\infty$ to ∞ are truncated from $-M$ to M . Similarly the index q and all sums that run from 0 to ∞ are truncated from 0 to $2M$. Finally equations (3.33b,d,f) are multiplied with -1 which makes the comparison with other solution algorithms easier later on. Summarising we get for $i = 1, \dots, K-1$

$$\nu_{\text{I}} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q}_{\text{B}0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ j\omega_0 \mathbf{K}_{\text{I},z} & -j\omega_0 \mathbf{K}_{\text{I},z} \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{c}_{\text{B}0}^- \end{bmatrix} = \mathcal{F}_{\text{B}1}(Z_0) \mathbf{c}_{\text{B}1}, \quad (3.35a)$$

$$\mathcal{F}_{\text{B}i}(Z_i) \mathbf{c}_{\text{B}i} = \mathcal{F}_{\text{B}i+1}(Z_i) \mathbf{c}_{\text{B}i+1}, \quad (3.35b)$$

$$\mathcal{F}_{\text{BK}}(Z_K) \mathbf{c}_{\text{BK}} = \nu_{\text{II}} \begin{bmatrix} \mathbf{Q}_{\text{BK}}^{\text{H}} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ j\omega_{K+1} \mathbf{K}_{\text{II},z} \end{bmatrix} \mathbf{c}_{\text{BK}+1}^+. \quad (3.35c)$$

Here \mathbf{I} is simply the identity matrix and $\mathbf{K}_{\text{I},z}$ and $\mathbf{K}_{\text{II},z}$ are diagonal matrices containing the terms $k_{\text{I},zm}/k_0$ and $k_{\text{II},zm}/k_0$ respectively. The matrices $\mathbf{Q}_{\text{B}0}$ and \mathbf{Q}_{BK} naturally contain the coupling coefficients from (3.34) where the superscript H denotes the complex conjugate transposed of a matrix. Moreover the terms $C_{\text{B}0,m}^-$, $C_{\text{BK}+1,m}^+$ and $\delta_{0,m}$ are collected in the vectors $\mathbf{c}_{\text{B}0}^-$, $\mathbf{c}_{\text{BK}+1}^+$ and \mathbf{d}_0 . Finally we introduced some shorthand notation for the solution in the grating layers evaluated at the different interfaces

$$\mathcal{F}_{\text{B}i}(Z_{i-1}) = \begin{bmatrix} \mathbf{Q}_{\text{B}i-1}^{\text{H}} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{M}_{\text{B}i} & -\mathbf{M}_{\text{B}i} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{X}_{\text{B}i} \end{bmatrix}, \quad (3.36a)$$

$$\mathcal{F}_{\text{B}i}(Z_i) = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q}_{\text{B}i} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{M}_{\text{B}i} & -\mathbf{M}_{\text{B}i} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{\text{B}i} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}, \quad (3.36b)$$

$$\mathbf{c}_{\text{B}i} = \begin{bmatrix} \mathbf{c}_{\text{B}i}^+ \\ \mathbf{c}_{\text{B}i}^- \end{bmatrix}, \quad (3.36c)$$

where $\mathbf{M}_{\text{B}i}$ and $\mathbf{X}_{\text{B}i}$ are diagonal matrices containing the terms $\mu_{\text{B}i,m}$ and $x_{\text{B}i,m}$ respectively and $\mathbf{Q}_{\text{B}i}$ naturally contains the other coupling coefficients from (3.34). The solution in the grating layers $\mathcal{F}_{\text{B}i}$ is sometimes referred to as the *fundamental solution*. Moreover the terms $C_{\text{B}i,m}^-$ and $C_{\text{B}i,m}^+$ are collected in the vectors $\mathbf{c}_{\text{B}i}^-$ and $\mathbf{c}_{\text{B}i}^+$ respectively. Elim-

3.3.2 Conical diffraction

For the conical diffraction case the necessary equations for the unknown expansion coefficients follow from the interface boundary conditions (2.36). From these equations it can be seen that both the x - and y -component of the electric and magnetic field are required. As mentioned before the Rayleigh expansions in the superstrate and substrate were derived for a general conical angle of incidence. Therefore we can simply use the x - and y -component of the expansions in (2.49) and (2.50). Additionally the solution for the x -component of the electric and magnetic field inside the grating layers is given by the Bloch mode expansion in (3.23). This means that inside the grating layers we still need to find an expression for the y -component of these fields before we can apply the interface boundary conditions. These expressions are derived from (2.34b) (after multiplying with ε_i^r) and (2.34d) which we can rewrite to

$$\left(\varepsilon_i^r \frac{\partial}{\partial x} \frac{1}{\varepsilon_i^r} \frac{\partial}{\partial x} + k_0^2 \varepsilon_i^r\right) H_{i,y} = -jk_y \varepsilon_i^r \frac{\partial}{\partial x} \left(\frac{1}{\varepsilon_i^r} H_{i,x}\right) + j\omega \varepsilon_0 \varepsilon_i^r \frac{\partial}{\partial z} E_{i,x}, \quad (3.40a)$$

$$\left(\frac{\partial^2}{\partial x^2} + k_0^2 \varepsilon_i^r\right) E_{i,y} = -j\omega \mu_0 \frac{\partial}{\partial z} H_{i,x} - jk_y \frac{\partial}{\partial x} E_{i,x}. \quad (3.40b)$$

The extra multiplication with the relative permittivity is not necessary but makes it easier to link the problem to the planar diffraction case later on. In order to write down the solution of (3.40) both a particular solution as well as a homogeneous solution needs to be found. First a particular solution is derived in terms of the solution components $\tilde{u}_{B,i,m}$ and $\tilde{v}_{B,i,m}$. This makes sense since these solution components were used to write down the solution of the x -component of the electromagnetic field which now appears in the inhomogeneous parts of the equation. By using (3.14) it can be verified that the particular solution is given by

$$H_{i,y}^{\text{part}} = -jY \left(\frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^{\text{h}} \tilde{u}_{B,i,m}^{\text{h}'} \tilde{v}_{B,i,m}^{\text{h}} + \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^{\text{e}} \varepsilon_i^r \tilde{u}_{B,i,m}^{\text{e}} \tilde{v}_{B,i,m}^{\text{e}'} \right), \quad (3.41a)$$

$$E_{i,y}^{\text{part}} = \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^{\text{h}} \tilde{u}_{B,i,m}^{\text{h}} \tilde{v}_{B,i,m}^{\text{h}'} + \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^{\text{e}} \tilde{u}_{B,i,m}^{\text{e}'} \tilde{v}_{B,i,m}^{\text{e}}, \quad (3.41b)$$

$$\tilde{\lambda}_{i,m} = \frac{k_0^2}{k_0^2 \tilde{\mu}_{B,i,m}^2 - k_y^2}. \quad (3.41c)$$

Here an extra superscript is introduced to distinguish between the x -component of the electric and magnetic field. From the expressions above we see that a problem arises when a conical eigenvalue $\tilde{\mu}_{B,i,m}$ is equal to k_y/k_0 and thus one of the denominators in (3.41c) is zero. Given a certain grating it is very unlikely that this will actually happen, although one should check the transcendental eigenvalue equation before doing any of the computations. In the remainder of this thesis we will assume not to be in this special case. Under this assumption we now show that the homogeneous solution of (3.40) is identical to the trivial solution. To this end first the boundary conditions for the homogeneous equation are investigated. Since the solution of the total electromagnetic

field consisting of the homogeneous solution plus the particular solution is pseudo-periodic, it is clear that having found an expression for the particular solution which is pseudo-periodic also the homogeneous solution must be pseudo-periodic. Moreover, the inner layer interface boundary conditions at $x = \pm X_i$ for the homogeneous solution reduce to the ones we already derived for the planar diffraction case in (3.3). This can be seen from (2.33) by using the continuity of the tangential electromagnetic fields in (2.17), the inner layer interface boundary conditions in (3.15) and the particular solution in (3.41). But this suggests that the x -dependent part of the homogeneous solution must be equal to the non-trivial eigenfunction in the planar diffraction case corresponding to the planar eigenvalue equal to zero. This in turn corresponds to a conical eigenvalue equal to k_y/k_0 , which was exactly the special case we would not consider. Thus, the homogeneous solution is identical to the trivial solution and the total solution of the y -component of the electromagnetic field is simply given by the particular solutions in (3.41).

Now that all components are available the interface boundary conditions (2.36) can be applied which after reordering give for $z = Z_0$

$$jn_1 \sum_{m=-\infty}^{\infty} (R_{xm}^h + \delta_{0,m}(s_x \cos \psi - p_x \sin \psi)) e^{-jk_{xm}x} = \sum_{m=0}^{\infty} \tilde{u}_{B1,m}^h \tilde{v}_{B1,m}^h(0), \quad (3.42a)$$

$$\sum_{m=-\infty}^{\infty} (R_{xm}^e + \delta_{0,m}(s_x \sin \psi + p_x \cos \psi)) e^{-jk_{xm}x} = \sum_{m=0}^{\infty} \tilde{u}_{B1,m}^e \tilde{v}_{B1,m}^e(0), \quad (3.42b)$$

$$jn_1 \sum_{m=-\infty}^{\infty} (R_{ym}^h + \delta_{0,m}(s_y \cos \psi - p_y \sin \psi)) e^{-jk_{ym}x} = \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{1,m}^h \tilde{u}_{B1,m}^{h'} \tilde{v}_{B1,m}^h(0) + \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{1,m}^e \epsilon_1^r \tilde{u}_{B1,m}^e \tilde{v}_{B1,m}^{e'}, \quad (3.42c)$$

$$\sum_{m=-\infty}^{\infty} (R_{ym}^e + \delta_{0,m}(s_y \sin \psi + p_y \cos \psi)) e^{-jk_{ym}x} = \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{1,m}^h \tilde{u}_{B1,m}^{h'} \tilde{v}_{B1,m}^{h'}(0) + \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{1,m}^e \tilde{u}_{B1,m}^{e'} \tilde{v}_{B1,m}^e(0), \quad (3.42d)$$

for $z = Z_i$ and $i = 1, \dots, K-1$

$$\sum_{m=0}^{\infty} \tilde{u}_{B i,m}^h \tilde{v}_{B i,m}^h(Z_i) = \sum_{m=0}^{\infty} \tilde{u}_{B i+1,m}^h \tilde{v}_{B i+1,m}^h(Z_i), \quad (3.42e)$$

$$\sum_{m=0}^{\infty} \tilde{u}_{B i,m}^e \tilde{v}_{B i,m}^e(Z_i) = \sum_{m=0}^{\infty} \tilde{u}_{B i+1,m}^e \tilde{v}_{B i+1,m}^e(Z_i), \quad (3.42f)$$

$$\begin{aligned} \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^h \tilde{u}_{B i,m}^{h'} \tilde{v}_{B i,m}^h(Z_i) + \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^e \varepsilon_i^r \tilde{u}_{B i,m}^e \tilde{v}_{B i,m}^{e'}(Z_i) = \\ \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i+1,m}^h \tilde{u}_{B i+1,m}^{h'} \tilde{v}_{B i+1,m}^h(Z_i) + \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i+1,m}^e \varepsilon_{i+1}^r \tilde{u}_{B i+1,m}^e \tilde{v}_{B i+1,m}^{e'}(Z_i), \end{aligned} \quad (3.42g)$$

$$\begin{aligned} \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^h \tilde{u}_{B i,m}^h \tilde{v}_{B i,m}^{h'}(Z_i) + \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^e \tilde{u}_{B i,m}^{e'} \tilde{v}_{B i,m}^e(Z_i) = \\ \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i+1,m}^h \tilde{u}_{B i+1,m}^h \tilde{v}_{B i+1,m}^{h'}(Z_i) + \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i+1,m}^e \tilde{u}_{B i+1,m}^{e'} \tilde{v}_{B i+1,m}^e(Z_i), \end{aligned} \quad (3.42h)$$

and for $z = Z_K$

$$\sum_{m=0}^{\infty} \tilde{u}_{BK,m}^h \tilde{v}_{BK,m}^h(Z_K) = j n_{II} \sum_{m=-\infty}^{\infty} T_{xm}^h e^{-jk_{xm}x}, \quad (3.42i)$$

$$\sum_{m=0}^{\infty} \tilde{u}_{BK,m}^e \tilde{v}_{BK,m}^e(Z_K) = \sum_{m=-\infty}^{\infty} T_{xm}^e e^{-jk_{xm}x}, \quad (3.42j)$$

$$\begin{aligned} \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{K,m}^h \tilde{u}_{BK,m}^{h'} \tilde{v}_{BK,m}^h(Z_K) + \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{K,m}^e \varepsilon_K^r \tilde{u}_{BK,m}^e \tilde{v}_{BK,m}^{e'}(Z_K) = \\ j n_{II} \sum_{m=-\infty}^{\infty} T_{ym}^h e^{-jk_{xm}x}, \end{aligned} \quad (3.42k)$$

$$\begin{aligned} \frac{1}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{K,m}^h \tilde{u}_{BK,m}^h \tilde{v}_{BK,m}^{h'}(Z_K) + \frac{k_y}{k_0} \frac{j}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{K,m}^e \tilde{u}_{BK,m}^{e'} \tilde{v}_{BK,m}^e(Z_K) = \\ \sum_{m=-\infty}^{\infty} T_{ym}^e e^{-jk_{xm}x}. \end{aligned} \quad (3.42l)$$

Like in the planar diffraction case the known expansion coefficients of the incident field are given by the discrete Dirac delta function. Note that also here the constant term $-jY$ appearing in front of all the equations involving the matching of the magnetic field is divided out. The same holds for the y -dependency of the electromagnetic field since this term appears in all of the equations and does not depend on the summation index. Now again the *method of moments* can be used to solve for the unknown expansion coefficients. Following the hybrid method for each interface we first define four sets of *test functions* equal to the eigenfunctions in the two layers adjacent to the interface. Then we multiply (3.42a,b,e,f,i,j) with the conjugates of $\tilde{u}_{B1,q}^h$, $\varepsilon_1^r \tilde{u}_{B1,q}^e$, $\tilde{u}_{B i+1,q}^h$, $\varepsilon_{i+1}^r \tilde{u}_{B i+1,q}^e$, $e^{-jk_{xp}x}$ and $e^{-jk_{xp}x}$ respectively and after dividing by the pitch we integrate over the pitch. Similarly we multiply (3.32c,d,g,h,k,l) with the conjugates of $e^{-jk_{xp}x}$, $e^{-jk_{xp}x}$, $\tilde{u}_{B i,q}^e$, $\tilde{u}_{B i,q}^h$, $\tilde{u}_{BK,q}^e$ and $\tilde{u}_{BK,q}^h$ respectively and after dividing by the pitch we again integrate over the pitch. Finally using the orthogonality of the eigenfunctions and substituting the solutions for $\tilde{v}_{B i,m}$

and $\tilde{v}_{B_i,m}^j$, the following set of equations can be derived for $p \in \mathbb{Z}$ and $q \in \mathbb{N}_0$

$$jn_1 \sum_{m=-\infty}^{\infty} (R_{xm}^h + \delta_{0,m}(s_x \cos \psi - p_x \sin \psi)) \tilde{q}_{B_0,qm}^h = \tilde{C}_{B_1,q}^{h-} \tilde{x}_{B_1,q}^h + \tilde{C}_{B_1,q}^{h+}, \quad (3.43a)$$

$$\sum_{m=-\infty}^{\infty} (R_{xm}^e + \delta_{0,m}(s_x \sin \psi + p_x \cos \psi)) \tilde{q}_{B_0,qm}^e = \tilde{C}_{B_1,q}^{e-} \tilde{x}_{B_1,q}^e + \tilde{C}_{B_1,q}^{e+}, \quad (3.43b)$$

$$\begin{aligned} jn_1 (R_{yp}^h + \delta_{0,p}(s_y \cos \psi - p_y \sin \psi)) = \\ \frac{k_y}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{1,m}^h (\tilde{C}_{B_1,m}^{h-} \tilde{x}_{B_1,m}^h + \tilde{C}_{B_1,m}^{h+}) \tilde{q}_{B_1,pm}^{h'} + \\ \sum_{m=0}^{\infty} \tilde{\mu}_{B_1,m}^e \tilde{\lambda}_{1,m}^e (\tilde{C}_{B_1,m}^{e-} \tilde{x}_{B_1,m}^e - \tilde{C}_{B_1,m}^{e+}) \overline{\tilde{q}_{B_0,mp}^e}, \end{aligned} \quad (3.43c)$$

$$\begin{aligned} R_{yp}^e + \delta_{0,p}(s_y \sin \psi + p_y \cos \psi) = \\ \sum_{m=0}^{\infty} \tilde{\mu}_{B_1,m}^h \tilde{\lambda}_{1,m}^h (\tilde{C}_{B_1,m}^{h-} \tilde{x}_{B_1,m}^h - \tilde{C}_{B_1,m}^{h+}) \overline{\tilde{q}_{B_0,mp}^h} + \\ \frac{k_y}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{1,m}^e (\tilde{C}_{B_1,m}^{e-} \tilde{x}_{B_1,m}^e + \tilde{C}_{B_1,m}^{e+}) \tilde{q}_{B_1,pm}^{e'}, \end{aligned} \quad (3.43d)$$

for $q \in \mathbb{N}_0$ and $i = 1, \dots, K-1$

$$\sum_{m=0}^{\infty} (\tilde{C}_{B_i,m}^{h-} + \tilde{C}_{B_i,m}^{h+} \tilde{x}_{B_i,m}^h) \tilde{q}_{B_i,qm}^h = \tilde{C}_{B_{i+1},q}^{h-} \tilde{x}_{B_{i+1},q}^h + \tilde{C}_{B_{i+1},q}^{h+}, \quad (3.43e)$$

$$\sum_{m=0}^{\infty} (\tilde{C}_{B_i,m}^{e-} + \tilde{C}_{B_i,m}^{e+} \tilde{x}_{B_i,m}^e) \tilde{q}_{B_i,qm}^e = \tilde{C}_{B_{i+1},q}^{e-} \tilde{x}_{B_{i+1},q}^e + \tilde{C}_{B_{i+1},q}^{e+}, \quad (3.43f)$$

$$\begin{aligned} \frac{k_y}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^h (\tilde{C}_{B_i,m}^{h-} + \tilde{C}_{B_i,m}^{h+} \tilde{x}_{B_i,m}^h) \tilde{q}_{B_i,qm}^* + \tilde{\mu}_{B_i,q}^e \tilde{\lambda}_{i,q}^e (\tilde{C}_{B_i,q}^{e-} - \tilde{C}_{B_i,q}^{e+} \tilde{x}_{B_i,q}^e) = \\ \frac{k_y}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i+1,m}^h (\tilde{C}_{B_{i+1},m}^{h-} \tilde{x}_{B_{i+1},m}^h + \tilde{C}_{B_{i+1},m}^{h+}) \tilde{q}_{B_{i+1},qm}^{h'} + \\ \sum_{m=0}^{\infty} \tilde{\mu}_{B_{i+1},m}^e \tilde{\lambda}_{i+1,m}^e (\tilde{C}_{B_{i+1},m}^{e-} \tilde{x}_{B_{i+1},m}^e - \tilde{C}_{B_{i+1},m}^{e+}) \overline{\tilde{q}_{B_i,mq}^e} \end{aligned} \quad (3.43g)$$

$$\begin{aligned} \tilde{\mu}_{B_i,q}^h \tilde{\lambda}_{i,q}^h (\tilde{C}_{B_i,q}^{h-} - \tilde{C}_{B_i,q}^{h+} \tilde{x}_{B_i,q}^h) + \frac{k_y}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i,m}^e (\tilde{C}_{B_i,m}^{e-} + \tilde{C}_{B_i,m}^{e+} \tilde{x}_{B_i,m}^e) \overline{\tilde{q}_{B_i,mq}^*} = \\ \sum_{m=0}^{\infty} \tilde{\mu}_{B_{i+1},m}^h \tilde{\lambda}_{i+1,m}^h (\tilde{C}_{B_{i+1},m}^{h-} \tilde{x}_{B_{i+1},m}^h - \tilde{C}_{B_{i+1},m}^{h+}) \overline{\tilde{q}_{B_i,mq}^h} + \\ \frac{k_y}{k_0} \sum_{m=0}^{\infty} \tilde{\lambda}_{i+1,m}^e (\tilde{C}_{B_{i+1},m}^{e-} \tilde{x}_{B_{i+1},m}^e + \tilde{C}_{B_{i+1},m}^{e+}) \tilde{q}_{B_{i+1},qm}^{e'}, \end{aligned} \quad (3.43h)$$

and for $p \in \mathbb{Z}$ and $q \in \mathbb{N}_0$

$$\sum_{m=0}^{\infty} (\tilde{\mathbf{C}}_{\text{BK},m}^{\text{h}-} + \tilde{\mathbf{C}}_{\text{BK},m}^{\text{h}+} \tilde{\mathbf{x}}_{\text{BK},m}^{\text{h}}) \tilde{q}_{\text{BK},pm}^{\text{h}} = j n_{\text{II}} T_{xp}^{\text{h}}, \quad (3.43i)$$

$$\sum_{m=0}^{\infty} (\tilde{\mathbf{C}}_{\text{BK},m}^{\text{e}-} + \tilde{\mathbf{C}}_{\text{BK},m}^{\text{e}+} \tilde{\mathbf{x}}_{\text{BK},m}^{\text{e}}) \tilde{q}_{\text{BK},pm}^{\text{e}} = T_{xp}^{\text{e}}, \quad (3.43j)$$

$$\begin{aligned} & \sum_{m=0}^{\frac{k_y}{k_0}} \tilde{\lambda}_{K,m}^{\text{h}} (\tilde{\mathbf{C}}_{\text{BK},m}^{\text{h}-} + \tilde{\mathbf{C}}_{\text{BK},m}^{\text{h}+} \tilde{\mathbf{x}}_{\text{BK},m}^{\text{h}}) \tilde{q}_{\text{BK},qm}^* + \\ & \tilde{\mu}_{\text{BK},q}^{\text{e}} \tilde{\lambda}_{K,q}^{\text{e}} (\tilde{\mathbf{C}}_{\text{BK},q}^{\text{e}-} - \tilde{\mathbf{C}}_{\text{BK},q}^{\text{e}+} \tilde{\mathbf{x}}_{\text{BK},q}^{\text{e}}) = j n_{\text{II}} \sum_{m=-\infty}^{\infty} T_{ym}^{\text{h}} \overline{\tilde{q}_{\text{BK},mq}^{\text{h}}}, \end{aligned} \quad (3.43k)$$

$$\begin{aligned} & \tilde{\mu}_{\text{BK},q}^{\text{h}} \tilde{\lambda}_{K,q}^{\text{h}} (\tilde{\mathbf{C}}_{\text{BK},q}^{\text{h}-} - \tilde{\mathbf{C}}_{\text{BK},q}^{\text{h}+} \tilde{\mathbf{x}}_{\text{BK},q}^{\text{h}}) + \\ & \sum_{m=0}^{\frac{k_y}{k_0}} \tilde{\lambda}_{K,m}^{\text{e}} (\tilde{\mathbf{C}}_{\text{BK},m}^{\text{e}-} + \tilde{\mathbf{C}}_{\text{BK},m}^{\text{e}+} \tilde{\mathbf{x}}_{\text{BK},m}^{\text{e}}) \tilde{q}_{\text{BK},mq}^* = \sum_{m=-\infty}^{\infty} T_{ym}^{\text{e}} \overline{\tilde{q}_{\text{BK},mq}^{\text{e}}}, \end{aligned} \quad (3.43l)$$

with the coupling coefficients between two adjacent layers for $i = 1, \dots, K-1$

$$\tilde{q}_{\text{B}0,qm} = \langle e^{-jk_{xm}x}, \tilde{\mathbf{u}}_{\text{B}1,q} \rangle_{1/\tilde{w}_1}, \quad \tilde{q}_{\text{B}i,qm} = \langle \tilde{\mathbf{u}}_{\text{B}i,m}, \tilde{\mathbf{u}}_{\text{B}i+1,q} \rangle_{1/\tilde{w}_{i+1}}, \quad \tilde{q}_{\text{B}K,pm} = \langle \tilde{\mathbf{u}}_{\text{B}K,m}, e^{-jk_{xp}x} \rangle, \quad (3.44a)$$

$$\tilde{q}_{\text{B}1,pm}^{\text{h}'} = \frac{j}{k_0} \langle \tilde{\mathbf{u}}_{\text{B}1,m}^{\text{h}'}, e^{-jk_{xp}x} \rangle, \quad \tilde{q}_{\text{B}i+1,qm}^{\text{h}'} = \frac{j}{k_0} \langle \tilde{\mathbf{u}}_{\text{B}i+1,m}^{\text{h}'}, \tilde{\mathbf{u}}_{\text{B}i,q}^{\text{e}} \rangle, \quad (3.44b)$$

$$\tilde{q}_{\text{B}1,pm}^{\text{e}'} = \frac{j}{k_0} \langle \tilde{\mathbf{u}}_{\text{B}1,m}^{\text{e}'}, e^{-jk_{xp}x} \rangle, \quad \tilde{q}_{\text{B}i+1,qm}^{\text{e}'} = \frac{j}{k_0} \langle \tilde{\mathbf{u}}_{\text{B}i+1,m}^{\text{e}'}, \tilde{\mathbf{u}}_{\text{B}i,q}^{\text{h}} \rangle, \quad (3.44c)$$

and inside the grating layers for $i = 1, \dots, K$

$$\tilde{q}_{\text{B}i,qm}^* = \frac{j}{k_0} \langle \tilde{\mathbf{u}}_{\text{B}i,m}^{\text{h}'}, \tilde{\mathbf{u}}_{\text{B}i,q}^{\text{e}} \rangle, \quad (3.44d)$$

and where $\tilde{\mathbf{x}}_{\text{B}i,m} = e^{-k_0 \tilde{\mu}_{\text{B}i,m} h_i}$. Note that only one set of coupling coefficients (3.44d) inside the grating layers is defined. When the superscripts in this inner product are reversed we can use partial integration and the pseudo-periodic boundary conditions to rewrite this inner product in terms of the coupling coefficients. In order to solve this algebraic linear system of equations for the unknown expansion coefficients we truncate the series. This means that the index p and all sums that run from $-\infty$ to ∞ are truncated from $-M$ to M . Similarly the index q and all sums that run from 0 to ∞ are truncated from 0 to $2M$. Summarising we get for $i = 1, \dots, K-1$

$$\begin{bmatrix} \mathbf{Q}_{\text{B}0} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathcal{F}_{\text{I},s} & \mathcal{F}_{\text{I},c} \\ -\mathcal{F}_{\text{I},c} & \mathcal{F}_{\text{I},s} \end{bmatrix} \begin{bmatrix} \mathcal{J} & \mathcal{J} \\ \frac{1}{n_1} \mathcal{K}_{\text{I},z} & -\frac{1}{n_1} \mathcal{K}_{\text{I},z} \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{c}_{\text{B}0}^- \end{bmatrix} = \tilde{\mathcal{F}}_{\text{B}1}(Z_0) \tilde{\mathbf{c}}_{\text{B}1}, \quad (3.45a)$$

$$\tilde{\mathcal{F}}_{\text{B}i}(Z_i) \tilde{\mathbf{c}}_{\text{B}i} = \tilde{\mathcal{F}}_{\text{B}i+1}(Z_i) \tilde{\mathbf{c}}_{\text{B}i+1}, \quad (3.45b)$$

$$\tilde{\mathcal{F}}_{\text{B}K}(Z_K) \tilde{\mathbf{c}}_{\text{B}K} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q}_{\text{B}K}^{\text{H}} \end{bmatrix} \begin{bmatrix} \mathcal{F}_{\text{II},s} & \mathcal{F}_{\text{II},c} \\ -\mathcal{F}_{\text{II},c} & \mathcal{F}_{\text{II},s} \end{bmatrix} \begin{bmatrix} \mathcal{J} \\ \frac{1}{n_{\text{II}}} \mathcal{K}_{\text{II},z} \end{bmatrix} \mathbf{c}_{\text{B}K+1}^+, \quad (3.45c)$$

where the block diagonal and block anti-diagonal matrices are given by

$$\begin{aligned}
\mathbf{I} &= \text{diag}(\mathbf{I}, \mathbf{I}), & \mathbf{Q}_{B0} &= \text{diag}(\tilde{\mathbf{Q}}_{B0}^h, \tilde{\mathbf{Q}}_{B0}^e), \\
\mathbf{J} &= \text{a-diag}(-\mathbf{I}, -\mathbf{I}), & \mathbf{Q}_{BK} &= \text{diag}(\tilde{\mathbf{Q}}_{BK}^e, \tilde{\mathbf{Q}}_{BK}^h), \\
\mathcal{F}_{*,c} &= \text{diag}(jn_*\mathbf{F}_c, \mathbf{F}_c), & \mathcal{K}_{*,z} &= \text{diag}(-\mathbf{K}_{*,z}, \mathbf{K}_{*,z}), \\
\mathcal{F}_{*,s} &= \text{diag}(jn_*\mathbf{F}_s, \mathbf{F}_s), & &
\end{aligned} \tag{3.46}$$

and $\star = \text{I, II}$. Here the coupling coefficients from (3.44a) are collected in the matrices $\tilde{\mathbf{Q}}_{B0}$ and $\tilde{\mathbf{Q}}_{BK}$. Additionally two diagonal matrices \mathbf{F}_c and \mathbf{F}_s are introduced for the conical diffraction case only and contain the terms $\cos \phi_m$ and $\sin \phi_m$ respectively. Moreover it is important to note that (3.45a) and (3.45c) are not a direct result of truncating the equations. As an extra step we eliminated the expansion coefficients of the Rayleigh expansions and replaced them with the s - and p -polarised counterparts using (2.53) and (2.54). The resulting system of equations now has the same structure as in the planar diffraction case where the block vectors are given by

$$\mathbf{c}_{B0}^- = \begin{bmatrix} \mathbf{r}_s \\ \mathbf{r}_p \end{bmatrix}, \quad \mathbf{c}_{BK+1}^+ = \begin{bmatrix} \mathbf{t}_s \\ \mathbf{t}_p \end{bmatrix}, \quad \mathbf{d}_0 = \begin{bmatrix} \sin \psi \mathbf{d}_0 \\ \cos \psi \mathbf{d}_0 \end{bmatrix}. \tag{3.47}$$

Here the reflected field coefficients R_{sm} and R_{pm} are collected in the vectors \mathbf{r}_s and \mathbf{r}_p . The transmitted field vectors are defined in a similar way. Finally we introduced some shorthand notation for the solution in the grating layers or fundamental solution evaluated at the different interfaces

$$\tilde{\mathcal{F}}_{Bi}(Z_{i-1}) = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathcal{K}_y \mathbf{Q}'_{Bi} \mathcal{L}_i & \mathcal{J} \mathbf{Q}_{Bi-1}^H \mathcal{L}_i \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathcal{M}_{Bi} & -\mathcal{M}_{Bi} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathcal{X}_{Bi} \end{bmatrix}, \tag{3.48a}$$

$$\tilde{\mathcal{F}}_{Bi}(Z_i) = \begin{bmatrix} \mathbf{Q}_{Bi} & \mathbf{O} \\ \mathcal{K}_y \mathbf{Q}_{Bi}^* \mathcal{L}_i & \mathcal{J} \mathcal{L}_i \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathcal{M}_{Bi} & -\mathcal{M}_{Bi} \end{bmatrix} \begin{bmatrix} \mathcal{X}_{Bi} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}, \tag{3.48b}$$

$$\tilde{\mathbf{c}}_{Bi} = \begin{bmatrix} \mathbf{c}_{Bi}^+ \\ \mathbf{c}_{Bi}^- \end{bmatrix}, \quad \mathbf{c}_{Bi}^\pm = \begin{bmatrix} \tilde{\mathbf{c}}_{Bi}^{\pm h} \\ \tilde{\mathbf{c}}_{Bi}^{\pm e} \end{bmatrix}, \tag{3.48c}$$

where again some block diagonal matrices are introduced

$$\begin{aligned}
\mathcal{M}_{Bi} &= \text{diag}(\tilde{\mathbf{M}}_{Bi}^h, \tilde{\mathbf{M}}_{Bi}^e), & \mathbf{Q}_{Bi} &= \text{diag}(\tilde{\mathbf{Q}}_{Bi}^h, \tilde{\mathbf{Q}}_{Bi}^e), \\
\mathcal{X}_{Bi} &= \text{diag}(\tilde{\mathbf{X}}_{Bi}^h, \tilde{\mathbf{X}}_{Bi}^e), & \mathbf{Q}'_{Bi} &= \text{diag}(\tilde{\mathbf{Q}}_{Bi}^{h'}, \tilde{\mathbf{Q}}_{Bi}^{e'}), \\
\mathcal{L}_i &= \text{diag}(\tilde{\mathbf{L}}_i^h, \tilde{\mathbf{L}}_i^e), & \mathbf{Q}_{Bi}^* &= \text{diag}(\tilde{\mathbf{Q}}_{Bi}^{*h}, \tilde{\mathbf{Q}}_{Bi}^{*e}), \\
\mathcal{K}_y &= \text{diag}(\mathbf{K}_y, \mathbf{K}_y). & &
\end{aligned} \tag{3.49}$$

The matrices $\tilde{\mathbf{M}}_{Bi}$, $\tilde{\mathbf{X}}_{Bi}$ and $\tilde{\mathbf{L}}_i$ are diagonal matrices containing the eigenvalues $\tilde{\mu}_{Bi,m}$ and derived quantities $\tilde{\alpha}_{Bi,m}$ and $\tilde{\lambda}_{i,m}$ respectively. Moreover the matrices $\tilde{\mathbf{Q}}_{Bi}$, $\tilde{\mathbf{Q}}'_{Bi}$ and $\tilde{\mathbf{Q}}_{Bi}^*$ simply consist of the coupling coefficients from (3.44). Naturally the diagonal matrix \mathbf{K}_y contains the wave vector components k_y/k_0 . Finally the terms $\tilde{\mathbf{C}}_{Bi,m}^{\pm h}$ and $\tilde{\mathbf{C}}_{Bi,m}^{\pm e}$ are collected in the vectors $\tilde{\mathbf{c}}_{Bi}^{\pm h}$ and $\tilde{\mathbf{c}}_{Bi}^{\pm e}$. Eliminating the truncated expansion coefficients

\mathbf{c}_{B0}^- and \mathbf{c}_{BK+1}^+ of the scattered field from (3.45a) and (3.45c) gives

$$\tilde{\mathbf{G}}_{B1}^{1:} \tilde{\mathcal{F}}_{B1}(Z_0) \tilde{\mathbf{c}}_{B1} = \tilde{\mathbf{d}}_B^1, \quad (3.50a)$$

$$\tilde{\mathcal{F}}_{Bi}(Z_i) \tilde{\mathbf{c}}_{Bi} = \tilde{\mathcal{F}}_{Bi+1}(Z_i) \tilde{\mathbf{c}}_{Bi+1}, \quad (3.50b)$$

$$\tilde{\mathbf{G}}_{BK}^{2:} \tilde{\mathcal{F}}_{BK}(Z_K) \tilde{\mathbf{c}}_{BK} = \tilde{\mathbf{d}}_B^2, \quad (3.50c)$$

with the auxiliary matrices and vectors

$$\tilde{\mathbf{G}}_{B1} = \begin{bmatrix} \mathbf{I} & -\mathbf{Q}_{B0}(\mathcal{K}_y^2 - n_I^2 \mathbf{I})^{-1}(\mathcal{K}_x \mathcal{K}_y - n_I \mathcal{K}_{I,z} \mathcal{J}_I) \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (3.51a)$$

$$\tilde{\mathbf{G}}_{BK} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{I} & (\mathcal{K}_y^2 - n_{II}^2 \mathbf{I})^{-1}(\mathcal{K}_x \mathcal{K}_y + n_{II} \mathcal{K}_{II,z} \mathcal{J}_{II}) \mathbf{Q}_{BK}^{-H} \end{bmatrix}, \quad (3.51b)$$

$$\tilde{\mathbf{d}}_B = \begin{bmatrix} -2\mathbf{Q}_{B0}(\mathcal{K}_y^2 - n_I^2 \mathbf{I})^{-1}(n_I \mathcal{F}_{I,c} \mathcal{K}_{I,z} + \mathcal{F}_{I,s} \mathcal{K}_{I,z}^2 \mathcal{J}) \mathbf{d}_0 \\ \mathbf{0} \end{bmatrix}, \quad (3.51c)$$

and

$$\mathcal{K}_x = \text{diag}(\mathbf{K}_x, \mathbf{K}_x), \quad \mathcal{J}_\star = \text{a-diag}(-jn_\star \mathbf{I}, -\frac{1}{jn_\star} \mathbf{I}), \quad (3.51d)$$

where $\star = \text{I, II}$. Like \mathbf{K}_y the diagonal matrix \mathbf{K}_x now contains the wave vector components k_{xm}/k_0 . Equation (3.50) can be combined into one system where the coefficient matrix is similar to the sparse block matrix in the planar diffraction case (3.38), except that all matrices and vectors need to be replaced with their conical counterparts provided with a tilde. As already mentioned we will discuss the stable solution algorithm of this system in Chapter 5.

3.4 Generalisation of the Bloch mode method

As mentioned before the main challenge with the Bloch mode method is finding the roots of the transcendental equation. In the previous sections we have shown how to solve this problem for a lossless symmetric grating with only two transitions (symmetrically positioned around the origin) in a planar and conical diffraction mount. For a general asymmetric grating with lossy materials or with more transitions it is much more difficult to use the Bloch mode method. For this reason we will have a look at a different discretisation strategy in the next chapter. However, we first give a short overview of some attempts to generalise the Bloch mode method for these more complicated situations.

The first generalisation deals with lossy materials for a symmetric grating with only two transitions in a planar diffraction mount. In [10,11] the general framework is given which looks very similar to the lossless case. The main differences are that the under-

lying operator is no longer self-adjoint and that therefore a secondary adjoint problem is solved before proceeding with the method of moments. Moreover the transcendental equation no longer has real-valued but complex-valued roots which requires a specialised root-finding algorithm. More information on two different approaches of such a root-finding algorithm can be found in [8,34]. The first approach uses Cauchy's integral formula to determine the number of roots lying within an annulus in the complex plane. If this number is sufficiently small, Cauchy's integral and Newton's formulae are used to construct a polynomial having the same zeros as the transcendental equation within the region. The zeros of the low-order polynomial are evaluated by Müller's method and are refined iteratively by the same method. If the transcendental equation has a large number of zeros within the region, a bisection method is used along the radial and if necessary in the angular direction. The second approach is based on a continuation principle. This means that you start by computing the roots of a simpler problem and then gradually change the parameters of the simpler problem to the original more difficult problem while tracking the roots. A possible starting point could simply be a homogeneous layer of which the eigenvalues are given by k_{xm}^2/k_0^2 in the Rayleigh expansions of Section 2.4. Another possibility is to start with the original symmetric grating but with only the real part of the refraction indices. We can then use the techniques of the previous sections to compute the roots and gradually add the imaginary part of the refraction indices. Finally the extension to the conical diffraction mount with lossy materials is discussed in [24].

The second generalisation deals with more transitions inside one layer. In order to keep the notation simple let us assume a one layer grating and therefore drop the index i . The offsets are then given by $X_{(L+2l-1)/2} = -X_{(L-2l+1)/2} =: \tilde{X}_l$ for $l = 1, \dots, \frac{L+1}{2}$ and $\tilde{X}_0 = 0$. The real-valued refraction indices in each of the intervals $\tilde{X}_{l-1} < x < \tilde{X}_l$ are given by n_l with corresponding weight function w_l . Under these assumptions the two linearly independent solutions of the problem in the periodic x -direction are

$$a(x) = \begin{cases} a_1 = \cos(\kappa_1 x), & |x| \leq \tilde{X}_1, \\ a_{l+1} = a_l(\tilde{X}_l) \cos(\kappa_{l+1}(|x| - \tilde{X}_l)) + \\ \quad \frac{w_l}{w_{l+1}} \frac{1}{\kappa_{l+1}} a'_l(\tilde{X}_l) \sin(\kappa_{l+1}(|x| - \tilde{X}_l)), & \tilde{X}_l \leq |x| \leq \tilde{X}_{l+1}, \end{cases} \quad (3.52a)$$

and

$$b(x) = \begin{cases} b_1 = \frac{1}{\kappa_1} \sin(\kappa_1 x), & |x| \leq \tilde{X}_1, \\ b_{l+1} = \text{sign}(x) \left(b_l(\tilde{X}_l) \cos(\kappa_{l+1}(|x| - \tilde{X}_l)) + \right. \\ \quad \left. \frac{w_l}{w_{l+1}} \frac{1}{\kappa_{l+1}} b'_l(\tilde{X}_l) \sin(\kappa_{l+1}(|x| - \tilde{X}_l)) \right), & \tilde{X}_l \leq |x| \leq \tilde{X}_{l+1}, \end{cases} \quad (3.52b)$$

where

$$\kappa_1^2 = k_0^2 n_1^2 + k_0 \mu_i^2, \quad \kappa_{l+1}^2 = \kappa_1^2 + \Delta \kappa_{l+1}^2, \quad \Delta \kappa_{l+1}^2 = k_0^2 n_{l+1}^2 - k_0^2 n_1^2. \quad (3.52c)$$

Combining these results with $\Delta \tilde{X}_l = \tilde{X}_l - \tilde{X}_{l-1}$ gives the following relation

$$\begin{bmatrix} a(\frac{\Lambda}{2}) & a'(\frac{\Lambda}{2}) \\ b(\frac{\Lambda}{2}) & b'(\frac{\Lambda}{2}) \end{bmatrix} = \begin{bmatrix} a_{\frac{l+1}{2}}(\tilde{X}_{\frac{l+1}{2}}) & a'_{\frac{l+1}{2}}(\tilde{X}_{\frac{l+1}{2}}) \\ b_{\frac{l+1}{2}}(\tilde{X}_{\frac{l+1}{2}}) & b'_{\frac{l+1}{2}}(\tilde{X}_{\frac{l+1}{2}}) \end{bmatrix} = K_1 \prod_{l=2}^{\frac{l+1}{2}} W_l K_l, \quad (3.53a)$$

$$K_l = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\kappa_l} \end{bmatrix} \begin{bmatrix} \cos(\kappa_l \Delta \tilde{X}_l) & -\sin(\kappa_l \Delta \tilde{X}_l) \\ \sin(\kappa_l \Delta \tilde{X}_l) & \cos(\kappa_l \Delta \tilde{X}_l) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \kappa_l \end{bmatrix}, \quad (3.53b)$$

$$W_l = \begin{bmatrix} 1 & 0 \\ 0 & \frac{w_{l-1}}{w_l} \end{bmatrix}. \quad (3.53c)$$

These relations are necessary when writing out the transcendental equation which is again given by (3.8). Although we still have the separation of the eigenvalues for a Littrow mount, it is no longer obvious whether we can derive an interval based on the asymptotes of a monotonically increasing tangent or decreasing cotangent formulation that guarantees exactly one root. For a dielectric asymmetric grating it is possible to derive the Bloch modes in a similar fashion although there we would lose all symmetry in the transcendental equation. Especially the separation of the eigenvalues as discussed before is no longer possible. Finding the roots in a Littrow mount might prove very challenging in this case.

Chapter 4

Discretisation with RCWA

In this chapter we discuss another discretisation strategy for Maxwell's equations with the interface, pseudo-periodic and radiation boundary conditions for the three diffraction cases. This discretisation technique is known as the *Rigorous Coupled-Wave Analysis* (RCWA) and is also based on a mode expansion method [29, 30]. Essentially the electromagnetic fields in all layers are expanded into the same set of basis functions. These basis functions are identical to the eigenfunctions of the superstrate and substrate as derived in Section 2.4. Therefore in Section 4.1 the solution of the field in the grating layers is derived for the various diffraction cases. Then in Section 4.2 the layers are connected through the interface boundary condition which again results in a linear system of equations. Since the structure of this system is very similar to the Bloch mode method, we again postpone the actual solving part until Chapter 5.

4.1 Derivation RCWA modes inside grating structure

Since RCWA is also a mode expansion method we can write the solution outside the grating layers in terms of the Rayleigh expansion as was derived in Section 2.4. Contrary to the Bloch mode method we do not write the solution inside the grating layers as a linear combination of eigenfunctions of that layer. Instead RCWA expands the electromagnetic field inside the grating layers in terms of the complete set of orthonormal eigenfunctions of the superstrate and substrate layer, the pseudo-periodic Fourier series (2.46). Moreover the periodic permittivity and reciprocal permittivity function are

expanded into a standard Fourier series for $i = 1, \dots, K$

$$\epsilon_i^r = \sum_{m=-\infty}^{\infty} \epsilon_{i,m} e^{j2\pi mx/\Lambda}, \quad \epsilon_{i,m} = \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} \epsilon_i^r e^{-j2\pi mx/\Lambda} dx, \quad (4.1a)$$

$$\frac{1}{\epsilon_i^r} = \sum_{m=-\infty}^{\infty} \pi_{i,m} e^{j2\pi mx/\Lambda}, \quad \pi_{i,m} = \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} \frac{1}{\epsilon_i^r} e^{-j2\pi mx/\Lambda} dx. \quad (4.1b)$$

Because all functions in each layer are expanded into the same set of basis function, the matching at the layer interfaces is much simpler and no projection matrices need to be computed. Also since no complicated transcendental equation needs to be solved it is possible to deal with a wider range of geometries without any extra difficulty. In this chapter we therefore do not restrict ourselves to a symmetric lossless dielectric grating but look at general multi-layered gratings with multiple transitions inside a layer and complex-valued refraction indices. However, since we essentially approximate the Bloch modes with a Fourier series and Fourier discretise the piecewise constant permittivity function we do not expect RCWA to perform as well as the Bloch mode method. This performance difference and convergence discussion will be postponed until Section 7.1.

4.1.1 Planar diffraction

We start by expanding the electromagnetic field in pseudo-periodic Fourier modes. Also now we will solve the problem in the layers $i = 1, \dots, K$ for both polarisations in parallel by introducing a field F_{Ri} which represents the y -component of either the electric or magnetic field for TE and TM polarisation respectively

$$F_{Ri}(x, z) = f \sum_{m=-\infty}^{\infty} u_{Ri,m}(x) v_{Ri,m}(z) = f \sum_{m=-\infty}^{\infty} v_{Ri,m} e^{-jk_{xm}x}, \quad (4.2)$$

where $f = 1$ for TE polarisation and $f = -jY$ for TM polarisation. Contrary to the Bloch mode method where both polarisations could be treated the same way, here we have to pay some special attention to the TM polarisation case. This means that, when needed, an extra superscript is introduced to distinguish between the y -component of the electric and magnetic field. Substituting the (pseudo-periodic) Fourier series into (2.25) and (2.30) and slightly rearranging the terms gives

$$\sum_{m=-\infty}^{\infty} \left(\frac{1}{k_0^2} v_{Ri,m}^{e''} - \frac{k_{xm}^2}{k_0^2} v_{Ri,m}^e + \sum_{p=-\infty}^{\infty} \epsilon_{i,m-p} v_{Ri,p}^e \right) e^{-jk_{xm}x} = 0, \quad (4.3a)$$

$$\sum_{m=-\infty}^{\infty} \left(\frac{1}{k_0^2} v_{Ri,m}^{h''} - \sum_{p=-\infty}^{\infty} \epsilon_{i,m-p} \left(\frac{k_{xp}}{k_0} \sum_{q=-\infty}^{\infty} \pi_{i,p-q} \frac{k_{xq}}{k_0} v_{Ri,q}^h - v_{Ri,p}^h \right) \right) e^{-jk_{xm}x} = 0. \quad (4.3b)$$

In the derivation of (4.3a) we used the relation $k_{xm} - 2\pi q/\Lambda = k_{x(m+q)}$ in the product of the exponentials coming from the electric field and permittivity expansions where in the

latter the index of summation was changed to q . Moreover, we made the substitution for the index of summation so that $p = m + q$ before arriving at (4.3a). Similar steps can be taken in the derivation of (4.3b) for the magnetic field. Since the pseudo-periodic Fourier series forms a complete orthonormal basis, all the coefficients must be zero for $m \in \mathbb{Z}$ resulting in the following two sequences of equations

$$\frac{1}{k_0^2} v_{Ri,m}^{e''} = \frac{k_{xm}^2}{k_0^2} v_{Ri,m}^e - \sum_{p=-\infty}^{\infty} \epsilon_{i,m-p} v_{Ri,p}^e, \quad (4.4a)$$

$$\frac{1}{k_0^2} v_{Ri,m}^{h''} = \sum_{p=-\infty}^{\infty} \epsilon_{i,m-p} \left(\frac{k_{xp}}{k_0} \sum_{q=-\infty}^{\infty} \pi_{i,p-q} \frac{k_{xq}}{k_0} v_{Ri,q}^h - v_{Ri,p}^h \right). \quad (4.4b)$$

An approximate solution for the electromagnetic field in the grating layers is derived by truncating the equations and solving a discrete eigenvalue problem. More specifically, truncating the two sequences of equations (4.4) and the involved series for $-M \leq m, p, q \leq M$ results in two second-order ordinary differential equations (ODEs) for $v_{Ri,m}$ collected in the vector \mathbf{v}_{Ri}

$$\frac{1}{k_0^2} \mathbf{v}_{Ri}^{e''} = (\mathbf{K}_x^2 - \mathbf{E}_i) \mathbf{v}_{Ri}^e =: \mathbf{A}_i \mathbf{v}_{Ri}^e, \quad (4.5a)$$

$$\frac{1}{k_0^2} \mathbf{v}_{Ri}^{h''} = \mathbf{E}_i (\mathbf{K}_x \mathbf{P}_i \mathbf{K}_x - \mathbf{I}) \mathbf{v}_{Ri}^h =: \mathbf{E}_i \mathbf{C}_i \mathbf{v}_{Ri}^h, \quad (4.5b)$$

where the Fourier coefficients $\epsilon_{i,m-p}$ and $\pi_{i,m-p}$ are collected in the Toeplitz matrices \mathbf{E}_i and \mathbf{P}_i . In the derivation of (4.5b) we used *Laurent's multiplication rule* throughout. Unfortunately this rule does not uniformly preserve the continuity of the appropriate electromagnetic field components across the discontinuities of the permittivity function. For example, in (4.4b) the product of the permittivity with the term within brackets comes from Fourier transforming (2.29b). Across a discontinuity of the permittivity function we know from the inner layer interface boundary conditions (3.3) that $H_{i,y}$ is continuous and also its derivative with respect to z is continuous. This means that $E_{i,x}$ and therefore also the term within the brackets of (4.4b) is discontinuous across this discontinuity and has a jump complementary to the jump in the permittivity. For these special cases where the product of two functions with concurrent complementary jumps is approximated with a finite number of Fourier modes an *inverse multiplication rule* was derived in [26, 27]. This rule does indeed uniformly preserve the continuity of the appropriate electromagnetic field components. From an implementation point of view the rule states that whenever the permittivity function (reciprocal permittivity function) is encountered in such a product with complementary jumps, the corresponding Toeplitz matrix is replaced by the inverse of the Toeplitz matrix corresponding to the reciprocal permittivity function (permittivity function). The easiest way to recognise these special products is by looking at the coupled PDE system in (2.29) before eliminating any of the field components. Keeping this in mind the following set of modified second-order

ODEs can be derived

$$\frac{1}{k_0^2} \mathbf{v}_{Ri}^{e''} = \mathbf{A}_i \mathbf{v}_{Ri}^e, \quad (4.6a)$$

$$\frac{1}{k_0^2} \mathbf{v}_{Ri}^{h''} = \mathbf{P}_i^{-1} (\mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{K}_x - \mathbf{I}) \mathbf{v}_{Ri}^h =: \mathbf{P}_i^{-1} \mathbf{B}_i \mathbf{v}_{Ri}^h. \quad (4.6b)$$

We assume the constant matrices in (4.6) to be diagonalisable: $\mathbf{A}_i = \mathbf{Q}_{Ri}^e \mathbf{M}_{Ri}^{e2} \mathbf{Q}_{Ri}^{e-1}$ and $\mathbf{P}_i^{-1} \mathbf{B}_i = \mathbf{Q}_{Ri}^h \mathbf{M}_{Ri}^{h2} \mathbf{Q}_{Ri}^{h-1}$ with \mathbf{Q}_{Ri} a normalised matrix containing the eigenvectors and \mathbf{M}_{Ri} a diagonal matrix containing the square root of the eigenvalues where the branch cut is chosen along the negative real axis (so all values have a non-negative real part). For dielectric layers where the refraction index is a piecewise constant real-valued function it can be shown that the matrix \mathbf{E}_i and therefore also \mathbf{A}_i is Hermitian in the TE polarisation case. It is well known that a Hermitian and thus normal matrix is unitarily diagonalisable and has real-valued eigenvalues: $\mathbf{A}_i = \mathbf{Q}_{Ri}^e \mathbf{M}_{Ri}^{e2} \mathbf{Q}_{Ri}^{eH}$. Moreover the inverse of the two Hermitian Toeplitz matrices \mathbf{E}_i and \mathbf{P}_i is again Hermitian but in general not Toeplitz anymore. Contrary to the planar diffraction case with TE polarisation, the matrix $\mathbf{P}_i^{-1} \mathbf{B}_i$ for TM polarisation is no longer Hermitian nor normal. Instead it is actually the product of two Hermitian matrices and from this one can show that the eigenvalues are still real-valued. Moreover we would like to point out that the inverse of the Toeplitz matrices \mathbf{E}_i and \mathbf{P}_i do exist since the spectrum of both matrices lies in between the minimum and maximum value of the strictly positive permittivity and reciprocal permittivity function. The general solution of (4.6) in terms of the eigenvalues and eigenvectors is given by

$$\mathbf{v}_{Ri} = \mathbf{Q}_{Ri} \left(e^{k_0 \mathbf{M}_{Ri}(z-Z_i)} \mathbf{c}_{Ri}^- + e^{-k_0 \mathbf{M}_{Ri}(z-Z_{i-1})} \mathbf{c}_{Ri}^+ \right), \quad (4.7)$$

where the vectors \mathbf{c}_{Ri}^- and \mathbf{c}_{Ri}^+ contain the unknown expansion coefficients $C_{Ri,m}^-$ and $C_{Ri,m}^+$ respectively. The approximate solution for the electric field in a grating layer is then given by

$$\begin{aligned} F_{Ri} &= f \sum_{m=-M}^M v_{Ri,m} e^{-jk_{xm}x} \\ &= f \sum_{m=-M}^M e^{-jk_{xm}x} \sum_{s=-M}^M q_{Ri,ms} \left(C_{Ri,s}^- e^{k_0 \mu_{Ri,s}(z-Z_i)} + C_{Ri,s}^+ e^{-k_0 \mu_{Ri,s}(z-Z_{i-1})} \right) \\ &= f \sum_{s=-M}^M \left(\sum_{m=-M}^M q_{Ri,ms} e^{-jk_{xm}x} \right) \left(C_{Ri,s}^- e^{k_0 \mu_{Ri,s}(z-Z_i)} + C_{Ri,s}^+ e^{-k_0 \mu_{Ri,s}(z-Z_{i-1})} \right). \end{aligned} \quad (4.8)$$

When comparing the RCWA solution in (4.8) with the Bloch solution in (3.12) for dielectric layers the following three observations are important. First we observe that the Bloch solution in (3.12) gives the exact solution of the electromagnetic field in terms of the exact eigenvalues and eigenfunctions of the underlying problem. Of course when

the layers are matched and the linear system of equations is derived we still have to truncate the expansions. On the other hand the RCWA solution in (4.8) is already an approximation of the electromagnetic field (since the series is already truncated) even before matching the layers. The RCWA eigenvalues $\mu_{Ri,m}^2$ only approximate the real Bloch eigenvalues $\mu_{Bi,m}^2$. Moreover the eigenvectors contain the approximate Fourier coefficients of the exact eigenfunctions projected on a finite number of pseudo-periodic Fourier modes. In particular the last line of (4.8) shows the approximate eigenfunction within the first set of brackets and also the z -dependent part within the second set of brackets, both can be compared with $u_{Bi,m}$ and $v_{Bi,m}$ respectively in (3.12). Secondly we observe that for TE polarisation the approximate eigenfunctions are orthonormal with respect to the inner product given in (3.10b) where $w_i = 1$. This follows directly from the fact that the matrix \mathbf{Q}_{Ri}^e containing the eigenvectors is unitary. Similarly for TM polarisation the approximate eigenfunctions are orthogonal with respect to the inner product given in (3.10b) but now with $w_i = \sum_{m=-\infty}^{\infty} \pi_{i,m} e^{j2\pi mx/\Lambda}$. This is because the matrix \mathbf{Q}_{Ri}^h containing the eigenvectors is no longer unitary but does satisfy the relation $\mathbf{Q}_{Ri}^{hH} \mathbf{P}_i \mathbf{Q}_{Ri}^h = \mathbf{D}_i$ a diagonal matrix. So like the real Bloch modes also the approximate Bloch modes are orthonormal (orthogonal) with respect to the standard (weighed) inner product for TE (TM) polarisation. Thirdly we observe that all remarks at the end of Section 3.1.1 concerning the solution component in the z -direction representing either propagating or evanescent waves are also valid here.

4.1.2 Conical diffraction

The conical diffraction case consists of two separate problems resembling the two independent polarisations of the planar diffraction case. Therefore we will pay special attention to the differences with the planar diffraction case as well as the similarities with the conical Bloch mode method. As in Section 4.1.1 we start by expanding the electromagnetic field in pseudo-periodic Fourier modes. In the layers $i = 1, \dots, K$ a field \tilde{F}_{Ri} is introduced which represents the x -component of either the magnetic or electric field respectively

$$\tilde{F}_{Ri}(x, z) = \tilde{f} \sum_{m=-\infty}^{\infty} \tilde{u}_{Ri,m}(x) \tilde{v}_{Ri,m}(z) = \tilde{f} \sum_{m=-\infty}^{\infty} \tilde{v}_{Ri,m} e^{-jk_{xm}x}, \quad (4.9)$$

where $\tilde{f} = -jY$ for the magnetic field and $\tilde{f} = 1$ for the electric field. Like in the Bloch mode method we purposely swapped the order of the magnetic and electric field in order to retain some symmetry later on. After substituting the (pseudo-periodic) Fourier series into (2.35) and rearranging the terms, the same techniques as in the planar diffraction can be applied to simplify the expressions. These techniques include the shift relation for the wave vector as well as changing the order of summation. In the simplified expressions again all the coefficients must be zero because the pseudo-periodic Fourier series forms a complete orthonormal basis. This results in the following two sequences

of equations for $i = 1, \dots, K$ and $m \in \mathbb{Z}$

$$\frac{1}{k_0^2} \tilde{v}_{Ri,m}^{h''} = \frac{k_y^2}{k_0^2} \tilde{v}_{Ri,m}^h + \frac{k_{xm}^2}{k_0^2} \tilde{v}_{Ri,m}^h - \sum_{p=-\infty}^{\infty} \epsilon_{i,m-p} \tilde{v}_{Ri,p}^h \quad (4.10a)$$

$$\begin{aligned} \frac{1}{k_0^2} \tilde{v}_{Ri,m}^{e''} &= \frac{k_y^2}{k_0^2} \tilde{v}_{Ri,m}^e + \frac{k_{xm}}{k_0} \sum_{p=-\infty}^{\infty} \pi_{i,m-p} \frac{k_{xp}}{k_0} \sum_{q=-\infty}^{\infty} \epsilon_{i,p-q} \tilde{v}_{Ri,q}^e - \\ &\quad \sum_{p=-\infty}^{\infty} \epsilon_{i,m-p} \tilde{v}_{Ri,p}^e. \end{aligned} \quad (4.10b)$$

Note that in the conical diffraction case, like in the Bloch mode method, the equation for the x -component of the magnetic (electric) field resembles the equation for the y -component of the electric (magnetic) field in the planar diffraction case with TE (TM) polarisation.

Two second-order ODEs for $\tilde{v}_{Ri,m}$ are derived by truncating the two sequences of equations (4.10) and the involved series for $-M \leq m, p, q \leq M$

$$\frac{1}{k_0^2} \tilde{\mathbf{v}}_{Ri}^{h''} = (\mathbf{K}_y^2 + \mathbf{K}_x^2 - \mathbf{E}_i) \tilde{\mathbf{v}}_{Ri}^h = (\mathbf{K}_y^2 + \mathbf{A}_i) \tilde{\mathbf{v}}_{Ri}^h, \quad (4.11a)$$

$$\frac{1}{k_0^2} \tilde{\mathbf{v}}_{Ri}^{e''} = \left(\mathbf{K}_y^2 + (\mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{K}_x - \mathbf{I}) \mathbf{P}_i^{-1} \right) \tilde{\mathbf{v}}_{Ri}^e = (\mathbf{K}_y^2 + \mathbf{B}_i \mathbf{P}_i^{-1}) \tilde{\mathbf{v}}_{Ri}^e, \quad (4.11b)$$

where the matrices in these expressions have already been introduced in Section 4.1.1 on planar RCWA diffraction and at the end of Section 3.1.2 on conical Bloch diffraction. Note that in (4.11b) the inverse multiplication rule has been applied instead of Laurent's multiplication rule. Similarly as for TM polarisation the resulting equation uniformly preserves the continuity of the appropriate electromagnetic field components across the discontinuities of the permittivity function. Again we assume the matrices in (4.11) to be diagonalisable: $\mathbf{K}_y^2 + \mathbf{A}_i = \tilde{\mathbf{Q}}_{Ri}^h \tilde{\mathbf{M}}_{Ri}^{h2} \tilde{\mathbf{Q}}_{Ri}^{h-1}$ and $\mathbf{K}_y^2 + \mathbf{B}_i \mathbf{P}_i^{-1} = \tilde{\mathbf{Q}}_{Ri}^e \tilde{\mathbf{M}}_{Ri}^{e2} \tilde{\mathbf{Q}}_{Ri}^{e-1}$ with $\tilde{\mathbf{Q}}_{Ri}$ normalised matrices containing the eigenvectors and $\tilde{\mathbf{M}}_{Ri}$ diagonal matrices containing the square root of the eigenvalues. Comparing the two eigenvalue problems of the conical diffraction case with those of the planar diffraction case, it is clear they are related when the matrices \mathbf{A}_i and \mathbf{B}_i of both diffraction cases are the same. The Toeplitz matrices and their inverse only contain the Fourier coefficients of the geometry and therefore do not change when going from the planar to the conical diffraction case. This means that the matrix \mathbf{K}_x needs to be the same for both cases which happens for all conical angles of incidence that have the same k_x as the corresponding planar angle of incidence. Since \mathbf{K}_y is actually just a scaled identity matrix, the eigenvalues of the conical diffraction case are simply shifted from the corresponding planar diffraction case. Moreover, the eigenvectors are then related through $\tilde{\mathbf{Q}}_{Ri}^h = \mathbf{Q}_{Ri}^h$ for TE polarisation and $\tilde{\mathbf{Q}}_{Ri}^e = \mathbf{P}_i \mathbf{Q}_{Ri}^e$ for TM polarisation. Recall that in Section 3.1.2 on conical Bloch a similar comment was made on the transcendental eigenvalue problem and corresponding eigenvalues and eigenfunctions. For dielectric gratings the diagonalisation step is again guaranteed using the

same arguments as in the planar diffraction case. The general solution of (4.11) in terms of the eigenvalues and eigenvectors is given by

$$\tilde{\mathbf{v}}_{Ri} = \tilde{\mathbf{Q}}_{Ri} \left(e^{k_0 \tilde{\mathbf{M}}_{Ri}(z-Z_i)} \tilde{\mathbf{c}}_{Ri}^- + e^{-k_0 \tilde{\mathbf{M}}_{Ri}(z-Z_{i-1})} \tilde{\mathbf{c}}_{Ri}^+ \right), \quad (4.12)$$

where the vectors $\tilde{\mathbf{c}}_{Ri}^-$ and $\tilde{\mathbf{c}}_{Ri}^+$ contain the unknown expansion coefficients $\tilde{C}_{Ri,m}^-$ and $\tilde{C}_{Ri,m}^+$ respectively. The approximate solution for the electromagnetic field in a grating layer is then given by

$$\begin{aligned} \tilde{F}_{Ri} &= \tilde{f} \sum_{m=-M}^M \tilde{v}_{Ri,m} e^{-jk_{xm}x} \\ &= \tilde{f} \sum_{p=-M}^M \left(\sum_{m=-M}^M \tilde{q}_{Ri,mp} e^{-jk_{xm}x} \right) \left(\tilde{C}_{Ri,p}^- e^{k_0 \tilde{\mu}_{Ri,p}(z-Z_i)} + \tilde{C}_{Ri,p}^+ e^{-k_0 \tilde{\mu}_{Ri,p}(z-Z_{i-1})} \right). \end{aligned} \quad (4.13)$$

All three observations from the end of Section 4.1.1 on the planar diffraction case carry over to the above conical diffraction case. However, having found the relations between the planar and conical eigenvectors after (4.11), the orthogonality relations in the second observation need to be changed accordingly.

4.2 Matching at the interfaces

The electromagnetic fields in the superstrate and substrate are described by the Rayleigh expansion, whereas the fields in the grating layers are described by the RCWA mode expansion. This section applies interface boundary conditions in order to get the necessary relations between the unknown expansion coefficients that still appear in these expansions. Since these expansions are all based on the pseudo-periodic Fourier series no projections or coupling matrices are required and the linear combinations can be matched coefficient-wise.

4.2.1 Planar diffraction

For the planar diffraction case the necessary equations for the unknown expansion coefficients follow from the interface boundary conditions (2.27) or (2.32) depending on the polarisation. From these equations that were also used in the Bloch mode method it can be seen that only the y -component of the electromagnetic fields is required. The solution in the superstrate and substrate is again given by the y -component of the Raleigh expansions in (2.49) and (2.50) for TE and TM polarisation respectively. This means that the left-hand side of (3.32a,b) and right-hand side of (3.32e,f) and all the variables

appearing there largely remain unchanged. However the subscript in the expansion coefficients differs and also the sums are already truncated. Additionally the solution inside the grating layers is now given by the RCWA mode expansion in (4.7) for both polarisations. Summarising we get for $z = Z_0$

$$\nu_I \sum_{m=-M}^M (C_{R0,m}^- + \delta_{0,m}) e^{-jk_{xm}x} = \sum_{m=-M}^M v_{R1,m}(0) e^{-jk_{xm}x}, \quad (4.14a)$$

$$j\omega_0 \nu_I \sum_{m=-M}^M \frac{k_{L,zm}}{k_0} (C_{R0,m}^- - \delta_{0,m}) e^{-jk_{xm}x} = \frac{1}{k_0} \sum_{m=-M}^M \left(\sum_{p=-M}^M \omega_{1,m-p} v'_{R1,p}(0) \right) e^{-jk_{xm}x}, \quad (4.14b)$$

for $z = Z_i$ and $i = 1, \dots, K-1$

$$\sum_{m=-M}^M v_{Ri,m}(Z_i) e^{-jk_{xm}x} = \sum_{m=-M}^M v_{Ri+1,m}(Z_i) e^{-jk_{xm}x}, \quad (4.14c)$$

$$\frac{1}{k_0} \sum_{m=-M}^M \left(\sum_{p=-M}^M \omega_{i,m-p} v'_{Ri,p}(Z_i) \right) e^{-jk_{xm}x} = \frac{1}{k_0} \sum_{m=-M}^M \left(\sum_{p=-M}^M \omega_{i+1,m-p} v'_{Ri+1,p}(Z_i) \right) e^{-jk_{xm}x}, \quad (4.14d)$$

and for $z = Z_K$

$$\sum_{m=-M}^M v_{RK,m}(Z_K) e^{-jk_{xm}x} = \nu_{II} \sum_{m=-M}^M C_{RK+1,m}^+ e^{-jk_{xm}x}, \quad (4.14e)$$

$$\frac{1}{k_0} \sum_{m=-M}^M \left(\sum_{p=-M}^M \omega_{K,m-p} v'_{RK,p}(Z_K) \right) e^{-jk_{xm}x} = -j\omega_{K+1} \nu_{II} \sum_{m=-M}^M \frac{k_{L,zm}}{k_0} C_{RK+1,m}^+ e^{-jk_{xm}x}. \quad (4.14f)$$

where we also Fourier transformed the weight function for $i = 1, \dots, K$

$$w_i = \sum_{m=-\infty}^{\infty} \omega_{i,m} e^{j2\pi m x / \Lambda}, \quad \omega_{i,m} = \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} w_i e^{-j2\pi m x / \Lambda} dx. \quad (4.15)$$

Note that the relation $k_{xm} - 2\pi q / \Lambda = k_{x(m+q)}$ is used once more in the product of the exponentials coming from the expansions of the electromagnetic field and weight function before arriving at (4.14). Recall that in the case of TE polarisation the weight function is equal to one and thus the Fourier coefficients are given by $\omega_{i,m} = \delta_{0,m}$. This means that (4.14b,d,f) simplify dramatically where the term within brackets in the grating layers then becomes $v'_{Ri,m}$ evaluated at the correct height. In the case of TM polarisation the weight function is equal to the piecewise constant reciprocal permittivity function

and thus the Fourier components are given by $\omega_{i,m} = \pi_{i,m}$. Naturally the interface boundary conditions do not simplify any further now. Similar as with the Bloch mode method the coefficients $C_{R0,m}^-$ represent either the expansion coefficients R_{ym}^e or R_{ym}^h depending on the polarisation. A similar remark can be made for the coefficients $C_{RK+1,m}^+$ in the substrate. The system of equations (4.14) in known expansion functions (the pseudo-periodic Fourier series) and with unknown expansion coefficients can be solved by matching the terms in front of these expansion functions which are now the same for all layers. In the framework of the method of moments we are simply choosing the *test functions* equal to the expansion functions. Substituting the solution for $v_{Ri,m}$ results in the following set of equations for $-M \leq m \leq M$

$$\nu_I (C_{R0,m}^- + \delta_{0,m}) = \sum_{s=-M}^M q_{R1,ms} (C_{R1,s}^- x_{R1,s} + C_{R1,s}^+), \quad (4.16a)$$

$$j\omega_0 \nu_I \frac{k_{I,zm}}{k_0} (C_{R0,m}^- - \delta_{0,m}) = \sum_{p=-M}^M \omega_{1,m-p} \sum_{s=-M}^M q_{R1,ps} \mu_{R1,s} (C_{R1,s}^- x_{R1,s} - C_{R1,s}^+), \quad (4.16b)$$

for $i = 1, \dots, K-1$

$$\sum_{s=-M}^M q_{Ri,ms} (C_{Ri,s}^- + C_{Ri,s}^+ x_{Ri,s}) = \sum_{s=-M}^M q_{Ri+1,ms} (C_{Ri+1,s}^- x_{Ri+1,s} + C_{Ri+1,s}^+), \quad (4.16c)$$

$$\begin{aligned} \sum_{p=-M}^M \omega_{i,m-p} \sum_{s=-M}^M q_{Ri,ps} \mu_{Ri,s} (C_{Ri,s}^- - C_{Ri,s}^+ x_{Ri,s}) = \\ \sum_{p=-M}^M \omega_{i+1,m-p} \sum_{s=-M}^M q_{Ri+1,ps} \mu_{Ri+1,s} (C_{Ri+1,s}^- x_{Ri+1,s} - C_{Ri+1,s}^+), \end{aligned} \quad (4.16d)$$

and

$$\sum_{s=-M}^M q_{RK,ms} (C_{RK,s}^- + C_{RK,s}^+ x_{RK,s}) = \nu_{II} C_{RK+1,m}^+, \quad (4.16e)$$

$$\sum_{p=-M}^M \omega_{K,m-p} \sum_{s=-M}^M q_{RK,ps} \mu_{RK,s} (C_{RK,s}^- - C_{RK,s}^+ x_{RK,s}) = -j\omega_{K+1} \nu_{II} \frac{k_{II,zm}}{k_0} C_{RK+1,m}^+, \quad (4.16f)$$

where $x_{Ri,m} = e^{-k_0 \mu_{Ri,m} h_i}$. Compared to the Bloch mode method in (3.33) we see that here the algebraic linear system of equations for the unknown expansion coefficients is already truncated. Like in the planar Bloch mode method equations (4.16b,d,f) are multiplied with -1 which makes the comparison with other solution algorithms easier

later on. Summarising we get for $i = 1, \dots, K-1$

$$\nu_I \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ jw_0 \mathbf{K}_{I,z} & -jw_0 \mathbf{K}_{I,z} \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{c}_{R0}^- \end{bmatrix} = \mathcal{F}_{R1}(Z_0) \mathbf{c}_{R1}, \quad (4.17a)$$

$$\mathcal{F}_{Ri}(Z_i) \mathbf{c}_{Ri} = \mathcal{F}_{Ri+1}(Z_i) \mathbf{c}_{Ri+1}, \quad (4.17b)$$

$$\mathcal{F}_{RK}(Z_K) \mathbf{c}_{RK} = \nu_{II} \begin{bmatrix} \mathbf{I} \\ jw_{K+1} \mathbf{K}_{II,z} \end{bmatrix} \mathbf{c}_{RK+1}^+. \quad (4.17c)$$

Here all matrices corresponding to the superstrate and substrate layers have already been defined in Section 3.3.1. Also the vectors are defined there with the exception that the subscript referring to the method needs to be changed. Only the shorthand notation for the solution in the grating layers evaluated at the different interfaces is slightly different

$$\mathcal{F}_{Ri}(Z_{i-1}) = \begin{bmatrix} \mathbf{Q}_{Ri} & \mathbf{O} \\ \mathbf{O} & \mathbf{W}_i \mathbf{Q}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{M}_{Ri} & -\mathbf{M}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{X}_{Ri} \end{bmatrix}, \quad (4.18a)$$

$$\mathcal{F}_{Ri}(Z_i) = \begin{bmatrix} \mathbf{Q}_{Ri} & \mathbf{O} \\ \mathbf{O} & \mathbf{W}_i \mathbf{Q}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{M}_{Ri} & -\mathbf{M}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{Ri} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}. \quad (4.18b)$$

Similarly as in the previous chapter the solution in the grating layers \mathcal{F}_{Ri} is referred to as the *fundamental solution*. Again most matrices and vectors are already defined in the previous chapter after changing the subscript. The matrix \mathbf{W}_i on the other hand contains the Fourier coefficients of the weight function and is equal to \mathbf{I} or \mathbf{P}_i depending on the polarisation which is either TE or TM respectively. Eliminating the truncated expansion coefficients \mathbf{c}_{R0}^- and \mathbf{c}_{RK+1}^+ of the scattered field from (4.17a) and (4.17c) gives

$$\mathcal{G}_{R1}^{1:} \mathcal{F}_{R1}(Z_0) \mathbf{c}_{R1} = \mathbf{d}_R^1, \quad (4.19a)$$

$$\mathcal{F}_{Ri}(Z_i) \mathbf{c}_{Ri} = \mathcal{F}_{Ri+1}(Z_i) \mathbf{c}_{Ri+1}, \quad (4.19b)$$

$$\mathcal{G}_{RK}^{2:} \mathcal{F}_{RK}(Z_K) \mathbf{c}_{RK} = \mathbf{d}_R^2, \quad (4.19c)$$

with the auxiliary matrices and vectors

$$\mathcal{G}_{R1} = \begin{bmatrix} jw_0 \mathbf{K}_{I,z} & \mathbf{I} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad \mathcal{G}_{RK} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ jw_{K+1} \mathbf{K}_{II,z} & -\mathbf{I} \end{bmatrix}, \quad (4.20a)$$

$$\mathbf{d}_R = \begin{bmatrix} 2jw_0 \nu_I \mathbf{K}_{I,z} \mathbf{d}_0 \\ \mathbf{0} \end{bmatrix}. \quad (4.20b)$$

Like before equation (4.19) can be combined into one system where the coefficient matrix is a sparse block matrix. For the sake of completeness the linear system and coefficient matrix is given below but except for the subscript index referring to the method

this is essentially the same as in (3.39) for the planar Bloch mode method

$$\begin{bmatrix} \mathcal{F}_{R1}(Z_1) & -\mathcal{F}_{R2}(Z_1) & & & & \\ & & \ddots & & \ddots & \\ & & & & & \\ & & & \mathcal{F}_{RK-1}(Z_{K-1}) & -\mathcal{F}_{RK}(Z_{K-1}) & \\ \mathcal{G}_{R1}\mathcal{F}_{R1}(Z_0) & & & & & \mathcal{G}_{RK}\mathcal{F}_{RK}(Z_K) \end{bmatrix} \begin{bmatrix} c_{R1} \\ \vdots \\ c_{RK-1} \\ c_{RK} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ d_R \end{bmatrix}. \quad (4.21)$$

Like with Bloch this sparse linear system is referred to as a *discrete boundary value problem* where the last block row is said to contain the boundary conditions of the discrete BVP. Also now we will postpone the actual solving part of this system until Chapter 5.

4.2.2 Conical diffraction

For the conical diffraction case the necessary equations for the unknown expansion coefficients follow from the interface boundary conditions (2.36). From these equations that were also used in the Bloch mode method it can be seen that both the x - and y -component of the electric and magnetic field are required. The solution in the superstrate and substrate is again given by the x - and y -component of the Raleigh expansions in (2.49) and (2.50). So similar as in the planar diffraction case this means that the left-hand side of (3.42a-d) and right-hand side of (3.42i-l) and all the variables appearing there largely remain unchanged. Naturally the subscript in the expansion coefficients differs and also now the sums are already truncated. Additionally the solution for the x -component of the electric and magnetic field inside the grating layers is given by the RCWA mode expansion in (4.12). This means that inside the grating layers we still need to find an expression for the y -component of these fields before we can apply the interface boundary conditions. Identically as in the Bloch mode method these are derived from (2.34b) and (2.34d) which simplify to (3.40). However, in the context of RCWA it is easier to divide out the relative permittivity again from (3.40a) when referring to this equation. In order to write down the RCWA solution of (3.40) we first need to introduce a pseudo-periodic Fourier expansion for the y -components of the electromagnetic field. Then, after substituting all (pseudo-periodic) Fourier series into the equation, using the shift relation that was explained in the planar diffraction case and truncating the

equations it can be verified that the solution is given by

$$H_{i,y} = -jY \sum_{m=-M}^M \tilde{v}_{Ri,m}^h e^{-jk_{xm}x}, \quad \tilde{\mathbf{v}}_{Ri}^h = \mathbf{B}_i^{-1} (\mathbf{K}_y \mathbf{K}_x \mathbf{E}_i^{-1} \tilde{\mathbf{v}}_{Ri}^h + \frac{1}{k_0} \tilde{\mathbf{v}}_{Ri}^{e'}), \quad (4.22a)$$

$$E_{i,y} = \sum_{m=-M}^M \tilde{v}_{Ri,m}^e e^{-jk_{xm}x}, \quad \tilde{\mathbf{v}}_{Ri}^e = \mathbf{A}_i^{-1} \left(\frac{1}{k_0} \tilde{\mathbf{v}}_{Ri}^{h'} + \mathbf{K}_y \mathbf{K}_x \tilde{\mathbf{v}}_{Ri}^e \right), \quad (4.22b)$$

where in the first equation again the inverse multiplication rule is used instead of Laurent's multiplication rule as explained in Section 4.1. From these expressions we see that a problem arises when one of the matrices \mathbf{A}_i or \mathbf{B}_i is not invertible. This is exactly the same as saying that an eigenvalue of one of these matrices is equal to zero. But this in turn means that a conical eigenvalue is equal to k_y/k_0 and that the corresponding planar diffraction case has an eigenvalue equal to zero. As already mentioned in the chapter on Bloch where a denominator could become zero it is very unlikely that this will actually happen for a certain grating. In the remainder of this thesis we will therefore assume not to be in this special case.

Now that all components are available the interface boundary conditions (2.36) can be applied which after reordering give for $z = Z_0$

$$jn_1 \sum_{m=-M}^M (R_{xm}^h + \delta_{0,m}(s_x \cos \psi - p_x \sin \psi)) e^{-jk_{xm}x} = \sum_{m=-M}^M \tilde{v}_{R1,m}^h(0) e^{-jk_{xm}x}, \quad (4.23a)$$

$$\sum_{m=-M}^M (R_{xm}^e + \delta_{0,m}(s_x \sin \psi + p_x \cos \psi)) e^{-jk_{xm}x} = \sum_{m=-M}^M \tilde{v}_{R1,m}^e(0) e^{-jk_{xm}x}, \quad (4.23b)$$

$$jn_1 \sum_{m=-M}^M (R_{ym}^h + \delta_{0,m}(s_y \cos \psi - p_y \sin \psi)) e^{-jk_{xm}x} = \sum_{m=-M}^M \left(\sum_{p=-M}^M \hat{b}_{1,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \hat{e}_{1,pq} \tilde{v}_{R1,q}^h(0) + \frac{1}{k_0} \tilde{v}_{R1,p}^{e'}(0) \right) \right) e^{-jk_{xm}x}, \quad (4.23c)$$

$$\sum_{m=-M}^M (R_{ym}^e + \delta_{0,m}(s_y \sin \psi + p_y \cos \psi)) e^{-jk_{xm}x} = \sum_{m=-M}^M \left(\sum_{p=-M}^M \hat{a}_{1,mp} \left(\frac{1}{k_0} \tilde{v}_{R1,p}^{h'}(0) + \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \tilde{v}_{R1,p}^e(0) \right) \right) e^{-jk_{xm}x}, \quad (4.23d)$$

for $z = Z_i$ and $i = 1, \dots, K-1$

$$\sum_{m=-M}^M \tilde{v}_{Ri,m}^h(Z_i) e^{-jk_{xm}x} = \sum_{m=-M}^M \tilde{v}_{Ri+1,m}^h(Z_i) e^{-jk_{xm}x}, \quad (4.23e)$$

$$\sum_{m=-M}^M \tilde{v}_{Ri,m}^e(Z_i) e^{-jk_{xm}x} = \sum_{m=-M}^M \tilde{v}_{Ri+1,m}^e(Z_i) e^{-jk_{xm}x}, \quad (4.23f)$$

$$\sum_{m=-M}^M \left(\sum_{p=-M}^M \widehat{b}_{i,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \widehat{e}_{i,pq} \widehat{v}_{Ri,q}^h(Z_i) + \frac{1}{k_0} \widehat{v}_{Ri,p}^{e'}(Z_i) \right) \right) e^{-jk_{xm}x} =$$

$$\sum_{m=-M}^M \left(\sum_{p=-M}^M \widehat{b}_{i+1,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \widehat{e}_{i+1,pq} \widehat{v}_{Ri+1,q}^h(Z_i) + \frac{1}{k_0} \widehat{v}_{Ri+1,p}^{e'}(Z_i) \right) \right) e^{-jk_{xm}x}, \quad (4.23g)$$

$$\sum_{m=-M}^M \left(\sum_{p=-M}^M \widehat{a}_{i,mp} \left(\frac{1}{k_0} \widehat{v}_{Ri,p}^{h'}(Z_i) + \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \widehat{v}_{Ri,p}^e(Z_i) \right) \right) e^{-jk_{xm}x} =$$

$$\sum_{m=-M}^M \left(\sum_{p=-M}^M \widehat{a}_{i+1,mp} \left(\frac{1}{k_0} \widehat{v}_{Ri+1,p}^{h'}(Z_i) + \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \widehat{v}_{Ri+1,p}^e(Z_i) \right) \right) e^{-jk_{xm}x}, \quad (4.23h)$$

and for $z = Z_K$

$$\sum_{m=-M}^M \widehat{v}_{RK,m}^h(Z_K) e^{-jk_{xm}x} = jn_{II} \sum_{m=-M}^M T_{xm}^h e^{-jk_{xm}x}, \quad (4.23i)$$

$$\sum_{m=-M}^M \widehat{v}_{RK,m}^e(Z_K) e^{-jk_{xm}x} = \sum_{m=-M}^M T_{xm}^e e^{-jk_{xm}x}, \quad (4.23j)$$

$$\sum_{m=-M}^M \left(\sum_{p=-M}^M \widehat{b}_{K,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \widehat{e}_{K,pq} \widehat{v}_{RK,q}^h(Z_K) + \frac{1}{k_0} \widehat{v}_{RK,p}^{e'}(Z_K) \right) \right) e^{-jk_{xm}x} =$$

$$jn_{II} \sum_{m=-M}^M T_{ym}^h e^{-jk_{xm}x}, \quad (4.23k)$$

$$\sum_{m=-M}^M \left(\sum_{p=-M}^M \widehat{a}_{K,mp} \left(\frac{1}{k_0} \widehat{v}_{RK,p}^{h'}(Z_K) + \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \widehat{v}_{RK,p}^e(Z_K) \right) \right) e^{-jk_{xm}x} =$$

$$\sum_{m=-M}^M T_{ym}^e e^{-jk_{xm}x}. \quad (4.23l)$$

Above we simply expanded the relations found in (4.22) in order to show the correspondences and differences with the Bloch mode method in (3.42). Moreover $\widehat{a}_{i,mp}$, $\widehat{b}_{i,mp}$ and $\widehat{e}_{i,pq}$ are the entries of the matrices \mathbf{A}_i^{-1} , \mathbf{B}_i^{-1} and \mathbf{E}_i^{-1} respectively. Similar as with the Bloch mode method the constant term $-jY$ appearing in front of all the equations involving the matching of the magnetic field is divided out. The same holds for the y -dependency of the electromagnetic field since this term appears in all of the equations and does not depend on the summation index. Like in the planar RCWA diffraction case the system of equations (4.23) in known expansions functions (the pseudo-periodic Fourier series) and with unknown expansion coefficients can be solved by matching the terms in front of these expansion functions which are the same for all layers. Substituting the solution for $\widehat{v}_{Ri,m}$ results in the following set of equations for $-M \leq m \leq M$.

$$jn_1(R_{xm}^h + \delta_{0,m}(s_x \cos \psi - p_x \sin \psi)) = \sum_{s=-M}^M \tilde{q}_{R1,ms}^h (\tilde{C}_{R1,s}^{h-} \tilde{x}_{R1,s}^h + \tilde{C}_{R1,s}^{h+}), \quad (4.24a)$$

$$R_{xm}^e + \delta_{0,m}(s_x \sin \psi + p_x \cos \psi) = \sum_{s=-M}^M \tilde{q}_{R1,ms}^e (\tilde{C}_{R1,s}^{e-} \tilde{x}_{R1,s}^e + \tilde{C}_{R1,s}^{e+}), \quad (4.24b)$$

$$\begin{aligned} jn_1(R_{ym}^h + \delta_{0,m}(s_y \cos \psi - p_y \sin \psi)) = \\ \sum_{p=-M}^M \hat{b}_{1,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \hat{e}_{1,pq} \sum_{s=-M}^M \tilde{q}_{R1,qs}^h (\tilde{C}_{R1,s}^{h-} \tilde{x}_{R1,s}^h + \tilde{C}_{R1,s}^{h+}) + \right. \\ \left. \sum_{s=-M}^M \tilde{q}_{R1,ps}^e \tilde{\mu}_{R1,s}^e (\tilde{C}_{R1,s}^{e-} \tilde{x}_{R1,s}^e - \tilde{C}_{R1,s}^{e+}) \right), \end{aligned} \quad (4.24c)$$

$$\begin{aligned} R_{ym}^e + \delta_{0,m}(s_y \sin \psi + p_y \cos \psi) = \\ \sum_{p=-M}^M \hat{a}_{1,mp} \left(\sum_{q=-M}^M \tilde{q}_{R1,pq}^h \tilde{\mu}_{R1,q}^h (\tilde{C}_{R1,q}^{h-} \tilde{x}_{R1,q}^h - \tilde{C}_{R1,q}^{h+}) + \right. \\ \left. \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \tilde{q}_{R1,pq}^e (\tilde{C}_{R1,q}^{e-} \tilde{x}_{R1,q}^e + \tilde{C}_{R1,q}^{e+}) \right), \end{aligned} \quad (4.24d)$$

for $i = 1, \dots, K-1$

$$\sum_{s=-M}^M \tilde{q}_{Ri,ms}^h (\tilde{C}_{Ri,s}^{h-} + \tilde{C}_{Ri,s}^{h+} \tilde{x}_{Ri,s}^h) = \sum_{s=-M}^M \tilde{q}_{Ri+1,ms}^h (\tilde{C}_{Ri+1,s}^{h-} \tilde{x}_{Ri+1,s}^h + \tilde{C}_{Ri+1,s}^{h+}), \quad (4.24e)$$

$$\sum_{s=-M}^M \tilde{q}_{Ri,ms}^e (\tilde{C}_{Ri,s}^{e-} + \tilde{C}_{Ri,s}^{e+} \tilde{x}_{Ri,s}^e) = \sum_{s=-M}^M \tilde{q}_{Ri+1,ms}^e (\tilde{C}_{Ri+1,s}^{e-} \tilde{x}_{Ri+1,s}^e + \tilde{C}_{Ri+1,s}^{e+}), \quad (4.24f)$$

$$\begin{aligned} \sum_{p=-M}^M \hat{b}_{i,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \hat{e}_{i,pq} \sum_{s=-M}^M \tilde{q}_{Ri,qs}^h (\tilde{C}_{Ri,s}^{h-} + \tilde{C}_{Ri,s}^{h+} \tilde{x}_{Ri,s}^h) + \right. \\ \left. \sum_{s=-M}^M \tilde{q}_{Ri,ps}^e \tilde{\mu}_{Ri,s}^e (\tilde{C}_{Ri,s}^{e-} - \tilde{C}_{Ri,s}^{e+} \tilde{x}_{Ri,s}^e) \right) = \\ \sum_{p=-M}^M \hat{b}_{i+1,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \hat{e}_{i+1,pq} \sum_{s=-M}^M \tilde{q}_{Ri+1,qs}^h (\tilde{C}_{Ri+1,s}^{h-} \tilde{x}_{Ri+1,s}^h + \tilde{C}_{Ri+1,s}^{h+}) + \right. \\ \left. \sum_{s=-M}^M \tilde{q}_{Ri+1,ps}^e \tilde{\mu}_{Ri+1,s}^e (\tilde{C}_{Ri+1,s}^{e-} \tilde{x}_{Ri+1,s}^e - \tilde{C}_{Ri+1,s}^{e+}) \right), \end{aligned} \quad (4.24g)$$

$$\begin{aligned}
& \sum_{p=-M}^M \hat{a}_{i,mp} \left(\sum_{q=-M}^M \tilde{q}_{Ri,pq}^h \tilde{\mu}_{Ri,q}^h (\tilde{C}_{Ri,q}^{h-} - \tilde{C}_{Ri,q}^{h+} \tilde{x}_{Ri,q}^h) + \right. \\
& \quad \left. \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \tilde{q}_{Ri,pq}^e (\tilde{C}_{Ri,q}^{e-} + \tilde{C}_{Ri,q}^{e+} \tilde{x}_{Ri,q}^e) \right) = \\
& \sum_{p=-M}^M \hat{a}_{i+1,mp} \left(\sum_{q=-M}^M \tilde{q}_{Ri+1,pq}^h \tilde{\mu}_{Ri+1,q}^h (\tilde{C}_{Ri+1,q}^{h-} \tilde{x}_{Ri+1,q}^h - \tilde{C}_{Ri+1,q}^{h+}) + \right. \\
& \quad \left. \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \tilde{q}_{Ri+1,pq}^e (\tilde{C}_{Ri+1,q}^{e-} \tilde{x}_{Ri+1,q}^e + \tilde{C}_{Ri+1,q}^{e+}) \right), \quad (4.24h)
\end{aligned}$$

and

$$\sum_{s=-M}^M \tilde{q}_{RK,ms}^h (\tilde{C}_{RK,s}^{h-} + \tilde{C}_{RK,s}^{h+} \tilde{x}_{RK,s}^h) = j n_{II} T_{xm}^h, \quad (4.24i)$$

$$\sum_{s=-M}^M \tilde{q}_{RK,ms}^e (\tilde{C}_{RK,s}^{e-} + \tilde{C}_{RK,s}^{e+} \tilde{x}_{RK,s}^e) = T_{xm}^e, \quad (4.24j)$$

$$\begin{aligned}
& \sum_{p=-M}^M \hat{b}_{K,mp} \left(\frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \hat{e}_{K,pq} \sum_{s=-M}^M \tilde{q}_{RK,qs}^h (\tilde{C}_{RK,s}^{h-} + \tilde{C}_{RK,s}^{h+} \tilde{x}_{RK,s}^h) + \right. \\
& \quad \left. \sum_{s=-M}^M \tilde{q}_{RK,ps}^e \tilde{\mu}_{RK,s}^e (\tilde{C}_{RK,s}^{e-} - \tilde{C}_{RK,s}^{e+} \tilde{x}_{RK,s}^e) \right) = j n_{II} T_{ym}^h, \quad (4.24k)
\end{aligned}$$

$$\begin{aligned}
& \sum_{p=-M}^M \hat{a}_{K,mp} \left(\sum_{q=-M}^M \tilde{q}_{RK,pq}^h \tilde{\mu}_{RK,q}^h (\tilde{C}_{RK,q}^{h-} - \tilde{C}_{RK,q}^{h+} \tilde{x}_{RK,q}^h) + \right. \\
& \quad \left. \frac{k_y}{k_0} \frac{k_{xp}}{k_0} \sum_{q=-M}^M \tilde{q}_{RK,pq}^e (\tilde{C}_{RK,q}^{e-} + \tilde{C}_{RK,q}^{e+} \tilde{x}_{RK,q}^e) \right) = T_{ym}^e, \quad (4.24l)
\end{aligned}$$

where $\tilde{x}_{Ri,m} = e^{-k_0 \tilde{\mu}_{Ri,m} h_i}$. Compared to the Bloch mode method in (3.43) we see that here the algebraic linear system of equations for the unknown expansion coefficients is already truncated. Summarising we get for $i = 1, \dots, K-1$

$$\begin{bmatrix} \mathcal{F}_{I,s} & \mathcal{F}_{I,c} \\ -\mathcal{F}_{I,c} & \mathcal{F}_{I,s} \end{bmatrix} \begin{bmatrix} \mathcal{J} & \mathcal{J} \\ \frac{1}{n_I} \mathcal{K}_{I,z} & -\frac{1}{n_I} \mathcal{K}_{I,z} \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{c}_{R0}^- \end{bmatrix} = \tilde{\mathcal{F}}_{R1}(Z_0) \tilde{\mathbf{c}}_{R1}, \quad (4.25a)$$

$$\tilde{\mathcal{F}}_{Ri}(Z_i) \tilde{\mathbf{c}}_{Ri} = \tilde{\mathcal{F}}_{Ri+1}(Z_i) \tilde{\mathbf{c}}_{Ri+1}, \quad (4.25b)$$

$$\tilde{\mathcal{F}}_{RK}(Z_K) \tilde{\mathbf{c}}_{RK} = \begin{bmatrix} \mathcal{F}_{II,s} & \mathcal{F}_{II,c} \\ -\mathcal{F}_{II,c} & \mathcal{F}_{II,s} \end{bmatrix} \begin{bmatrix} \mathcal{J} \\ \frac{1}{n_{II}} \mathcal{K}_{II,z} \end{bmatrix} \mathbf{c}_{RK+1}^+, \quad (4.25c)$$

Similarly as in the planar diffraction case all matrices corresponding to the superstrate and substrate layers have already been defined in Section 3.3.2. Moreover all vectors are defined there as well with the exception that the subscript referring to the method needs to be changed. Also now (4.25a) and (4.25c) are not a direct result of summarising the previous set of equations. As an extra step the expansion coefficients of the Rayleigh

expansions are eliminated and replaced with the s - and p -polarised counterparts resulting in a system of equations that has the same structure as the planar diffraction case. Finally the shorthand notation for the fundamental solution or solution in the grating layers evaluated at the different interfaces changes to

$$\tilde{\mathcal{F}}_{Ri}(Z_{i-1}) = \begin{bmatrix} \mathbf{Q}_{Ri} & \mathbf{O} \\ \mathcal{A}_i \mathcal{K}_y \mathcal{K}_x \mathcal{W}_i \mathbf{Q}_{Ri} & \mathcal{A}_i \mathcal{J} \mathbf{Q}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathcal{M}_{Ri} & -\mathcal{M}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathcal{X}_{Ri} \end{bmatrix}, \quad (4.26a)$$

$$\tilde{\mathcal{F}}_{Ri}(Z_i) = \begin{bmatrix} \mathbf{Q}_{Ri} & \mathbf{O} \\ \mathcal{A}_i \mathcal{K}_y \mathcal{K}_x \mathcal{W}_i \mathbf{Q}_{Ri} & \mathcal{A}_i \mathcal{J} \mathbf{Q}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathcal{M}_{Ri} & -\mathcal{M}_{Ri} \end{bmatrix} \begin{bmatrix} \mathcal{X}_{Ri} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}. \quad (4.26b)$$

Again most (block diagonal) matrices are defined in the previous chapter after changing the subscript. The block diagonal matrices that have not yet been defined are

$$\mathcal{A}_i = \text{diag}(\mathbf{B}_i^{-1}, \mathbf{A}_i^{-1}), \quad \mathcal{W}_i = \text{diag}(\mathbf{E}_i^{-1}, \mathbf{I}), \quad \mathbf{Q}_{Ri} = \text{diag}(\tilde{\mathbf{Q}}_{Ri}^h, \tilde{\mathbf{Q}}_{Ri}^e). \quad (4.27)$$

Eliminating the truncated expansion coefficients $\tilde{\mathbf{c}}_{R0}^-$ and $\tilde{\mathbf{c}}_{RK+1}^+$ of the scattered field from (4.25a) and (4.25c) gives

$$\tilde{\mathcal{G}}_{R1}^1: \tilde{\mathcal{F}}_{R1}(Z_0) \tilde{\mathbf{c}}_{R1} = \tilde{\mathbf{d}}_R^1, \quad (4.28a)$$

$$\tilde{\mathcal{F}}_{Ri}(Z_i) \tilde{\mathbf{c}}_{Ri} = \tilde{\mathcal{F}}_{Ri+1}(Z_i) \tilde{\mathbf{c}}_{Ri+1}, \quad (4.28b)$$

$$\tilde{\mathcal{G}}_{RK}^2: \tilde{\mathcal{F}}_{RK}(Z_K) \tilde{\mathbf{c}}_{RK} = \tilde{\mathbf{d}}_R^2, \quad (4.28c)$$

with the auxiliary matrices and vectors

$$\tilde{\mathcal{G}}_{R1} = \begin{bmatrix} \mathbf{I} & -(\mathcal{K}_y^2 - n_I^2 \mathbf{I})^{-1} (\mathcal{K}_x \mathcal{K}_y - n_I \mathcal{K}_{I,z} \mathcal{J}_I) \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad (4.29a)$$

$$\tilde{\mathcal{G}}_{RK} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ -\mathbf{I} & (\mathcal{K}_y^2 - n_{II}^2 \mathbf{I})^{-1} (\mathcal{K}_x \mathcal{K}_y + n_{II} \mathcal{K}_{II,z} \mathcal{J}_{II}) \end{bmatrix}, \quad (4.29b)$$

$$\tilde{\mathbf{d}}_R = \begin{bmatrix} -2(\mathcal{K}_y^2 - n_I^2 \mathbf{I})^{-1} (n_I \mathcal{F}_{I,c} \mathcal{K}_{I,z} + \mathcal{F}_{I,s} \mathcal{K}_{I,z}^2 \mathcal{J}) \mathbf{d}_0 \\ \mathbf{O} \end{bmatrix}. \quad (4.29c)$$

Equation (4.28) can be combined into one system where the coefficient matrix is similar to the sparse block matrix in the planar diffraction case (4.21), except that all matrices and vectors need to be replaced with their conical counterparts provided with a tilde. Clearly the structure for all diffraction cases is the same and therefore we will again postpone the stable solution algorithm of this system until Chapter 5.

Chapter 5

Solution strategies for the truncated linear system

In the previous chapters on Bloch and RCWA a linear system of equations was derived when connecting the layers through the interface boundary conditions. This chapter discusses several solution strategies for the truncated linear system. This system can be solved using Gaussian elimination with pivoting which is stable but not very efficient. Due to its special sparse block structure one can also solve it using *condensation algorithms* and speed up the computation. In Section 5.1 a simple but unstable condensation algorithm known as the *transfer matrix* or *T-matrix algorithm* is briefly outlined. The instability of this algorithm is addressed before continuing with stable condensation algorithms. Over the past 20 years several stable condensation algorithms have been suggested. However from the original papers it is not always mathematically clear where the stability comes from. Therefore in Section 5.2 it is explained how a decoupling of waves by applying a Riccati transformation to the original sparse system introduces the necessary property needed for stability. Finally in Section 5.3 the link to a *multiple shooting* approach is discussed, a technique frequently used in mathematics for stably solving boundary value problems.

5.1 Unstable transfer matrix algorithm

Probably the simplest condensation algorithm and most intuitive way of solving the truncated linear system in (3.39) or (4.21) is by eliminating all but one unknown. This

results in what is known as the standard transfer matrix or T -matrix algorithm

$$\left(\mathcal{G}_1 \prod_{i=1}^{K-1} \left(\mathcal{F}_i(Z_{i-1}) \mathcal{F}_i^{-1}(Z_i) \right) \mathcal{F}_K(Z_{K-1}) + \mathcal{G}_K \mathcal{F}_K(Z_K) \right) \mathbf{c}_K = \mathbf{d}, \quad (5.1)$$

where for the moment the subscript B or R referring to the method is dropped. Recall that for the conical diffraction case the expression is similar except that then all matrices and vectors are provided with a tilde. Because there is no incoming field at the bottom interface, the bottom half of the right-hand side vector \mathbf{d} is equal to zero and (5.1) can be reduced further to

$$\mathbf{d}^1 = \mathcal{G}_1^1 \prod_{i=1}^{K-1} \left(\mathcal{F}_i(Z_{i-1}) \mathcal{F}_i^{-1}(Z_i) \right) \mathcal{F}_K(Z_{K-1}) \begin{bmatrix} \mathbf{I} \\ \mathbf{H} \end{bmatrix} \mathbf{c}_K^1, \quad (5.2a)$$

$$\mathbf{H} = - \left(\mathcal{G}_K^2 \mathcal{F}_K^2(Z_K) \right)^{-1} \left(\mathcal{G}_K^2 \mathcal{F}_K^1(Z_K) \right). \quad (5.2b)$$

Having solved the smaller system (5.2a) one could then compute $\mathbf{c}_K^2 = \mathbf{H} \mathbf{c}_K^1$ followed by a backward sweep to solve for the other expansion coefficients in the grating layers. For example, in the RCWA planar diffraction case (4.17b) is used in this backward sweep and additionally (4.17a) and (4.17c) to compute the expansion coefficients of the scattered field. In some applications where only the expansion coefficients of the scattered field are of interest, the T -matrix algorithm can be optimised slightly so that only one system needs to be solved without an additional sweep. For example, in the RCWA planar diffraction case one would eliminate all the expansion coefficients of the grating layers from (4.17) and solve the system

$$\left(\mathcal{G}_{R1}^* + \frac{\gamma_{II}}{\nu_1} \prod_{i=1}^K \left(\mathcal{F}_i(Z_{i-1}) \mathcal{F}_i^{-1}(Z_i) \right) \mathcal{G}_{RK}^* \right) \mathbf{c}_R^* = \mathbf{d}_R^*, \quad (5.3a)$$

where

$$\mathcal{G}_{R1}^* = \begin{bmatrix} -\mathbf{I} & \mathbf{O} \\ jw_0 \mathbf{K}_{I,z} & \mathbf{O} \end{bmatrix}, \quad \mathcal{G}_{RK}^* = \begin{bmatrix} \mathbf{O} & \mathbf{I} \\ \mathbf{O} & jw_{K+1} \mathbf{K}_{II,z} \end{bmatrix}, \quad (5.3b)$$

$$\mathbf{c}_R^* = \begin{bmatrix} \mathbf{c}_{R0} \\ \mathbf{c}_{RK+1}^+ \end{bmatrix}, \quad \mathbf{d}_R^* = \begin{bmatrix} \mathbf{d}_0 \\ jw_0 \mathbf{K}_{I,z} \mathbf{d}_0 \end{bmatrix}. \quad (5.3c)$$

It is known that the standard T -matrix algorithm is numerically unstable for a large truncation parameter or for thick layers. As an example consider Table 5.1 where one can see values of the diffraction efficiency computed with the stable Riccati algorithm still to be discussed and with the unstable T -matrix algorithm. Whereas an increasing number of harmonics should produce more accurate results the T -matrix algorithm starts to diverge at some point. A similar behaviour can be seen when the grating height is increased and the T -matrix algorithm again starts to deviate from the Riccati results. As it turns out the reason for this behaviour is the propagation of round-off errors caused by the finite precision of the PC, in this case double-precision arithmetic.

M	Riccati	T -matrix
10	3.78420e-001	3.78420e-001
20	3.78550e-001	3.78550e-001
30	3.78562e-001	3.75624e-001
40	3.78565e-001	8.01873e-001
50	3.78567e-001	1.03869e+000

(a) $D = 0.25$

D	Riccati	T -matrix
0.1	5.33485e-001	5.33485e-001
0.2	4.32316e-001	4.32316e-001
0.3	3.31330e-001	3.45350e-001
0.4	2.71950e-001	1.00489e+000
0.5	2.33254e-001	9.60999e-001

(b) $M = 25$

Table 5.1: Diffraction efficiency of the 0th reflected order for varying (a) harmonics M and (b) total grating height D . The grating is a silicon ($n = 3.77$) trapezoid approximated with 5 equally thick layers on a silicon substrate. The top most layer of the trapezoid has width 0.25 and the bottom most layer of the trapezoid has width 0.75. Moreover $\Lambda = \lambda_0 = 1$ and the incident field is TE polarised in a planar diffraction mount with $\theta = \pi/3$.

In order to understand in more detail why the T -matrix algorithm is unstable and when this becomes a problem one first has to analyse the solutions in the grating layers. The z -dependent part of these solutions is given by (3.11) and (4.6). From the explanation given there it follows that the solution consists of two parts: waves moving in the positive z -direction and waves moving in the negative z -direction. Moreover these waves are either propagating (i.e. solution components with constant amplitude) or evanescent (i.e. exponentially decreasing solution components). This clear distinction of solution behaviour in the positive and negative z -direction is called *dichotomy*. When looking at the fundamental solutions in (3.36) and (4.18) the dichotomy is also visible in the block structure: the left columns represent the non-increasing solutions in the positive z -direction and the right columns represent the non-increasing solutions in the negative z -direction (i.e. non-decreasing solutions in the positive z -direction). As can be seen from (5.2) or (5.3) no effort is made to split the growing and decaying modes. Generally speaking this causes the coefficient matrix of the condensed linear system to become very ill-conditioned. This means that even a small error of the order of the machine precision (typically 10^{-16} for double precision arithmetic) can grow exponentially large resulting in completely wrong expansion coefficients as was demonstrated in Table 5.1. Note that this happens long before overflow can possibly cause a problem in the inverse of the exponential matrices that are required to set up the linear system. As an indication of when the T -matrix algorithm becomes unstable one can look at the smallest exponential term per layer. If the product of all these exponential terms is of the order of the machine precision one should not expect the T -matrix algorithm to be stable anymore. For the example given above this actually happens when the number of harmonics is larger than 23 or when the grating height is greater than 0.23.

5.2 Stable recursion with Riccati transformations

The previous section has shown that simply condensing the block structured sparse linear system of equations without paying any attention to the individual solution components can be disastrous. Therefore in this section a stable condensation algorithm is derived based on Riccati transformations. In literature several other and sometimes similar (stable) condensation algorithms have been suggested like the R -matrix algorithm [24], S -matrix algorithm [14, 22] or enhanced transmittance matrix approach [30]. In [25] the connection between the R - and S -matrix algorithm has been thoroughly examined. Finally in [33] the enhanced transmittance matrix approach has been recognised as a special case of the S -matrix algorithm. All these condensation algorithms somehow exploit the dichotomy that is present in the solution components by separating the exponentially growing and decaying modes. This decoupling of waves is the crucial property needed for stability. Unfortunately a detailed description of the actual separation starting from (3.39) or (4.21) is missing in the aforementioned papers.

Riccati methods are frequently used to separate solution components [2, 28] although often they are applied to the original ODE before discretisation (in this case along the z -direction). However, apart from the expansion coefficients, here the solution in the z -direction is known in closed form (either by solving a transcendental eigenvalue equation in the case of Bloch or by solving an eigenvalue problem in the case of RCWA). A discrete equivalent of the Riccati method for continuous ODEs can be used to get the desired decoupling and stably condense the system. The key ingredients are constant Riccati transformation matrices. Both a one-stage and a two-stage Riccati approach is discussed. Although mathematically equivalent the latter has a slight implementation advantage as will be explained later. For notational convenience both approaches are derived using the planar diffraction case as a reference. The conical diffraction case then simply follows by adding another block structure having shown already that the structure of the linear system is the same. Finally, also the connection with the enhanced transmittance matrix approach is given thereby confirming that this frequently used approach is indeed stable.

5.2.1 One-stage Riccati approach

First the sparse linear system in (3.39) or (4.21) is rewritten into normal form with identity matrices on the main diagonal by premultiplying by the block diagonal matrix

$\text{diag}(\mathcal{F}_1^{-1}(Z_1), \dots, \mathcal{F}_{K-1}^{-1}(Z_{K-1}), I)$

$$\begin{bmatrix} I & & -\mathcal{F}_1^{-1}(Z_1)\mathcal{F}_2(Z_1) & & \\ & \ddots & & \ddots & \\ & & I & & -\mathcal{F}_{K-1}^{-1}(Z_{K-1})\mathcal{F}_K(Z_{K-1}) \\ \mathcal{G}_1\mathcal{F}_1(Z_0) & & & & \mathcal{G}_K\mathcal{F}_K(Z_K) \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_{K-1} \\ \mathbf{c}_K \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix}, \quad (5.4)$$

where for the moment the subscript B or R referring to the method is dropped again. Then (5.4) is transformed using a series of Riccati matrices. This is done by premultiplying by the block diagonal matrix $\text{diag}(\mathcal{R}_1^{-1}, \dots, \mathcal{R}_{K-1}^{-1}, I)$ and by introducing new scaled unknowns $\underline{\mathbf{c}}_i := \mathcal{R}_i^{-1} \mathbf{c}_i$

$$\begin{bmatrix} I & & -\mathcal{U}_1 & & \\ & \ddots & & \ddots & \\ & & I & & -\mathcal{U}_{K-1} \\ \mathcal{G}_1\mathcal{F}_1(Z_0)\mathcal{R}_1 & & & & \mathcal{G}_K\mathcal{F}_K(Z_K)\mathcal{R}_K \end{bmatrix} \begin{bmatrix} \underline{\mathbf{c}}_1 \\ \vdots \\ \underline{\mathbf{c}}_{K-1} \\ \underline{\mathbf{c}}_K \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix}. \quad (5.5a)$$

Here the Riccati transformation matrices \mathcal{R}_i and auxiliary matrices \mathcal{U}_i are given by

$$\mathcal{R}_i = \begin{bmatrix} I & \mathbf{O} \\ \mathbf{R}_i \mathbf{X}_i & I \end{bmatrix}, \quad \mathcal{U}_i = \mathcal{R}_i^{-1} \mathcal{F}_i^{-1}(Z_i) \mathcal{F}_{i+1}(Z_i) \mathcal{R}_{i+1}. \quad (5.5b)$$

Note that the inverse of a Riccati transformation matrix simply follows from putting a minus sign in front of the lower left block. Moreover this lower left block is scaled with the matrix \mathbf{X}_i containing the non-increasing exponential terms. Although not essential, it makes the comparison with the enhanced transmittance matrix approach easier and it makes the recursion to be derived properly normalised. Recall that the last block row in the (transformed) sparse linear system was said to contain the boundary conditions of the discrete BVP. It is important to realise that these boundary conditions are completely *separated*. This follows directly from the matrices \mathcal{G}_1 and \mathcal{G}_K where the bottom half and top half respectively contain only zeros. This clear separation of boundary conditions is a key ingredient in the stable recursion to be derived in this section: half of these separated boundary conditions control the non-increasing modes in the positive z -direction while the other half control the non-increasing modes in the negative z -direction. From the bottom part of the right-hand side vector it can be seen that the most efficient recursion is derived by starting at the bottom interface. This follows from the fact that there

is no incident field or source term coming from the substrate and that therefore the bottom half of \mathbf{d} is simply zero. So the bottom half of the last block row in the transformed equation reads

$$\mathbf{G}_K^{2:} \mathcal{F}_K(Z_K) \mathcal{R}_K \underline{\mathbf{c}}_K = \mathbf{0}. \quad (5.6)$$

From (5.6) and using the notation introduced in (5.2b), the initial Riccati submatrix is chosen such that $\underline{\mathbf{c}}_K^2 = \mathbf{0}$

$$\mathbf{R}_K \mathbf{X}_K = \mathbf{H}. \quad (5.7a)$$

The backward recursion for the other Riccati submatrices \mathbf{R}_i is found by introducing a decoupling. This is accomplished by requiring the auxiliary matrices \mathbf{U}_i to be block upper triangular. Setting the part \mathbf{U}_i^{21} equal to zero gives the desired recursion

$$\begin{aligned} \mathbf{R}_i \mathbf{X}_i = & \left(\left(\mathcal{F}_i^{-1}(Z_i) \right)^{21} \left(\mathcal{F}_{i+1}^{1:}(Z_i) \mathcal{R}_{i+1}^{:1} \right) + \left(\mathcal{F}_i^{-1}(Z_i) \right)^{22} \left(\mathcal{F}_{i+1}^{2:}(Z_i) \mathcal{R}_{i+1}^{:1} \right) \right) \cdot \\ & \left(\left(\mathcal{F}_i^{-1}(Z_i) \right)^{11} \left(\mathcal{F}_{i+1}^{1:}(Z_i) \mathcal{R}_{i+1}^{:1} \right) + \left(\mathcal{F}_i^{-1}(Z_i) \right)^{12} \left(\mathcal{F}_{i+1}^{2:}(Z_i) \mathcal{R}_{i+1}^{:1} \right) \right)^{-1}. \end{aligned} \quad (5.7b)$$

Since the sparse block matrix in (5.5a) now contains all block upper triangular matrices \mathbf{U}_i and the initial condition of the recursion guaranteed that $\underline{\mathbf{c}}_K^2 = \mathbf{0}$, it is easy to see that $\underline{\mathbf{c}}_i^2 = \mathbf{0}$ for all $i = K-1, \dots, 1$. So only half the number of scaled unknowns needs to be determined. This also means that only the part \mathbf{U}_i^{11} is required. Having found the final Riccati submatrix \mathbf{R}_1 , the top half of the last block row in the transformed equation reads

$$\mathbf{d}^1 = \mathbf{G}_1^{1:} \mathcal{F}_1(Z_0) \mathcal{R}_1^{:1} \underline{\mathbf{c}}_1^1. \quad (5.8)$$

Having solved this equation for $\underline{\mathbf{c}}_1^1$ the other scaled unknowns $\underline{\mathbf{c}}_i^1$ for $i = 2, \dots, K$ can be found with a forward recursion by solving the system

$$\underline{\mathbf{c}}_i^1 = \mathbf{U}_i^{11} \underline{\mathbf{c}}_{i+1}^1, \quad (5.9a)$$

where

$$\mathbf{U}_i^{11} = \left(\mathcal{F}_i^{-1}(Z_i) \right)^{11} \left(\mathcal{F}_{i+1}^{1:}(Z_i) \mathcal{R}_{i+1}^{:1} \right) + \left(\mathcal{F}_i^{-1}(Z_i) \right)^{12} \left(\mathcal{F}_{i+1}^{2:}(Z_i) \mathcal{R}_{i+1}^{:1} \right). \quad (5.9b)$$

As a final step the original expansion coefficients can be found by transforming the scaled unknowns back using the relation $\mathbf{c}_i = \mathcal{R}_i \underline{\mathbf{c}}_i = \mathcal{R}_i^{:1} \underline{\mathbf{c}}_i^1$. As already indicated in the beginning the expressions derived thus far are also valid for the conical diffraction case, except that then the Riccati submatrix \mathbf{R}_i is by itself a block matrix. As an example (and in order to explain the link with the enhanced transmittance matrix approach) the expressions for the planar RCWA fundamental solutions \mathcal{F}_{Ri} are substituted into (5.7)-(5.9) while dropping the subscript R. The stable one-stage condensation algorithm with Riccati transformations then simplifies to

Algorithm:

- (i). Compute *backward* recursion for Riccati submatrices for $i = K - 1, \dots, 1$

$$\mathbf{R}_K = \left(\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K + jw_{K+1} \mathbf{K}_{\text{I},z} \mathbf{Q}_K \right)^{-1} \left(\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K - jw_{K+1} \mathbf{K}_{\text{I},z} \mathbf{Q}_K \right), \quad (5.10a)$$

$$\mathbf{R}_i = \left(\mathbf{Q}_i^{-1} \mathbf{F}_{i+1} - (\mathbf{W}_i \mathbf{Q}_i \mathbf{M}_i)^{-1} \mathbf{G}_{i+1} \right) \left(\mathbf{Q}_i^{-1} \mathbf{F}_{i+1} + (\mathbf{W}_i \mathbf{Q}_i \mathbf{M}_i)^{-1} \mathbf{G}_{i+1} \right)^{-1}, \quad (5.10b)$$

with

$$\mathbf{F}_{i+1} = \mathbf{Q}_{i+1} (\mathbf{I} + \mathbf{X}_{i+1} \mathbf{R}_{i+1} \mathbf{X}_{i+1}),$$

$$\mathbf{G}_{i+1} = \mathbf{W}_{i+1} \mathbf{Q}_{i+1} \mathbf{M}_{i+1} (\mathbf{I} - \mathbf{X}_{i+1} \mathbf{R}_{i+1} \mathbf{X}_{i+1}).$$

- (ii). Solve linear system for scaled unknown

$$2jw_0 \gamma_1 \mathbf{K}_{\text{I},z} \mathbf{d}_0 = (\mathbf{G}_1 + jw_0 \mathbf{K}_{\text{I},z} \mathbf{F}_1) \underline{\mathbf{c}}_1^1. \quad (5.11)$$

- (iii). Compute *forward* recursion for other scaled unknowns for $i = 1, \dots, K - 1$

$$\underline{\mathbf{c}}_{i+1}^1 = 2 \left(\mathbf{Q}_i^{-1} \mathbf{F}_{i+1} + (\mathbf{W}_i \mathbf{Q}_i \mathbf{M}_i)^{-1} \mathbf{G}_{i+1} \right)^{-1} \mathbf{X}_i \underline{\mathbf{c}}_i^1. \quad (5.12)$$

- (iv). Transform scaled unknowns back to original expansion coefficients for $i = 1, \dots, K$

$$\mathbf{c}_i^+ = \underline{\mathbf{c}}_i^1, \quad (5.13a)$$

$$\mathbf{c}_i^- = \mathbf{R}_i \mathbf{X}_i \underline{\mathbf{c}}_i^1. \quad (5.13b)$$

5.2.2 Two-stage Riccati approach

The two-stage Riccati approach does not bring the sparse linear system in (3.39) or (4.21) to a normal form with identity matrices on the main diagonal. Instead the original system is transformed directly using two different series of Riccati matrices. This is done by premultiplying by the block diagonal matrix $\text{diag}(\mathcal{S}_1^{-1}, \dots, \mathcal{S}_{K-1}^{-1}, \mathbf{I})$ and by introducing

new scaled unknowns $\underline{c}_i := \mathcal{R}_i^{-1} \mathbf{c}_i$

$$\begin{bmatrix} \mathcal{V}_1 & & -\mathcal{U}_1 & & & \\ & & & & & \\ & & \ddots & & \ddots & \\ & & & & \mathcal{V}_{K-1} & & -\mathcal{U}_{K-1} \\ \mathcal{G}_1 \mathcal{F}_1(Z_0) \mathcal{R}_1 & & & & \mathcal{G}_K \mathcal{F}_K(Z_K) \mathcal{R}_K & & \end{bmatrix} \begin{bmatrix} \underline{c}_1 \\ \vdots \\ \underline{c}_{K-1} \\ \underline{c}_K \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix}. \quad (5.14a)$$

Here the Riccati transformation matrices \mathcal{S}_i , \mathcal{R}_i and auxiliary matrices \mathcal{U}_i , \mathcal{V}_i are given by

$$\mathcal{R}_i = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{R}_i \mathbf{X}_i & \mathbf{I} \end{bmatrix}, \quad \mathcal{U}_i = \mathcal{S}_i^{-1} \mathcal{F}_{i+1}(Z_i) \mathcal{R}_{i+1}, \quad (5.14b)$$

$$\mathcal{S}_i = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathcal{S}_i & \mathbf{I} \end{bmatrix}, \quad \mathcal{V}_i = \mathcal{S}_i^{-1} \mathcal{F}_i(Z_i) \mathcal{R}_i. \quad (5.14c)$$

Again the lower left block of one of the Riccati transformation matrices is scaled with the matrix \mathbf{X}_i in order to make an easy comparison later on. Following the same procedure as before the recursion is started at the bottom interface for efficiency reasons. Because the bottom half of the last row in the transformed equation has not changed compared to the one-stage Riccati approach, the initial Riccati submatrix is also given by

$$\mathbf{R}_K \mathbf{X}_K = \mathbf{H}. \quad (5.15a)$$

The backward recursion for the other Riccati submatrices \mathcal{S}_i and \mathbf{R}_i is again found by introducing a decoupling. This is accomplished by requiring both auxiliary matrices \mathcal{V}_i and \mathcal{U}_i to be block upper triangular. Setting the parts \mathcal{V}_i^{21} and \mathcal{U}_i^{21} equal to zero gives the desired two-stage recursion

$$\mathcal{S}_i = \left(\mathcal{F}_{i+1}^{2:}(Z_i) \mathcal{R}_{i+1}^{:1} \right) \left(\mathcal{F}_{i+1}^{1:}(Z_i) \mathcal{R}_{i+1}^{:1} \right)^{-1}, \quad (5.15b)$$

$$\mathbf{R}_i \mathbf{X}_i = - \left((\mathcal{S}_i^{-1})^{2:} \mathcal{F}_i^{:2}(Z_i) \right)^{-1} \left((\mathcal{S}_i^{-1})^{2:} \mathcal{F}_i^{:1}(Z_i) \right). \quad (5.15c)$$

From the resulting block upper triangular structure and the initial condition of the recursion that guaranteed $\underline{c}_K^2 = \mathbf{0}$ it follows that again $\underline{c}_i^2 = \mathbf{0}$ for all $i = K, \dots, 1$. Therefore only the parts \mathcal{V}_i^{11} and \mathcal{U}_i^{11} need to be computed. Having found the final Riccati submatrix \mathbf{R}_1 , the same equation as for the one-stage Riccati approach needs to be solved

$$\mathbf{d}^1 = \mathcal{G}_1^{1:} \mathcal{F}_1(Z_0) \mathcal{R}_1^{:1} \underline{c}_1^1. \quad (5.16)$$

The other scaled unknowns \underline{c}_i^1 for $i = 2, \dots, K$ can be found with a forward recursion by solving the system

$$\mathcal{V}_i^{11} \underline{c}_i^1 = \mathcal{U}_i^{11} \underline{c}_{i+1}^1, \quad (5.17a)$$

where

$$\mathcal{V}_i^{11} = \mathcal{F}_i^{1:}(Z_i) \mathcal{R}_i^{1:}, \quad \mathcal{U}_i^{11} = \mathcal{F}_{i+1}^{1:}(Z_i) \mathcal{R}_{i+1}^{1:}. \quad (5.17b)$$

Finally the original expansion coefficients follow from transforming the scaled unknowns back using the relation $\underline{c}_i = \mathcal{R}_i \underline{c}_i^1 = \mathcal{R}_i^{1:} \underline{c}_i^1$. For the conical diffraction case all expressions remain valid while taking care of the extra block structure appearing in the Riccati transformation matrices. In order to see the similarities with the one-stage Riccati approach the expressions for the planar RCWA fundamental solutions \mathcal{F}_R are substituted into (5.15)-(5.17) while dropping the subscript R. The stable two-stage condensation algorithm with Riccati transformations then simplifies to

Algorithm:

- (i). Compute *backward* recursion for Riccati submatrices for $i = K - 1, \dots, 1$

$$\mathbf{R}_K = \left(\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K + jw_{K+1} \mathbf{K}_{\Pi,z} \mathbf{Q}_K \right)^{-1} \left(\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K - jw_{K+1} \mathbf{K}_{\Pi,z} \mathbf{Q}_K \right), \quad (5.18a)$$

$$\mathbf{S}_i = \mathbf{G}_{i+1} \mathbf{F}_{i+1}^{-1}, \quad (5.18b)$$

$$\mathbf{R}_i = \left(\mathbf{W}_i \mathbf{Q}_i \mathbf{M}_i + \mathbf{S}_i \mathbf{Q}_i \right)^{-1} \left(\mathbf{W}_i \mathbf{Q}_i \mathbf{M}_i - \mathbf{S}_i \mathbf{Q}_i \right), \quad (5.18c)$$

with

$$\begin{aligned} \mathbf{F}_{i+1} &= \mathbf{Q}_{i+1} (\mathbf{I} + \mathbf{X}_{i+1} \mathbf{R}_{i+1} \mathbf{X}_{i+1}), \\ \mathbf{G}_{i+1} &= \mathbf{W}_{i+1} \mathbf{Q}_{i+1} \mathbf{M}_{i+1} (\mathbf{I} - \mathbf{X}_{i+1} \mathbf{R}_{i+1} \mathbf{X}_{i+1}). \end{aligned}$$

- (ii). Solve system for scaled unknown

$$2jw_0 \gamma_1 \mathbf{K}_{L,z} \mathbf{d}_0 = \left(\mathbf{G}_1 + jw_0 \mathbf{K}_{L,z} \mathbf{F}_1 \right) \underline{c}_1^1. \quad (5.19)$$

- (iii). Compute *forward* recursion for other scaled unknowns for $i = 1, \dots, K - 1$

$$\underline{c}_{i+1}^1 = \mathbf{F}_{i+1}^{-1} \mathbf{Q}_i (\mathbf{I} + \mathbf{R}_i) \mathbf{X}_i \underline{c}_i^1. \quad (5.20)$$

- (iv). Transform scaled unknowns back to original expansion coefficients for $i = 1, \dots, K$

$$\mathbf{c}_i^+ = \underline{\mathbf{c}}_i^1, \quad (5.21a)$$

$$\mathbf{c}_i^- = \mathbf{R}_i \mathbf{X}_i \underline{\mathbf{c}}_i^1. \quad (5.21b)$$

5.2.3 Riccati versus enhanced transmittance matrix approach

Although the two Riccati approaches look slightly different at first glance, the four steps described in (5.10)-(5.13) are actually mathematically equivalent to those in (5.18)-(5.21). In step (i) both approaches initialise the backward recursion identically. Moreover, after substituting (5.18b) into (5.18c) and slightly rewriting the expressions it follows that the recursive relations for \mathbf{R}_i are also the same. Clearly the linear system in step (ii) and post-processing in step (iv) are identical. Showing the equivalence of step (iii) is slightly more challenging. It can be shown that after substituting (5.18c) into (5.20) and with some basic algebraic manipulation that also both forward recursions are indeed equal.

Having shown the equivalence of both approaches, the most efficient implementation can now be chosen. To do this one can look at all unique operations in the forward recursion that are of cubic complexity: LU-factorisations where the inverse of a matrix is required, solving the subsequent system where the right-hand (or left-hand) side is a matrix and the product of two full (not diagonal) matrices. The backward recursion does not contribute significantly in this complexity analysis. This is because it requires the same LU factorisations already computed in the forward recursion and the subsequent solve has quadratic complexity, the right-hand side being a vector. Moreover, the initialisation of the backward recursion, the solving of the linear system for the scaled unknown and the post-processing are the same for both approaches. Therefore only equation (5.10b) needs to be compared with (5.18b) and (5.18c). Adding all operations and realising that an LU factorisation is more expensive than a product of two full matrices favours the two-stage Riccati approach as can be seen from Table 5.2.

	<i>LU</i>	<i>Solve</i>	<i>Product</i>
One-stage	3	3	1
Two-stage	2	2	2

Table 5.2: Number of unique operations of cubic complexity per layer

As far as we know, the stable condensation algorithms as derived in Sections 5.2.1 and 5.2.2 have not been studied before in literature in combination with RCWA or Bloch. Instead, people often use the enhanced transmittance matrix approach to stably con-

dense the linear system. The enhanced transmittance matrix approach was originally proposed in [30] and a detailed algorithmic description can be found therein. However, it is not very clear how the decoupling of waves which is the crucial property needed for stability enters this algorithm. Having shown this for the one-stage and two-stage Riccati approach we will now explain why the enhanced transmittance matrix approach is actually identical to our suggested approach so that we can therefore conclude that it is stable.

In this paragraph both approaches are compared on the algorithmic level and as a result contain a lot of detail. Without loss of continuity the reader can skip this part. In Section 6 "Full Solution Approach" of [30] the enhanced transmittance approach is given for a planar diffraction case. In that section the matrix \mathbf{W} contains the eigenvectors and should therefore be compared with our matrix \mathbf{Q} . Similarly the matrix \mathbf{V} corresponds to the product \mathbf{WQM} in our notation. In the article the corresponding backward recursion is started by initialising two matrices \mathbf{f}_{L+1} and \mathbf{g}_{L+1} with the identity matrix and $j\mathbf{Z}_{\Pi}$ where the latter corresponds to our matrix $jw_{K+1}\mathbf{K}_{\Pi,z}$. From this one computes two other matrices \mathbf{a}_L and \mathbf{b}_L using the relation below (29). It is exactly these two matrices that should be compared with how the Riccati recursion is initialised in (5.10a), more precisely $\mathbf{b}_L\mathbf{a}_L^{-1} = \mathbf{R}_K$. The updating of the matrices \mathbf{f}_l and \mathbf{g}_l from layer to layer is given by equation (30) and is the same as how we update the matrices \mathbf{F}_i and \mathbf{G}_i just below (5.10b). Similarly the updating of the matrices \mathbf{a}_i and \mathbf{b}_i just below equation (29) should be compared with our updating of the Riccati submatrix \mathbf{R}_i in (5.10a) through the relation $\mathbf{b}_i\mathbf{a}_i^{-1} = \mathbf{R}_i$. The scaled unknown \mathbf{T}_1 is identical to $\underline{\mathbf{c}}_1^I$ and the linear system that is solved in equation (31) matches our linear system in (5.11) after eliminating the scattered field amplitude. Finally the forward recursion as given in (32) and thus the matrix $\mathbf{a}_i^{-1}\mathbf{X}_i$ corresponds to our forward recursion in (5.13) and the matrix $\left(\mathbf{u}_i^{II}\right)^{-1}$. So indeed the enhanced transmittance matrix approach is identical to the Riccati approach.

5.3 Connection with multiple shooting and stabilised march

A standard way of solving a linear BVP, by which we mean a linear (system of) first-order ODE(s) with prescribed boundary conditions, is the *single shooting method*. Here the solution of the BVP is given by a linear combination of solutions of associated *initial value problems* (IVP). The solutions of these IVPs are typically found by using some sort of numerical integration scheme. When the underlying ODE contains both rapidly increasing and decreasing modes, the single shooting method runs into stability issues. These issues are partly alleviated by using a *multiple shooting method*. Here the integration interval is partitioned into several smaller subintervals on which again IVPs are solved. By matching the solution at the boundary points of these subintervals (i.e. the shooting points) the solution of the original BVP can be found. The linear system that comes from such a multiple shooting method typically looks like the linear system in

(4.21). If the unstable condensation algorithm or T -matrix algorithm would have been used to condense the coefficient matrix coming from such a multiple shooting method, again the same instability would have been observed as for the single shooting method. However, since the IVPs are solved on smaller intervals the instability would not have been as severe as for single shooting, but nevertheless still too large as Table 5.1 already indicated. Therefore the linear system of a multiple shooting method also requires a stable condensation algorithm as derived in Section 5.2 for example. A more detailed overview of shooting methods for BVPs can be found in [2].

In order to put RCWA into the framework of this classical multiple shooting method the underlying ODE per layer should simply be integrated numerically instead of writing down the solution in terms of eigenvalues and eigenvectors. Thus the shooting points simply coincide with the layer interfaces. Since most numerical integration schemes solve a first-order ODE one can not directly start from (4.6) or (4.11). In case of planar diffraction the starting point is (2.24) or (2.29) which after eliminating the z -component of the field result in the desired coupled first-order system for the tangential fields. In case of conical diffraction the coupled first-order system is given by (2.34). Naturally the fields and permittivity function are again expanded in (pseudo-) periodic Fourier series, substituted in these first-order systems and truncated. Suppose the truncated system is of order $2M$, then also $2M$ IVPs need to be solved per layer. The initial conditions of these IVPs usually correspond to unit vectors (i.e. columns of the identity matrix of the same order). The resulting linear system is then given by (4.21) where $\mathcal{F}_{R_i}(Z_i) = I$ and where $\mathcal{F}_{R_i}(Z_{i-1})$ contains the numerically integrated solutions. The overall complexity of this multiple shooting approach is still cubic since the stable condensation algorithm involves LU factorisations. However, the eigenvalue computation per layer which also has cubic complexity is now replaced by numerically solving $2M$ IVPs. For a typical grating as described in Table 5.1 it turns out that solving these IVPs is actually (much) more expensive than computing a full eigendecomposition. Here we should mention that we have not extensively investigated different numerical integration schemes and mainly focused on some of Matlab's built-in ODE solvers (e.g. Runge-Kutta). Three observations are worth mentioning after having tested these solvers for several cases. First we observed that sometimes extra shooting points inside a layer are needed for stability reasons making the subsequent linear system larger and thus even more expensive. Secondly this suggests that the multiple shooting approach might prove advantageous when the grating is approximated with a lot of (thin) layers. For standard RCWA the workload is pretty much constant per layer whereas with shooting the workload per layer becomes less when the height of a layer is reduced (i.e. typically when more layers are used to approximate the grating). Although in our application space we have not encountered many of these smooth profiles that require many layers, it might prove useful for others. Thirdly we have not made any attempt to exploit the special structure that is present in the coefficient matrix of the first-order ODE. Some of the blocks are simply zero, others are diagonal and many of the full blocks have some sort of underlying Toeplitz structure. By exploiting this structure one could define a matrix-vector product, which usually forms the basis of many of these numerical integrators, that is very efficient.

The classical multiple shooting method can also be modified using a stabilised march in combination with reduced superposition. A *stabilised march* is an alternative for the stable condensation algorithm based on Riccati transformations as discussed in the previous two sections. The *reduced superposition principle* allows for a more efficient implementation of the shooting method in the sense that less IVPs have to be solved. A detailed description of this marching technique and reduced superposition principle for a general BVP can be found in section 4.4 of [2]. Here we would just like to mention some key features and show the connection with RCWA. Recall that for standard RCWA the second-order ODEs per layer are completely independent and can in principle be solved in parallel. Similarly for standard multiple shooting the first-order ODEs per layer are completely independent and thus all IVPs can also be solved in parallel. Afterwards the stable condensation algorithm is a sequential process by design which then requires all these intermediate solution components of all layers. A stabilised march on the other hand combines both the solving of the ODE per layer with a stable decoupling algorithm. This means that not all IVPs can be solved in parallel anymore, only those that belong to one and the same layer. Essentially one starts at the bottom interface with an orthogonal set of initial conditions. Then, as with standard multiple shooting, these initial conditions are numerically integrated up to the next interface or shooting point. At this point a classical QU-decomposition is performed on the matrix containing the numerically integrated solution components. Here Q is an orthogonal matrix and U is an upper triangular matrix. The reorthogonalisation restores the linear independence of the solution components and the columns of the matrix Q become the new initial conditions for the next layer. The upper triangular matrix U introduces the necessary decoupling needed for stability. This process is repeated for all layers until the top interface of the top layer is reached. Then a small linear system can be solved for the expansion coefficients in the top layer after which all other coefficients can be found with a backward recursion using the upper triangular matrices. It is worth mentioning here that the essential part is the upper triangular matrix U taking care of the decoupling. The fact that the matrix Q is orthogonal is (very) nice to have but not essential here. Therefore also a slightly different decomposition could be used where U is block upper triangular and Q close to being orthogonal. In order to speed up the computations one can apply the reduced superposition principle. Because of the completely separated boundary conditions of the discrete BVP actually only half the number of IVPs need to be solved. This also means that half the number of columns need to be reorthogonalised resulting in a QU-decomposition where Q is a rectangular matrix of size $2M \times M$ and U is a square matrix of size $M \times M$. Instead of using the columns of the identity matrix as initial conditions, the reduced superposition principle requires the initial conditions to be chosen such that they already satisfy the boundary condition of the discrete BVP belonging to the bottom interface. So with these modifications the classical multiple shooting method is approximately twice as fast, but still requires rather smooth profiles in order to be competitive.

We will now show that the stable condensation algorithm based on Riccati transformations can also be interpreted as a stabilised march with reduced superposition but

without a classical QU-decomposition. In order to keep the notation as simple as possible we will use the planar diffraction case as a reference and drop the subscript R again. First a set of orthogonal initial conditions are defined at the bottom interface

$$\widehat{\mathcal{F}}_K(Z_K) = \begin{bmatrix} \mathbf{I} \\ j\omega_{K+1}\mathbf{K}_{II,z} \end{bmatrix} =: \begin{bmatrix} \mathbf{F}_{K+1} \\ \mathbf{G}_{K+1} \end{bmatrix}. \quad (5.22)$$

Since $\mathcal{G}_K^2 \widehat{\mathcal{F}}_K(Z_K) = \mathbf{d}^2 = 0$ they indeed satisfy the boundary conditions of the discrete BVP belonging to the bottom interface. Instead of numerically integrating these initial conditions up to the next layer, the fundamental solutions derived in (4.18) can be used to write down the solution at the next interface or shooting point

$$\begin{aligned} \widehat{\mathcal{F}}_K(Z_{K-1}) &= \mathcal{F}_K(Z_{K-1})\mathcal{F}_K^{-1}(Z_K)\widehat{\mathcal{F}}_K(Z_K) \\ &= \frac{1}{2} \begin{bmatrix} \mathbf{Q}_K & \mathbf{Q}_K \\ \mathbf{W}_K\mathbf{Q}_K\mathbf{M}_K & -\mathbf{W}_K\mathbf{Q}_K\mathbf{M}_K \end{bmatrix} \begin{bmatrix} \mathbf{x}_K^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_K \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{Q}_K^{-1} & (\mathbf{W}_K\mathbf{Q}_K\mathbf{M}_K)^{-1} \\ \mathbf{Q}_K^{-1} & -(\mathbf{W}_K\mathbf{Q}_K\mathbf{M}_K)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{K+1} \\ \mathbf{G}_{K+1} \end{bmatrix}. \end{aligned} \quad (5.23)$$

Unfortunately this expression cannot be used directly since the product $\mathcal{F}_K(Z_{K-1})\mathcal{F}_K^{-1}(Z_K)$ contains both the exponentially growing and decaying terms which caused all the problems in the first place. Equation (5.23) as such is just another way of writing down the first step in the unstable compactification algorithm or T -matrix algorithm as discussed in 5.1. If on the other hand one is able to decompose the right-hand side of (5.23) as a QU-decomposition without having to build this right-hand side explicitly (which would result in disastrous round-off error accumulation), the RCWA equivalent of a stabilised march could continue. A traditional QU-decomposition is not an option, however it is possible to come up with a decomposition where \mathbf{Q} is a rectangular matrix of size $2M \times M$ with properly scaled independent columns and where \mathbf{U} is a full (not upper triangular) matrix of size $M \times M$. Recall that without reduced superposition \mathbf{U} needs to be a (block) upper triangular matrix of size $2M \times 2M$ so that in the case of reduced superposition we only need the upper left block which can thus be upper triangular or

simply full, but of size $M \times M$. To this end (5.23) is rewritten as follows

$$\begin{aligned}
\widehat{\mathcal{F}}_K^1(Z_{K-1}) &= \frac{1}{2} \mathbf{Q}_K \left(\mathbf{X}_K^{-1} \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} + (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) + \right. \\
&\quad \left. \mathbf{X}_K \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} - (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) \right) \\
&= \mathbf{Q}_K \left(\mathbf{I} + \mathbf{X}_K \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} - (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) \cdot \right. \\
&\quad \left. \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} + (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right)^{-1} \mathbf{X}_K \right) \\
&\quad \frac{1}{2} \mathbf{X}_K^{-1} \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} + (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) \\
&= \mathbf{Q}_K (\mathbf{I} + \mathbf{X}_K \mathbf{R}_K \mathbf{X}_K) \mathbf{u}_K^{11} = \mathbf{F}_K \mathbf{u}_K^{11}, \tag{5.24a}
\end{aligned}$$

$$\begin{aligned}
\widehat{\mathcal{F}}_K^2(Z_{K-1}) &= \frac{1}{2} \mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K \left(\mathbf{X}_K^{-1} \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} + (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) - \right. \\
&\quad \left. \mathbf{X}_K \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} - (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) \right) \\
&= \mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K \left(\mathbf{I} - \mathbf{X}_K \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} - (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) \cdot \right. \\
&\quad \left. \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} + (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right)^{-1} \mathbf{X}_K \right) \\
&\quad \frac{1}{2} \mathbf{X}_K^{-1} \left(\mathbf{Q}_K^{-1} \mathbf{F}_{K+1} + (\mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K)^{-1} \mathbf{G}_{K+1} \right) \\
&= \mathbf{W}_K \mathbf{Q}_K \mathbf{M}_K (\mathbf{I} - \mathbf{X}_K \mathbf{R}_K \mathbf{X}_K) \mathbf{u}_K^{11} = \mathbf{G}_K \mathbf{u}_K^{11}. \tag{5.24b}
\end{aligned}$$

So the decomposition that RCWA makes is simply

$$\widehat{\mathcal{F}}_K(Z_{K-1}) = \begin{bmatrix} \mathbf{F}_K \\ \mathbf{G}_K \end{bmatrix} \mathbf{u}_K^{11}, \tag{5.24c}$$

where all the matrices are already defined in the stable backward one-stage Riccati recursion (5.10). Note that in the derivation above a different relation was used for \mathbf{R}_K instead of (5.10a). Namely, one can also use (5.10b) with $i = K$ in combination with the extended definition of \mathbf{F}_{K+1} and \mathbf{G}_{K+1} in (5.22). The decomposition derived in (5.24) for the bottom layer also holds for all other layers and one simply needs to replace the subscript K with i . So having shown already that the enhanced transmittance matrix approach was equal to the stable compactification algorithm based on Riccati transformations, we have now also shown that both approaches can be interpreted as a stabilised march with reduced superposition.

Chapter 6

Modifications and improvements for RCWA

Although RCWA is a simpler and more flexible mode expansion method than Bloch, this comes at the price of not discretising the transitions inside a grating layer very well. The jump in the permittivity function in combination with the (pseudo-periodic) Fourier series is not optimal. In this chapter we will therefore discuss two possible improvements for the standard RCWA method with the goal of improving the convergence. The first improvement replaces the standard Fourier discretisation with a layer-specific stretched Fourier discretisation and is known as *adaptive spatial resolution (ASR)*. ASR first applies a coordinate transformation in the periodic x -direction before Fourier discretising the transformed Maxwell's equations again. This coordinate transformation squeezes the spatial variable near a transition in a layer (thereby increasing the resolution near this point) and stretches the spatial variable far away from this transition. Because each layer in principle requires a different coordinate transformation, again a coupling matrix needs to be computed when the layers are connected through the interface boundary conditions. The second improvement removes the Fourier discretisation completely and uses a *finite difference (FD)* discretisation. In order to maintain the general structure of the linear system and the stable compactification algorithm, finite differences is applied only in the periodic x -direction and not in the other spatial coordinates. Two types of FD stencils are derived while taking special care of the inner layer interface boundary conditions. Unfortunately being unable to fully exploit the sparseness of the FD discretisation this approach is not competitive for square like grating structures but might still be useful for smooth profiles.

6.1 Adaptive spatial resolution

The name ASR and the idea of improving the convergence of RCWA using a coordinate transformation was originally suggested in [17]. In this article the geometry consists of one grating layer containing a binary structure in between a superstrate and substrate halfspace. A single coordinate transformation is then used to transform the equations in the grating layer as well as in both halfspaces followed by the standard (pseudo-periodic) Fourier discretisation. This approach ensures that no coupling matrix needs to be computed when applying the interface boundary conditions between different layers. However, two additional eigenvalue problems have to be solved in both halfspaces thereby creating additional overhead. Moreover extending this approach with a single coordinate transformation that works well for all layers of a multi-layered grating is not really feasible. Therefore in [36] the ASR method is improved with a layer specific coordinate transformation that does not require extra work in the superstrate and substrate halfspaces. The drawback here is that now a coupling matrix has to be computed when the layers are connected through the interface boundary conditions. In this section we will follow the latter approach and briefly repeat the necessary changes to the existing RCWA model in a planar diffraction case for both polarisations. The starting point is the coordinate transformation in [36] for which we show that the coefficients of the coupling matrix can be computed efficiently with the help of Bessel-related special functions. Then using the exact eigenfunctions as a reference we explain that the ASR improvement simply tries to approximate these Bloch modes better than the standard RCWA method and that therefore the convergence is also comparable to the Bloch mode method. Finally also two other coordinate transformations are investigated in order to get an idea of the criteria for a good transformation. These alternative transformations explained in more detail in Appendix C are either less smooth resulting in a coupling matrix that is easier to compute or stretch on only one side of the grating interface. This last idea is motivated by the fact that some eigenfunctions as computed by Bloch show an exponential decay on only one side of the grating interface.

6.1.1 Overview of necessary changes to RCWA

The ASR improvement that we discuss here does not change the Helmholtz equation (2.37) outside the grating layers. So there the solution is still given by the classical Rayleigh expansion as was derived in (2.49) and (2.50). However inside the grating layers we no longer solve (2.25) or (2.30) directly. Instead ASR first applies a coordinate transformation in these layers along the periodic x -direction. This transformation increases the resolution around the offsets $X_{i,l}$ where the piecewise constant permittivity function is discontinuous and where its truncated Fourier approximation is worst. A possible change of variable in layer i for $x \in (X_{i,l-1}, X_{i,l})$ and $l = 1, \dots, L_i$ is the

following

$$x(s) = \alpha_{i,l} + \beta_{i,l}s + \frac{\gamma_{i,l}}{2\pi} \sin\left(2\pi \frac{s - S_{i,l-1}}{S_{i,l} - S_{i,l-1}}\right), \quad \text{for } s \in (S_{i,l-1}, S_{i,l}), \quad (6.1)$$

where

$$\alpha_{i,l} = \frac{S_{i,l}X_{i,l-1} - S_{i,l-1}X_{i,l}}{S_{i,l} - S_{i,l-1}}, \quad (6.2a)$$

$$\beta_{i,l} = \frac{X_{i,l} - X_{i,l-1}}{S_{i,l} - S_{i,l-1}}, \quad (6.2b)$$

$$\gamma_{i,l} = G_i(S_{i,l} - S_{i,l-1}) - (X_{i,l} - X_{i,l-1}). \quad (6.2c)$$

This coordinate transformation is very general and still has some parameters that can be tuned. The additional parameters $S_{i,l}$ are the offsets in the transformed space and can in principle be chosen differently from the original offsets $X_{i,l}$ (naturally the pitch should remain the same and thus $S_{i,0} = -\Lambda/2$ and $S_{i,L_i} = \Lambda/2$). By choosing for example a uniform distribution of offsets in the transformed space one can enhance certain small features in the original problem. The parameter G_i controls the amount of stretching and squeezing that is present inside a grating layer and should lie in the interval $(0, \min_{l=1,\dots,L_i} \beta_{i,l}]$ for the transformation to be one-to-one. Note that if $S_{i,l} = X_{i,l}$ and $G_i = \beta_{i,l} = 1$ the original RCWA formulation is found again. It is important to realise that the transformation given by (6.1) is continuously differentiable and that $dx/ds = G_i$ at the offsets $s = S_{i,l}$. These criteria are important when later on other transformations are discussed. Finally it should be pointed out that (6.1) should not be used directly for the grating in Figure 2.3. This is because we do not want to refine near the edges of such a unit cell since the permittivity does not jump there. One possible option is that each layer of such a grating is temporarily shifted to the left until the permittivity function really jumps when crossing the unit cell. An additional phase factor for the fields can then compensate for this shift. Alternatively the general transformation is reformulated by combining the first and last interval for such a unit cell. Here the latter approach is chosen for the symmetric grating discussed in Chapter 3 corresponding to Figure 2.3 with $X_{i,2} = -X_{i,1} = X_i$. Setting the transformed offsets equal to the real offsets simplifies the coordinate transformation to

$$x(s) = \begin{cases} s + \frac{(G_i - 1)2X_i}{2\pi} \sin\left(2\pi \frac{s + X_i}{2X_i}\right), & |s| \leq X_i, \\ s + \frac{(G_i - 1)(\Lambda - 2X_i)}{2\pi} \text{sign}(s) \sin\left(2\pi \frac{|s| - X_i}{\Lambda - 2X_i}\right), & X_i \leq |s| \leq \frac{\Lambda}{2}. \end{cases} \quad (6.3)$$

Figure 6.1 shows an example of this transformation for $X_i = 0.2$ and $\Lambda = 1$ when no stretching ($G_i = 1$), some stretching ($G_i = 0.5$) and a reasonable amount of stretching ($G_i = 0.1$) is applied.

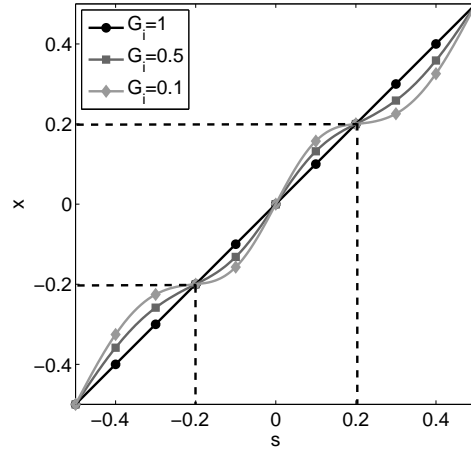


Figure 6.1: Coordinate transformation for three different values of the stretching parameter G_i for a layer in a symmetric binary grating with $X_i = 0.2$ and $\Lambda = 1$.

The coordinate transformation is used to transform the coupled system of first-order partial differential equations in the grating layers given by (2.24) and (2.29) or equivalently the second-order partial differential equations (2.25) and (2.30). Since the transformation only has an impact along the x -direction, one can simply replace $\frac{\partial}{\partial x}$ with $\frac{ds}{dx} \frac{\partial}{\partial s}$. The transformed equations for the planar diffraction case become for $i = 1, \dots, K$

$$\frac{1}{k_0^2} \left(\frac{1}{h_i} \frac{\partial}{\partial s} \frac{1}{h_i} \frac{\partial}{\partial s} + \frac{\partial^2}{\partial z^2} + k_0^2 \hat{\varepsilon}_i^r \right) \hat{E}_{i,y} = 0, \quad (6.4a)$$

$$\frac{1}{k_0^2} \left(\hat{\varepsilon}_i^r \frac{\partial}{\partial s} \frac{1}{\hat{\varepsilon}_i^r h_i} \frac{\partial}{\partial s} + \frac{\partial^2}{\partial z^2} + k_0^2 \hat{\varepsilon}_i^r \right) \hat{H}_{i,y} = 0, \quad (6.4b)$$

where

$$h_i(s) = \frac{dx}{ds} = \begin{cases} 1 + (G_i - 1) \cos \left(2\pi \frac{s + X_i}{2X_i} \right), & |s| \leq X_i, \\ 1 + (G_i - 1) \cos \left(2\pi \frac{|s| - X_i}{\Lambda - 2X_i} \right), & X_i \leq |s| \leq \frac{\Lambda}{2}, \end{cases} \quad (6.4c)$$

and

$$\hat{\varepsilon}_i^r = \hat{\varepsilon}_i^r(s), \quad \hat{E}_{i,y} = \hat{E}_{i,y}(s, z), \quad \hat{H}_{i,y} = \hat{H}_{i,y}(s, z). \quad (6.4d)$$

In principle one can now expand the transformed (reciprocal) permittivity function, the derivative of the coordinate transformation and the electromagnetic fields in (pseudo-periodic) Fourier series as was done in Section 4.1. However, in order to improve the

convergence and take full advantage of ASR, first some functions are grouped together before discretising them. Equation (6.4) is therefore rewritten into

$$\frac{1}{k_0^2} \left(\frac{\partial^2}{\partial z^2} + \frac{1}{h_i} \left(\frac{\partial}{\partial s} \frac{1}{h_i} \frac{\partial}{\partial s} + k_0^2 \widehat{\varepsilon}'_i \right) \right) \widehat{E}_{i,y} = 0, \quad (6.5a)$$

$$\frac{1}{k_0^2} \left(\frac{\partial^2}{\partial z^2} + \frac{1}{\widehat{p}'_i} \left(\frac{\partial}{\partial s} \frac{1}{\widehat{\varepsilon}'_i} \frac{\partial}{\partial s} + k_0^2 h_i \right) \right) \widehat{H}_{i,y} = 0, \quad (6.5b)$$

where

$$\widehat{\varepsilon}'_i = \widehat{\varepsilon}_i h_i, \quad \widehat{p}'_i = \frac{h_i}{\widehat{\varepsilon}_i}. \quad (6.5c)$$

Because the coordinate transformation was continuously differentiable this means that the Fourier coefficients of its derivative typically go faster to zero than those of the piecewise constant permittivity function. When a function is smooth its Fourier coefficients go to zero fast and thus less coefficients are required to get a good approximation. By Fourier transforming these scaled (reciprocal) permittivity functions instead of all the individual functions the convergence is improved significantly which is supported by the numerical tests in Section 7.1. Figure 6.2 illustrates the effect of Fourier transforming the different functions for a typical grating layer.

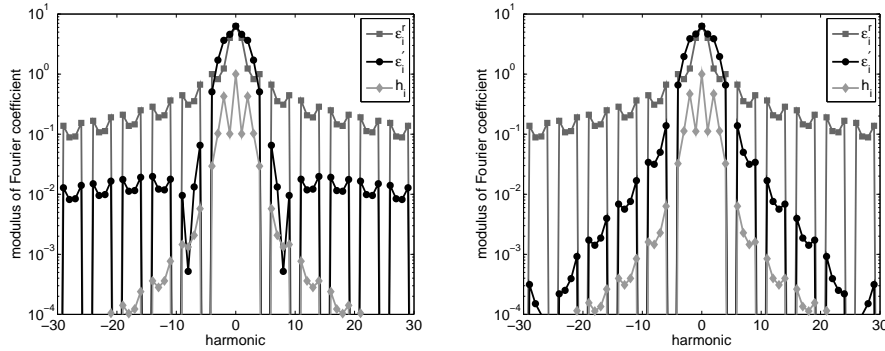


Figure 6.2: Modulus of the Fourier coefficients of the permittivity function, scaled permittivity function and derivative of the coordinate transformation. The grating layer corresponds to a symmetric binary grating with $n_{\text{I}} = 1$, $n_{\text{II}} = 3.77$, $X_i = 0.2$, $\Lambda = 1$ and $G_i = 0.1$ (left), $G_i = 0.01$ (right).

The discretisation for ASR remains the same as in standard RCWA. Thus the derivative of the coordinate transformation h_i and scaled (reciprocal) permittivity functions $\widehat{\varepsilon}'_i$ and \widehat{p}'_i are again expanded in regular Fourier series as in (4.1). Similarly the transformed electromagnetic fields $\widehat{E}_{i,y}$ and $\widehat{H}_{i,y}$ are expanded in pseudo-periodic Fourier series as in (4.2). Substituting these expansions into (6.5) and truncating them while taking care of

applying either Laurent's multiplication rule or the inverse multiplication rule results in the following two second-order ODEs

$$\frac{1}{k_0^2} \widehat{\mathbf{v}}_{Ri}^{e''} = \widehat{\mathbf{H}}_i^{-1} (\mathbf{K}_x \widehat{\mathbf{H}}_i^{-1} \mathbf{K}_x - \widehat{\mathbf{E}}_i') \widehat{\mathbf{v}}_{Ri}^e =: \widehat{\mathbf{A}}_i' \widehat{\mathbf{v}}_{Ri}^e, \quad (6.6a)$$

$$\frac{1}{k_0^2} \widehat{\mathbf{v}}_{Ri}^{h''} = \widehat{\mathbf{E}}_i'^{-1} (\mathbf{K}_x \widehat{\mathbf{P}}_i'^{-1} \mathbf{K}_x - \widehat{\mathbf{H}}_i) \widehat{\mathbf{v}}_{Ri}^h =: \widehat{\mathbf{E}}_i'^{-1} \widehat{\mathbf{B}}_i' \widehat{\mathbf{v}}_{Ri}^h, \quad (6.6b)$$

where the matrices $\widehat{\mathbf{H}}_i$, $\widehat{\mathbf{E}}_i'$ and $\widehat{\mathbf{P}}_i'$ are Toeplitz and contain the Fourier coefficients of the respective functions h_i , $\widehat{\varepsilon}_i'$ and \widehat{p}_i' . When no stretching is applied and the transformed and real offsets are chosen equal then $\widehat{\mathbf{H}}_i = \mathbf{I}$, $\widehat{\mathbf{E}}_i' = \mathbf{E}_i$ and $\widehat{\mathbf{P}}_i' = \mathbf{P}_i$ and the standard RCWA equations (4.6) are recovered. Note that in (6.6a) the inverse of the Toeplitz matrix $\widehat{\mathbf{H}}_i$ appears suggesting the use of the inverse multiplication rule. However, since the derivative of the coordinate transformation is continuous actually both multiplication rules can be applied. In this case the inverse rule is used because then all Fourier coefficients can be computed analytically. The solution of these two ODEs can be written in terms of the eigenvalues and eigenvectors of the constant coefficient matrices like in (4.7). Finally the solution of the transformed electromagnetic field in each layer is then given by (4.8) except that now the variable x in the pseudo-periodic Fourier modes needs to be replaced with s .

The unknown expansion coefficients in the solution of the (transformed) electromagnetic fields in each of the grating layers, superstrate and substrate can be found by applying the interface boundary conditions. Since in principle all grating layers can have different coordinate transformations, one can no longer simply match the coefficients as was done with standard RCWA in 4.2.1. Similar as with Bloch in 3.3.1 a method of moments technique can be used to derive the necessary relations. Here we choose a homogeneous method and use the standard pseudo-periodic Fourier modes in (2.46) as test functions. Note that these are also used as the expansion functions in the superstrate and substrate. The reason for choosing a homogeneous method is that computing the coefficients of the corresponding coupling or projection matrix \mathbf{S}_i are relatively expensive making a hybrid method less attractive. For each grating layer the coefficients of this coupling matrix are given by

$$\begin{aligned} \mathbf{S}_{i,pm} &= \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-jk_{xm}s(x)} e^{jk_{xp}x} dx \\ &= \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} h_i(s) e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds. \end{aligned} \quad (6.7)$$

Since the coordinate transformation in (6.3) is given as a function of the transformed variable s and is not that easily inverted, the second line in (6.7) can be used to compute all coefficients. However, it is not completely clear how these coupling coefficients can be computed stably and efficiently for this specific coordinate transformation. At least three different approaches can be used to compute these coefficients. Perhaps the simplest option is to approximate the integral and use some sort of numerical quadrature.

The second option is to rewrite the integral into computing the Fourier coefficients of some function and then use an FFT (Fast Fourier Transform) library routine. The third option is to rewrite the integral into the sum of some standard integrals resulting in the evaluation of Bessel related special functions. In Appendix B these three approaches are explained in more detail. Having computed the coupling matrix in (6.7) the matching boundary conditions again result in a system of equations identical to (4.17), except that the fundamental solutions in the grating layers are slightly different. Therefore (4.18) needs to be replaced with

$$\mathcal{F}_{Ri}(Z_{i-1}) = \begin{bmatrix} \mathbf{S}_i & \mathbf{O} \\ \mathbf{O} & \mathbf{S}_i \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{Q}}_{Ri} & \mathbf{O} \\ \mathbf{O} & \widehat{\mathbf{W}}_i \widehat{\mathbf{Q}}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \widehat{\mathbf{M}}_{Ri} & -\widehat{\mathbf{M}}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \widehat{\mathbf{X}}_{Ri} \end{bmatrix}, \quad (6.8a)$$

$$\mathcal{F}_{Ri}(Z_i) = \begin{bmatrix} \mathbf{S}_i & \mathbf{O} \\ \mathbf{O} & \mathbf{S}_i \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{Q}}_{Ri} & \mathbf{O} \\ \mathbf{O} & \widehat{\mathbf{W}}_i \widehat{\mathbf{Q}}_{Ri} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \widehat{\mathbf{M}}_{Ri} & -\widehat{\mathbf{M}}_{Ri} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{X}}_{Ri} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}. \quad (6.8b)$$

The matrices $\widehat{\mathbf{Q}}_{Ri}$, $\widehat{\mathbf{M}}_{Ri}$ and $\widehat{\mathbf{X}}_{Ri}$ are similar to standard RCWA and thus contain the eigenvectors, square root of the eigenvalues and exponential terms but now of the coefficient matrices in (6.6). Finally the matrix $\widehat{\mathbf{W}}_i$ is again related to the weight function coming from the different matching boundary condition between TE and TM polarisation. For TE polarisation this matrix is simply the identity, but for TM polarisation this matrix is equal to the Toeplitz matrix $\widehat{\mathbf{P}}_i$ containing the Fourier coefficients of the transformed reciprocal permittivity. This means that in case of TM polarisation first the multiplication with the reciprocal permittivity is taken in the transformed space before coupling or projecting the solution to the next layer. The remaining steps are identical to standard RCWA, meaning that the large linear system for the expansion coefficients in the grating layers can be derived and solved using a stable condensation algorithm.

6.2 Finite difference discretisation

The Fourier discretisation used in both standard RCWA and ASR does not model the material interfaces inside a grating layer as well as the Bloch modes. Although ASR alleviates the problem by stretching near such an interface one could also remove the Fourier discretisation completely and replace it with a finite difference (FD) discretisation in the grating layers. An FD discretisation can model this material interface better either by discretising the inner layer interface boundary conditions directly or by incorporating this interface into a special finite difference stencil. Unlike Bloch this approach (as well as ASR) does not get more complicated when the materials are lossy or when more material interfaces are present in a grating layer. In order to remain in the RCWA framework as much as possible the discretisation along the other spatial coordinates remains the same. This means that along the z -direction the geometry is still approximated by layers in which a second-order ODE is solved using an eigendecomposition. Also in the superstrate and substrate halfspaces the Rayleigh expansion is used to write

down the solution. The subsequent transmission problem can then be solved with the classical enhanced transmittance matrix approach or the stable condensation algorithm based on Riccati transformation as discussed in Section 5.2.

Focusing on the planar diffraction case a typical starting point when applying an FD discretisation are the coupled first-order partial differential equations in (2.24) or (2.29). By eliminating the x -component of the field only tangential components to the material interface inside a grating layer remain. Then, like with a Finite Difference Time Domain (FDTD) method, some sort of Yee cell can be introduced resulting in a sampling of the electric and magnetic field on a staggered grid. Having approximated the first-order derivatives with respect to the periodic x -direction with typically a central difference scheme, one can then eliminate the remaining z -component of the field. This finally results in a second-order ODE for the tangential y -component of the field. In [23] such an approach is chosen where the resulting second-order ODE is optimised for the specific grating diffraction problem by some sort of averaging of the (reciprocal) permittivity function when crossing a material interface. In this section we start directly from the second-order PDE (2.25) or (2.30) requiring no staggered grid making the derivation somewhat easier. Moreover, the averaging of the permittivity is replaced by discretising the inner layer interface boundary conditions (3.2a) directly. Finally also a non-uniform mesh different from [23] is investigated that is based on geometric refinement closely matching the behaviour of the ASR coordinate transformation in (6.1).

6.2.1 Partitioned domain approach

A first-order accurate finite difference scheme for the second derivative of a function using a non-uniform mesh with geometric refinement is derived in several steps. First consider the interval $[0, \Lambda]$ which is sampled with the points $\{x_n\}_{n=0}^{n=N+1}$ such that $x_0 = 0$, $x_{N+1} = \Lambda$ with geometric refinement near the right boundary as indicated in Figure 6.3. The geometric refinement is constructed using the auxiliary function g

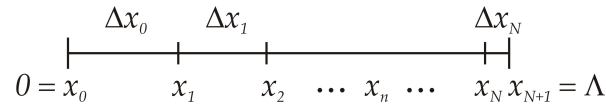


Figure 6.3: Non-uniform mesh with geometric refinement near the right boundary

$$g(x) = \frac{\delta^{1-x} - \delta}{1 - \delta} \Lambda, \quad \text{with } \delta = \frac{1}{G^{N+1}}. \quad (6.9)$$

The location of the discretisation points or mesh is now given by

$$x_n = g \left(\frac{n}{N+1} \right), \quad (6.10)$$

for $n = 0, \dots, N+1$. Due to its construction the points x_0 and x_{N+1} indeed correspond with the left and right boundary of the interval and moreover

$$\begin{aligned} \frac{\Delta x_n}{\Delta x_{n-1}} &= \frac{x_{n+1} - x_n}{x_n - x_{n-1}} = \frac{\delta^{1-\frac{n+1}{N+1}} - \delta^{1-\frac{n}{N+1}}}{\delta^{1-\frac{n}{N+1}} - \delta^{1-\frac{n-1}{N+1}}} \\ &= \frac{\delta^{-\frac{1}{N+1}} - 1}{1 - \delta^{\frac{1}{N+1}}} = \frac{G - 1}{1 - G^{-1}} = G. \end{aligned} \quad (6.11)$$

If the parameter G which controls the level of refinement is chosen positive and smaller than 1, then the mesh refines geometrically to the right boundary. Note that a similar parameter is present in the ASR improvement that controls the amount of stretching and squeezing in a grating layer. Clearly the standard uniform mesh is obtained when setting $G = 1$. From (6.11) it follows that for $n \in 0, \dots, N$

$$\Delta x_n = G^n \Delta x_0, \quad (6.12a)$$

and

$$\Lambda = \sum_{n=0}^N \Delta x_n = \Delta x_0 \sum_{n=0}^N G^n = \Delta x_0 \frac{1 - G^{N+1}}{1 - G}, \quad (6.12b)$$

so that

$$\Delta x_0 = \frac{1 - G}{1 - G^{N+1}} \Lambda, \quad (6.12c)$$

for $G \neq 1$. For a uniform mesh (6.12c) should be replaced by $\Delta x_0 = \Lambda/(N+1)$. Having introduced a mesh with geometric refinement towards one boundary one can now construct a first-order accurate difference scheme for the second derivative of a function, say $u(x)$, on all interior points. The standard way of deriving such a scheme is by properly combining the Taylor expansions of $u(x_{n-1})$ and $u(x_{n+1})$ around the point x_n with $u(x_n)$ resulting in

$$u''(x_n) = 2 \frac{\Delta x_n u(x_{n-1}) - (\Delta x_{n-1} + \Delta x_n) u(x_n) + \Delta x_{n-1} u(x_{n+1})}{\Delta x_{n-1} \Delta x_n (\Delta x_{n-1} + \Delta x_n)} + err, \quad (6.13a)$$

$$\begin{aligned} err &= \frac{1}{3} (\Delta x_{n-1} - \Delta x_n) u'''(x_n) - \\ &\quad \frac{1}{12} (\Delta x_{n-1}^2 - \Delta x_{n-1} \Delta x_n + \Delta x_n^2) u''''(x_n) + \dots \end{aligned} \quad (6.13b)$$

The three-point relation or stencil (6.13a) for the second derivative on a non-uniform mesh is only first-order accurate as can be seen from the error term in (6.13b). In the special case of a uniform mesh this finite difference scheme reduces to the standard symmetric three-point stencil which is second-order accurate. Combining (6.11) with (6.13a) gives the following approximation of the second derivative on a mesh with geometric refinement

$$\begin{aligned} u''(x_n) &\doteq \frac{1}{\Delta x_{n-1}^2} \frac{2}{G(G+1)} (Gu(x_{n-1}) - (G+1)u(x_n) + u(x_{n+1})) \\ &= \frac{1}{\Delta x_n^2} \frac{2G}{G+1} (Gu(x_{n-1}) - (G+1)u(x_n) + u(x_{n+1})). \end{aligned} \quad (6.14)$$

Apart from this approximation on the interior points also the boundary points need to be considered. There typically boundary conditions are approximated which in the RCWA context means that a first derivative is required near the right boundary point with refinement. As will become clear later on the left boundary point actually can be treated as an interior point and is therefore neglected in this part. For the non-uniform mesh a first-order approximation of the first derivative can be used making the overall scheme also first-order accurate, whereas for the uniform mesh a second-order accurate scheme is probably best. Again Taylor expansions of $u(x_N)$ (and $u(x_{N-1})$) around the point x_{N+1} can be combined with $u(x_{N+1})$ resulting in the following single sided difference schemes

$$u'(x_{N+1}) = \frac{u(x_{N+1}) - u(x_N)}{\Delta x_N} + \frac{1}{2}\Delta x_N u''(x_{N+1}) + \dots, \quad (6.15a)$$

$$\begin{aligned} u'(x_{N+1}) &= \frac{\Delta x_N^2 u(x_{N-1}) - (\Delta x_{N-1} + \Delta x_N)^2 u(x_N)}{\Delta x_{N-1} \Delta x_N (\Delta x_{N-1} + \Delta x_N)} + \\ &\quad \frac{\Delta x_{N-1} (\Delta x_{N-1} + 2\Delta x_N) u(x_{N+1})}{\Delta x_{N-1} \Delta x_N (\Delta x_{N-1} + \Delta x_N)} + \frac{1}{6} \Delta x_N (\Delta x_{N-1} + \Delta x_N) u'''(x_{N+1}) + \dots \end{aligned} \quad (6.15b)$$

Again combining (6.11) with (6.15) gives two possible approximations of the first derivative at the right boundary of a mesh with geometric refinement

$$u'(x_{N+1}) \doteq \frac{1}{\Delta x_N} (u(x_{N+1}) - u(x_N)), \quad (6.16a)$$

$$\begin{aligned} u'(x_{N+1}) &\doteq \frac{1}{\Delta x_{N-1}} \frac{1}{G(G+1)} (G^2 u(x_{N-1}) - (G+1)^2 u(x_N) + (2G+1)u(x_{N+1})) \\ &= \frac{1}{\Delta x_N} \frac{1}{G+1} (G^2 u(x_{N-1}) - (G+1)^2 u(x_N) + (2G+1)u(x_{N+1})). \end{aligned} \quad (6.16b)$$

Having derived all necessary relations on the interior and boundary points of a non-uniform mesh with geometric refinement, one can now apply these stencils to the spe-

cific grating diffraction problem. Again the symmetric grating discussed in Chapter 3 corresponding to Figure 2.3 with $X_{i,2} = -X_{i,1} = X_i$ is used as an example. The geometry is simplified even further by assuming only a single layer grating. This means that the layer index subscript i can be dropped. Like the coordinate transformation in ASR here the mesh refines on both sides of the material interface as indicated in Figure 6.4. More precisely the mesh actually consists of four copies of the auxiliary mesh in

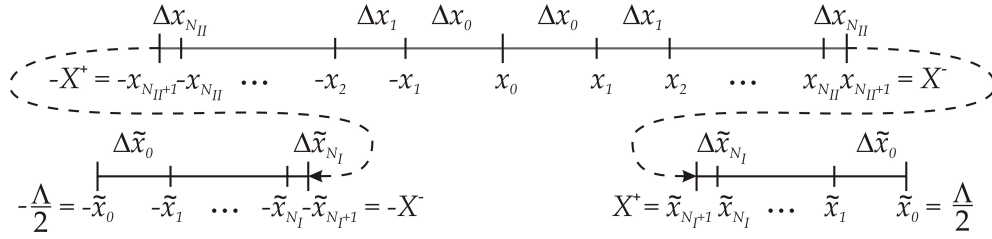


Figure 6.4: Non-uniform mesh with geometric refinement near material interfaces

Figure 6.3 in a symmetric configuration. The top part of Figure 6.4 shows the mesh in between the offsets $(-X, X)$ where the permittivity is constant and equal to $\epsilon_{II} = n_{II}^2$. Here $x_0 = 0$ is centred in the origin whereas $-x_{N_{II}+1} = -X^+$ and $x_{N_{II}+1} = X^-$ are located just right and left of the material interface respectively. The bottom part of Figure 6.4 shows the mesh in between $(-\frac{\Lambda}{2}, -X)$ and $(X, \frac{\Lambda}{2})$ where the permittivity is again constant but now equal to $\epsilon_I = n_I^2$. Here $\pm x̃_0 = \pm \Lambda/2$ corresponds with the end points of the interval. Moreover $-x̃_{N_I+1} = -X^-$ and $x̃_{N_I+1} = X^+$ are now located just to the left and right of the material interface respectively. The four points located at the material interfaces are considered the boundary points of the problem. There the inner layer interface boundary conditions (3.2a) are discretised using the single sided finite difference stencils derived in (6.15). All other points are considered interior points including those located at $x = \pm \frac{\Lambda}{2}$. This is because there is no material interface located at these edges. Instead the pseudo-periodic boundary condition implies that the interior point located at $-x̃_0$ actually overlaps with $x̃_0$ and can therefore be removed completely. On the remaining interior points the second-order PDE (2.25) or (2.30) is semi-discretised using the three-point stencil as derived in (6.14). Thus the pseudo-periodic boundary condition should be incorporated into this stencil at the locations $-x̃_1$ and $x̃_0$.

Let $f_{\pm n}(z) = F(\pm x_n, z)$ for $n = 1, \dots, N_{II} + 1$ and $f_0(z) = F(x_0, z)$ (and similarly for \tilde{f} related to $F(\tilde{x}_n, z)$) represent the sampled y -component of either the electric or magnetic field for TE or TM polarisation respectively. Then on the interior points the following system of second-order ODEs can be derived for $n = 2, \dots, N_I$

$$\begin{cases} \frac{1}{k_0^2} \frac{d^2}{dz^2} \tilde{f}_{-1} = \frac{1}{k_0^2} \frac{1}{\Delta \tilde{x}_1^2} \frac{2G_I}{G_I + 1} (-\tau^{-1} G_I \tilde{f}_0 + (G_I + 1) \tilde{f}_{-1} - \tilde{f}_{-2}) - \epsilon_I \tilde{f}_{-1} \\ \frac{1}{k_0^2} \frac{d^2}{dz^2} \tilde{f}_{-n} = \frac{1}{k_0^2} \frac{1}{\Delta \tilde{x}_n^2} \frac{2G_I}{G_I + 1} (-G_I \tilde{f}_{-n+1} + (G_I + 1) \tilde{f}_{-n} - \tilde{f}_{-n-1}) - \epsilon_I \tilde{f}_{-n} \end{cases} \quad (6.17a)$$

and for $n = 1, \dots, N_{\text{II}}$

$$\begin{cases} \frac{1}{k_0^2} \frac{d^2}{dz^2} f_{-n} = \frac{1}{k_0^2} \frac{1}{\Delta x_n^2} \frac{2G_{\text{II}}}{G_{\text{II}} + 1} (-f_{-n-1} + (G_{\text{II}} + 1)f_{-n} - G_{\text{II}}f_{-n+1}) - \epsilon_{\text{II}} f_{-n} \\ \frac{1}{k_0^2} \frac{d^2}{dz^2} f_0 = \frac{1}{k_0^2} \frac{1}{\Delta x_0^2} (-f_{-1} + 2f_0 - f_1) - \epsilon_{\text{II}} f_0 \\ \frac{1}{k_0^2} \frac{d^2}{dz^2} f_n = \frac{1}{k_0^2} \frac{1}{\Delta x_n^2} \frac{2G_{\text{II}}}{G_{\text{II}} + 1} (-G_{\text{II}}f_{n-1} + (G_{\text{II}} + 1)f_n - f_{n+1}) - \epsilon_{\text{II}} f_n \end{cases} \quad (6.17b)$$

and for $n = N_{\text{I}}, \dots, 1$

$$\begin{cases} \frac{1}{k_0^2} \frac{d^2}{dz^2} \tilde{f}_n = \frac{1}{k_0^2} \frac{1}{\Delta \tilde{x}_n^2} \frac{2G_{\text{I}}}{G_{\text{I}} + 1} (-\tilde{f}_{n+1} + (G_{\text{I}} + 1)\tilde{f}_n - G_{\text{I}}\tilde{f}_{n-1}) - \epsilon_{\text{I}} \tilde{f}_n \\ \frac{1}{k_0^2} \frac{d^2}{dz^2} \tilde{f}_0 = \frac{1}{k_0^2} \frac{1}{\Delta \tilde{x}_0^2} (-\tilde{f}_1 + 2\tilde{f}_0 - \tau \tilde{f}_{-1}) - \epsilon_{\text{I}} \tilde{f}_0 \end{cases} \quad (6.17c)$$

where $\tau = e^{-jk_x \Lambda}$ comes from the pseudo-periodic boundary condition. Before this system of second-order ODEs can be solved (which up to now is the same for both TE and TM polarisation), first the boundary points located at both sides of the material interfaces have to be eliminated. As already mentioned the necessary relations are derived from discretising the inner layer interface boundary conditions (3.2a) (which are different for TE and TM polarisation) using single sided finite difference stencils. The first-order approximations are

$$\begin{cases} \tilde{f}_{-N_{\text{I}}-1} = f_{-N_{\text{II}}-1} \\ \frac{w_{\text{I}}}{\Delta \tilde{x}_{N_{\text{I}}}} (-\tilde{f}_{-N_{\text{I}}} + \tilde{f}_{-N_{\text{I}}-1}) = \frac{w_{\text{II}}}{\Delta x_{N_{\text{II}}}} (-f_{-N_{\text{II}}-1} + f_{-N_{\text{II}}}) \end{cases} \quad (6.18a)$$

$$\begin{cases} f_{N_{\text{II}}+1} = \tilde{f}_{N_{\text{I}}+1} \\ \frac{w_{\text{II}}}{\Delta x_{N_{\text{II}}}} (-f_{N_{\text{II}}} + f_{N_{\text{II}}+1}) = \frac{w_{\text{I}}}{\Delta \tilde{x}_{N_{\text{I}}}} (-\tilde{f}_{N_{\text{I}}+1} + \tilde{f}_{N_{\text{I}}}) \end{cases} \quad (6.18b)$$

whereas the second-order approximations are

$$\left\{ \begin{array}{l} \tilde{f}_{-N_I-1} = f_{-N_{II}-1} \\ \frac{w_I}{\Delta \tilde{x}_{N_I}} \frac{1}{G_I + 1} (G_I^2 \tilde{f}_{-N_I+1} - (G_I + 1)^2 \tilde{f}_{-N_I} + (2G_I + 1) \tilde{f}_{-N_I-1}) = \\ \frac{w_{II}}{\Delta x_{N_{II}}} \frac{1}{G_{II} + 1} (-2G_{II} + 1) f_{-N_{II}-1} + (G_{II} + 1)^2 f_{-N_{II}} - G_{II}^2 f_{-N_{II}+1} \end{array} \right. \quad (6.19a)$$

$$\left\{ \begin{array}{l} f_{N_{II}+1} = \tilde{f}_{N_I+1} \\ \frac{w_{II}}{\Delta x_{N_{II}}} \frac{1}{G_{II} + 1} (G_{II}^2 f_{N_{II}-1} - (G_{II} + 1)^2 f_{N_{II}} + (2G_{II} + 1) f_{N_{II}+1}) = \\ \frac{w_I}{\Delta \tilde{x}_{N_I}} \frac{1}{G_I + 1} (-2G_I + 1) \tilde{f}_{N_I+1} + (G_I + 1)^2 \tilde{f}_{N_I} - G_I^2 \tilde{f}_{N_I-1} \end{array} \right. \quad (6.19b)$$

Using the first-order approximations (6.18) to eliminate the field components sampled on the boundary points from (6.17) results in the following first-order accurate system of second-order ODEs

$$\frac{1}{k_0^2} \frac{d^2}{dz^2} \begin{bmatrix} \tilde{\mathbf{f}} \\ \mathbf{f} \end{bmatrix} = \begin{bmatrix} \mathcal{A}^{11} & \mathcal{A}^{12} \\ \mathcal{A}^{21} & \mathcal{A}^{22} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{f}} \\ \mathbf{f} \end{bmatrix}, \quad (6.20a)$$

where

$$\tilde{\mathbf{f}} = [\tilde{f}_{N_I}, \dots, \tilde{f}_{-N_I}]^T, \quad \mathbf{f} = [f_{-N_{II}}, \dots, f_{N_{II}}]^T. \quad (6.20b)$$

The constant coefficient matrix \mathcal{A} consists of two square tri-diagonal matrices \mathcal{A}^{11} and \mathcal{A}^{22} and two (in general) rectangular matrices \mathcal{A}^{12} and \mathcal{A}^{21} with non-zero components in the lower left and top right corner. Here the partitioning of the matrix matches the size of the vectors in (6.20b). More specifically the matrices on the diagonal can be written as

$$\mathcal{A}^{11} = \mathbf{X}_I \mathbf{D}_I - \epsilon_I \mathbf{I}, \quad \mathcal{A}^{22} = \mathbf{X}_{II} \mathbf{D}_{II} - \epsilon_{II} \mathbf{I}, \quad (6.21a)$$

where

$$\mathbf{X}_I = \frac{1}{k_0^2} \frac{2G_I}{G_I + 1} \text{diag} \left(\frac{1}{\Delta \tilde{x}_{N_I}^2}, \dots, \frac{1}{\Delta \tilde{x}_1^2}, \frac{G_I + 1}{2G_I} \frac{1}{\Delta \tilde{x}_0^2}, \frac{1}{\Delta \tilde{x}_1^2}, \dots, \frac{1}{\Delta \tilde{x}_{N_I}^2} \right), \quad (6.21b)$$

$$\mathbf{X}_{II} = \frac{1}{k_0^2} \frac{2G_{II}}{G_{II} + 1} \text{diag} \left(\frac{1}{\Delta x_{N_{II}}^2}, \dots, \frac{1}{\Delta x_1^2}, \frac{G_{II} + 1}{2G_{II}} \frac{1}{\Delta x_0^2}, \frac{1}{\Delta x_1^2}, \dots, \frac{1}{\Delta x_{N_{II}}^2} \right), \quad (6.21c)$$

and the off-diagonal blocks have slightly more fill-in and are given by

$$\mathcal{A}^{12} = \frac{1}{k_0^2} \frac{1}{\Delta x_0^2} \begin{bmatrix} & \frac{1}{3}\bar{w}_{\text{II}} & -\frac{4}{3}\bar{w}_{\text{II}} \\ -\frac{4}{3}\bar{w}_{\text{II}} & \frac{1}{3}\bar{w}_{\text{II}} & \end{bmatrix}, \quad (6.22d)$$

$$\mathcal{A}^{21} = \frac{1}{k_0^2} \frac{1}{\Delta x_0^2} \begin{bmatrix} & \frac{1}{3}\bar{w}_{\text{I}} & -\frac{4}{3}\bar{w}_{\text{I}} \\ -\frac{4}{3}\bar{w}_{\text{I}} & \frac{1}{3}\bar{w}_{\text{I}} & \end{bmatrix}. \quad (6.22e)$$

Having derived the second-order ODE which is either first-order accurate with geometric refinement or second-order accurate on a regular uniform mesh, one can now return to the standard RCWA approach for solving it. By computing the eigenvalues and eigenvectors of the constant coefficient matrix the z -dependent part of the solution can be written down like in (4.7). Initially the idea was to keep the matrix as sparse as possible and find an optimised eigendecomposition routine which exploits this sparseness (e.g. library routines from ARPACK [39] or EISPACK [40]). This is the reason why only low-order approximations are considered resulting in less fill-in. Unfortunately the periodic boundary conditions make the coefficient matrix not banded whereas most of these optimised routines require this property. Moreover all eigenvalues and eigenvectors are required and not just a few. Therefore all numerical results were obtained using the same eigendecomposition routine that was used in standard RCWA based on dense and full matrices (e.g. library routines from LAPACK [41]). Clearly this does not result in optimal computation times and perhaps other strategies could be employed improving the performance. One alternative could be to solve the second-order ODE numerically using for example a multiple shooting approach as discussed in Section 5.3. Because the matrix is very sparse an efficient matrix-vector product could be defined which is possibly more efficient than having to compute all eigenvectors and eigenvalues. Having solved the ODE the unknown expansion coefficients are again found by applying the interface boundary conditions (2.27) or (2.32). Here also a small modification is required since the solution in the grating layer is only available at discrete points corresponding to the mesh along the x -direction, whereas in the superstrate and substrate a continuous description is available in the form of the Rayleigh expansions (which uses the pseudo-periodic Fourier series along the x -direction). One possible solution is to sample these Rayleigh expansions at the mesh points of the grating layer when applying the interface boundary conditions. The remaining steps are again identical to standard RCWA, meaning that the resulting linear system can be solved using

a stable condensation algorithm. When the grating consist of more layers each having their own mesh these interface boundary conditions become slightly more involved. Then an approach as discussed in [23] could be considered where with each mesh-point a local expansion functions is associated followed by a moment matching technique as discussed in 3.3.1.

6.2.2 Single domain approach

Contrary to the partitioned domain approach consisting of a mesh with both interior and boundary points one can also consider a single domain approach consisting of a mesh with interior points only. In that case a special finite difference stencil needs to be derived for those points in the mesh that lie just left and right of a material interface (when a material interface coincides with a mesh point only there a special stencil is required). This approach has been used before in the context of waveguides and in [37] a derivation of such a stencil can be found. Because for waveguide computations the periodic boundary condition is typically replaced with a transparent boundary condition the resulting matrix is not only sparse but also banded. So then it is worthwhile to invest in a stencil that results in a tri-diagonal matrix for which a lot of efficient library routines are available. Recall that in the partitioned domain approach as discussed in 6.2.1 the single sided difference scheme needed for the boundary points increased the bandwidth of the matrix. The stencil as derived in [37] assumes a uniform mesh and is second-order accurate for TE polarisation whilst only first-order accurate for TM polarisation. However by exploiting the special structure of the equations to be discretised the accuracy can be improved to fourth and second-order for TE and TM polarisation respectively without increasing the bandwidth of the matrix [38]. Such a single domain approach on a uniform mesh can also be used in the context of grating diffraction problems. However in combination with an eigendecomposition routine that is not able to exploit the sparseness because of the periodic boundary conditions, the added value of a single domain approach over a partitioned domain approach seems limited.

Chapter 7

Numerical results

Having discussed various discretisation strategies and modifications in the previous chapters, this chapter compares and evaluates these methods by presenting numerical results. In Section 7.1 some typical test structures are considered consisting of a relatively easy low contrast resist application, followed by a more challenging high contrast silicon application. Section 7.2 explains how these methods can be incorporated in a metrology tool for critical dimension (CD) measurements.

7.1 Comparison of the forward diffraction models and their improvements

First the two mode expansion methods Bloch and RCWA as discussed in Chapters 3 and 4 are compared for a resist and silicon type application. The resist case consists of a resist pillar lying on a homogeneous Backward Anti-Reflective Coating (BARC) layer on a silicon substrate with the superstrate simply air. The refraction indices of air, resist, BARC and silicon are fixed at 1.0, 1.51, 1.55 and 3.88 respectively. The geometrical parameters like pitch, resist width, resist height and BARC height are fixed at 500nm, 80nm or 250nm (iso or dense), 120nm and 30nm. Note that two different resist widths are investigated corresponding to two typical configurations, namely isolated and dense structures. Moreover the resist pillar is modelled as a single layer structure in case of a perfectly vertical sidewall angle, and later also as a 6-layered trapezoid in case of a sloped sidewall angle of 85 degrees (in which case all layers of the trapezoid are equally thick, thus 20nm each, and the mid-CD of the trapezoid then corresponds to the specified iso or dense widths of the single layered example). Finally the incident field is characterised by a linearly TE or TM polarised plane wave at a wavelength of 632.8nm and polar angle of 25 degrees (thus planar incidence). Although all geometri-

cal parameters are specified in nanometres here (because that is typically the preferred unit used in optical lithography), both methods internally normalise with respect to the wavelength. Figure 7.1 shows the absolute error in the diffraction efficiency, typically the quantity of interest since that is what a metrology tool often measures, as a function of the number of modes for the single layered resist case. The top figure on the left (right) corresponds to the 0^{th} diffraction order while the bottom figure on the left (right) corresponds to the 1^{st} diffraction order for the isolated (dense) resist pillar. The solid lines all correspond to incident TE polarisation whilst the dashed lines correspond to incident TM polarisation. The colours represent the different methods: RCWA (red), Bloch with homogeneous projection on plane waves (green) and Bloch with hybrid projection on mixed plane waves and Bloch modes (blue). The number of modes along the x -axis are actually *single-sided*, meaning that the *total* number of modes is always twice that plus one (also for Bloch). Finally the reference values are obtained using the hybrid Bloch approach with 400 modes. As can be seen from the figure the lines cor-

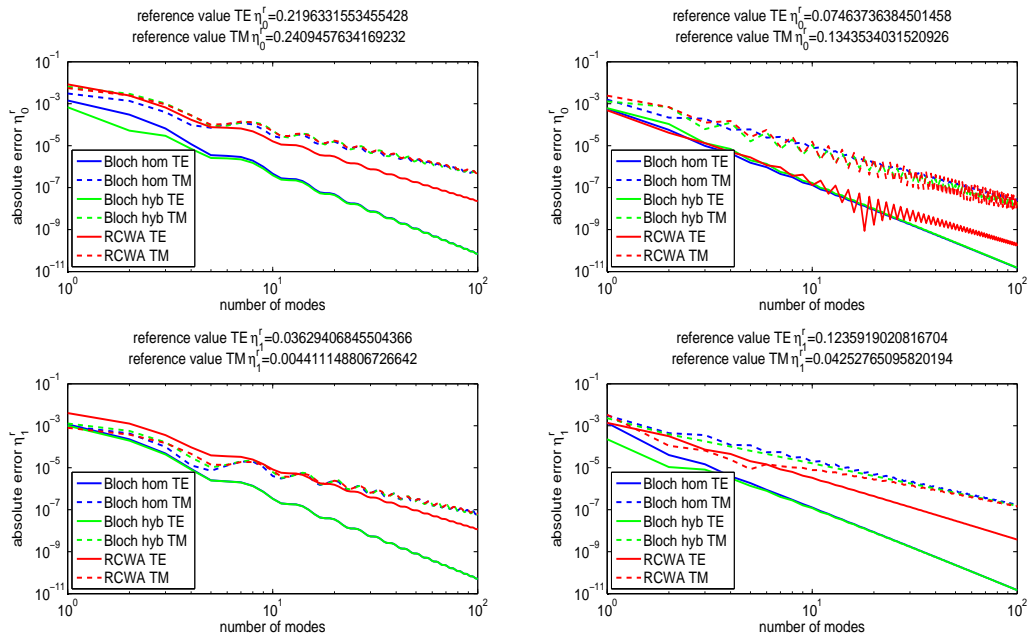


Figure 7.1: Absolute error in the diffraction efficiency of the 0^{th} (top) and 1^{st} (bottom) diffraction order for an isolated (left) and dense (right) single layer resist grating as a function of the number of modes. The solid lines correspond to incident TE polarisation, the dashed lines to incident TM polarisation and the colours represent the different discretisations Bloch homogeneous, Bloch hybrid and RCWA (blue, green, red).

responding to the homogeneous and hybrid Bloch mode method are most of the times indistinguishable. Note that even with only a single mode (thus for RCWA the harmonics run from $-1, \dots, 1$ and for Bloch 3 modes are used in total) already an error

of approximately 10^{-3} is obtained for both the isolated as well as the dense pillars and for both polarisations. This shows that low contrast applications like resist are indeed amongst the easier cases. Even so, when a higher accuracy is required the advantage of Bloch becomes clear, that is for TE polarisation. For this polarisation the Bloch lines not only show a slight offset with respect to the RCWA lines, also the slope is slightly steeper for Bloch. This means that when an accuracy of 10^{-6} is required the number of modes for Bloch is 3 times smaller than for RCWA. For TM polarisations however the differences between the methods almost disappears and Bloch does not show a real advantage anymore over RCWA. The different slopes between both polarisations can be explained from the different smoothness the field possesses. Recall that for TE polarisation the electric field was continuously differentiable across a material interface whereas for TM polarisation the magnetic field was only continuous across such an interface. This could explain why the results of Bloch and RCWA actually overlap for TM polarisation. Although Bloch is in principle capable of computing the real eigenmodes inside a layer whereas RCWA is only able to approximate them, Bloch cannot deal with the corner points in the resist pillar any better than RCWA can. For TM polarisation the magnetic field in these corner points is singular which could explain why both methods behave the same.

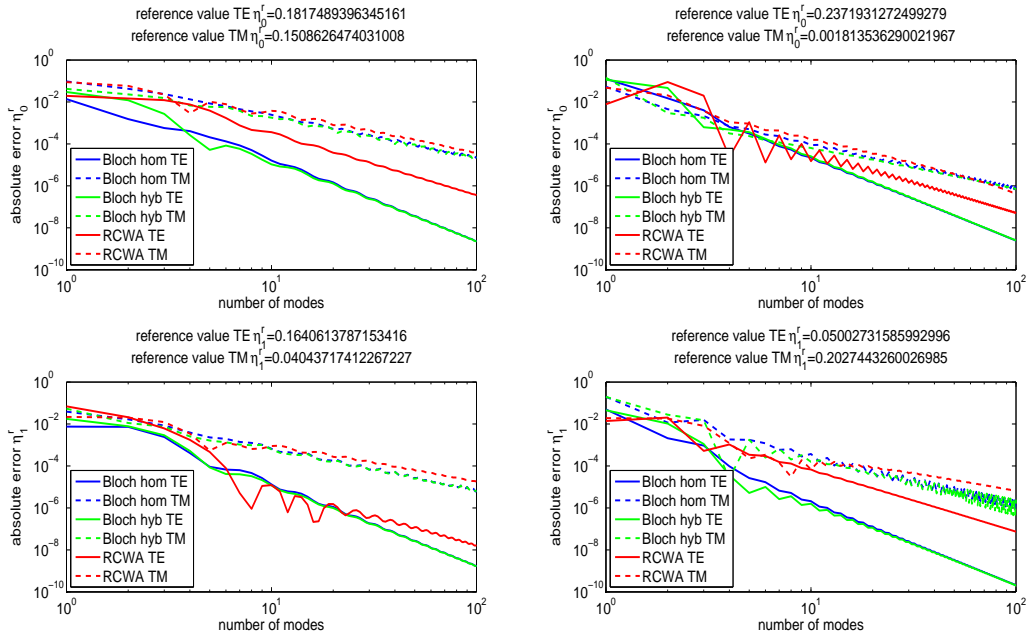


Figure 7.2: Absolute error in the diffraction efficiency of the 0th (top) and 1st (bottom) diffraction order for an isolated (left) and dense (right) single layer silicon grating as a function of the number of modes. The solid lines correspond to incident TE polarisation, the dashed lines to incident TM polarisation and the colours represent the different discretisations Bloch homogeneous, Bloch hybrid and RCWA (blue, green, red).

The silicon case consists of a silicon pillar directly lying on a silicon substrate with the superstrate simply air. The refraction indices of air and silicon are fixed at 1.0 and 3.88 respectively. The geometrical parameters like pitch, silicon width and silicon height are identical to the resist case and thus fixed at 500nm, 80nm or 250nm (iso or dense) and 120nm. Also now the silicon pillar is modelled as a single layer structure and later as 6-layered trapezoid in case of a perfectly vertical and sloped sidewall angle of 85 degrees respectively. Finally the incident field is also identical to the resist case, thus linearly TE and TM polarised at 632.8nm and incident under 25 degrees. Figure 7.2 is the counterpart of Figure 7.1 but now shows the absolute error in the diffraction efficiency for the silicon case. Overall the results are very similar to the resist case, although here the initial error when using only a single mode is in the order of $10^{-2} - 10^{-1}$. Not surprisingly the higher contrast in the refraction index for this silicon case requires more modes than the resist case to obtain a similar error level. When higher accuracy levels are required the use of Bloch modes clearly show an advantage over RCWA for TE polarisation, although again for TM polarisation the differences between both methods are minor.

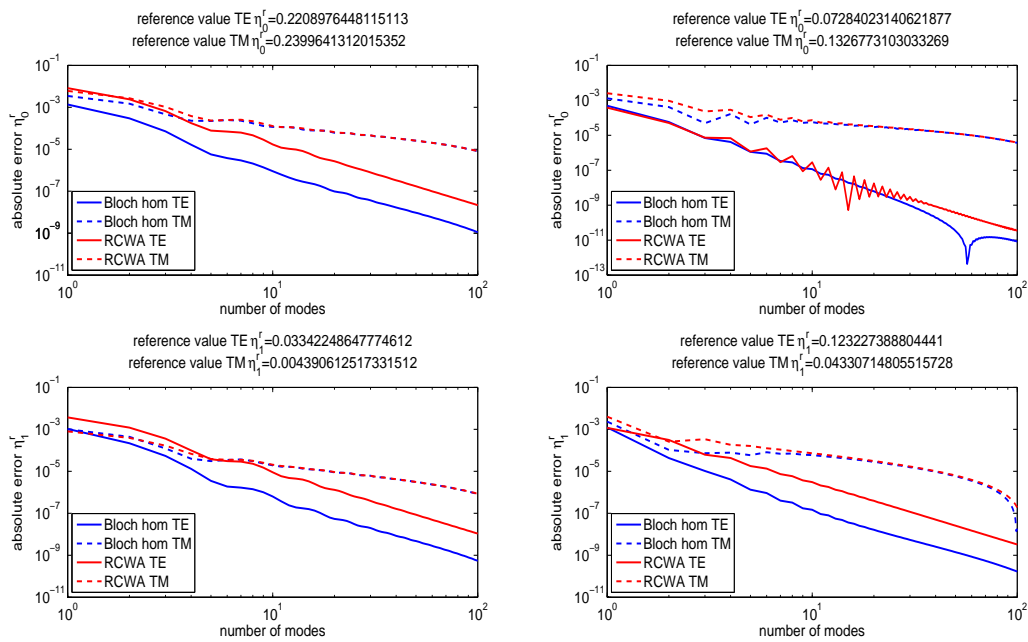


Figure 7.3: Absolute error in the diffraction efficiency of the 0th (top) and 1st (bottom) diffraction order for an isolated (left) and dense (right) 6-layer trapezoid resist grating as a function of the number of modes. The solid lines correspond to incident TE polarisation, the dashed lines to incident TM polarisation and the colours represent the different discretisations Bloch homogeneous and RCWA (blue, red).

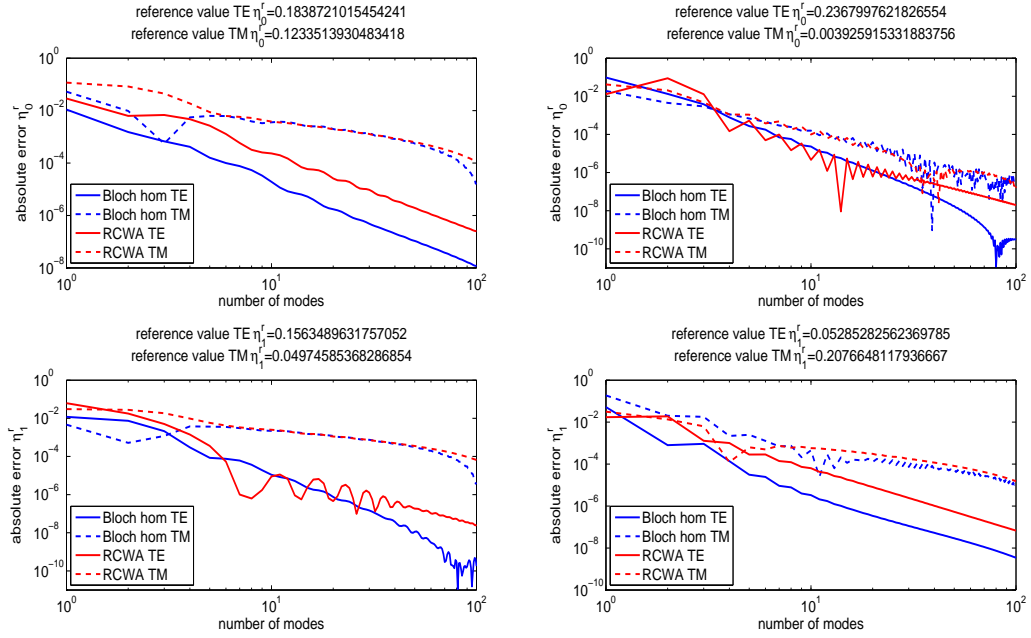


Figure 7.4: Absolute error in the diffraction efficiency of the 0th (top) and 1st (bottom) diffraction order for an isolated (left) and dense (right) 6-layer trapezoid silicon grating as a function of the number of modes. The solid lines correspond to incident TE polarisation, the dashed lines to incident TM polarisation and the colours represent the different discretisations Bloch homogeneous and RCWA (blue, red).

Having presented the results for a single layer pillar of resist and silicon, it is interesting to see that for a more realistic scenario of a 6-layered trapezoid the results roughly remain unchanged as can be seen in Figure 7.3 for the resist case and in Figure 7.4 for the silicon case. Since the single layer results showed hardly any difference between the hybrid and homogeneous Bloch approach, here only the homogeneous Bloch is shown (which is also used to generate the reference solution). Another reason for choosing the homogeneous Bloch approach is that for a trapezoid its implementation is somewhat easier, since the elements of the projection or coupling matrix are always computed in the same way and no Bloch to Bloch mode coupling needs to be computed. The fact that some graphs tend to curve downwards near the maximum number of 100 modes can be explained by the fact that the reference solution was not computed accurately enough. For the dense silicon case the error plot for the 0th TM order looks somewhat noisy but this seems to be the case for both the Bloch as well as RCWA results. Finally it is worth mentioning that these graphs by themselves confirm that the stable recursion with Riccati transformations as discussed in Chapter 5.2 indeed works and is required. The silicon case presented here is very similar to the input used to generate the data in Table 5.1a. Clearly using the *T-matrix algorithm* here would have been disastrous when more than 20 modes were used (let alone to compute a reference solution at 400 modes).

In Section 6.1 the technique adaptive spatial resolution (ASR) was introduced to improve RCWA. In Figure 7.5 the results of this technique are shown for the single layer silicon case. The ASR-stretching parameter is fixed at $G = 0.01$ in combination with the coordinate transformation of (6.3). These results can be directly compared with those in Figure 7.2 as can be seen from the identical reference values and RCWA curves. Furthermore the legend shows three types of ASR results (ASR, ASR+, ASR 2x) which correspond to three slightly different simulation settings or model inputs. The first (ASR)

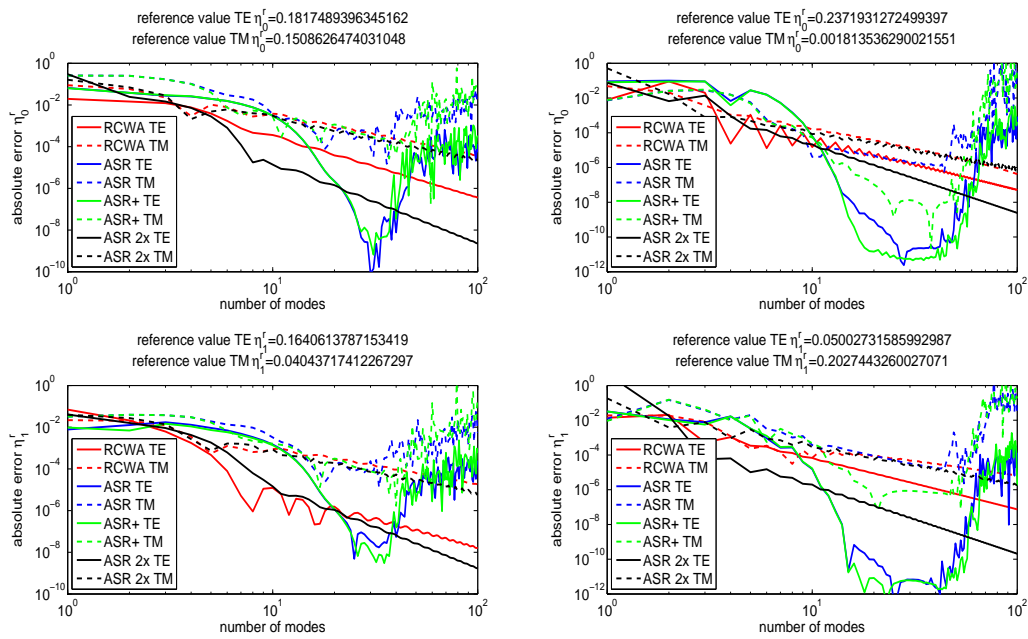


Figure 7.5: Absolute error in the diffraction efficiency of the 0th (top) and 1st (bottom) diffraction order for an isolated (left) and dense (right) single layer silicon grating as a function of the number of modes. The solid lines correspond to incident TE polarisation, the dashed lines to incident TM polarisation and the colours represent the different discretisations RCWA, ASR, ASR+, ASR 2x (red, blue, green, black).

type corresponds to the standard approach as described in Section 6.1 which means that only in the single silicon grating layer a coordinate transformation is used. The second (ASR+) type introduces two artificial homogeneous layers just above and below the single silicon grating layer. Thus in between the grating layer and superstrate of air there is an additional 10nm thick homogeneous layer of air in which two refinement points are chosen near the corner points of the underlying silicon pillar. Similarly in between the grating layer and silicon substrate an additional 10nm thick homogeneous layer of silicon is placed with again two refinement points near the corner points of the silicon pillar. So instead of a single layer in between two halfspaces, now three layers each with the same coordinate transformation are positioned in between two halfspaces. The idea

is that for TM polarisation the corner singularities in the field are modelled better when the refinement is also present just above and below these singularities. The third (ASR 2x) type actually does the same as the first type except that for the grating layer a 2 times larger eigenvalue problem is solved of which only half of the (converged) eigenvalues and eigenvectors are used. Several observations can be made from Figure 7.5:

- First note that upon close inspection of the ASR 2x results it becomes clear that these black curves are (almost) identical to the Bloch curves in Figure 7.2. Apparently the stretching is able to approximate the Bloch modes, which are the exact eigenfunctions and thus in a way optimal, much better than standard RCWA. This is supported by Figure 7.6 which shows the absolute error in the eigenvalues for a fixed number of 20 modes (thus for 41 modes in total). Whereas the standard RCWA method shows a clear difference between TE and TM polarisation, ASR approximates the eigenvalues almost equally well for both polarisations. Although ASR shows a deterioration in the larger eigenvalues, the smaller eigenvalues are approximated much better compared to RCWA. Since the smaller eigenvalues correspond to the most important modes, the improvement obtained there outweighs the increased error in the less important higher frequency modes. Moreover increasing the eigenvalue system in standard RCWA by a factor 2 does not give the same improvement as for ASR 2x.

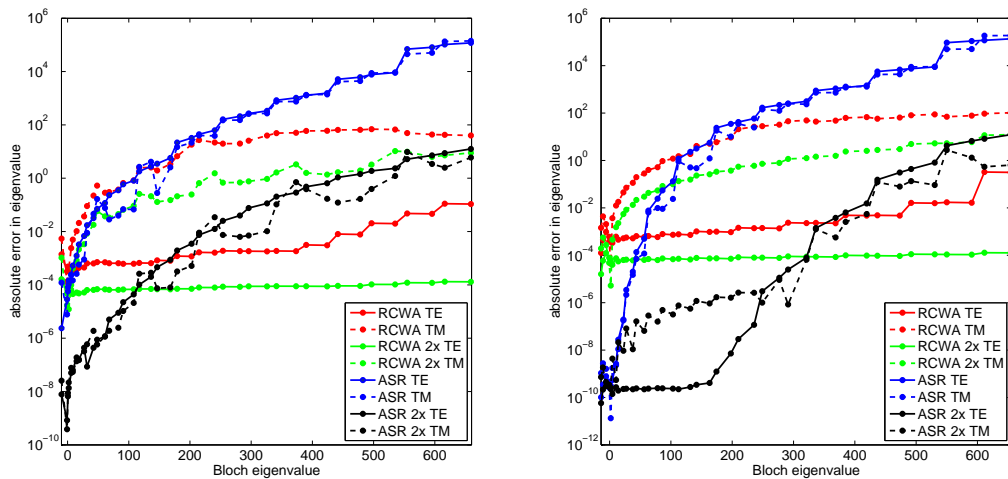


Figure 7.6: Absolute error in the eigenvalues for an isolated (left) and dense (right) single layer silicon grating. The reference values along the x-axis are obtained from the Bloch mode method. The solid lines correspond to incident TE polarisation, the dashed lines to incident TM polarisation and the colours represent the different discretisations RCWA, RCWA 2x, ASR, ASR 2x (red, green, blue, black). The number of modes is fixed at 20.

- Secondly what is perhaps most striking and not well understood is the plateau in Figure 7.2 (around 20-40 modes) that can be reached for TE polarisation with standard ASR. This plateau is slightly larger for the dense grating and in this regime the error in the diffraction efficiency is even smaller than obtained with Bloch. On the basis of the approximated eigenvalues it is not clear why ASR would do better than Bloch (or for that matter better than ASR 2x which seems to approximate Bloch very well). Moreover, increasing the number of modes beyond this plateau only gives unstable and inaccurate results. The reason for this behaviour is partly related to ill-conditioned matrices in the algorithm which can be traced back to the eigenvectors and eigenvalues that were not approximated well (and thus not a stability issue like the T -matrix algorithm). This plateau is *not* present when the alternative coordinate transformations are used from Appendix C which were missing the extra smoothness and fast decaying Fourier coefficients. Clearly a better understanding of this plateau could prove extremely useful since one is able to do better than the exact eigenfunctions from Bloch, and one does not have to solve a larger eigensystem and throw away half of the eigenvectors and eigenvalues. It should be pointed out that this plateau is not present for TM polarisation in combination with the standard ASR even though the error in the spectrum looks almost similar to TE polarisation.
- Thirdly the auxiliary layers introduced by ASR+ only seem to show a limited advantage. As expected for TE polarisation there is no improvement since also no corner singularities are expected there. However, for TM polarisation on the dense silicon stack again some sort of plateau can be observed where the error is smaller than both RCWA and Bloch. By construction RCWA and Bloch simply cannot refine near a corner whereas with the artificial layers of ASR+ one is able to capture the higher gradients in the field around sharp corners of the geometry. The isolated case shows that adding these artificial layers does not always improve the convergence of the diffraction efficiencies. An additional complication is introduced when a trapezoid grating is approximated with layers. Either multiple refinement points per layer should be introduced and/or additional thin layers could subdivide existing layers. For such a trapezoid the additional overhead generated by the extra layers quickly throws away most of the improved convergence, especially since for these trapezoid structures the plateau typically lies somewhat closer to the standard RCWA and Bloch results.
- Finally the ASR results obtained in Figure 7.2 for the silicon case are almost identical for the resist case. Although for resist the initial error was already pretty small, also here a plateau can be found for standard ASR in combination with TE polarisation reducing the error even further. Changing the angle of incidence also does not significantly change the outcome of the results, making a good coordinate transformation thus pretty robust for a wide range of incident angles. Alternative values for the stretching parameter G also did not improve the results any further and also for various other cases a value of 0.01 seemed to be rather optimal. Lastly the more general stretching of (6.1) did not improve the result of the isolated silicon pillar either (e.g. choosing the offsets S such that the transformed pillar effectively becomes dense).

In section 6.2 a second alternative to standard RCWA was discussed based on finite difference discretisations. Here we briefly summarise the results obtained with this method. Not having found a proper eigendecomposition routine that can exploit the sparseness in the coefficient matrix, this method is simply not competitive to ASR or even standard RCWA. For example repeating the test in Figure 7.6 for the dense silicon case using 42 meshpoints without refinement (which is roughly the equivalent of 20 single-sided modes resulting in the same size of the coefficient matrix and thus equal computational complexity) shows a graph for both polarisations with the same trend as RCWA TM. However, this graph lies roughly two orders above the RCWA graph meaning that the eigenvalues are approximated much worse. Moreover the TE polarisation now performs equally bad as TM whereas for standard RCWA the TE eigenvalues are approximated much better than TM. When adding a geometric refinement factor $G_I = G_{II} = 0.75$ which seems to perform the best for this geometry and number of meshpoints, the eigenvalues are approximated only slightly better (not even a full order on the smallest eigenvalues) and by no means comparable to the improvement seen with ASR. Although the geometric refinement technique resulted in a relatively simple structured coefficient matrix, perhaps other mesh refinements (which require a different finite difference stencil to be derived) might perform better and show a similar improvement as ASR. Nevertheless if one is not able to exploit the sparseness (except from a storage point of view of the coefficient matrix) this method will simply not be competitive enough.

7.2 Application of the forward diffraction model to grating reconstruction

As explained in the introduction a metrology tool [4,5] typically measures overlay or CD. Here we focus on the latter since measuring CD typically requires a close cooperation with Maxwell solvers. A simplified schematic overview of the metrology tool for a CD measurement can be found in Figure 7.7. A light source illuminates the periodic grating on the wafer after which its reflected signal or intensity pattern I is measured on a CCD camera. This intensity pattern is typically a function of the incident wavelength λ , angle of incidence (θ, ϕ) and polarisation ψ for a certain grating. The incident wavelength and polarisation can be selected using colour and polarisation filters in the illumination branch. Instead of using only a single angle of incidence and measuring spectroscopically (i.e. varying the wavelength continuously) here the wavelength is fixed and instead the angles of incidence are varied using a high numerical aperture (NA) microscope objective. Each point on the objective lens essentially corresponds to a unique angle of incidence. By illuminating the entire microscope a lot of different angles of incidence can be measured at the same time making high throughput times possible. After reflection of the grating the light passes back through the microscope objective and finally is detected on a camera.

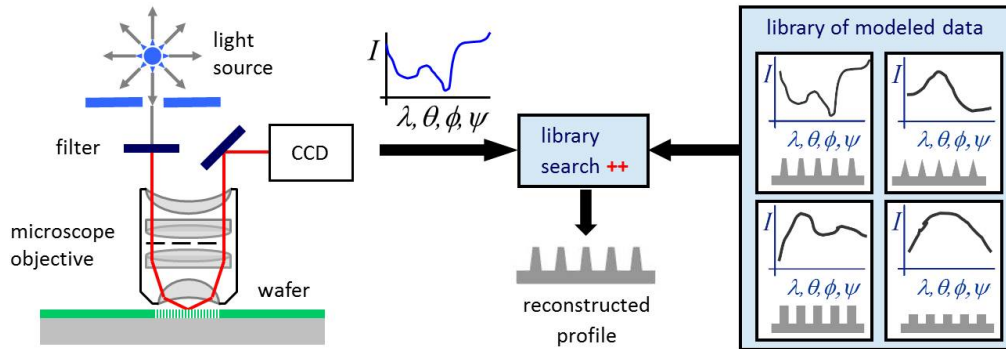


Figure 7.7: Schematic overview of metrology tool and reconstruction process

Unfortunately the measured signal does not directly translate into the CD parameters like height, width or sidewall angle of a grating. Instead this information is somehow encrypted in the measured intensity pattern. Therefore an additional post-processing step or reconstruction is required to extract or decrypt this information from the measurement. The right part of Figure 7.7 depicts such a reconstruction strategy which here simply consists of a library search. A library simply contains pre-computed intensity patterns for all kinds of different gratings with different CD configurations. A library search then looks for the closest match between the measured and computed intensity patterns in this library. The CD configuration that belongs to the best match then corresponds to the actual grating structure on the wafer. More advanced library search algorithms can be used that for example interpolate between patterns to get an even more accurate description of the grating. This library option clearly has an obvious drawback, namely it only works if a dense enough library (in combination with a smart sampling scheme) has been pre-computed. This does not necessarily have to be a problem if for example the library can be reused for multiple measurements. If this is not the case a second reconstruction strategy can be considered which consist of a direct reconstruction. A direct reconstruction or real-time regression typically minimises a cost function (e.g. the 2-norm of a vector where each entry corresponds to a camera pixel containing the difference of the measured and simulated intensity) using some optimisation technique. Some examples of these optimisation techniques include (damped) Gauss-Newton, Levenberg-Marquardt or Broyden-Fletcher-Goldfarb-Shann (BFGS) possibly extended with some form of regularisation. In [1] these optimisation techniques have been investigated in more detail for this specific grating reconstruction problem. Moreover [1] also discusses a technique to get CD parameter derivatives from RCWA very efficiently which are required by some of these optimisation techniques (clearly a finite difference approximation can also be used to approximate these derivatives at the cost of additional forward computations). Most of the times this direct reconstruction strategy is still supported by a small library that is used to extract an initial guess for the real-time regression. Clearly this library only needs to be coarse and therefore the setup time is much smaller than the full fledged library solution.

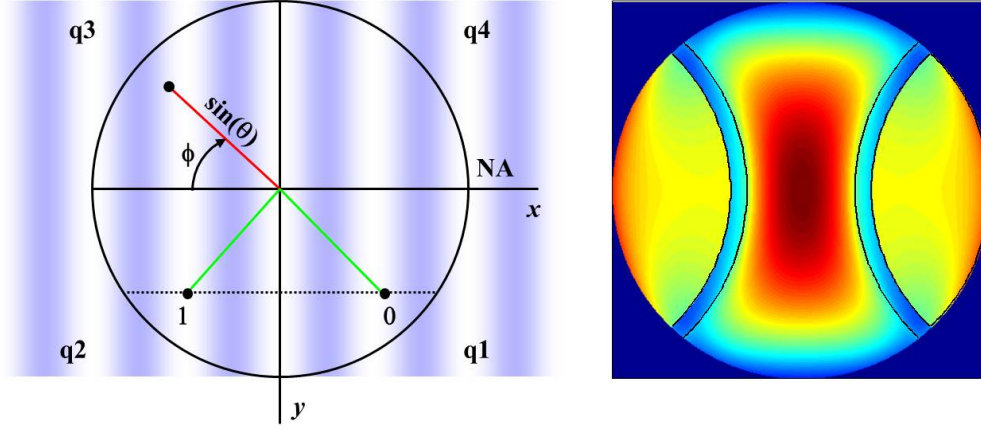


Figure 7.8: Single angle of incidence and corresponding outgoing diffraction orders (left). Superposition over all angles of incidence results in typical simulated intensity pattern (right).

It is clear that both reconstruction strategies require an accurate and fast forward model. Library fill times should remain in the order of hours whereas a real-time regressions is more in the order of seconds. In the remainder of this section we discuss in more detail how to compute these intensity patterns, how to incorporate the diffraction algorithms like RCWA or Bloch and how to exploit some additional symmetries of the problem. Figure 7.8 shows a top view of the microscope objective while looking down towards a grating which is aligned with the y -axis. As mentioned at the beginning of this section each point on the high NA objective lens corresponds to a unique angle of incidence and thus a single forward model computation. The angles of incidence (θ, ϕ) which also correspond to those in Figure 2.2 can be related to positions on the lens as indicated in the left part of Figure 7.8. After the incident plane wave interacts with the grating several diffraction orders are reflected and (some of) the propagating orders pass back through the lens. In the figure these are indicated with the 0th and 1st order. The location of these outgoing plane waves follows from tracking the outgoing wave vector. If for example the position of the incoming point corresponds to (x^i, y^i) and the position of the outgoing point of the m^{th} reflected order corresponds to (x_m^r, y_m^r) then

$$(x^i, y^i) = (-k_x, -k_y)/k_0, \quad (7.1a)$$

$$(x_m^r, y_m^r) = (k_{x_m}, k_y)/k_0, \quad (7.1b)$$

assuming the superstrate is air and $n_1 = 1$. Equation (2.44) gives the desired connection between the incoming and outgoing points

$$(x_m^r, y_m^r) = (-x^i - m\lambda_0/\Lambda, -y^i) \quad (7.1c)$$

In order to arrive at the final simulated intensity pattern one simply needs to superimpose the output from all individual incoming angles. Because all incoming angles are

independent from each other, they can be computed in parallel allowing for a dramatic speed improvement when using a multi-core hardware platform. After superposition one can clearly see ring-like areas in the intensity pattern. These rings mark out the areas where higher diffraction orders are located which also follows from the constant offset along the x -direction in (7.1c) for these higher orders.

Apart from the level of parallelism between different angles of incidence also additional symmetries can be exploited in the computation. For example when the grating is symmetric (which was the case in Figure 7.8) only one quadrant of the objective lens needs to be sampled, the others simply follow from mirror symmetry afterwards reducing the overall computation time by a factor 4. In case the grating is asymmetric only two quadrants (either q1 and q2 or alternatively q3 and q4) need to be sampled, the other two follow from mirror-symmetry in the x -axis in which case the computation time can be halved. Sometimes it is advantageous for the reconstruction to repeat the measurement for a second polarisation state (e.g. by rotating the polarisation filter). Then for each angle of incidence actually two forward problems are solved (note that because Maxwell's equations are linear at most two independent polarisations need to be computed, all other polarisation states follow from taking a proper linear combination). From an implementation point of view it is important to keep these two forward problems grouped together. This is because for all conical angles of incidence both polarisation states share the same transcendental eigenvalue equation in Bloch or equivalently solve the same eigenvalue problem in RCWA. This can also be seen from (4.29c) and (3.51c) where the incident polarisation only appears in the right-hand side vector \mathbf{d}_0 and not in the coefficient matrix of the linear system. It is actually possible to reuse the eigenvalue information (either Bloch or RCWA) in more places. For all incoming points that lie on a line parallel to the y -axis the eigenvalue problem is essentially the same apart from a constant shift related to the y -component of the wave vector (this has also been addressed earlier below Equations (3.18) and (4.11)). Thus by sampling the objective lens such that points fall on lines parallel to the y -axis, additional speed improvements are possible. Note that this trick does have certain implications when distributing work packages on a multi-core hardware platform. Apart from these symmetry properties that can speed up the computations for certain individual sample points, a final speed improvement can be realised by selecting as few sample points as possible. Clearly if one wishes to compare the simulated intensity pattern one-to-one with the measurement, the sample points must coincide with the camera pixels. Typically this is in the order of several hundreds of pixels along the radius of the microscope objective. Even with the speed improvements discussed above this number is simply too large to make real-time reconstruction possible. Therefore a subset of these pixels is chosen as representative sample points which can then be compared with some averaged (because of better signal to noise) measured pixels. If no a-priori knowledge is available on what a good subset could be, typically a uniform sampling of the microscope objective is suggested. If during the reconstruction more information becomes available (e.g. areas in the intensity pattern with large sensitivities or minimal correlation between different CD parameters) a smarter subset of points can be selected which could speed up the reconstruction or make it more robust.

Chapter 8

Conclusions and future work

In this thesis two mode expansion methods are presented that solve the plane wave incident scattering problem of periodic (staircase approximated) gratings. The Bloch mode method describes the electromagnetic field in the exact eigenfunctions of each grating layer. By using a symmetric unit cell description two clearly separated sub-problems can be derived, the periodic eigenvalue problem and semi-periodic eigenvalue problem. For lossless symmetric grating layers an efficient strategy is described to compute all eigenvalues of the (semi-)periodic eigenvalue problem without running the risk of skipping one. The individual layers can still be stacked on top of each other to create a more general asymmetric profile. Moreover the extension to deal with multiple material interfaces inside a (symmetric) grating layer has been explained, thereby enabling more complicated gratings to be modelled. However, generalising and implementing the Bloch mode method for lossy and/or asymmetric grating layers has proven to be rather difficult, which led us to investigate an alternative mode expansion method. The Rigorous Coupled-Wave Analysis describes the grating profile (and electromagnetic field) in (pseudo-periodic) Fourier series for each grating layer. This industry standard method is much more flexible than Bloch since its implementation does not differentiate between lossy, lossless, symmetric or asymmetric gratings.

Both methods after computing the (approximate) eigenfunctions of each grating layer match the solution at the grating layer interfaces. The resulting linear system has a large but sparse block structured coefficient matrix. Standard transfer matrix algorithms are not able to solve the condensed linear system stably. Other condensation algorithms like for example the enhanced transmittance matrix approach or S -matrix algorithm are stable but in the derivation the key aspect of separating the exponentially growing and decaying terms is often missing. Therefore a decoupling algorithm based on Riccati transformations has been presented to condense and solve this linear system stably and efficiently. This technique has been used before in a much wider context of solving

general boundary value problems and its stability results have a solid mathematical foundation. Two implementation variants of the Riccati approach are derived of which the two-stage variant is the most efficient. Moreover from an algorithmic point of view we have shown that the Riccati approach can be linked one-to-one to the frequently used enhanced transmittance matrix approach (thereby confirming its stability as well).

Although RCWA is more flexible than Bloch, the sharp material interfaces inside a grating layer (i.e. the piecewise constant geometry description) are not that well approximated using truncated Fourier series. For both isolated and dense gratings of high and low contrast Bloch typically outperforms (at least is not worse than) RCWA in the sense that with the same number of modes a higher accuracy is obtained. Not wanting to give up on the extra flexibility and relatively straightforward implementation of RCWA, two modifications have been investigated that could improve its convergence. Adaptive Spatial Resolution applies an additional layer specific coordinate transformation before Fourier discretising the problem again like standard RCWA. By increasing the resolution with this transformation near a material interface inside a grating layer, a higher accuracy is obtained. Numerical examples show that a good coordinate transformation should not only refine near an interface but also do so in a smooth way. Moreover we have shown that ASR approaches the accuracy of Bloch when only a subset of the (converged) eigenvalues are used. However, numerical tests have also revealed some unexpected behaviour of ASR. In particular for TE polarisation a dramatic improvement in the accuracy is observed which remains present over a wide range of inputs. For TM polarisation also some improvement is observed when adding artificial layers to the staircase approximated grating. These extra layers allow for extra refinement around the corners of the grating where singularities are expected in the field. The second modification is based on a finite difference discretisation that completely replaces the Fourier discretisation. This approach is able to properly take care of the material transitions inside a grating layer by incorporating them in the FD-stencils. Unfortunately the sparsity of the resulting coefficient matrix could not be exploited to make this implementation competitive to standard RCWA or Bloch. Only when combined with either an efficient sparse-eigendecomposition routine or alternatively a numerical integrator which uses an efficient matrix-vector multiplication could this approach perhaps become competitive.

Finally we have shown how these diffraction models are combined with a metrology tool when applied to a grating CD reconstruction application. Because such a reconstruction, whether based on libraries or real-time regression, typically requires a lot of calls to these forward diffraction models, it is essential to have the most accurate and fastest implementation available allowing for high throughput times. Apart from simply having a fast forward diffraction model, symmetries in the grating geometry can be exploited to minimise the number of calls. Also we have shown how to exploit additional symmetries (or opportunities to reuse information) inside the diffraction models themselves. Integration of this work into the overall CD reconstruction software of the metrology tool has contributed in making real-time (typically in the order of seconds) grating CD reconstruction possible on a multi-core hardware platform.

Several topics have actually raised new questions that would require more research but in the end could result in even faster throughput times for the metrology tool. For example, Bloch typically shows a comparable or faster convergence when compared to RCWA. The main challenge would be to implement a robust root finding algorithm for the transcendental eigenvalue equation to deal with both lossy (i.e. find complex-valued roots) and asymmetric gratings (i.e. separate roots when close together). Alternatively one could also invest in further developing the ASR technique. Although smoothness of the coordinate transformation plays an important role, a more extensive search for alternative transformations (or tuning of the existing transformation parameters) could improve the accuracy even more. Moreover, the unexpected but dramatic improvement for TE polarisation (even outperforming Bloch) is still not understood. Insights into this mechanism could prove extremely useful when high precision is required at minimal computational cost. Finally CD reconstruction on 2D periodic gratings has also received a great deal of attention. For Bloch modes this extension might not be easy to derive (or even feasible) except for perhaps very specific structure like holes in perfectly electric conductors. For RCWA we already implemented a 2D periodic version that can deal with staircase approximated rectangular like structures (i.e. the material interfaces of the geometry inside a grating layer should be aligned with one of the two vectors spanning the orthogonal unit cell). Since the computational complexity of the algorithm is now $O(M^6)$, a real-time reconstruction with RCWA would still be challenging. In order to reduce computation time and memory usage it is therefore essential to exploit as much symmetry as possible. Also the combination with ASR could further improve the convergence.

Appendix A

Quantities related to the Rayleigh expansion

A.1 Derivation of the perpendicular s -polarised and p -polarised parts

The tangential components in the direction perpendicular to \mathbf{s}_m and \mathbf{p}_m can be obtained in several steps. These steps make use of some vector identities and previously derived relations. For example, for the tangential component of the electric field the perpendicular s -polarised part is given by

$$\begin{aligned} \mathbf{R}_m^e \cdot (\mathbf{s}_m \times \mathbf{e}_z) &= -\frac{1}{k_0 n_1} (\mathbf{k}_m^r \times \mathbf{R}_m^h) \cdot (\mathbf{s}_m \times \mathbf{e}_z) \\ &= -\frac{1}{k_0 n_1} \mathbf{e}_z \cdot \left((\mathbf{k}_m^r \times \mathbf{R}_m^h) \times \mathbf{s}_m \right) \\ &= -\frac{1}{k_0 n_1} \mathbf{e}_z \cdot \left(\mathbf{R}_m^h (\mathbf{k}_m^r \cdot \mathbf{s}_m) - \mathbf{k}_m^r (\mathbf{R}_m^h \cdot \mathbf{s}_m) \right) \\ &= \frac{1}{k_0 n_1} R_{pm} \mathbf{e}_z \cdot \mathbf{k}_m^r = -\frac{k_{1,zm}}{k_0 n_1} R_{pm}. \end{aligned} \tag{A.1}$$

The first equality simply follows from relation (2.51a) whereas the second equality is the result of applying the vector identity $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})$. Applying another vector identity $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{A}(\mathbf{B} \cdot \mathbf{C})$ gives the third equality. Finally if one realises that (2.52) is simply $\arctan(k_y/k_{xm})$ but extended to four quadrants, then the last equality is obtained from the fact that $\mathbf{k}_m^r \cdot \mathbf{s}_m = 0$. The other perpendicular components can be derived in a similar way.

A.2 Derivation of the energy flow

The time-averaged energy flow for the general conical diffraction case in terms of the s -polarised and p -polarised parts can be obtained in several steps. These steps involve introducing a new orthogonal coordinate system, applying some vector identities and making use of previously derived relations. For example, the energy of a reflected diffraction order in the superstrate in the direction of $-\mathbf{e}_z$ is given by

$$\begin{aligned}
 P_{0,m}^r &= \int_0^1 \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} \mathbf{S}_{0,m}^r \cdot -\mathbf{e}_z \, dx dy \\
 &= -\frac{1}{2} Y_I \Lambda e^{-2\text{Im}[k_{1,zm}]z} \left(\mathbf{R}_m^e \times \overline{\mathbf{R}_m^h} \right) \cdot \mathbf{e}_z \\
 &= -\frac{1}{2} Y_I \Lambda e^{-2\text{Im}[k_{1,zm}]z} \frac{1}{k_0 n_I} \left(\mathbf{R}_m^e \times \overline{(\mathbf{k}_m^r \times \mathbf{R}_m^e)} \right) \cdot \mathbf{e}_z, \tag{A.2}
 \end{aligned}$$

where equations (2.59a) and (2.51a) were used in deriving the second and third equality respectively. In order to arrive at an expression in terms of the s -polarised and p -polarised parts, the cross products are computed with respect to the orthogonal coordinate system $(\mathbf{s}_m \times \mathbf{e}_z, \mathbf{s}_m, \mathbf{e}_z)$. Therefore the wave vector and Rayleigh amplitude need to be rewritten in terms of this new coordinate system

$$\begin{aligned}
 \mathbf{k}_m^r &= (\mathbf{k}_m^r \cdot (\mathbf{s}_m \times \mathbf{e}_z)) (\mathbf{s}_m \times \mathbf{e}_z) + (\mathbf{k}_m^r \cdot \mathbf{s}_m) \mathbf{s}_m + (\mathbf{k}_m^r \cdot \mathbf{e}_z) \mathbf{e}_z \\
 &= (\mathbf{k}_m^r \cdot (\mathbf{s}_m \times \mathbf{e}_z)) (\mathbf{s}_m \times \mathbf{e}_z) - k_{1,zm} \mathbf{e}_z, \tag{A.3a}
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{R}_m^e &= (\mathbf{R}_m^e \cdot (\mathbf{s}_m \times \mathbf{e}_z)) (\mathbf{s}_m \times \mathbf{e}_z) + (\mathbf{R}_m^e \cdot \mathbf{s}_m) \mathbf{s}_m + (\mathbf{R}_m^e \cdot \mathbf{e}_z) \mathbf{e}_z \\
 &= -\frac{k_{1,zm}}{k_0 n_I} R_{pm} (\mathbf{s}_m \times \mathbf{e}_z) + R_{sm} \mathbf{s}_m + (\mathbf{R}_m^e \cdot \mathbf{e}_z) \mathbf{e}_z. \tag{A.3b}
 \end{aligned}$$

Here in the last line of (A.3a) the relation $\mathbf{k}_m^r \cdot \mathbf{s}_m = 0$ was used as was already derived in Appendix A.1. Moreover in the last line of (A.3b) the definition of the s -polarised part (2.53a) and the relation for the perpendicular s -polarised part (2.54a) were substituted. The last term in (A.3b) still needs to be expressed in terms of the s -polarised and p -polarised parts which is accomplished by following the same steps as in Appendix A.1

$$\begin{aligned}
 \mathbf{R}_m^e \cdot \mathbf{e}_z &= -\frac{1}{k_0 n_I} \left(\mathbf{k}_m^r \times \mathbf{R}_m^h \right) \cdot ((\mathbf{s}_m \times \mathbf{e}_z) \times \mathbf{s}_m) \\
 &= -\frac{1}{k_0 n_I} (\mathbf{s}_m \times \mathbf{e}_z) \cdot \left(\mathbf{s}_m \times \left(\mathbf{k}_m^r \times \mathbf{R}_m^h \right) \right) \\
 &= -\frac{1}{k_0 n_I} (\mathbf{s}_m \times \mathbf{e}_z) \cdot \left(\mathbf{k}_m^r (\mathbf{s}_m \cdot \mathbf{R}_m^h) - \mathbf{R}_m^h (\mathbf{s}_m \cdot \mathbf{k}_m^r) \right) \\
 &= -\frac{1}{k_0 n_I} R_{pm} (\mathbf{k}_m^r \cdot (\mathbf{s}_m \times \mathbf{e}_z)). \tag{A.4}
 \end{aligned}$$

In the first equality (2.51a) and the orthogonality of the new coordinate system were used followed by the vector identity $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})$ in the second equality. Applying another vector identity $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ gives the third equality. The last equality is again obtained from the fact that $\mathbf{s}_m \cdot \mathbf{k}_m^r = 0$. Now all the ingredients are there to compute the cross products with respect to the new coordinate system

$$\begin{aligned} \mathbf{k}_m^r \times \mathbf{R}_m^e &= k_{I,zm} R_{sm} (\mathbf{s}_m \times \mathbf{e}_z) + \frac{1}{k_0 n_I} R_{pm} \left(k_{I,zm}^2 + (\mathbf{k}_m^r \cdot (\mathbf{s}_m \times \mathbf{e}_z))^2 \right) \mathbf{s}_m + \\ &\quad R_{sm} (\mathbf{k}_m^r \cdot (\mathbf{s}_m \times \mathbf{e}_z)) \mathbf{e}_z \\ &= k_{I,zm} R_{sm} (\mathbf{s}_m \times \mathbf{e}_z) + k_0 n_I R_{pm} \mathbf{s}_m + R_{sm} (\mathbf{k}_m^r \cdot (\mathbf{s}_m \times \mathbf{e}_z)) \mathbf{e}_z, \end{aligned} \quad (\text{A.5})$$

where the term within the largest set of brackets is nothing more than the squared absolute value of the wave vector and thus equal to $k_0^2 n_I^2$. Finally the total cross product becomes

$$\begin{aligned} \left(\mathbf{R}_m^e \times \overline{(\mathbf{k}_m^r \times \mathbf{R}_m^e)} \right) \cdot \mathbf{e}_z &= -\frac{k_{I,zm}}{k_0 n_I} R_{pm} \overline{k_0 n_I R_{pm}} - R_{sm} \overline{k_{I,zm} R_{sm}} \\ &= -\overline{k_{I,zm}} |R_{sm}|^2 - k_{I,zm} |R_{pm}|^2. \end{aligned} \quad (\text{A.6})$$

The energy flow is thus given by

$$P_{0,m}^r = \frac{1}{2} Y_I \Lambda e^{-2\text{Im}[k_{I,zm}]z} \left(\frac{\overline{k_{I,zm}}}{k_0 n_I} |R_{sm}|^2 + \frac{k_{I,zm}}{k_0 n_I} |R_{pm}|^2 \right). \quad (\text{A.7})$$

Following the same steps an expression can be derived for the energy flow of a transmitted diffraction order in the substrate in the direction of \mathbf{e}_z

$$P_{K+1,m}^t = \frac{1}{2} \overline{Y_{II}} \Lambda e^{2\text{Im}[k_{II,zm}](z-Z_K)} \left(\left(\frac{\overline{k_{II,zm}}}{k_0 n_{II}} \right) |T_{sm}|^2 + \frac{k_{II,zm}}{k_0 n_{II}} |T_{pm}|^2 \right). \quad (\text{A.8})$$

Appendix B

ASR coupling matrix

The coefficients of the coupling matrix in (6.7) can be computed in several ways. Here three different approaches with some of their (dis)advantages are discussed for the general coordinate transformation in (6.1). They are based on either numerical quadrature, FFT computations or the evaluation of Bessel related special functions. It is worth mentioning that the first two approaches do not really require any information on the specific coordinate transformation and can therefore also be applied to transformations other than (6.1). First the integral can be simplified somewhat by applying partial integration

$$\begin{aligned} \mathbf{S}_{i,pm} &= \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} h_i(s) e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds \\ &= \frac{1}{\Lambda} \frac{1}{jk_{xp}} \left[e^{-jk_{xm}s} e^{jk_{xp}x(s)} \right]_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} + \frac{1}{\Lambda} \frac{k_{xm}}{k_{xp}} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds, \end{aligned} \quad (\text{B.1})$$

for $k_{xp} \neq 0$. When $k_{xp} = 0$ there is no need to apply partial integration since then the original integral is trivial. Therefore in the remainder of this section $k_{xp} \neq 0$ and one can focus on the integral appearing in the right-hand side of (B.1).

B.1 Numerical quadrature

Probably an obvious choice when computing the integral in (B.1) is simply to approximate it using some sort of numerical quadrature. Naturally all kinds of different quadratures can be evaluated but in this thesis we have not attempted to find an optimal quadrature with optimal settings. Instead, having tried a couple, the best results were

obtained with Matlab's built in 'quad' function which uses a recursive adaptive Simpson quadrature. It turns out that even for a simple coordinate transformation with only two jumps as given by (6.3) this numerical quadrature is (much) more expensive than the other two approaches discussed in B.2 and B.3. Partly this is due to the fact that the numerical quadrature has to be restarted for all coefficients of the coupling matrix since the integrand simply changes. Moreover the error tolerance needs to be set pretty strict in order to get enough correct digits. For example, in order to get 6 digits matching with the other approaches for the simple binary example the absolute error tolerance was set to 10^{-6} . Finally it should be noted that the integrand consists of complex-valued exponentials and sines making a function evaluation in the numerical quadrature not that cheap. So although the overall complexity scales with M^2 , the time spent in the numerical quadrature makes this approach not very attractive.

B.2 Fast Fourier Transform

Another way of computing the integral in (B.1) is by making use of FFTs. To this end rewrite the integral as follows

$$\frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds = \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{j(k_{xp}x(s) - k_0 n_1 \sin \theta s)} e^{j2\pi ms/\Lambda} ds. \quad (\text{B.2})$$

This means that the integral can also be interpreted as computing the Fourier coefficients of the function $e^{j(k_{xp}x(s) - k_0 n_1 \sin \theta s)}$. By sampling this function equidistantly on the interval $[-\frac{\Lambda}{2}, \frac{\Lambda}{2}]$ and using an FFT library routine one can approximate these Fourier coefficients. For a fixed value of p and thus for a complete row in the coupling matrix only one FFT is required! Therefore the overall complexity of computing the coupling matrix is $O(M^2 \log M)$ resulting in a very efficient implementation. For the simple coordinate transformation of (6.3) the integral was computed accurately up to at least 6 correct digits by making the FFT size (and thus the amount of samples) at least twice as large as the size of the coupling matrix. In order to fully exploit the speed of an FFT library routine one should make sure that the FFT size equals a power of two or the product of small prime numbers. Clearly this approach requires less implementation effort than the numerical quadrature and as a bonus it is also much faster.

B.3 Bessel related special functions

A third way of computing the integral in (B.1) is by making use of Bessel related special functions. However this approach only applies to the specific coordinate transformation

of (6.1). First the integral is split up into the following sum

$$\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds = \sum_{l=1}^{L_i-1} \int_{S_{i,l-1}}^{S_{i,l}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds, \quad (\text{B.3})$$

where the focus now lies on one of these new integrals with bounds between two off-sets in the transformed space. Substituting the specific coordinate transformation while dropping, for notational convenience, the subscript i denoting the layer index gives

$$\begin{aligned} \int_{S_{l-1}}^{S_l} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds &= e^{jk_{xp}\alpha_l} \int_{S_{l-1}}^{S_l} e^{j(k_{xp}\beta_l - k_{xm})s + jk_{xp} \frac{\gamma_l}{2\pi} \sin\left(2\pi \frac{s-S_{l-1}}{S_l-S_{l-1}}\right)} ds \\ &= e^{jk_{xp}\alpha_l} \frac{S_l - S_{l-1}}{2\pi} \int_0^{2\pi} e^{j(k_{xp}\beta_l - k_{xm})\left(\frac{S_l - S_{l-1}}{2\pi}t + S_{l-1}\right) + jk_{xp} \frac{\gamma_l}{2\pi} \sin t} dt \\ &= e^{j(k_{xp}\alpha_l + (k_{xp}\beta_l - k_{xm})S_{l-1})} \frac{S_l - S_{l-1}}{2\pi} \int_0^{2\pi} e^{j(\nu_{l,pm}t + z_{l,p} \sin t)} dt \end{aligned} \quad (\text{B.4a})$$

where

$$\nu_{l,pm} = (k_{xp}\beta_l - k_{xm}) \frac{S_l - S_{l-1}}{2\pi}, \quad z_{l,p} = k_{xp} \frac{\gamma_l}{2\pi}. \quad (\text{B.4b})$$

When $\gamma_l = 0$ this implies that $z_{l,p} = 0$ making the integral in (B.4a) again trivial. Therefore assume that the coordinate transformation is not trivial and thus $z_{l,p} \neq 0$. In this case a standard integral can be obtained by splitting up once more the integration interval

$$\begin{aligned} \int_0^{2\pi} e^{j(\nu_{l,pm}t + z_{l,p} \sin t)} dt &= \int_0^{\pi} e^{j(\nu_{l,pm}t + z_{l,p} \sin t)} dt + \int_{\pi}^{2\pi} e^{j(\nu_{l,pm}t + z_{l,p} \sin t)} dt \\ &= \int_0^{\pi} e^{j(\nu_{l,pm}t + z_{l,p} \sin t)} dt + e^{j\nu_{l,pm}\pi} \int_0^{\pi} e^{j(\nu_{l,pm}t - z_{l,p} \sin t)} dt \end{aligned} \quad (\text{B.5})$$

Since both $\nu_{l,pm}$ and $z_{l,p}$ are real-valued one can relate the integrals appearing in (B.5) to so-called Anger and Weber functions

$$\mathbf{J}_\nu(z) = \int_0^{\pi} \frac{\cos(\nu t - z \sin t)}{\pi} dt, \quad (\text{B.6a})$$

$$\mathbf{E}_\nu(z) = \int_0^{\pi} \frac{\sin(\nu t - z \sin t)}{\pi} dt, \quad (\text{B.6b})$$

so that

$$\mathbf{J}_\nu(z) + j\mathbf{E}_\nu(z) = \int_0^{\pi} \frac{e^{j(\nu t - z \sin t)}}{\pi} dt. \quad (\text{B.6c})$$

Using the symmetry properties of these special functions only $\nu \geq 0$ needs to be considered. When ν is an integer the Anger function reduces to the classical Bessel function

whereas the Weber function can then be related to a Struve function and a sum with Gamma functions. For non-integer values of ν either power series for small arguments or asymptotic expansions for large arguments can be used. Here again several special functions are required namely the Lommel functions, Lommel functions of the second kind, Bessel functions of the first and second kind of fractional order. A more detailed overview of the Anger and Weber functions and a possible implementation in C can be found in [35]. Also here the overall complexity scales with M^2 because each coefficient of the coupling matrix results in different values for $\nu_{l,pm}$ and $z_{l,p}$. Although $z_{l,p}$ remains constant for a complete row in the coupling matrix, this unfortunately does not result in a lot of reuse. Moreover, large values of $\nu_{l,pm}$ and $z_{l,p}$ (for example when increasing the number of harmonics) sometimes resulted in inaccurate results where only 3 significant digits were obtained. This could possibly be alleviated somewhat by having a closer look at the specific implementation and used formulas. All things considered it turns out that the amount of function calls per coefficient of the coupling matrix make this third approach not competitive enough. Although faster than the numerical quadrature discussed in B.1 the implementation based on [35] cannot compete with the FFT implementation in B.2 which is faster, more accurate and easier to implement.

Appendix C

Alternative ASR coordinate transformations

The coordinate transformation in (6.1) is of course not unique but already shows very good results. Without going into an optimisation process of finding the best possible coordinate transformation, it is important to find certain criteria that, if satisfied, result in a good transformation. Because of the larger gradients in the field near a material interface, one of these criteria should be that the coordinate transformation refines near such an interface resulting in a better approximation when Fourier series are used. Clearly (6.1) does exactly that and the two alternatives discussed here also satisfy this criterium.

The first alternative tries to mimic the exact same behaviour as the original coordinate transformation but uses only piecewise low order polynomials. The reason for doing so is that then the coupling coefficients can be computed slightly easier while trying to preserve the same good convergence. The transformation is given by

$$x(s) = \alpha_{i,l} + \beta_{i,l}s + 2\gamma_{i,l} \frac{s - S_{i,l-\frac{1}{2}}}{S_{i,l} - S_{i,l-1}} \left(\left| \frac{s - S_{i,l-\frac{1}{2}}}{S_{i,l} - S_{i,l-1}} \right| - \frac{1}{2} \right),$$

for $s \in (S_{i,l-1}, S_{i,l})$, (C.1)

where $S_{i,l-\frac{1}{2}} = (S_{i,l} + S_{i,l-1})/2$ and where the coefficients $\alpha_{i,l}$, $\beta_{i,l}$ and $\gamma_{i,l}$ are the same as in (6.2). A plot of this coordinate transformation would look the same as in Figure 6.1. The coupling coefficients can be computed using the same strategies as in Appendix B except that no Bessel related special functions are required. Now equation (B.3) can be

rewritten into

$$\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds = \sum_{l=1}^{L_i-1} \left(\int_{S_{i,l-1}}^{S_{i,l}-\frac{1}{2}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds + \int_{S_{i,l}-\frac{1}{2}}^{S_{i,l}} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds \right). \quad (\text{C.2})$$

Focusing on one term in this sum and substituting the specific coordinate transformation into the second integral (the first integral can be simplified in a similar way) while dropping the subscript i gives

$$\int_{S_{l-\frac{1}{2}}}^{S_l} e^{-jk_{xm}s} e^{jk_{xp}x(s)} ds = e^{j(k_{xp}\alpha_l + (k_{xp}\beta_l - k_{xm})\frac{1}{2}S_{l-\frac{1}{2}})} \frac{S_l - S_{l-1}}{2} \int_0^1 e^{j(\nu_{l,pm}t + z_{l,p}t^2)} dt \quad (\text{C.3a})$$

where

$$\nu_{l,pm} = (k_{xp}\beta_l - k_{xm}) \frac{S_l - S_{l-1}}{2} - k_{xp} \frac{\gamma_l}{2}, \quad z_{l,p} = k_{xp} \frac{\gamma_l}{2}. \quad (\text{C.3b})$$

For a non-trivial coordinate transformation with $\gamma_l \neq 0$ and thus also $z_{l,p} \neq 0$ the remaining integral in (C.3a) can be linked to the complex-valued error function or its related Fresnel integrals

$$C(z) = \int_0^z \cos\left(\frac{\pi}{2}t^2\right) dt, \quad (\text{C.4a})$$

$$S(z) = \int_0^z \sin\left(\frac{\pi}{2}t^2\right) dt, \quad (\text{C.4b})$$

so that

$$C(z) + jS(z) = \int_0^z e^{j\frac{\pi}{2}t^2} dt. \quad (\text{C.4c})$$

These Fresnel integrals can be implemented somewhat easier than the Anger and Weber functions and do not require any additional special functions [35]. However, still based on asymptotic expansions and an overall complexity of M^2 (because each coefficient of the coupling matrix results in a different integration interval) makes this implementation still not as competitive as the FFT trick discussed in B.2. Figure C.1 illustrates the effect of Fourier transforming the different ASR related functions for a typical grating layer using the alternative coordinate transform. These results can be compared directly with those of Figure 6.2. The original transformation based on a sine is smoother than the piecewise low order polynomial transformation resulting in a faster drop-off of the Fourier coefficients. Increasing the stretching seems to have much less impact on the Fourier coefficients with the alternative transformation.

The second alternative tries to accommodate certain specific Bloch modes which are the exact eigenfunctions that standard RCWA tries to approximate as was explained just below (4.8). Recall that in Section 3.2 it was explained that for a lossless binary grating as depicted in Figure 2.3 with $n_{II} > n_I$ that κ_{II} is always positive real-valued and κ_{Ii} is

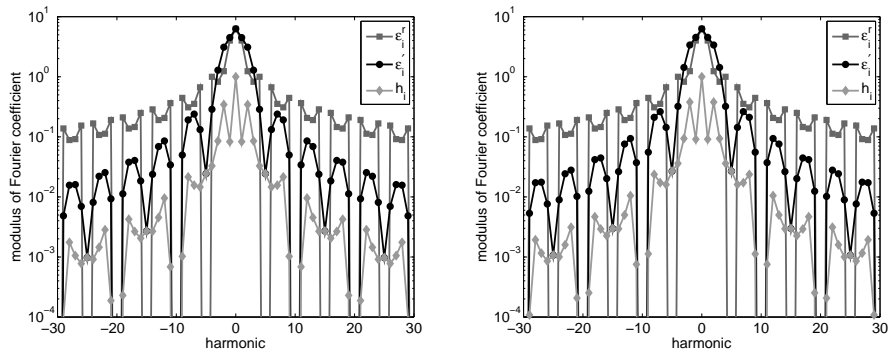


Figure C.1: Modulus of the Fourier coefficients of the permittivity function, scaled permittivity function and derivative of the coordinate transformation. The grating layer corresponds to a symmetric binary grating with $n_{\text{I}} = 1$, $n_{\text{II}} = 3.77$, $X_i = 0.2$, $\Lambda = 1$ and $G_i = 0.1$ (left), $G_i = 0.01$ (right).

either positive real- or imaginary-valued. But this in turn means that the corresponding eigenfunction either consists of sines and cosines in both media, or sines and cosines in the part with refraction index n_{II} and hyperbolic sines and cosines in the part with refraction index n_{I} . In other words the eigenfunction either has a propagating behaviour in both media or a propagating behaviour in the part with the larger refraction index and an evanescent or exponentially decaying behaviour in the part with smaller refraction index. If the ASR coordinate transformation is used to get a better approximation of these eigenfunctions, typically the exponentially decaying modes are difficult to capture with RCWA and its standard Fourier series. Therefore a good coordinate transformation should probably focus on capturing this exponentially decaying behaviour. Since this behaviour only takes place on one side of the material interface, this could suggest that one only needs to refine on one side too. The corresponding coordinate transformation is therefore largely the same as the original or first alternative, except that in the part with the lower refraction index the parameter $\gamma_{i,l}$ is set to zero. For example, for the binary grating the first line in equation (6.3) would only contain the linear term and not the sine. In Figure 6.1 this would mean that with a stretching of $G_i = 0.1$ the graph would correspond with the blue line in the interval $(-0.2, 0.2)$ and with the green line otherwise. Clearly such a coordinate transformation is still continuous but not differentiable.

Bibliography

- [1] N. van der Aa. *Sensitivity analysis for grating reconstruction*. PhD thesis, Technische Universiteit Eindhoven, 2007.
- [2] U.M. Ascher, R.M.M. Mattheij, and R.D. Russell. *Numerical solutions of boundary value problems for ordinary differential equations*. Philadelphia : Society for Industrial and Applied mathematics, 1995.
- [3] M. van Beurden. *Integro-differential equations for electromagnetic scattering*. PhD thesis, Technische Universiteit Eindhoven, 2003.
- [4] A.J.M. den Boef, A.J. Bleeker, Y.J.L.M. van Dommelen, M. Dusa, A.G.M. Kiers, P.F. Luehrmann, H.P.M. Pellemans, M. van der Schaar, C.D. Grouwstra, and M.G.M.M. van Kraaij. Method and apparatus for angular-resolved spectroscopic lithography characterisation. European Patent EP1628164, February 2006.
- [5] A.J.M. den Boef, W.M. Corbeij, M. Dusa, and M.G.M.M. van Kraaij. A method of characterising the transmission losses of an optical system. European Patent EP1868035, December 2007.
- [6] M. Born and E. Wolf. *Principles of Optics*. Cambridge University Press, reprinted with corrections, seventh (expanded) edition, 2002.
- [7] L.C. Botten. Completeness and modal expansion methods in diffraction theory. *Opt. Acta.*, 32(12):1479–1488, 1985.
- [8] L.C. Botten, M.S. Craig, and R.C. McPhedran. Complex zeros of analytic functions. *Computer Phys. Commun.*, 29:245–259, 1983.
- [9] L.C. Botten, M.S. Craig, R.C. McPhedran, J.L. Adams, and J.R. Andrewartha. The dielectric lamellar diffraction grating. *Opt. Acta.*, 28(3):413–428, 1981.
- [10] L.C. Botten, M.S. Craig, R.C. McPhedran, J.L. Adams, and J.R. Andrewartha. The finitely conducting lamellar diffraction grating. *Opt. Acta.*, 28(8):1087–1102, 1981.
- [11] L.C. Botten, M.S. Craig, R.C. McPhedran, J.L. Adams, and J.R. Andrewartha. Highly conducting lamellar diffraction grating. *Opt. Acta.*, 28(8):1103–1106, 1981.

-
- [12] E.A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. McGraw-Hill book Company, 1955.
- [13] R.E. Collin and F.J. Zucker. *Antenna theory*. McFrw-Hill, 1969.
- [14] N.P.K. Cotter, T.W. Preist, and J.R. Sambles. Scattering-matrix approach to multi-layer diffraction. *J. opt. Soc. Am. A*, 12(5):1097–1103, May 1995.
- [15] M.S.P. Eastham. *The spectral theory of periodic differential equations*. Scottish Academic Press, 1973.
- [16] G.E. Forsythe, M.A. Malcolm, and C.B. Moler. *Computer Methods for Mathematical Computations*. Prentice-Hall, 1976.
- [17] G. Granet. Reformulation of the lamellar grating problem through the concept of adaptive spatial resolution. *J. opt. Soc. Am. A*, 16(10):2510–2516, October 1999.
- [18] R.F. Harrington. *Field computation by moment methods*. IEEE Press, 1993.
- [19] A. Ishimura. *Electromagnetic Wave Propagation, Radiation, and Scattering*. Prentice-Hall, 1991.
- [20] J.D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, 3rd edition, 1999.
- [21] D.S. Jones. *The Theory of Electromagnetism*. Pergamon Press, 1964.
- [22] D.Y.K. Ko and J.R. Sambles. Scattering matrix method for propagation of radiation in stratified media: attenuated total reflection studies of liquid crystals. *J. opt. Soc. Am. A*, 5(11):1863–1866, November 1988.
- [23] P. Lalanne and J. Hugonin. Numerical performance of finite-difference modal methods for the electromagnetic analysis of one-dimensional lamellar gratings. *J. Opt. Soc. Am. A*, 17(6):1033–1042, June 2000.
- [24] L. Li. Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity. *J. opt. Soc. Am. A*, 10(12):2581–2591, December 1993.
- [25] L. Li. Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings. *J. opt. Soc. Am. A*, 13(5):1024–1035, May 1996.
- [26] L. Li. Use of fourier series in the analysis of discontinuous periodic structures. *J. Opt. Soc. Am. A*, 13(9):1870–1876, September 1996.
- [27] L. Li. *Mathematical modeling in optical science*, chapter 4. Use of Fourier series in the analysis of discontinuous periodic structures, pages 111–139. Number 22 in *Frontiers in Applied Mathematics*. SIAM, 2001.
- [28] R.M.M. Mattheij. Decoupling and stability of algorithms for boundary value problems. *Society for Industrial and Applied Mathematics*, 27(1):1–44, March 1985.

-
- [29] M.G. Moharam, D.A. Pommet, and E.B. Grann. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *J. Opt. Soc. Am.*, 12(5):1068–1076, May 1995.
- [30] M.G. Moharam, D.A. Pommet, and E.B. Grann. Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach. *J. Opt. Soc. Am. A*, 12(5):1077–1086, May 1995.
- [31] R. Petit, editor. *Electromagnetic Theory of Gratings*. Springer-Verlag, 1980.
- [32] J.Y. Suratteau, M. Cadilhac, and R. Petit. Sur la détermination numérique des efficacités de certain réseaux diélectriques profonds. *J. Optics (Paris)*, 14(6):273–288, 1983.
- [33] E.L. Tan. Note on formulation of the enhanced scattering-(transmittance-) matrix approach. *J. opt. Soc. Am. A*, 19(6):1157–1161, June 2002.
- [34] G. Tayeb and R. Petit. On the numerical study of deep conducting lamellar diffraction gratings. *Opt. Acta*, 31:1361–1365, 1984.
- [35] W.J. Thompson. *Atlas for computing mathematical functions*. Wiley-Interscience, 1997.
- [36] T. Vallius and M. Honkanen. Reformulation of the fourier modal method with adaptive spatial resolution: application to multilevel profiles. *Opt. Express*, 10(1):24–34, January 2001.
- [37] C. Vassallo. Improvement of the finite difference method for step-index optical waveguides. *IEE Proceedings-J*, 139(2):137–142, April 1992.
- [38] C. Vassallo. Interest of improved three-point formuleas for finite-difference modeling of optical devices. *J. opt. Soc. Am. A*, 14(12):3273–3284, December 1997.
- [39] ARPACK – ARnoldi PACKAge. <http://www.caam.rice.edu/software/arpack/>.
- [40] EISPACK – EIgenSystem PACKAge. <http://www.netlib.org/eispack/>.
- [41] LAPACK – Linear Algebra PACKAge. <http://www.netlib.org/lapack/>.

Index

- T*-matrix algorithm, *see* transfer matrix algorithm
- p*-polarised, 24
- s*-polarised, 24

- adaptive spatial resolution, 85–91
- admittance, 16
- angular frequency, 10
- ASR, *see* adaptive spatial resolution
- azimuthal angle, 15

- Bloch mode expansion, 27–33
- boundary conditions, 12

- condensation algorithm, 69
- conductivity, 11
- conductor, 11
- conical diffraction, 16
- constitutive relations, 10
- continuity equation, 9

- dichotomy, 71
- dielectric, 11
- dielectric constant, 11
- diffraction, 8
 - efficiency, 26
 - order, 23
 - plane, 24
 - theory, 7
- discrete boundary value problem, 42, 63
- dispersion free, 11
- doping, 4

- electric
 - charge density, 9
 - current density, 9
 - displacement, 9
 - field, 9
 - flux density, 9
 - surface charge density, 12
 - surface current density, 12
- energy flow density, 25
- evanescent wave, 23

- FD, *see* finite differences
- finite difference, 85
- free space
 - permeability, 11
 - permittivity, 11
- fundamental solution, 41, 62

- grating, 4
- grating layer, 15

- Helmholtz equation, 12
- Huygens' principle, 8

- initial value problem, 79
- inner layer interface
 - boundary conditions, 28, 31
- insulator, 11
- integrated circuit, 2
- interconnects, 2
- interface boundary conditions, 12
- interference, 8
- isotropic, 11

- law of reflection, 8
- law of refraction, 8
- linearly reacting, 11
- lithography, 2–4
- Littrow mount, 34

- magnetic
 - field, 9
 - flux density, 9
 - induction, 9
- Maxwell's equations, 9–10
 - Ampère's law, 9
 - Faraday's law, 9
 - Gauss' law electric, 10
 - Gauss' law magnetic, 10
- method of moments, 40, 45
 - homogeneous method, 40
 - hybrid method, 40
- metrology
 - critical dimension, 4
 - overlay, 4
- monochromatic field, 10
- Moore's law, 1
- multiple shooting, 69
- multiple shooting method, 79
- multiplication rule
 - inverse, 55
 - Laurent, 55
- perfectly electric conductor, 11
- periodic eigenvalue problem, 34
- permeability, 11
- permittivity, 11
- photoresist, 3
- pitch, 14
- planar diffraction, 15
- plane of incidence, 15
- plane wave, 15
- polar angle, 15
- polarisation, 8
 - angle, 15
 - TE, transverse electric, 16
 - TM, transverse magnetic, 16
- power flux density, 25
- Poynting vector, 25
- propagating wave, 23
- pseudo-periodic
 - boundary condition, 16
 - Fourier series, 22
- radiation boundary conditions, 12
- Rayleigh expansion, 20
- Rayleigh radiation condition, 13
- RCWA, *see* Rigorous Coupled-Wave Analysis
- RCWA mode expansion, 53–59
- reduced superposition principle, 81
- refraction index, 12
- Rigorous Coupled-Wave Analysis, 53
 - semi-periodic eigenvalue problem, 34
 - semiconductor, 2
 - separated boundary conditions, 73
 - separation of variables, 21, 28
 - silicon, 2
 - single shooting method, 79
 - Sommerfeld radiation condition, 13
 - source free, 11
 - stabilised march, 81
 - substrate, 15
 - superstrate, 15
- test functions, 40, 45, 61
- time-harmonic field, 10
- time-invariant, 11
- transcendental equation, 30, 34–38
- transfer matrix algorithm, 69–71, 107
- transistor, 2
- wafer, 2
- wave vector, 15
- wavelength, 15
- wavenumber, 12
- weighed inner product, 21

Summary

Forward Diffraction Modelling: Analysis and Application to Grating Reconstruction

The semiconductor industry uses lithography machines for manufacturing complex integrated circuits (also called ICs) onto wafers. Because an IC is built up layer by layer and feature sizes get smaller and smaller, tight control of the lithography process is required to guarantee a fast production of working ICs. Typically a lot of information on the lithography process can be obtained by measuring test structures or gratings which are scattered over the wafer. These gratings are tiny periodic structures much smaller than ICs. First these gratings are illuminated and its response (a scattered intensity) is measured. For certain applications like overlay metrology the asymmetry in this measured signal (due to an offset between two gratings) can be used to align the lithographic process. For other applications like critical dimension (CD) metrology one is interested in the shape of the grating lines that produced the measured signal. Since this information is not directly available but encrypted in the measurement, a reconstruction algorithm is used to extract it. The reconstructed values like height, width and sidewall angle can then be related to machine settings like dose and focus which control the lithographic process. In particular the CD metrology application requires rigorous mathematical models that solve optical diffraction problems for periodic gratings in combination with advanced reconstruction algorithms.

This thesis focuses on the optical diffraction problem for 1D periodic gratings. Starting from Maxwell's equations a reduced model is derived by simplifying both the grating and the incident electromagnetic field. The former is approximated with an infinitely periodic layered structure with isotropic non-magnetic materials. The latter is approximated with a time-harmonic incident plane wave. The reduced model is discretised using two different mode expansion methods, Bloch and the Rigorous Coupled-Wave Analysis (RCWA). Bloch expands the electromagnetic field in each layer in terms of the exact eigenfunctions whereas RCWA only uses approximate eigenfunctions. After truncation of the involved series a transmission problem is derived by matching the fields at the layer interfaces. Having solved the resulting linear system, the scattered field can be computed easily.

Both mode expansion methods solve a similar linear system containing a large but sparse, block structured coefficient matrix. However, special care needs to be taken when solving this system stably and efficiently. Therefore a stable condensation algorithm is derived based on Riccati transformations that decouples the exponentially growing and decaying terms that are present in the solution. This separation or decoupling is the key feature explaining the stability which is not always clear in alternative condensation algorithms. Furthermore the algorithm is optimised for speed by using a two-stage approach. Finally it is shown that the resulting stable recursions are identical to those used in the “enhanced transmittance matrix approach” (a frequently used condensation algorithm), thereby confirming its stability as well.

This thesis also examines and extends both mode expansions methods. The Bloch method is generalised to deal with multiple material transitions inside a grating layer covering a wider range of applications. However, lossy or fully asymmetric gratings are still hard to solve. On the other hand the Fourier discretisation used in RCWA is much more flexible but only approximates the more exact discretisation of Bloch. Therefore two RCWA modifications have been investigated to improve the accuracy while keeping its flexibility and relatively straightforward implementation. Adaptive Spatial Resolution applies an additional layer specific coordinate transformation before Fourier discretising the problem again. A good transformation not only refines near a material interface but also does this in a smooth way. A significant improvement in accuracy is observed that approaches and sometimes outperforms the results obtained with the Bloch method. The second modification removes the Fourier discretisation completely and uses a finite difference approximation in the periodic direction. Although this approach allows for a better discretisation near a material interface, the sparsity of the resulting matrices could not be exploited to make a competitive implementation within the standard RCWA framework.

Finally the integration of the forward diffraction model in the CD reconstruction application is discussed. Either a library based or real-time regressions approach can be used for this reconstruction. Both approaches rely heavily on having an accurate and fast forward model. By exploiting additional symmetries and smart reuse of information, acceptable library fill times and real-time reconstructions are now feasible.

Samenvatting

De halfgeleiderindustrie maakt gebruik van lithografiemachines om complexe geïntegreerde circuits (ook wel IC's) te produceren op wafers. Omdat een IC laag voor laag wordt opgebouwd en de afmetingen van de structuren alsmaar kleiner worden, is het belangrijk het lithografisch proces zo goed mogelijk onder controle te krijgen om een snelle productie van IC's te kunnen garanderen. Informatie over dit proces kan verkregen worden door specifieke structuren genaamd gratings te meten die verspreid liggen over de wafer. Deze gratings zijn extreem kleine, periodiek herhalende structuren die zelfs nog kleiner zijn dan een IC. Eerst wordt een grating belicht om vervolgens de respons, bestaande uit een verstrooid veld, te meten. Voor bepaalde toepassingen zoals overlay metrologie kan de asymmetrie in dit gemeten signaal (als gevolg van een verschuiving tussen twee gratings) gebruikt worden om het lithografische proces uit te lijnen. Voor andere toepassingen zoals "critical dimension" (CD) metrologie is men meer geïnteresseerd in de daadwerkelijke vorm van de grating lijnen behorende bij dit gemeten signaal. Omdat vorminformatie niet direct beschikbaar maar gecodeerd is in de meting, wordt een reconstructie-algoritme gebruikt om deze informatie te decoderen. De gereconstrueerde vormparameters zoals hoogte, breedte en hoeken van schuine wanden kunnen vervolgens gerelateerd worden aan instellingen van de machine, zoals bijvoorbeeld de focusering of hoeveelheid licht, die het lithografisch proces sturen. Voornamelijk de CD metrologie applicatie vereist rigoureuze wiskundige modellen voor het oplossen van optische diffractie problemen voor periodieke gratings in combinatie met geavanceerde reconstructie algoritmen.

Dit proefschrift focust zich op het optische diffractie probleem voor 1D periodieke gratings. Beginnende bij de Maxwell vergelijkingen wordt een gereduceerd model afgeleid dat zowel de grating alsook het invallende elektromagnetische veld vereenvoudigd. De grating wordt oneindig periodiek verondersteld en trapsgewijs benaderd met lagen waarin de materialen isotroop en niet-magnetisch zijn. Het invallende veld wordt gemodelleerd met een tijdharmonische vlakke golf. Het gereduceerde model wordt vervolgens gediscretiseerd met behulp van twee verschillende expansiemethoden, Bloch en de "Rigorous Coupled-Wave Analysis" (RCWA). Bloch expandeert het elektromagnetische veld in termen van de exacte eigenfuncties behorende bij elke laag, daar waar RCWA slechts benaderde eigenfuncties gebruikt. Na het trunceren van de betreffende

reeksen wordt een transmissieprobleem opgesteld door de velden aan elkaar gelijk te stellen op de grensvlakken tussen de verschillende lagen. Nadat het resulterende lineaire systeem is opgelost kan het verstrooide veld eenvoudig worden berekend.

Beide expansiemethoden lossen een vergelijkbaar lineair systeem op waarvan de coëfficiënten matrix groot en blok-gestructureerd is en veelal nullen bevat. Echter, het stabiel en efficiënt oplossen van dit systeem vereist speciale zorg. Daarom is een stabiel condensatie algoritme afgeleid op basis van Riccati transformaties dat de exponentieel groeiende en dalende termen die aanwezig zijn in de oplossing ontkoppelt. Deze ont koppeling is de cruciale eigenschap dat de stabiliteit van dit algoritme verklaart en wat vaak niet duidelijk naar voren komt in alternatieve condensatie algoritmen. Daarnaast is het algoritme geoptimaliseerd voor snelheid door gebruik te maken van een twee-stappen-implementatie. Ten slotte wordt aangetoond dat de resulterende stabiele recursie identiek is aan de “enhanced transmittance matrix approach” (een veelgebruikt condensatie algoritme), waarmee diens stabiliteit bevestigd is.

In dit proefschrift worden beide expansiemethoden nader onderzocht en uitgebreid. De Bloch methode wordt gegeneraliseerd zodat deze ook kan omgaan met meerdere materiaalovergangen per laag waardoor een breder scala van toepassingen afgedekt kan worden. Echter, verliezende of volledig asymmetrische gratings zijn nog steeds moeilijk door te rekenen. Daarentegen is de Fourier discretisatie binnen RCWA een stuk flexibeler, maar geeft deze alleen een benadering van de meer exacte discretisatie in Bloch. Daarom zijn er twee RCWA gerelateerde modificaties onderzocht om diens nauwkeurigheid te verbeteren met behoud van de flexibiliteit en relatief eenvoudige implementatie. “Adaptive Spatial Resolution” past een additionele laag-specifieke coördinatentransformatie toe alvorens het probleem opnieuw te Fourier discretiseren. Een goede transformatie verfijnt niet alleen in de buurt van een materiaalovergang, maar doet dit ook op een gladde manier. Een significante verbetering in de nauwkeurigheid is waargenomen die niet alleen in de buurt komt van de met Bloch verkregen resultaten maar soms zelfs beter is. De tweede modificatie verwijdert de Fourier discretisatie volledig en vervangt deze door een eindige differentie discretisatie in de periodieke richting. Hoewel deze aanpak beter om kan gaan met een materiaalovergang, kan de ijheid van de resulterende matrices niet uitgebuit worden om een concurrerende implementatie te maken binnen het standaard RCWA kader.

Tenslotte wordt de integratie van het voorwaartse diffractie model binnen de CD reconstructie applicatie besproken. Zowel een op bibliotheken gebaseerde of een real-time regressie aanpak kan gebruikt worden voor deze reconstructie. Beide aanpakken leunen sterk op het hebben van een accuraat en snel voorwaarts model. Door gebruik te maken van extra symmetrieën en slim hergebruik van informatie, zijn acceptabele bibliotheek vultijden en real-time reconstructies vandaag de dag mogelijk.

Acknowledgements

This PhD project which in a way already started during my final Master's project in 2004 has been a close cooperation between university and industry. It has been a pleasure to meet so many new friends and colleagues in the *CASA* group (*Centre for Analysis, Scientific computing and Applications*) of the Eindhoven University of Technology and in ASML over the past couple of years. I would like to take this opportunity to thank some of them in particular.

First of all I am indebted to my promotor Bob Mattheij and supervisor at the university Jos Maubach for all the fruitful discussions and guidance. I would like to thank them especially for their patience and continued support. Also I want to express my gratitude to my other supervisors Irwan Setija and Arie den Boef from ASML with whom I had many helpful discussion which led to a better understanding of the ASML lithography challenges and applications. Finally I am grateful to Anton Tjihuis, Paul Urbach and Jaap Molenaar for being part of my defense committee and I would like to thank Irwan Setija (once more) and Wim Coene for completing the extended committee.

Because this PhD project has been part of a much larger project at ASML, I would like to give my sincere gratitude to some colleagues that were closely involved in this research. Teaming up with fellow PhD student Nico van der Aa and scientific consultants Marc Noot and Jürgen Tas from *CASA* has proven to be invaluable for moving this project forward. Also my other colleagues at ASML that provided an enjoyable and stimulating working environment I hereby thank you.

Being part of *CASA* has been a wonderful experience and in particular I would like to thank Miguel Patrício with whom I shared an office for almost four years. Also my other colleagues and friends at the university that joined in on our regular lunches at Kennispoort (food, food, food, 12 o'clock sharp), social events, infamous poker matches, evenings out in town, *CASA* outings and so much more, you were all part of an unforgettable experience in my life. Thank you Laura Astola, Evgeniya Balmashnova, Dragan Bezanovic, Mirela Darau, Willem Dijkstra, Remco Duits, Ali Etaati, Carlo De Falco,

Yabin Fan, Yves van Gennip, Christina Giannopapa, Hans Groot, Shruti Gumaste, Andriy Hlod, Davit Harutyunyan, Pieter Heres, Michiel Hochstenbach, Qingzhi (Darcy) Hou, Zoran Ilievski, Roxana Ionutiu, Bart Janssen, Godwin Kakuba, Jan Willem Knopper, Jan Kroot, Kundan Kumar, Agnieszka Lutowska, Bas van der Linden, Kamyar Malakpoor, Temesgen Markos, Remo Minero, Jos and Peter in 't panhuis, Maxim Pisarenco, Matthias Röger, Ronald Rook, Patricio Rosen Esquivel, Maria Rudnaya, Valeriu Savcenco, Berkan Sesen, Evgueni Shcherbakov, Sudhir Srivastava, Maria Ugryumova, Marco Veneroni, Arie Verhoeven, Erwin Vondenhoff and all other poker friends I made down the road. I am also very grateful for all the help I received from Enna van Dijk who not only took care of all the administrative tasks but with whom I also had great fun in organising the Sinterklaas events in our group.

Lastly, but by no means least, I would like to thank all my friends and family who have supported me throughout the years and brought joy to my life. Especially my parents, Annie and Paul, and sister Karin who helped me through sometimes hectic times by showing enthusiasm and offering a place of rest when needed. Thank you!

Mark van Kraaij

Eindhoven, February 2011

Curriculum vitae

Mark van Kraaij was born on February 10th, 1980, in Mierlo, the Netherlands. After finishing his pre-university education at the Strabrecht College (Geldrop) in 1998, he started studying applied mathematics at the Eindhoven University of Technology where he received a *fl.*5000,- scholarship award. With a specialisation in numerical mathematics, he obtained his Master's title *cum laude* in 2004. His thesis was entitled "A more Rigorous Coupled-Wave Analysis" under supervision of prof.dr. R.M.M. Mattheij, dr. J.M.L. Maubach and in close cooperation with ASML, world's leading provider of lithography systems for the semiconductor industry.

From 2004 to 2008 he has been working as a PhD student at the Eindhoven University of Technology within the CASA group (*Centre for Analysis, Scientific computing and Applications*). His PhD project was a continuation of his Master's project, again sponsored by and in close cooperation with ASML. At the end of 2008 he started working for ASML as a design engineer in the Research group.

