

Sensitivity analysis for grating reconstruction

Citation for published version (APA):

Aa, van der, N. P. (2007). *Sensitivity analysis for grating reconstruction*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Mathematics and Computer Science]. Technische Universiteit Eindhoven.
<https://doi.org/10.6100/IR630497>

DOI:

[10.6100/IR630497](https://doi.org/10.6100/IR630497)

Document status and date:

Published: 01/01/2007

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.

Sensitivity analysis for 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 donderdag 8 november 2007 om 16.00 uur

door

Nicolaas Petrus van der Aa

geboren te Helmond

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. R.M.M. Mattheij

Copromotor:

dr. H.G. ter Morsche

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Aa, Nicolaas Petrus van der

Sensitivity for grating reconstruction / door Nicolaas Petrus van der Aa. - Eindhoven :
Technische Universiteit Eindhoven, 2007.

Proefschrift. - ISBN 978-90-386-1139-6

NUR 919

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

2000 Mathematics Subject Classification: 65N25, 65N21, 65F15, 78A45, 78A46, 90C30

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

 **ASML** - Veldhoven, The Netherlands.

Copyright ©2007 by Nicolaas Petrus van der Aa, Eindhoven, The Netherlands.

All rights are reserved. Reproduction in whole or in part is prohibited without written consent
of the copyright owner.

Dedicated to my sons Aaron and Bram
and to my Cindy

Contents

1	Introduction	1
1.1	Lithography	2
1.2	Problem description	5
1.3	Outline	8
2	Modelling	11
2.1	Physical model	11
2.2	Mathematical model	14
2.2.1	Maxwell's equations	14
2.2.2	Incident field	17
2.2.3	Boundary conditions	18
2.3	Rayleigh expansion	20
2.4	Diffraction efficiencies	23
2.5	Propagating and evanescent diffraction orders	26
3	RCWA	27
3.1	The key feature of RCWA	28
3.2	General theory of truncating Fourier series	29
3.3	The field inside a layer	30
3.3.1	Planar diffraction: TE polarization	31
3.3.2	Planar diffraction: TM polarization	32
3.3.3	Conical diffraction	33
3.4	Boundary conditions	36
3.4.1	Planar diffraction	36
3.4.2	Conical diffraction	38
3.5	Final system	40
3.6	Enhanced transmittance matrix approach	43
3.7	Influence of the number of harmonics and layers	44
4	The C method	47
4.1	The key feature of the C method	48
4.2	The details of the C method	49

4.2.1	Coordinate transformation of the Helmholtz equation	49
4.2.2	The shape of the solution	50
4.2.3	Boundary conditions	52
4.3	RCWA vs. C method	57
4.4	Extension to overhanging gratings	58
5	Sensitivity	61
5.1	Finite differences	61
5.2	Analytical RCWA sensitivity	62
5.2.1	Sensitivity with respect to RCWA shape parameters	62
5.2.2	Sensitivity with respect to physical shape parameters	68
5.2.3	Sensitivity with respect to other parameters	72
5.3	Complexity of the algorithms	74
5.4	Comparison of analytical method vs. finite differences	75
5.4.1	Accuracy	75
5.4.2	Speed	77
6	Eigenvalue and eigenvector sensitivity	83
6.1	Existence of eigenvalue and eigenvector derivatives	84
6.2	Simple cases	86
6.2.1	Distinct eigenvalues	88
6.2.2	Repeated eigenvalues with distinct eigenvalue derivatives	89
6.2.3	Repeated eigenvalue with repeated eigenvalue derivatives	90
6.3	Generalization	91
6.4	Examples	95
6.5	Consequences for RCWA	99
6.5.1	The occurrence of repeated eigenvalues	100
6.5.2	The redundance of normalizing the eigenvectors	101
6.5.3	Eigenvalue and eigenvector derivatives in RCWA	105
7	Shape parameter reconstruction	107
7.1	Method of inversion	108
7.2	Preliminary results	110
8	Conclusions and future work	115
8.1	Conclusions	115
8.2	Future work	116
A	Mathematical derivations	119
A.1	Conversion of x - and y - to s - and p -components	119
A.1.1	Proof of (2.43a)	119
A.1.2	Proof of (2.49)	120
A.2	Solving systems of differential equations with constant coefficients	121

A.2.1	1st order differential equations with constant coefficients	122
A.2.2	2nd order differential equations with constant coefficients	123
A.3	Properties of the eigenvalues	124
Bibliography		127
Index		131
Summary		133
Samenvatting		135
Acknowledgement		137
Curriculum vitae		139

Chapter 1

Introduction

In 1946 the first electronic computer, the so-called ENIAC (see Figure 1.1), was built using vacuum tubes. It occupied 167 m², weighed 30 tonnes, consumed 160 kilowatts of electrical power and costed almost \$ 487,000 (in 1946). Nevertheless, its computing power was less than that of a simple programmable engineering calculator of today. Vacuum tubes show several limitations such as the fact that they are fragile, they burn out after several thousand hours of use, are bulky and give off considerable heat. In reaction, the *transistor* was invented by Bardeen, Brattain and Shockley at ATT Bell Labs in 1947 [8, 49]. In 1956 the inventors received the Nobel Prize in physics for their invention. Because a transistor enables the flow of electrons in a solid-material (a so-called semiconductor) instead of in vacuum, they can be made much smaller and more reliable. Until the invention of the *integrated circuit (IC)* in 1958, many transistors were made on a *wafer* (a round substrate

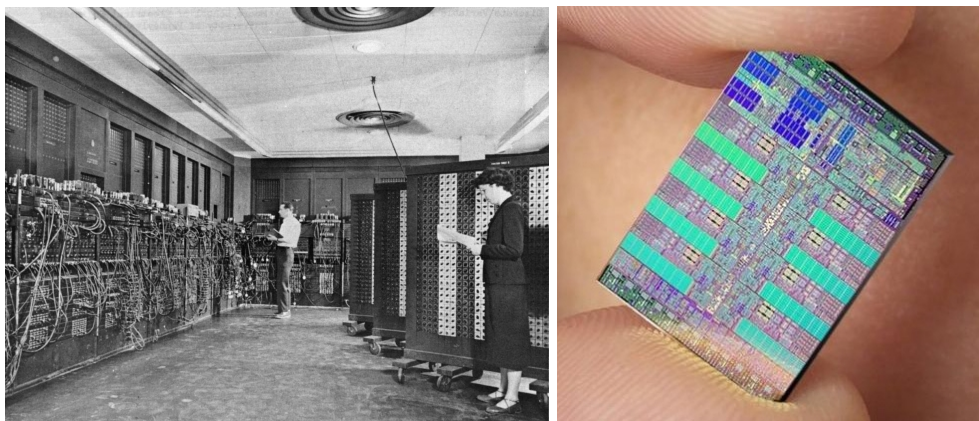


Figure 1.1: The ENIAC (*left*) vs. the Cell processor (*right*).

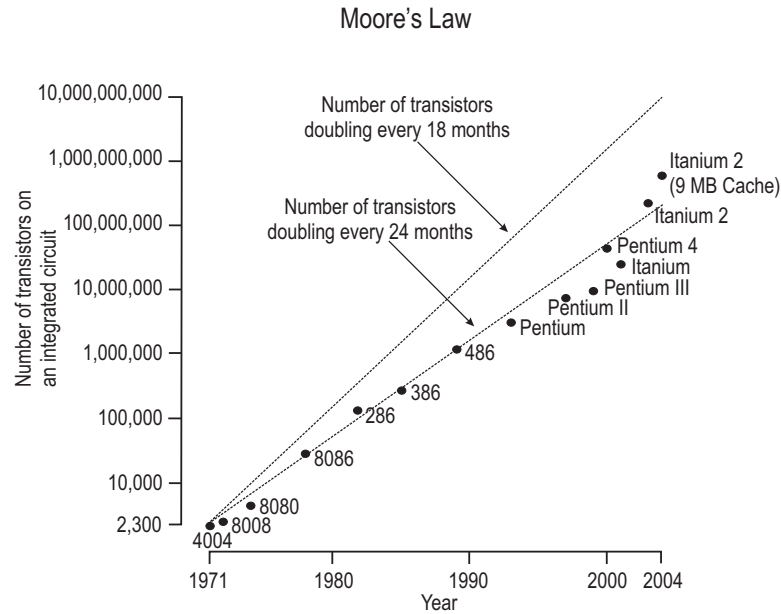


Figure 1.2: Moore's law.

made of silicon) with the intention to cut them from this wafer and to package them individually. The IC leaves all transistors on the wafer where they are interconnected. The most important reason for the IC industry to have grown so rapidly is that the decreasing feature size has continuously decreased the cost of chips, relative to their performance and functionality. In 1965 Gordon Moore predicted that the number of transistors on a chip would roughly double every 18 months [38]. This so-called *Moore's law* is still very accurate, although the doubling occurs every 24 months as illustrated in Figure 1.2. Nowadays one of the drivers behind the decreasing feature size is the gaming industry. For example, Sony PlayStation 3 makes use of the IBM/Toshiba/Sony Cell processor. This processor contains 234 million transistors on 221 mm² using only some watts of electrical power.

The decreasing feature sizes of computer chips have largely been possible due to advances in photolithography. To understand the motivation of this project, we will briefly dwell on why lithography is the most critical part of the chip-making process.

1.1 Lithography

An IC is the result of a multi-step photochemical manufacturing process. Figure 1.3 shows all steps within the chip making process in the order of occurrence. The production of an IC starts with a cylinder of mono-crystalline silicon that is cut into slices (1). Next, the slice

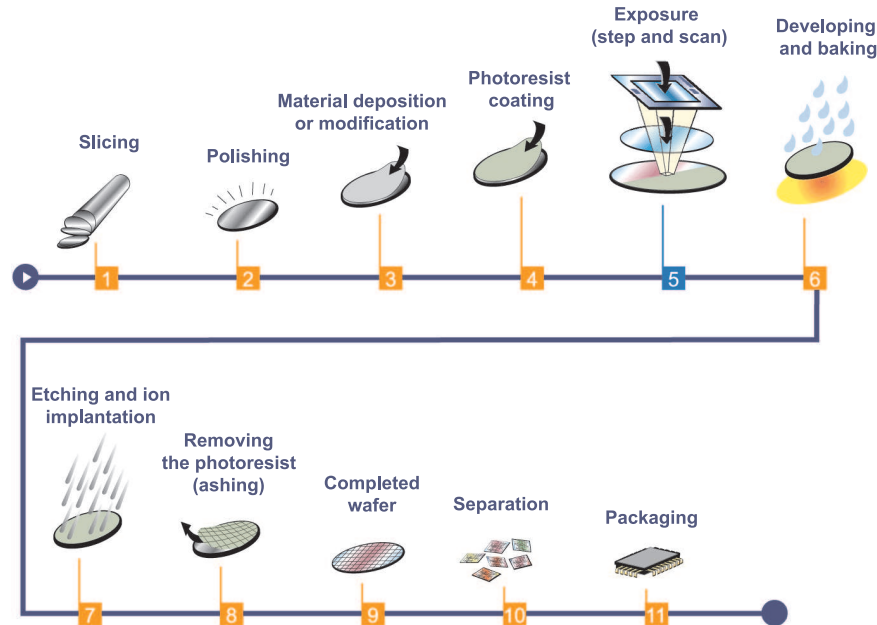


Figure 1.3: The chip making process.

is polished to obtain an ultra-flat wafer (2). This wafer is the basis on which the circuit elements like transistors, resistors and capacitors are built. In step (3) a *BARC* (Bottom Anti-Reflective Coating) layer is deposited on the wafer to reduce unwanted light reflections in the lithography step. After a deposition of a thin layer of light sensitive material called *photoresist* (4), the actual *lithography* step occurs (5), where a circuit pattern is projected onto a section of the wafer. During the exposure, the photoresist undergoes a chemical reaction. Depending on whether *positive* or *negative* resist is used, the composition of the resist becomes less or more chemically robust. After exposure, the wafer is developed, which implies that the remaining solvents are removed and the photoresist is hardened. The exposed resist is washed away (6). In this way the photolithography step transfers the circuit image onto the wafer. Etching and implantation of ions are done to create vertical and horizontal paths between layers on the wafer (7). The last process is to remove the less chemically robust part of the photoresist (8). Steps (3) to (8) are repeated 20-30 times in order to build up an entire IC. An example of such an IC is shown in Figure 1.4. The wafer contains several ICs (9) and therefore it should be cut into individual ICs (10). Finally, the ICs are packaged and connector pins are added to produce the final chip (11). For a more detailed description about the processes to make computer chips see [52].

In process steps (3) to (8) it is crucial that each new pattern is placed on top of the previously made pattern as accurately as possible. The only process for which the position on the wafer is critical is the photolithography step (step (5)). All other processes affect the wafer as a whole. For example, resist spinning (step (3)) is a process where a droplet

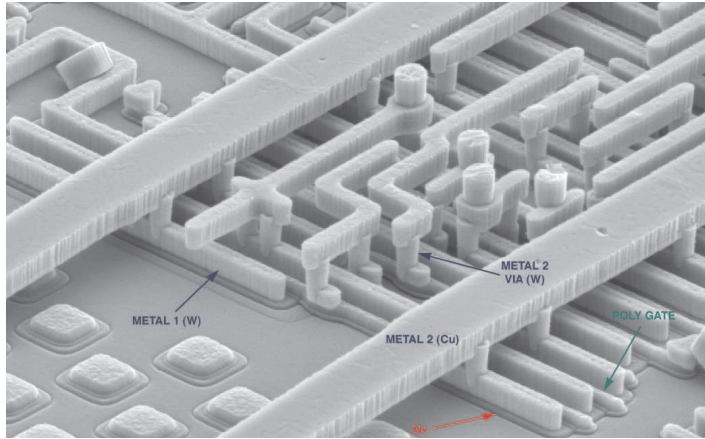


Figure 1.4: An enhanced view of an IC as it is manufactured nowadays. The lines have typical dimensions of a couple of nanometers. (Bron: IBM)

of liquid resist is put in the centre of a wafer while it is spinning at high speed such that the resist spreads out over the wafer. Photolithography creates a pattern onto a wafer by imaging a *mask*, which is a sheet of glass, partially covered by an opaque material. By shining light onto the mask, it exposes a pattern onto the photoresist. Since an IC is built up layer after layer, the patterns of every layer have to be placed exactly on top of each other to ensure that the IC performs well. The quality of this positioning is expressed by the term *overlay*, which describes the positional accuracy of a new lithographic pattern relative to the existing pattern on the wafer. To minimize overlay, accurate knowledge about the position on the wafer should be available.

To measure a position on a wafer with nanometer precision, special structures are put on the wafer. An example of such a structure is a *diffraction grating*, which is a repetitive array of either apertures or obstacles that has the effect of producing periodic alterations in the phase, the amplitude or both of a scattered wave. In this thesis, we will consider *phase gratings* and in practice this phase grating will typically be a *binary diffraction grating* (see Figure 1.5). There are several reasons to use such a diffraction grating and particularly

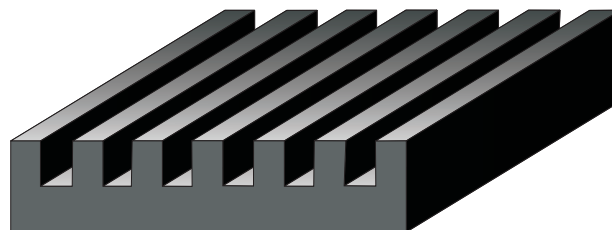


Figure 1.5: Representation of a binary diffraction grating.

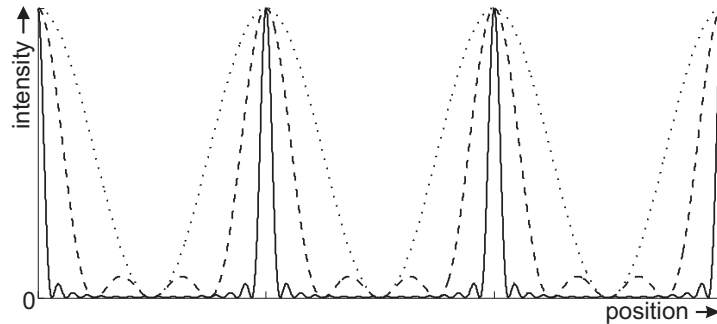


Figure 1.6: Illustration of normalized intensity profiles for 2, 4 and 16 slits.

a binary one. Gratings have the property that a plane wave will be diffracted in multiple *diffraction orders* that propagate in well-defined directions determined by the wavelength of the light and the period of the grating. The diffraction orders have a phase, which ensures that the position of this periodic structure can be found with a high accuracy [11]. By scanning the grating structure one is able to find the center of symmetry with an accuracy in the order of nanometers. Note that a pre-align step is required to indicate roughly the position of the grating, since the signal will also be periodic. The reason to prefer a periodic grating over a single structure is that, the more periods the grating has, the sharper the peaks of the intensity profile of the diffraction orders become as is illustrated in Figure 1.6. This implies that we can apply *spatial filtering* where only the peaks of the signal are allowed to pass and other parts are blocked to reduce the influence of noise. The reason why the grating usually is chosen to be a binary one is that it is relatively easy to manufacture. Finally, the reason why the grating structure should be small is that wafers are very expensive and, since gratings are not part of the final chip, a chip manufacturer does not want to use too much space on them.

1.2 Problem description

The objective of lithography is to produce an IC image. Although this process faces many difficulties, this thesis will treat only two of them. The first problem concerns measuring the position on a wafer as discussed in the previous section. In this context the gratings are called *alignment markers*. The second problem is the *resolution* of the lithography process, which is defined as the ability to produce a resist line that meets its requirements.

The first challenge is to correct for the effects on alignment markers caused by other processes such as etching or polishing. Ideally, an alignment marker is symmetric and the position found by all diffraction orders is its center of symmetry. But these markers are influenced by other processes, which cause the actual shape of the marker to deviate from its presumed shape. In Figure 1.7 a processed grating is shown where locally the symme-

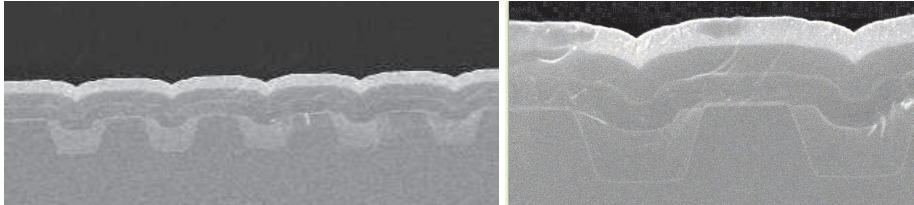


Figure 1.7: A representation of an alignment marker (*left*) and a magnification of one of its periods (*right*) (Source: ASML).

try is affected. When the marker becomes asymmetric, one cannot define the center of symmetry anymore and all diffraction orders will give different "positions" of the marker. Therefore, we would like to find the actual shape of these damaged gratings. Together with the knowledge of what kind of process caused the damage, we can reconstruct the shape of the original grating and estimate its position more accurately.

The second challenge concerns the resolution of a lithography process. The features to determine this resolution are the linewidth (midCD), the side-wall-angle (SWA) and the height of a trapezoidal grating structure (see Figure 1.8). For the same reasons as for the alignment markers, we have a periodic array of these trapezoids such that we can use the benefits of diffraction gratings here as well. Two important parameters that influence the shape of the trapezoidal lines are the distance between the image plane and the wafer (*focus*) and the intensity of the light (*dose*). The dose has an effect on the linewidth of the resist line after development, because the photoresist only reacts to light when this light has a certain intensity. Focus has an effect on the side-wall-angle, but also on the linewidth of the trapezoid (see Figure 1.9). To know whether the focus and dose were good enough to let the resist structure satisfy our requirements, we need to know its shape. For more details on focus and dose, see [24].

In this thesis we focus on the question whether we can reconstruct a diffraction grating.

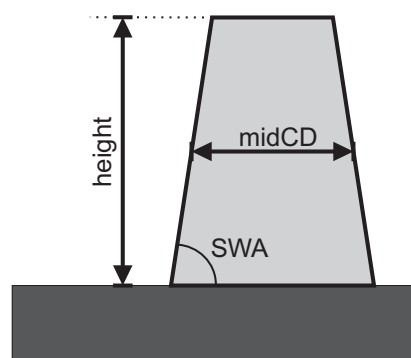


Figure 1.8: The definition of the shape parameters for a symmetric trapezoid.

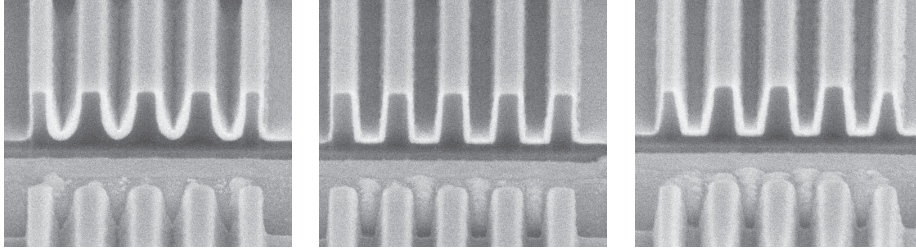


Figure 1.9: The effect of focus on the shape of the resist line. The figure in the middle is the result of producing trapezoidal lines when the wafer is in focus, while the left and right figures are the results of printing the lines when the wafer was out of focus because it has been placed above and below the imaging plane, respectively (Source: ASML).

This question has three aspects which we want to answer, namely

- **Possibility**
Can we reconstruct the actual shape?
- **Efficiency**
Can we do this reconstruction in a reasonable short time?
- **Stability**
Can we ensure that our method is mathematically robust?

The last two questions are equally important, because a stable method which takes too much time is useless in practice, but a fast method with wrong answers is useless as well.

To have an answer in the first place, we have to make some additional assumptions. Of course, this answer still has to represent the physics. Let us assume that

- the shape parameters behave periodically to ensure that we can apply periodic boundary conditions. This implies that also damages of the alignment mark or misprints of the resist lines are assumed to be identical over all periods.
- the initial shape parameters for the inverse process are estimates of the actual grating configuration which are close enough to the actual shape parameters, such that convergence of the iterative optimization process is ensured. In other words, a part of the robustness of our method is ensured by this assumption.

1.3 Outline

Finding the shape parameters from a measured scattered field can be defined as an inverse problem. It is easier to find the scattered field for a given grating structure, which we will define as the *forward problem*. This forward problem is used iteratively in a so-called *optimization process* or *inverse problem* as illustrated in Figure 1.10. This process starts with an initial guess of the shape parameters. Substitution of this initial guess in the forward problem gives a computed scattered field, which is compared to a measured one. The difference between them and the computed gradient of the scattered field with respect to the shape parameters lead to an improved guess of the shape parameters. This iterative process is repeated until the computed scattered field matches the measured one.

Chapters 2, 3 and 4 deal with the forward model. The fundamental physics is described in Chapter 2, where the incident field is defined and the equations which the field has to satisfy together with the boundary conditions are discussed. Next, two methods are discussed to solve the forward problem to find the scattered field for a given incident field and grating structure. These two methods are the rigorous coupled-wave analysis (RCWA) and the C method, which are explained in Chapter 3 and 4, respectively. Both methods are discussed in detail and both methods are preferred for a certain type of grating. However, RCWA is applicable to any type of grating while the C method cannot. This is the reason why the remaining chapters of this thesis will use RCWA as the forward model.

The inverse model requires knowledge about how the field reacts to small changes of the shape parameters. Chapter 5 describes two methods to compute first-order derivatives of the scattered field with respect to a shape parameter, namely finite differences and a

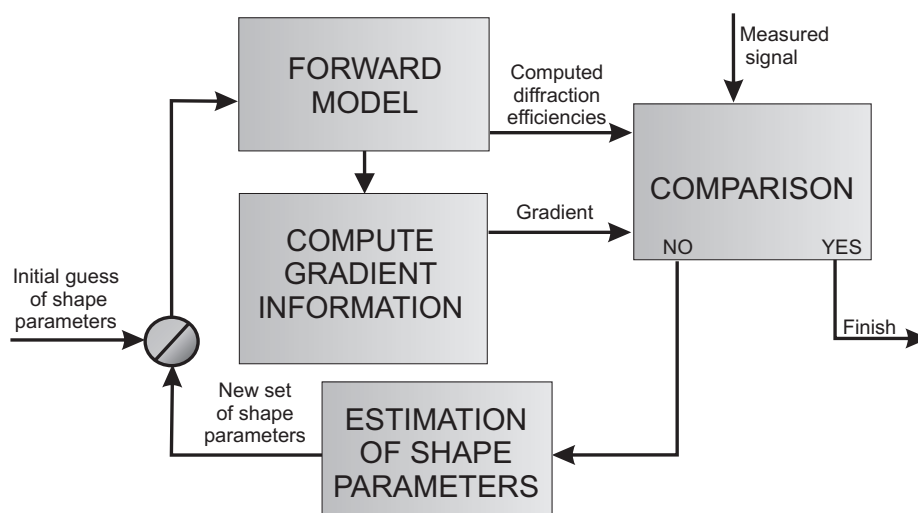


Figure 1.10: Schematic representation of the inverse process.

new method based on straightforward differentiating the relations within RCWA. The new method is derived to avoid solving additional eigenvalue problems, which are the most computationally expensive part of the RCWA algorithm. Instead, we need the derivatives of the eigenvalues and eigenvectors. Chapter 6 shows how these derivatives can be found in a mathematically correct way.

The final step of the inverse problem is the actual optimization. Chapter 7 will address several important issues in the optimization methods to perform an reconstruction of the shape parameters. Finally, Chapter 8 summarizes the main results of this thesis and discusses the issues that have to be investigated in the future.

Chapter 2

Modelling

2.1 Physical model

A typical grating we consider is infinitely periodic in the x -direction, while it is independent of the y -direction. In practice, this implies that the laser does not see the edges of the grating. Although all grating structures are three-dimensional in space, we speak of a *one-dimensional diffraction grating* when the periodicity of the grating is in one direction only. Examples of such one-dimensional diffraction gratings are given in Figure 2.1.

Let a grating be illuminated by a laser beam, which is approximated by a plane wave. The scattering of the wave by a grating structure can be explained by the wave nature of

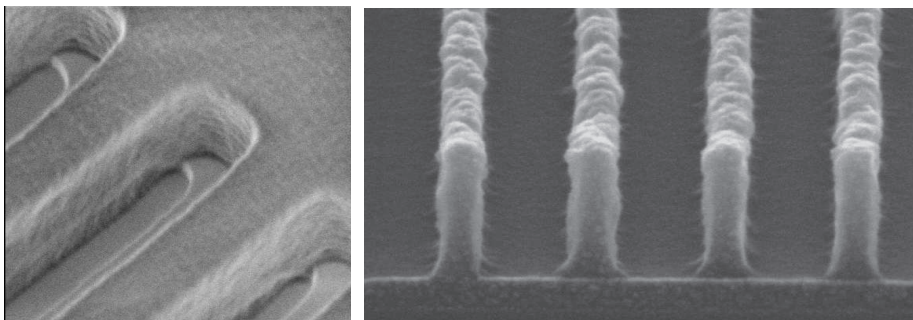


Figure 2.1: Examples of one-dimensional diffraction gratings. The left picture shows a top view of an etched grating profile and the right picture shows a cross-section of a grating consisting of resist lines (Source: ASML).

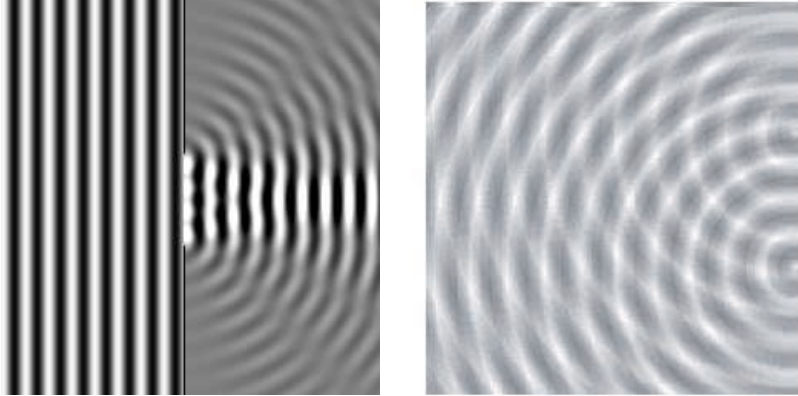


Figure 2.2: Illustration of the wave phenomena diffraction (*left*) and interference (*right*).

light. Since we want to consider light as an electromagnetic phenomenon, we will use the term *field* instead of light. A grating causes *interference* and *diffraction*, which are both typical phenomena of waves. Interference is the superposition of multiple wave fronts, while diffraction is the deviation of light from rectilinear propagation, which occurs whenever a portion of a wavefront is somehow obstructed. To have a clear picture of what happens physically, one can imagine the dips of the reflection grating as apertures of a transmission grating. The phenomena diffraction and interference are depicted in Figure 2.2. Diffraction of a wave can be explained by *Huygens's principle*, which states that every point on a propagating wavefront serves as the source of spherical secondary wavelets, such that the wavefront at some later time is the envelope of these wavelets [20]. Fresnel was able to account for interference by supplementing Huygens's principle that became famous as the *Huygens-Fresnel principle*, which states that the amplitude of the optical field at any point is the superposition of all these wavelets (considering their amplitudes and relative phases).

Let the incident field be given as a monochromatic plane wave with wavelength λ_0 , which is linearly polarized. The latter implies that the orientation of the field is constant. Also let the incident field be *time-harmonic*, which means that the time-dependence of the field is given by a periodic function. The angle under which the field is incident upon the grating structure can be specified by three quantities, namely the *polar angle* θ and the *azimuthal angle* ϕ , which indicate the direction of the field, and the *polarization angle* ψ , which is the angle denoting the orientation of the electric field (see Figure 2.3). If $\phi = 0$ we speak of *planar diffraction* and if $\phi \neq 0$ we call it *conical diffraction*. The *plane of incidence* is defined as that plane which contains the incident field and the scattered field. If the electric field component of a linearly polarized wave is perpendicular to the plane of incidence, that is $\psi = \pi/2$, we speak of a *transverse electric wave*. Similarly a field is a *transverse magnetic wave* if the magnetic field component of the linearly polarized wave is perpendicular to the plane of incidence. The TE-component is also referred to as *s-component* and the

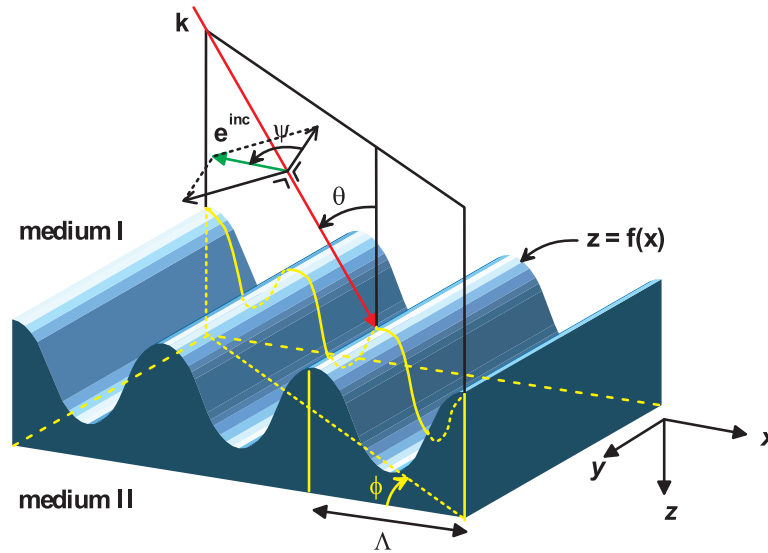


Figure 2.3: Graphical representation of the polar angle θ , the azimuthal angle ϕ and the angle between the incident electric field vector and the plane of incidence ψ .

TM-component is the p -component. Any other polarization state can be described as a linear combination of both the TE- and TM-component.

By considering a periodic grating, the secondary waves of the reflected field will cause constructive and destructive interference. In terms of superposition this means that in the first case all waves add up, while in the latter case the waves will cancel each other out. The infinite number of periods of the grating ensures that the regions of constructive interference are narrowed to certain angles only. For planar diffraction these angles can be found by the so-called *grating equation* given by

$$\Lambda(\sin \theta_{\text{out}} - \sin \theta) = n\lambda_0, \quad (2.1)$$

where Λ is the period of the grating, θ and θ_{out} are the angles of the incident and scattered field, respectively, n denotes the diffraction order and λ_0 is the wavelength. This equation is illustrated in Figure 2.4, where the optical path length $AB - CD$ is equal to a multiple of the wavelength λ_0 . From (2.1) it can be observed that large wavelengths have less diffraction orders than small wavelengths.

The grating equation only gives the angle of the diffraction order. What we would like to know is the amplitude and phase of the field in the direction of these diffraction orders. Among others, the Huygens-Fresnel principle was given a mathematical formulation by Kirchhoff which resulted in *Kirchhoff's scalar diffraction theory* [10]. It uses the fact that if the field and its normal derivative are known on a surface, the field at any point can be

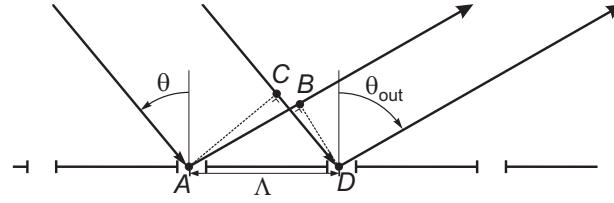


Figure 2.4: Definition of the parameters for deriving the grating equation. It holds that the optical path length $AB - CD = \Lambda (\sin \theta_{\text{out}} - \sin \theta)$.

calculated. However, the exact evaluation of the field on a surface, such as the field in an aperture is difficult. When the aperture size is large in terms of the wavelength *Kirchhoff's approximation* can be applied, which states that the aperture field may be approximated by that of the incident field. Because diffraction gratings are used with a period of the same order of magnitude as the wavelength of the light, we can only use rigorous methods which solve Maxwell's equations directly. Only then we can be sure that diffraction effects are taken into account and we obtain the correct amplitude and phase.

In the next section we will derive a mathematical model to obtain the field above the grating starting from Maxwell's equations.

2.2 Mathematical model

The mathematical model consists of a field description inside the domain, a description of the incident field and boundary conditions, which are discussed in in this section.

2.2.1 Maxwell's equations

A general electromagnetic field is described by Maxwell's equations [19, 48]. In the setting of diffraction grating theory we can simplify these equations significantly as will be shown in this section. The time-dependent Maxwell equations in Cartesian coordinates are given by

$$\nabla \times \mathbf{e}(\mathbf{x}, t) = -\frac{\partial \mathbf{h}(\mathbf{x}, t)}{\partial t}, \quad (\text{Faraday's law}) \quad (2.2a)$$

$$\nabla \times \mathbf{h}(\mathbf{x}, t) = \mathbf{j}(\mathbf{x}, t) + \frac{\partial \mathbf{d}(\mathbf{x}, t)}{\partial t}, \quad (\text{Ampère's law}) \quad (2.2b)$$

$$\nabla \cdot \mathbf{b}(\mathbf{x}, t) = 0, \quad (\text{Gauss's law for magnetic field}) \quad (2.2c)$$

$$\nabla \cdot \mathbf{d}(\mathbf{x}, t) = \rho(\mathbf{x}, t). \quad (\text{Gauss's law for electric field}) \quad (2.2d)$$

Faraday's law (2.2a) describes how an electric field \mathbf{e} (in V/m) is created when the magnetic flux density \mathbf{b} (in Vs/m²) changes. Similarly, Ampère's law (2.2b) describes how a magnetic field \mathbf{h} (in A/m) is created by a current density \mathbf{j} (in A/m²) and a change in the electric flux density \mathbf{d} (in As/m²). Gauss's law for electric fields gives the relationship between the electric flux density and the sources of that flux, the charge density ρ (in As/m³). The magnetic equivalence of Gauss's law (2.2c) states that no free magnetic poles exist. Furthermore, \mathbf{x} is the position vector in Cartesian coordinates and t denotes the time variable.

To formulate Maxwell's equations in terms of the electric and magnetic fields only, we will introduce some additional relations. These so-called *constitutive relations* describe the influence of the material under consideration on the electromagnetic fields. We assume the materials to be source-free, linear with respect to the electromagnetic fields, isotropic, time-invariant, dispersion-free and non-magnetic. The *source-free* property means that no external currents or charges are present, which implies that the current density \mathbf{j} and the charge density ρ consist only of those (internal) parts caused by the electromagnetic fields, which are denoted by \mathbf{j}_{int} and ρ_{int} respectively. In other words:

$$\mathbf{j}(\mathbf{x}, t) = \mathbf{j}_{\text{int}}(\mathbf{x}, t), \quad \rho(\mathbf{x}, t) = \rho_{\text{int}}(\mathbf{x}, t). \quad (2.3)$$

The *linearity* with respect to the fields means that the electric flux density \mathbf{d} and the internal current density \mathbf{j}_{int} are proportional to the electric field \mathbf{e} and the magnetic flux density \mathbf{b} is proportional to the magnetic field \mathbf{h} . Materials are *isotropic* when their properties are independent of the direction and polarization. By letting the material properties to be independent of time, the materials are *time-invariant*. Finally, assume that the materials are *dispersion-free*, which means that the response of the medium is instantaneous. All these properties are summarized in the following constitutive relations, viz.

$$\begin{cases} \mathbf{d}(\mathbf{x}, t) = \varepsilon(\mathbf{x})\mathbf{e}(\mathbf{x}, t), \\ \mathbf{b}(\mathbf{x}, t) = \mu_0\mathbf{h}(\mathbf{x}, t), \\ \mathbf{j}(\mathbf{x}, t) = \mathbf{j}_{\text{int}}(\mathbf{x}, t) = \sigma(\mathbf{x})\mathbf{e}(\mathbf{x}, t). \end{cases} \quad (2.4)$$

Here, ε is the *permittivity* (in As/Vm), μ_0 is the *permeability* of vacuum (in Vs/Am) and σ is the *conductivity* (in A/Vm). In (2.4) we used the property that all materials are *non-magnetic*, which says that the permeability is equal to the permeability of vacuum $\mu = \mu_0$ for all materials.

When these constitutive relations are substituted in Faraday's law (2.2a) and Ampère's law (2.2b) we obtain two relations with only the electric and magnetic field as unknowns.

$$\nabla \times \mathbf{e}(\mathbf{x}, t) = -\mu_0 \frac{\partial \mathbf{h}(\mathbf{x}, t)}{\partial t}, \quad (2.5a)$$

$$\nabla \times \mathbf{h}(\mathbf{x}, t) = \sigma(\mathbf{x})\mathbf{e}(\mathbf{x}, t) + \varepsilon(\mathbf{x}) \frac{\partial \mathbf{e}(\mathbf{x}, t)}{\partial t}. \quad (2.5b)$$

Let the time dependency of the electromagnetic fields be *time-harmonic*, that is

$$\mathbf{e}(\mathbf{x}, t) = \text{Re} \{ \tilde{\mathbf{e}}(\mathbf{x}) \exp[i\omega t] \}, \quad \mathbf{h}(\mathbf{x}, t) = \text{Re} \{ \tilde{\mathbf{h}}(\mathbf{x}) \exp[i\omega t] \}, \quad (2.6)$$

where $\tilde{\mathbf{e}}$ and $\tilde{\mathbf{h}}$ denote the part of the electric and magnetic field that only depend on the position \mathbf{x} . The variable ω is the *angular temporal frequency* (in rad/s). Introducing the time-harmonic property for the electromagnetic fields in (2.5) gives

$$\nabla \times \tilde{\mathbf{e}}(\mathbf{x}) = -i\omega\mu_0\tilde{\mathbf{h}}(\mathbf{x}), \quad (2.7a)$$

$$\nabla \times \tilde{\mathbf{h}}(\mathbf{x}) = i\omega \left(\varepsilon(\mathbf{x}) - i\frac{\sigma(\mathbf{x})}{\omega} \right) \tilde{\mathbf{e}}(\mathbf{x}). \quad (2.7b)$$

Note that we omitted to take the real part, since taking the real part of the complex-valued solution of (2.7) gives the same as when we solve (2.7) for only the real parts of the electromagnetic fields [46]. For notational convenience, the tilde in the electromagnetic fields is discarded and $\varepsilon - i(\sigma/\omega)$ is written as $\tilde{\varepsilon}$.

Since Maxwell's equations in the form of (2.7) will be used frequently in the remainder of this thesis, we will give these equations in Cartesian coordinates, viz.

$$\frac{\partial}{\partial y} e_z(x, y, z) - \frac{\partial}{\partial z} e_y(x, y, z) = -i\omega\mu_0 h_x(x, y, z), \quad (2.8a)$$

$$\frac{\partial}{\partial z} e_x(x, y, z) - \frac{\partial}{\partial x} e_z(x, y, z) = -i\omega\mu_0 h_y(x, y, z), \quad (2.8b)$$

$$\frac{\partial}{\partial x} e_y(x, y, z) - \frac{\partial}{\partial y} e_x(x, y, z) = -i\omega\mu_0 h_z(x, y, z), \quad (2.8c)$$

$$\frac{\partial}{\partial y} h_z(x, y, z) - \frac{\partial}{\partial z} h_y(x, y, z) = i\omega\tilde{\varepsilon}(x, y, z)e_x(x, y, z), \quad (2.8d)$$

$$\frac{\partial}{\partial z} h_x(x, y, z) - \frac{\partial}{\partial x} h_z(x, y, z) = i\omega\tilde{\varepsilon}(x, y, z)e_y(x, y, z), \quad (2.8e)$$

$$\frac{\partial}{\partial x} h_y(x, y, z) - \frac{\partial}{\partial y} h_x(x, y, z) = i\omega\tilde{\varepsilon}(x, y, z)e_z(x, y, z). \quad (2.8f)$$

We can decouple the electric and magnetic field by taking the curl of both (2.7a) and (2.7b). Using the identities $\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2 \mathbf{a}$ and $\nabla \cdot (\nabla \times \mathbf{a}) = 0$, the relation between the wave number k_0 and the angular temporal frequency ω given by

$$k_0 = \omega\sqrt{\varepsilon_0\mu_0}, \quad (2.9)$$

and introducing the *relative permittivity* $\tilde{\varepsilon}^r$ as

$$\tilde{\varepsilon}^r(x, z) := \tilde{\varepsilon}(x, z)/\varepsilon_0, \quad (2.10)$$

we can write

$$\nabla^2 \mathbf{e}(\mathbf{x}) + k_0^2 \tilde{\varepsilon}^r(\mathbf{x}) \mathbf{e}(\mathbf{x}) = \nabla \left(\frac{1}{\tilde{\varepsilon}^r(\mathbf{x})} ((\nabla \tilde{\varepsilon}^r(\mathbf{x})) \cdot \mathbf{e}(\mathbf{x})) \right), \quad (2.11a)$$

$$\nabla^2 \mathbf{h}(\mathbf{x}) + k_0^2 \tilde{\varepsilon}^r(\mathbf{x}) \mathbf{h}(\mathbf{x}) = \tilde{\varepsilon}^r(\mathbf{x}) \left(\left(\nabla \frac{1}{\tilde{\varepsilon}^r(\mathbf{x})} \right) \times (\nabla \times \mathbf{h}(\mathbf{x})) \right). \quad (2.11b)$$

If our domain consists of a homogeneous material as is the case above and below the grating structure, identities (2.11a) and (2.11b) become *Helmholtz equations*, viz.

$$\nabla^2 \mathbf{e}(\mathbf{x}) + k_0^2 \tilde{\varepsilon}^r \mathbf{e}(\mathbf{x}) = \mathbf{0}, \quad (2.12a)$$

$$\nabla^2 \mathbf{h}(\mathbf{x}) + k_0^2 \tilde{\varepsilon}^r \mathbf{h}(\mathbf{x}) = \mathbf{0}. \quad (2.12b)$$

Note that (2.12a) and (2.12b) are exactly the same for both the electric and magnetic field.

In case of planar diffraction, both the incident field and the grating structure are independent of the y -coordinate and therefore the entire field is independent of the y -coordinate. As a result, (2.11) gives two relations when we consider the y -components of both electric and magnetic field, viz.

$$\frac{\partial^2}{\partial x^2} e_y(x, z) + \frac{\partial^2}{\partial z^2} e_y(x, z) = -k_0^2 \tilde{\epsilon}^r(x, z) e_y(x, z), \quad (2.13a)$$

$$\frac{\partial}{\partial x} \left(\frac{1}{\tilde{\epsilon}^r(x, z)} \frac{\partial}{\partial x} h_y(x, z) \right) + \frac{\partial}{\partial z} \left(\frac{1}{\tilde{\epsilon}^r(x, z)} \frac{\partial}{\partial z} h_y(x, z) \right) = -k_0^2 h_y(x, z). \quad (2.13b)$$

TE polarization implies that only the y -component of the electric field is present and therefore, only (2.13a) should be considered. Similarly, for TM polarization only (2.13b) has to be solved. Any other case of planar diffraction can be considered as a linear combination of TE and TM polarization. For conical diffraction, where the azimuthal angle is nonzero, it is not possible to uncouple the x - or y -component of a field directly and we have to solve the combination of (2.11a) and (2.11b).

2.2.2 Incident field

In Section 2.1 the angles describing the incident field have been defined (see Figure 2.3). For conical diffraction the normalized incident electric field is given by

$$\begin{aligned} \mathbf{e}^{\text{inc}}(x, y, z) &= \exp[-i\mathbf{k} \cdot \mathbf{x}] \mathbf{u} \\ &= \exp[-ik_0 n_1 (x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta)] \mathbf{u}, \end{aligned} \quad (2.14)$$

with

$$\begin{aligned} \mathbf{u} &= \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= (\cos \psi \cos \theta \cos \phi - \sin \psi \sin \phi) \mathbf{u}_x \\ &\quad + (\cos \psi \cos \theta \sin \phi + \sin \psi \cos \phi) \mathbf{u}_y - \cos \psi \sin \theta \mathbf{u}_z. \end{aligned} \quad (2.15)$$

The matrices are rotation matrices with the polar angle θ , the azimuthal angle ϕ and the polarization angle ψ , respectively. They describe the way the incident electric field is decomposed in x -, y - and z -direction. The vectors \mathbf{u}_x , \mathbf{u}_y and \mathbf{u}_z are unit vectors in the x -, y - and z -direction, respectively.

Although planar diffraction is merely a special case of conical diffraction, the TE and TM polarizations are frequently used in the context of one-dimensional diffraction grating theory, because of their simplicity and because any other polarization for planar diffraction is a linear combination of them. For *TE polarization* ($\psi = \pi/2$ and $\phi = 0$) only the y -component of

the electric field e_y^{inc} is present, which is given by

$$e_y^{\text{inc}}(x, z) = \exp[-ik_0 n_1(x \sin \theta + z \cos \theta)]. \quad (2.16)$$

Similarly for *TM polarization* ($\psi = 0$ and $\phi = 0$) the incident magnetic field is perpendicular to the plane of incidence and therefore, only the y -component of the magnetic field h_y^{inc} is present. This follows from substitution of $\psi = 0$ and $\phi = 0$ in (2.14), which results in

$$\mathbf{e}^{\text{inc}}(x, z) = (\cos \theta \mathbf{u}_x - \sin \theta \mathbf{u}_z) \exp[-ik_0 n_1(x \sin \theta + z \cos \theta)]. \quad (2.17)$$

The magnetic field can be found by using Maxwell's equation (2.8b).

$$\begin{aligned} h_y^{\text{inc}}(x, z) &= \frac{i}{\omega \mu_0} \left(\frac{\partial e_x^{\text{inc}}(x, z)}{\partial z} - \frac{\partial e_z^{\text{inc}}(x, z)}{\partial x} \right) \\ &= n_1 \sqrt{\varepsilon_0 / \mu_0} \exp[-ik_0 n_1(x \sin \theta + z \cos \theta)]. \end{aligned} \quad (2.18)$$

2.2.3 Boundary conditions

Diffraction of a grating is a boundary value problem, which means that boundary conditions are required to find a unique solution. Since the boundary conditions are similar for both the electric and magnetic field, let us define \mathbf{f} as

$$\mathbf{f}(x, y, z) := \begin{cases} \mathbf{e}(x, y, z) \\ \mathbf{h}(x, y, z) \end{cases}, \quad (2.19)$$

depending on which field is under consideration. The one-dimensional diffraction grating considered here has three types of boundary conditions:

- **Outgoing wave condition**

As Sommerfeld said [50]:

The energy which is radiated from the sources must scatter to infinity; no energy may be radiated from infinity into ... the field.

This implies that the reflected field, which is defined as the total field minus the incident field, and the transmitted field, which is the field that passes through the grating, should be outgoing waves or decaying waves only. Note that in the halfspace above the grating the outgoing wave condition holds for the reflected field only and not for the incident field.

- **Pseudo-periodic boundary condition**

The periodicity of the grating allows us to restrict the domain to one period only. The idea of this boundary condition is given in Figure 2.5. The distance between the wave fronts at (x, y, z) and $(x + \Lambda, y, z)$ is equal to $\Lambda \sin \theta \cos \phi$. For the phase, we

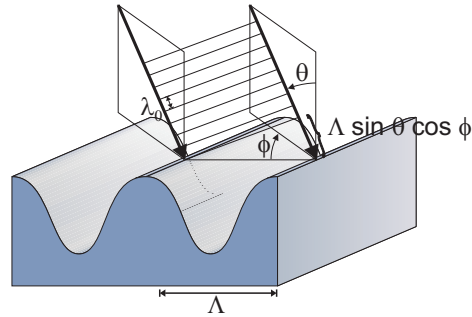


Figure 2.5: Illustration of the pseudo-periodic boundary condition or Floquet condition.

have to express this distance in terms of the number of wavelengths, which is equal to $(\Lambda/\lambda_0) \sin \theta \cos \phi$. The phase difference is zero if this number is a natural number. The pseudo-periodic boundary condition or *Floquet condition* is given by [43]

$$\mathbf{f}(x + \Lambda, y, z) = \mathbf{f}(x, y, z) \exp[-ik_0 n_1 \Lambda \sin \theta \cos \phi]. \quad (2.20)$$

For a positive polar angle θ the phase of the wavefront at $(x + \Lambda, y, z)$ leads the wavefront at (x, y, z) .

- **Continuity at the interface**

Since the tangential field components of the electric field should always be continuous and the same holds for the magnetic field when the materials are assumed to be non-magnetic, the boundary condition at an interface is given by

$$\mathbf{n}(x, y, z) \times (\mathbf{f}^+(x, y, z) - \mathbf{f}^-(x, y, z)) = \mathbf{0}, \quad \text{for } (x, y, z) \text{ on the interface,} \quad (2.21)$$

where \mathbf{f}^+ and \mathbf{f}^- are the fields above and below the interface, respectively, and \mathbf{n} is the normal pointing in the positive z -direction as illustrated in Figure 2.6. A mathematical derivation of this boundary condition can be found in [22].

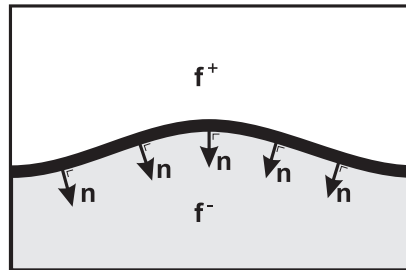


Figure 2.6: Definition of the normal and the field above and below an interface.

With the outgoing wave condition, the Floquet condition and the continuity of the tangential field components a unique solution exists to the problem of diffraction by a periodic structure [43].

2.3 Rayleigh expansion

To find the reflected and the transmitted field, we use the fact that above and below the grating profile (media I and II of Figure 2.7) the so-called *Rayleigh expansion* is valid [43]. For the halfspaces above and below the grating structure, only one material is present and therefore the fields satisfy the Helmholtz equations (2.12a) and (2.12b). The field also has to obey the outgoing wave condition and the Floquet condition (2.20). This will be sufficient to derive the Rayleigh expansion.

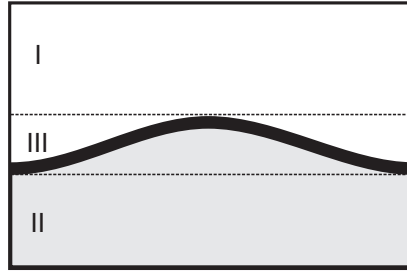


Figure 2.7: Identification of the domains where the Rayleigh expansion is valid. The Rayleigh expansion holds in media I and II, but not in III.

Consider the electric or magnetic field for conical diffraction. Because the Helmholtz equations (2.12) are identical for both fields, we use \mathbf{f} instead as has been defined in (2.19). Since \mathbf{f} satisfies the pseudo-periodic boundary condition (2.20), we can define $\tilde{\mathbf{f}}$ as

$$\tilde{\mathbf{f}}(x, y, z) := \mathbf{f}(x, y, z) \exp[ik_0 n_l \sin \theta \cos \phi x], \quad (2.22)$$

which is a periodic function in x with period Λ . We can represent $\tilde{\mathbf{f}}$ by its *Fourier series* as

$$\tilde{\mathbf{f}}(x, y, z) = \sum_{n=-\infty}^{\infty} \sigma_n(y, z) \exp\left[i\frac{2\pi n}{\Lambda}x\right]. \quad (2.23)$$

Here σ_n are the Fourier coefficients of $\tilde{\mathbf{f}}$. The field \mathbf{f} is then given by

$$\mathbf{f}(x, y, z) = \sum_{n=-\infty}^{\infty} \sigma_n(y, z) \exp[-ik_{x,n}x], \quad (2.24)$$

with

$$k_{x,n} := k_0 n_l \sin \theta \cos \phi - \frac{2\pi n}{\Lambda}. \quad (2.25)$$

When we substitute this expansion of \mathbf{f} in the Helmholtz equation (2.12), we find by using separation of variables that both the part depending on y and the one depending on z should be exponential functions. The shape of the exponential function depending on y is determined by the boundary condition at the grating interface which states that all tangential field components should be continuous at this interface for all y . Because of this boundary condition the exponential function inherits the exponential function depending on y of the incident field. Thus, the field can be given by

$$\mathbf{f}(x, y, z) = \sum_{n=-\infty}^{\infty} \boldsymbol{\gamma}_n(z) \exp[-i(k_{x,n}x + k_y y)], \quad (2.26)$$

with

$$k_y := k_0 n_l \sin \theta \sin \phi. \quad (2.27)$$

Finally, substitution of (2.26) in the Helmholtz equation yields

$$\sum_{n=-\infty}^{\infty} \left\{ \frac{\partial^2 \boldsymbol{\gamma}_n(z)}{\partial z^2} + (k_0^2 n_m^2 - k_{x,n}^2 - k_y^2) \boldsymbol{\gamma}_n(z) \right\} \exp[-i(k_{x,n}x + k_y y)] = \mathbf{0}, \quad (2.28)$$

where $m = I, II$ denotes the upper or lower medium. This relation is valid for any value of x and y and therefore every term of the Fourier series must be equal to zero.

$$\frac{d^2 \boldsymbol{\gamma}_n(z)}{dz^2} + (k_0^2 n_m^2 - k_{x,n}^2 - k_y^2) \boldsymbol{\gamma}_n(z) = \mathbf{0}. \quad (2.29)$$

Let

$$k_{m,z,n} := \sqrt{k_0^2 n_m^2 - k_{x,n}^2 - k_y^2}, \quad (2.30)$$

so that $\boldsymbol{\gamma}_n(z)$ can be written as

$$\boldsymbol{\gamma}_n(z) = \mathbf{r}_n^- \exp[-ik_{m,z,n}z] + \mathbf{r}_n^+ \exp[ik_{m,z,n}z]. \quad (2.31)$$

The coefficients \mathbf{r}_n^- represent incoming plane waves propagating toward the grating surface and \mathbf{r}_n^+ represent the outgoing plane waves. Because of the outgoing wave condition, the incoming waves should be zero (except for the wave representing the incident field). Thus the field \mathbf{f}_I can be expressed as a series of (decaying) plane waves, viz.

$$\mathbf{f}_I(x, y, z) := \mathbf{f}^{inc}(x, y, z) + \sum_{n=-\infty}^{\infty} \mathbf{r}_n \exp[-i(k_{x,n}x + k_y y - k_{I,z,n}z)]. \quad (2.32)$$

Analogously the transmitted field \mathbf{f}_{II} valid for $z \geq D$, where D represents the thickness of the grating, is given by

$$\mathbf{f}_{II}(x, y, z) := \sum_{n=-\infty}^{\infty} \mathbf{t}_n \exp[-i(k_{x,n}x + k_y y + k_{II,z,n}(z - D))]. \quad (2.33)$$

The quantities \mathbf{r}_n and \mathbf{t}_n are called the *reflected* and *transmitted field amplitudes*, respectively. This concludes the derivation of the Rayleigh expansion for the reflected and

transmitted field.

For planar TE and TM polarization, the Rayleigh expansions reduce to

$$e_{l,y}(x, z) = e_y^{\text{inc}}(x, z) + \sum_{n=-\infty}^{\infty} r_n \exp[-i(k_{x_n}x - k_{l,z_n}z)], \quad (2.34a)$$

$$h_{l,y}(x, z) = h_y^{\text{inc}}(x, z) + \sum_{n=-\infty}^{\infty} \tilde{r}_n \exp[-i(k_{x_n}x - k_{l,z_n}z)], \quad (2.34b)$$

with r_n and \tilde{r}_n are the reflected field amplitudes belonging to the electric and magnetic field, respectively. As a direct consequence of Maxwell's equations, they are related by

$$\tilde{r}_n := -\frac{1}{\omega\mu_0} (k_{l,z_n}r_{x_n} + k_{x_n}r_{z_n}). \quad (2.35)$$

In a similar way the Rayleigh expansions for the transmitted orders are given by

$$e_{ll,y}(x, z) = \sum_{n=-\infty}^{\infty} t_n \exp[-i(k_{x_n}x + k_{ll,z_n}(z - D))], \quad (2.36a)$$

$$h_{ll,y}(x, z) = \sum_{n=-\infty}^{\infty} \tilde{t}_n \exp[-i(k_{x_n}x + k_{ll,z_n}(z - D))], \quad (2.36b)$$

with

$$\tilde{t}_n := \frac{1}{\omega\mu_0} (k_{ll,z_n}t_{x_n} - k_{x_n}t_{z_n}). \quad (2.37)$$

For conical diffraction we can find r_{z_n} when r_{x_n} and r_{y_n} are known, because both the total electric field above the grating and the incident electric field are divergence-free, viz.

$$\nabla \cdot \mathbf{e}_l = -i(k_{x_n}r_{x_n} + k_y r_{y_n} - k_{l,z_n}r_{z_n}) \exp[-i(k_{x_n}x + k_y y - k_{l,z_n}z)] = 0. \quad (2.38)$$

Thus

$$r_{z_n} = \frac{k_{x_n}r_{x_n} + k_y r_{y_n}}{k_{l,z_n}}. \quad (2.39)$$

Similarly, we can derive that

$$t_{z_n} = -\frac{k_{x_n}t_{x_n} + k_y t_{y_n}}{k_{ll,z_n}}. \quad (2.40)$$

More often we are not interested in r_{x_n} or r_{y_n} , but in r_{s_n} and r_{p_n} , which are the components of the amplitudes of the electric and magnetic fields normal to the *diffraction plane*, which is the xz -plane rotated by the angle ϕ_n given by

$$\phi_n := \arctan(k_y/k_{x_n}). \quad (2.41)$$

Note that r_{sn} and r_{pn} and analogously t_{sn} and t_{pn} may be considered as the TE and TM components of the reflected and transmitted field. They are defined as

$$r_{sn} := r_{yn} \cos \phi_n - r_{xn} \sin \phi_n, \quad (2.42a)$$

$$r_{pn} := -\frac{1}{k_0 n_l} \{ (k_{l,zn} r_{xn} + k_{xn} r_{zn}) \cos \phi_n + (k_y r_{zn} + k_{l,zn} r_{yn}) \sin \phi_n \}, \quad (2.42b)$$

$$t_{sn} := t_{yn} \cos \phi_n - t_{xn} \sin \phi_n, \quad (2.42c)$$

$$t_{pn} := \frac{1}{k_0 n_l} \{ (k_{l,zn} t_{xn} - k_{xn} t_{zn}) \cos \phi_n - (k_y t_{zn} - k_{l,zn} t_{yn}) \sin \phi_n \}. \quad (2.42d)$$

The relation of r_{pn} and t_{pn} can be written in terms of r_{xn} and r_{yn} or t_{xn} and t_{yn} , viz.

$$r_{pn} = -\frac{k_0 n_l}{k_{l,zn}} (r_{xn} \cos \phi_n + r_{yn} \sin \phi_n). \quad (2.43a)$$

$$t_{pn} = \frac{k_0 n_l^2}{k_{l,zn} n_l} (t_{xn} \cos \phi_n + t_{yn} \sin \phi_n). \quad (2.43b)$$

The derivation of (2.43a) can be found in Appendix A.1. The derivation of (2.43b) is similar. Notice that for planar diffraction r_{sn} coincides with the y -component of the electric field and r_{pn} coincides with the y -component of the magnetic field.

2.4 Diffraction efficiencies

The quantity that is frequently used for comparison with actual measurements is the diffraction efficiency. The *diffraction efficiency* for the reflected or transmitted field is defined as the modulus of the ratio of the flux of the Poynting vector of the reflected or transmitted field and the flux of the Poynting vector of the incident field through a surface parallel to the mean plane of the grating [41]. The reflected electric field above the grating structure for conical diffraction is given by

$$\mathbf{e}^{\text{ref}}(x, y, z) := \sum_n \mathbf{r}_n \exp[-i(k_{xn}x + k_y y - k_{l,zn}z)]. \quad (2.44)$$

From Maxwell's equations (2.8) it follows that the reflected magnetic field \mathbf{h}^{ref} can be derived from the reflected electric field.

$$h_x^{\text{ref}}(x, y, z) = \frac{1}{\omega \mu_0} \sum_{n=-N}^N (k_y r_{zn} + k_{l,zn} r_{yn}) \exp[-i(k_{xn}x + k_y y - k_{l,zn}z)], \quad (2.45a)$$

$$h_y^{\text{ref}}(x, y, z) = \frac{1}{\omega \mu_0} \sum_{n=-N}^N -(k_{l,zn} r_{xn} + k_{xn} r_{zn}) \exp[-i(k_{xn}x + k_y y - k_{l,zn}z)], \quad (2.45b)$$

$$h_z^{\text{ref}}(x, y, z) = \frac{1}{\omega \mu_0} \sum_{n=-N}^N (k_{xn} r_{yn} - k_y r_{xn}) \exp[-i(k_{xn}x + k_y y - k_{l,zn}z)]. \quad (2.45c)$$

To compute the diffraction efficiency, we need the *Poynting vector* \mathbf{s} , which is defined as the power per unit of area with unit Watt/m² [10]. In mathematical terms \mathbf{s} is given by

$$\mathbf{s} := \text{Re} \{ \mathbf{e} \times \bar{\mathbf{h}} \}. \quad (2.46)$$

The bar indicates the conjugate of a complex number or complex vector. The *power* p through a surface defined by $0 \leq x \leq \Lambda$, $0 \leq y \leq 1$ and $z = \text{constant}$, which is parallel to the mean plane of the grating (parallel to the x - and y -axis), can be computed by taking the inner product of the Poynting vector with the normal of this surface and integrating over it. The normal points in the positive z -direction and therefore we take the inner product with the z -component of the Poynting vector which is given by

$$\begin{aligned} s_z^{\text{ref}}(x, y, z) &= \text{Re} \left\{ e_x^{\text{ref}}(x, y, z) \bar{h}_y^{\text{ref}}(x, y, z) - e_y^{\text{ref}}(x, y, z) \bar{h}_x^{\text{ref}}(x, y, z) \right\} \\ &= -\text{Re} \left\{ \frac{1}{\omega \mu_0} \sum_{n=-N}^N \sum_{m=-N}^N (\bar{k}_{l,zn} r_{xn} \bar{r}_{xm} + k_{xm} r_{xn} \bar{r}_{zm} + k_y r_{yn} \bar{r}_{zm} \right. \\ &\quad \left. + \bar{k}_{l,zn} r_{yn} \bar{r}_{ym}) \exp[-i(k_{xn} - k_{xm})x + i(k_{l,zn} - \bar{k}_{l,zn})z] \right\}. \end{aligned} \quad (2.47)$$

When we integrate $s_z^{\text{ref}}(x, y, z)$ over the specified area, we see that the terms where $m \neq n$ drop out because of the exponential term in x . To distinguish between orders consider the power p_n^{ref} of diffraction order n , viz.

$$\begin{aligned} p_n^{\text{ref}}(z) &:= \text{Re} \left\{ \frac{1}{\omega \mu_0} (-\bar{k}_{l,zn} r_{xn} \bar{r}_{xn} - \bar{k}_{l,zn} r_{yn} \bar{r}_{yn} - k_{l,zn} r_{zn} \bar{r}_{zn}) \cdot \right. \\ &\quad \left. \cdot \exp[i(k_{l,zn} - \bar{k}_{l,zn})z] \right\}. \end{aligned} \quad (2.48)$$

If the refractive index in the upper halfspace is a real number, the dependence on z drops out. If we have an absorbing material with a complex-valued refractive index the dependence on z remains. As already indicated in the previous section, the field amplitudes are normally not expressed in the x -, y - and z coordinates, but by their s - and p -polarized counterparts, viz.

$$p_n^{\text{ref}}(z) = \text{Re} \left\{ \frac{\Lambda}{\omega \mu_0} (-\bar{k}_{l,zn} r_{sn} \bar{r}_{sn} - k_{l,zn} r_{pn} \bar{r}_{pn}) \exp[i(k_{l,zn} - \bar{k}_{l,zn})z] \right\}. \quad (2.49)$$

The derivation of this relation is given in Appendix A.1.

The diffraction efficiency also requires that the power of the incident field is computed. The incident field is given by (2.14) and (2.15). The incident magnetic field \mathbf{h}^{inc} can be

computed as follows

$$h_x^{\text{inc}}(x, y, z) = -\frac{k_0 n_l}{\omega \mu_0} (\cos \psi \sin \phi + \sin \psi \cos \theta \cos \phi) \cdot \exp[-i k_0 n_l (x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta)] \quad (2.50a)$$

$$h_y^{\text{inc}}(x, y, z) = \frac{k_0 n_l}{\omega \mu_0} (\cos \psi \cos \phi - \sin \psi \cos \theta \sin \phi) \cdot \exp[-i k_0 n_l (x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta)] \quad (2.50b)$$

$$h_z^{\text{inc}}(x, y, z) = \frac{k_0 n_l}{\omega \mu_0} \sin \psi \sin \theta \cdot \exp[-i k_0 n_l (x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta)] \quad (2.50c)$$

As in the case of the reflected field, the power of the incident field after integration of the z-component of the Poynting vector, is given by

$$s_z^{\text{inc}}(x, y, z) = \text{Re} \left\{ e_x^{\text{inc}}(x, y, z) \bar{h}_y^{\text{inc}}(x, y, z) - e_y^{\text{inc}}(x, y, z) \bar{h}_x^{\text{inc}}(x, y, z) \right\} = \frac{k_0 n_l}{\omega \mu_0} \cos \theta. \quad (2.51)$$

Integration gives

$$p^{\text{inc}} = \frac{\Lambda k_0 n_l}{\omega \mu_0} \cos \theta. \quad (2.52)$$

Now we can derive an expression for the diffraction efficiency of the reflected field η_{rn} . We compute them for $z = 0$:

$$\eta_{rn} := \left| \frac{p_n^{\text{ref}}(0)}{p^{\text{inc}}} \right| = r_{sn} \bar{r}_{sn} \text{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\} + r_{pn} \bar{r}_{pn} \text{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\}. \quad (2.53)$$

Analogously, the diffraction efficiency η_{tn} for the transmitted orders can be computed for $z = D$ by

$$\eta_{tn} := \left| \frac{p_n^{\text{trans}}(D)}{p^{\text{inc}}} \right| = t_{sn} \bar{t}_{sn} \text{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\} + t_{pn} \bar{t}_{pn} \text{Re} \left\{ \frac{k_{l,zn} n_l}{k_0 n_l^2 \cos \theta} \right\}. \quad (2.54)$$

For planar TE polarization the diffraction efficiencies (2.53) and (2.54) reduce to

$$\eta_{rn} = r_n \bar{r}_n \text{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\}, \quad (2.55a)$$

$$\eta_{tn} = t_n \bar{t}_n \text{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\}. \quad (2.55b)$$

Similar, for planar TM polarization the diffraction efficiencies (2.53) and (2.54) reduce to

$$\eta_{rn} = \tilde{r}_n \bar{\tilde{r}}_n \text{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\}, \quad (2.56a)$$

$$\eta_{tn} = \tilde{t}_n \bar{\tilde{t}}_n \text{Re} \left\{ \frac{k_{l,zn} n_l}{k_0 n_l^2 \cos \theta} \right\}. \quad (2.56b)$$

Diffraction efficiencies have a nice property. Because of conservation of energy it holds that if the diffraction efficiencies of all reflected and transmitted orders are summed, the result is 1 if all materials have a real-valued refractive index and is smaller than 1 if at least one of the materials has a complex-valued refractive index.

2.5 Propagating and evanescent diffraction orders

In Section 2.1 the grating equation has been derived, which provides a way to find the angle for the reflected diffraction orders. However, this equation is only given for planar diffraction. Although it is possible to extend this grating equation for conical diffraction, we will treat it here from a different point of view.

When n_1 is a real number, it holds from its definition (2.25) that $k_{1,zn}$ is either real or purely imaginary. When $k_{1,zn}$ is purely imaginary, $\eta_{r,n}$ is equal to 0. We speak of *evanescent diffraction orders* in that case. If $k_{1,zn}$ is real, we speak of *propagating diffraction orders*. Let us find that n for which $k_{1,zn} = 0$. Because of the square root in (2.30) consider $k_{1,zn}^2$ instead.

$$\begin{aligned} k_{1,zn}^2 &= k_0^2 n_1^2 - k_{xn}^2 - k_y^2 \\ &= k_0^2 (n_1^2 - (n_1 \sin \theta \cos \phi - n(\lambda_0/\Lambda))^2 - (n_1 \sin \theta \sin \phi)^2) \equiv 0. \end{aligned} \quad (2.57)$$

Since this is a quadratic equation with argument n , the solution after some goniometric simplifications is given by

$$n = -n_1 \frac{\Lambda}{\lambda_0} \left(\sin \theta \cos \phi \pm \sqrt{1 - \sin^2 \theta \sin^2 \phi} \right). \quad (2.58)$$

Since $n = 0$ always results in a real-valued $k_{1,zn}$, the interval where $k_{1,zn}$ is real-valued and therefore a non-zero diffraction efficiency for the reflected field is obtained, is given by

$$\begin{aligned} n_1 \frac{\Lambda}{\lambda_0} \left(\sin \theta \cos \phi - \sqrt{1 - \sin^2 \theta \sin^2 \phi} \right) &\leq n \\ &\leq n_1 \frac{\Lambda}{\lambda_0} \left(\sin \theta \cos \phi + \sqrt{1 - \sin^2 \theta \sin^2 \phi} \right). \end{aligned} \quad (2.59)$$

For planar diffraction ($\phi = 0$) this inequality is reduced to:

$$n_1 \frac{\Lambda}{\lambda_0} (-1 - \sin \theta) \leq n \leq n_1 \frac{\Lambda}{\lambda_0} (1 - \sin \theta). \quad (2.60)$$

In conclusion, the propagating diffraction orders must satisfy (2.59) in case of conical diffraction and (2.60) in case of planar diffraction.

Chapter 3

RCWA

In 1981 the rigorous coupled-wave analysis (RCWA) was introduced [35] to compute a solution of Maxwell's equations for diffraction by grating structures. This method uses the fact that the field above and below the grating can be described by the Rayleigh expansion as derived in Section 2.3. To find the reflected and transmitted field amplitudes of these Rayleigh expansions, an expression should be found for the field inside the grating. Since the tangential field components are continuous, the field inside the grating can be connected to the Rayleigh expansions such that we can find the field amplitudes. The main problem inside the grating is that the refractive index depends on both x and z , while the periodicity of the grating is only in the x -direction. Applying Fourier techniques to obtain an algebraic system is not possible unless we eliminate this z -dependence. A possible way to eliminate this direction is to divide the grating domain into thin horizontal layers in which the permittivity can be assumed piecewise constant. Expansions in Fourier series of both the field and the permittivity lead to an algebraic eigenvalue system for every layer. Connecting the solutions of the fields at the layer interfaces by means of the continuity of the tangential field components ensures that a linear system can be found with only the reflected and transmitted field amplitudes as unknowns. Solving this system directly appears to be unstable for thick layers and it took until 1995 to find a stable way of solving this system by the so-called enhanced transmittance matrix approach [36, 37].

Since the introduction of the RCWA algorithm, questions have arisen how many terms or *harmonics* in the Fourier series should be taken into account and how many layers are required to approximate the geometry sufficiently accurate. In 1996 a correction was introduced to handle products of truncated series [25]. This improved the convergence of the solution for TM polarization with respect to the number of harmonics. Nowadays RCWA is an easy to use algorithm, which is applicable to a wide range of multi-layer grating structures.

In this chapter the RCWA algorithm will be discussed for both planar and conical diffraction. The convergence issues are treated at the end of this chapter.

3.1 The key feature of RCWA

To obtain algebraic eigenvalue systems, the dependence of all other directions than the periodic x -direction should be removed. RCWA removes this dependency by introducing K horizontal layers in the grating domain (see Figure 3.1) such that the relative permittivity $\tilde{\epsilon}^r$ inside such a layer is approximated by a piecewise constant function of x only. This implies that the field is computed for a staircase approximation of the grating profile. In general, the permittivity of an arbitrary grating is given by

$$\tilde{\epsilon}_i^r(x) = \begin{cases} \tilde{\epsilon}_{i,1}^r, & -\Lambda/2 \leq x \leq t_{i1}, \\ \tilde{\epsilon}_{i,2}^r, & t_{i1} \leq x \leq t_{i2}, \\ \vdots & \vdots \\ \tilde{\epsilon}_{i,M}^r, & t_{i(M-1)} \leq x \leq \Lambda/2, \end{cases} \quad (3.1)$$

for $-\Lambda/2 < t_{i1} < \dots < t_{i(M-1)} < \Lambda/2$, where t_{ik} is the k th position of a material transition inside layer i and M is the number of blocks inside a layer. The elimination of the z -coordinate ensures that the expansions in Fourier series of the relative permittivity and its

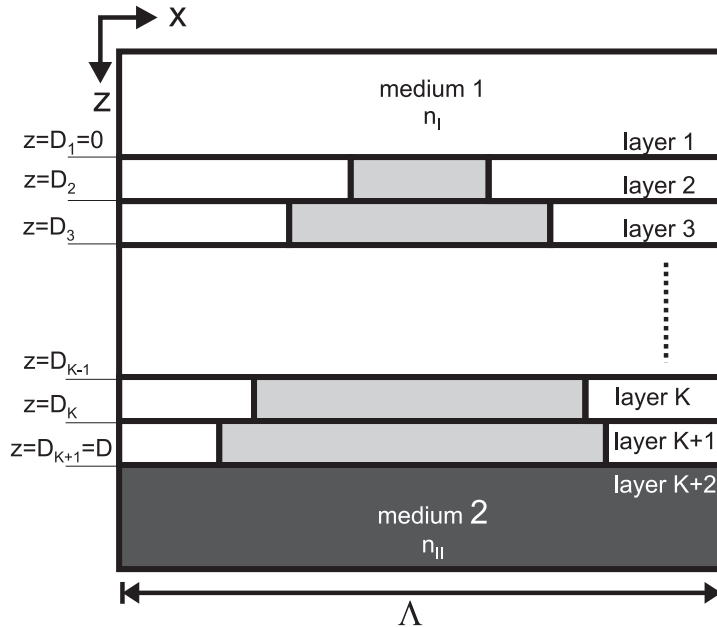


Figure 3.1: The definition of the RCWA shape parameters.

reciprocal have constant coefficients, viz.

$$\tilde{\epsilon}_i^r(x) = \sum_n \epsilon_{i,n} \exp[i \frac{2\pi}{\Lambda} nx], \quad \epsilon_{i,n} = \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} \tilde{\epsilon}_i^r(x) \exp[-i \frac{2\pi}{\Lambda} nx] dx, \quad (3.2a)$$

$$\frac{1}{\tilde{\epsilon}_i^r(x)} = \sum_n \pi_{i,n} \exp[i \frac{2\pi}{\Lambda} nx], \quad \pi_{i,n} = \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} \frac{1}{\tilde{\epsilon}_i^r(x)} \exp[-i \frac{2\pi}{\Lambda} nx] dx. \quad (3.2b)$$

Note that the sum over n can represent an infinite sum in theory, but in practice it will be a finite sum. Because of the piecewise constant approximation of the permittivity the convergence of these Fourier series will be similar to that of the series $1/n$.

Also the electric and magnetic field can be defined for every layer, viz.

$$\mathbf{e}(\mathbf{x}) = \mathbf{e}_i(\mathbf{x}) \text{ and } \mathbf{h}(\mathbf{x}) = \mathbf{h}_i(\mathbf{x}), \quad (3.3)$$

where $-\Lambda/2 \leq x \leq \Lambda/2$, $y \in \mathbb{R}$ and $D_{i-1} < z < D_i$ and $i = 2, \dots, K + 1$. The layer thickness of layer i is denoted by D_i . Note that the fields still depend on all coordinates. In contrast to the permittivity, the fields are not periodic, but pseudo-periodic as indicated in Section 2.2.3. This implies that we can perform an expansion in a Fourier series with an additional phase factor in the x -direction. Because the grating structure is independent of the y -direction, the incident field ensures a phase factor in this direction. As a result the fields can be described by

$$\mathbf{e}_i(x, y, z) = \sum_n \mathbf{s}_{i,n}(z) \exp[-i(k_x n x + k_y y)], \quad (3.4a)$$

$$\mathbf{h}_i(x, y, z) = -i\sqrt{\epsilon_0/\mu_0} \sum_n \mathbf{u}_{i,n}(z) \exp[-i(k_x n x + k_y y)], \quad (3.4b)$$

where $\mathbf{s}_{i,n}$ and $\mathbf{u}_{i,n}$ are called the *electric* or *magnetic field components*. The correction factor for the magnetic field is from Maxwell's equations.

Before we discuss how the RCWA algorithm finds the field amplitudes, a mathematical issue has to be addressed first. When dealing with finite series, special attention has to be given to multiplications. This issue will be discussed in the next section.

3.2 General theory of truncating Fourier series

Any periodic function can be expanded in a Fourier series, resulting in an (infinite) sum of modes. In practice only finite sums can be handled and special care should be taken when products occur between two truncated Fourier series. Let us introduce two definitions.

Definition 3.1 *Two piecewise continuous functions f and g are said to have a pair of concurrent jump discontinuities at x_0 if*

$$\lim_{x \downarrow x_0} f(x) \neq \lim_{x \uparrow x_0} f(x) \text{ and } \lim_{x \downarrow x_0} g(x) \neq \lim_{x \uparrow x_0} g(x). \quad (3.5)$$

Definition 3.2 If two functions f and g have a pair of concurrent jump discontinuities, the jumps are said to be complementary if

$$\lim_{x \uparrow x_0} f(x)g(x) = \lim_{x \downarrow x_0} f(x)g(x). \quad (3.6)$$

The use of Laurent's rule, which is always valid when dealing with infinite series, is not always the best choice for truncated Fourier series as indicated by the following rules [25].

Property 3.3 If f or g has no concurrent jump discontinuities, then $h(x) = f(x)g(x)$ can be expanded in a Fourier series by Laurent's rule, viz.

$$h_n = \lim_{N \rightarrow \infty} \sum_{m=-N}^N f_{n-m} g_m, \quad (3.7)$$

where f_n , g_n and h_n are the Fourier coefficients of f , g and h , respectively.

Property 3.4 If all the concurrent jump discontinuities of f and g are pairwise complementary, then $h(x) = f(x)g(x)$ can be expanded in a Fourier series by the inverse rule, viz.

$$h_n = \lim_{N \rightarrow \infty} \sum_{m=-N}^N \tilde{f}_{nm} g_m, \quad (3.8)$$

where \tilde{f}_{nm} is the (n, m) th element of the inverse matrix of the Toeplitz matrix where the (n, m) th element equals f_{n-m} , provided that the inverse matrix exists.

Property 3.5 If f and g have concurrent but noncomplementary jump discontinuities, then $h(x) = f(x)g(x)$ cannot be expanded in a Fourier series by either Laurent's rule or the inverse rule.

The proof of these rules can be found in [28]. These rules play a crucial role in the RCWA algorithm as will be shown in the next section.

3.3 The field inside a layer

In Section 2.2.1 relations were derived to describe the electric and magnetic field for both planar and conical diffraction. This section discusses the substitution of the Fourier series of the fields and material properties into relations (2.11) or (2.13), for conical or planar diffraction, respectively. The way to derive an algebraic eigensystem for every layer differs significantly for both cases. Therefore we will treat them separately.

3.3.1 Planar diffraction: TE polarization

For TE polarization the relation for the y -component of the electric field is given by (2.13a). Let us introduce the slicing of the grating domain, which means that (3.1) is used for the permittivity and (3.3) for the electric field. As a result the governing equation for the electric field becomes

$$\frac{\partial^2}{\partial x^2} e_{i,y}(x, z) + \frac{\partial^2}{\partial z^2} e_{i,y}(x, z) = -k_0^2 \tilde{\epsilon}_i^r(x) e_{i,y}(x, z), \quad (3.9)$$

for $i = 2, \dots, K+1$. The right-hand side of this relation contains a product of two functions of which only the relative permittivity is a discontinuous function. Thus, no concurrent jump discontinuities occur and we can apply Laurent's rule to the finite Fourier series of the y -component of the electric field (3.4a) and the relative permittivity, viz.

$$\begin{aligned} & \sum_{n=-N}^N \frac{d^2}{dz^2} s_{i,n}(z) \exp[-ik_x n x] \\ &= -k_0^2 \sum_{n=-N}^N \sum_{m=-N}^N \epsilon_{i,n-m} s_{i,m} \exp[-ik_x n x] + \sum_{n=-N}^N k_{x,n}^2 s_{i,n}(z) \exp[-ik_x n x]. \end{aligned} \quad (3.10)$$

Since (3.10) holds for all x in $-\Lambda/2 \leq x \leq \Lambda/2$, it follows that

$$\frac{d^2}{dz^2} s_{i,n}(z) = -k_0^2 \sum_{m=-N}^N \epsilon_{i,n-m} s_{i,m}(z) + k_{x,n}^2 s_{i,n}(z), \quad (3.11)$$

for $-N \leq n \leq N$. The relation above can be read as a matrix relation of dimension $(2N+1) \times (2N+1)$ for the unknown electric field coefficients $s_{i,n}$ collected in a $(2N+1)$ -vector $\mathbf{s}_i(z)$. Introducing a new variable $z' := k_0 z$ and dividing all terms by k_0^2 results in the following matrix equation

$$\frac{d^2}{dz'^2} \mathbf{s}_i(z') = \mathcal{A}_i \mathbf{s}_i(z'), \quad (3.12)$$

with the auxiliary matrix

$$\mathcal{A}_i := \mathbf{K}_x^2 - \mathbf{E}_i, \quad (3.13)$$

where \mathbf{K}_x is a diagonal matrix with the entries $k_{x,n}/k_0$ on its diagonal for $-N \leq n \leq N$ and \mathbf{E}_i is a Toeplitz matrix where the (n, m) th entry equals $\epsilon_{i,n-m}$ for $-N \leq n, m \leq N$.

Relation (3.12) is a system of second-order differential equations with constant coefficients. In Appendix A.2.2 we derive the solution of such a second-order differential equation. As a result, the field component vector $\mathbf{s}_{i,n}(z)$ is given by

$$\mathbf{s}_i(z) = \sum_{n=1}^{2N+1} \mathbf{w}_{i,n} (c_{i,n}^+ \exp[-k_0 q_{i,n}(z - D_{i-1})] + c_{i,n}^- \exp[k_0 q_{i,n}(z - D_i)]), \quad (3.14)$$

where $q_{i,n}$ is the square root with a positive real part of the n th eigenvalue of \mathbf{A}_i and $\mathbf{w}_{i,n}$ is its n th eigenvector. The field constants $c_{i,n}^+$ and $c_{i,n}^-$ are still unknown. Let us define matrix \mathbf{W}_i as the matrix with the eigenvectors $\mathbf{w}_{i,n}$ as its columns and \mathbf{Q}_i as a diagonal matrix with the square roots $q_{i,n}$ on its diagonal. When all field constants $c_{i,n}^+$ and $c_{i,n}^-$ are collected in the vectors \mathbf{c}_i^+ and \mathbf{c}_i^- , respectively, (3.14) can be written as

$$\mathbf{s}_i(z) = \mathbf{W}_i \left(\exp[-k_0 \mathbf{Q}_i (z - D_{i-1})] \mathbf{c}_i^+ + \exp[k_0 \mathbf{Q}_i (z - D_i)] \mathbf{c}_i^- \right). \quad (3.15)$$

Thus, the problem of finding the field inside a layer has been reduced to finding the field constants.

3.3.2 Planar diffraction: TM polarization

For TM polarization, relation (2.13b) holds for the y -component of the magnetic field. After the substitution of the slicing of the grating domain, which means that (3.1) is used for the permittivity and (3.3) for the magnetic field, the governing equation for the magnetic field becomes

$$\frac{\partial^2}{\partial z^2} h_{i,y}(x, z) = -k_0^2 \tilde{\epsilon}_i^r(x) h_{i,y}(x, z) - \tilde{\epsilon}_i^r(x) \frac{\partial}{\partial x} \left(\frac{1}{\tilde{\epsilon}_i^r(x)} \frac{\partial}{\partial x} h_{i,y}(x, z) \right), \quad (3.16)$$

for $i = 2, \dots, K + 1$. When we introduce the Fourier series, we have to take care of the theory of products of truncated series as given in Section 3.2. On the right-hand side of (3.16) some discontinuous quantities appear such as $\tilde{\epsilon}_i^r$ and $(\partial h_{i,y})/(\partial x)$. Since Maxwell's equation (2.8f) says that

$$e_{i,z}(x, z) = \frac{1}{i\omega \epsilon_0 \tilde{\epsilon}_i^r(x)} \frac{\partial}{\partial x} h_{i,y}(x, z), \quad (3.17)$$

and since $e_{i,z}$ is continuous, we have complementary concurrent jump discontinuities for the reciprocal of the permittivity and the partial derivative of the y -component of the magnetic field with respect to x . Also $\partial^2 h_{i,y}/\partial z^2$ is continuous, which implies that the entire right-hand side of (3.16) has concurrent jump discontinuities that are complementary. Applying the inverse rule (3.8) twice to (3.16) gives

$$\begin{aligned} & \sum_{n=-N}^N \frac{d^2}{dz^2} u_{i,n}(z) \exp[-ik_x n x] \\ &= \sum_{n=-N}^N \sum_{m=-N}^N \tilde{\pi}_{i,nm} \left(-k_0^2 u_{i,m}(z) + \sum_{r=-N}^N k_{xm} \tilde{\epsilon}_{i,mr} k_{xr} u_{i,r}(z) \right) \exp[-ik_x n x], \end{aligned} \quad (3.18)$$

where $\tilde{\epsilon}_{i,nm}$ is the (n, m) th element of the inverse matrix of \mathbf{E}_i with entries $(\mathbf{E}_i)_{nm} = \epsilon_{i,n-m}$ and $\tilde{\pi}_{i,nm}$ is similarly defined as the (n, m) th element of the inverse matrix of \mathbf{P}_i with entries $(\mathbf{P}_i)_{nm} = \pi_{i,n-m}$. Since (3.18) holds for $-\Lambda/2 < x < \Lambda/2$, we have

$$\frac{d^2}{dz^2} u_{i,n}(z) = \sum_{m=-N}^N \tilde{\pi}_{i,nm} \left(-k_0^2 u_{i,m}(z) + \sum_{r=-N}^N k_{xm} \tilde{\epsilon}_{i,mr} k_{xr} u_{i,r}(z) \right), \quad (3.19)$$

for $-N \leq n \leq N$. Similar to the TE polarization case, the relation above can be read as a matrix relation of dimension $(2N+1) \times (2N+1)$ for the unknown magnetic field coefficients $u_{i,n}(z)$ collected in a $(2N+1)$ -vector $\mathbf{u}_i(z)$. Introducing a new variable $z' := k_0 z$ and dividing all terms by k_0^2 gives the following matrix relation

$$\frac{d^2}{dz'^2} \mathbf{u}_i(z') = \mathbf{P}_i^{-1} \mathbf{B}_i \mathbf{u}_i(z'), \quad (3.20)$$

with the auxiliary matrix

$$\mathbf{B}_i := (\mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{K}_x - \mathbf{I}). \quad (3.21)$$

Similar to TE polarization, the field component vector $\mathbf{u}_{i,n}(z)$ can be found by using the method described in Appendix A.2.2 to obtain the solution of a second-order differential equation with constant coefficients, viz.

$$\mathbf{u}_i(z) = \sum_{n=1}^{2N+1} \mathbf{w}_{i,n} (c_{i,n}^+ \exp[-k_0 q_{i,n}(z - D_{i-1})] + c_{i,n}^- \exp[k_0 q_{i,n}(z - D_i)]), \quad (3.22)$$

where $q_{i,n}$ is the square root with a positive real part of the n th eigenvalue and $\mathbf{w}_{i,n}$ is the n th eigenvector of $\mathbf{P}_i^{-1} \mathbf{B}_i$. The field constants $c_{i,n}^+$ and $c_{i,n}^-$ are still unknown. Similarly to TE polarization (3.22) can be written as a matrix relation, viz.

$$\mathbf{u}_i(z) = \mathbf{W}_i (\exp[-k_0 \mathbf{Q}_i(z - D_{i-1})] \mathbf{c}_i^+ + \exp[k_0 \mathbf{Q}_i(z - D_i)] \mathbf{c}_i^{-1}). \quad (3.23)$$

The problem of finding the field inside a layer has been reduced to finding the field constants.

3.3.3 Conical diffraction

The decoupling of the electric and magnetic field is not as straightforward as in planar diffraction. Only after the introduction of the slicing of the grating domain we are able to find two relations for only one component of the electric and magnetic field. When (3.1) is used for the permittivity and (3.3) for the electromagnetic fields, (2.11) give two decoupled relations for the x -components of the electric and magnetic field. In other words,

$$\begin{aligned} \frac{\partial^2}{\partial x^2} e_{i,x}(x, y, z) + \frac{\partial^2}{\partial y^2} e_{i,x}(x, y, z) + \frac{\partial^2}{\partial z^2} e_{i,x}(x, y, z) \\ = -k_0^2 \tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z) - \frac{\partial}{\partial x} \left(\frac{1}{\tilde{\epsilon}_i^r(x)} \frac{\partial}{\partial x} \tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z) \right), \end{aligned} \quad (3.24a)$$

$$\begin{aligned} \frac{\partial^2}{\partial x^2} h_{i,x}(x, y, z) + \frac{\partial^2}{\partial y^2} h_{i,x}(x, y, z) + \frac{\partial^2}{\partial z^2} h_{i,x}(x, y, z) \\ = -k_0^2 \tilde{\epsilon}_i^r(x) h_{i,x}(x, y, z), \end{aligned} \quad (3.24b)$$

for $i = 2, \dots, K+1$. The relations of the y - or z -components of the electric and magnetic field are still coupled to other components. The first relation has to be reorganized to

ensure that we handle products of truncated Fourier series in a correct way. Thus,

$$\begin{aligned} \frac{\partial^2}{\partial z^2} e_{i,x}(x, y, z) &= -\frac{\partial^2}{\partial y^2} e_{i,x}(x, y, z) - k_0^2 \tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z) \\ &\quad + \frac{\partial}{\partial x} \left(\frac{1}{\tilde{\epsilon}_i^r(x)} \frac{\partial}{\partial x} (\tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z)) \right). \end{aligned} \quad (3.25)$$

The right-hand side of the relation above contains complementary concurrent jump discontinuities. That $\tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z)$ is continuous, is a direct consequence of using Maxwell's equation (2.8d). To see that also $(1/\tilde{\epsilon}_i^r(x))(\partial/\partial x)(\tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z))$ is continuous, first substitute (2.8d) and then use (2.8e) and (2.8f) to show that

$$\frac{1}{\tilde{\epsilon}_i^r(x)} \frac{\partial}{\partial x} (\tilde{\epsilon}_i^r(x) e_{i,x}(x, y, z)) = \frac{\partial e_{i,y}(x, y, z)}{\partial y} + \frac{\partial e_{i,z}(x, y, z)}{\partial z}. \quad (3.26)$$

Since the right-hand side of this relation consists only of continuous functions, the left-hand side has a complementary pair of discontinuities. To conclude, we can apply the inverse rule (3.8) to the products of the Fourier series. Similarly as for planar diffraction, introduce $\tilde{\epsilon}_{i,nm}$ as the (n, m) th elements of the inverse matrix of \mathbf{E}_i and $\tilde{\pi}_{i,nm}$ as the (n, m) th elements of the inverse matrix of \mathbf{P}_i . Then (3.25) becomes

$$\begin{aligned} &\sum_{n=-N}^N \frac{d^2}{dz^2} s_{i,xn}(z) \exp[-i(k_{xn}x + k_y y)] \\ &= k_y^2 \sum_{n=-N}^N s_{i,xn}(x, y, z) \exp[-i(k_{xn}x + k_y y)] \\ &\quad - k_0^2 \sum_{n=-N}^N \sum_{m=-N}^N \tilde{\pi}_{i,nm} s_{i,xm}(z) \exp[-i(k_{xn}x + k_y y)] \\ &\quad - \sum_{n=-N}^N \sum_{m=-N}^N \sum_{r=-N}^N k_{xn} \tilde{\epsilon}_{i,nm} k_{xm} \tilde{\pi}_{i,mr} s_{i,xr} \exp[-i(k_{xn}x + k_y y)]. \end{aligned} \quad (3.27)$$

Similar to planar diffraction, the relation above can be read as a matrix relation of dimension $(2N+1) \times (2N+1)$ for the unknown electric field coefficients $s_{i,xn}(z)$ collected in a $(2N+1)$ -vector $\mathbf{s}_{i,x}(z)$. Introducing a new variable $z' := k_0 z$ and dividing all terms by k_0^2 gives the following matrix equation

$$\frac{d^2}{dz'^2} \mathbf{s}_{i,x}(z') = \mathbf{C}_i \mathbf{s}_{i,x}(z'), \quad (3.28)$$

with auxiliary matrix

$$\mathbf{C}_i := \mathbf{K}_y^2 + \mathbf{B}_i \mathbf{P}_i^{-1}, \quad (3.29)$$

where \mathbf{B}_i is given by (3.21) and $\mathbf{K}_y = (k_y/k_0)\mathbf{I}$.

Relation (3.24b) contains a product of two functions where only one function is discontinuous. This implies that Laurent's rule can be applied to all products.

$$\begin{aligned}
& \sum_{n=-N}^N u_{i,xn}(z) \exp[-i(k_x n x + k_y y)] \\
&= \sum_{n=-N}^N k_x^2 u_{i,xn}(z) \exp[-i(k_x n x + k_y y)] + k_y^2 \sum_{n=-N}^N u_{i,xn}(z) \exp[-i(k_x n x + k_y y)] \\
&\quad - k_0^2 \sum_{n=-N}^N \sum_{m=-N}^N \epsilon_{i,nm} u_{i,xm}(z) \exp[-i(k_x n x + k_y y)]. \tag{3.30}
\end{aligned}$$

Introduce a $(2N + 1)$ -vector $\mathbf{u}_{i,x}$ with $u_{i,xn}$ as its entries for $-N \leq n \leq N$. Let $z' := k_0 z$ and divide (3.30) by k_0^2 to obtain

$$\frac{d^2}{dz'^2} \mathbf{u}_{i,x}(z') = \mathcal{D}_i \mathbf{u}_{i,x}(z'), \tag{3.31}$$

with auxiliary matrix

$$\mathcal{D}_i := (\mathbf{K}_y^2 + \mathcal{A}_i). \tag{3.32}$$

Matrix \mathcal{A}_i is defined in (3.13). Like for planar diffraction, the field component vectors $\mathbf{s}_{i,x}(z)$ and $\mathbf{u}_{i,x}(z)$ can be found by using Appendix A.2.2, where we derive a form of the solution of a second-order differential equation with constant coefficients, viz.

$$\begin{aligned}
\mathbf{s}_{i,x}(z) = & \sum_{n=-N}^N \mathbf{w}_{2,i,n} (c_{2,i,n}^+ \exp[-k_0 q_{2,i,n}(z - D_{i-1})] \\
& + c_{2,i,n}^- \exp[k_0 q_{2,i,n}(z - D_i)]), \tag{3.33}
\end{aligned}$$

$$\begin{aligned}
\mathbf{u}_{i,x}(z) = & \sum_{n=-N}^N \mathbf{w}_{1,i,n} (c_{1,i,n}^+ \exp[-k_0 q_{1,i,n}(z - D_{i-1})] \\
& + c_{1,i,n}^- \exp[k_0 q_{1,i,n}(z - D_i)]), \tag{3.34}
\end{aligned}$$

where $q_{i,1,n}$ and $q_{i,2,n}$ are the square roots with a positive real part of the n th eigenvalue and $\mathbf{w}_{1,i,n}$ and $\mathbf{w}_{2,i,n}$ are the n th eigenvectors of \mathcal{C}_i and \mathcal{D}_i , which are defined by (3.29) and (3.32), respectively. The field constants $c_{1,i,n}^+$, $c_{1,i,n}^-$, $c_{2,i,n}^+$ and $c_{2,i,n}^-$ are still unknown. Let all eigenvectors $\mathbf{w}_{1,i,n}$ and $\mathbf{w}_{2,i,n}$ be collected in the matrices $\mathbf{W}_{1,i}$ and $\mathbf{W}_{2,i}$, respectively, and all eigenvalues $q_{1,i,n}$ and $q_{2,i,n}$ be the diagonal entries of $\mathbf{Q}_{1,i}$ and $\mathbf{Q}_{2,i}$, respectively. When the unknown field constants $c_{1,i,n}^+$, $c_{1,i,n}^-$, $c_{2,i,n}^+$ and $c_{2,i,n}^-$ are collected in vectors $\mathbf{c}_{1,i}^+$, $\mathbf{c}_{1,i}^-$, $\mathbf{c}_{2,i}^+$ and $\mathbf{c}_{2,i}^-$, respectively, the field components $\mathbf{s}_{i,x}$ and $\mathbf{u}_{i,x}$ can be given as

$$\mathbf{s}_{i,x}(z) = \mathbf{W}_{2,i} (\exp[-k_0 \mathbf{Q}_{2,i}(z - D_{i-1})] \mathbf{c}_{2,i}^+ + \exp[k_0 \mathbf{Q}_{2,i}(z - D_i)] \mathbf{c}_{2,i}^-), \tag{3.35}$$

$$\mathbf{u}_{i,x}(z) = \mathbf{W}_{1,i} (\exp[-k_0 \mathbf{Q}_{1,i}(z - D_{i-1})] \mathbf{c}_{1,i}^+ + \exp[k_0 \mathbf{Q}_{1,i}(z - D_i)] \mathbf{c}_{1,i}^-). \tag{3.36}$$

This concludes the description of the fields inside a layer. The field constants remain to be determined and for this purpose we will use boundary conditions at the interface as will be discussed in the next section.

3.4 Boundary conditions

To connect the solutions of the fields inside every layer to each other, boundary conditions are required. From the boundary conditions discussed in Section 2.2.3 we have not yet used the continuity of the tangential field components. When the layers are connected, the tangential field components are the x - and y -components of both the electric and magnetic field. Thus,

$$e_{i,x}(x, y, D_i) = e_{i+1,x}(x, y, D_i), \quad (3.37a)$$

$$e_{i,y}(x, y, D_i) = e_{i+1,y}(x, y, D_i), \quad (3.37b)$$

$$h_{i,x}(x, y, D_i) = h_{i+1,x}(x, y, D_i), \quad (3.37c)$$

$$h_{i,y}(x, y, D_i) = h_{i+1,y}(x, y, D_i), \quad (3.37d)$$

for $i = 1, \dots, K+1$. Because the fields above and below the grating structure are described by their Rayleigh expansion, we have to reformulate the boundary conditions in terms of field amplitudes (from the Rayleigh expansions) and the field component (from the expansions of the fields inside a layer). We will again make a distinction between planar and conical diffraction.

3.4.1 Planar diffraction

For planar diffraction only the y -component of the electric or magnetic field is present. From Maxwell's equations (2.8b) and (2.8e) we can find the x -component of the magnetic field when the y -component of the electric field is present and the x -component of the electric field if we have a y -component of the magnetic field, viz.

$$h_{i,x}(x, z) = -\frac{i}{\omega\mu_0} \frac{\partial}{\partial z} e_{i,y}(x, z), \quad (3.38a)$$

$$e_{i,x}(x, z) = \frac{i}{\omega\tilde{\epsilon}_i(x)} \frac{\partial}{\partial z} h_{i,y}(x, z). \quad (3.38b)$$

For TE polarization only $e_{i,y}$ and $h_{i,x}$ are present and therefore boundary conditions (3.37b) and (3.37c) are the only relevant ones. Since we have only the field components $s_{i,n}$ available that belong to the y -component of the electric field, we use (3.38a) to eliminate the x -component of the magnetic field from boundary condition (3.37c). As a result, boundary conditions (3.37b) and (3.37c) become

$$e_{i,y}(x, D_i) = e_{i+1,y}(x, D_i), \quad (3.39a)$$

$$\frac{\partial}{\partial z} e_{i,y}(x, D_i) = \frac{\partial}{\partial z} e_{i+1,y}(x, D_i). \quad (3.39b)$$

Since for TE polarization the y -component of the electric field has been expressed either by the Rayleigh expansion in the upper and lower halfspace or by its expansion in Fourier

series inside a layer, we can express the boundary equations (3.39) in terms of the field amplitudes and field components, viz.

$$\mathbf{d}_0 + \mathbf{r} = \mathbf{s}_2(0), \quad (3.40a)$$

$$-i\mathbf{d}_0 n \cos \theta + i\mathbf{Y}_I \mathbf{r} = \frac{1}{k_0} \frac{d}{dz} \mathbf{s}_2(0), \quad (3.40b)$$

$$\mathbf{s}_i(D_i) = \mathbf{s}_{i+1}(D_i), \quad (3.40c)$$

$$\frac{1}{k_0} \frac{d}{dz} \mathbf{s}_i(D_i) = \frac{1}{k_0} \frac{d}{dz} \mathbf{s}_{i+1}(D_i), \quad (3.40d)$$

$$\mathbf{s}_{K+1}(D) = \mathbf{t}, \quad (3.40e)$$

$$\frac{1}{k_0} \frac{d}{dz} \mathbf{s}_{K+1}(D) = -i\mathbf{Y}_{II} \mathbf{t}, \quad (3.40f)$$

where \mathbf{Y}_I and \mathbf{Y}_{II} are diagonal matrices with $k_{i,zn}/k_0$ and $k_{||,zn}/k_0$ on its diagonal for $-N \leq n \leq N$. The vector \mathbf{d}_0 is an all-zero vector of size $(2N + 1)$ except for entry $N + 1$ which is equal to 1.

Similarly, for TM polarization only solutions of the field components $u_{i,n}$ are available that belong to the y -component of the magnetic field. Therefore, the x -component of the electric field should be removed from boundary condition (3.37a). This is done by substituting (3.38b) in this boundary condition. Together with the boundary condition (3.37d) they are written as

$$h_{i,y}(x, D_i) = h_{i+1,y}(x, D_i), \quad (3.41a)$$

$$\frac{1}{\tilde{\epsilon}_i^r(x)} \frac{\partial}{\partial z} h_{i,y}(x, D_i) = \frac{1}{\tilde{\epsilon}_{i+1}^r(x)} \frac{\partial}{\partial z} h_{i+1,y}(x, D_i). \quad (3.41b)$$

Note that $\tilde{\epsilon}_i^r$ is a discontinuous function in x , but $\partial h_{i,y}/\partial z$ is not. Therefore Laurent's rule (3.7) can be applied here. Like for TE polarization, we can express the y -component of the magnetic field by either its Rayleigh expansion or its Fourier series. As a result, we can write boundary equations (3.41) in terms of the field amplitudes and field components, viz.

$$\mathbf{d}_0 + \mathbf{r} = \mathbf{u}_2(0), \quad (3.42a)$$

$$-i\mathbf{d}_0 \frac{\cos \theta}{n_1} + i\mathbf{Z}_I \mathbf{r} = \frac{1}{k_0} \mathbf{P}_2 \frac{d}{dz} \mathbf{u}_2(0), \quad (3.42b)$$

$$\mathbf{u}_i(D_i) = \mathbf{u}_{i+1}(D_i), \quad (3.42c)$$

$$\frac{1}{k_0} \mathbf{P}_i \frac{d}{dz} \mathbf{u}_i(D_i) = \frac{1}{k_0} \mathbf{P}_{i+1} \frac{d}{dz} \mathbf{u}_{i+1}(D_i), \quad (3.42d)$$

$$\mathbf{u}_{K+1}(D) = \mathbf{t}, \quad (3.42e)$$

$$\frac{1}{k_0} \mathbf{P}_{K+1} \frac{d}{dz} \mathbf{u}_{K+1}(D) = -i \mathbf{Z}_{\parallel} \mathbf{t}, \quad (3.42f)$$

where \mathbf{Z}_{\parallel} and \mathbf{Z}_{\parallel} are diagonal matrices with $k_{i,zn}/k_0 n_i^2$ and $k_{\parallel,zn}/k_0 n_{\parallel}^2$ on their diagonal for $-N \leq n \leq N$.

3.4.2 Conical diffraction

To implement the boundary conditions (3.37b) and (3.37d) we have to find an expression for the y -components of both electric and magnetic field, $e_{i,y}$ and $h_{i,y}$, respectively. As stated before, Maxwell's equations (2.8c) and (2.8f) enable us to remove the z -component of the electric and magnetic field. As a result Maxwell's equations (2.8b) and (2.8e) become

$$\begin{aligned} \frac{\partial}{\partial z} e_{i,x}(x, y, z) &= -i\omega\mu_0 h_{i,y}(x, y, z) \\ &+ \frac{i}{\omega} \frac{\partial}{\partial x} \left(\frac{1}{\tilde{\epsilon}_i(x)} \frac{\partial}{\partial y} h_{i,x}(x, y, z) - \frac{1}{\tilde{\epsilon}_i(x)} \frac{\partial}{\partial x} h_{i,y}(x, y, z) \right), \end{aligned} \quad (3.43a)$$

$$\begin{aligned} \frac{\partial}{\partial z} h_{i,x}(x, y, z) &= i\omega\tilde{\epsilon}_i(x) e_{i,y}(x, y, z) \\ &+ \frac{i}{\omega\mu_0} \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} e_{i,y}(x, y, z) - \frac{\partial}{\partial y} e_{i,x}(x, y, z) \right), \end{aligned} \quad (3.43b)$$

which are two coupled equations with only the x - and y -components of the electric and magnetic fields. As can be seen it is not possible to isolate the x - or y -components of the electric and magnetic field as we did in the planar diffraction case. This implies that all four boundary conditions (3.37) remain in their original form. Substituting the expansions in Fourier series for the x - and y -components of the electric and magnetic field into (3.37) and collecting all field components in the $(2N + 1)$ -vectors $\mathbf{s}_{i,x}$, $\mathbf{s}_{i,y}$, $\mathbf{u}_{i,x}$ and $\mathbf{u}_{i,y}$, respectively, gives

$$\mathbf{d}_0 (\cos \psi \cos \theta \cos \phi - \sin \psi \sin \phi) + \mathbf{r}_x = \mathbf{s}_{2,x}(0), \quad (3.44a)$$

$$\mathbf{d}_0 (\cos \psi \cos \theta \sin \phi + \sin \psi \cos \phi) + \mathbf{r}_y = \mathbf{s}_{2,y}(0), \quad (3.44b)$$

$$i\mathbf{d}_0 n_1 (-\cos \psi \sin \phi - \sin \psi \cos \theta \cos \phi) + i\mathbf{K}_y \mathbf{r}_z + i\mathbf{Y}_1 \mathbf{r}_y = \mathbf{u}_{2,x}(0), \quad (3.44c)$$

$$i\mathbf{d}_0 n_1 (\cos \psi \cos \phi - \sin \psi \cos \theta \sin \phi) - i\mathbf{Y}_1 \mathbf{r}_x - i\mathbf{K}_x \mathbf{r}_z = \mathbf{u}_{2,y}(0), \quad (3.44d)$$

$$\mathbf{s}_{i,x}(D_i) = \mathbf{s}_{i+1,x}(D_i), \quad (3.44e)$$

$$\mathbf{s}_{i,y}(D_i) = \mathbf{s}_{i+1,y}(D_i), \quad (3.44f)$$

$$\mathbf{u}_{i,x}(D_i) = \mathbf{u}_{i+1,x}(D_i), \quad (3.44g)$$

$$\mathbf{u}_{i,y}(D_i) = \mathbf{u}_{i+1,y}(D_i), \quad (3.44h)$$

$$\mathbf{s}_{K+1,x}(D) = \mathbf{t}_x, \quad (3.44i)$$

$$\mathbf{s}_{K+1,y}(D) = \mathbf{t}_y, \quad (3.44j)$$

$$\mathbf{u}_{K+1,x}(D) = i\mathbf{K}_y \mathbf{t}_z - i\mathbf{Y}_{||} \mathbf{t}_y, \quad (3.44k)$$

$$\mathbf{u}_{K+1,y}(D) = i\mathbf{Y}_{||} \mathbf{t}_x - i\mathbf{K}_x \mathbf{t}_z. \quad (3.44l)$$

The field components $s_{i,y,n}$ and $u_{i,y,n}$ can be computed from (3.43a) and (3.43b) after substitution of the Fourier series. The first equation (3.43a) gives

$$\begin{aligned} & \sum_{n=-N}^N \frac{d}{dz} s_{i,x,n}(z) \exp[-i(k_{x,n}x + k_y y)] \\ &= -k_0 \sum_{n=-N}^N u_{i,y,n}(z) \exp[-i(k_{x,n}x + k_y y)] \\ & \quad - \frac{1}{k_0} \left(k_y \sum_{n=-N}^N \sum_{m=-N}^N k_{x,n} \tilde{\epsilon}_{i,nm} u_{i,x,m} \exp[-i(k_{x,n}x + k_y y)] \right. \\ & \quad \left. + \sum_{n=-N}^N \sum_{m=-N}^N k_{x,n} \tilde{\epsilon}_{i,nm} k_{x,m} u_{i,y,m} \exp[-i(k_{x,n}x + k_y y)] \right), \end{aligned} \quad (3.45)$$

where the inverse rule (3.8) has been used, since the right-hand side of (3.43a) contains several complementary concurrent jump discontinuities. Since (3.45) holds for $-\Lambda/2 \leq x \leq \Lambda/2$ and $y \in \mathbb{R}$, it holds that

$$\begin{aligned} \frac{d}{dz} s_{i,x,n}(z) &= -k_0 u_{i,y,n}(z) - \frac{k_y}{k_0} \sum_{m=-N}^N k_{x,n} \tilde{\epsilon}_{i,nm} u_{i,x,m} \\ & \quad + \frac{1}{k_0} \sum_{m=-N}^N k_{x,n} \tilde{\epsilon}_{i,nm} k_{x,m} u_{i,y,m}, \end{aligned} \quad (3.46)$$

for $-N \leq n \leq N$. When this relation is written in matrix form and rearranged to obtain a matrix relation for the magnetic field component vector $\mathbf{u}_{i,y}$, we have

$$\mathbf{u}_{i,y}(z) = \mathbf{K}_y \mathbf{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{u}_{i,x}(z) + \frac{1}{k_0} \mathbf{B}_i^{-1} \frac{d}{dz} \mathbf{s}_{i,x}(z). \quad (3.47)$$

In a similar way we can derive a matrix relation for the electric field components $s_{i,y}$. Substitute the expansions in Fourier series into (3.43b) to obtain

$$\begin{aligned} & \sum_{n=-N}^N \frac{d}{dz} u_{i,x,n}(z) \exp[-i(k_{x,n}x + k_y y)] \\ &= -k_0 \sum_{n=-N}^N \sum_{m=-N}^N \epsilon_{i,n-m} s_{i,y,m}(z) \exp[-i(k_{x,n}x + k_y y)] \\ & \quad + \frac{1}{k_0} \sum_{n=-N}^N (k_{x,n}^2 s_{i,y,n} - k_{x,n} s_{i,x,n}) \exp[-i(k_{x,n}x + k_y y)]. \end{aligned} \quad (3.48)$$

Note that (3.43b) contains no concurrent jump discontinuities. Therefore, Laurent's rule (3.7) has been applied. Since (3.48) holds for $-\Lambda/2 \leq x \leq \Lambda/2$ and $y \in \mathbb{R}$, we have

$$\frac{d}{dz} u_{i,xn}(z) = -k_0 \sum_{m=-N}^N \epsilon_{i,n-m} s_{i,ym}(z) + \frac{1}{k_0} (k_{xn}^2 s_{i,yn} - k_y k_{xn} s_{i,xm}), \quad (3.49)$$

for $-N \leq n \leq N$. When this relation is written in matrix form and rearranged to obtain a matrix relation for the electric field component vector $\mathbf{s}_{i,y}$, we have

$$\mathbf{s}_{i,y}(z) = \frac{1}{k_0} \mathbf{A}_i^{-1} \frac{d}{dz} \mathbf{u}_{i,x}(z) + \mathbf{K}_y \mathbf{A}_i^{-1} \mathbf{K}_x \mathbf{s}_{i,x}(z). \quad (3.50)$$

The substitution of (3.47) and (3.50) into the boundary conditions (3.44) leaves us with only $4N + 4$ unknown vectors and $4N + 4$ matrix relations.

3.5 Final system

The field components that are found in Section 3.3 and the boundary conditions derived in the previous section are used to obtain a system of equations to find the field amplitudes. For planar diffraction, we use the solution of \mathbf{s}_i , given by (3.15), and the solution of \mathbf{u}_i , given by (3.23), in the boundary conditions (3.40) and (3.42), respectively. Let us introduce the matrix \mathbf{V}_i for every layer $i = 2, \dots, K + 1$ as

$$\mathbf{V}_i := \begin{cases} \mathbf{W}_i \mathbf{Q}_i & \text{for TE polarization,} \\ \mathbf{P}_i \mathbf{W}_i \mathbf{Q}_i & \text{for TM polarization,} \end{cases} \quad (3.51)$$

and define matrix \mathbf{X}_i as

$$\mathbf{X}_i := \exp[-k_0 \mathbf{Q}_i d_i], \quad (3.52)$$

with d_i the thickness of layer i . The boundary conditions (3.40) and (3.42) can now be written in a unified manner as

for $z = 0$

$$\mathfrak{d} + \begin{bmatrix} \mathbf{I} \\ \mathfrak{M} \end{bmatrix} \mathbf{r} = \begin{bmatrix} \mathbf{W}_2 & \mathbf{W}_2 \mathbf{X}_2 \\ \mathbf{V}_2 & -\mathbf{V}_2 \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_2^+ \\ \mathbf{c}_2^- \end{bmatrix}, \quad (3.53a)$$

for $i = 3, \dots, K + 1$ and $z = D_{i-1}$

$$\begin{bmatrix} \mathbf{W}_{i-1} \mathbf{X}_{i-1} & \mathbf{W}_{i-1} \\ \mathbf{V}_{i-1} \mathbf{X}_{i-1} & -\mathbf{V}_{i-1} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{i-1}^+ \\ \mathbf{c}_{i-1}^- \end{bmatrix} = \begin{bmatrix} \mathbf{W}_i & \mathbf{W}_i \mathbf{X}_i \\ \mathbf{V}_i & -\mathbf{V}_i \mathbf{X}_i \end{bmatrix} \begin{bmatrix} \mathbf{c}_i^+ \\ \mathbf{c}_i^- \end{bmatrix}, \quad (3.53b)$$

for $z = D$

$$\begin{bmatrix} \mathbf{W}_{K+1} \mathbf{X}_{K+1} & \mathbf{W}_{K+1} \\ \mathbf{V}_{K+1} \mathbf{X}_{K+1} & -\mathbf{V}_{K+1} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{K+1}^+ \\ \mathbf{c}_{K+1}^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathfrak{M} \end{bmatrix} \mathbf{t}, \quad (3.53c)$$

where the vector \mathbf{d} and matrices \mathfrak{M} and \mathfrak{N} depend on the kind of diffraction. For planar diffraction with TE polarization they are given by

$$\mathbf{d} := \begin{bmatrix} \mathbf{d}_0 \\ i\mathbf{d}_0 n_l \cos \theta \end{bmatrix}, \quad (3.54a)$$

$$\mathfrak{M} := -\mathbf{Y}_l, \quad (3.54b)$$

$$\mathfrak{N} := \mathbf{Y}_l, \quad (3.54c)$$

while for planar diffraction with TM polarization they are defined as

$$\mathbf{d} := \begin{bmatrix} \mathbf{d}_0 \\ i\mathbf{d}_0 \frac{\cos \theta}{n_l} \end{bmatrix}, \quad (3.55a)$$

$$\mathfrak{M} := -\mathbf{Z}_l, \quad (3.55b)$$

$$\mathfrak{N} := \mathbf{Z}_l. \quad (3.55c)$$

For conical diffraction, we have the solutions of $\mathbf{s}_{i,x}$, given by (3.35), and $\mathbf{u}_{i,x}$, given by (3.36). In Section 3.4.2 we derived expressions of $\mathbf{s}_{i,y}$ and $\mathbf{u}_{i,y}$ in terms of $\mathbf{s}_{i,x}$ and $\mathbf{u}_{i,x}$. Therefore, we can also express $\mathbf{s}_{i,y}$ and $\mathbf{u}_{i,y}$ in terms of the field constants by using (3.47) and (3.50), viz.

$$\begin{aligned} \mathbf{u}_{i,y}(z) &= \mathbf{K}_y \mathbf{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{W}_{1,i} \\ &\quad \cdot (\exp[-k_0 \mathbf{Q}_{1,i}(z-D_{i-1})] \mathbf{c}_{1,i}^+ + \exp[k_0 \mathbf{Q}_{1,i}(z-D_i)] \mathbf{c}_{1,i}^-) \\ &\quad - \mathbf{B}_i^{-1} \mathbf{W}_{2,i} \mathbf{Q}_{2,i} \\ &\quad \cdot (\exp[-k_0 \mathbf{Q}_{2,i}(z-D_{i-1})] \mathbf{c}_{2,i}^+ - \exp[k_0 \mathbf{Q}_{2,i}(z-D_i)] \mathbf{c}_{2,i}^-), \end{aligned} \quad (3.56)$$

$$\begin{aligned} \mathbf{s}_{i,y}(z) &= -\mathbf{A}_i^{-1} \mathbf{W}_{1,i} \mathbf{Q}_{1,i} \\ &\quad \cdot (\exp[-k_0 \mathbf{Q}_{1,i}(z-D_{i-1})] \mathbf{c}_{1,i}^+ - \exp[k_0 \mathbf{Q}_{1,i}(z-D_i)] \mathbf{c}_{1,i}^-) \\ &\quad + \mathbf{K}_y \mathbf{A}_i^{-1} \mathbf{K}_x \mathbf{W}_{2,i} \mathbf{Q}_{2,i} \\ &\quad \cdot (\exp[-k_0 \mathbf{Q}_{2,i}(z-D_{i-1})] \mathbf{c}_{2,i}^+ + \exp[k_0 \mathbf{Q}_{2,i}(z-D_i)] \mathbf{c}_{2,i}^-). \end{aligned} \quad (3.57)$$

Let us introduce some matrices to keep the notation short after substitution of the latter relations in the boundary conditions (3.44). First introduce matrices $\mathbf{X}_{1,i}$ and $\mathbf{X}_{2,i}$ as

$$\mathbf{X}_{1,i} := \exp[-k_0 \mathbf{Q}_{1,i} d_i], \quad (3.58a)$$

$$\mathbf{X}_{2,i} := \exp[-k_0 \mathbf{Q}_{2,i} d_i], \quad (3.58b)$$

and combine them in one overall matrix \mathbf{X}_i as

$$\mathbf{X}_i := \begin{bmatrix} \mathbf{X}_{1,i} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{2,i} \end{bmatrix}. \quad (3.59)$$

Secondly, define matrices \mathbf{W}_i and \mathbf{V}_i as

$$\mathbf{W}_i := \begin{bmatrix} \mathbf{0} & \mathbf{W}_{2,i} \\ \mathbf{V}_{11,i} & \mathbf{V}_{12,i} \end{bmatrix}, \quad (3.60)$$

$$\mathbf{V}_i := - \begin{bmatrix} \mathbf{W}_{1,i} & \mathbf{0} \\ \mathbf{V}_{21,i} & \mathbf{V}_{22,i} \end{bmatrix}. \quad (3.61)$$

with

$$\mathbf{V}_{11,i} := \mathcal{A}_i^{-1} \mathbf{W}_{1,i} \mathbf{Q}_{1,i}, \quad (3.62a)$$

$$\mathbf{V}_{12,i} := \mathbf{K}_y \mathcal{A}_i^{-1} \mathbf{K}_x \mathbf{W}_{2,i}, \quad (3.62b)$$

$$\mathbf{V}_{21,i} := \mathbf{K}_y \mathcal{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{W}_{1,i}, \quad (3.62c)$$

$$\mathbf{V}_{22,i} := \mathcal{B}_i^{-1} \mathbf{W}_{2,i} \mathbf{Q}_{2,i}, \quad (3.62d)$$

where the definitions of $\mathbf{V}_{11,i}$, $\mathbf{V}_{12,i}$, $\mathbf{V}_{21,i}$ and $\mathbf{V}_{22,i}$ are direct consequences of relations (3.56) and (3.57). Finally, let us define vector \mathfrak{d} and matrices \mathfrak{M} and \mathfrak{N} as

$$\mathfrak{d} := \begin{bmatrix} \mathbf{d}_0 (\cos \psi \cos \theta \cos \phi - \sin \psi \sin \phi) \\ \mathbf{d}_0 (\cos \psi \cos \theta \sin \phi + \sin \psi \cos \phi) \\ i \mathbf{d}_0 n_l (-\cos \psi \sin \phi - \sin \psi \cos \theta \cos \phi) \\ i \mathbf{d}_0 n_l (\cos \psi \cos \phi - \sin \psi \cos \theta \sin \phi) \end{bmatrix}, \quad (3.63a)$$

$$\mathfrak{M} := \begin{bmatrix} \mathbf{K}_y \mathbf{Y}_l^{-1} \mathbf{K}_x & (\mathbf{Y}_l + \mathbf{K}_y \mathbf{Y}_l^{-1} \mathbf{K}_y) \\ -(\mathbf{Y}_l + \mathbf{K}_x \mathbf{Y}_l^{-1} \mathbf{K}_x) & -\mathbf{K}_x \mathbf{Y}_l^{-1} \mathbf{K}_y \end{bmatrix}, \quad (3.63b)$$

$$\mathfrak{N} := \begin{bmatrix} -\mathbf{K}_y \mathbf{Y}_{ll}^{-1} \mathbf{K}_x & -(\mathbf{Y}_{ll} + \mathbf{K}_y \mathbf{Y}_{ll}^{-1} \mathbf{K}_y) \\ (\mathbf{Y}_{ll} + \mathbf{K}_x \mathbf{Y}_{ll}^{-1} \mathbf{K}_x) & \mathbf{K}_x \mathbf{Y}_{ll}^{-1} \mathbf{K}_y \end{bmatrix}. \quad (3.63c)$$

The matrices \mathfrak{M} and \mathfrak{N} are the result of eliminating \mathbf{r}_z and \mathbf{t}_z by means of (2.39) and (2.40) from the boundary conditions for conical diffraction. When the field coefficients and the field amplitudes are collected in

$$\mathbf{c}_i^+ := \begin{bmatrix} -\mathbf{c}_{1,i}^+ \\ \mathbf{c}_{2,i}^+ \end{bmatrix}, \quad \mathbf{c}_i^- := \begin{bmatrix} \mathbf{c}_{1,i}^- \\ \mathbf{c}_{2,i}^- \end{bmatrix}, \quad \mathbf{r} := \begin{bmatrix} \mathbf{r}_x \\ \mathbf{r}_y \end{bmatrix}, \quad \mathbf{t} := \begin{bmatrix} \mathbf{t}_x \\ \mathbf{t}_y \end{bmatrix}, \quad (3.64)$$

the boundary conditions (3.44) give the $K + 1$ systems given by (3.53) such that we unified the notation for planar and conical diffraction and we can discuss the solution procedure for both diffraction cases simultaneously.

From the $K + 1$ systems given by (3.53) we can eliminate the vectors \mathbf{c}_i^+ and \mathbf{c}_i^- for $i = 2, \dots, K + 1$ such that we have sufficiently many equations for the two unknown vectors consisting of the reflected and transmitted field amplitudes, \mathbf{r} and \mathbf{t} , respectively.

$$\mathfrak{d} + \begin{bmatrix} \mathbf{I} \\ \mathfrak{M} \end{bmatrix} \mathbf{r} = \prod_{i=2}^{K+1} \left(\begin{bmatrix} \mathbf{W}_i & \mathbf{W}_i \mathbf{X}_i \\ \mathbf{V}_i & -\mathbf{V}_i \mathbf{X}_i \end{bmatrix} \begin{bmatrix} \mathbf{W}_i \mathbf{X}_i & \mathbf{W}_i \\ \mathbf{V}_i \mathbf{X}_i & -\mathbf{V}_i \end{bmatrix}^{-1} \right) \begin{bmatrix} \mathbf{I} \\ \mathfrak{N} \end{bmatrix} \mathbf{t}. \quad (3.65)$$

Note that, although the notation for planar and conical diffraction are similar, the matrices are twice as large for conical diffraction as for planar diffraction. Unfortunately, obtaining a direct solution of \mathbf{r} and \mathbf{t} is unstable for thick layers. The reason for the instability can be found by studying the matrix \mathbf{X}_i , which is the only matrix where the layer thickness appears. Although the inverse matrix of \mathbf{X}_i can be computed analytically, round-off errors cause numerical instabilities. Therefore, taking the inverse of this matrix should be avoided and this can be achieved by the enhanced transmittance matrix approach which is discussed in the next section.

3.6 Enhanced transmittance matrix approach

To avoid the usage of \mathbf{X}_i^{-1} , the *enhanced transmittance matrix approach* has been developed which uses a well chosen substitution of the transmitted field amplitudes \mathbf{t} , namely

$$\mathbf{t}_{i+1} := \mathbf{A}_i^{-1} \mathbf{X}_i \mathbf{t}_i \text{ for } i = 2, \dots, K + 2, \quad (3.66)$$

where $\mathbf{t}_{K+2} := \mathbf{t}$. After the substitution of (3.66) in (3.65), we can define matrices \mathbf{A}_{K+1} , \mathbf{B}_{K+1} , \mathbf{F}_{K+1} and \mathbf{G}_{K+1} , viz.

$$\begin{aligned} \mathfrak{d} + \begin{bmatrix} \mathbf{I} \\ \mathfrak{M} \end{bmatrix} \mathbf{r} = & \prod_{i=2}^K \left(\begin{bmatrix} \mathbf{W}_i & \mathbf{W}_i \mathbf{X}_i \\ \mathbf{V}_i & -\mathbf{V}_i \mathbf{X}_i \end{bmatrix} \begin{bmatrix} \mathbf{X}_i^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{W}_i & \mathbf{W}_i \\ \mathbf{V}_i & -\mathbf{V}_i \end{bmatrix}^{-1} \right) \\ & \cdot \underbrace{\left(\begin{bmatrix} \mathbf{W}_{K+1} & \mathbf{W}_{K+1} \mathbf{X}_{K+1} \\ \mathbf{V}_{K+1} & -\mathbf{V}_{K+1} \mathbf{X}_{K+1} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{K+1}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{W}_{K+1} & \mathbf{W}_{K+1} \\ \mathbf{V}_{K+1} & -\mathbf{V}_{K+1} \end{bmatrix}^{-1} \right) \begin{bmatrix} \mathbf{F}_{K+2} \\ \mathbf{G}_{K+2} \end{bmatrix} \mathbf{A}_{K+1}^{-1} \mathbf{X}_{K+1} \mathbf{t}_{K+1}}_{=:\begin{bmatrix} \mathbf{A}_{K+1} \\ \mathbf{B}_{K+1} \end{bmatrix}} \\ & \underbrace{\hspace{15em}}_{=:\begin{bmatrix} \mathbf{F}_{K+1} \\ \mathbf{G}_{K+1} \end{bmatrix}} \end{aligned} \quad (3.67)$$

where

$$\mathbf{F}_{K+2} := \mathbf{I} \text{ and } \mathbf{G}_{K+2} := \mathfrak{M}. \quad (3.68)$$

In the same manner matrices \mathbf{A}_i , \mathbf{B}_i , \mathbf{F}_i and \mathbf{G}_i can be computed for all layers $i = 2, \dots, K + 1$.

$$\mathbf{A}_i = \frac{1}{2} (\mathbf{W}_i^{-1} \mathbf{F}_{i+1} + \mathbf{V}_i^{-1} \mathbf{G}_{i+1}), \quad (3.69a)$$

$$\mathbf{B}_i = \frac{1}{2} (\mathbf{W}_i^{-1} \mathbf{F}_{i+1} - \mathbf{V}_i^{-1} \mathbf{G}_{i+1}), \quad (3.69b)$$

$$\mathbf{F}_i = \mathbf{W}_i (\mathbf{X}_i^{-1} \mathbf{A}_i + \mathbf{X}_i \mathbf{B}_i) \mathbf{A}_i^{-1} \mathbf{X}_i = \mathbf{W}_i (\mathbf{I} + \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i), \quad (3.69c)$$

$$\mathbf{G}_i = \mathbf{V}_i (\mathbf{X}_i^{-1} \mathbf{A}_i - \mathbf{X}_i \mathbf{B}_i) \mathbf{A}_i^{-1} \mathbf{X}_i = \mathbf{V}_i (\mathbf{I} - \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i). \quad (3.69d)$$

As can be seen, the computations of these four matrices do not involve \mathbf{X}_i^{-1} , but reduce (3.65) at each substitution of (3.66). In other words,

$$\mathfrak{d} + \begin{bmatrix} \mathbf{I} \\ \mathfrak{M} \end{bmatrix} \mathbf{r} = \prod_{i=2}^k \left(\begin{bmatrix} \mathbf{W}_i & \mathbf{W}_i \mathbf{X}_i \\ \mathbf{V}_i & -\mathbf{V}_i \mathbf{X}_i \end{bmatrix} \begin{bmatrix} \mathbf{W}_i \mathbf{X}_i & \mathbf{W}_i \\ \mathbf{V}_i \mathbf{X}_i & -\mathbf{V}_i \end{bmatrix}^{-1} \right) \begin{bmatrix} \mathbf{F}_{k+1} \\ \mathbf{G}_{k+1} \end{bmatrix} \mathbf{t}_{k+1}, \quad (3.70)$$

and after the substitution of (3.66) for all $i = 2, \dots, K + 1$ in (3.65) the final system we have to solve is

$$\mathfrak{d} + \begin{bmatrix} \mathbf{I} \\ \mathfrak{M} \end{bmatrix} \mathbf{r} = \begin{bmatrix} \mathbf{F}_2 \\ \mathbf{G}_2 \end{bmatrix} \mathbf{t}_2. \quad (3.71)$$

Finally, we can find the reflected field amplitudes and the auxiliary \mathbf{t}_2 when we solve the following system

$$\begin{bmatrix} -\mathbf{I} & \mathbf{F}_2 \\ -\mathfrak{M} & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{t}_2 \end{bmatrix} = \mathfrak{D}. \quad (3.72)$$

If the transmitted field amplitudes are important to know, we can find them by using (3.66) recursively, viz.

$$\mathbf{t} = \mathbf{A}_{K+1}^{-1} \mathbf{X}_{K+1} \dots \mathbf{A}_i^{-1} \mathbf{X}_i \dots \mathbf{A}_2^{-1} \mathbf{X}_2 \mathbf{t}_2. \quad (3.73)$$

By using the enhanced transmitted matrix approach we can compute the reflected and transmitted field amplitudes in a stable way.

3.7 Influence of the number of harmonics and layers

Until now we have not mentioned what values the numbers K (number of layers) and N (number of harmonics) should have. Intuitively, the more harmonics one takes into account in the Fourier series of both the electromagnetic field and the permittivity, the more accurate the solution. Similarly, the more layers to approximate the grating structure will give more accurate solutions. On the other hand, an increasing number of harmonics and layers will also rapidly increase the computational effort for obtaining such solution. Since eigenvalues have to be solved for every layer, the complexity of the RCWA algorithm is $\mathcal{O}(KN^3)$. Since we want good results fast, it is important to find a criterion for the number of harmonics and layers to ensure that we have an answer that still matches with the physics.

We will discuss the convergence with respect to the number of layers and harmonics by considering a symmetric trapezoid on two homogeneous layers (testcase 7 of Table 5.1). When we consider the reflected field amplitudes obtained by the RCWA algorithm with a high number of harmonics ($N = 100$) and a high number of layers ($K = 100$) as the true reflected field amplitudes, we can compare this "true" value with our simulations for a lower amount of harmonics and/or layers. In Figure 3.2 we have displayed the logarithm of the absolute error between the computed diffraction efficiencies and our reference values as a function of the number of layers and the number of harmonics. The error also depends on which order you consider. As long the ratio between K and N is kept constant, the error decreases when more layers and harmonics are taken into account. It is a waste of computation time to take 100 layers for 2 harmonics and vice versa.

The difference in convergence behaviour between TE and TM polarization is generally presumed to be caused by the convergence properties of the Fourier series of the permittivity and the field [32]. In general we can say that the Fourier series converge as the series $1/n$. From numerical simulations, we can show that the eigenvalues of the auxiliary matrices \mathbf{A}_i converge like $1/n^3$, while the eigenvalues of $\mathbf{P}_i^{-1} \mathbf{B}_i$ converge like $1/n^2$. This is seen as the major contribution to the difference in convergence of the diffraction efficiencies for planar

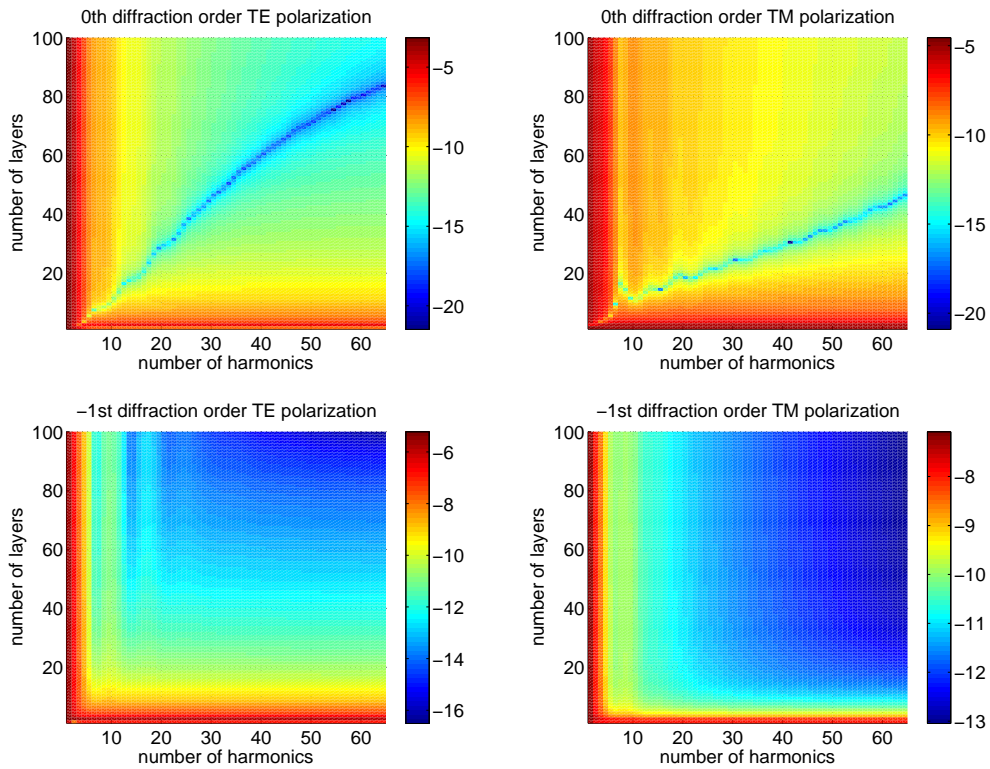


Figure 3.2: Logarithmic representation of the absolute error made in the computation of the zeroth and first order diffraction efficiencies of the reflected field compared to the diffraction efficiencies for 100 harmonics and 100 layers. This error is given as a function of the number of layers and the number of harmonics.

TE and TM polarization.

For the zeroth diffraction order we observe that for a specific combination of K and N we have an even better accuracy than when we would take more layers or harmonics into account. This phenomenon is still a mystery, but it offers potential to reduce the number of layers and harmonics drastically when only the zeroth order is of importance.

Chapter 4

The C method

The C method is an alternative method to compute the diffracted field coming from a grating structure. The method is named after its inventor Jean Chandezon, who came up with this method in the late '80's [13, 30]. The method is also known as the *coordinate transformation method*. At that time, the RCWA algorithm was already widely used and tested. However, especially for smooth grating profiles such as sinusoidal gratings the RCWA algorithm has the drawback that it has to use many layers to approximate the grating structure sufficiently accurately. The essential resemblance between the C method and RCWA is that both methods try to eliminate the dependence of the refractive index on directions other than the direction the periodicity is in. As RCWA cuts its region into slices, the C method uses a coordinate transformation to obtain this. Both methods have their advantages and disadvantages and in the end, the C method is not meant as a substitute for RCWA, but an alternative for gratings with a smooth profile [2].

Since the C method has similar features as the RCWA method, it profitted from the progress of the RCWA algorithm. For example, the mathematics of taking a product of truncated series as discussed in Section 3.2 can also be implemented in the C method [29].

Although the C method is meant for gratings with a smooth profile, the method as it is known today can only be applied to profiles that can be described by a function of the x -direction. In practice we are also dealing with overhanging grating profiles, e.g. when a resist line is exposed when it was out of focuss. When the C method is not applicable to all grating profiles common in practice, the method is not useable for reconstruction purposes. In this thesis we will only study planar diffraction with TE and TM polarization for a two media problem only. This means that we have one interface between two homogeneous layers with refractive indices n_i and n_{II} , respectively, as is illustrated in Figure 4.1.

The extension to more media is treated in among others [31] and conical diffraction has been discussed in [27]. Although generalizations of the coordinate transformation have

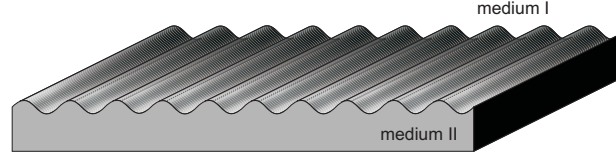


Figure 4.1: Illustration of a two medium grating structure.

been addressed in the literature [17, 44], the C method can still not solve diffraction by overhanging grating profiles. In Section 4.4 this extension to overhanging grating profiles will be discussed briefly.

4.1 The key feature of the C method

The C method does not introduce artificial interfaces; the only interface is the grating profile itself. The two media are *homogeneous*, which implies that each medium has a constant refractive index. Therefore, the electric and magnetic field satisfy the Helmholtz equations (2.12), the outgoing wave condition and the pseudo-periodic boundary condition as discussed in Section 2.2.3.

The key feature of the C method is the coordinate transformation. To eliminate the z -dependence of the refractive index, introduce the coordinate transformation where the interface is described as $z = a(x)$, viz.

$$\begin{cases} u := x, \\ v := y, \\ w := z - a(x). \end{cases} \quad (4.1)$$

To use this coordinate transformation in the Helmholtz equation, we have to know its effect on unit vectors, partial derivatives and field magnitudes. These identities are summarized below. For notational convenience the derivative of $a(u)$ with respect to u will be denoted as $a'(u)$.

- Unit vectors

$$\mathbf{u}_x = \frac{\partial u}{\partial x} \mathbf{u}_u + \frac{\partial v}{\partial x} \mathbf{u}_v + \frac{\partial w}{\partial x} \mathbf{u}_w = \mathbf{u}_u - a'(u) \mathbf{u}_w, \quad (4.2a)$$

$$\mathbf{u}_y = \frac{\partial u}{\partial y} \mathbf{u}_u + \frac{\partial v}{\partial y} \mathbf{u}_v + \frac{\partial w}{\partial y} \mathbf{u}_w = \mathbf{u}_v, \quad (4.2b)$$

$$\mathbf{u}_z = \frac{\partial u}{\partial z} \mathbf{u}_u + \frac{\partial v}{\partial z} \mathbf{u}_v + \frac{\partial w}{\partial z} \mathbf{u}_w = \mathbf{u}_w. \quad (4.2c)$$

- Partial derivatives

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial}{\partial u} + \frac{\partial v}{\partial x} \frac{\partial}{\partial v} + \frac{\partial w}{\partial x} \frac{\partial}{\partial w} = \frac{\partial}{\partial u} - a'(u) \frac{\partial}{\partial w}, \quad (4.3a)$$

$$\frac{\partial}{\partial z} = \frac{\partial u}{\partial z} \frac{\partial}{\partial u} + \frac{\partial v}{\partial z} \frac{\partial}{\partial v} + \frac{\partial w}{\partial z} \frac{\partial}{\partial w} = \frac{\partial}{\partial w}. \quad (4.3b)$$

- Magnitudes of the fields

$$\begin{aligned} e_x = \mathbf{e} \cdot \mathbf{u}_x &= \{e_u \mathbf{u}_u + e_v \mathbf{u}_v + e_w \mathbf{u}_w\} \cdot (\mathbf{u}_u - a'(u) \mathbf{u}_w) \\ &= e_u - a'(u) e_w, \end{aligned} \quad (4.4a)$$

$$e_y = \mathbf{e} \cdot \mathbf{u}_y = \{e_u \mathbf{u}_u + e_v \mathbf{u}_v + e_w \mathbf{u}_w\} \cdot \mathbf{u}_v = e_v, \quad (4.4b)$$

$$e_z = \mathbf{e} \cdot \mathbf{u}_z = \{e_u \mathbf{u}_u + e_v \mathbf{u}_v + e_w \mathbf{u}_w\} \cdot \mathbf{u}_w = e_w. \quad (4.4c)$$

Just like the electric field, the magnetic field components can be described in this way:

$$h_x = h_u - a'(u) h_w, \quad h_y = h_v, \quad h_z = h_w. \quad (4.5)$$

Note that the components in the u - and v -directions are tangential to the interface, but the w -component is not normal to the interface. The periodicity is preserved in the u -direction.

Next, the way the C method uses this coordinate transformation to solve the diffraction problem for a two media grating will be discussed.

4.2 The details of the C method

4.2.1 Coordinate transformation of the Helmholtz equation

Because the C method considers the media separately and inside a medium the refractive index is constant, Maxwell's equations reduce to the Helmholtz equations (2.13a) and (2.13b) for planar diffraction. Since both variants are identical, introduce f_y to represent the y -component of either the electric field or the magnetic field, cf. (2.19), such that

$$\frac{\partial^2}{\partial x^2} f_y(x, z) + \frac{\partial^2}{\partial z^2} f_y(x, z) = -k_0^2 n_m^2(x, z) f_y(x, z), \quad (4.6)$$

for $m = \text{I, II}$ and where we use the refractive index instead of the relative permittivity, which are related by

$$n_m^2(x, z) = \tilde{\epsilon}_m^r(x, z). \quad (4.7)$$

Substituting the coordinate transformation as given by (4.1) in the Helmholtz equation and using the identities for the derivatives (4.3) and for the field amplitudes (4.4) result in

$$\left\{ \left(\frac{\partial}{\partial u} - a'(u) \frac{\partial}{\partial w} \right)^2 + \frac{\partial^2}{\partial w^2} + k_0^2 n_m^2 \right\} f_v(u, w) = 0. \quad (4.8)$$

4.2.2 The shape of the solution

Since the electric and magnetic field satisfy the Floquet condition (2.20) and $u = x$, the field in the new coordinate system has to satisfy

$$f_v(u + \Lambda, w) = f_v(u, w) \exp[-ik_0 n_l \Lambda \sin \theta], \quad (4.9)$$

which implies that the field f_v can be expanded in a Fourier series with respect to the u -coordinate as long as the phase correction is taken into account. Thus

$$f_v(u, w) = \sum_{n=-N}^N \tilde{f}_{vn}(w) \exp[-ik_{un}u], \quad (4.10)$$

where \tilde{f}_{vn} are the Fourier coefficients of f_v and k_{un} are defined identically as k_{xn} (2.25) for $-N \leq n \leq N$, thus

$$k_{un} := k_0 n_l \Lambda \sin \theta - n \frac{2\pi}{\Lambda}. \quad (4.11)$$

Since (4.8) contains $a'(u)$ we also need the expansion in Fourier series of the derivative of the profile function, which is given by

$$a'(u) = \sum_{n=-N}^N a_n \exp[in \frac{2\pi}{\Lambda} u], \quad a_n = \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} a'(u) \exp[-in \frac{2\pi}{\Lambda} u] du. \quad (4.12)$$

Substitution of the Fourier series (4.10) and (4.12) in the transformed Helmholtz equation (4.8) is the first step of transforming the partial differential equation into an algebraic eigenvalue system. When we truncate the Fourier series, we have to make sure we can apply Laurent's rule (3.7) or the inverse rule (3.8) to the products of these Fourier series. Inside a medium the refractive index does not change, but since the profile function a occurs in the coordinate transformation (4.1), discontinuities in this profile function also have an effect inside a medium. Physically all components of the electric and magnetic field are continuous, but the introduction of the coordinate transformation can cause artificial discontinuities. When the profile function a' has discontinuities like for a trapezoidal grating, the u -component of the field has a discontinuity because of the coordinate transformation. Similarly, the derivative of any field component with respect to u is discontinuous. However, the Helmholtz equation (4.8) contains the term $\partial/\partial u - a'(u)(\partial/\partial w)$, which is continuous

because it is the same as the derivative with respect to x . To conclude, computing products of the truncated Fourier series can be performed by using Laurent's rule (3.7).

The substitution of the Fourier series in the Helmholtz equation (4.8) gives

$$\begin{aligned}
& - \sum_{n=-N}^N k_{un}^2 f_{vn}(w) \exp[-ik_{un}u] + i \sum_{n=-N}^N \sum_{m=-N}^N a_{m-n} k_{um} \frac{d}{dw} f_{vm}(w) \exp[-ik_{un}u] \\
& + i \sum_{n=-N}^N k_{un} \sum_{m=-N}^N a_{n-m} \frac{d}{dw} f_{vm}(w) \exp[-ik_{un}u] \\
& + \sum_{n=-N}^N \sum_{m=-N}^N a_{n-m} \sum_{r=-N}^N a_{m-r} \frac{d^2}{dw^2} f_{vr}(w) \exp[-ik_{un}u] \\
& + \sum_{n=-N}^N \frac{d^2}{dw^2} f_{vn}(w) \exp[-ik_{un}u] + k_0^2 n_m^2 \sum_{n=-N}^N f_{vn}(w) \exp[-ik_{un}u] = 0. \quad (4.13)
\end{aligned}$$

Projection on the exponential basis function yields

$$\begin{aligned}
& - k_{un}^2 f_{vn}(w) + i \sum_{m=-N}^N a_{m-n} k_{um} \frac{d}{dw} f_{vm}(w) + i \sum_{m=-N}^N k_{un} a_{n-m} \frac{d}{dw} f_{vm}(w) \\
& + \sum_{m=-N}^N \sum_{r=-N}^N a_{n-m} a_{m-r} \frac{d^2}{dw^2} f_{vr}(w) + \frac{d^2}{dw^2} f_{vn}(w) + k_0^2 n_m^2 f_{vn}(w) = 0. \quad (4.14)
\end{aligned}$$

Let us define $k_{m,wn}$ for $m = I, II$ and $-N \leq n \leq N$ identically to $k_{m,zn}$ as

$$k_{m,wn} := \sqrt{k_0^2 n_m^2 - k_{un}^2}. \quad (4.15)$$

We collect f_{vn} for $-N \leq n \leq N$ in the vector \mathbf{f}_v . Moreover, \mathbf{K}_u is defined as a diagonal matrix consisting of k_{un} , $\mathbf{K}_{m,w}$ as a diagonal matrix consisting of $k_{m,wn}$ and \mathbf{A} as a Toeplitz matrix consisting of the Fourier coefficients of the profile function derivative $a'(u)$. As a result (4.14) can be written as

$$-\mathbf{K}_{m,w}^2 \mathbf{f}_v(w) + (\mathbf{A} \mathbf{K}_u + \mathbf{K}_u \mathbf{A}) \frac{d}{dw} \mathbf{f}_v(w) + (\mathbf{I} + \mathbf{A}^2) \frac{d^2}{dw^2} \mathbf{f}_v(w) = \mathbf{0}. \quad (4.16)$$

Introduce a new vector \mathbf{q}_v defined as

$$\mathbf{q}_v(w) := \frac{d}{dw} \mathbf{f}_v(w), \quad (4.17)$$

such that we can write (4.16) as a system of first-order derivatives only, viz.

$$\mathfrak{H} \begin{bmatrix} \mathbf{f}_v(w) \\ \mathbf{q}_v(w) \end{bmatrix} = \frac{d}{dw} \begin{bmatrix} \mathbf{f}_v(w) \\ \mathbf{q}_v(w) \end{bmatrix}, \quad (4.18)$$

with

$$\mathfrak{H} := \begin{bmatrix} \mathbf{K}_{m,w}^{-2} (\mathbf{K}_u \mathbf{A} - \mathbf{A} \mathbf{K}_u) & \mathbf{K}_{m,w}^{-2} (\mathbf{I} + \mathbf{A}^2) \\ \mathbf{I} & \mathbf{0} \end{bmatrix}^{-1}. \quad (4.19)$$

This is a system of first-order differential equations of which the general solution can be found in Appendix A.2.1. The eigenvalues of \mathfrak{S} can be divided in two sets of size $2N+1$ each, such that the first set contains all eigenvalues which are real and positive or imaginary with a positive imaginary part. The other consists of all eigenvalues which are real and negative or imaginary with a negative imaginary part. To let the solution satisfy the outgoing wave condition, medium I only uses the former, while medium II only uses the latter. The real eigenvalues correspond to the propagating orders, while the imaginary eigenvalues belong to the evanescent orders. To describe the field components \mathbf{f}_v let us define matrix \mathbf{W}_I which contains the first $2N + 1$ entries of those eigenvectors that belong to the real negative eigenvalues or the imaginary eigenvalues with a negative imaginary part. The eigenvalues are collected in the $2N + 1$ diagonal matrix $\mathbf{\Lambda}_I$ and therefore

$$\mathbf{f}_{I,v}(w) = \mathbf{W}_I \mathbf{c}_I \exp[\mathbf{\Lambda}_I w], \quad (4.20)$$

where \mathbf{c}_I are the unknown field coefficients of medium I. Similarly we can define the matrix \mathbf{W}_{II} which contains the first $2N + 1$ entries of those eigenvectors that belong to the real positive eigenvalues or the imaginary eigenvalues with a positive imaginary part. The eigenvalues are collected in the $2N + 1$ diagonal entries of the matrix $\mathbf{\Lambda}_{II}$ and therefore

$$\mathbf{f}_{II,v}(w) = \mathbf{W}_{II} \mathbf{c}_{II} \exp[\mathbf{\Lambda}_{II} w]. \quad (4.21)$$

where \mathbf{c}_{II} contains the unknown field coefficients of medium II. The eigenvalues of matrix \mathfrak{S} have to be linked to their diffraction order. It can be shown numerically [12] and analytically [26], that the real eigenvalues of \mathfrak{S} converge to the values $k_{m,wn}$ as the dimensions of \mathfrak{S} grow. Since the values of $k_{m,wn}$ are already linked to a diffraction order, this property ensures that the propagating orders are also assigned. As a consequence the eigenvectors belonging to the propagating orders can be computed analytically. Still the eigenvalues and eigenvectors of the evanescent orders are needed for the boundary condition. However, since we are only interested in the amplitudes and phases of the propagating orders, it is irrelevant to which diffraction order these evanescent waves belong.

4.2.3 Boundary conditions

The previous section described the field of the upper and lower medium in terms of unknown field coefficients. In this section the boundary conditions are discussed to find these field coefficients. The boundary condition we have not used so far is the condition that the tangential components are continuous at the interface, viz.

$$\mathbf{n} \times (\mathbf{e}_I(u, 0) - \mathbf{e}_{II}(u, 0)) = 0, \quad \mathbf{n} \times (\mathbf{h}_I(u, 0) - \mathbf{h}_{II}(u, 0)) = 0 \quad (4.22)$$

with $\mathbf{n} = a'(x)\mathbf{u}_x + \mathbf{u}_z$. Note that $\mathbf{n} \neq \mathbf{u}_w$ and \mathbf{u}_u is tangential to the interface. If \mathbf{n} is substituted in equation (4.22) and transformed to the (u, v, w) -coordinates, the results is

$$f_{I,v}(u, 0) = f_{II,v}(u, 0), \quad (4.23a)$$

$$g_{I,u}(u, 0) = g_{II,u}(u, 0), \quad (4.23b)$$

where g_u represents the u -component of the magnetic field for TE polarization and the u -component of the electric field for TM polarization. To use the continuity of the fields at the interface, also the incident field has to be transformed, viz.

$$\begin{aligned} f_v^{\text{inc}}(u, w) &= \exp[-i(k_{u0}u + k_{l,w0}(w + a(u)))] \\ &= \exp[-ik_{l,w0}a(u)] \exp[-i(k_{u0}u + k_{l,w0}w)]. \end{aligned} \quad (4.24)$$

To obtain an algebraic system, expand the exponential term containing $a(u)$ in its Fourier series. For notational convenience introduce

$$\mathcal{H}_n(\xi) := \frac{1}{\Lambda} \int_0^\Lambda \exp[i\xi a(u) - in\frac{2\pi}{\Lambda}u] du, \quad (4.25)$$

to describe the Fourier coefficients. For the incident field this implies that

$$f_v^{\text{inc}}(u, w) = \sum_n \mathcal{H}_n(-k_{l,w0}) \exp[-i(k_{un}u + k_{l,w0}w)]. \quad (4.26)$$

For the two-media problem, the interface lies at $w = 0$ and therefore, the dependence of the w -coordinate can be ignored.

Let us use the fact that the real eigenvalues of \mathfrak{H} converge to the values k_{m,w_n} as the dimensions of \mathfrak{H} grow. This implies that for the propagating diffraction orders the eigenvalue and its eigenvector may be replaced. Thus for the propagating orders, the Rayleigh expansion (2.36a) should also be transformed. Let us denote the set of propagating orders by \mathcal{U}_m with $m = \text{I, II}$. Then, the propagating orders are transformed according to

$$\begin{aligned} &\sum_{n \in \mathcal{U}_l} r_n \exp[-i(k_{x_n}x - k_{l,z_n}z)] \\ &= \sum_{n \in \mathcal{U}_l} r_n \exp[-i(k_{un}u - k_{l,w_n}(w + a(u)))] \\ &= \sum_{n \in \mathcal{U}_l} r_n \exp[ik_{l,w_n}a(u)] \exp[-i(k_{un}u - k_{l,w_n}w)] \\ &= \sum_{n \in \mathcal{U}_l} \sum_{m=-N}^N r_n \mathcal{H}_m(k_{l,w_n}) \exp[-i(k_{u(n+m)}u - k_{l,w_n}w)] \\ &= \sum_{n \in \mathcal{U}_l} \sum_{m=-N}^N r_n \mathcal{H}_{m-n}(k_{l,w_n}) \exp[-i(k_{um}u - k_{l,w_n}w)]. \end{aligned} \quad (4.27)$$

When we define \mathcal{V}_m as the set of evanescent orders inside medium m , the total field of medium I at the interface where $w = 0$ is given by

$$\begin{aligned} f_{l,v}(u, 0) &= \sum_n \mathcal{H}_n(-k_{l,w0}) \exp[-ik_{un}u] \\ &\quad + \sum_{n \in \mathcal{U}_l} \sum_{m=-N}^N r_n \mathcal{H}_{m-n}(k_{l,w_n}) \exp[-ik_{um}u] \\ &\quad + \sum_{n \in \mathcal{V}_l} \sum_{m=-N}^N w_{l,mn} c_{l,n} \exp[-ik_{un}u]. \end{aligned} \quad (4.28)$$

In a similar way the field of medium II at the interface where $w = 0$ is given by

$$\begin{aligned} f_{II,v}(u, 0) &= \sum_{n \in \mathcal{U}_I} \sum_{m=-N}^N t_n \mathcal{H}_{m-n}(-k_{II,wn}) \exp[-ik_{um}u] \\ &+ \sum_{n \in \mathcal{V}_I} \sum_{m=-N}^N w_{II,mn} c_{II,n} \exp[-ik_{un}u]. \end{aligned} \quad (4.29)$$

Substituting the expansions (4.28) and (4.29) in this boundary condition with $w = 0$ gives

$$\begin{aligned} \mathcal{H}_m(-k_{I,w0}) + \sum_{n \in \mathcal{U}_I} \mathcal{H}_{m-n}(k_{I,wn}) r_n + \sum_{r \in \mathcal{V}_I} w_{I,mq} c_{I,r} \\ = \sum_{n \in \mathcal{U}_I} \mathcal{H}_{m-n}(-k_{II,wn}) t_n + \sum_{r \in \mathcal{V}_I} w_{II,mr} c_{II,r}. \end{aligned} \quad (4.30)$$

This equation can be written in matrix form. Introduce M_I and M_{II} as the number of propagating diffraction orders in media I and II, respectively. Define the following matrices

- $\mathbf{F}_I^{\text{prop}} \in \mathbb{C}^{2N+1 \times M_I}$ for the propagating orders in medium I;
- $\mathbf{F}_I^{\text{evan}} \in \mathbb{C}^{2N+1 \times 2N+1-M_I}$ for the evanescent orders in medium I;
- $\mathbf{F}_{II}^{\text{prop}} \in \mathbb{C}^{2N+1 \times M_{II}}$ for the propagating orders in medium II;
- $\mathbf{F}_{II}^{\text{evan}} \in \mathbb{C}^{2N+1 \times 2N+1-M_{II}}$ for the evanescent orders in medium II;
- $\mathbf{f}_I^{\text{inc}} \in \mathbb{C}^{2N+1 \times 1}$ for the incident field.

The entries of these matrices are given by

$$(\mathbf{F}_I^{\text{prop}})_{np} = \mathcal{H}_{n-p}(k_{I,wp}), \quad p \in \mathcal{U}_I, \quad (4.31a)$$

$$(\mathbf{F}_I^{\text{evan}})_{nq} = w_{I,nq}, \quad q \in \mathcal{V}_I, \quad (4.31b)$$

$$(\mathbf{F}_{II}^{\text{prop}})_{ns} = \mathcal{H}_{n-s}(-k_{II,ws}), \quad s \in \mathcal{U}_{II}, \quad (4.31c)$$

$$(\mathbf{F}_{II}^{\text{evan}})_{nt} = w_{II,nt}, \quad t \in \mathcal{V}_{II}, \quad (4.31d)$$

$$(\mathbf{f}_I^{\text{inc}})_{n1} = \mathcal{H}_n(-k_{I,w0}), \quad (4.31e)$$

with $-N \leq n \leq N$. The first boundary condition in matrix form then becomes

$$\begin{bmatrix} \mathbf{F}_I^{\text{prop}} & \mathbf{F}_I^{\text{evan}} & -\mathbf{F}_{II}^{\text{prop}} & -\mathbf{F}_{II}^{\text{evan}} \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{c}_I \\ \mathbf{t} \\ \mathbf{c}_{II} \end{bmatrix} = -\mathbf{f}_I^{\text{inc}}, \quad (4.32)$$

where \mathbf{r} and \mathbf{t} contain those field amplitudes that belong to the propagating orders. In other words, the vector \mathbf{r} consists of r_n where $n \in \mathcal{U}_I$ and the vector \mathbf{t} consists of t_n where $n \in \mathcal{U}_{II}$. Vectors \mathbf{c}_I and \mathbf{c}_{II} are the unknown field coefficients corresponding to the evanescent orders

inside medium I and II, respectively. This boundary condition is valid for both TE and TM polarization.

The second boundary condition is more troublesome. The only field available tangential to the interface is the u -component, but there is no u -component of the \mathbf{f} -field. However, by means of the Maxwell equations (2.8), the electric field can be written as a function of magnetic field components and vice versa. By using the identities for the unit vectors (4.2), the partial derivatives (4.3) and the field magnitudes (4.4) expressions are derived for e_u and h_u , viz.

$$e_u(u, w) = \frac{1}{i\omega\tilde{\epsilon}_m} \left[a'(u) \frac{\partial h_v(u, w)}{\partial u} - (1 + (a'(u))^2) \frac{\partial h_v(u, w)}{\partial w} \right], \quad (4.33a)$$

$$h_u(u, w) = -\frac{1}{i\omega\mu_0} \left[a'(u) \frac{\partial e_v(u, w)}{\partial u} - (1 + (a'(u))^2) \frac{\partial e_v(u, w)}{\partial w} \right]. \quad (4.33b)$$

Since both expressions have a similar form, introduce field g_u , which is defined as

$$g_u(u, w) := \begin{cases} \sqrt{\epsilon_0/\mu_0}\tilde{\epsilon}_m^r e_u(u, w), & \text{for TM polarization,} \\ -\sqrt{\mu_0/\epsilon_0} h_u(u, w), & \text{for TE polarization,} \end{cases} \quad (4.34)$$

which is similar as in (4.23b) with some correction terms for notational convenience. If the expansions for f_v are substituted in (4.33) and w is set to zero, this gives

$$\begin{aligned} & g_{I,un}(u, 0) \\ &= \frac{1}{k_0} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} + \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) k_{I,w0} \right\} \mathcal{H}_m(-k_{I,w0}) \\ &+ \frac{1}{k_0} \sum_{p \in \mathcal{U}_I} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} - \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) k_{I,wp} \right\} \mathcal{H}_{m-p}(k_{I,wp}) a_p \\ &+ \frac{1}{k_0} \sum_{q \in \mathcal{V}_I} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} + \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) \lambda_{I,q} \right\} w_{I,mq} c_{I,q}. \end{aligned} \quad (4.35)$$

Similarly for medium II:

$$\begin{aligned} & g_{II,un}(u, 0) \\ &= \frac{1}{k_0} \sum_{s \in \mathcal{U}_{II}} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} + \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) k_{II,ws} \right\} \mathcal{H}_{m-s}(-k_{II,ws}) \\ &+ \frac{1}{k_0} \sum_{t \in \mathcal{V}_{II}} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} - \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) \lambda_{II,t} \right\} w_{II,mt}. \end{aligned} \quad (4.36)$$

Since $w = 0$ is already implemented, the two equations (4.35) and (4.36) are identical. This equality can be given in matrix form. Therefore introduce the following matrices:

- $\mathbf{G}_I^{\text{prop}} \in \mathbb{C}^{2N+1 \times M}$ for the propagating orders in medium I;

- $\mathbf{G}_I^{\text{evan}} \in \mathbb{C}^{2N+1 \times 2N+1-M_I}$ for the evanescent orders in medium I;
- $\mathbf{G}_{II}^{\text{prop}} \in \mathbb{C}^{2N+1 \times M_{II}}$ for the propagating orders in medium II;
- $\mathbf{G}_{II}^{\text{evan}} \in \mathbb{C}^{2N+1 \times 2N+1-M_{II}}$ for the evanescent orders in medium II;
- $\mathbf{g}_I^{\text{inc}} \in \mathbb{C}^{2N+1 \times 1}$ for the incident field.

The entries of these matrices are given by

$$\begin{aligned} (\mathbf{G}_I^{\text{prop}})_{np} &= \frac{1}{k_0} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} \mathcal{H}_{m-p}(k_{I,wp}) \right. \\ &\quad \left. - \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) \mathcal{H}_{m-p}(k_{I,wp}) k_{I,wp} \right\}, p \in \mathcal{U}_I \end{aligned} \quad (4.37a)$$

$$\begin{aligned} (\mathbf{G}_I^{\text{evan}})_{nq} &= \frac{1}{k_0} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} w_{I,mq} \right. \\ &\quad \left. - \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) w_{I,mq} \lambda_{I,q} \right\}, q \in \mathcal{V}_I \end{aligned} \quad (4.37b)$$

$$\begin{aligned} (\mathbf{G}_{II}^{\text{prop}})_{ns} &= \frac{1}{k_0} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} \mathcal{H}_{m-s}(k_{II,ws}) \right. \\ &\quad \left. + \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) \mathcal{H}_{m-s}(k_{II,ws}) k_{II,ws} \right\}, s \in \mathcal{U}_{II} \end{aligned} \quad (4.37c)$$

$$\begin{aligned} (\mathbf{G}_{II}^{\text{evan}})_{nt} &= \frac{1}{k_0} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} w_{II,mt} \right. \\ &\quad \left. - \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) w_{II,mt} \lambda_{II,t} \right\}, t \in \mathcal{V}_{II}, \end{aligned} \quad (4.37d)$$

$$\begin{aligned} (\mathbf{g}_I^{\text{inc}})_{n1} &= \frac{1}{k_0} \sum_{m=-N}^N \left\{ a_{n-m} k_{um} \mathcal{H}_m(-k_{I,w0}) \right. \\ &\quad \left. + \left(\delta_{nm} + \sum_{r=-N}^N a_{n-r} a_{r-m} \right) \mathcal{H}_m(-k_{I,w0}) k_{I,w0} \right\}. \end{aligned} \quad (4.37e)$$

The matrix equation then becomes

$$\begin{bmatrix} \mathbf{G}_I^{\text{prop}} & \mathbf{G}_I^{\text{evan}} & -\mathbf{G}_{II}^{\text{prop}} & -\mathbf{G}_{II}^{\text{evan}} \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{c}_I \\ \mathbf{t} \\ \mathbf{c}_{II} \end{bmatrix} = -\mathbf{g}_I^{\text{inc}}. \quad (4.38)$$

Combining equations (4.32) and (4.38) gives

$$\begin{bmatrix} \mathbf{F}_I^{\text{prop}} & \mathbf{F}_I^{\text{evan}} & -\mathbf{F}_{II}^{\text{prop}} & -\mathbf{F}_{II}^{\text{evan}} \\ \mathbf{G}_I^{\text{prop}} & \mathbf{G}_I^{\text{evan}} & -\mathbf{G}_{II}^{\text{prop}} & -\mathbf{G}_{II}^{\text{evan}} \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{c}_I \\ \mathbf{t} \\ \mathbf{c}_{II} \end{bmatrix} = - \begin{bmatrix} \mathbf{f}_I^{\text{inc}} \\ \mathbf{g}_I^{\text{inc}} \end{bmatrix}. \quad (4.39)$$

Solving this last equation gives the reflected and transmitted field for a two-media problem in case of planar diffraction.

4.3 RCWA vs. C method

The main difference between RCWA and the C method is the way they remove the z -dependence as summarized in Figure 4.2. As a consequence of this difference RCWA solves one eigenvalue system per *layer* of size $2N + 1$ if planar diffraction is considered, while the C method solves one per *medium* of size $4N + 2$. This is because RCWA can decouple the relations for the field components, while the C method cannot. For conical diffraction the size of the eigensystem for RCWA is also $4N + 2$. Note that N is the number of harmonics of the truncated Fourier series, but the Fourier series in the RCWA algorithm are expansions of the permittivity distribution inside a layer, while for the C method they are expansions of the interface function.

The RCWA algorithm approximates the grating structure by slicing it, while the C method does not. But the C method is restricted to interfaces which can be described by a function of the periodicity coordinate, while the RCWA algorithm can be applied to any type of grating. Because the C method uses the grating profile itself as an interface, the C method can also handle a *perfect electric conductor* or *PEC* material, which is very hard to do with RCWA.

An example of a grating structure that is ideal for the C method is a *sinusoidal grating*. Let the grating be illuminated by a field with wavelength $\lambda_0 = 0.7\Lambda$ and polar angle $\theta = 15^\circ$. The lower medium consists of silicon ($n_{II} = 3.77 - 0.01i$) and the upper medium of air

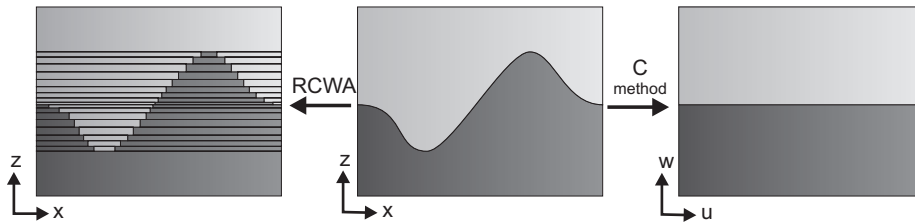


Figure 4.2: Illustration of how RCWA and the C method remove the z -dependency.

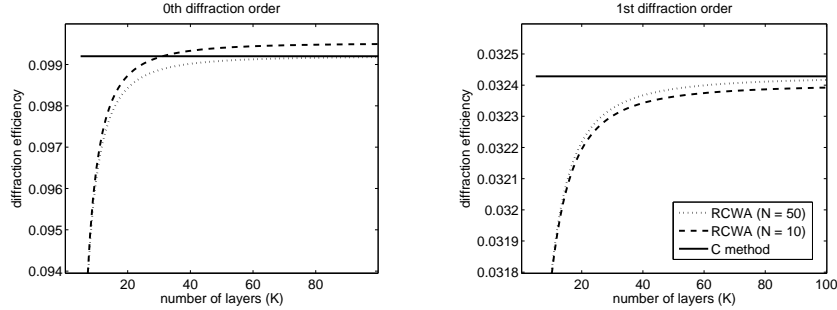


Figure 4.3: The diffraction efficiencies of the zeroth and first diffraction order, where for the RCWA algorithm the convergence with respect to the number of layers is shown.

($n_1 = 1$). Let us consider planar diffraction with TE polarization. The C method uses $N = 20$ to obtain the reflected diffraction efficiencies. To illustrate the results of RCWA, we use $N = 10$ and $N = 50$ for the number of harmonics inside the truncated Fourier series and $K = 5, \dots, 100$ to approximate the sinusoidal profile. The results are illustrated in Figure 4.3. To have 4 significant digits for both methods, we need approximately 50 layers for the RCWA algorithm. For this number of layers, the RCWA algorithm is more than 4 times slower with $N = 10$ than the C method with $N = 20$. Note that the convergence behaviour with respect to the number of layers could be improved significantly when the layers no longer have identical layer thicknesses.

We showed here that the C method has the potential to be faster and more accurate for the situation where the grating profile can be described by a continuously differentiable function. However, when we want to compute the inverse problem, the C method should be compatible with any type of grating structure. With the coordinate transformation as defined in (4.1) we cannot use the C method for binary grating structures or overhanging gratings. This issue is addressed in the next section.

4.4 Extension to overhanging gratings

An overhanging diffraction grating cannot be described by a function of the coordinate the periodicity is in, but by a parametrization. Examples of overhanging gratings are given in Figure 4.4. The binary grating can be considered as a limit case for the C method, since a binary structure can be approximated very closely by a trapezoidal structure as illustrated in [17]. To show an example of how an overhanging grating is described, let us consider case (d) of Figure 4.4 of which its profile is given by

$$\begin{cases} x(t) = au + b\sin(2Kt), \\ z(t) = c/2 + c/2\cos(Kt). \end{cases} \quad (4.40)$$

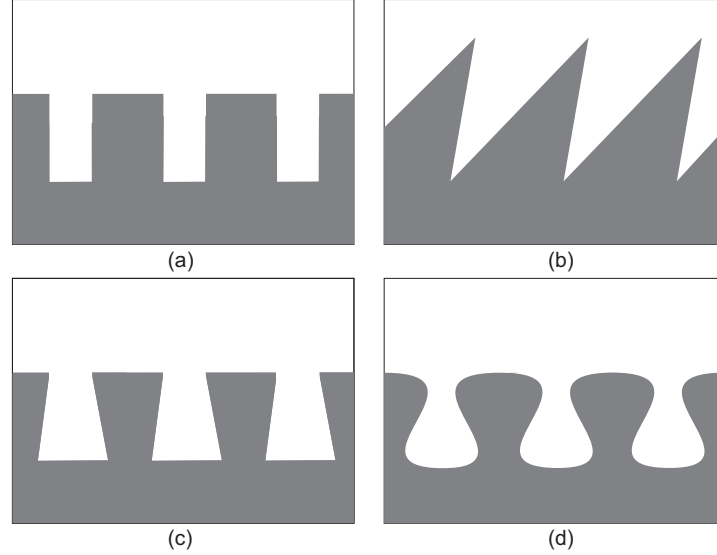


Figure 4.4: Examples of overhanging gratings: the binary grating (a), a one-side overhanging grating (b), an overhanging trapezoidal grating (c) and a smooth overhanging grating (d).

The coordinate transformation (4.1) is a general applicable coordinate transformation as long as $a(x)$ is a function and is available. Of course the number of harmonics that has to be taken into account to obtain an accurate solution depends on the smoothness of this function. The general coordinate transformation can be given by

$$\begin{cases} u = f(x, z), \\ v = y, \\ w = g(x, z), \end{cases} \quad (4.41)$$

where we have to find the functions f and g . In [45] gratings have been considered that are only overhanging on one side only, such that we can rotate the grating in order to describe the grating profile by a function again. In [16, 44] attempts have been made to generalize the coordinate transformation (4.1) by assuming a special case of (4.41), viz.

$$\begin{cases} u = f(x) + c_u z, \\ v = y, \\ w = g(x) + c_w z. \end{cases} \quad (4.42)$$

However, diffraction of a grating structure as described by (4.40) still cannot be computed. Figure 4.5 shows the main idea of the original coordinate transformation (4.1) where for each value of w the grating profile function appears but translated. If we apply the same idea to the overhanging grating described above, the lines will intersect, which implies that we can have multiple values of (u, w) to describe a point (x, z) , or in other words, the

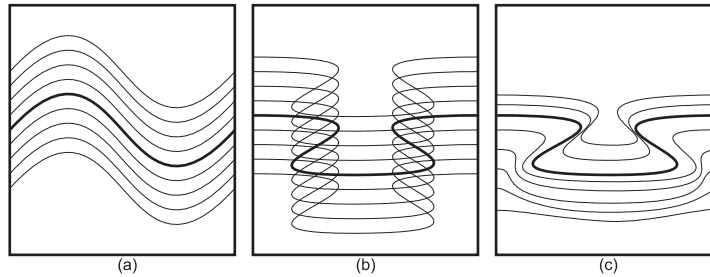


Figure 4.5: Illustration of the coordinate transformation if the grating profile can be described by a function of the coordinate the periodicity is in (a) and how applying the same idea to an overhanging grating does not give a bijective transformation (b). Figure (c) gives an example of how the generalized coordinate transformation should look like.

transformation is no longer *bijective*. For overhanging gratings we have to find a coordinate transformation that is generally applicable and is bijective.

As a feasibility study, [53] tried to find such a coordinate transformation for two special cases, namely the overhanging trapezoidal grating and the overhanging grating described by (4.40). The algorithm of the C method does not show any limitations to the coordinate transformation except that this transformation must be available in an analytical form and it should ensure that we have a bijective match between (x, y, z) and (u, v, w) .

Chapter 5

Sensitivity

Chapter 3 describes the RCWA algorithm to obtain the diffracted field when a grating structure and incident field is given. In practice, however, the grating shape can deviate from its idealized form. To find out how the deviations of the grating shape parameters affect the diffracted field, the behavior of the field with respect to such a grating shape should be found.

A method to compute the sensitivity of a diffracted field with respect to a grating shape parameter is finite differences (e.g. [33]) where RCWA is evaluated once more for slightly different settings of the parameters. The differences between the fields give approximations of the field derivatives. Remark that the choice of the number of layers should be fixed. If finite differences is performed with two different numbers of layers, an additional discretisation error is introduced. Since RCWA solves eigenvalue problems to obtain the diffracted field, every other setting requires additional eigenvalue problems to be solved. Reevaluating eigensystems is computationally expensive, so undesirable and an alternative has to be found. Differentiation of the matrix relations of the RCWA algorithm directly with respect to a shape parameter, provides a way to circumvent those reevaluations of eigensystems [1, 3].

This chapter discusses finite differences in Section 5.1. Next, the outline of the analytical method is presented in Section 5.2 and finally, both methods are compared on accuracy of the outcomes and the computational speed.

5.1 Finite differences

Finite differences is one of the two methods discussed in this chapter to obtain first-order derivatives of the field amplitudes with respect to some grating shape parameter. A physical

model is usually available to describe the shape of a grating profile. An example is the symmetric trapezoidal grating, where the shape is given by a height, width and angle. Although RCWA introduces its own set of shape parameters (see Figure 3.1), finite differences can be applied directly to the physical model. The result will be approximations of the first-order derivatives with respect to some physical parameter.

Finite differences is an intuitive and simple to implement method to compute sensitivity information of the field amplitudes. This method computes a quantity at neighbouring points to obtain first-order derivatives. It is therefore directly applicable to the physical model. We restrict ourselves to using *forward finite differences* for approximating first-order derivatives of the field amplitudes \mathbf{r} and \mathbf{t} with respect to some shape parameter p , viz.

$$\frac{\partial \mathbf{r}}{\partial p} = \frac{\mathbf{r}(p + \Delta p) - \mathbf{r}(p)}{\Delta p} + \mathcal{O}(\Delta p), \quad \frac{\partial \mathbf{t}}{\partial p} = \frac{\mathbf{t}(p + \Delta p) - \mathbf{t}(p)}{\Delta p} + \mathcal{O}(\Delta p). \quad (5.1)$$

To have higher order accuracy, we can also use *central finite differences*, which has $\mathcal{O}(\Delta p^2)$ accuracy. However, central finite differences require two additional RCWA evaluations instead of one. Because all variants of finite differences require solving additional eigensystems and solving an eigensystem is the most time-consuming part of the RCWA algorithm, it is desirable to keep these additional evaluations to a minimum.

5.2 Analytical RCWA sensitivity

The alternative is an analytical approach yielding first-order derivatives of the field amplitudes with respect to the RCWA shape parameters such as a layer thickness or a block width. This implies that the conversion to the physical parameters has to be made afterwards. However, the method avoids reevaluations of eigensystems at neighbouring values of p because it differentiates (3.72) directly with respect to a shape parameter p at $p = p_0$, viz.

$$\begin{bmatrix} -\mathbf{I} & \mathbf{F}_2 \\ -\mathfrak{M} & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{r}}{\partial p} \\ \frac{\partial \mathbf{t}_2}{\partial p} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \mathbf{F}_2}{\partial p} \\ \frac{\partial \mathbf{G}_2}{\partial p} \end{bmatrix} \mathbf{t}_2. \quad (5.2)$$

Since \mathbf{t}_2 is already available from the RCWA algorithm, the right-hand side of (5.2) is known if we are able to compute $(\partial \mathbf{F}_2)/(\partial p)$ and $(\partial \mathbf{G}_2)/(\partial p)$. Note that solving system (5.2) requires the same inverse matrix as in (3.72).

5.2.1 Sensitivity with respect to RCWA shape parameters

In order to compute $(\partial \mathbf{F}_2)/(\partial p)$ and $(\partial \mathbf{G}_2)/(\partial p)$, we have to compute all derivatives of \mathbf{A}_i , \mathbf{B}_i , \mathbf{F}_i and \mathbf{G}_i with respect to p for $i = 2, \dots, K + 1$. It is crucial for the computational

speed to know on which shape parameter(s) each matrix depends before rushing into the differentiation. Matrices \mathbf{A}_i , \mathbf{B}_i , \mathbf{F}_i and \mathbf{G}_i are updated per layer as in (3.69) and the dependencies are given by

$$\mathbf{A}_i = \mathbf{A}_i(\mathbf{x}_i, \dots, \mathbf{x}_{K+1}, d_{i+1}, \dots, d_{K+1}), \quad \mathbf{W}_i = \mathbf{W}_i(\mathbf{x}_i), \quad (5.3a)$$

$$\mathbf{B}_i = \mathbf{B}_i(\mathbf{x}_i, \dots, \mathbf{x}_{K+1}, d_{i+1}, \dots, d_{K+1}), \quad \mathbf{V}_i = \mathbf{V}_i(\mathbf{x}_i), \quad (5.3b)$$

$$\mathbf{F}_i = \mathbf{F}_i(\mathbf{x}_i, \dots, \mathbf{x}_{K+1}, d_i, \dots, d_{K+1}), \quad \mathbf{Q}_i = \mathbf{Q}_i(\mathbf{x}_i), \quad (5.3c)$$

$$\mathbf{G}_i = \mathbf{G}_i(\mathbf{x}_i, \dots, \mathbf{x}_{K+1}, d_i, \dots, d_{K+1}), \quad \mathbf{X}_i = \mathbf{X}_i(\mathbf{x}_i, d_i), \quad (5.3d)$$

where d_i is the layer thickness of layer i and where the vector \mathbf{x}_i consists of all block widths inside layer i . Differentiating \mathbf{A}_i , \mathbf{B}_i , \mathbf{F}_i and \mathbf{G}_i with respect to d_j or x_{jk} (the k th entry of \mathbf{x}_j) differs significantly because of the way the matrices depend on both parameters. Additionally, the derivatives also depend on the layer to which the RCWA shape parameter corresponds. First consider the derivatives with respect to layer thicknesses, viz.

$$\frac{\partial \mathbf{A}_i}{\partial d_j} = \begin{cases} \frac{1}{2} \left(\mathbf{W}_i^{-1} \frac{\partial \mathbf{F}_{i+1}}{\partial d_j} + \mathbf{V}_i^{-1} \frac{\partial \mathbf{G}_{i+1}}{\partial d_j} \right) & \text{if } j > i, \\ \mathbf{0} & \text{if } j \leq i, \end{cases} \quad (5.4a)$$

$$\frac{\partial \mathbf{B}_i}{\partial d_j} = \begin{cases} \frac{1}{2} \left(\mathbf{W}_i^{-1} \frac{\partial \mathbf{F}_{i+1}}{\partial d_j} - \mathbf{V}_i^{-1} \frac{\partial \mathbf{G}_{i+1}}{\partial d_j} \right) & \text{if } j > i, \\ \mathbf{0} & \text{if } j \leq i, \end{cases} \quad (5.4b)$$

$$\frac{\partial \mathbf{F}_i}{\partial d_j} = \begin{cases} \mathbf{W}_i \mathbf{X}_i \left(\frac{\partial \mathbf{B}_i}{\partial d_j} \mathbf{A}_i^{-1} + \mathbf{B}_i \frac{\partial \mathbf{A}_i^{-1}}{\partial d_j} \right) \mathbf{X}_i & \text{if } j > i, \\ \mathbf{W}_i \left(\frac{\partial \mathbf{X}_i}{\partial d_i} \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i + \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \frac{\partial \mathbf{X}_i}{\partial d_i} \right) & \text{if } j = i, \\ \mathbf{0} & \text{if } j < i, \end{cases} \quad (5.4c)$$

$$\frac{\partial \mathbf{G}_i}{\partial d_j} = \begin{cases} -\mathbf{V}_i \mathbf{X}_i \left(\frac{\partial \mathbf{B}_i}{\partial d_j} \mathbf{A}_i^{-1} + \mathbf{B}_i \frac{\partial \mathbf{A}_i^{-1}}{\partial d_j} \right) \mathbf{X}_i & \text{if } j > i, \\ -\mathbf{V}_i \left(\frac{\partial \mathbf{X}_i}{\partial d_i} \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i + \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \frac{\partial \mathbf{X}_i}{\partial d_i} \right) & \text{if } j = i, \\ \mathbf{0} & \text{if } j < i, \end{cases} \quad (5.4d)$$

where the derivative of \mathbf{X}_i with respect to d_i is given by

$$\frac{\partial \mathbf{X}_i}{\partial d_i} = -k_0 \mathbf{Q}_i \mathbf{X}_i. \quad (5.5)$$

In a similar way we can find the derivatives of \mathbf{A}_i , \mathbf{B}_i , \mathbf{F}_i and \mathbf{G}_i with respect to the block widths inside a layer, viz.

$$\frac{\partial \mathbf{A}_i}{\partial x_{jk}} = \begin{cases} \frac{1}{2} \left(\mathbf{W}_i^{-1} \frac{\partial \mathbf{F}_{i+1}}{\partial x_{jk}} + \mathbf{V}_i^{-1} \frac{\partial \mathbf{G}_{i+1}}{\partial x_{jk}} \right) & \text{if } j > i, \\ \frac{1}{2} \left(\frac{\partial \mathbf{W}_i^{-1}}{\partial x_{jk}} \mathbf{F}_{i+1} + \frac{\partial \mathbf{V}_i^{-1}}{\partial x_{jk}} \mathbf{G}_{i+1} \right) & \text{if } j = i, \\ \mathbf{0} & \text{if } j < i, \end{cases} \quad (5.6a)$$

$$\frac{\partial \mathbf{B}_i}{\partial x_{jk}} = \begin{cases} \frac{1}{2} \left(\mathbf{W}_i^{-1} \frac{\partial \mathbf{F}_{i+1}}{\partial x_{jk}} - \mathbf{V}_i^{-1} \frac{\partial \mathbf{G}_{i+1}}{\partial x_{jk}} \right) & \text{if } j > i, \\ \frac{1}{2} \left(\frac{\partial \mathbf{W}_i^{-1}}{\partial x_{jk}} \mathbf{F}_{i+1} - \frac{\partial \mathbf{V}_i^{-1}}{\partial x_{jk}} \mathbf{G}_{i+1} \right) & \text{if } j = i, \\ \mathbf{0} & \text{if } j < i, \end{cases} \quad (5.6b)$$

$$\frac{\partial \mathbf{F}_i}{\partial x_{jk}} = \begin{cases} \mathbf{W}_i \mathbf{X}_i \left(\frac{\partial \mathbf{B}_i}{\partial x_{jk}} \mathbf{A}_i^{-1} + \mathbf{B}_i \frac{\partial \mathbf{A}_i^{-1}}{\partial x_{jk}} \right) \mathbf{X}_i & \text{if } j > i, \\ \frac{\partial \mathbf{W}_i}{\partial x_{ik}} (\mathbf{I} + \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i) + \mathbf{W}_i \left(\frac{\partial \mathbf{X}_i}{\partial x_{ik}} \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i \right. \\ \left. + \mathbf{X}_i \frac{\partial \mathbf{B}_i}{\partial x_{ik}} \mathbf{A}_i^{-1} \mathbf{X}_i + \mathbf{X}_i \mathbf{B}_i \frac{\partial \mathbf{A}_i^{-1}}{\partial x_{ik}} \mathbf{X}_i + \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \frac{\partial \mathbf{X}_i}{\partial x_{ik}} \right) & \text{if } j = i, \\ \mathbf{0} & \text{if } j < i, \end{cases} \quad (5.6c)$$

$$\frac{\partial \mathbf{G}_i}{\partial x_{jk}} = \begin{cases} -\mathbf{V}_i \mathbf{X}_i \left(\frac{\partial \mathbf{B}_i}{\partial x_{jk}} \mathbf{A}_i^{-1} + \mathbf{B}_i \frac{\partial \mathbf{A}_i^{-1}}{\partial x_{jk}} \right) \mathbf{X}_i & \text{if } j > i, \\ \frac{\partial \mathbf{V}_i}{\partial x_{ik}} (\mathbf{I} - \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i) - \mathbf{V}_i \left(\frac{\partial \mathbf{X}_i}{\partial x_{ik}} \mathbf{B}_i \mathbf{A}_i^{-1} \mathbf{X}_i \right. \\ \left. + \mathbf{X}_i \frac{\partial \mathbf{B}_i}{\partial x_{ik}} \mathbf{A}_i^{-1} \mathbf{X}_i + \mathbf{X}_i \mathbf{B}_i \frac{\partial \mathbf{A}_i^{-1}}{\partial x_{ik}} \mathbf{X}_i + \mathbf{X}_i \mathbf{B}_i \mathbf{A}_i^{-1} \frac{\partial \mathbf{X}_i}{\partial x_{ik}} \right) & \text{if } j = i, \\ \mathbf{0} & \text{if } j < i, \end{cases} \quad (5.6d)$$

where x_{jk} denotes the k th block width inside layer j . The computation of the derivatives of \mathbf{W}_i , \mathbf{V}_i and \mathbf{X}_i and their inverses with respect to block width k are discussed later in this section and in the next chapter.

To find the right-hand side of (5.2) we consider two different approaches. The first method is to update matrices $(\partial \mathbf{F}_i)/(\partial p)$ and $(\partial \mathbf{G}_i)/(\partial p)$ for each layer according to (5.4) or (5.6) simultaneously as \mathbf{F}_i and \mathbf{G}_i . Finally we obtain \mathbf{F}_2 , \mathbf{G}_2 , $(\partial \mathbf{F}_2)/(\partial p)$ and $(\partial \mathbf{G}_2)/(\partial p)$. Next, compute \mathbf{r} and \mathbf{t}_2 first and use \mathbf{t}_2 to compute the right-hand side of (5.2). The second method takes advantage of the fact that \mathbf{t}_2 has to be known before we can evaluate the right-hand side. Therefore \mathbf{r} and \mathbf{t}_2 can be computed by an RCWA evaluation. Next, the products $(\partial \mathbf{F}_2)/(\partial p) \mathbf{t}_2$ and $(\partial \mathbf{G}_2)/(\partial p) \mathbf{t}_2$ can be found by first finding $(\partial \mathbf{F}_j)/(\partial p_j) \mathbf{t}_j$ and $(\partial \mathbf{G}_j)/(\partial p_j) \mathbf{t}_j$ for $j = 2, \dots, K+1$ and using the matrices already computed in the ordinary RCWA evaluation to link them to each other. This second method uses matrix-vector products, while the first method only has matrix-matrix products. For this reason we will call the first method the *matrix-matrix approach* and the second the *matrix-vector approach*. The order in which both approaches are executed is summarized in Figure 5.1.

Let us consider the matrix-vector approach in more detail. If we write out the formulas of $(\partial \mathbf{F}_2)/(\partial p_j) \mathbf{t}_2$ and $(\partial \mathbf{G}_2)/(\partial p_j) \mathbf{t}_2$, we have the following result for the derivative

$$\begin{bmatrix} \frac{\partial \mathbf{F}_2}{\partial p_j} \\ \frac{\partial \mathbf{G}_2}{\partial p_j} \end{bmatrix} \mathbf{t}_2 = \frac{1}{2^{j-2}} \prod_{i=2}^{j-1} \left\{ \begin{bmatrix} \mathbf{W}_i \\ -\mathbf{V}_i \end{bmatrix} \mathbf{X}_i \left(\begin{bmatrix} \mathbf{W}_i^{-1} & -\mathbf{V}_i^{-1} \end{bmatrix} - \mathbf{B}_i \mathbf{A}_i^{-1} \begin{bmatrix} \mathbf{W}_i^{-1} & \mathbf{V}_i^{-1} \end{bmatrix} \right) \right\} \\ \cdot \begin{bmatrix} \mathfrak{A}_j \\ \mathfrak{B}_j \end{bmatrix} \prod_{i=j-1}^2 \{ \mathbf{A}_i^{-1} \mathbf{X}_i \} \mathbf{t}_2, \quad (5.7)$$

where p_j is a shape parameter inside layer j . Matrices \mathfrak{A}_j and \mathfrak{B}_j depend on whether the shape parameter is a layer thickness d_j or a block width x_{jk} . The matrices \mathfrak{A}_j and \mathfrak{B}_j for

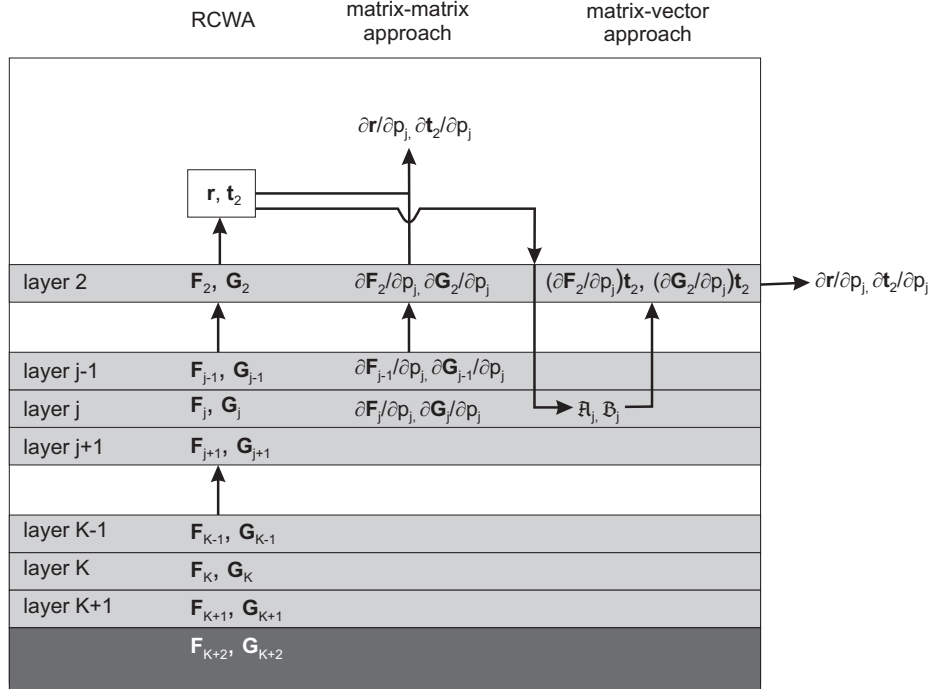


Figure 5.1: Illustration of the order in which the matrix-matrix and matrix-vector approach computes its derivatives of the field amplitudes.

the layer thickness are given by

$$\begin{bmatrix} \mathfrak{A}_j \\ \mathfrak{B}_j \end{bmatrix} := \begin{bmatrix} \mathbf{W}_j \\ -\mathbf{V}_j \end{bmatrix} \left(\frac{\partial \mathbf{X}_j}{\partial d_j} \mathbf{B}_j \mathbf{A}_j^{-1} \mathbf{X}_j + \mathbf{X}_j \mathbf{B}_j \mathbf{A}_j^{-1} \frac{\partial \mathbf{X}_j}{\partial d_j} \right). \quad (5.8)$$

For a block width x_{jk} the relation becomes more complicated, because more matrices depend on this parameter (see (5.3)), viz.

$$\begin{aligned} \begin{bmatrix} \mathfrak{A}_j \\ \mathfrak{B}_j \end{bmatrix} &:= \begin{bmatrix} \mathbf{W}_j \\ -\mathbf{V}_j \end{bmatrix} \left\{ \frac{1}{2} \mathbf{X}_j \left((\mathbf{I} - \mathbf{B}_j \mathbf{A}_j^{-1}) \frac{\partial \mathbf{W}_j^{-1}}{\partial x_{jk}} \mathbf{F}_{j+1} - (\mathbf{I} + \mathbf{B}_j \mathbf{A}_j^{-1}) \frac{\partial \mathbf{V}_j^{-1}}{\partial x_{jk}} \mathbf{G}_{j+1} \right) \right. \\ &\quad \left. + \frac{\partial \mathbf{X}_j}{\partial x_{jk}} \mathbf{B}_j \mathbf{A}_j^{-1} \mathbf{X}_j + \mathbf{X}_j \mathbf{B}_j \mathbf{A}_j^{-1} \frac{\partial \mathbf{X}_j}{\partial x_{jk}} \right\} + \begin{bmatrix} \frac{\partial \mathbf{W}_j}{\partial x_{jk}} (\mathbf{I} + \mathbf{X}_j \mathbf{B}_j \mathbf{A}_j^{-1} \mathbf{X}_j) \\ \frac{\partial \mathbf{V}_j}{\partial x_{jk}} (\mathbf{I} - \mathbf{X}_j \mathbf{B}_j \mathbf{A}_j^{-1} \mathbf{X}_j) \end{bmatrix}. \quad (5.9) \end{aligned}$$

When the right-hand side of (5.2) has been found, the derivatives of the field amplitudes with respect to the shape parameter can be computed by solving (5.2). To find the derivatives of the transmitted field amplitudes, (3.73) should be differentiated with respect to the shape

parameter p , viz.

$$\begin{aligned}
\frac{\partial \mathbf{t}}{\partial p} &= \mathbf{A}_{K+1}^{-1} \mathbf{X}_{K+1} \dots \mathbf{A}_{i+1}^{-1} \mathbf{X}_{i+1} \left(\frac{\partial \mathbf{A}_i^{-1}}{\partial p} \mathbf{X}_i + \mathbf{A}_i^{-1} \frac{\partial \mathbf{X}_i}{\partial p} \right) \mathbf{A}_{i-1}^{-1} \mathbf{X}_{i-1} \dots \mathbf{A}_2^{-1} \mathbf{X}_2 \mathbf{t}_2 \\
&+ \mathbf{A}_{K+1}^{-1} \mathbf{X}_{K+1} \dots \mathbf{A}_i^{-1} \mathbf{X}_i \left(\left(\frac{\partial \mathbf{A}_{i-1}^{-1}}{\partial p} \mathbf{X}_{i-1} \right) \mathbf{A}_{i-2}^{-1} \mathbf{X}_{i-2} \dots \mathbf{A}_2^{-1} \mathbf{X}_2 + \dots \right. \\
&\quad \left. + \mathbf{A}_{i-1}^{-1} \mathbf{X}_{i-1} \dots \mathbf{A}_3^{-1} \mathbf{X}_3 \left(\frac{\partial \mathbf{A}_2^{-1}}{\partial p} \mathbf{X}_2 \right) \right) \mathbf{t}_2 + \dots \\
&+ \mathbf{A}_{K+1}^{-1} \mathbf{X}_{K+1} \dots \mathbf{A}_2^{-1} \mathbf{X}_2 \frac{\partial \mathbf{t}_2}{\partial p}, \tag{5.10}
\end{aligned}$$

where the parameter p is assumed to be defined for layer i with $2 \leq i \leq K+1$. Regardless of whether the matrix-matrix or the matrix-vector approach has been used to obtain the derivatives of the field amplitudes, the final step is to compute the derivatives of the diffraction efficiencies as defined in (2.53). Straightforward differentiation with respect to a shape parameter p gives

$$\frac{\partial \eta_{rn}}{\partial p} = 2 \operatorname{Re} \left\{ \frac{\partial r_{sn}}{\partial p} \bar{r}_{sn} \right\} \operatorname{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\} + 2 \operatorname{Re} \left\{ \frac{\partial r_{pn}}{\partial p} \bar{r}_{pn} \right\} \operatorname{Re} \left\{ \frac{k_{l,zn}}{k_0 n_l \cos \theta} \right\}, \tag{5.11a}$$

$$\frac{\partial \eta_{tn}}{\partial p} = 2 \operatorname{Re} \left\{ \frac{\partial t_{sn}}{\partial p} \bar{t}_{sn} \right\} \operatorname{Re} \left\{ \frac{k_{ll,zn}}{k_0 n_l \cos \theta} \right\} + 2 \operatorname{Re} \left\{ \frac{\partial t_{pn}}{\partial p} \bar{t}_{pn} \right\} \operatorname{Re} \left\{ \frac{k_{ll,zn} n_l}{k_0 n_l^2 \cos \theta} \right\}. \tag{5.11b}$$

Note that if the shape parameter under consideration is a layer thickness, we have now derived the complete method. For a block width we still have to find the derivatives of \mathbf{W}_i , \mathbf{V}_i , \mathbf{X}_i and their inverse matrices.

When the RCWA shape parameter under consideration is a block width, first-order derivatives of the eigenvalues and eigenvectors have to be computed. It is important to realize that the matrices discussed from this point are only defined for one layer only. Therefore, they only depend on the shape parameters defined for that layer. The actual computation of the eigenvalue and eigenvector derivatives are postponed to Chapter 6, but we will do some preliminary work. As will be shown in Chapter 6, the computation of eigenvalues and eigenvectors require the derivative of the matrix from which the eigenvalues and eigenvectors have to be computed. This implies that we need the derivatives of \mathbf{A}_i , \mathbf{B}_i , \mathbf{C}_i and \mathbf{D}_i , which are defined by (3.13), (3.21), (3.29) and (3.32). These derivatives in turn require the derivatives of the Toeplitz matrices \mathbf{E}_i and \mathbf{P}_i . RCWA allows for analytical expressions of the Fourier coefficients of the relative permittivity and/or its reciprocal. Since the formulas for \mathbf{P}_i are similar to those of \mathbf{E}_i , we restrict ourselves to \mathbf{E}_i . The piecewise-constant relative permittivity $\tilde{\epsilon}_i^r(x)$ is given by

$$\tilde{\epsilon}_i^r(x) := \begin{cases} \tilde{\epsilon}_{i,1}^r, & -\Lambda/2 \leq x \leq t_{i1}, \\ \tilde{\epsilon}_{i,2}^r, & t_{i1} \leq x \leq t_{i2}, \\ \vdots & \vdots \\ \tilde{\epsilon}_{i,M}^r, & t_{i(M-1)} \leq x \leq \Lambda/2, \end{cases} \tag{5.12}$$

with $-\Lambda/2 < t_{i1} < \dots < t_{i(M-1)} < \Lambda/2$, where t_{ik} is the k th position of a material transition inside layer i . Note that the number of transitions M can vary between layers. The Fourier

coefficients $\epsilon_{i,m}$ of this relative permittivity can be found by

$$\epsilon_{i,m} = \begin{cases} \frac{i}{2\pi m} \tilde{\epsilon}_{i,1}^r (\exp[-i\frac{2\pi}{\Lambda} m t_{i1}] - \exp[i\pi m]) \\ \quad + \frac{i}{2\pi m} \tilde{\epsilon}_{i,2}^r (\exp[-i\frac{2\pi}{\Lambda} m t_{i2}] - \exp[-i\frac{2\pi}{\Lambda} m t_{i1}]) + \dots \\ \quad + \frac{i}{2\pi m} \tilde{\epsilon}_{i,M}^r (\exp[-i\pi m] - \exp[-i\frac{2\pi}{\Lambda} m t_{i(M-1)}]) , & m \neq 0, \\ \frac{1}{\Lambda} \tilde{\epsilon}_{i,1}^r (t_{i1} + \Lambda/2) + \frac{1}{\Lambda} \tilde{\epsilon}_{i,2}^r (t_{i2} - t_{i1}) \\ \quad + \dots + \frac{1}{\Lambda} \tilde{\epsilon}_{i,M}^r (\Lambda/2 - t_{i(M-1)}), & m = 0. \end{cases} \quad (5.13)$$

The derivatives of the Fourier coefficients with respect to a position of a material transition can now easily be found, viz.

$$\frac{\partial \epsilon_{i,m}}{\partial t_{ir}} = \frac{1}{\Lambda} \exp[-i\frac{2\pi}{\Lambda} m t_r] (\tilde{\epsilon}_{ir}^r - \tilde{\epsilon}_{i(r+1)}^r). \quad (5.14)$$

Instead of positions of material transitions, we prefer to use block widths instead. It reduces the number of significant shape parameters. For example, a binary grating can be described by one block, while in general it has 2 transitions. Because a block width x_{ik} is related to the transitions as $x_{ik} := t_{ik} - t_{i(k-1)}$, we have

$$\frac{\partial \epsilon_{i,m}}{\partial x_{ik}} = \frac{\partial \epsilon_{i,m}}{\partial t_{ik}} - \frac{\partial \epsilon_{i,m}}{\partial t_{i(k-1)}}, \quad (5.15)$$

for $k = 1, \dots, M$. For the RCWA algorithm the Fourier coefficients are collected in the Toeplitz matrix \mathbf{E}_i with element (m, n) equal to $\epsilon_{i,m-n}$. Similarly the derivative of this matrix with respect to a block width x_{ik} has element (m, n) equal to $(\partial \epsilon_{i,n-m}) / (\partial x_{ik})$. To find the derivatives with respect to a block width x_{ik} of the matrices of which the eigenvalues and eigenvectors have to be found, first the derivatives of the matrices \mathbf{A}_i and \mathbf{B}_i as defined in (3.13) and (3.21), respectively, will be computed. For notational convenience denote the derivative with respect to a block width x_{ik} by a prime.

$$\mathbf{A}_i' = -\mathbf{E}_i', \quad (5.16a)$$

$$\mathbf{B}_i' = \mathbf{K}_x (\mathbf{E}_i^{-1})' \mathbf{K}_x. \quad (5.16b)$$

The derivative of the inverse matrix \mathbf{E}_i^{-1} can easily be found by differentiating the matrix relation $\mathbf{I} = \mathbf{E}_i^{-1} \mathbf{E}_i$, viz.

$$(\mathbf{E}_i^{-1})' = -\mathbf{E}_i^{-1} \mathbf{E}_i' \mathbf{E}_i^{-1}. \quad (5.17)$$

Since \mathbf{A}_i is already the matrix of which eigenvalues and eigenvectors have to be computed for planar TE polarization, we only have to consider the matrices for TM polarization and conical diffraction as defined by (3.20), (3.29) and (3.32), respectively.

$$(\mathbf{P}_i^{-1} \mathbf{B}_i)' = (\mathbf{P}_i^{-1})' \mathbf{B}_i + \mathbf{P}_i^{-1} \mathbf{B}_i', \quad (5.18a)$$

$$\mathbf{C}_i' = \mathbf{B}_i' \mathbf{P}_i^{-1} + \mathbf{B}_i (\mathbf{P}_i^{-1})', \quad (5.18b)$$

$$\mathbf{D}_i' = \mathbf{A}_i'. \quad (5.18c)$$

Here, (5.17) can be used to compute the derivatives of the inverse matrices.

Although the eigenvalue and eigenvector derivatives will be treated in Chapter 6, let us assume that the derivatives of the eigenvalue and eigenvector matrix Λ'_i and \mathbf{W}'_i , respectively, are available. To compute the derivatives of \mathbf{V}_i and \mathbf{X}_i with respect to a block width x_{ik} , first the derivative of \mathbf{Q}_i should be computed. Since \mathbf{Q}_i is a diagonal matrix consisting of the square roots with a positive real part of the eigenvalue matrix Λ_i , its derivative with respect to a block width x_{ik} is given by

$$\mathbf{Q}'_i = \frac{1}{2} \mathbf{Q}_i^{-1} \Lambda'_i. \quad (5.19)$$

Matrix \mathbf{X}'_i can now be computed by

$$\mathbf{X}'_i = -k_0 d_i \mathbf{X}_i \mathbf{Q}'_i. \quad (5.20)$$

Note that for conical diffraction, \mathbf{X}_i and \mathbf{Q}_i consist of $\mathbf{X}_{1,i}$, $\mathbf{X}_{2,i}$, $\mathbf{Q}_{1,i}$ and $\mathbf{Q}_{2,i}$ and their derivatives are computed similarly.

Computation of \mathbf{V}'_i for planar diffraction yields

$$\mathbf{V}'_i = \begin{cases} \mathbf{W}'_i \mathbf{Q}_i + \mathbf{W}_i \mathbf{Q}'_i & \text{for TE polarization,} \\ \mathbf{P}'_i \mathbf{W}_i \mathbf{Q}_i + \mathbf{P}_i \mathbf{W}'_i \mathbf{Q}_i + \mathbf{P}_i \mathbf{W}_i \mathbf{Q}'_i & \text{for TM polarization.} \end{cases} \quad (5.21)$$

For conical diffraction the matrices \mathbf{W}_i and \mathbf{V}_i contain four block matrices of which the derivatives are given by

$$\mathbf{V}'_{11,i} = (\mathcal{A}_i^{-1})' \mathbf{W}_{1,i} \mathbf{Q}_{1,i} + \mathcal{A}_i^{-1} \mathbf{W}'_{1,i} \mathbf{Q}_{1,i} + \mathcal{A}_i^{-1} \mathbf{W}_{1,i} \mathbf{Q}'_{1,i}, \quad (5.22a)$$

$$\mathbf{V}'_{12,i} = \mathbf{K}_y (\mathcal{A}_i^{-1})' \mathbf{K}_x \mathbf{W}_{2,i} + \mathbf{K}_y \mathcal{A}_i^{-1} \mathbf{K}_x \mathbf{W}'_{2,i}, \quad (5.22b)$$

$$\mathbf{V}'_{21,i} = \mathbf{K}_y (\mathcal{B}_i^{-1})' \mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{W}_{1,i} + \mathbf{K}_y \mathcal{B}_i^{-1} \mathbf{K}_x (\mathbf{E}_i^{-1})' \mathbf{W}_{1,i} + \mathbf{K}_y \mathcal{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{W}'_{1,i}, \quad (5.22c)$$

$$\mathbf{V}'_{22,i} = (\mathcal{B}_i^{-1})' \mathbf{W}_{2,i} \mathbf{Q}_{2,i} + \mathcal{B}_i^{-1} \mathbf{W}'_{2,i} \mathbf{Q}_{2,i} + \mathcal{B}_i^{-1} \mathbf{W}_{2,i} \mathbf{Q}'_{2,i}. \quad (5.22d)$$

The derivatives of the inverse matrices of \mathbf{W}_i , \mathbf{V}_i and \mathbf{X}_i can be found by using (5.17).

This completes the description of the analytical approach. By avoiding solving additional eigensystems, the analytical approach is faster than finite differences. Another feature of the analytical approach is that it does not need any additional approximations other than those already made within RCWA itself.

5.2.2 Sensitivity with respect to physical shape parameters

In practice a grating shape is described by a set of physical shape parameters. Although RCWA needs a conversion step to the RCWA shape parameters and we can compute the first-order derivatives of the field amplitudes with respect to those shape parameters, we would like to find a method to compute these derivatives with respect to the physical shape

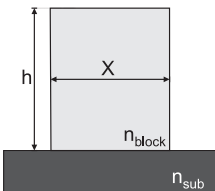
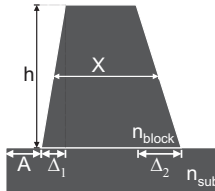
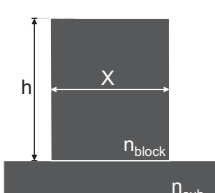
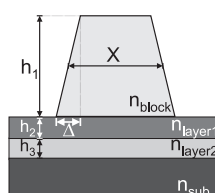
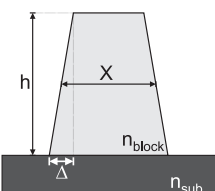
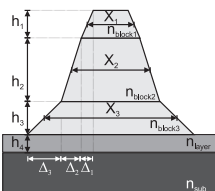
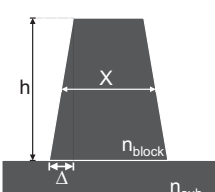
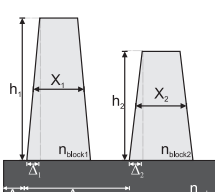
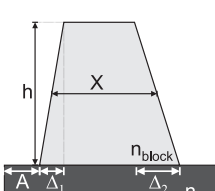
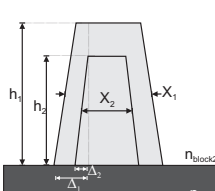
<p>Test case 1: Binary grating (resist)</p> 	<p>Parameter values $h = 0.5\Lambda$ $X = 0.5\Lambda$ $n_{\text{block}} = 1.51$ $n_{\text{sub}} = 3.77 - 0.01i$</p>	<p>Test case 6: Asymmetric trapezoid (silicon)</p> 	<p>Parameter values $h = 0.5\Lambda$ $X = 0.5\Lambda$ $\Delta_1 = 0.05\Lambda$ $\Delta_2 = 0.10\Lambda$ $A = 0$ $n_{\text{block}} = 3.77 - 0.01i$ $n_{\text{sub}} = 3.77 - 0.01i$</p>
<p>Test case 2: Binary grating (silicon)</p> 	<p>Parameter values $h = 0.5\Lambda$ $X = 0.5\Lambda$ $n_{\text{block}} = 3.77 - 0.01i$ $n_{\text{sub}} = 3.77 - 0.01i$</p>	<p>Test case 7: Symmetric trapezoid on homogeneous layers</p> 	<p>Parameter values $h_1 = 0.4\Lambda$ $n_{\text{block}} = 1.51$ $X = 0.14\Lambda$ $n_{\text{layer1}} = 1.56 - 0.01i$ $\Delta = 0.04\Lambda$ $n_{\text{layer2}} = 1.46$ $h_2 = 0.18\Lambda$ $n_{\text{sub}} = 3.77 - 0.01i$ $h_3 = 0.004\Lambda$</p>
<p>Test case 3: Symmetric trapezoid (resist)</p> 	<p>Parameter values $h = 0.5\Lambda$ $X = 0.5\Lambda$ $\Delta = 0.05\Lambda$ $n_{\text{block}} = 1.51$ $n_{\text{sub}} = 3.77 - 0.01i$</p>	<p>Test case 8: Stack of symmetric trapezoids</p> 	<p>Parameter values $h_1 = 0.02\Lambda$ $n_{\text{block1}} = 1.51$ $h_2 = 0.15\Lambda$ $n_{\text{block2}} = 1.51$ $h_3 = 0.03\Lambda$ $n_{\text{block3}} = 1.51$ $h_4 = 0.1\Lambda$ $n_{\text{layer}} = 1.56 - 0.01i$ $X_1 = 0.075\Lambda$ $n_{\text{sub}} = 3.77 - 0.01i$ $X_2 = 0.1\Lambda$ $\Delta_1 = 0.02\Lambda$ $X_3 = 0.135\Lambda$ $\Delta_2 = 0.005\Lambda$ $\Delta_3 = 0.01\Lambda$</p>
<p>Test case 4: Symmetric trapezoid (silicon)</p> 	<p>Parameter values $h = 0.5\Lambda$ $X = 0.5\Lambda$ $\Delta = 0.05\Lambda$ $n_{\text{block}} = 3.77 - 0.01i$ $n_{\text{sub}} = 3.77 - 0.01i$</p>	<p>Test case 9: Two symmetric trapezoids next to each other</p> 	<p>Parameter values $h_1 = 0.78\Lambda$ $h_2 = 0.80\Lambda$ $X_1 = 0.25\Lambda$ $X_2 = 0.25\Lambda$ $\Delta_1 = 0.05\Lambda$ $\Delta_2 = 0.05\Lambda$ $A_1 = 0$ $n_{\text{block1}} = 1.51$ $A_2 = 0.5\Lambda$ $n_{\text{block2}} = 1.51$ $n_{\text{sub}} = 3.77 - 0.01i$</p>
<p>Test case 5: Asymmetric trapezoid (resist)</p> 	<p>Parameter values $h = 0.5\Lambda$ $X = 0.5\Lambda$ $\Delta_1 = 0.05\Lambda$ $\Delta_2 = 0.10\Lambda$ $A = 0$ $n_{\text{block}} = 1.51$ $n_{\text{sub}} = 3.77 - 0.01i$</p>	<p>Test case 10: Coated symmetric trapezoid</p> 	<p>Parameter values $h_1 = 0.55\Lambda$ $h_2 = 0.50\Lambda$ $X_1 = 0.35\Lambda$ $X_2 = 0.25\Lambda$ $\Delta_1 = 0.05\Lambda$ $\Delta_2 = 0.05\Lambda$ $A_1 = 0$ $n_{\text{block1}} = 1.51$ $A_2 = 0$ $n_{\text{block2}} = 1.75$ $n_{\text{sub}} = 3.77 - 0.01i$</p>

Table 5.1: Definition of the shape parameters and their values for ten grating profiles to test the sensitivity theory.

parameters. In Table 5.1 ten test cases are defined which will be used in this chapter and in Chapter 7. Each test case has a physical model of the grating structure.

The binary grating (test cases 1 and 2) can be approximated geometrically by one layer only and the physical parameters X and h are also the RCWA shape parameters x_{22} and d_2 . The other test cases involve more extensive physical models. For the symmetric trapezoid (test cases 3 and 4) the RCWA shape parameters can be described in terms of the physical parameters, i.e.

$$d_i := \frac{h}{K}, \quad (5.23a)$$

$$x_{i2} := X + \frac{K - 2i + 3}{K} \Delta, \quad (5.23b)$$

for $i = 2, \dots, K + 1$. We assume here that all K layers have the same layer thickness. Note that x_{i1} and x_{i3} are not considered because they are irrelevant for describing the symmetric trapezoidal grating because of the symmetry. By using (5.23) and the quotient rule for differentiation we can express the derivatives with respect to the physical parameters in terms of derivatives with respect to RCWA shape parameters, viz.

$$\frac{\partial}{\partial h} = \sum_{i=2}^{K+1} \frac{\partial d_i}{\partial h} \frac{\partial}{\partial d_i} + \frac{\partial x_{i2}}{\partial h} \frac{\partial}{\partial x_{i2}} = \frac{1}{K} \sum_{i=2}^{K+1} \frac{\partial}{\partial d_i}, \quad (5.24a)$$

$$\frac{\partial}{\partial X} = \sum_{i=2}^{K+1} \frac{\partial}{\partial x_{i2}}, \quad (5.24b)$$

$$\frac{\partial}{\partial \Delta} = \sum_{i=2}^{K+1} \frac{K - 2i + 3}{K} \frac{\partial}{\partial x_{i2}}. \quad (5.24c)$$

As can be seen from (5.24b) and (5.24c) the derivatives with respect to X and Δ are closely related. As a consequence of the choice of the physical parameters eigenvalue and eigenvector derivatives are not involved in the derivatives with respect to the height of the trapezoid. The test cases 5 until 8 are similar to the symmetric trapezoid.

For test cases 9 and 10 we have to do an additional step to express the field amplitude derivatives with respect to the physical parameters in terms of the derivatives with respect to the RCWA shape parameters. We define two symmetric trapezoids to have two trapezoids next to each other or to have a coated trapezoid. Since RCWA introduces horizontal layers over the entire period, both trapezoids share the same set of layers. The highest trapezoid has some additional layers. By dividing the highest trapezoid into two smaller trapezoids as has been done in Figure 5.2 we can express the RCWA shape parameters in terms of these intermediate parameters defined by

$$h_a = h_1 - h_2, \quad h_b = h_2, \quad (5.25a)$$

$$X_a = X_1 - \frac{h_2}{h_1} \Delta_1, \quad X_b = X_2 - \frac{h_1 - h_2}{h_1} \Delta_1, \quad X_c = X_2, \quad (5.25b)$$

$$\Delta_a = \frac{h_1 - h_2}{h_1} \Delta_1, \quad \Delta_b = \frac{h_2}{h_1} \Delta_1, \quad \Delta_c = \Delta_2. \quad (5.25c)$$

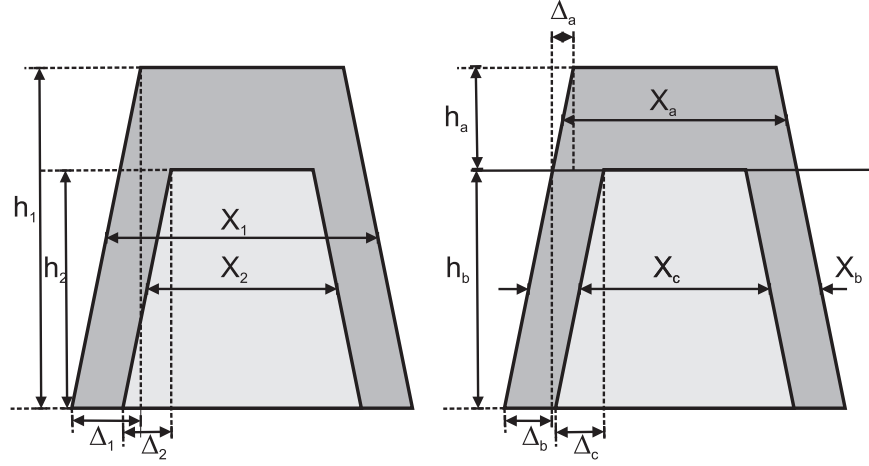


Figure 5.2: Definition of the physical shape parameters (*left*) and intermediate shape parameters (*right*) of two trapezoids next to each other.

However, eventually we want to have the derivatives with respect to the actual physical parameters and not these intermediate parameters. Therefore, we have to express the derivatives with respect to the physical shape parameters in terms of the intermediate parameters as

$$\frac{\partial}{\partial h_1} = \frac{\partial}{\partial h_a} + \frac{h_2}{h_1^2} \Delta_1 \left(\frac{\partial}{\partial \Delta_a} - \frac{\partial}{\partial \Delta_b} + \frac{\partial}{\partial X_a} + \frac{\partial}{\partial X_b} \right), \quad (5.26a)$$

$$\frac{\partial}{\partial h_2} = -\frac{\partial}{\partial h_a} + \frac{\partial}{\partial h_b} + \frac{\Delta_1}{h_1} \left(-\frac{\partial}{\partial \Delta_a} + \frac{\partial}{\partial \Delta_b} - \frac{\partial}{\partial X_a} - \frac{\partial}{\partial X_b} \right), \quad (5.26b)$$

$$\frac{\partial}{\partial X_1} = \frac{\partial}{\partial X_a} + \frac{\partial}{\partial X_b}, \quad (5.26c)$$

$$\frac{\partial}{\partial X_2} = \frac{\partial}{\partial X_c}, \quad (5.26d)$$

$$\frac{\partial}{\partial \Delta_1} = \frac{h_1 - h_2}{h_1} \left(\frac{\partial}{\partial \Delta_a} + \frac{\partial}{\partial X_b} \right) + \frac{h_2}{h_1} \left(\frac{\partial}{\partial \Delta_b} - \frac{\partial}{\partial X_a} \right), \quad (5.26e)$$

$$\frac{\partial}{\partial \Delta_2} = \frac{\partial}{\partial \Delta_c}. \quad (5.26f)$$

The computation of the field amplitude derivatives with respect to the intermediate parameters can be found similarly as when we compute the field amplitude derivatives of a symmetric trapezoid. We use (5.26) to find the derivatives with respect to the physical parameters.

We are now ready to compare the analytical method to the finite differences method, since finite differences find the field amplitude derivatives immediately with respect to the physical shape parameters. The test cases from Table 5.1 will be used in Section 5.4 to compare both methods.

5.2.3 Sensitivity with respect to other parameters

Besides derivatives with respect to shape parameters also derivatives with respect to the refractive index, period of the grating or the angle of incidence can be computed. The idea of the analytical method remains the same except for the derivative of the matrix from which the eigenvalues and eigenvectors have to be computed. In this section we will discuss only the derivatives with respect to the refractive index and the angle of incidence.

The refractive index is encapsulated inside the Fourier coefficients of the relative permittivity and its reciprocal. Therefore, the derivative with respect to the refractive index appears in matrices \mathbf{E}_i and \mathbf{P}_i . The derivative will be considered with respect to the real and imaginary part of the refractive index, $n_{i,r}$ and $k_{i,r}$ respectively. The relative permittivity $\tilde{\epsilon}_{i,r}^r$ can be related to the refractive index according to

$$\tilde{\epsilon}_{i,r}^r = (n_{i,r} - ik_{i,r})^2. \quad (5.27)$$

The first-order derivatives of the Toeplitz matrix \mathbf{E}_i are given by

$$\left(\frac{\partial \mathbf{E}_i}{\partial n_{i,r}} \right)_{nm} = \begin{cases} \frac{i}{\pi(n-m)} (n_{i,r} - ik_{i,r}) \\ \cdot (\exp[-i\frac{2\pi}{\lambda} m x_r] - \exp[-i\frac{2\pi}{\lambda} (n-m)x_{r-1}]), & m \neq 0, \\ 2(n_{i,r} - ik_{i,r})(x_r - x_{r-1}), & m = 0, \end{cases} \quad (5.28)$$

$$\left(\frac{\partial \mathbf{E}_i}{\partial k_{i,r}} \right)_{nm} = \begin{cases} \frac{1}{\pi(n-m)} (n_{i,r} - ik_{i,r}) \\ \cdot (\exp[-i\frac{2\pi}{\lambda} m x_r] - \exp[-i\frac{2\pi}{\lambda} (n-m)x_{r-1}]), & m \neq 0, \\ -2i(n_{i,r} - ik_{i,r})(x_r - x_{r-1}), & m = 0, \end{cases} \quad (5.29)$$

and similarly the first-order derivatives of \mathbf{P}_i are given by

$$\left(\frac{\partial \mathbf{P}_i}{\partial n_{i,r}} \right)_{nm} = \begin{cases} -\frac{i}{\pi(n-m)} \frac{1}{(n_{i,r} - ik_{i,r})^3} \\ \cdot (\exp[-i\frac{2\pi}{\lambda} m x_r] - \exp[-i\frac{2\pi}{\lambda} (n-m)x_{r-1}]), & m \neq 0, \\ -\frac{2}{(n_{i,r} - ik_{i,r})^3} (x_r - x_{r-1}), & m = 0, \end{cases} \quad (5.30)$$

$$\left(\frac{\partial \mathbf{P}_i}{\partial k_{i,r}} \right)_{nm} = \begin{cases} -\frac{1}{\pi(n-m)} \frac{1}{(n_{i,r} - ik_{i,r})^3} \\ \cdot (\exp[-i\frac{2\pi}{\lambda} m x_r] - \exp[-i\frac{2\pi}{\lambda} (n-m)x_{r-1}]), & m \neq 0, \\ \frac{2i}{(n_{i,r} - ik_{i,r})^3} (x_r - x_{r-1}), & m = 0. \end{cases} \quad (5.31)$$

To compute the derivatives of the reflected and transmitted field amplitudes, we can substitute these derivatives in (5.16) and (5.18) and proceed similar as when a block width is under consideration.

Another parameter that can be considered is the angle of incidence. This can be the azimuthal angle ϕ or the polar angle θ . These angles only occur at matrices \mathbf{K}_x and \mathbf{K}_y inside the layers, but also in \mathbf{Y}_I , \mathbf{Y}_{II} , \mathbf{Z}_I and \mathbf{Z}_{II} . This implies that when such an angle is under consideration, the matrix of which the eigenvalues have to be computed will change for every layer, including layers 1 and $K+2$. Let us first consider the derivatives of matrices \mathbf{K}_x and \mathbf{K}_y with respect to an angle of incidence. Since matrix \mathbf{K}_x consists of entries $k_{x,n}$

and matrix \mathbf{K}_y of entries k_y , let us focus on those entries:

$$k_{x,n} = n_l \sin \theta \cos \phi - n \frac{\lambda}{\Lambda}, \quad (5.32a)$$

$$\frac{\partial k_{x,n}}{\partial \theta} = n_l \cos \theta \cos \phi, \quad (5.32b)$$

$$\frac{\partial k_{x,n}}{\partial \phi} = -n_l \sin \theta \sin \phi, \quad (5.32c)$$

$$k_y = n_l \sin \theta \sin \phi, \quad (5.32d)$$

$$\frac{\partial k_y}{\partial \theta} = n_l \cos \theta \sin \phi, \quad (5.32e)$$

$$\frac{\partial k_y}{\partial \phi} = -n_l \sin \theta \cos \phi. \quad (5.32f)$$

The angle of incidence also plays a role above and below the grating profile, namely in the matrices \mathbf{Y}_I , \mathbf{Y}_{II} , \mathbf{Z}_I and \mathbf{Z}_{II} , which contain $k_{I,zn}$ or $k_{II,zn}$:

$$k_{m,zn} = \sqrt{k_0^2 n_m^2 - k_{x,n}^2 - k_y^2}, \quad (5.33a)$$

$$\frac{\partial k_{m,zn}}{\partial \theta} = -\frac{k_{x,n} \frac{\partial k_{x,n}}{\partial \theta} + k_y \frac{\partial k_y}{\partial \theta}}{k_{m,zn}} = -\frac{n_l \cos \theta (k_{x,n} \cos \phi + k_y \sin \phi)}{k_{m,zn}}, \quad (5.33b)$$

$$\frac{\partial k_{m,zn}}{\partial \phi} = -\frac{k_{x,n} \frac{\partial k_{x,n}}{\partial \phi} + k_y \frac{\partial k_y}{\partial \phi}}{k_{m,zn}} = \frac{n_l \cos \theta (k_{x,n} \sin \phi + k_y \cos \phi)}{k_{m,zn}}. \quad (5.33c)$$

A difficulty arises when a propagating order becomes evanescent. Since the term of which the square root is taken in the definition of $k_{m,zn}$ (5.33a) can be either positive or negative, the derivative of $k_{m,zn}$ does not exist for those values of θ and ϕ for which $k_{m,zn}$ is zero. These values indicate the transition of a propagating order to an evanescent order and vice versa. In Section 2.5 we showed when an order is propagating or evanescent for a specific value of θ and ϕ . Similarly, we can find values of θ as a function of ϕ for which $k_{m,zn}$ is zero, viz.

$$\sin \theta = \frac{n \lambda \cos \phi \pm \sqrt{-n^2 \lambda^2 \sin^2 \phi + \Lambda^2 n_m^2}}{\Lambda n_l}. \quad (5.34)$$

This relation gives θ as a function of ϕ for different orders. Figure 5.3 gives an example of those combinations of θ and ϕ for which (5.34) holds when we consider test case 7, the symmetric trapezoid on two homogeneous layers. As can be seen, the diffraction efficiencies have several discontinuities and several of them can be found directly from (5.34). There are more discontinuities, but these are caused by the grating profile and the used materials. We should be careful not to be too close to such a line when we differentiate.

Also derivatives with respect to the wavelength of the field or the period of the grating can be considered, but in general these parameters are given and will therefore not be considered here.

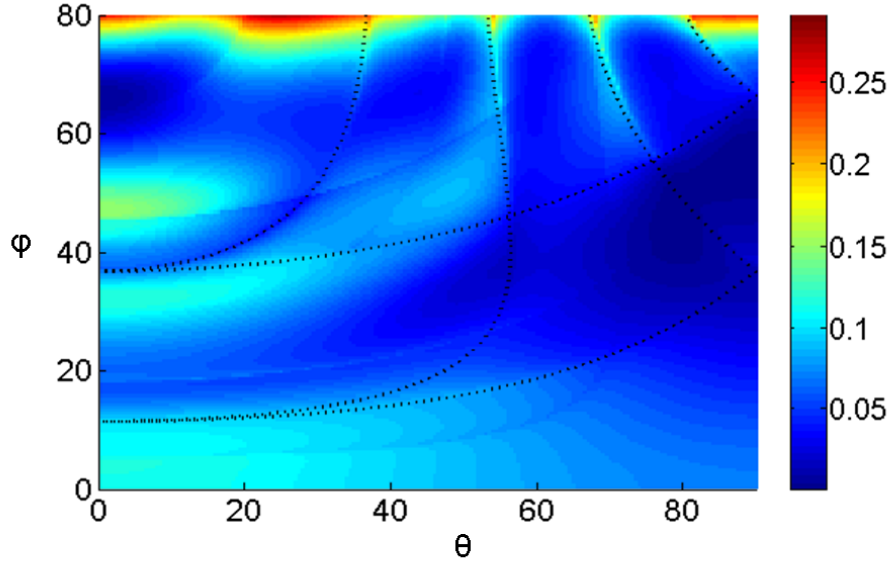


Figure 5.3: Illustration of the diffraction efficiency computed for test case 7 of Table 5.1 for polar angle $0 \leq \theta \leq 80^\circ$ and $0 \leq \phi \leq 90^\circ$ and the combinations of θ and ϕ for which $k_{i,zn} = 0$.

5.3 Complexity of the algorithms

To understand the reason why the matrix-vector approach will be faster than the matrix-matrix approach or finite differences, the complexity of the algorithm has to be considered. A plain RCWA computation requires the computation of \mathbf{F}_2 and \mathbf{G}_2 before the linear system (3.72) can be solved. These two matrices are found by updating matrices \mathbf{A}_i , \mathbf{B}_i , \mathbf{F}_i and \mathbf{G}_i for every layer i ($i = 2, \dots, K + 1$) as in (3.69). Besides the multiplications and summations of matrices, every layer requires matrices \mathbf{W}_i , \mathbf{V}_i and \mathbf{X}_i which are the result of solving an eigensystem. Also their inverse matrices should be available. Both the matrix-matrix multiplications and solving the eigensystems ensures that the complexity of a plain RCWA evaluation takes $\mathcal{O}(KN^3)$ flops where K is the number of layers and N is the number of harmonics after the truncation of the series.

For the field amplitude derivatives we have to solve the linear system (5.2). To do so we have to compute the matrices $(\partial\mathbf{F}_2)/(\partial p)$ and $(\partial\mathbf{G}_2)/(\partial p)$. As indicated in Section 5.2.1, there are two ways of finding the field amplitude derivatives. The first one is the matrix-matrix approach where first the matrices $(\partial\mathbf{F}_2)/(\partial p)$ and $(\partial\mathbf{G}_2)/(\partial p)$ are computed

and next they are multiplied by the vector \mathbf{t}_2 which is already available from a plain RCWA evaluation. This approach requires matrix-matrix multiplications which require $\mathcal{O}(N^3)$ flops. The second approach (matrix-vector approach) makes use of the fact that we have to multiply by this vector \mathbf{t}_2 anyway and therefore we make use of matrix-vector products only. These matrix-vector products have a complexity of $\mathcal{O}(N^2)$.

Another feature from Section 5.2.1 is that finding the derivative with respect to a layer thickness essentially differs from finding the derivative with respect to a block width. The main reason for this is that the eigensystems to be solved in a plain RCWA evaluation do not depend on the layer thickness. This implies that for a layer thickness only the derivative of \mathbf{X}_i has to be computed. However, this computation requires no more than $\mathcal{O}(N)$ flops and therefore finding the derivative with respect to the layer thickness by using the matrix-vector approach has a complexity of $\mathcal{O}(K N^2)$.

Finding the derivative with respect to a block width requires the derivatives of \mathbf{W}_i , \mathbf{V}_i and \mathbf{X}_i and of their inverse. As will be discussed in Chapter 6, the eigenvector derivative computations still require matrix-matrix multiplications and therefore finding the derivative with respect to a block width, even with the matrix-vector approach, takes $\mathcal{O}(K N^3)$ flops. However, the factor C in $C K N^3$ will be reduced because solving eigensystems will not be carried out.

5.4 Comparison of analytical method vs. finite differences

The analytical method with the matrix-matrix and matrix-vector approach will be compared to finite differences on two different aspects. The first one is accuracy because a method that is not accurate is useless. The second one is speed, because this is the main reason why we considered the analytical approach in the first place. The accuracy of the analytical method vs. finite differences will be considered first in Section 5.4.1 and the speed of the method in Section 5.4.2.

5.4.1 Accuracy

In the comparison of the methods to compute first order derivatives, we do not have to make a distinction between the matrix-vector and matrix-matrix approach since the results for the diffraction efficiencies will be the same. Table 5.2 gives the results of the comparison between the diffraction efficiency derivatives with respect to all physical parameters obtained by both the analytical method and finite differences. Since the refractive index of the substrate is complex-valued for all test cases of Table 5.1, no diffraction efficiencies will be present for the transmitted orders. For the chosen wavelength ($\lambda = 0.7\lambda$) and the angle of incidence we have three propagating orders, which are the -1 st, 0 th and 1 st diffraction

order. We have used multiple step sizes δ in finite differences. The optimal step size for finite differences equals the square root of the machine precision ($\approx 10^{-16}$) [18], but to show the effect of other step sizes we included $\delta = 10^{-4}\Lambda$ and $10^{-6}\Lambda$. We see that the difference in accuracy between the two methods for a fixed number of layers to approximate the grating structure can be found in the 5th or 6th decimal when δ approximates the square root of the machine precision. Since the analytical method does not make any additional approximations besides the approximations already made by the RCWA algorithm itself, Table 5.2 actually validates how well finite differences performs. To conclude the accuracy comparison, we can state that the accuracy is no issue for both methods as long as we take the step size small enough for finite differences.

An important warning is at place here. Finite differences evaluate RCWA once more for every shape parameter and it is crucial that for both evaluations the number of layers to approximate the grating structure remain the same. The analytical method computes the derivatives of a layered structure instead of the actual structure. When we vary the number of layers in finite differences, we are also dealing with the approximation errors of the RCWA algorithm itself. The effects are serious. When we consider test case 3: the symmetric trapezoid made of resist, we see that using finite differences with 10 layers for the first evaluation and 11 for the second one gives a maximum error of 10^3 compared to the derivative of the analytical method. If we had used finite differences with a fixed number of layers 10 or 11, this error was of the order 10^{-5} .

test case	Maximum absolute error		
	$\delta = 10^{-4}\Lambda$	$\delta = 10^{-6}\Lambda$	$\delta = 10^{-8}\Lambda$
1	1.0187×10^{-3}	1.0190×10^{-5}	1.3371×10^{-6}
2	7.5013×10^{-3}	7.5007×10^{-5}	1.2534×10^{-6}
3	1.0400×10^{-3}	1.0403×10^{-5}	1.6798×10^{-6}
4	2.7523×10^{-3}	2.7526×10^{-5}	1.0886×10^{-6}
5	1.0303×10^{-3}	1.0305×10^{-5}	1.0730×10^{-6}
6	1.0792×10^{-3}	1.0790×10^{-5}	1.3760×10^{-6}
7	2.9156×10^{-3}	2.9146×10^{-5}	4.9997×10^{-7}
8	5.4494×10^{-3}	5.4458×10^{-5}	5.1701×10^{-7}
9	1.8953×10^{-3}	1.8910×10^{-5}	1.5488×10^{-6}
10	3.7740×10^{-3}	3.7748×10^{-5}	8.2360×10^{-7}

Table 5.2: Maximum absolute error between the derivatives of the diffraction efficiencies of all propagating orders obtained by the analytical method and those obtained by finite differences for various step sizes δ . These derivatives are computed for planar diffraction TM polarization with polar angle $\theta = 10^\circ$, wavelength $\lambda = 0.7\Lambda$ and $N = 25$.

5.4.2 Speed

To compare the computational speed of the two methods we measure three different times: the time one RCWA evaluation takes, the time the analytical approach takes to compute both the diffraction efficiencies of the reflected field and their derivatives with respect to all shape parameters and the time finite differences take to compute the same. These times are denoted by t_{RCWA} , $t_{\text{analytical}}$ and t_{FD} , respectively. To compare the speed of the analytical method to finite differences, we define a factor f indicating the ratio between the time the analytical method takes to compute the derivatives and the time finite differences take to do the same. In mathematical terms, the factor is defined by

$$f := \frac{t_{\text{analytical}} - t_{\text{RCWA}}}{t_{\text{FD}} - t_{\text{RCWA}}}. \quad (5.35)$$

The computational time depends on the number of harmonics taken into account. For a binary grating given by a width and a height, Figure 5.4 gives this factor as a function of the number of harmonics for both the matrix-matrix and matrix-vector approach as presented in Section 5.2.1. Although the matrix-vector approach is the absolute winner, the matrix-matrix approach is also faster than finite differences. For an increasing number of harmonics we see that the matrix-vector approach becomes increasingly faster, while the matrix-matrix approach becomes slower with respect to the finite difference approach. The main reason of the decreasing factor f as a function of the number of harmonics for the

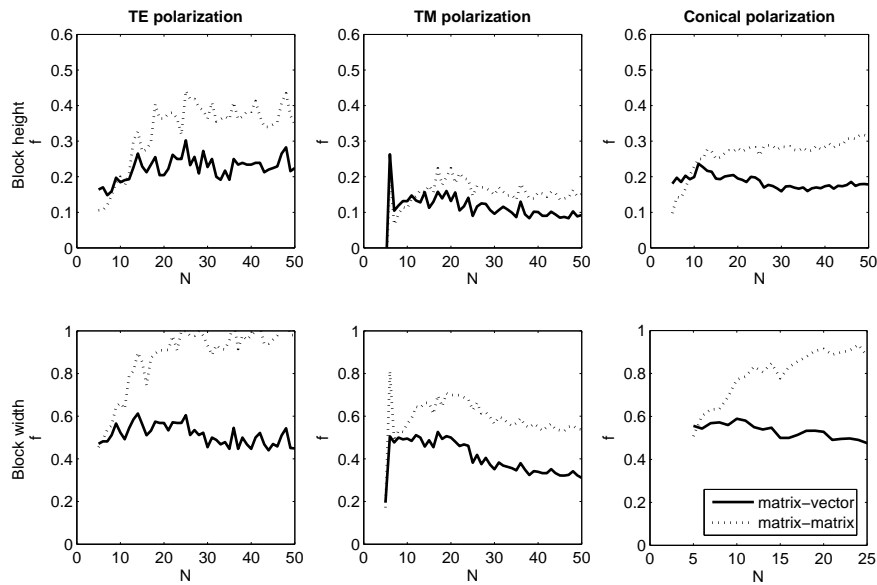


Figure 5.4: The factor f as a function of the number of harmonics (N) for a binary grating using both the matrix-matrix and matrix-vector approach for planar (TE and TM) and conical diffraction.

matrix-vector approach is the complexity $\mathcal{O}(K N^2)$ to find the right-hand-side of (5.2) when we discard the computations of the eigenvalue and eigenvector derivatives. This agrees with the complexity analysis of the previous section.

The binary grating is approximated geometrically by one layer only and is the only grating where the discretization does not introduce additional errors. To see how the number of layers influences the speed, we have to look at more complex structures such as the other eight test cases of Table 5.1. Before the diffraction efficiencies can be computed by RCWA a conversion has to be made from the physical shape parameters to RCWA shape parameters as has been done in Section 5.2.2.

In Figure 5.5 the factor f as defined in (5.35) has been plotted for the symmetric trapezoid as a function of the number of harmonics for the analytical method using both the matrix-matrix and the matrix-vector approach. To consider the influence of the number of layers RCWA uses to approximate the trapezoid, Figure 5.5 gives the results in case the trapezoid is sliced up in 10 and 25 layers, respectively. As can be seen from Section 5.3 and Figure 5.5 the number of layers does not influence f , since the computations are linear in the number of layers. Also, a similar trend like in the binary grating can be seen for the trapezoid.

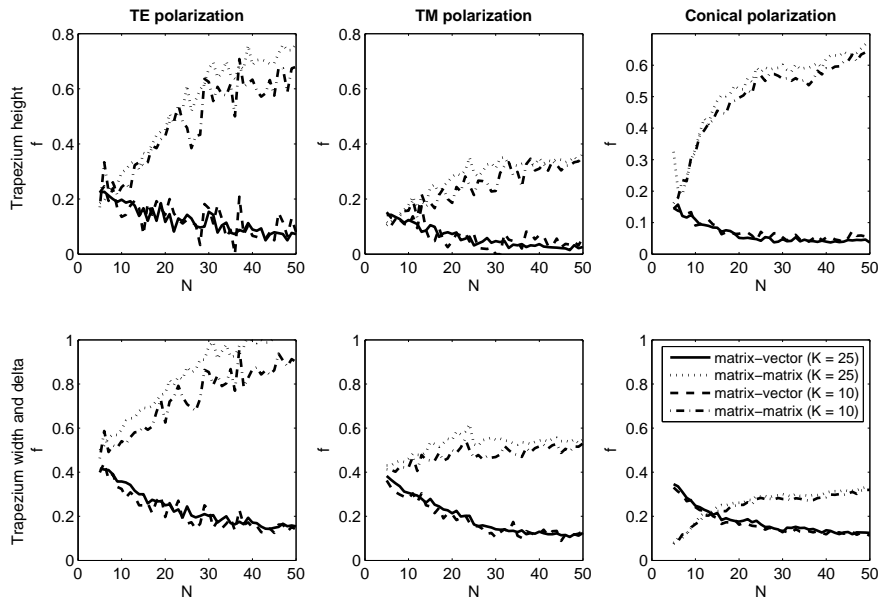


Figure 5.5: The factor f as a function of the number of harmonics (N) for a symmetric trapezoidal grating, which is approximated by 10 and 25 layers. Both the matrix-matrix and matrix-vector approach has been used for planar (TE and TM) and conical diffraction.

In the reconstruction of the shape parameters as will be discussed in Chapter 7 derivatives of the diffraction efficiencies with respect to all shape parameters have to be computed simultaneously. Since the matrix-vector approach is faster than the matrix-matrix approach, we will only consider the speed results of the matrix-vector approach. When we consider the factor f for all test cases defined in Table 5.1 as a function of the number of harmonics N , we get the results of Figure 5.6. From the test cases, we can make several observations, namely

- The more shape parameters are required to describe the grating structure, the smaller the factor will be. This implies that the analytical method becomes relatively faster than finite differences when more parameters are under consideration. The reason for this is that all derivatives with respect to the RCWA shape parameters are computed anyway, but they are distributed over more parameters. For finite differences this implies that more RCWA evaluations have to be carried out. Notice that for the test cases 9 and 10 of Table 5.1 the derivatives with respect to the physical parameters are computed by dividing the highest trapezoid into two trapezoids. Although finite differences computes only 6 derivatives, the analytical method has to compute 8 derivatives as is shown in Section 5.2.2. These derivatives are translated afterwards to the 6 derivatives with respect to the actual physical parameters. Because of this the factor f for test cases 9 and 10 will be higher than expected.
- The more harmonics are taken into account, the more gain, or lower factor f , we have by using the matrix-vector approach compared to finite differences. This is because the eigensystems grow larger and become more and more the dominant part of the computation.
- The factor f will be lower (and thus the analytical method faster than finite differences) for a structure with a complex-valued refractive index than for a structure with a real-valued index. Although the differences are minimal in general, this difference is larger for planar diffraction with TE polarization than for the other diffraction types. This is because the eigensystem for TE polarization has a special structure. The matrix is Hermitian if the refractive indices inside a layer are real-valued, which speeds up the computation of the eigenvalues and eigenvectors. The amount of work of the analytical method stays the same.
- For conical diffraction the factor f will be smaller than for planar diffraction. This is due to the fact that for conical diffraction the eigensystems are twice as large as for planar diffraction.

To conclude we see that using analytical over numerical derivatives of the field amplitudes results in a considerable reduction in computational time, no matter what structure one considers. To quantify the reduction we can say that for simple structures as in test cases 1 to 7 of Table 5.1 and $N = 20$ we have

$$f \approx \frac{0.11 \times (\# \text{ parameters heights}) + 0.28 \times (\# \text{ parameters other})}{\# \text{ total parameters}}, \quad (5.36)$$

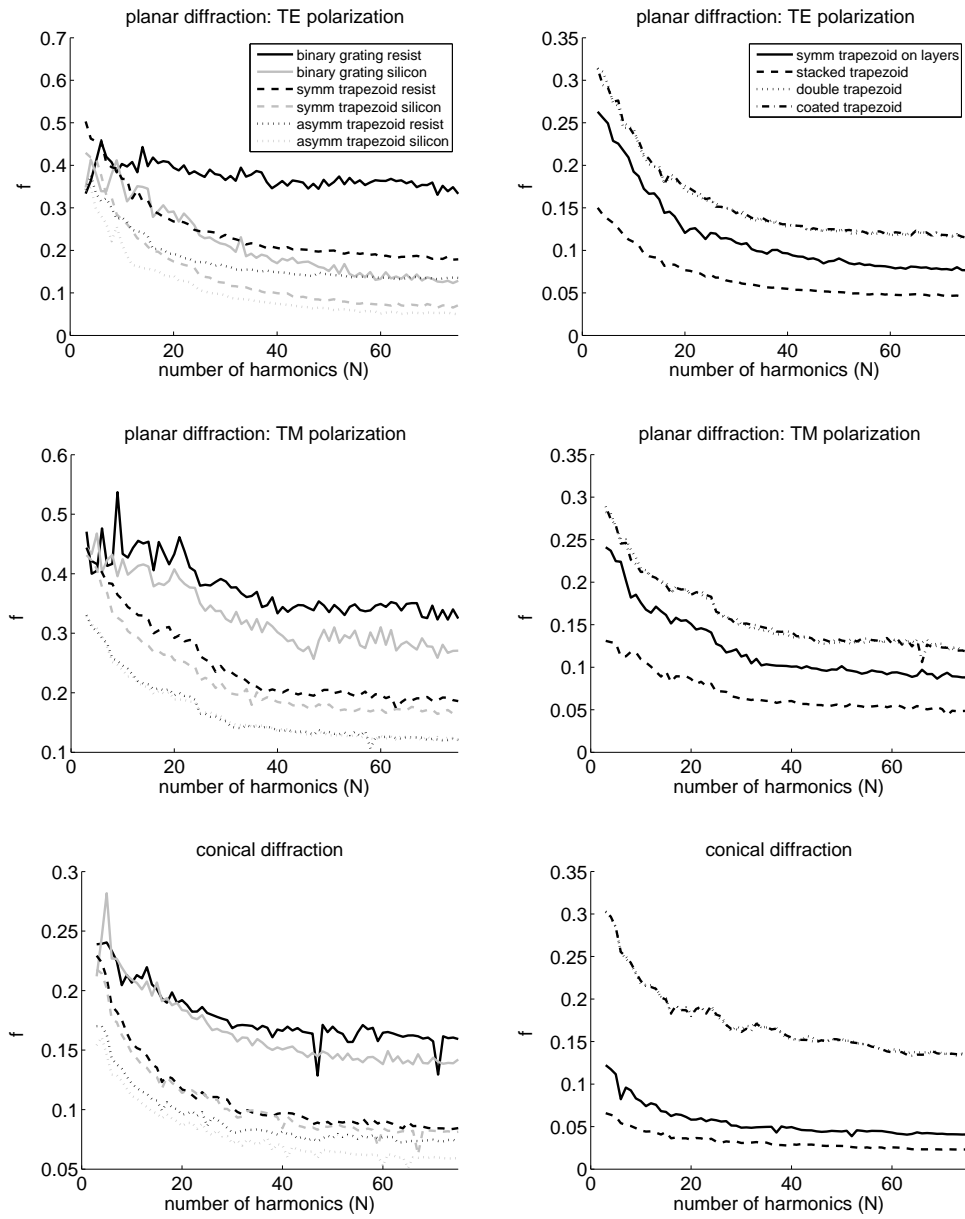


Figure 5.6: Time comparison for planar and conical diffraction of the simple structures (*left*) and the complex structures (*right*), where the computation time to compute the derivatives with the analytical method is compared to the computation time finite differences takes to compute the same.

where "# parameters heights" are those physical parameters that only depend on layer thicknesses and "# parameters other" are those physical parameters that depend also on block widths. For example, for the symmetric trapezoid, the height h depends only on the layer thicknesses, while the parameters X and Δ depend on the block widths. For the more complex structures such as those defined in test cases 8 to 10, we have a coupling between the structures. For example, the stacked grating consists of three trapezoidal structures on top of each other. Because of this coupling, the actual factor f will be lower than the estimate given by (5.36). Note that the number of layers do not affect the factor f significantly as is shown in Figure 5.5.

The matrix-vector approach of the analytical method will be used in Chapter 7, where the shape parameters will be reconstructed from a measured signal.

Chapter 6

Eigenvalue and eigenvector sensitivity

Eigenvalue and eigenvector derivatives are crucial for the analytical method discussed in the previous chapter. In this chapter we will discuss how they can be computed. Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ be a *non-defective* matrix given as a function of a certain parameter p . The non-defectiveness means that \mathbf{A} has N linearly independent eigenvectors. Let $\mathbf{\Lambda} \in \mathbb{C}^{N \times N}$ be the eigenvalue matrix of \mathbf{A} and $\mathbf{W} \in \mathbb{C}^{N \times N}$ a corresponding eigenvector matrix of \mathbf{A} , i.e.

$$\mathbf{A}(p)\mathbf{W}(p) = \mathbf{W}(p)\mathbf{\Lambda}(p). \quad (6.1)$$

Our goal is to compute $\mathbf{\Lambda}'(p_0)$ and $\mathbf{W}'(p_0)$ for one specific value p_0 of parameters p . In the remainder of this chapter, the p -dependence is discarded for notational convenience. It is well-known that eigenvectors are not uniquely determined. If the eigenvalues are distinct, the eigenvectors are determined up to a constant multiplier, whereas if an eigenvalue has geometric multiplicity higher than 1, any linear combination of the associated eigenvectors will be an eigenvector again. This implies that an eigenvector derivative cannot be computed uniquely as long as the eigenvectors are not fixed.

If all eigenvalues are *distinct*, one of the first methods to compute eigenvector derivatives analytically for *real-valued* \mathbf{A} at $p = p_0$ can be found in [40]. A numerical way of retrieving the eigenvector derivatives is given in [6, 47]. This numerical method is fast when only a small number of eigenvalue derivatives and their corresponding eigenvector derivatives have to be computed. If all eigenvalues and eigenvectors of \mathbf{A} are already available and all corresponding derivatives should be computed, an analytical approach will be much faster. For an overview of the methods available for a complex-valued matrix with distinct eigenvalues see [39].

Repeated eigenvalues have also been analyzed before, though not for an arbitrary complex-

valued matrix. For real-valued matrices, the eigenvector derivatives for a symmetric real-valued matrix can be found [34]. However, it is assumed that the eigenvalue derivatives are distinct. In [15] the theory is extended by assuming that there are still eigenvalue derivatives which are repeated, but the second-order derivatives of the eigenvalues have to be distinct. In [7] this theory is generalized to Hermitian matrices, where it is assumed that for an eigenvalue with multiplicity r , all its eigenvalue derivatives up to k th-order are all assumed to be equal and the $(k + 1)$ st-order derivatives are distinct again.

The generalization towards a *general* non-defective matrix has been provided in [4], which presents an algorithm to compute the first-order analytical derivatives of both eigenvalues and eigenvectors for such a general non-defective matrix. It is assumed that all eigenvalues and their associated eigenvectors are known either analytically or from a numerical procedure and that the eigenvector derivatives exist.

Before we start computing the derivatives, we will show that they exist indeed. Section 6.2 will discuss the simple cases known from the literature, Section 6.3 presents the generalization of the theory of computing eigenvalue and eigenvector derivatives and Section 6.4 will show examples of the theory presented in this chapter. Finally, the need of this theory for the RCWA algorithm will be discussed.

6.1 Existence of eigenvalue and eigenvector derivatives

Because the eigenvectors are not uniquely determined, they have to be fixed in some way, which we will call the *normalization* of the eigenvectors. Let for $k = 1, \dots, N$ the m th element of eigenvector \mathbf{w}_k be denoted by w_{mk} . In this thesis we will choose for every eigenvector m an element k equal to 1 for all values of p in a neighbourhood \mathcal{D} of $p = p_0$, viz.

$$w_{mk}(p) = 1, \quad k = 1, \dots, N, \quad 1 \leq m \leq N, \quad p \in \mathcal{D}. \quad (6.2)$$

To explain the choice of m and for the theory presented in this chapter, we need the *left-eigenvector matrix*, which is defined as

$$\mathbf{Y} := (\mathbf{W}^{-1})^*. \quad (6.3)$$

Note that $*$ denotes the complex conjugate transpose, that is $\mathbf{Y}^* = \mathbf{Y}^T$. When the m th element of the k th left-eigenvector is denoted by y_{mk} , the index m is chosen as follows

$$|w_{mk}||y_{mk}| = \max_{1 \leq j \leq N} |w_{jk}||y_{jk}|. \quad (6.4)$$

This prevents both w_{mk} and y_{mk} from becoming very small. In general, the choice of m will not be unique.

Let us first consider the existence and uniqueness of the eigenvalue and eigenvector derivatives for distinct eigenvalues only.

Theorem 6.1 Let $\mathbf{w}(p_0)$ be an eigenvector belonging to a distinct eigenvalue $\lambda(p_0)$ of $\mathbf{A}(p_0)$. Then for every p in a neighbourhood of p_0 there exist functions $\mathbf{w}(p)$ and $\lambda(p)$ such that they are continuously differentiable.

We will prove this theorem by using the *Implicit Function Theorem*, which is stated.

Theorem 6.2 (Implicit Function Theorem) Let \mathcal{D} be an open subset of \mathbb{C}^{N+1} , and suppose that the mapping $\mathbf{F} : \mathcal{D} \rightarrow \mathbb{C}^{N+1}$ is continuously differentiable. Suppose that $\mathbf{F}(\mathbf{z}_0, p_0) = \mathbf{0}$ and that the partial derivative matrix $(\partial\mathbf{F})(\partial p)(\mathbf{z}_0, p_0)$ is nonsingular at the point $(\mathbf{z}_0, p_0) \in \mathcal{D}$. Then there is a continuously differentiable mapping $\mathbf{z}(p)$ such that

$$\mathbf{F}(\mathbf{z}(p), p) = \mathbf{0} \text{ for all } p \in \mathcal{D}, \quad (6.5)$$

and

$$\frac{\partial\mathbf{z}(p_0)}{\partial p} = - \left[\frac{\partial\mathbf{F}}{\partial\mathbf{z}(p)}(\mathbf{z}(p_0), p_0) \right]^{-1} \frac{\partial\mathbf{F}}{\partial p}(\mathbf{z}(p_0), p_0). \quad (6.6)$$

The proof of this theorem can be found in [14]. Let us consider the proof of Theorem 6.1.

Proof. Proof of Theorem 6.1
Define a function \mathbf{F} as

$$\mathbf{F}(\mathbf{z}, p) := \begin{bmatrix} \mathbf{A}(p)\mathbf{w} - \lambda\mathbf{w} \\ \mathbf{u}^T\mathbf{w} - 1 \end{bmatrix} \text{ and } \mathbf{z} := \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix}. \quad (6.7)$$

It is clear that with $\mathbf{z} = (\mathbf{w}(p_0)^T, \lambda(p_0))^T$ it holds that

$$\mathbf{F}(\mathbf{z}, p_0) = \mathbf{0}. \quad (6.8)$$

The *Jacobian* $(\partial\mathbf{F}/\partial\mathbf{z})(\mathbf{z}(p), p)$ is given by

$$\frac{\partial\mathbf{F}}{\partial\mathbf{z}}(\mathbf{z}, p) = \begin{bmatrix} (\mathbf{A}(p) - \lambda\mathbf{I}) & \mathbf{w} \\ \mathbf{u}^T & 0 \end{bmatrix}. \quad (6.9)$$

To prove that the Jacobian is nonsingular for $p = p_0$, let us assume that it is singular, thus

$$\begin{bmatrix} (\mathbf{A}(p_0) - \lambda\mathbf{I}) & \mathbf{w} \\ \mathbf{u}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 0 \end{bmatrix}. \quad (6.10)$$

Since $(\mathbf{A}(p_0) - \lambda\mathbf{I})\mathbf{w} + \mathbf{w}\lambda = \mathbf{0}$ it holds that $\mathbf{A}(p_0)\mathbf{w} = \mathbf{0}$ and thus $\mathbf{w} \in \mathcal{N}(\mathbf{A}(p_0))$. The second equation provides

$$\mathbf{u}^T\mathbf{w} = 0 \quad (6.11)$$

Since $\mathbf{w} \neq \mathbf{0}$ we have a contradiction, which shows that the Jacobian is nonsingular.

The Implicit Function Theorem gives us the existence and uniqueness of the eigenvalue and

eigenvector derivatives for distinct eigenvalues. Because $\mathbf{F}(\mathbf{z}_0, p_0) = \mathbf{0}$ and $(\partial\mathbf{F}/(\partial\mathbf{z}))(\mathbf{z}_0, p_0)$ is nonsingular, this theorem says that there exists a unique continuously differentiable $\mathbf{z}(p)$ such that $\mathbf{F}(\mathbf{z}(p), p) = \mathbf{0}$ and $(\partial\mathbf{F}/\partial\mathbf{z})(\mathbf{z}(p), p)$ is nonsingular. The derivative of $\mathbf{z}(p)$ can be found by

$$\frac{\partial\mathbf{z}(p_0)}{\partial p} = - \left[\frac{\partial\mathbf{F}}{\partial\mathbf{z}(p)}(\mathbf{z}(p_0), p_0) \right]^{-1} \frac{\partial\mathbf{F}}{\partial p}(\mathbf{z}(p_0), p_0), \quad (6.12)$$

which completes the existence and uniqueness of the eigenvalue and eigenvector derivatives for distinct eigenvalues. \square

The generalization of the existence of eigenvalue and eigenvector derivatives for repeated eigenvalues are given in [21]. In the following sections we will focus on the mathematical method to actually compute the eigenvalue and eigenvector derivatives.

6.2 Simple cases

Before rushing into the generalization, let us consider the basic features of computing eigenvalue and eigenvector derivatives in the context of some simple cases. Let the non-defective matrix \mathbf{A} and its eigenvalue and eigenvector matrix $\mathbf{\Lambda}$ and \mathbf{W} be differentiable in a neighbourhood of $p = p_0$. Differentiate the eigensystem (6.1) directly with respect to p . The derivative with respect to p is denoted by a prime. So from (6.1) we have

$$\mathbf{A}'\mathbf{W} - \mathbf{W}\mathbf{\Lambda}' = -\mathbf{A}\mathbf{W}' + \mathbf{W}'\mathbf{\Lambda}. \quad (6.13)$$

In (6.13) both the eigenvalue derivative matrix $\mathbf{\Lambda}'$ and the eigenvector derivative matrix \mathbf{W}' occur. To find an expression for $\mathbf{\Lambda}'$, premultiply (6.13) by the left-eigenvector matrix as defined in (6.3), viz.

$$\mathbf{Y}^*\mathbf{A}'\mathbf{W} - \mathbf{\Lambda}' = -\mathbf{Y}^*\mathbf{A}\mathbf{W}' + \mathbf{Y}^*\mathbf{W}'\mathbf{\Lambda}. \quad (6.14)$$

Since \mathbf{A} is assumed to be non-defective, the eigenvectors span the complete $\mathbb{C}^{N \times N}$. For ease of notation let us define a matrix $\mathbf{C} := \mathbf{W}^{-1}\mathbf{W}'$, i.e.

$$\mathbf{W}' = \mathbf{W}\mathbf{C}. \quad (6.15)$$

Rather than determining \mathbf{W}' , we will try to find \mathbf{C} . From (6.14) we derive

$$\mathbf{Y}^*\mathbf{A}'\mathbf{W} - \mathbf{\Lambda}' = -\mathbf{\Lambda}\mathbf{C} + \mathbf{C}\mathbf{\Lambda}. \quad (6.16)$$

In Section 6.1 we showed that there exists a continuously differentiable eigenvector matrix, but the eigenvector matrix is computed for a fixed value of p and therefore the continuous differentiability is not guaranteed. This is shown by the following example.

Example 6.3 Let

$$\mathbf{A}(\rho) := \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \quad (6.17)$$

Then the eigenvalue matrix $\mathbf{\Lambda}(\rho)$ and an eigenvector matrix $\mathbf{W}(\rho)$ can be found as

$$\mathbf{\Lambda}(\rho) = \begin{bmatrix} 1-\rho & 0 \\ 0 & 1+\rho \end{bmatrix}, \quad \mathbf{W}(\rho) = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \quad (6.18)$$

respectively. For $\rho = 0$, the eigenvalues become repeated and a valid eigenvector matrix would be

$$\mathbf{W}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (6.19)$$

Note that for $\rho = 0$ the right-hand side of (6.16) vanishes completely and therefore

$$\mathbf{\Lambda}'(0) = \mathbf{Y}^*(0)\mathbf{A}'(0)\mathbf{W}(0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (6.20)$$

but $\mathbf{\Lambda}'(0)$ should be a diagonal matrix. Of course, the problem is caused by a wrong choice of \mathbf{W} . \square

The main focus of this and the following section is choosing an eigenvector matrix such that $\mathbf{\Lambda}'$ is continuous with respect to parameter ρ . To do this let us trivially assume that an eigenvector matrix $\tilde{\mathbf{W}}$ has already been found, e.g. by Lapack [5]. Then there exists a non-unique matrix $\mathbf{\Gamma} \in \mathbb{C}^{N \times N}$ such that

$$\mathbf{W} =: \tilde{\mathbf{W}}\mathbf{\Gamma}, \quad (6.21)$$

with $\mathbf{\Gamma} \in \mathbb{C}^{N \times N}$. Since \mathbf{W} should also be an eigenvector matrix of \mathbf{A} it has to satisfy (6.1) and therefore $\mathbf{\Gamma}$ has to satisfy the following condition

$$\mathbf{\Lambda}\mathbf{\Gamma} = \mathbf{\Gamma}\mathbf{\Lambda}. \quad (6.22)$$

Substitution of (6.21) into (6.16) results in

$$\tilde{\mathbf{Y}}^*\mathbf{A}'\tilde{\mathbf{W}}\mathbf{\Gamma} - \mathbf{\Gamma}\mathbf{\Lambda}' = -\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{C} + \mathbf{\Gamma}\mathbf{C}\mathbf{\Lambda}. \quad (6.23)$$

To find the eigenvector derivative matrix, we first have to determine $\mathbf{\Gamma}$. Relation (6.23) will be the starting point for the computation of eigenvalue and eigenvector derivatives. Since the derivation of a general formulation is rather involved, we start with the three simplest situations: distinct eigenvalues, repeated eigenvalues with distinct derivatives and, finally, repeated eigenvalues with repeated first-order derivatives and distinct second-order derivatives.

6.2.1 Distinct eigenvalues

For distinct eigenvalues, it follows from (6.22) that $\mathbf{\Gamma}$ must be a diagonal matrix. If (6.23) is considered element by element, we have

$$\tilde{\mathbf{y}}_k^* \mathbf{A}' \tilde{\mathbf{w}}_l \gamma_l - \delta_{kl} \gamma_l \lambda_l' = \gamma_k (\lambda_l - \lambda_k) c_{kl}, \quad (6.24)$$

where λ_k is the k th element on the diagonal of \mathbf{A} , $\tilde{\mathbf{y}}_k^*$ and $\tilde{\mathbf{w}}_k$ are the k th left- and right-eigenvector, respectively, c_{kl} is the entry (k, l) of matrix \mathbf{C} and $\delta_{kl} = 1$ if $k = l$ and 0 otherwise.

$$\lambda_k' = \tilde{\mathbf{y}}_k^* \mathbf{A}' \tilde{\mathbf{w}}_k, \quad k = 1, \dots, N, \quad (6.25)$$

$$c_{kl} = \frac{\tilde{\mathbf{y}}_k^* \mathbf{A}' \tilde{\mathbf{w}}_l \gamma_l}{\gamma_k (\lambda_l - \lambda_k)}, \quad k, l = 1, \dots, N, \quad k \neq l. \quad (6.26)$$

Note that $\gamma_k \neq 0$ has been used in (6.25), which is justified since $\mathbf{\Gamma}$ is a regular matrix. For distinct eigenvalues all eigenvalue derivatives can be found and the off-diagonal entries of matrix \mathbf{C} are determined when the entries of $\mathbf{\Gamma}$ are known. There are no relations for the diagonal entries of $\mathbf{\Gamma}$ and therefore they can be chosen arbitrary. This corresponds to the fact that eigenvectors are unique up to a multiplicative constant for distinct eigenvalues, i.e. these constants can be chosen freely. As mentioned in the previous section, we will use (6.2) to choose the diagonal entries of $\mathbf{\Gamma}$. For all $k = 1, \dots, N$ there is a m chosen by using condition (6.2) such that

$$w_{mk} = 1 \Rightarrow \tilde{w}_{mk} \gamma_k = 1 \Rightarrow \gamma_k = \frac{1}{\tilde{w}_{mk}}. \quad (6.27)$$

With this normalization condition there is always an m such that eigenvector \mathbf{w}_k is uniquely determined, as $\mathbf{w}_k \neq \mathbf{0}$. At this point the eigenvectors are uniquely determined and since the normalization condition (6.27) applies to each eigenvector and all values of p , the following holds

$$w'_{mk}(p) = 0 \quad \text{for all } p. \quad (6.28)$$

By using (6.15) an expression can be found for the coefficient c_{kk} in terms of the off-diagonal entries of the matrix \mathbf{C} .

$$w'_{mk} = \sum_{l=1}^N w_{ml} c_{lk} = 0 \quad \Rightarrow \quad c_{kk} = -\frac{1}{w_{mk}} \sum_{\substack{l=1 \\ l \neq k}}^N w_{ml} c_{lk} = -\sum_{\substack{l=1 \\ l \neq k}}^N w_{ml} c_{lk}. \quad (6.29)$$

In terms of the given eigenvectors this implies that

$$c_{kk} = -\sum_{\substack{l=1 \\ l \neq k}}^N \tilde{w}_{ml} \gamma_l c_{lk}. \quad (6.30)$$

Thus also the diagonal coefficients c_{kk} can be determined uniquely and the matrix \mathbf{W}' can be computed by (6.15).

6.2.2 Repeated eigenvalues with distinct eigenvalue derivatives

For repeated eigenvalues, condition (6.22) indicates that matrix $\mathbf{\Gamma}$ does not need to be diagonal. Every linear combination of such eigenvectors is still an eigenvector. Let now e.g. $\lambda_1 = \dots = \lambda_M$ for some $M \leq N$. Introduce the following partitioning of $\mathbf{\Lambda}$ and of the corresponding matrices $\tilde{\mathbf{W}}$, $\tilde{\mathbf{Y}}^*$ and $\mathbf{\Gamma}$:

$$\begin{aligned} \mathbf{\Lambda} &= \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix}, \quad \tilde{\mathbf{W}} = [\tilde{\mathbf{W}}_1 \quad \tilde{\mathbf{W}}_2], \quad \tilde{\mathbf{Y}}^* = \begin{bmatrix} \tilde{\mathbf{Y}}_1^* \\ \tilde{\mathbf{Y}}_2^* \end{bmatrix}, \\ \mathbf{\Gamma} &= \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}, \end{aligned} \quad (6.31)$$

where $\mathbf{\Lambda}_2 = \lambda \mathbf{I} \in \mathbb{C}^{M \times M}$ consists of all M repeated eigenvalues. The other eigenvalues are in $\mathbf{\Lambda}_1 \in \mathbb{C}^{(N-M) \times (N-M)}$. Note that $\mathbf{\Lambda}_1$ can still contain repeated eigenvalues, but none of the eigenvalues of this matrix is equal to λ . The block $\tilde{\mathbf{W}}_2 \in \mathbb{C}^{N \times M}$ consists of the right-eigenvectors belonging to the repeated eigenvalues and $\tilde{\mathbf{W}}_1 \in \mathbb{C}^{N \times (N-M)}$ consists of the right-eigenvectors belonging to the other eigenvalues. Likewise the matrix $\tilde{\mathbf{Y}}^*$ is divided into two block rows $\tilde{\mathbf{Y}}_1^* \in \mathbb{C}^{(N-M) \times N}$ and $\tilde{\mathbf{Y}}_2^* \in \mathbb{C}^{M \times N}$ respectively. The matrices $\mathbf{\Gamma}$ and \mathbf{C} are partitioned in a similar way, but note that (6.22) implies that the off-diagonal blocks of $\mathbf{\Gamma}$ are $\mathbf{0}$. As a result (6.23) can be described by a set of four equations:

$$\tilde{\mathbf{Y}}_1^* \mathbf{A}' \tilde{\mathbf{W}}_1 \mathbf{\Gamma}_1 - \mathbf{\Gamma}_1 \mathbf{\Lambda}'_1 = -\mathbf{\Gamma}_1 \mathbf{\Lambda}_1 \mathbf{C}_{11} + \mathbf{\Gamma}_1 \mathbf{C}_{11} \mathbf{\Lambda}_1, \quad (6.32a)$$

$$\tilde{\mathbf{Y}}_1^* \mathbf{A}' \tilde{\mathbf{W}}_2 \mathbf{\Gamma}_2 = \mathbf{\Gamma}_1 (\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{12}, \quad (6.32b)$$

$$\tilde{\mathbf{Y}}_2^* \mathbf{A}' \tilde{\mathbf{W}}_1 \mathbf{\Gamma}_1 = -\mathbf{\Gamma}_2 \mathbf{C}_{21} (\lambda \mathbf{I} - \mathbf{\Lambda}_1), \quad (6.32c)$$

$$\tilde{\mathbf{Y}}_2^* \mathbf{A}' \tilde{\mathbf{W}}_2 \mathbf{\Gamma}_2 - \mathbf{\Gamma}_2 \mathbf{\Lambda}'_2 = \mathbf{0}. \quad (6.32d)$$

Clearly (6.32d) is an eigenvalue problem itself, $\mathbf{\Lambda}'_2$ containing the eigenvalues of $\tilde{\mathbf{Y}}_2^* \mathbf{A}' \tilde{\mathbf{W}}_2$ and $\mathbf{\Gamma}_2$ being the eigenvector matrix. Assuming that all eigenvalues of $\mathbf{\Lambda}'_2$ are distinct, the column vectors of $\mathbf{\Gamma}_2$ are determined up to a constant multiplier. Since $\tilde{\mathbf{\Gamma}}_2$ can be computed by e.g. Lapack [5], we have $\mathbf{\Gamma}_2 = \tilde{\mathbf{\Gamma}}_2 \mathbf{\Psi}$, with $\mathbf{\Psi}$ a diagonal matrix. Now, the choice (6.2) gives an expression for the diagonal entries of $\mathbf{\Gamma}_2$:

$$w_{mk} = 1 \Rightarrow \left(\sum_{l=1}^M \tilde{w}_{ml} \gamma_{lk} \right) \psi_k = 1 \Rightarrow \psi_k = \frac{1}{\sum_{l=1}^M \tilde{w}_{ml} \gamma_{lk}}. \quad (6.33)$$

If $\mathbf{\Lambda}_1$ still contains repeated eigenvalues, the same procedure can be performed until $\mathbf{\Gamma}_1$ has been determined uniquely.

Since $\mathbf{\Gamma}$ is known, the continuously differentiable eigenvectors are determined uniquely and the eigenvector derivatives can be computed. Matrices \mathbf{C}_{12} and \mathbf{C}_{21} can be found from (6.32b) and (6.32c) and by using $\mathbf{Y}_m^* = \mathbf{\Gamma}_m^{-1} \tilde{\mathbf{Y}}_m^*$ for $m = 1, 2$, viz.

$$\mathbf{C}_{12} = (\lambda \mathbf{I} - \mathbf{\Lambda}_1)^{-1} \mathbf{Y}_1^* \mathbf{A}' \mathbf{W}_2, \quad (6.34a)$$

$$\mathbf{C}_{21} = -\mathbf{Y}_2^* \mathbf{A}' \mathbf{W}_1 (\lambda \mathbf{I} - \mathbf{\Lambda}_1)^{-1}. \quad (6.34b)$$

Since the eigenvalues of $\mathbf{\Lambda}_2$ are repeated, but their derivatives are not, (6.13) needs to be differentiated once more to determine \mathbf{C}_{22} . When the result is premultiplied by $\tilde{\mathbf{Y}}^*$, (6.21) and (6.15) are used for \mathbf{W} and \mathbf{W}' , respectively, and matrix \mathbf{D} is introduced to write \mathbf{W}'' in a similar way as in (6.15) such that

$$\mathbf{W}'' = \mathbf{W}\mathbf{D}, \quad (6.35)$$

we have

$$\mathbf{Y}_2^* \mathbf{A}'' \mathbf{W}_2 - \mathbf{\Lambda}_2'' = -2(\mathbf{\Lambda}_2' \mathbf{C}_{22} - \mathbf{C}_{22} \mathbf{\Lambda}_2') - 2\mathbf{Y}_2^* \mathbf{A}' \mathbf{W}_1 \mathbf{C}_{12}. \quad (6.36)$$

Then (6.36) gives an expression for the off-diagonal entries of \mathbf{C}_{22} , since \mathbf{C}_{12} is already available from (6.34a).

6.2.3 Repeated eigenvalue with repeated eigenvalue derivatives

If eigenvalue system (6.32d) has repeated eigenvalues too, which implies that some of the derivatives of the repeated eigenvalue under consideration are also equal, the column vectors of $\mathbf{\Gamma}_2$ are not unique either. As a consequence introduce a partitioning for $\mathbf{\Lambda}$, $\tilde{\mathbf{W}}$, $\tilde{\mathbf{Y}}^*$, $\mathbf{\Gamma}$ and \mathbf{C} :

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Lambda}}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{\Lambda}}_3 \end{bmatrix}, \quad \tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}}_1 & \hat{\mathbf{W}}_2 & \hat{\mathbf{W}}_3 \end{bmatrix}, \quad \tilde{\mathbf{Y}}^* = \begin{bmatrix} \tilde{\mathbf{Y}}_1^* \\ \hat{\mathbf{Y}}_2^* \\ \hat{\mathbf{Y}}_3^* \end{bmatrix}, \quad (6.37)$$

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Gamma}}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{\Gamma}}_3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \hat{\mathbf{C}}_{12} & \hat{\mathbf{C}}_{13} \\ \hat{\mathbf{C}}_{21} & \hat{\mathbf{C}}_{22} & \hat{\mathbf{C}}_{23} \\ \hat{\mathbf{C}}_{31} & \hat{\mathbf{C}}_{32} & \hat{\mathbf{C}}_{33} \end{bmatrix},$$

where $\hat{\mathbf{\Lambda}}_3 \in \mathbb{C}^{P \times P}$, with $P \leq M$, is the eigenvalue matrix containing those eigenvalues whose derivatives are equal. Then $\hat{\mathbf{\Lambda}}_2 \in \mathbb{C}^{(M-P) \times (M-P)}$ is the eigenvalue matrix that contains the remaining repeated eigenvalues.

The primary goal is to determine $\mathbf{\Gamma}$. To determine $\hat{\mathbf{\Gamma}}_3$, relation (6.13) is differentiated once more with respect to ρ and is premultiplied by the left-eigenvector matrix $\tilde{\mathbf{Y}}^*$. Next, relations (6.21), (6.15) and (6.35) are substituted, which are the expressions for \mathbf{W} , \mathbf{W}' and \mathbf{W}'' , respectively. To find $\hat{\mathbf{\Gamma}}_3$ we use the partitioning given by (6.37).

$$\hat{\mathbf{Y}}_3^* \mathbf{A}'' \hat{\mathbf{W}}_3 \hat{\mathbf{\Gamma}}_3 - \hat{\mathbf{\Gamma}}_3 \mathbf{\Lambda}_3'' = 2\hat{\mathbf{\Gamma}}_3 \hat{\mathbf{C}}_{31} (\lambda \mathbf{I} - \mathbf{\Lambda}_1) \hat{\mathbf{C}}_{13}. \quad (6.38)$$

For simplicity of notation, which will be of high importance in the generalization of the theory, we omit the hat again. Unfortunately \mathbf{C}_{31} and \mathbf{C}_{13} cannot be computed since $\mathbf{\Gamma}_3$ is not yet determined. However, (6.32c) and (6.32b) give expressions for \mathbf{C}_{13} and \mathbf{C}_{31} in terms of $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_3$. Insertion of these expressions into (6.38) again gives an eigenproblem with $\mathbf{\Gamma}_3$ the eigenvector matrix and $\mathbf{\Lambda}_3''$ the eigenvalue matrix. Hence

$$\left(\tilde{\mathbf{Y}}_3^* \mathbf{A}'' \tilde{\mathbf{W}}_3 + 2\tilde{\mathbf{Y}}_3^* \mathbf{A}' \tilde{\mathbf{W}}_1 (\lambda \mathbf{I} - \mathbf{\Lambda}_1)^{-1} \tilde{\mathbf{Y}}_1^* \mathbf{A}' \tilde{\mathbf{W}}_3 \right) \mathbf{\Gamma}_3 - \mathbf{\Gamma}_3 \mathbf{\Lambda}_3'' = \mathbf{0}. \quad (6.39)$$

If the second-order derivatives of the eigenvalues are distinct, $\mathbf{\Gamma}_3$ consists of column vectors that are unique up to a multiplicative constant. Again, (6.2) ensures an expression for the diagonal entries of $\mathbf{\Gamma}_3$. Since $\mathbf{\Gamma}_2$ has already been determined in (6.32d) and for $\mathbf{\Gamma}_1$ the same procedure can be performed until $\mathbf{\Gamma}_1$ has been determined uniquely, the desired continuously differentiable eigenvector matrix has been found.

To determine the eigenvector derivatives, the coefficient matrix \mathbf{C} needs to be computed. The blocks \mathbf{C}_{12} , \mathbf{C}_{13} , \mathbf{C}_{21} , \mathbf{C}_{31} are computed like in (6.32b) and (6.32c), respectively. The matrices \mathbf{C}_{23} and \mathbf{C}_{32} can be found from

$$\mathbf{Y}_2^* \mathbf{A}'' \mathbf{W}_3 = 2(\lambda' \mathbf{I} - \mathbf{\Lambda}'_2) \mathbf{C}_{23} + 2\mathbf{C}_{21}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{13}, \quad (6.40a)$$

$$\mathbf{Y}_3^* \mathbf{A}'' \mathbf{W}_2 = -2\mathbf{C}_{32}(\lambda' \mathbf{I} - \mathbf{\Lambda}'_2) + 2\mathbf{C}_{31}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{12}. \quad (6.40b)$$

The matrix \mathbf{C}_{33} follows from differentiating equation (6.13) twice, viz.

$$\begin{aligned} \mathbf{Y}_3^* \mathbf{A}''' \mathbf{W}_3 - \mathbf{\Lambda}_3''' &= -6\mathbf{C}_{31}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{11} \mathbf{C}_{13} - 6\mathbf{C}_{31}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{12} \mathbf{C}_{23} \\ &\quad - 6\mathbf{C}_{31}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{13} \mathbf{C}_{33} + 3\mathbf{D}_{31}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{C}_{13} \\ &\quad + 3\mathbf{C}_{31}(\lambda \mathbf{I} - \mathbf{\Lambda}_1) \mathbf{D}_{13} + 6\mathbf{C}_{31}(\lambda' \mathbf{I} - \mathbf{\Lambda}'_1) \mathbf{C}_{13} \\ &\quad + 6\mathbf{C}_{32}(\lambda' \mathbf{I} - \mathbf{\Lambda}'_2) \mathbf{C}_{23} + 3(\mathbf{C}_{33} \mathbf{\Lambda}_3'' - \mathbf{\Lambda}_3'' \mathbf{C}_{33}). \end{aligned} \quad (6.41)$$

Note that \mathbf{D} has the same partitioning as \mathbf{C} . Now that all off-diagonal entries of \mathbf{C} have been found, we use (6.29) to determine the diagonal entries to fill the coefficient matrix completely.

6.3 Generalization

The cases discussed in the previous section provide us a way to generalize the theory of finding the first-order derivatives of the eigenvectors for eigenvalues with the property that (some of) their derivatives up to the k th order are repeated, but the $(k+1)$ st-order derivatives are distinct. This generalization will be given in this section and for readability we try to keep the notation as simple as possible.

Let \mathbf{A} , its eigenvalue and eigenvector matrix $\mathbf{\Lambda}$ and \mathbf{W} , be n times continuously differentiable. Then from *Leibniz' rule* we have

$$\mathbf{A}^{(n)} \mathbf{W} - \mathbf{W} \mathbf{\Lambda}^{(n)} = - \sum_{k=1}^n \binom{n}{k} \left(\mathbf{A}^{(n-k)} \mathbf{W}^{(k)} - \mathbf{W}^{(k)} \mathbf{\Lambda}^{(n-k)} \right). \quad (6.42)$$

As before, let the non-uniqueness of \mathbf{W} be expressed by

$$\mathbf{W} = \tilde{\mathbf{W}} \mathbf{\Gamma}, \quad (6.43)$$

where $\mathbf{\Gamma}$ is a nonsingular matrix to be determined. Note that $\mathbf{\Gamma}$ has to satisfy (6.22). Since \mathbf{A} is non-defective, the k th-order derivative of the eigenvector matrix can be expressed in

terms of the eigenvectors of \mathbf{A} .

$$\mathbf{W}^{(k)} := \mathbf{W}\mathbf{C}_k, \quad k = 1, \dots, n. \quad (6.44)$$

Upon substituting (6.43) and (6.44) into (6.42) we obtain

$$\begin{aligned} \tilde{\mathbf{Y}}^* \mathbf{A}^{(n)} \tilde{\mathbf{W}} \mathbf{\Gamma} - \mathbf{\Gamma} \mathbf{\Lambda}^{(n)} = & \\ & - \mathbf{\Gamma} \sum_{k=0}^{n-1} \beta_0 \left(\mathbf{\Lambda}^{(k)} \mathbf{C}_{n-k} - \mathbf{C}_{n-k} \mathbf{\Lambda}^{(k)} \right) \\ & + \mathbf{\Gamma} \sum_{k=0}^{n-1} \sum_{\gamma=1}^n \sum_{m_1=1}^{\alpha_1} \dots \sum_{m_\gamma=1}^{\alpha_{\gamma-1}} (-1)^\gamma \beta_\gamma \left(\mathbf{\Lambda}^{(k)} \mathbf{C}_{m_1} - \mathbf{C}_{m_1} \mathbf{\Lambda}^{(k)} \right) \mathbf{C}_{m_2} \dots \mathbf{C}_{m_\gamma} \mathbf{C}_{m_{\gamma+1}}, \end{aligned} \quad (6.45)$$

where

$$\alpha_r := n - k - \sum_{j=1}^{r-1} m_j - 1, \quad r = 1, \dots, \gamma, \quad (6.46a)$$

$$\beta_0 := \binom{n}{k}, \quad (6.46b)$$

$$\beta_\gamma := \binom{n}{k} \binom{n-k}{m_1} \binom{n-k-m_1}{m_2} \dots \binom{n-k-m_1-\dots-m_{\gamma-1}}{m_\gamma}, \quad (6.46c)$$

$$m_{\gamma+1} := n - k - \sum_{j=1}^{\gamma} m_j. \quad (6.46d)$$

To determine $\mathbf{\Gamma}$, a generalization of the partitioning of the matrices \mathbf{A} , \mathbf{W} , \mathbf{Y}^* and \mathbf{C}_k for $k = 1, \dots, n-1$ has to be introduced. Thus, let the eigenvalue matrix \mathbf{A} be partitioned as

$$\mathbf{A} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{\Lambda}_{n+1} \end{bmatrix}, \quad (6.47)$$

such that

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \lambda \mathbf{I} \end{bmatrix}, \mathbf{A}' = \begin{bmatrix} \mathbf{\Lambda}'_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}'_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \lambda' \mathbf{I} \end{bmatrix}, \dots, \\ \mathbf{A}^{(n)} &= \begin{bmatrix} \mathbf{\Lambda}_1^{(n)} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2^{(n)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_3^{(n)} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{\Lambda}_{n+1}^{(n)} \end{bmatrix}. \end{aligned} \quad (6.48)$$

Notice that $\mathbf{\Lambda}_1$ in (6.48) can still contain repeated eigenvalues, but not the eigenvalue λ . On the other hand nothing is prescribed anymore for the derivatives, which means that

some diagonal entry of Λ'_1 might be equal to λ' . According to the partitioning of \mathbf{A} the following partitioning of $\tilde{\mathbf{W}}$, $\tilde{\mathbf{Y}}^*$ and \mathbf{C}_k is introduced.

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}}_1 & \cdots & \tilde{\mathbf{W}}_{n+1} \end{bmatrix}, \quad \tilde{\mathbf{Y}}^* = \begin{bmatrix} \tilde{\mathbf{Y}}_1^* \\ \vdots \\ \tilde{\mathbf{Y}}_{n+1}^* \end{bmatrix},$$

$$\mathbf{C}_m = \begin{bmatrix} \mathbf{C}_{m,11} & \mathbf{C}_{m,12} & \cdots & \mathbf{C}_{m,1(n+1)} \\ \mathbf{C}_{m,21} & \mathbf{C}_{m,22} & \cdots & \mathbf{C}_{m,2(n+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{m,(n+1)1} & \mathbf{C}_{m,(n+1)2} & \cdots & \mathbf{C}_{m,(n+1)(n+1)} \end{bmatrix}. \quad (6.49)$$

To determine Γ_{n+1} , the partitioning of \mathbf{A} , $\tilde{\mathbf{W}}$, $\tilde{\mathbf{Y}}^*$ and \mathbf{C} has to be substituted in (6.45). Then Γ_{n+1} will follow from the $(n+1, n+1)$ block.

$$\begin{aligned} & \tilde{\mathbf{Y}}_{n+1}^* \mathbf{A}^{(n)} \tilde{\mathbf{W}}_{n+1} \Gamma_{n+1} - \Gamma_{n+1} \Lambda_{n+1}^{(n)} = \\ & \Gamma \sum_{k=0}^{n-1} \sum_{\gamma=1}^n \sum_{m_1=1}^{\alpha_1} \cdots \sum_{m_\gamma=1}^{\alpha_\gamma} \sum_{p_1=1}^{n+1} \cdots \sum_{p_\gamma=1}^n (-1)^\gamma \beta_\gamma \\ & \quad \cdot \mathbf{C}_{m_1, (n+1)p_1} \left(\lambda^{(k)} \mathbf{I} - \Lambda_{p_1}^{(k)} \right) \mathbf{C}_{m_2, p_1 p_2} \cdots \mathbf{C}_{m_{\gamma+1}, p_\gamma (n+1)}. \end{aligned} \quad (6.50)$$

Not all coefficient matrices can be computed since the eigenvectors are still not fixed. However, an analytical expression for each coefficient matrix can be derived from (6.45) by considering the right derivative and block matrices. As a result we obtain the following.

$$\tilde{\mathbf{Y}}_{n+1}^* \mathbf{A}^{(n)} \tilde{\mathbf{W}}_{n+1} \Gamma_{n+1} - \Gamma_{n+1} \Lambda_{n+1}^{(n)} = -\mathcal{L}(n, n-1, n+1, n+1) \Gamma_{n+1}, \quad (6.51)$$

where the right-hand-side is found by using the following recurrence relation for \mathcal{L} :

$$\begin{aligned} \mathcal{L}(n, m, p, q) & := \\ & \sum_{\ell=1}^m \sum_{k=\ell}^{n-1} \frac{\gamma_{k\ell}}{\ell} \left(\hat{\mathbf{Y}}_p^* \mathbf{A}^{(k)} \hat{\mathbf{W}}_\ell - \delta_{p\ell} \lambda^{(k)} \mathbf{I} + \mathcal{L}(k, \ell-1, p, \ell) \right) \left(\lambda^{(\ell-1)} \mathbf{I} - \Lambda_\ell^{(\ell-1)} \right)^{-1} \\ & \quad \cdot \left(\mathcal{L}(n-k+\ell-1, \ell, \ell, q) + \hat{\mathbf{Y}}_\ell^* \mathbf{A}^{(n-k+\ell-1)} \hat{\mathbf{W}}_q - \delta_{\ell q} \lambda^{(n-k+\ell-1)} \mathbf{I} \right). \end{aligned} \quad (6.52)$$

In this relation the constant $\gamma_{k\ell}$ is defined as

$$\gamma_{k\ell} := \begin{cases} \binom{n}{k} & \text{if } k \geq n - k + \ell - 1, \\ \binom{n}{n-k+\ell-1} & \text{if } k < n - k + \ell - 1, \end{cases} \quad (6.53)$$

and

$$\hat{\mathbf{W}}_q := \begin{cases} \tilde{\mathbf{W}}_1 & \text{if } q = 1, \\ \tilde{\mathbf{W}}_q \Gamma_q & \text{if } 1 < q < n+1, \\ \tilde{\mathbf{W}}_{n+1} & \text{if } q = n+1. \end{cases} \quad (6.54)$$

By computing the eigenvalues and eigenvectors of (6.51) the matrix Γ_{n+1} can be determined for every n .

To determine an eigenvector matrix that is continuously differentiable for repeated eigenvalues, (6.51) has to be solved several times for different values of n . When repeated eigenvalues occur, (6.51) has to be evaluated for the first-order derivative. If (some of) the eigenvalues of the resulting matrix are still the same, (6.51) has to be evaluated once more for the second-order derivative. This process has to be repeated until (6.51) does not result in any repeated eigenvalues anymore. At each iteration step the eigenvectors have to be updated by multiplication of the Γ matrix. This iterative process is illustrated in Figure 6.1. After that all the blocks of Γ have been determined, the only thing left is the normalization of the eigenvectors as in (6.2) taking (6.4) into account and computing the coefficient matrix \mathbf{C}_1 .

After a continuously differentiable eigenvector matrix has been found, i.e. in fact Γ , the coefficient matrix \mathbf{C}_1 can be determined. For this the same recurrence relation as in the conjecture is used. Also these relations are a result of extending the procedure of the simple cases presented in the previous section.

$$\mathbf{C}_{1,mn} := \frac{1}{m} \left(\lambda^{(m-1)} \mathbf{I} - \Lambda_m^{(m-1)} \right)^{-1} \cdot \left\{ \mathbf{Y}_m^* \mathbf{A}^{(m)} \mathbf{W}_n + \mathcal{L}(m, m-1, m, n) \right\} \text{ if } m < n, \quad (6.55a)$$

$$\mathbf{C}_{1,mn} := \frac{1}{n} \left\{ \mathbf{Y}_m^* \mathbf{A}^{(n)} \mathbf{W}_n + \mathcal{L}(m, m-1, m, n) \right\} \cdot \left(\lambda^{(n-1)} \mathbf{I} - \Lambda_n^{(n-1)} \right)^{-1} \text{ if } m > n. \quad (6.55b)$$

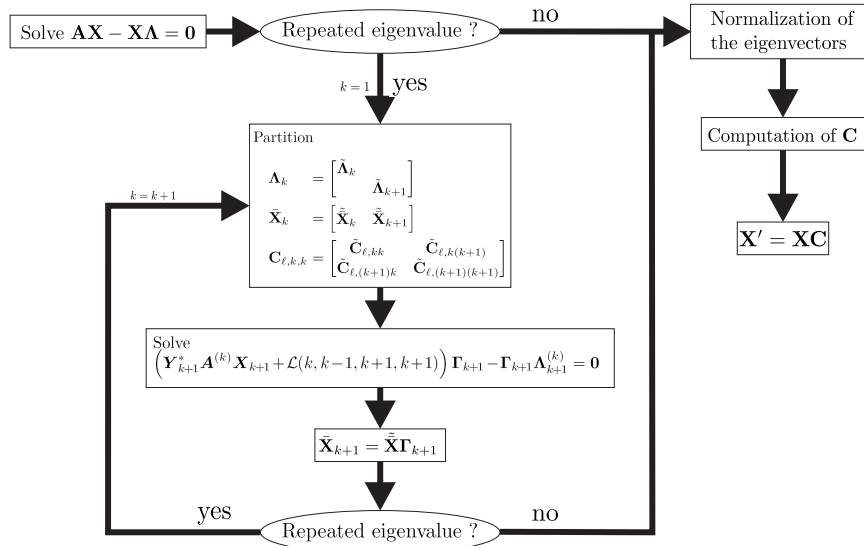


Figure 6.1: A flowchart representation of the way to compute Γ .

For the diagonal matrices $\mathbf{C}_{1,mm}$ for $m = 1, \dots, n + 1$ it holds that only the off-diagonal entries can be determined by the following formula

$$\begin{aligned} \Lambda_m^{(m-1)} \mathbf{C}_{1,mm} - \mathbf{C}_{1,mm} \Lambda_m^{(m-1)} := \\ \frac{1}{m} \left\{ \mathbf{Y}_m^* \mathbf{A}^{(m)} \mathbf{W}_m - \Lambda_m^{(m)} + \mathcal{L}(m, m-1, m, m) \right\}. \end{aligned} \quad (6.56)$$

Finally, the diagonal entries of \mathbf{C}_1 can then be found by

$$c_{kk} = - \sum_{\substack{l=1 \\ l \neq k}}^N w_{ml} c_{lk}, \quad (6.57)$$

analogously as for the simple cases in the previous section. After the determination of \mathbf{C}_1 , the first-order derivative of the eigenvector matrix \mathbf{W}' can be found.

The only thing that still has to be done is to find Λ'_1 . If Λ_1 still contains repeated eigenvalues, the whole process described in this section has to be performed for Λ_1 again. If Λ_1 only consists of distinct eigenvalues, Λ_1 is computed by determining the diagonal of $\mathbf{Y}_1^* \mathbf{A}' \mathbf{W}_1$, as is done similarly in (6.25).

Finally, a last remark has to be given. The fact that Λ_1 can still contain repeated eigenvalues ensures that Γ_1 does not have to be a diagonal matrix, but this matrix is unknown. However, the fact that Γ commutes with Λ and that no higher-order derivative of Λ_1 occurs, Γ_1 disappears from (6.51) and therefore definition (6.54) is justified.

6.4 Examples

In this section the generalization of the method presented in the previous section will be illustrated by some examples. The first three examples show that the three simple cases discussed in Section 6.2 fit with the general algorithm. The fourth example is a numerical example.

Example 6.4 If all eigenvalues of Λ are distinct, the only thing left for the eigenvectors is to normalize them according to (6.2) taking (6.4) into account. To compute the coefficient matrix \mathbf{C}_1 and the eigenvalue derivative matrix Λ' , we have to write out (6.56) in coordinates for $m = 1$. This results in (6.25) and (6.26), but now the values of γ_k for $k = 1, \dots, N$ are already known from the normalization. The diagonal entries of \mathbf{C}_1 do not follow from (6.56), but from (6.57) which is identical to (6.29) where again all γ_k are known. \square

Example 6.5 If $\mathbf{\Lambda}$ contains a repeated eigenvalue, a partitioning (6.47) has to be performed on $\mathbf{\Lambda}$ for $n = 1$ and a similar one on the eigenvector matrix $\tilde{\mathbf{W}}$ and coefficient matrix \mathbf{C} . This partitioning is equal to (6.31). Next, (6.51) has to be evaluated for $n = 1$. In this case \mathcal{L} is equal to zero and therefore (6.51) indeed reduces to (6.32d). Since this eigensystem does not yield any repeated eigenvalues again, also here the normalization of the eigenvectors takes place and the computation of \mathbf{C}_1 can be carried out by using (6.55b), (6.55a), (6.56) and (6.57) again. \square

Example 6.6 If both $\mathbf{\Lambda}$ and the eigenvalues of system (6.32d) contain repeated eigenvalues, partitioning (6.47) has to be performed for $n = 2$. Similarly, partitioning (6.49) is applied to the eigenvector matrix $\tilde{\mathbf{W}}$ and coefficient matrix \mathbf{C} . The result resembles the partitioning of (6.37). The evaluation of (6.51) returns the first nontrivial case, since \mathcal{L} is nonzero, viz.

$$\begin{aligned}\mathcal{L}(2, 1, 3, 3) &= 2\mathbf{Y}_3^* \mathbf{A}' \mathbf{W}_1 (\lambda \mathbf{I} - \mathbf{\Lambda}_1)^{-1} \mathbf{Y}_1^* \mathbf{A}' \mathbf{W}_3 \\ &= 2\tilde{\mathbf{Y}}_3^* \mathbf{A}' \tilde{\mathbf{W}}_1 (\lambda \mathbf{I} - \mathbf{\Lambda}_1)^{-1} \tilde{\mathbf{Y}}_1^* \mathbf{A}' \tilde{\mathbf{W}}_3 \Gamma_3.\end{aligned}\quad (6.58)$$

With this \mathcal{L} , (6.51) reduces to (6.39). Because this eigensystem is assumed to have only distinct eigenvalues, we can proceed by normalizing the eigenvectors and computing \mathbf{C}_1 in a similar way as in the previous simple case. \square

Example 6.7 The last example considers the computation of eigenvalue and eigenvector derivatives with respect to parameter p at p_0 for a matrix \mathbf{A} for which

$$\lambda_1(p_0) \neq \lambda_2(p_0) = \lambda_3(p_0) = \lambda_4(p_0) = \lambda_5(p_0), \quad (6.59a)$$

$$\lambda_2'(p_0) \neq \lambda_3'(p_0) = \lambda_4'(p_0) = \lambda_5'(p_0), \quad (6.59b)$$

$$\lambda_3''(p_0) \neq \lambda_4''(p_0) = \lambda_5''(p_0), \quad (6.59c)$$

$$\lambda_4^{(3)}(p_0) \neq \lambda_5^{(3)}(p_0). \quad (6.59d)$$

We construct a matrix \mathbf{A} by constructing an eigenvalue matrix $\mathbf{\Lambda}$ with these properties and defining an eigenvector matrix \mathbf{W} by

$$\mathbf{\Lambda}(p) = \text{diag} \begin{pmatrix} 1 \\ -1 + 11i + \sin(2\pi p) + \sin(6\pi p) \\ (1 - 11i) \cos(6\pi p) - 3 \sin(2\pi p) + \sin(6\pi p) \\ 7i + (1 - 4i) \cos(2\pi p) - 3 \sin(2\pi p) + \sin(6\pi p) \\ 7i + (6 - 4i) \cos(2\pi p) + 8 \cos(4\pi p) + 3 \cos(6\pi p) \end{pmatrix}, \quad (6.60)$$

$$\mathbf{W}(p) = \begin{bmatrix} i & 0 & 0 & 1 & 2p \\ 1 & 0 & i & 2p & 0 \\ 1 & i & 2p & 0 & 0 \\ 1 & 2p & 0 & 0 & 1 \\ 2p & 0 & 0 & 1 & 0 \end{bmatrix}. \quad (6.61)$$

Then the matrix \mathbf{A} is defined as

$$\mathbf{A} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^{-1}. \quad (6.62)$$

Clearly, we have analytical expressions for the eigenvalues and eigenvectors. Therefore, the results of the numerical method presented in this chapter can be verified immediately, because the derivatives of the eigenvalues and eigenvectors are also available analytically.

In general, the analytical expressions for the eigenvalues and eigenvectors are hard, if not impossible to obtain. However, the repeated eigenvalues only occur at $p = 0.5$ and \mathbf{A} for this value becomes

$$\mathbf{A}(0.5) = \begin{bmatrix} -6.5 + 10.0i & -5.5 - 1.0i & -1.0 + 5.5i & 5.5 + 1.0i & 11.0 + 2.0i \\ -1.0 + 5.5i & -2.0 + 16.5i & 5.5 + 1.0i & 1.0 - 5.5i & 2.0 - 11.0i \\ -1.0 + 5.5i & -1.0 + 5.5i & 4.5 + 12.0i & 1.0 - 5.5i & 2.0 - 11.0i \\ -1.0 + 5.5i & -1.0 + 5.5i & 5.5 + 1.0i & 5.5i & 2.0 - 11.0i \\ -1.0 + 5.5i & -1.0 + 5.5i & 5.5 + 1.0i & 1.0 - 5.5i & 1.0 \end{bmatrix}$$

To find a continuously differentiable eigenvector matrix we use the algorithm illustrated in Figure 6.1. The ingredients of this algorithm such as $\mathbf{A}'(0.5)$, $\mathbf{A}''(0.5)$, etc. can be evaluated since \mathbf{A} is known as a function of p . Note that in this example outcomes are given in 4 decimals.

Initialization step

$$\mathbf{\Lambda}_1 = 1, \mathbf{\Lambda}_2 = (-1 + 11i)\mathbf{I}^{4 \times 4}, \mathbf{W}_1(0.5) = \begin{bmatrix} i \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

$$\tilde{\mathbf{W}}_2(0.5) = \begin{bmatrix} 0.0640 + 0.3378i & 0.3344 - 0.2241i & 0.7454 & 0.8347 \\ 0.2113 - 0.0292i & 0.0417 + 0.4298i & 0.2981 - 0.1491i & -0.5274 - 0.0162i \\ 0.0783 + 0.5921i & 0.6093 & 0.2981 - 0.1491i & 0.0314 + 0.0154i \\ 0.6762 & 0.3739 + 0.2131i & 0.2981 - 0.1491i & 0.1076 - 0.0159i \\ 0.0956 + 0.1152i & 0.0011 - 0.3083i & 0.2981 - 0.1491i & 0.1076 - 0.0159i \end{bmatrix}.$$

Step 2

$$\mathbf{\Lambda}'_2(0.5) = -8\pi, \mathbf{\Lambda}'_3(0.5) = \mathbf{0}^{3 \times 3}, \mathbf{W}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -i \\ 0 \end{bmatrix},$$

$$\tilde{\mathbf{W}}_3(0.5) = \begin{bmatrix} 0.1121 - 0.0237i & 0.6333 + 0.1628i & 0.0069 + 0.0062i \\ -0.0689 + 0.0204i & 0.1039 - 0.1054i & -0.3404 - 0.1994i \\ 0.2719 - 0.0366i & -0.1329 + 0.1467i & -0.3422 + 0.2390i \\ 0.2175 + 0.2279i & 0.3827 + 0.1353i & 0.1082 - 0.1366i \\ -0.1054 - 0.2515i & 0.2506 + 0.0275i & -0.1014 + 0.1428i \end{bmatrix}.$$

Step 3

$$\mathbf{\Lambda}''_3(0.5) = (36 - 396i)\pi^2, \mathbf{\Lambda}''_4(0.5) = (4 - 16i)\pi^2 \mathbf{I}^{2 \times 2},$$

$$\mathbf{W}_3(0.5) = \begin{bmatrix} 0 \\ 1 \\ -i \\ 0 \\ 0 \end{bmatrix}, \tilde{\mathbf{W}}_4(0.5) = \begin{bmatrix} 0.2591 + 0.3698i & 0.0886 - 0.0603i \\ -0.0133 - 0.0239i & -0.1805 - 0.1802i \\ 0 & 0 \\ 0.2725 + 0.3937i & 0.2691 + 0.1200i \\ -0.0133 - 0.0239i & -0.1805 - 0.1802i \end{bmatrix}.$$

Step 4

$$\mathbf{\Lambda}_4^{(3)}(0.5) = 192\pi^3, \mathbf{\Lambda}_5^{(3)}(0.5) = \mathbf{0}^{2 \times 2}, \mathbf{W}_4(0.5) = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

$$\tilde{\mathbf{W}}_5(0.5) = \begin{bmatrix} 0.2730 - 0.2679i & 0.1196 - 0.1276i \\ -0.1154 - 0.2560i & 0.1529 + 0.0069i \\ 0 & 0 \\ 0 & 0 \\ 0.0170 - 0.1525i & 0.1265 - 0.2805i \\ 0 & 0 \end{bmatrix}.$$

Step 5

$$\mathbf{\Lambda}_5^{(4)}(0.5) = \begin{bmatrix} 0 & 0 \\ 0 & 192\pi^3 \end{bmatrix}, \mathbf{W}_5 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Since no repeated eigenvalues occur anymore, the continuously differentiable eigenvector matrix has been determined, which is equal to (6.61) for $p = 0.5$. Note that the normalization (6.2) has already been performed and the eigenvector matrix that is continuously differentiable at $p = 1/2$ taking into account the normalization is given by

$$\mathbf{W}(0.5) = \begin{bmatrix} i & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -i & 0 & 0 \\ 1 & -i & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (6.63)$$

The next step is to determine \mathbf{C} . The blocks \mathbf{C}_{mk} for $m, k = 1, \dots, 6$ with $m \neq k$ and the off-diagonal entries of \mathbf{C}_{mm} can be determined by using (6.55a), (6.55b) and (6.56), respectively. The diagonal entries of \mathbf{C} can be computed by using (6.57). As a consequence \mathbf{C} equals

$$\mathbf{C} = \begin{bmatrix} 0 & -i & 1 & -1 & -1 \\ -2 - 2i & 0 & -1 - i & 1 & 1 + 2i \\ -2 & 0 & 0 & 0 & 2 \\ -2i & -1 - i & 1 - i & 0 & -1 + i \\ 0 & i & -1 & 1 & 0 \end{bmatrix}. \quad (6.64)$$

To conclude this example, the eigenvalue and eigenvector derivative matrix are given by

$$\mathbf{\Lambda}'(0.5) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -8\pi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (6.65)$$

$$\mathbf{W}'(0.5) = \mathbf{W}(0.5)\mathbf{C} = \begin{bmatrix} -2i & 0 & 0 & 1-i & -1 \\ -2 & 0 & 0 & 0 & 1 \\ -2 & -i & -2i & 0 & 0 \\ -2 & -1 & -2i & 0 & -1-i & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \quad (6.66)$$

Note that this eigenvector derivative matrix belongs to the continuously differentiable eigenvector matrix with the normalization (6.2) taken into account. \square

6.5 Consequences for RCWA

A mathematical theory has been derived in this chapter to obtain all eigenvectors of \mathbf{A} with the property that they have to be continuously differentiable at $p = p_0$. In this section we will discuss its relevance to the RCWA algorithm by means of two questions, namely whether repeated eigenvalues occur inside RCWA at all and secondly whether the normalization of eigenvectors is needed.

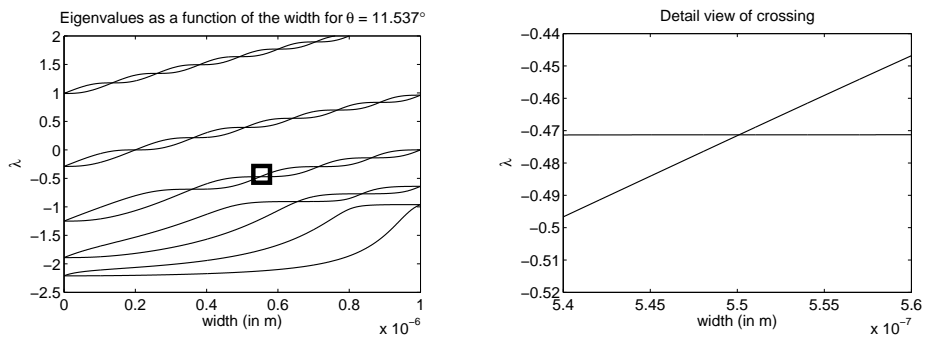


Figure 6.2: Illustration of the crossing of eigenvalues inside a layer consisting of a block with refractive index 1.5 and a block with air. The incident field is TE polarization with a wavelength of 0.4λ and the angle of incidence satisfies the relation (6.68). The figure on the right gives the eigenvalues as a function of the block width, while the left figure is a close-up of the crossing of two eigenvalues.

6.5.1 The occurrence of repeated eigenvalues

Let us first discuss whether repeated eigenvalues occur in the RCWA algorithm at all. For planar and conical diffraction the matrix of which the eigenvalues have to be computed are defined in (3.13), (3.21), (3.29) and (3.32). Appendix A.3 shows that when the refractive index inside a layer is real-valued, the eigenvalues will also be real-valued. Because of this fact, we use real-valued refractive indices for illustration purposes.

For TE polarization, the matrix of which the eigenvalues and eigenvectors have to be computed is defined in (3.13). It consists of a sum of two matrices. The first one, \mathbf{K}_x^2 , is diagonal, while the second one, \mathbf{E}_j , is a Toeplitz matrix. For homogeneous layers, \mathbf{E}_j becomes a constant times the identity matrix and it is easily shown that repeated eigenvalues occur when $k_{xn}^2 = k_{x(n+p)}^2$, which gives

$$2n_1 p \sin \theta = (2np + p^2) \frac{\lambda}{\Lambda}. \quad (6.67)$$

Since two diagonal entries should be equal, it holds that $p \neq 0$ and we can divide both sides of (6.67) by p . As a result, we find an expression for $\sin \theta$ in terms of the ratio between the

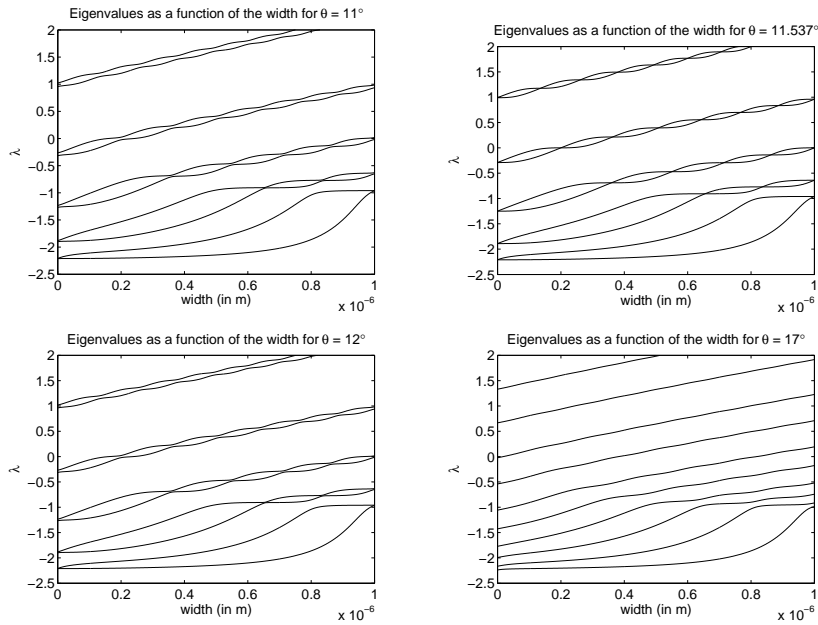


Figure 6.3: Illustration of the phenomena avoidance of crossing of eigenvalues. Four evaluations with different angles of incidence (11° , 11.537° , 12° and 17°) have been carried out to show the eigenvalues as a function of the block width of a binary grating structure with refractive index 1.5.

wavelength and the period of the grating, viz.

$$\sin \theta = \frac{(2n + p) \lambda}{2n_1 \Lambda}. \quad (6.68)$$

The question whether repeated eigenvalues still occur when we have material transitions inside a layer, can be answered affirmatively. From Figure 6.2 we can see that repeated eigenvalues only occur when θ satisfies (6.68). Furthermore, the eigenvalues show the *avoidance of crossing* phenomena as described by [23]. This phenomena is shown in Figure 6.3, where we deviate just slightly from the optimal θ satisfying (6.68).

In the TM polarization case no repeated eigenvalues occur due to the matrix definition (3.21), unless the layer is homogeneous. The matrix of which the eigenvalues have to be computed for conical diffraction is a compilation of the matrices for TE and TM polarization and, therefore, repeated eigenvalues can occur there as well. The following section discusses whether repeated eigenvalues really give problems.

6.5.2 The redundance of normalizing the eigenvectors

Within RCWA the goal is to find the derivatives of the field amplitudes or the diffraction efficiencies. Although we tried to make the necessity of normalizing the eigenvectors clear in this chapter, Chapter 3 manages to describe the complete RCWA algorithm without using this at all. This is because the normalization matrix $\mathbf{\Gamma}$ drops out. A physical argument for this is that within RCWA the computations of eigenvalues and eigenvectors are merely a mathematical way to obtain the field amplitudes, which are physical quantities that can be measured. The choice of the eigenvectors should not and also does not matter for the values of the field amplitudes.

Theorem 6.8 (Redundancy of $\mathbf{\Gamma}$)

The reflected or transmitted field amplitudes, obtained by solving (3.65), are independent of the choice of eigenvectors and therefore are independent of the choice of $\mathbf{\Gamma}$.

To give a formal proof, two lemmas will be needed.

Lemma 6.9

If $\mathbf{\Gamma}$ commutes with $\mathbf{\Lambda}_i$, it also commutes with \mathbf{X}_i and \mathbf{Q}_i .

Proof. Let us start with \mathbf{Q}_i , which is a diagonal matrix with diagonal elements $q_{i,n} = \sqrt{\lambda_{i,n}}$. Because $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$ commute, we have that $\lambda_{i,n} \gamma_{nm} = \gamma_{nm} \lambda_{i,m}$. This gives two possibilities, namely $\gamma_{nm} = 0$ or $\lambda_{i,n} = \lambda_{i,m}$. Both possibilities ensure that $q_{i,n} \gamma_{nm} = \gamma_{nm} q_{i,m}$ and therefore \mathbf{Q}_i commutes with $\mathbf{\Gamma}$.

The matrix \mathbf{X}_i is also a diagonal matrix with elements $x_{i,n} = \exp[-k_0 d q_{i,n}]$. By using

a Taylor expansion $\mathbf{X}_i = \mathbf{I} + k_0 d \mathbf{Q}_i + (k_0 d)^2 \mathbf{Q}_i^2 + \dots + (k_0 d)^n \mathbf{Q}_i^n + \dots$ and using the fact that $\mathbf{Q}_i \mathbf{\Gamma} = \mathbf{\Gamma} \mathbf{Q}_i$ repetitively, proves that $\mathbf{X}_i \mathbf{\Gamma} = \mathbf{\Gamma} \mathbf{X}_i$. \square

For planar diffraction we can write the eigenvector matrix \mathbf{W}_i as in (6.21), that is $\mathbf{W}_i = \tilde{\mathbf{W}}_i \mathbf{\Gamma}$. The matrix \mathbf{V}_i is defined for planar diffraction by (3.51). For conical diffraction matrices \mathbf{W}_i and \mathbf{V}_i are defined by (3.60) and (3.61). These matrices consist of block matrices $\mathbf{W}_{1,i}$ and $\mathbf{W}_{2,i}$. Both of them can be written as $\mathbf{W}_{m,i} = \tilde{\mathbf{W}}_{m,i} \mathbf{\Gamma}_m$ for $m = 1, 2$. Matrices \mathbf{W}_i and \mathbf{V}_i also contain block matrices $\mathbf{V}_{11,i}$, $\mathbf{V}_{12,i}$, $\mathbf{V}_{21,i}$ and $\mathbf{V}_{22,i}$, where we can also extract $\mathbf{\Gamma}$ as shown by the following lemma:

Lemma 6.10

For planar diffraction define $\tilde{\mathbf{V}}_i := \tilde{\mathbf{W}}_i \mathbf{Q}_i$ for TE polarization and $\tilde{\mathbf{V}}_i := \mathbf{P}_i \tilde{\mathbf{W}}_i \mathbf{Q}_i$ for TM polarization. For conical diffraction let us define $\tilde{\mathbf{W}}_i$ and $\tilde{\mathbf{V}}_i$ as

$$\tilde{\mathbf{W}}_i = \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{W}}_{2,i} \\ \tilde{\mathbf{V}}_{11,i} & \tilde{\mathbf{V}}_{12,i} \end{bmatrix} \text{ and } \tilde{\mathbf{V}}_i = - \begin{bmatrix} \tilde{\mathbf{W}}_{1,i} & \mathbf{0} \\ \tilde{\mathbf{V}}_{21,i} & \tilde{\mathbf{V}}_{22,i} \end{bmatrix}, \quad (6.69)$$

with

$$\tilde{\mathbf{V}}_{11,i} := \mathbf{A}_i^{-1} \tilde{\mathbf{W}}_{1,i} \mathbf{Q}_{1,i}, \quad \tilde{\mathbf{V}}_{12} := \mathbf{K}_y \mathbf{A}_i^{-1} \mathbf{K}_x \tilde{\mathbf{W}}_{2,i}, \quad (6.70)$$

$$\tilde{\mathbf{V}}_{21,i} := \mathbf{K}_y \mathbf{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \tilde{\mathbf{W}}_{1,i}, \quad \tilde{\mathbf{V}}_{22,i} := \mathbf{B}_i^{-1} \tilde{\mathbf{W}}_{2,i} \mathbf{Q}_{2,i}. \quad (6.71)$$

Then we can write for both types of diffraction \mathbf{W}_i as $\tilde{\mathbf{W}}_i \mathbf{\Gamma}$ and \mathbf{V}_i as $\tilde{\mathbf{V}}_i \mathbf{\Gamma}$.

Proof. The lemma is a result of using (6.21) and Lemma 6.9.

$$\text{TE polarization : } \mathbf{V}_i = \mathbf{W}_i \mathbf{Q}_i = \tilde{\mathbf{W}}_i \mathbf{\Gamma} \mathbf{Q}_i = \tilde{\mathbf{W}}_i \mathbf{Q}_i \mathbf{\Gamma} \equiv \tilde{\mathbf{V}}_i \mathbf{\Gamma}, \quad (6.72)$$

$$\text{TM polarization : } \mathbf{V}_i = \mathbf{P}_i \mathbf{W}_i \mathbf{Q}_i = \mathbf{P}_i \tilde{\mathbf{W}}_i \mathbf{\Gamma} \mathbf{Q}_i = \mathbf{P}_i \tilde{\mathbf{W}}_i \mathbf{Q}_i \mathbf{\Gamma} \equiv \tilde{\mathbf{V}}_i \mathbf{\Gamma}, \quad (6.73)$$

$$\text{Conical diffraction : } \mathbf{W}_i = \begin{bmatrix} \mathbf{0} & \mathbf{W}_{2,i} \\ \mathbf{V}_{11,i} & \mathbf{V}_{12,i} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{W}}_{2,i} \\ \tilde{\mathbf{V}}_{11,i} & \tilde{\mathbf{V}}_{12,i} \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_2 \end{bmatrix}, \quad (6.74)$$

$$\mathbf{V}_i = - \begin{bmatrix} \mathbf{W}_{1,i} & \mathbf{0} \\ \mathbf{V}_{21,i} & \mathbf{V}_{22,i} \end{bmatrix} = - \begin{bmatrix} \tilde{\mathbf{W}}_{1,i} & \mathbf{0} \\ \tilde{\mathbf{V}}_{21,i} & \tilde{\mathbf{V}}_{22,i} \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_2 \end{bmatrix}, \quad (6.75)$$

since

$$\begin{cases} \mathbf{V}_{11,i} = \mathbf{A}_i^{-1} \mathbf{W}_{1,i} \mathbf{Q}_{1,i} = \mathbf{A}_i^{-1} \tilde{\mathbf{W}}_{1,i} \mathbf{\Gamma}_1 \mathbf{Q}_{1,i} = \mathbf{A}_i^{-1} \tilde{\mathbf{W}}_{1,i} \mathbf{Q}_{1,i} \mathbf{\Gamma}_1 \equiv \tilde{\mathbf{V}}_{11,i} \mathbf{\Gamma}_1 \\ \mathbf{V}_{12,i} = \mathbf{K}_y \mathbf{A}_i^{-1} \mathbf{K}_x \mathbf{W}_{2,i} = \mathbf{K}_y \mathbf{A}_i^{-1} \mathbf{K}_x \tilde{\mathbf{W}}_{2,i} \mathbf{\Gamma}_2 \equiv \tilde{\mathbf{V}}_{12,i} \mathbf{\Gamma}_2, \\ \mathbf{V}_{21,i} = \mathbf{K}_y \mathbf{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{W}_{1,i} = \mathbf{K}_y \mathbf{B}_i^{-1} \mathbf{K}_x \mathbf{E}_i^{-1} \tilde{\mathbf{W}}_{1,i} \mathbf{\Gamma}_1 \equiv \tilde{\mathbf{V}}_{21,i} \mathbf{\Gamma}_1 \\ \mathbf{V}_{22,i} = \mathbf{B}_i^{-1} \mathbf{W}_{2,i} \mathbf{Q}_{2,i} = \mathbf{B}_i^{-1} \tilde{\mathbf{W}}_{2,i} \mathbf{Q}_{2,i} \mathbf{\Gamma}_2 \equiv \tilde{\mathbf{V}}_{22,i} \mathbf{\Gamma}_2, \end{cases} \quad (6.76)$$

which completes the proof. \square

Let us return to the proof of Theorem 6.8.

Proof. Consider (3.65). Consider the eigenvalues and eigenvectors inside layer k , which implies that only the following block has to be considered, viz.

$$\begin{aligned} & \begin{bmatrix} \mathbf{W}_i & \mathbf{W}_i \mathbf{X}_i \\ \mathbf{V}_i & -\mathbf{V}_i \mathbf{X}_i \end{bmatrix} \begin{bmatrix} \mathbf{W}_i \mathbf{X}_i & \mathbf{W}_i \\ \mathbf{V}_i \mathbf{X}_i & -\mathbf{V}_i \end{bmatrix}^{-1} \\ &= \frac{1}{2} \begin{bmatrix} \mathbf{W}_i (\mathbf{X}_i^{-1} + \mathbf{X}_i) \mathbf{W}_i^{-1} & \mathbf{W}_i (\mathbf{X}_i^{-1} - \mathbf{X}_i) \mathbf{V}_i^{-1} \\ \mathbf{V}_i (\mathbf{X}_i^{-1} - \mathbf{X}_i) \mathbf{W}_i^{-1} & \mathbf{V}_i (\mathbf{X}_i^{-1} + \mathbf{X}_i) \mathbf{V}_i^{-1} \end{bmatrix}. \end{aligned} \quad (6.77)$$

Discard the i -dependence for notational convenience. Assume the eigenvector matrix $\tilde{\mathbf{W}}$ has been found, then by using (6.21) we have $\mathbf{W} = \tilde{\mathbf{W}}\mathbf{\Gamma}$. Because of Lemma 6.10 only the first block of (6.77) has to be considered, the other blocks are similar. Using Lemma 6.9 this block can be written without the occurrence of $\mathbf{\Gamma}$, viz.

$$\tilde{\mathbf{W}}\mathbf{\Gamma} (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{\Gamma}^{-1} \tilde{\mathbf{W}}^{-1} = \tilde{\mathbf{W}} (\mathbf{X}^{-1} + \mathbf{X}) \tilde{\mathbf{W}}^{-1}. \quad (6.78)$$

Therefore, it can be seen that $\mathbf{\Gamma}$ is redundant. \square

Since both the ordinary and enhanced transmittance matrix approach both have relation (3.65) as their starting point, the results of either algorithm do not depend on the choice of the eigenvectors.

In the computations of the derivatives of the field amplitudes, the eigenvector derivatives are also required. Intuitively, the normalization of the eigenvectors should not play a role here as well. However, the eigenvectors computed by the RCWA algorithm do not ensure that they are continuously differentiable as shown in Section 6.1. Before we give the proof of the normalization redundancy, we need another lemma first.

Lemma 6.11

Define $\tilde{\mathbf{\Lambda}}$, $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{X}}$ as follows

$$(\tilde{\mathbf{\Lambda}})_{kl} := \begin{cases} (\tilde{\mathbf{W}}^{-1})_k \mathbf{A}' \tilde{\mathbf{W}}_l & \text{if } \lambda_k = \lambda_l \\ 0 & \text{if } \lambda_k \neq \lambda_l \end{cases}, \quad (6.79)$$

$$\tilde{\mathbf{Q}} := \frac{1}{2} \mathbf{Q}^{-1} \tilde{\mathbf{\Lambda}} \text{ and } \tilde{\mathbf{X}} := -\frac{k_0 d}{2} \mathbf{Q}^{-1} \tilde{\mathbf{\Lambda}} \mathbf{X}. \quad (6.80)$$

Then

$$\mathbf{\Gamma} \mathbf{\Lambda}' = \tilde{\mathbf{\Lambda}} \mathbf{\Gamma}, \mathbf{\Gamma} \mathbf{Q}' = \tilde{\mathbf{Q}} \mathbf{\Gamma} \text{ and } \mathbf{\Gamma} \mathbf{X}' = \tilde{\mathbf{X}} \mathbf{\Gamma}. \quad (6.81)$$

Proof. Differentiate the original eigensystem with respect to the shape parameter

$$(\tilde{\mathbf{W}}^{-1}) \mathbf{A}' \tilde{\mathbf{W}} \mathbf{\Gamma} - \mathbf{\Gamma} \mathbf{\Lambda}' = -\mathbf{\Gamma} (\mathbf{\Lambda} \mathbf{C} - \mathbf{C} \mathbf{\Lambda}) \quad (6.82)$$

If the right-hand side is written out in components, we can see that

$$(\mathbf{\Lambda} \mathbf{C} - \mathbf{C} \mathbf{\Lambda})_{kl} = 0 \text{ if } \lambda_k = \lambda_l \quad (6.83)$$

and since $\mathbf{\Gamma}$ only has nonzero elements at those positions, the right-hand-side has zeros exactly at those positions where $\mathbf{\Gamma}\mathbf{\Lambda}'$ has not.

The matrix \mathbf{Q} consists of the square roots of $\mathbf{\Lambda}$ with a positive real part and therefore its derivative with respect to the shape parameter is

$$\mathbf{Q}' = \frac{1}{2}\mathbf{Q}^{-1}\mathbf{\Lambda}', \quad (6.84)$$

and thus

$$\mathbf{\Gamma}\mathbf{Q}' = \frac{1}{2}\mathbf{\Gamma}\mathbf{Q}^{-1}\mathbf{\Lambda}' = \frac{1}{2}\mathbf{Q}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}' = \frac{1}{2}\mathbf{Q}^{-1}\tilde{\mathbf{\Lambda}}\mathbf{\Gamma} = \tilde{\mathbf{Q}}\mathbf{\Gamma}. \quad (6.85)$$

In a similar way, matrix \mathbf{X} consists of the exponential of \mathbf{Q} , therefore

$$\mathbf{X}' = -k_0 d\mathbf{Q}'\mathbf{X} \quad (6.86)$$

and therefore

$$\mathbf{\Gamma}\mathbf{X}' = -k_0 d\mathbf{\Gamma}\mathbf{Q}'\mathbf{X} = -k_0 d\tilde{\mathbf{Q}}\mathbf{\Gamma}\mathbf{X} = -k_0 d\tilde{\mathbf{Q}}\mathbf{X}\mathbf{\Gamma} = \tilde{\mathbf{X}}\mathbf{\Gamma}, \quad (6.87)$$

which completes the proof. \square

Theorem 6.12 (Redundancy of $\mathbf{\Gamma}$ in RCWA sensitivity)

The first order derivatives of the reflected or transmitted field amplitudes, obtained by solving the differentiated version of (3.65), are independent of the choice of eigenvectors and therefore independent of the choice of $\mathbf{\Gamma}$.

Proof. As in the proof of Theorem 6.8, consider only (6.77). Differentiate this block with respect to a shape parameter of layer i and denote this derivative by '.

$$\mathbf{W}'(\mathbf{X}^{-1} + \mathbf{X})\mathbf{W}^{-1} + \mathbf{W}((\mathbf{X}^{-1})' + \mathbf{X}')\mathbf{W}^{-1} + \mathbf{W}(\mathbf{X}^{-1} + \mathbf{X})(\mathbf{W}^{-1})'. \quad (6.88)$$

Assume that $\mathbf{\Lambda}$ can be partitioned into m blocks with $m = N$ if all eigenvalues are distinct and $m < N$ if repeated eigenvalues occur.

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & & & \\ & \mathbf{\Lambda}_2 & & \\ & & \ddots & \\ & & & \mathbf{\Lambda}_m \end{bmatrix} \quad \text{where } \mathbf{\Lambda}_k = \lambda_k \mathbf{I}. \quad (6.89)$$

Then matrix \mathbf{C} can be partitioned similarly as $\mathbf{\Lambda}$ such that

$$\mathbf{C}_{kl} = \frac{1}{\lambda_l - \lambda_k} \mathbf{y}_k^* \mathbf{A}' \mathbf{w}_l = \frac{1}{\lambda_l - \lambda_k} \mathbf{\Gamma}_k^{-1} \tilde{\mathbf{Y}}_k^* \mathbf{A}' \tilde{\mathbf{W}}_l \mathbf{\Gamma}_l. \quad (6.90)$$

The equality above only holds for $k \neq l$, thus we need another relation for \mathbf{C}_{kk} . Analyze the first and third term of (6.88) together:

$$\begin{aligned} & \mathbf{W}' (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{W}^{-1} + \mathbf{W} (\mathbf{X}^{-1} + \mathbf{X}) (\mathbf{W}^{-1})' \\ &= \tilde{\mathbf{W}} \mathbf{\Gamma} (\mathbf{C} (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{\Gamma}^{-1} \tilde{\mathbf{W}}^{-1} - \tilde{\mathbf{W}} \mathbf{\Gamma} (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{C} \mathbf{\Gamma}^{-1} \tilde{\mathbf{W}}^{-1}) \\ &= \tilde{\mathbf{W}} \mathbf{\Gamma} (\mathbf{C} (\mathbf{X}^{-1} + \mathbf{X}) - (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{C}) \mathbf{\Gamma}^{-1} \tilde{\mathbf{W}}^{-1}. \end{aligned} \quad (6.91)$$

Consider the following term

$$\begin{aligned} & (\mathbf{C} (\mathbf{X}^{-1} + \mathbf{X}) - (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{C})_{kl} \\ &= \begin{cases} \frac{1}{\lambda_l - \lambda_k} \left(\frac{1}{x_l} - \frac{1}{x_k} - x_l + x_k \right) \mathbf{\Gamma}_k^{-1} \tilde{\mathbf{Y}}_k^* \mathbf{A}' \tilde{\mathbf{W}}_l \mathbf{\Gamma}_l & \text{if } k \neq l \\ \mathbf{0} & \text{if } k = l \end{cases}. \end{aligned} \quad (6.92)$$

Note that $\mathbf{X}_k = x_k \mathbf{I}$ as for the eigenvalues.

When we define a new matrix $\tilde{\mathbf{C}}$ by

$$\tilde{\mathbf{C}}_{kl} := \begin{cases} \frac{1}{\lambda_l - \lambda_k} \tilde{\mathbf{W}}_k^{-1} \mathbf{A}' \tilde{\mathbf{W}}_l & \text{if } k \neq l \\ 0 & \text{if } k = l \end{cases}, \quad (6.93)$$

the matrix $\tilde{\mathbf{C}}$ does not depend in any way on the normalization matrix $\mathbf{\Gamma}$. Take another look at (6.91). We see now that

$$\begin{aligned} & \tilde{\mathbf{W}} \mathbf{\Gamma} (\mathbf{C} (\mathbf{X}^{-1} + \mathbf{X}) - (\mathbf{X}^{-1} + \mathbf{X}) \mathbf{C}) \mathbf{\Gamma}^{-1} \tilde{\mathbf{W}}^{-1} \\ &= \tilde{\mathbf{W}} (\tilde{\mathbf{C}} (\mathbf{X}^{-1} + \mathbf{X}) - (\mathbf{X}^{-1} + \mathbf{X}) \tilde{\mathbf{C}}) \tilde{\mathbf{W}}^{-1}, \end{aligned} \quad (6.94)$$

and clearly the result has no dependence on $\mathbf{\Gamma}$ anymore.

For the second block we need Lemma 6.11. Because the product of $\mathbf{\Gamma}$ and \mathbf{X}' can be rewritten, the dependency of $\mathbf{\Gamma}$ will drop out.

Also for the other blocks of (3.65) it can be shown that the $\mathbf{\Gamma}$ drops out after using Lemma 6.11. This completes the proof. \square

To conclude this chapter, we can state that by deriving a mathematically sound theory for the computation of eigenvalue and eigenvector derivatives, we can formally prove that we can compute sensitivity information in RCWA much less complicated. Furthermore, repeated eigenvalues do not give additional problems if they arise.

6.5.3 Eigenvalue and eigenvector derivatives in RCWA

In Chapter 5 we promised that this chapter provides the formulae for the eigenvalue and eigenvector derivatives. Let $\tilde{\mathbf{W}}$ be the eigenvector matrix computed by the RCWA algorithm. The eigenvalue derivatives λ'_i are given by

$$\lambda'_i = \tilde{\mathbf{y}}_i^* \mathbf{A}' \tilde{\mathbf{w}}_i, \quad i = 1, \dots, N. \quad (6.95)$$

For the eigenvector derivatives, we have to compute \mathbf{C} . Since entries (k, l) of \mathbf{C} for which $\lambda_k = \lambda_l$ are redundant for the derivative computations, let those entries be chosen equal to zero, viz.

$$c_{kl} = \begin{cases} \frac{\tilde{y}_k^* A'(p_0) \tilde{w}_l}{\lambda_l - \lambda_k} & \text{if } \lambda_k \neq \lambda_l, \\ \mathbf{0} & \text{if } \lambda_k = \lambda_l. \end{cases} \quad (6.96)$$

Finally, the eigenvector derivative matrix \mathbf{W}' can be found by

$$\mathbf{W}' = \tilde{\mathbf{W}}\mathbf{C}. \quad (6.97)$$

Chapter 7

Shape parameter reconstruction

Since inversion of the RCWA algorithm is not possible, we have to use optimization methods to reconstruct shape parameters from a measured field. These optimization methods find the shape parameters iteratively by a process as illustrated in Figure 7.1. The process starts with an initial guess of the shape parameters under consideration. One of the assumptions is that we are close enough to the actual solution, such that we can guarantee that the optimization algorithm provides us the true shape parameters. The RCWA algorithm is used to compute the diffraction efficiencies. When these computed diffraction efficiencies are compared with the measured ones and their differences are smaller than a user-defined

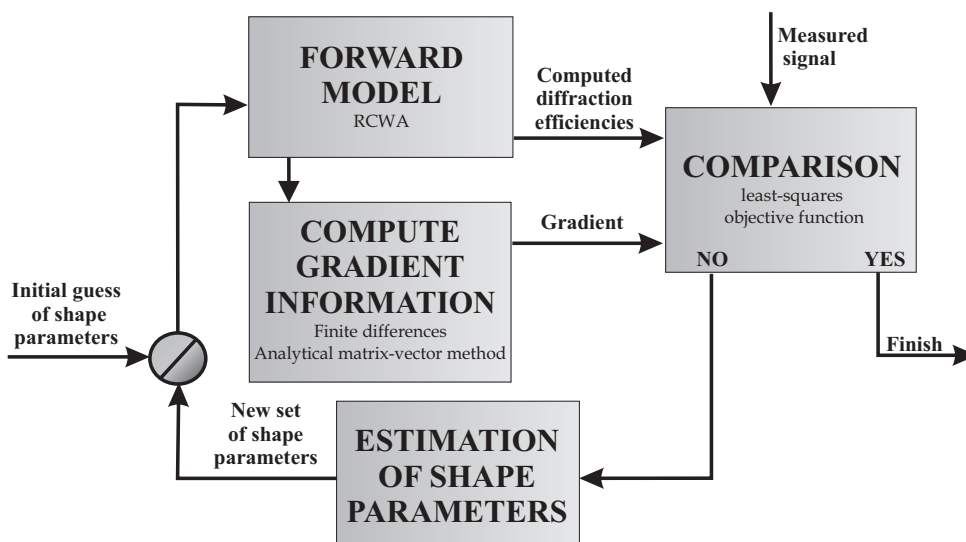


Figure 7.1: Schematic representation of the inverse process.

threshold, we have an approximation of the true shape parameters, assuming a proper formulation of the inverse problem. When the diffraction efficiencies do not match, the initial guess has to be adapted. The first-order derivatives of the field amplitudes with respect to the shape parameters determine the direction of the adaptation. Thus a new estimate of the shape parameters is obtained. The whole procedure is repeated until the computed diffraction efficiencies match the measured ones.

Let the measured field be taken such that we take a set of shape parameters satisfying a physical model and use this set to compute the diffraction efficiencies. This implies that the parameters that are searched for have to match them and the "measured" signal is free of noise. In this chapter we will show that even with this crude assumption, we already encounter some robustness issues.

7.1 Method of inversion

There are many optimization algorithms available and all of them have their advantages and disadvantages. To show what kind of algorithm is best for the diffraction problem described in this thesis, we will compare several of them. The way these algorithms are different from each other is the way they generate steps towards to optimum.

The algorithms are either *line search methods* or *trust-region methods*, which both generate steps with the help of a quadratic model of the objective function. Line search models use this model to generate a search direction, and then find a suitable step length along this direction. Trust-region methods define a region around the current iterate within which they trust the model to be an adequate representation of the objective function, and then choose the step to be the approximate minimizer of the model in this region. If a step is not acceptable, they reduce the size of the region and find a new minimizer.

In the optimization we want to compare the given diffraction efficiencies, which in practice will be measured, with computed diffraction efficiencies using our estimates of the shape parameters as input. Let the *objective function* be given by

$$f(\mathbf{p}) := \|\mathbf{r}(\mathbf{p})\|_2 = \frac{1}{2} \sum_{j=-N}^N r_j(\mathbf{p})^2, \quad (7.1)$$

with

$$r(\mathbf{p}) := \eta_r^{\text{sim}}(\mathbf{p}) - \eta_r^{\text{measured}}. \quad (7.2)$$

Since this objective function is a sum of squares, we speak of *least-squares optimization*. Although the function can be minimized using a general unconstrained minimization technique, certain characteristics of the problem can be exploited to improve the efficiency of the solution procedure. The gradient and Hessian matrix of (7.1) have a special structure.

Denoting the Jacobian matrix of $r(\mathbf{p})$ as $\mathbf{J}(\mathbf{p})$, we have

$$\nabla f(\mathbf{p}) = \sum_{j=1}^m r_j(\mathbf{p}) \nabla r_j(\mathbf{p}) = \mathbf{J}(\mathbf{p})^T \mathbf{r}(\mathbf{p}), \quad (7.3)$$

$$\begin{aligned} \nabla^2 f(\mathbf{p}) &= \sum_{j=1}^m \nabla r_j(\mathbf{p}) \nabla r_j(\mathbf{p})^T + \sum_{j=1}^m r_j(\mathbf{p}) \nabla^2 r_j(\mathbf{p}) \\ &= \mathbf{J}(\mathbf{p})^T \mathbf{J}(\mathbf{p}) + \sum_{j=1}^m r_j(\mathbf{p}) \nabla^2 r_j(\mathbf{p}). \end{aligned} \quad (7.4)$$

We will consider two methods for nonlinear least-squares problems, namely the Gauss-Newton method and the Levenberg-Marquardt method. For comparison we will also look at Nelder-Mead, which is a derivative-free method, steepest descent and the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) method, which is a quasi-Newton method. Without going into subtle details of every method, we would only like to state the basic features.

The *Nelder-Mead method* is a derivative-free method that keeps track of $N + 1$ points in \mathbb{R}^N , whose convex hull forms a simplex. In a single iteration of the algorithm, we remove that vertex (point of the simplex) with the worst function value and replace it with another point with a better value. The new point is obtained by reflecting, expanding, or contracting the simplex along the line joining the worst vertex with the centroid of the remaining vertices. If we cannot find a better point in this manner, we retain only the vertex with the best function value, and we shrink the simplex by moving all other vertices toward this value.

The *steepest-descent method*, *BFGS method* and *Gauss-Newton method* are line search methods. These methods use a Taylor expansion to obtain a search direction p and a step-length parameter α , viz.

$$f(x_k + \alpha p) = f(x_k) + \alpha p^T \nabla f(x_k) + \frac{1}{2} \alpha^2 p^T \nabla^2 f(x_k) p + \mathcal{O}(\alpha^3). \quad (7.5)$$

The steepest-descent method has descent direction $-\nabla f_k$, while the other two methods have descent direction $-(\nabla^2 f_k)^{-1} \nabla f_k$. The last one follows from setting $\alpha = 1$, differentiating (7.5) with respect to p , neglecting the higher order terms of α and setting the result equal to zero. The difference between the BFGS method and the Gauss-Newton method lies in the way they compute the Hessian $\nabla^2 f(x_k)$. The BFGS uses the so-called *secant equation* $B_{k+1}(x_{k+1} - x_k) = \nabla f_{k+1} - \nabla f_k$, which mimics the property $\nabla^2 f_k(x_{k+1} - x_k) \approx \nabla f_{k+1} - \nabla f_k$. The Gauss-Newton method exploits the fact that the objective function is a least-squares problem and therefore uses the special structure of the Hessian as given in (7.4). The Hessian is in this case approximated by $\mathbf{J}_k^T \mathbf{J}_k$. Finally, the step size α has to be computed such that it satisfies the *Wolfe conditions* which ensure that this step size gives a substantial reduction of f .

Finally, we also consider the *Levenberg-Marquardt method*. This method uses the same Hessian approximation as the Gauss-Newton method, but the line search is replaced by a trust-region strategy. For an overview of these and other optimization methods, including

the details, see [42]. We used the implementation of these methods as in the Optimization Toolbox V3.1 of Matlab V7.3 (R2006b).

7.2 Preliminary results

In this section we will only show some preliminary results. The reason for calling them "preliminary" is that we made two assumptions in this thesis, which are not strictly valid in practice. The first assumption was given in Section 1.2, where we assumed that the initial estimates of the shape parameters are close "enough" to the actual shape parameters. With this assumption we ensure that the minimum found by the optimization method under consideration gives the actual shape parameters. Another assumption made here is that we measure the diffraction efficiencies noise-free. In this section we will discuss some issues we encounter in the inverse problem of reconstructing the shape parameters.

Let us first look at the assumption of the initial guess being close enough to the actual shape. Consider testcase 2: the binary grating consisting of silicon (see Table 5.1). Let the incident field have a wave length $\lambda_0 = 0.7\lambda$. When we consider the diffraction efficiencies as a function of the height and width of the binary grating for only one angle of incidence ($\theta = 0$, $\phi = 0$ and $\psi = \pi/2$), we see that there are two optima very close to each other as is illustrated in Figure 7.2.

Figure 7.2 immediately gives insight in the seriousness of noise in our problem. Most likely, any value in the blue bar near the actual shape will be a candidate for a local optimum when noise is considered. A remedy can be to consider more angles of incidence simultaneously. Other angles of incidence give rise to other local minima and combining the diffraction efficiencies for several angles simultaneously, makes it more likely that the minimum we obtain is the actual shape of the grating profile. There are several questions to be asked at this point. How many angles should we take to ensure that the minimum is the actual shape of the grating and which angles should we take? The first question is a fundamental one and depends on the amount of independent information of a measurement. From a mathematical point of view one should take as many angles of incidence as possible to ensure that the minimum is the actual shape. Especially for the comparison with measurements which are not free of noise, taking more points will reduce the influence of noise. However, from a more practical point of view, one would like to use as few angles of incidence as possible to keep the computation time per iteration to a minimum without increasing the number of iterations in the optimization routine. The latter occurs for example when noise is dominant. Therefore, the practical application will determine the number of angles.

The second question involves more insight. Figure 7.4 illustrates the objective function as defined in (7.1) for several settings of the polar angle θ , the azimuthal angle ϕ and the polarization ψ for test case 2 of Table 5.1: the binary grating consisting of silicon. Using angles which are very close to each other only makes sense when the influence of noise has to be reduced, but in this thesis we have no noise. The angles that reduce the number of

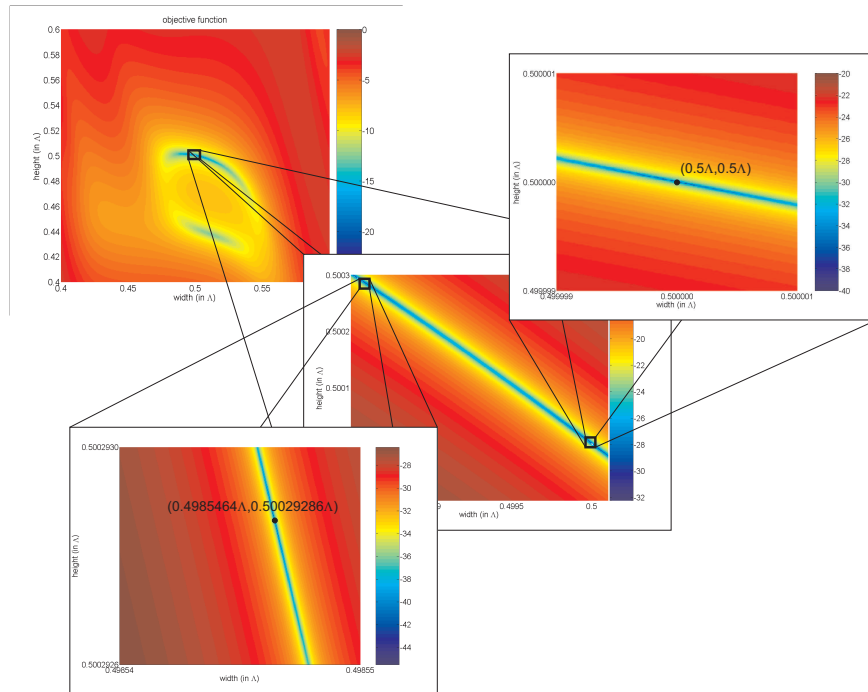


Figure 7.2: Illustration of the diffraction efficiencies as a function of the parameter space of a binary grating for one angle of incidence ($\theta = 0$, $\phi = 0$ and $\psi = \pi/2$) and its local minima.

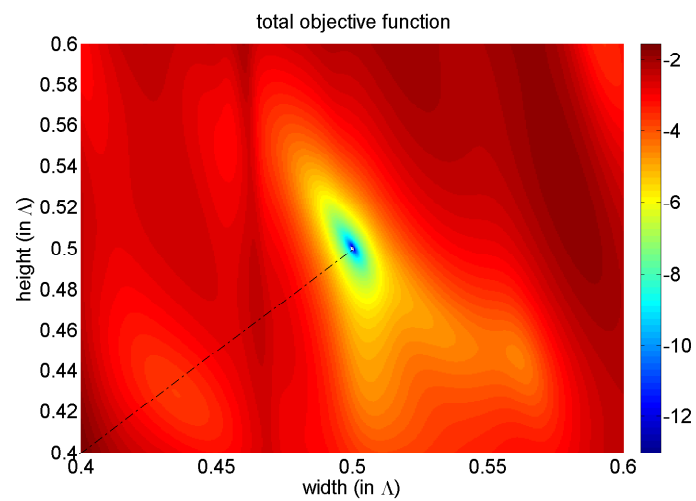


Figure 7.3: Logarithm of the objective function for all angles used in Figure 7.4 as a function of the parameters of the binary diffraction grating and the direction the deviations are admitted for the initial guesses of the shape parameters.

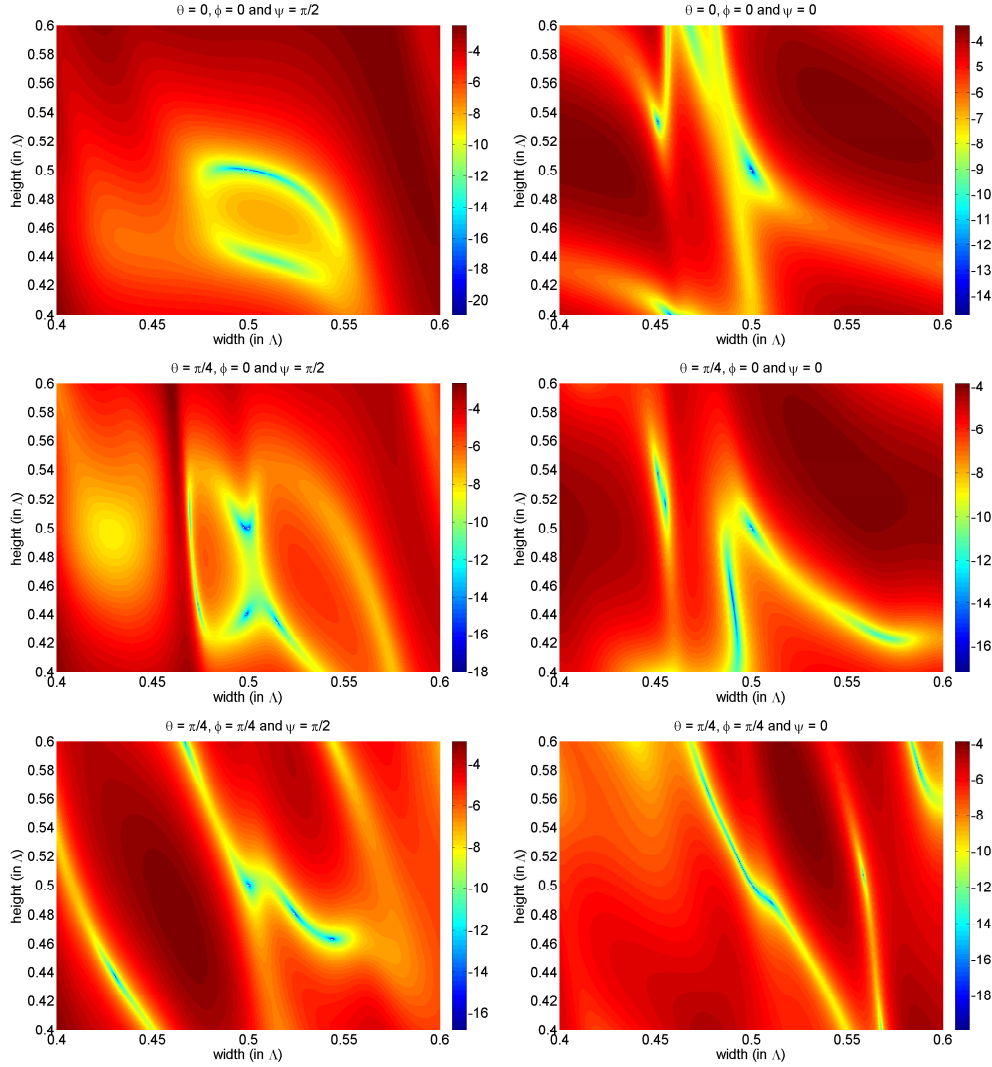


Figure 7.4: Logarithm of the objective function for one angle of incidence at the time as a function of the parameters of the binary diffraction grating.

local minima should be further away from each other. These angles can be chosen such that different diffraction orders are present. From relation (2.59) we can select different areas to choose these angles. The redistribution of the energy will cause the distinction for those angles as is done in Figure 7.4.

In the previous section we discussed five optimization methods, each with its own features. As an example to show the convergence behaviour, we change the initial shape parameter along the line indicated in Figure 7.3. When we apply those methods to the binary grating

structure made of silicon with only one angle of incidence ($\theta = 0$, $\phi = 0$ and $\psi = \pi/2$), we see that even for a deviation of 0.001, all methods except Nelder-Mead converge to another local minimum, since there are two local minima situated very close to each other as illustrated in Figure 7.2. When we take the angles of incidence into account as used in Figure 7.4, all methods converge to the actual shape parameters as long as the deviation is less than 0.01. Besides the convergence to the actual shape parameters, an important feature is how fast a method reaches this convergence. Consider Figure 7.5 where the convergence behaviour of all methods is illustrated when the three angles of incidence and two polarization states from Figure 7.4 are used. The convergence results show what the methods predict: when we are close enough to a minimum there is fast convergence for the Gauss-Newton and Levenberg-Marquardt method, the BFGS method is a little slower, but still faster than steepest descent and the slowest method is the Nelder-Mead method. Also the convergence behaviour of the method when we move further away with our initial guess from the actual shape is intuitive, since the quadratic convergence of Gauss-Newton, Levenberg-Marquardt and BFGS only starts when we are close enough to a minimum.

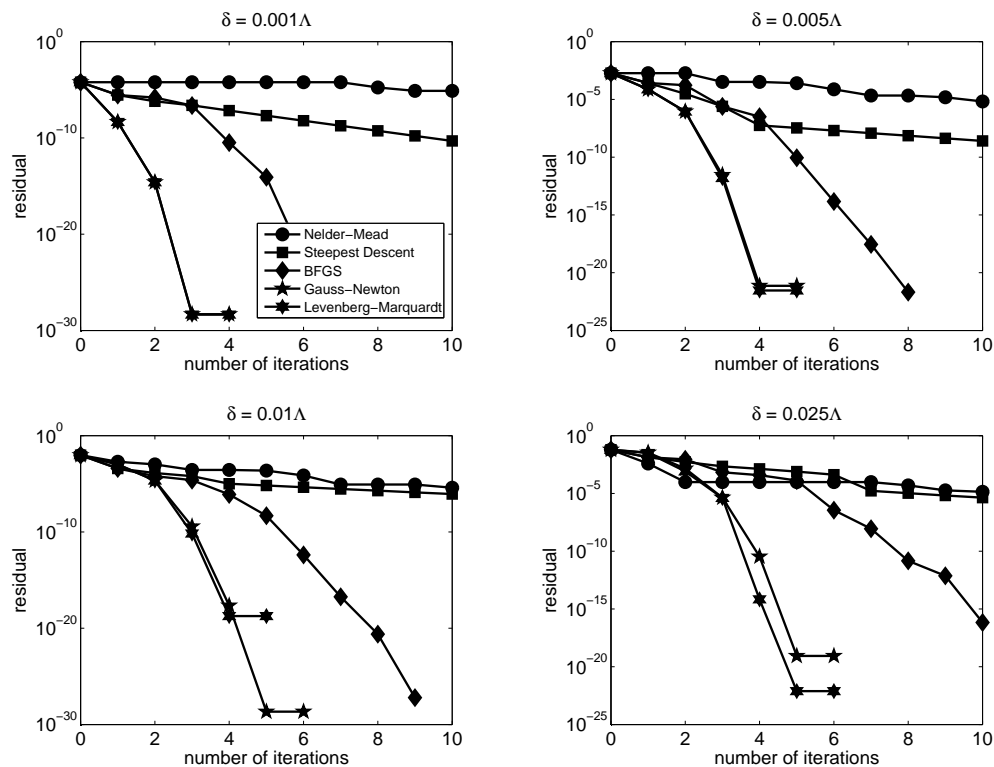


Figure 7.5: Convergence behaviour of the residual for several initial guesses by using the three angles of incidences each with two polarization states of Figure 7.4. The initial guesses are defined by δ such that the width is the actual width - δ and the height is the height - δ .

ing multiple angles does not ensure there are no other local minima, but the more angles one takes into account, the less the change is that two local minima are very close to each other. But if we are too far away from the actual shape, we can have convergence to other values.

In Chapter 5 we derived an analytical method to speed up the computations of the field amplitude derivatives with respect to the grating shape parameters. Since the optimization process requires that for every iteration the derivatives have to be computed with respect to all shape parameters, the analytical approach will decrease the computation time for each iteration drastically. Similar to the results shown in Section 5.4.2, the speed increase compared to finite differences will be more when we have a higher number of harmonics and shape parameters.

Chapter 8

Conclusions and future work

8.1 Conclusions

In the reconstruction of the shape parameters of a grating structure, the computation of the field derivatives with respect to those parameters plays an important role. An intuitive and simple to implement method is finite differences, which recomputes the field with RCWA for a slightly changed shape parameter. However, RCWA incorporates solving eigensystems, which is computationally expensive. In this thesis we provide an alternative method to compute first-order derivatives of the field faster and more accurately. The speed increase is mainly due to avoiding the solution of additional eigensystems. The better accuracy is due to the fact that no additional approximations are made besides those of the RCWA algorithm itself.

The analytical method computes the derivatives of the field with respect to the shape parameters by straightforward differentiating all relations within RCWA, including all eigenvalues and eigenvectors. In general, eigenvectors are not continuously differentiable, because they are not even uniquely determined. In this thesis a mathematical basis is provided for finding that eigenvector that is differentiable irrespective of the eigenvalues being distinct or repeated. Employing this mathematical basis, we show that the normalization of the eigenvectors is redundant in the field derivative computations. As a result we can use the eigenvectors obtained in the RCWA algorithm immediately in our sensitivity theory.

Both finite differences and the analytical approach compute the sensitivity for a discretized grating structure. This implies that the number of layers should be kept the same for both RCWA evaluations in finite differences and in the comparison between both methods to avoid additional errors.

A grating shape is described by a number of physical parameters. For finite differences

every derivative with respect to such a parameter requires at least one additional RCWA evaluation. The analytical method computes the derivatives of the field with respect to the RCWA parameters first and converts them to derivatives with respect to the physical parameters. If a structure is described by many (physical) parameters, the computational effort is significantly lower than that of finite differences. For the binary structure the analytical method computes the derivatives twice as fast as finite differences do, while for a stacked grating consisting of a homogeneous layer with three symmetric trapezoids, the analytical method is more than ten times faster.

Under the assumptions that the initial guess is close enough to actual shape parameters, we are able to reconstruct the shape parameters. To make the optimization process more robust, we have to take more angles of incidence into account to eliminate other local minima. Because of the structure of the objective function, least-squares optimization methods like Gauss-Newton and Levenberg-Marquardt converge fastest to the actual shape parameters (< 5 iterations).

The RCWA algorithm is applicable to any type of grating. The C method shows faster convergence than RCWA for smooth grating profiles, but this method is not applicable to overhanging gratings. The main reason why the C method fails for overhanging gratings is that it does not remove the dependence of the material properties on all other coordinates than the coordinate the periodicity is in. Therefore the homogeneity of the medium cannot be used and the relations of the C method are not solvable.

8.2 Future work

As Arthur Bloch said: "*Every solution breeds new problems*" [9]. This thesis made a start with some of the issues encountered in the inverse diffraction problem. There are still many questions to be answered and many other questions raised during the project. We can categorize the questions under three headings: the forward model, the sensitivity analysis and the optimization procedure.

One major question in the *forward model* is how many terms we have to take into account in the Fourier series of the relative permittivity or its reciprocal and of the field to obtain a sufficiently accurate answer of the field amplitudes. Related to this issue is the question whether we can speed up the RCWA algorithm by performing other expansions as is done in the *adaptive spatial resolution* technique [51]. Another issue is to avoid the assumption that the grating structures are periodic, since damages or misprints can also occur at one location only and gratings have a finite number of periods. For this reason we have to consider aperiodic gratings as well.

Next consider the *sensitivity analysis*. For the reconstruction of parameters, it could be worthwhile to investigate second-order derivatives of the field with respect to the shape parameters. If these second-order derivatives can be computed fast and reliable, it can

improve the convergence rate of the optimization process. Another advantage of having this second-order derivative is that it is an indication of how the field behaves around an optimum. An important grating structure in practice is the *two-dimensional diffraction gratings* which is periodic in two directions. They have already been investigated in the literature and are important for applications in lithography. Since the computation time for two-dimensional gratings grows as the sixth power of the number of harmonics instead of the third power for one-dimensional gratings, the analytical sensitivity theory should be extended to two-dimensional gratings to reduce the computation time drastically.

The major questions are dealing with the *optimization*. Our basic assumption is that our initial guess is "close enough" to the actual solution to guarantee convergence of the solution. We have not given a formal definition of "enough", which requires a good insight in the solution domain. The objective function can have multiple optima and the distances between those optima can be very different. Global optimization should be researched to overcome the assumption of "close enough". A second issue is the objective function, which is chosen such that it has the advantage that it characterizes the optimization as a least-squares optimization problem. But the question whether this is the best one, or whether some diffraction orders are more important than others have not been considered yet. Related to this topic is the number of angles one should take into account for the incident field and their positions to ensure that the local minima in the vicinity of the actual shape parameters are removed. It might be possible that using other wavelengths ensures the actual shape parameters are found. Finally, the simulated diffraction efficiencies are compared to "measured" data. In this thesis we assumed that the measured data is coming from an ideal sensor, that is the data has no noise. Also, our measured data are derived from a shape that satisfies the physical model, but in practice the structure will contain features that are not modeled. Therefore among others stop criteria have to be investigated such as the relative residual (indication of progress in the objective function) and regularization techniques should be used to match the solution to known a priori information.

Appendix A

Mathematical derivations

A.1 Conversion of x - and y - to s - and p -components

A.1.1 Proof of (2.43a)

Let r_{xn} , r_{yn} , r_{zn} be given. Although r_{sn} and r_{pn} are defined already in Section 2.3, we will state them here again.

$$r_{sn} = -\sin \phi_n r_{xn} + \cos \phi_n r_{yn}, \quad (\text{A.1a})$$

$$r_{pn} = -\frac{1}{k_0 n_l} \{ (k_{l,zn} r_{xn} + k_{xn} r_{zn}) \cos \phi_n + (k_y r_{zn} + k_{l,zn} r_{yn}) \sin \phi_n \}. \quad (\text{A.1b})$$

Eliminate r_{zn} from (A.1b) by using (2.39).

$$\begin{aligned} r_{pn} &= -\frac{1}{k_0 n_l} \left((k_{l,zn} r_{xn} + k_{xn} r_{zn}) \cos \phi_n + (k_y r_{zn} + k_{l,zn} r_{yn}) \sin \phi_n \right) \\ &= -\frac{1}{k_0 n_l} \left(\left(k_{l,zn} r_{xn} + k_{xn} \frac{k_{xn} r_{xn} + k_y r_{yn}}{k_{l,zn}} \right) \cos \phi_n \right. \\ &\quad \left. + \left(k_y \frac{k_{xn} r_{xn} + k_y r_{yn}}{k_{l,zn}} + k_{l,zn} r_{yn} \right) \sin \phi_n \right) \\ &= -\frac{1}{k_0 n_l k_{l,zn}} \left(k_{l,zn}^2 r_{xn} \cos \phi_n + k_{xn}^2 r_{xn} \cos \phi_n + k_{xn} k_y r_{yn} \cos \phi_n \right. \\ &\quad \left. + k_y k_{xn} r_{xn} \sin \phi_n + k_y^2 r_{yn} \sin \phi_n + k_{l,zn}^2 r_{yn} \sin \phi_n \right) \end{aligned}$$

$$\begin{aligned}
&\stackrel{\ddagger}{=} -\frac{1}{k_0 n_l k_{l,zn}} \left(k_{l,zn}^2 r_{xn} \cos \phi_n + k_{xn}^2 r_{xn} \cos \phi_n + k_{xn}^2 r_{yn} \sin \phi_n \right. \\
&\quad \left. + k_y^2 r_{xn} \cos \phi_n + k_y^2 r_{yn} \sin \phi_n + k_{l,zn}^2 r_{yn} \sin \phi_n \right) \\
&= -\frac{1}{k_0 n_l k_{l,zn}} \left(k_{xn}^2 + k_y^2 + k_{l,zn}^2 \right) \left(r_{xn} \cos \phi_n + r_{yn} \sin \phi_n \right) \\
&\stackrel{\ddagger}{=} -\frac{k_0 n_l}{k_{l,zn}} \left(r_{xn} \cos \phi_n + r_{yn} \sin \phi_n \right). \tag{A.2}
\end{aligned}$$

For the equality indicated by †, the definition of the angle ϕ_n is used in the identity

$$k_{xn} \sin \phi_n = k_y \cos \phi_n, \tag{A.3}$$

and for ‡ the identity

$$k_0^2 n_l^2 = k_{xn}^2 + k_y^2 + k_{l,zn}^2, \tag{A.4}$$

has been used. This completes the proof of relation (2.43a).

A.1.2 Proof of (2.49)

For notational convenience define A as

$$A := -\frac{k_0 n_l}{k_{l,zn}}. \tag{A.5}$$

For the diffraction efficiency we have to write (2.48) in terms of r_{sn} and r_{pn} . We can do this in parts. The first part is

$$\begin{aligned}
&r_{xn} \bar{r}_{xn} + r_{yn} \bar{r}_{yn} \\
&= \left(-r_{sn} \sin \phi_n + \frac{1}{A} r_{pn} \cos \phi_n \right) \left(-\bar{r}_{sn} \sin \phi_n + \frac{1}{A} \bar{r}_{pn} \cos \phi_n \right) \\
&\quad + \left(r_{sn} \cos \phi_n + \frac{1}{A} r_{pn} \sin \phi_n \right) \left(\bar{r}_{sn} \cos \phi_n + \frac{1}{A} \bar{r}_{pn} \sin \phi_n \right) \\
&= r_{sn} \bar{r}_{sn} \sin^2 \phi_n - \frac{1}{A} r_{sn} \bar{r}_{pn} \sin \phi_n \cos \phi_n - \frac{1}{A} r_{pn} \bar{r}_{sn} \sin \phi_n \cos \phi_n \\
&\quad + \frac{1}{A \bar{A}} r_{pn} \bar{r}_{pn} \cos^2 \phi_n + r_{sn} \bar{r}_{sn} \cos^2 \phi_n + \frac{1}{A} r_{sn} \bar{r}_{pn} \sin \phi_n \cos \phi_n \\
&\quad + \frac{1}{A} r_{pn} \bar{r}_{sn} \sin \phi_n \cos \phi_n + \frac{1}{A \bar{A}} r_{pn} \bar{r}_{pn} \sin^2 \phi_n \\
&= r_{sn} \bar{r}_{sn} + \frac{1}{A \bar{A}} r_{pn} \bar{r}_{pn}. \tag{A.6}
\end{aligned}$$

The part where r_{zn} is in is more involved, since r_{pn} and r_{sn} are given in terms of r_{xn} and r_{yn} only.

$$\begin{aligned}
r_{zn}\bar{r}_{zn} &= \frac{k_{xn}r_{xn} + k_y r_{yn}}{k_{l,zn}} \cdot \frac{k_{xn}\bar{r}_{xn} + k_y \bar{r}_{yn}}{\bar{k}_{l,zn}} \\
&= \frac{k_{xn}^2 r_{xn} \bar{r}_{xn} + k_{xn} k_y r_{xn} \bar{r}_{yn} + k_y k_{xn} r_{yn} \bar{r}_{xn} + k_y^2 r_{yn} \bar{r}_{yn}}{k_{l,zn} \bar{k}_{l,zn}} \\
&= \frac{1}{k_{l,zn} \bar{k}_{l,zn}} \\
&\quad \cdot \left\{ (k_{xn}^2 \sin^2 \phi_n - 2k_y k_{xn} \sin \phi_n \cos \phi_n + k_y^2 \cos^2 \phi_n) r_{sn} \bar{r}_{sn} \right. \\
&\quad + \frac{1}{A} ((k_y^2 - k_{xn}^2) \sin \phi_n \cos \phi_n + k_y k_{xn} (\cos^2 \phi_n - \sin^2 \phi_n)) r_{pn} \bar{r}_{sn} \\
&\quad + \frac{1}{A} ((k_y^2 - k_{xn}^2) \sin \phi_n \cos \phi_n + k_{xn} k_y (\cos^2 \phi_n - \sin^2 \phi_n)) r_{sn} \bar{r}_{pn} \\
&\quad \left. + \frac{1}{A\bar{A}} (k_{xn}^2 \cos^2 \phi_n + 2k_{xn} k_y \sin \phi_n \cos \phi_n + k_y^2 \sin^2 \phi_n) r_{pn} \bar{r}_{pn} \right\} \\
&= \frac{1}{A\bar{A}} \frac{k_{xn}^2 + k_y^2}{k_{l,zn} \bar{k}_{l,zn}} r_{pn} \bar{r}_{pn}. \tag{A.7}
\end{aligned}$$

The latter equality is the result of using identity (A.3). Combining the two result gives

$$\begin{aligned}
&\bar{k}_{l,zn} (r_{xn} \bar{r}_{xn} + r_{yn} \bar{r}_{yn}) + k_{l,zn} r_{zn} \bar{r}_{zn} \\
&= \bar{k}_{l,zn} \left(r_{sn} \bar{r}_{sn} + \frac{1}{A\bar{A}} r_{pn} \bar{r}_{pn} \right) + \frac{1}{A\bar{A}} \frac{k_{xn}^2 + k_y^2}{\bar{k}_{l,zn}} r_{pn} \bar{r}_{pn} \\
&= \bar{k}_{l,zn} r_{sn} \bar{r}_{sn} + \frac{1}{A\bar{A}} \frac{1}{\bar{k}_{l,zn}} (k_{l,zn}^2 + k_{xn}^2 + k_y^2) r_{pn} \bar{r}_{pn} \\
&= \bar{k}_{l,zn} r_{sn} \bar{r}_{sn} + \frac{1}{A\bar{A}} \frac{k_0^2 n_l^2}{\bar{k}_{l,zn}} r_{pn} \bar{r}_{pn}. \tag{A.8}
\end{aligned}$$

Substitution of the definition of A (A.5) in the latter relation completes the proof of (2.49).

A.2 Solving systems of differential equations with constant coefficients

Both RCWA and the C method have to solve a system of differential equations with constant coefficients at some point. In this section a general shape of the solution of such a system will be derived.

A.2.1 1st order differential equations with constant coefficients

Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ be a matrix with constant entries and $\mathbf{s}(z) \in \mathbb{C}^{N \times 1}$ be the column vector we have to solve for which depends on a parameter z . The system under consideration is given by

$$\frac{d}{dz} \mathbf{s}(z) = \mathbf{A} \mathbf{s}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.9})$$

Since \mathbf{A} is a non-defective matrix, it can be diagonalized as $\mathbf{A} = \mathbf{W} \mathbf{\Lambda} \mathbf{W}^{-1}$ with $\mathbf{\Lambda}$ the eigenvalue matrix of \mathbf{A} and \mathbf{W} the eigenvector matrix. Substitution of this diagonalization in (A.9) and premultiplying with the inverse eigenvector matrix \mathbf{W}^{-1} results in

$$\mathbf{W}^{-1} \frac{d}{dz} \mathbf{s}(z) = \mathbf{\Lambda} \mathbf{W}^{-1} \mathbf{s}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.10})$$

Since \mathbf{A} consists of only constant coefficients, also the eigenvalue and eigenvector matrix consists merely of constants. Therefore,

$$\frac{d}{dz} (\mathbf{W}^{-1} \mathbf{s}(z)) = \mathbf{\Lambda} \mathbf{W}^{-1} \mathbf{s}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.11})$$

The next step is to substitute $\mathbf{y}(z) := \mathbf{W}^{-1} \mathbf{s}(z)$ into (A.11), viz.

$$\frac{d}{dz} \mathbf{y}(z) = \mathbf{\Lambda} \mathbf{y}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.12})$$

This relation results in N first-order differential equations

$$\frac{d}{dz} y_k(z) = \lambda_k y_k(z), \quad z_0 \leq z \leq z_1, \quad (\text{A.13})$$

for $k = 1, \dots, N$. Element $y_k(z)$ is the k th element of $\mathbf{y}(z)$ and λ_k is the k th diagonal element of $\mathbf{\Lambda}$. The solution of (A.13) is given by

$$y_k(z) = c \exp[\lambda_k z], \quad k = 1, \dots, N. \quad (\text{A.14})$$

In matrix form this is

$$\mathbf{y}(z) = \mathbf{c} \exp[\mathbf{\Lambda} z]. \quad (\text{A.15})$$

Because we want the shape of \mathbf{s} instead of \mathbf{y} , we have to substitute $\mathbf{s}(z) = \mathbf{W} \mathbf{y}(z)$ again, viz.

$$\mathbf{s}(z) = \mathbf{W} \mathbf{c} \exp[\mathbf{\Lambda} z] = \sum_{k=1}^N \mathbf{w}_k c_k \exp[\lambda_k z]. \quad (\text{A.16})$$

A.2.2 2nd order differential equations with constant coefficients

For a system of second-order differential equations, the system under consideration is given by

$$\frac{d^2}{dz^2} \mathbf{s}(z) = \mathbf{A} \mathbf{s}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.17})$$

Similar as in the first-order differential equations, \mathbf{A} can be diagonalized as $\mathbf{A} = \mathbf{W} \mathbf{\Lambda} \mathbf{W}^{-1}$. Substitution of this diagonalization in (A.17) and premultiplying with the inverse eigenvector matrix \mathbf{W}^{-1} results in

$$\mathbf{W}^{-1} \frac{d^2}{dz^2} \mathbf{s}(z) = \mathbf{\Lambda} \mathbf{W}^{-1} \mathbf{s}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.18})$$

Since \mathbf{A} consists of only constant coefficients, also the eigenvalue and eigenvector matrix consists merely of constants. Therefore,

$$\frac{d^2}{dz^2} (\mathbf{W}^{-1} \mathbf{s}(z)) = \mathbf{\Lambda} \mathbf{W}^{-1} \mathbf{s}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.19})$$

The next step is to substitute $\mathbf{y}(z) := \mathbf{W}^{-1} \mathbf{s}(z)$ into (A.19), viz.

$$\frac{d^2}{dz^2} \mathbf{y}(z) = \mathbf{\Lambda} \mathbf{y}(z), \quad z_0 \leq z \leq z_1. \quad (\text{A.20})$$

This relation results in N second-order differential equations

$$\frac{d^2}{dz^2} y_k(z) = \lambda_k y_k(z), \quad z_0 \leq z \leq z_1, \quad (\text{A.21})$$

for $k = 1, \dots, N$. Element $y_k(z)$ is the k th element of $\mathbf{y}(z)$ and λ_k is the k th diagonal element of $\mathbf{\Lambda}$. The characteristic equations of these second-order differential equations are given by

$$r^2 - \lambda_k = 0, \quad k = 1, \dots, N, \quad (\text{A.22})$$

with $r_1 = \sqrt{\lambda_k}$ and $r_2 = -\sqrt{\lambda_k}$ as its roots. In this thesis we choose the roots with a positive real part. The general solution of (A.21) can be given as

$$y_k(z) = c_1 \exp[\sqrt{\lambda_k} z] + c_2 \exp[-\sqrt{\lambda_k} z], \quad k = 1, \dots, N. \quad (\text{A.23})$$

In matrix form this is

$$\mathbf{y}(z) = \mathbf{c}_1 \exp[\sqrt{\mathbf{\Lambda}} z] + \mathbf{c}_2 \exp[-\sqrt{\mathbf{\Lambda}} z]. \quad (\text{A.24})$$

Because we want the shape of \mathbf{s} instead of \mathbf{y} , we have to substitute $\mathbf{s}(z) = \mathbf{W} \mathbf{y}(z)$ again, viz.

$$\begin{aligned} \mathbf{s}(z) &= \mathbf{W} \left(\mathbf{c}_1 \exp[\sqrt{\mathbf{\Lambda}} z] + \mathbf{c}_2 \exp[-\sqrt{\mathbf{\Lambda}} z] \right) \\ &= \sum_{k=1}^N \mathbf{w}_k \left(c_{1k} \exp[\sqrt{\lambda_k} z] + c_{2k} \exp[-\sqrt{\lambda_k} z] \right). \end{aligned} \quad (\text{A.25})$$

This is the general shape of the solution of the system.

A.3 Properties of the eigenvalues

A complex-valued matrix \mathbf{A} is called *Hermitian* if $\mathbf{A}^* = \mathbf{A}$. For planar diffraction where the field is TE polarized the matrix of which the eigenvalues and eigenvectors are computed is given by (3.13):

$$\mathbf{A}_i = \mathbf{K}_x^2 - \mathbf{E}_i, \quad (\text{A.26})$$

where \mathbf{K}_x is a diagonal matrix with real-valued entries $k_{x,n}/k_0$ on its diagonal for $-N \leq n \leq N$. The matrix \mathbf{E}_i is a Toeplitz matrix of which the entries are given by

$$(\mathbf{E}_i)_{nm} = \epsilon_{i,n-m}, \quad 1 \leq n, m \leq 2N + 1, \quad (\text{A.27})$$

where $\epsilon_{i,nm}$ are the Fourier components of the complex-valued relative permittivity expansion of layer i given by (5.13). To be Hermitian, the off-diagonal entries $\epsilon_{i,m}$ and $\epsilon_{i,-m}$ have to be the complex conjugate of each other for $m = 1, \dots, N$. So, it holds that

$$\begin{aligned} \epsilon_{i,-m}^* &= \left\{ \frac{i}{2\pi(-m)} \tilde{\epsilon}_{i,1}^r \left(\exp[-i\frac{2\pi}{\Lambda}(-m)t_1] - \exp[i\pi(-m)] \right) \right. \\ &\quad + \frac{i}{2\pi(-m)} \tilde{\epsilon}_{i,2}^r \left(\exp[-i\frac{2\pi}{\Lambda}(-m)t_2] - \exp[-i\frac{2\pi}{\Lambda}(-m)t_1] \right) + \dots \\ &\quad \left. + \frac{i}{2\pi(-m)} \tilde{\epsilon}_{i,M}^r \left(\exp[-i\pi(-m)] - \exp[-i\frac{2\pi}{\Lambda}(-m)t_{M-1}] \right) \right\}^* \\ &= \frac{i}{2\pi m} (\tilde{\epsilon}_{i,1}^r)^* \left(\exp[-i\frac{2\pi}{\Lambda}mt_1] - \exp[i\pi m] \right) \\ &\quad + \frac{i}{2\pi m} (\tilde{\epsilon}_{i,2}^r)^* \left(\exp[-i\frac{2\pi}{\Lambda}mt_2] - \exp[-i\frac{2\pi}{\Lambda}mt_1] \right) + \dots \\ &\quad + \frac{i}{2\pi m} (\tilde{\epsilon}_{i,M}^r)^* \left(\exp[-i\pi m] - \exp[-i\frac{2\pi}{\Lambda}mt_{M-1}] \right) \\ &= \epsilon_{i,m} \text{ if } (\tilde{\epsilon}_{i,p}^r)^* = \tilde{\epsilon}_{i,p}^r \text{ for } p = 1, \dots, M \end{aligned} \quad (\text{A.28})$$

Since the relative complex permittivity is related to the refractive index for a non-magnetic medium as $\tilde{\epsilon}_{i,p}^r = n_{i,p}^2$, the refractive index may have a real or purely imaginary value to ensure $\tilde{\epsilon}_{i,p}^r$ to be real-valued.

Another feature the matrix has to obey to be Hermitian, is to have real diagonal entries. The terms for which $m = 0$ from equation (5.13) imply that the diagonal elements are real only if all materials have a real relative permittivity or, in other words, are dielectric.

For planar diffraction where the field is TM polarized, the matrix of which the eigenvalues and eigenvectors are computed is given by $\mathbf{P}_i^{-1}\mathbf{B}_i$ of which \mathbf{B}_i is given by (3.21)

$$\mathbf{B}_i = (\mathbf{K}_x \mathbf{E}_i^{-1} \mathbf{K}_x - \mathbf{I}). \quad (\text{A.29})$$

and the matrix \mathbf{P}_i is a Toeplitz matrix similar to \mathbf{E}_i of which the entries are given by

$$(\mathbf{P}_i)_{nm} = \pi_{i,n-m}, \quad 1 \leq n, m \leq 2N + 1. \quad (\text{A.30})$$

In a similar way, we can show that \mathbf{P}_i is Hermitian. However, the matrix of the eigensystem consists of a product of two Hermitian matrices and in general such a product is not Hermitian anymore. However, we can prove that the eigenvalues of a product of two Hermitian matrices have real-valued eigenvalues.

Let \mathbf{A} and \mathbf{B} be Hermitian matrices. It holds that

$$\mathbf{ABw} = \lambda \mathbf{w}. \quad (\text{A.31})$$

Taking the conjugate transpose of both sides, we have

$$\mathbf{w}^* \mathbf{B}^* \mathbf{A}^* = \mathbf{w}^* \lambda^* \Rightarrow \mathbf{w}^* \mathbf{BA} = \lambda^* \mathbf{w}^*. \quad (\text{A.32})$$

By premultiplying (A.31) by \mathbf{w}^* and postmultiply (A.32) by \mathbf{w} , we have

$$\lambda \mathbf{w}^* \mathbf{w} = \lambda^* \mathbf{w}^* \mathbf{w}. \quad (\text{A.33})$$

Therefore it holds that $\lambda = \lambda^*$ and thus $\lambda \in \mathbb{R}$.

Bibliography

- [1] Nico van der Aa. RCWA sensitivity computations of the scattered field with respect to grating shape parameters. In *EOS Topical Meeting, Diffractive Optics 2005, Warsaw, Poland, September 3-7, 2005*, pages 17–18, 2005.
- [2] N.P. van der Aa. Diffraction grating theory with RCWA or the C method. In *Progress in Industrial Mathematics at ECMI 2004*, pages 99–103. Springer, 2006.
- [3] N.P. van der Aa and R.M.M. Mattheij. Computing shape parameter sensitivity of the field of 1D surface-relief gratings by using an analytical approach based on RCWA. *J. Opt. Soc. Am. A*, 24(9):2692–2700, September 2007.
- [4] N.P. van der Aa, H.G. ter Morsche, and R.R.M. Mattheij. Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem. *Electronic Journal of Linear Algebra*, 16:300–314, October 2007.
- [5] E. Anderson *et al.* *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, third edition, 1999.
- [6] A.L. Andrew. Convergence of an iterative method for derivatives of eigensystems. *Journal of Computational Physics*, 26:107–112, 1978.
- [7] A.L. Andrew and R.C.E. Tan. Computation of derivatives of repeated eigenvalues and the corresponding eigenvectors of symmetric matrix pencils. *SIAM J. Matrix Anal. Appl.*, 20:78–100, 1998.
- [8] J. Bardeen and W. Brattain. *Three-electrode circuit element utilizing semiconductor materials*. U.S. Patent 2,524,035, October 1950.
- [9] Arthur Bloch. *The Complete Murphy's Law*. Price Stern Sloan, 1991.
- [10] Max Born and Emil Wolf. *Principles of Optics*. Cambridge University Press, 7th (expanded) edition, 2002.
- [11] G. Bouwhuis and S. Wittekoek. Automatic alignment system for optical projection printing. *IEEE Transactions on Electron Devices*, 26(4):723–728, April 1979.
- [12] J. Chandezon, M.T. Dupuis, G. Cornet, and D. Maystre. Multicoated gratings: a differential formalism applicable in the entire optical region. *J. opt. Soc. Am.*, 72(7):839–846, July 1982.
- [13] J. Chandezon, D. Maystre, and G. Raoult. A new theoretical method for diffraction gratings and its numerical application. *Journal of Optics (Paris)*, 11(4):235–241, 1980.
- [14] Patrick M. Fitzpatrick. *Advanced Calculus - A Course in Mathematical Analysis*, chapter 17, pages 399–427. PWS Publishing Company, 1995.
- [15] M.I. Friswell. The derivatives of repeated eigenvalues and their associated eigenvectors. *Journal of Vibration and Acoustics*, 118:390–397, 1996.

- [16] Gérard Granet and Jean Chandezon. The method of curvilinear coordinates applied to the problem of scattering from surface-relief gratings defined by parametric equations: application to scattering from a cycloidal grating. *Pure Appl. Opt.*, 6:727–740, 1997.
- [17] Gérard Granet, Jean Chandezon, and Jean-Pierre Plumey. Reformulation of the coordinate transformation method through the concept of adaptive spatial resolution. Application to trapezoidal gratings. *J. Opt. Soc. Am. A*, 18(9):2102–2108, September 2001.
- [18] Michael T. Heath. *Scientific Computing: An Introductory Survey*. McGraw-Hill, second edition, 2002.
- [19] Eugene Hecht. *Optics*. Addison Wesley, 4th edition, 2002.
- [20] Christiaan Huygens. *Traité de la Lumière*. 1690.
- [21] Sun Ji-Guang. Eigenvalues and eigenvectors of a matrix dependent on several parameters. *Journal of Computational Mathematics*, 3(4):351–364, October 1985.
- [22] John D. Kraus and Daniel A. Fleish. *Electromagnetics with Applications*. McGraw-Hill, fifth edition, 1999.
- [23] Peter D. Lax. *Linear Algebra*. John Wiley & Sons, 1997.
- [24] Harry J. Levinson. *Principles of Lithography*. The Society of Photo-Optical Instrumentation Engineers, 2nd edition, 2005.
- [25] Lifeng Li. Use of Fourier series in the analysis of discontinuous periodic structures. *J. Opt. Soc. Am. A*, 13(9):1870–1876, September 1996.
- [26] Lifeng Li. Justification of matrix truncation in the modal methods of diffraction gratings. *J. Opt. A: Pure Appl. Opt.*, 1:531–536, 1999.
- [27] Lifeng Li. Oblique-coordinate-system-based Chandezon method for modeling one-dimensionally periodic, multilayer, inhomogeneous, anisotropic gratings. *J. Opt. Soc. Am. A*, 16(10):2521–2531, October 1999.
- [28] Lifeng Li. *Mathematical reflections on the Fourier modal method in grating theory*, chapter 4, pages 111–139. *Mathematical Modeling in Optical Science*. SIAM, 2001.
- [29] Lifeng Li and Jean Chandezon. Improvement of the coordinate transformation method for surface-relief gratings with sharp edges. *J. Opt. Soc. Am. A*, 13(11):2247–2255, November 1996.
- [30] Lifeng Li, Jean Chandezon, Gérard Granet, and Jean-Pierre Plumey. Rigorous and efficient grating-analysis method made easy for optical engineers. *Applied Optics*, 38(2):304–313, January 1999.
- [31] Lifeng Li, G Granet, J P Plumey, and J Chandezon. Some topics in extending the C method to multilayer gratings of different profiles. *Pure and Applied Optics*, 5(2):141–156, March 1996.
- [32] Lifeng Li and Charles W. Haggans. Convergence of the coupled-wave method for metallic lamellar diffraction gratings. *J. Opt. Soc. Am. A*, 10(6):1184–1189, June 1993.
- [33] R.M.M. Mattheij and J. Molenaar. *Ordinary Differential Equations in Theory and Practice*. SIAM Society for Industrial & Applied Mathematics, 2002.
- [34] W.C. Mills-Curran. Calculation of eigenvector derivatives for structures with repeated eigenvalues. *AIAA Journal*, 26:867–871, 1988.
- [35] M.G. Moharam. Rigorous coupled-wave analysis of planar-grating diffraction. *J. Opt. Soc. Am.*, 71(7):811–818, July 1981.

- [36] M.G. Moharam, E.B. Grann, D. A. Pommet, and T.K. Gaylord. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis for binary gratings. *J. Opt. Soc. Am. A*, 12(5):1068–1076, May 1995.
- [37] M.G. Moharam, E.B. Grann, D. A. Pommet, and T.K. Gaylord. 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.
- [38] Gordon E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38(8), April 1965.
- [39] D.V. Murthy and R.T. Haftka. Derivatives of eigenvalues and eigenvectors of a general complex matrix. *International Journal For Numerical Methods in Engineering*, 26:293–311, 1988.
- [40] R.B. Nelson. Simplified calculation of eigenvector derivatives. *AIAA Journal*, 14:1201–1205, 1976.
- [41] Michel Nevière and Evgeny Popov. *Light Propagation in Periodic Media - Differential Theory and Design*. Marcel Dekker, Inc., 2003.
- [42] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer, second edition, 2006.
- [43] R. Petit. *Electromagnetic Theory of Gratings*. Topics in Current Physics. Springer-Verlag Berlin Heidelberg, 1980.
- [44] J.P. Plumey and G. Granet. Generalization of the coordinate transformation method with application to surface-relief gratings. *J. Opt. Soc. Am. A*, 16(3):508–516, March 1999.
- [45] J.P. Plumey, B. Guizal, and J. Chandezon. Coordinate transformation method as applied to asymmetric gratings with vertical facets. *J. Opt. Soc. Am. A*, 14(3):610–617, March 1997.
- [46] Edward J. Rothwell and Michael J. Cloud. *Electromagnetics*. CRC Press, 2001.
- [47] C.S. Rudisill and Y.-Y. Chu. Numerical methods for evaluating the derivatives of eigenvalues and eigenvectors. *AIAA Journal*, 13:834–837, 1975.
- [48] Matthew N.O. Sadiku. *Elements of Electromagnetism*. Oxford University Press, third edition, 2001.
- [49] W. Shockley. *Semiconductor amplifier*. U.S. Patent 2,502,488, April 1950.
- [50] A. Sommerfeld. *Partial Differential Equations in Physics*. Academic Press, New York, 1949.
- [51] T. Vallius and M. Honkanen. Reformulation of the Fourier modal method with adaptive spatial resolution: application to multilevel profiles. *Optics Express*, 10(1):24–34, January 2002.
- [52] S. Wolf. *Microchip Manufacturing*. Lattice Press, 2004.
- [53] Yuhang Zheng. Extension of the C Method for Overhanging Gratings. Master's thesis, Technische Universiteit Eindhoven, March 2006.

Index

- p -component, 13
- s -component, 12
- C method, 47–60

- adaptive spatial resolution, 116
- alignment marker, 5
- Ampère's law, 14, 15
- angular temporal frequency, 16
- avoidance of crossing, 100
- azimuthal angle, 12, 17, 72–73, 110

- BARC, 3
- BFGS method, 109
- bijective transformation, 60
- boundary condition
 - continuity at the interface, 19, 36–40, 52–57
 - pseudo-periodic, 18, 50
- concurrent jump discontinuity, 29
 - complementary, 30
- conductivity, 15
- conical diffraction, 12
- constitutive relations, 15
- coordinate transformation method, 47

- diffraction, 12
 - efficiency, 23–26, 66
 - grating, 4
 - asymmetric trapezoidal, 69
 - binary, 4, 69
 - coated, 69
 - sinusoidal, 57
 - stacked, 69
 - order, 5
 - plane, 22
- dispersion-free, 15
- dose, 6

- eigenvalue derivative, 83–105
- eigenvector derivative, 83–105
- eigenvector normalization, 84
- enhanced transmittance matrix approach, 43
- evanescent diffraction order, 26

- Faraday's law, 14, 15
- field, 12
 - incident, 17–18
 - reflected, 18
 - transmitted, 18
- field amplitudes
 - reflected, 21
 - transmitted, 21
- field components
 - electric, 29
 - magnetic, 29
- finite differences, 61–62
 - central, 62
 - forward, 62
- Floquet condition, 19
- focus, 6
- forward problem, 8
- Fourier series, 20

- Gauss's law
 - electric field, 14
 - magnetic field, 14
- Gauss-Newton method, 109
- grating equation, 13

- harmonics, 27
- Helmholtz equation, 16, 49–50
- Hermitian matrix, 124
- homogeneous, 48
- Huygens's principle, 12
- Huygens-Fresnel principle, 12

- Implicit Function Theorem, 85
- integrated circuit (IC), 1
- interference, 12
- inverse problem, 8
- inverse rule, 30

- isotropic, 15
- Jacobian, 85
- Kirchhoff
 - approximation, 14
 - scalar diffraction theory, 13
- Laurent's rule, 30
- least-squares optimization, 108
- left-eigenvector matrix, 84
- Leibniz' rule, 91
- Levenberg-Marquardt method, 109
- line search method, 108
- linearity assumption, 15
- lithography, 3
- mask, 4
- matrix-matrix approach, 64
- matrix-vector approach, 64
- Maxwell's equations, 14–17
- Moore's law, 2
- negative resist, 3
- Nelder-Mead method, 109
- non-defective, 83
- non-magnetic, 15
- objective function, 108
- one-dimensional diffraction grating, 11
- optimization, 8
- outgoing wave condition, 18
- overlay, 4
- perfect electric conductor (PEC), 57
- permeability, 15
- permittivity, 15
- phase grating, 4
- photoresist, 3
- planar diffraction, 12
- plane of incidence, 12
- polar angle, 12, 17, 72–73, 110
- polarization
 - TE, 17
 - TM, 18
- polarization angle, 12, 17, 110
- positive resist, 3
- power, 24
- Poynting vector, 24
- propagating diffraction order, 26
- Rayleigh expansion, 20–23
- RCWA, 27–45
- reflected field, 18
- relative permittivity, 16
- resolution, 5
- secant equation, 109
- source-free, 15
- spatial filtering, 5
- steepest-descent method, 109
- TE polarization, 17
- time-harmonic, 12, 15
- time-invariant, 15
- TM polarization, 18
- transistor, 1
- transmitted field, 18
- transverse electric wave, 12
- transverse magnetic wave, 12
- trust-region method, 108
- two-dimensional diffraction grating, 117
- wafer, 1
- Wolfe conditions, 109

Summary

Periodic structures, called diffraction gratings, play an important role in optical lithography. The diffraction of the incident field in multiple diffraction orders provides a way to accurately determine a position on a wafer on one hand and on the other hand it provides a test method to determine the quality of the photolithographic process. For both applications it is crucial to be able to find the actual shape of the structure to correct for damages or imperfections.

When besides the incident field also the shape of a diffraction grating is known, we can compute the diffracted field by using the rigorous coupled-wave analysis (RCWA) or the C method. These methods solve Maxwell's equations for time-harmonic fields directly, which is required because such a grating typically has a period smaller than the wavelength of the incident field. The basic idea of both methods is that they transform Maxwell's equations into algebraic eigensystems, which have to be solved in order to obtain the diffracted field.

The reconstruction of the grating shape is carried out by first making an initial guess of its shape. Next the computed diffracted field is compared to actual measurements and the difference between them determines how the shape parameters should be adjusted. For the reconstruction we make use of standard optimization techniques such as quasi-Newton methods to find local optima. We assume that the initial guess of the grating shape is close enough to its actual shape such that the optimum that is found is the actual shape and take more angles of incidence to make the optimization more robust.

The focus of this thesis is finding the first-order derivative information of the diffracted field with respect to the shape parameters. This is possible using finite differences where the diffracted field is computed again for a neighbouring value of the shape parameter under consideration. However, straightforward differentiation of the relations within RCWA or the C method gives a more accurate, but also faster way to find this derivative information. When straightforward differentiation is used, we also have to find eigenvalue and eigenvector derivatives, but to determine these derivatives no additional eigenvalue systems have to be solved. This implies that the reconstruction process can be performed faster and more accurate. Besides the speed-up of the reconstruction, we also provide a firm mathematical basis to this sensitivity theory.

The sensitivity of RCWA is tested for some specific grating structures, such as the binary grating, the trapezoidal grating and more advanced structures as the coated trapezoid and a stacked grating of multiple trapezoids. The simulations show that for the most simple structure, the binary grating, we have the derivatives with respect to shape parameters up to twice as fast as obtained with finite differences, depending on the truncation number of the Fourier series. When the number of physical shape parameters increases, the analytical method becomes increasingly faster than finite differences. For the stacked trapezoids, the analytical method is more than 10 times faster than finite differences. In practice, the grating shapes will be more and more complex and therefore, the analytical approach offers a more and more significant speed increase in the computations of the derivatives without loss of accuracy.

Samenvatting

Diffraactieroosters zijn periodieke structuren die een grote rol spelen in de optische lithografie. Doordat het rooster het invallende veld verstrooit over meerdere diffractie-orde, geeft het aan de ene kant een zeer nauwkeurige manier om een positie op een wafer te bepalen en aan de andere kant de mogelijkheid om de kwaliteit van het resultaat van de lithografie stap te kwantificeren. Voor beide applicaties is het belangrijk dat de vormparameters nauwkeurig gereconstrueerd kunnen worden om imperfecties of beschadigingen te identificeren.

Als behalve het invallende veld ook de vorm van het diffractierooster bekend is, kunnen we het verstrooide veld berekenen met behulp van het RCWA-algoritme of de C-methode. Beide methoden lossen rechtstreeks de Maxwell vergelijkingen voor tijdharmonische velden op, omdat de roosters typisch een periode hebben die kleiner is dan de golflengte van het invallende veld. Het basis idee van beide methoden is dat de Maxwell vergelijkingen getransformeerd worden in een algebraïsch eigenwaarde probleem dat opgelost moet worden om het verstrooide veld te berekenen.

De reconstructie van het diffractierooster begint met het maken van een initiële schatting van de vorm. Vervolgens wordt het verstrooide veld berekend en vergeleken met het gemeten verstrooide veld. Het verschil tussen beiden bepaald hoe de initiële schatting van de vormparameters aangepast dient te worden. Voor het reconstrueren van de vormparameters maken we gebruik van standaard optimalisatie routines als het quasi-Newton algoritme om de lokale minima te vinden. We veronderstellen dat de initiële vormparameters dicht genoeg bij de echte vormparameters zitten zodanig dat het gevonden optimum ook de daadwerkelijke vormparameters zijn. Verder nemen we meerdere hoeken van inval mee om de optimalisatie meer robuust te maken.

Het belangrijkste deel van dit proefschrift gaat over het berekenen hoe het verstrooide veld reageert op kleine veranderingen in de vormparameters, d.w.z. het vinden van de eerste orde afgeleiden van het verstrooide veld naar de vormparameters. Deze afgeleiden kunnen berekend worden door middel van eindige differenties waar het verstrooide veld nogmaals wordt berekend voor een iets gewijzigde parameter en het verschil tussen beide verstrooide velden geeft informatie over de afgeleide met betrekking tot deze parameter. Echter, door rechtstreeks de formules in het RCWA-algoritme te differentiëren, kunnen we de afgeleiden nauwkeuriger en sneller berekenen. Rechtstreeks differentiëren wil ook zeggen dat we

afgeleiden van de eigenwaarden en eigenvectoren moeten berekenen. Dit kan zonder het oplossen van extra eigenwaarde systemen als in eindige differenties. Dit betekent dat de reconstructie van de vormparameters sneller en nauwkeurig kan worden gedaan. Naast de snelheids- en nauwkeurigheidswinst, is hiermee ook een wiskundige basis voor deze methode gelegd.

De gevoeligheid van het RCWA-algoritme is getest aan de hand van specifieke diffractieroosters, zoals binaire roosters, trapeziumvormige roosters en meer geavanceerde structuren als gelaagde trapezia of opgestapelde trapezia. De simulaties laten zien dat voor de meest simpele structuur, het binaire rooster, we de eerste orde afgeleiden met betrekking tot de vormparameters tot twee keer zo snel kunnen berekenen met onze analytische methode dan met eindige differenties, afhankelijk van het aantal termen in de Fourierreeksen. Als het aantal fysische parameters toeneemt, wordt de analytische methode relatief sneller vergeleken met eindige differenties. Voor het samengestelde rooster waarbij elke periode uit een opstapeling van trapezia bestaat, is de analytische methode meer dan 10 keer sneller dan eindige differenties. In de praktijk zullen de diffractieroosters steeds complexer worden en voor deze gevallen zal de analytische methode een steeds hogere snelheidswinst opleveren in het berekenen van de eerste orde afgeleiden zonder verlies in nauwkeurigheid.

Acknowledgement

It has already been 10 years ago that I started my study of mathematics at the Technische Universiteit Eindhoven. During my study I always said I would never become a PhD student. However, during my master's project at ASML I realized that I had the wrong image of being a PhD student. As a result I started my PhD project at ASML in the beginning of 2003. During my master's and my PhD project, many people supported me in both personal and professional point of view. I would like to take the opportunity to thank them all and some of them in particular.

First of all I would like to thank my promotor professor Bob Mattheij for giving me the opportunity to show and improve my qualities in the CASA group (Centre of Analysis, Scientific computing and Applications). I would like to thank him and my copromotor Hennie ter Morsche for their guidance during both my master's and PhD project. If I look back at my time in CASA and think about the person I asked the most questions, it would definitely be Enna van Dijk, who was always very kind and helpful.

My gratitude goes out to ASML, world leader in the manufacturing of advanced machinery for the semiconductor industry, who fully financed both my master's and my PhD project. It was very pleasant to work in a hightech, multidisciplinary environment with many smart and helpful colleagues. I would like to thank especially my supervisors Arie den Boef and Irwan Setija for their clear explanations of the physics behind the project and their critical remarks.

My PhD project was not a stand-alone project. Special thanks go out to my colleague PhD student Mark van Kraaij for the helpful discussions and advices. Also Jos Maubach, Marc Noot and Jurgen Tas, who were all staff members of CASA, participated in this project for ASML. I would like to thank them for their help and interest.

Of course a PhD project cannot be finished without a reading committee. I would like to thank professors Anton Tjihuis, Paul Urbach and Joseph Braat for their time and their efforts to improve the contents of my thesis.

Besides the scientific research, my time as a PhD student was also the time many things changed in my personal life. The most important change I am very grateful for is that I became (and still are) a proud father of two boys, Aaron en Bram. Whenever I enter my

home after a day of hard work, they come to me smiling and give me a hug. Those are the moments you know it is all worth it. Last but not least, I would also take the opportunity to thank my Cindy for all the special moments we spend together and the support she gave me when I needed it.

Curriculum vitae

Nico van der Aa was born June 13th, 1979, in Helmond, the Netherlands. After finishing his preuniversity education at the Dr. Knippenbergcollege (Helmond) in 1997, he started studying applied mathematics at Eindhoven University of Technology. With a specialization in numerical mathematics, he obtained his Master's title in 2003. The final project was carried out for ASML, world leader in the manufacturing of advanced technology systems for the semiconductor industry, where a statistical method was investigated to improve the position estimates of diffraction gratings on a wafer such that the quality of a lithography step can be increased. The resulting thesis, written under the supervision of prof.dr. R.M.M. Mattheij and dr. H.G. ter Morsche, was titled *A basis for the Order-to-Order method - Theory and validation*. Because of the pleasant interaction with ASML and his supervisors, Nico joined CASA (Centre of Analysis, Scientific Computing and Applications) from 2003 to 2007 and started on the same problem as in his Master's project, but used an electromagnetic approach instead.



