

Optimization of a 3-beam splitter by means of the Rigorous Coupled Wave Analysis

Astrid Averkamp
0157635
April 29, 2011

Supervisors:

dr. M. Hammer
dr. M. Maksimovic

Committee:

prof. dr. ir. E.W.C. van Groesen
dr. M. Maksimovic
dr. M. Hammer
dr. M.A. Bochev

UNIVERSITEIT TWENTE.

Applied Mathematics

Chair AAMP

Summary

This master thesis introduces and explains an algorithm called the Rigorous Coupled Wave Analysis. It was developed in the early 80ies to simulate electromagnetic waves meeting a periodic grating structure. Foundation are the Maxwell equations: a set of partial differential equations describing the behaviour of electrical and magnetical fields in space and time. These are to be solved in the frequency domain.

The periodic grating is approximated by means of a Fourier expansion. We use three different ansätze: One for the incoming and reflected part of the field, one for the grating area and the last one for the transmitted waves. These are inserted in the Maxwell equations, resulting in the coupled wave equations for the grating area. This system of second order differential equations can be solved by means of its eigenvalues. At the end the reflected and transmitted fields can be computed by the continuous interface conditions. The details of the mathematical formulation of the algorithm are explained, followed by the stable implementation in Matlab. The program is now able to calculate the fields of binary multi-layer systems with incoming waves of TE- and TM-polarization.

This algorithm is then used to design a periodic grating in such a way, that an incoming laser beam is split into three fractions of equal power. The design parameters are optimized until the relative output power approaches $1/3$ for each diffraction order. An almost perfect configuration was found, but it appeared to be difficult to produce, due to fine, highly relevant details. Therefore also gratings with slightly worse performances but more realistic parameters are presented.

Contents

1	Introduction	3
2	The Rigorous Coupled Wave Analysis	5
2.1	TE polarization	6
2.2	TM polarization	10
2.3	Multilayer systems	11
3	Benchmarking results	15
3.1	TE polarization	15
3.2	TM polarization	20
3.3	Multilayer systems	22
4	Design of a three-beam splitter	26
4.1	Determination of the number of orders	26
4.2	Optimization	28
5	Conclusion	33

1 Introduction

When modeling the propagation of light one has to distinguish two different approaches. One is called geometrical optics and approximates light as rays. If we look at light on a smaller scale, for example when we look at two interfering light sources, one observes a pattern that is explainable by a wave nature. Optics dealing with this is called diffractive. As in this thesis we are considering a grating on the micrometer scale, we are interested in the latter. For that, we will need to solve the Maxwell's equations, a set of partial differential equations which describe the electric and magnetic parts of the optical fields. In diffraction optics, there are certain devices called Diffractive Optical Elements (DOE). They are used, for example, to form laser beams. This is to be seen in figure 1, a beam propagates through an arbitrary DOE and diffracted light comes out. To predict the actual influence of such an element on the beamshape we have to make use of computational simulation tools. Common methods are the Finite Element method (FEM) [24], Finite Difference Time Domain (FDTD) [2] and the Rigorous Coupled Wave Analysis (RCWA) [1]. In this thesis we focus on the RCWA. It is stable and has been used successfully now since many years. The method is especially meant for periodic structures.

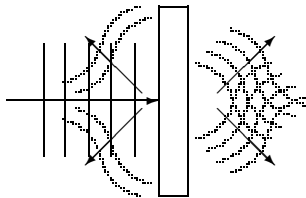


Figure 1: A Diffractive Optical Element. Light waves come in from the left side and the device changes the shape of the reflected and transmitted waves.

The Rigorous Coupled Wave Analysis was first developed by M. G. Moharam and T. K. Gaylord in the early 80ies. The algorithm is based on the assumption of an infinite and periodic structure. The Maxwell equations of classical electrodynamics are to be solved. The relative permittivity and the electric and magnetic fields are approximated by a Fourier series. The amplitudes of these fields satisfy certain coupled-wave equations. The latter are obtained by combining the ansatz of the fields with the Maxwell equations. This leads to a system of differential equations of second order with constant coefficients, which can be solved by means of their eigenvalues. The result describes the diffracted field, i.e. the reflected and transmitted waves of the grating.

Reference [1] explains the algorithm step by step. This article is cited by many other papers about this topic. Two of the authors were the developers of the RCWA [3].

In this first paper only the two-dimensional Transverse Electric (TE) light is considered. Transverse Magnetic (TM) waves (two-dimensional) follow in [9]. While there are some other names for RCWA, i.e. Fourier Modal Method and Moharam-Gaylord method [22], apparently RCWA is still the most used term. Among other related articles, reference [4] contains practical information concerning the programming, and reference [20] uses a different notation than all the other papers but contains all the theory and dedicates much text to the algorithm. The subject is also covered by text books [21], [5]. Reference [23] contains the essentials of the method, the accentuation lies on numerical efficiency. There are several articles dedicated to special topics: [7] considers anisotropic material, [8] deformable structures, [14] a photorefractive lens and [12] a pure reflection grating. Apparently,

for metallic gratings there are difficulties with the convergence with TM-modes. Two studies, [16] and [17], try to solve this problem: the last one more successfully. Not many authors deal with cylindrical, elliptical and azimuthally systems, exceptions are e.g. [10] and [11]. Reference [15] provides a short description of the RCWA for a 3-dimensional, inhomogeneous object in spherical coordinates. In [13] the RCWA is used for holograms.

Work on this master thesis was carried out partly at the engineering company Mecal Focal B.V., situated in Enschede. This company consists of three branches, concerned with wind energy, semiconductor industry and vision & optronics. Amongst other things the latter division offers roughness and flatness inspection, a cornea topographer and fluid process monitoring. The intention of this project was to understand the general idea of the RCWA and to gain experience with it. There already is a software at hand, but its code is not accessible. With a deeper insight the company hopes to broaden the useability.

As an example, the self-written RCWA-program is used to design a beam splitter. With this device an incoming laser beam can be parted. The number of outgoing beams depends on the grating design: its parameters determine which diffractive orders are let through. Such features can be used for measuring purposes. For example, one can examine the flatness of a plain: a regular grid of laser points is projected on a testing area. Now on an even surface this laser grid is regular in both directions, if distances between the points appear changed, there might be a buckle in the area.

2 The Rigorous Coupled Wave Analysis

To simulate DOEs we will make use of the RCWA. The intention is to identify and calculate the reflected and transmitted waves, together they form the diffracted field. In order to do so the RCWA needs to solve the Maxwell equations, equations which govern electric and magnetic fields and waves, so also light. We'll describe the details of the algorithm according to the paper [1]. For a sketch of the situation see figure 2.

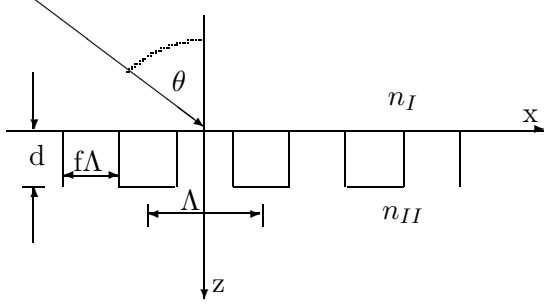


Figure 2: A grating in two-dimensional view. The wave comes in at an angle θ , it passes from material I (with refractive index n_I) into material II . The height of the grating is marked by the letter d , further there is the period Λ and the duty cycle of the ridge f .

We start with two Maxwell equations:

$$\nabla \times \mathbf{E} = \frac{-\partial \mathbf{B}}{\partial t} \quad (1)$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}. \quad (2)$$

Here \mathbf{E} is the electric field, \mathbf{B} the magnetic induction, \mathbf{H} the magnetic field, defined as $\mathbf{H} = \mu \mathbf{B}$, and \mathbf{D} the electric displacement, defined as $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$. μ is the permeability of the region and $\epsilon_0 \epsilon$ is the permittivity. We assume that the material in question is nonmagnetic at optical frequencies such that $\mu_0 = \mu$, with μ_0 the vacuum permeability. The permittivity is written as $\epsilon_0 \epsilon$, where the relative permittivity $\epsilon = n^2$ defines the refractive index n . We are interested in the frequency domain, i.e. we examine the behaviour of the fields for one defined frequency ω . So \mathbf{H} is of the general form

$$\mathbf{H}(r, t) = \text{Re}[\mathbf{H}(r) \exp(j\omega t)] \quad (3)$$

and \mathbf{E} is

$$\mathbf{E}(r, t) = \text{Re}[\mathbf{E}(r) \exp(j\omega t)]. \quad (4)$$

When we fill those in the given Maxwell equations (1) and (2) we get

$$\mathbf{E} = \frac{-j}{\epsilon \omega} \nabla \times \mathbf{H}, \quad (5)$$

$$\mathbf{H} = \frac{j}{\mu \omega} \nabla \times \mathbf{E}, \quad (6)$$

what we write out in components. This results in six equations. We make the assumption that there is no variation in the y -direction, so $\partial y = 0$, see figure 2. One observes that the system of six equations splits into two separate sets: One corresponding to the TE polarization, the other to TM (both for the 2-dimensional case).

2.1 TE polarization

In the TE-case we are only interested in three of the equations mentioned above, namely those containing the y -component of the electric field (E_y):

$$\frac{\partial E_y}{\partial z} = j\omega\mu_0 H_x, \quad (7)$$

$$\frac{\partial H_x}{\partial z} = j\omega\epsilon_0\epsilon E_y + \frac{H_z}{\partial x}, \quad (8)$$

$$\frac{\partial E_y}{\partial x} = -j\omega\mu_0 H_z. \quad (9)$$

These three equations can be put together to form the Helmholtz equation:

$$\nabla^2 E_y + \omega^2\mu\epsilon E_y = 0. \quad (10)$$

We specialize to a DOE with a given piecewise constant shape with period Λ (see figure 2). We focus first on the grating region $0 < z < d$. In this region the relative permittivity is a piecewise constant function of x only. In order to get a continuous function in the grating region we expand ϵ into a (complex) Fourier series:

$$\epsilon(x) = \sum_{h=-\infty}^{\infty} \tilde{\epsilon}_h \exp(j\frac{2\pi h}{\Lambda}x), \quad (11)$$

where the Fourier coefficients are

$$\tilde{\epsilon}_0 = n_{rd}^2 f + n_{gr}^2 (1 - f) \quad (12)$$

$$\tilde{\epsilon}_h = (n_{rd}^2 - n_{gr}^2) \frac{\sin(\pi h f)}{\pi h} \quad (13)$$

The parameters in these expressions can be found in figure 2, $n_{rd} = n_{II}$ is the refractive index of the ridge and $n_{gr} = n_I$ the refractive index of the groove. Note that this algorithm doesn't work for complex refractive indices. The period spans from $x = -\Lambda/2$ to $x = \Lambda/2$. In view of the periodicity of the DOE, we choose the following Fourier ansatz for the electric and magnetic field for the grating region (i.e.: $0 < z < d$):

$$E_{gy} = \sum_{i=-\infty}^{\infty} S_{yi}(z) \exp(-jk_{xi}x), \quad (14)$$

$$H_{gx} = -j\sqrt{\frac{\epsilon_0}{\mu_0}} \sum_i U_{xi}(z) \exp(-jk_{xi}x), \quad (15)$$

where S_{yi} is the yet unknown amplitude of the wave and U_{xi} follows from Maxwell's equations and is given by

$$k_0 U_{xi} = \frac{\partial S_{yi}}{\partial z}, \quad (16)$$

and

$$k_{xi} = k_0[n_I \sin \theta - i(\lambda_0/\Lambda)] = k_0 n_I \sin \theta - i \frac{2\pi}{\Lambda}. \quad (17)$$

with the vacuum wavenumber $k_0 = \omega\sqrt{\epsilon_0\mu_0} = 2\pi/\lambda_0$, where λ_0 is the wavelength in vacuum. The y-component of the electric field in the grating region is required to satisfy the Helmholtz equation (10):

$$\nabla^2 E_{gy} + k_0^2 \epsilon E_{gy} = 0. \quad (18)$$

Now the ansatz (14), (15) is inserted and we receive:

$$- \sum_i S_{yi} k_{xi}^2 \exp(-jk_{xi}x) + \sum_i \partial_z^2 S_{yi} \exp(-jk_{xi}x) + \omega^2 \mu_0 \epsilon_0 \sum_i \sum_h S_{yi} \tilde{\epsilon}_h \exp(-jk_{x(h+i)}x) = 0. \quad (19)$$

From now on the i in the last term is called p :

$$- \sum_i S_{yi} k_{xi}^2 \exp(-jk_{xi}x) + \sum_i \partial_z^2 S_{yi} \exp(-jk_{xi}x) + \omega^2 \mu_0 \epsilon_0 \sum_p S_{yp} \sum_h \tilde{\epsilon}_h \exp(-jk_{x(h+p)}x) = 0. \quad (20)$$

We call $i = h + p$, but shift the indices in the former \sum_h : we replace $i - p$ by i . That's possible because h extends from $-\infty$ to $+\infty$:

$$- \sum_i S_{yi} k_{xi}^2 \exp(-jk_{xi}x) + \sum_i \partial_z^2 S_{yi} \exp(-jk_{xi}x) + \omega^2 \mu_0 \epsilon_0 \sum_p \sum_i S_{yp} \tilde{\epsilon}_{(i-p)} \exp(-jk_{xi}x) = 0. \quad (21)$$

we write:

$$\sum_i (-S_{yi} k_{xi}^2 + \partial_z^2 S_{yi} + \omega^2 \mu_0 \epsilon_0 \sum_p S_{yp} \tilde{\epsilon}_{(i-p)}) \exp(-jk_{xi}x) = 0. \quad (22)$$

what results in:

$$\frac{\partial^2 S_{yi}}{\partial(z')^2} = \left(\frac{k_{xi}}{k_0}\right)^2 S_{yi} - \sum_p \tilde{\epsilon}_{(i-p)} S_{yp}, \quad (23)$$

for all i where $z' = k_0 z$. These are called the coupled wave equations. We want to write this in matrix form with index i , but an infinite matrix is of no numerical use. So we have to assume convergence and choose an appropriate range $i = -N, \dots, N$, with N a number which is big enough for convergence and is yet to choose. So the size of the matrix becomes $2N + 1$, starting with $i = -N$ as the first entry, followed by $i = -N + 1$ until the last entry $i = N$. This leads to the matrix equation

$$[\partial^2 \mathbf{S}_y / \partial(z')^2] = [\mathbf{K}_x^2 - \mathbf{E}][\mathbf{S}_y]. \quad (24)$$

Here \mathbf{K}_x is a diagonal matrix with the diagonal element at position i being equal to k_{xi}/k_0 and \mathbf{E} consists of the permittivity components with the entries $E_{i,p} = \tilde{\epsilon}_{(i-p)}$.

Now we've got a system of second-order differential equations with a system matrix $[\mathbf{K}_x^2 - \mathbf{E}]$ with constant coefficients. A solution [25] in terms of the eigenvalues of the matrix $[\mathbf{K}_x^2 - \mathbf{E}]$ is:

$$S_{yi}(z) = \sum_{m=1}^n w_{i,m} (c_m^+ \exp(-k_0 q_m z) + c_m^- \exp[k_0 q_m (z - d)]), \quad (25)$$

$$U_{xi}(z) = \sum_{m=1}^n v_{i,m} (-c_m^+ \exp(-k_0 q_m z) + c_m^- \exp[k_0 q_m (z - d)]), \quad (26)$$

where c_m^+ and c_m^- are yet to define constants, $w_{i,m}$ is the i^{th} component of the eigenvector related to an eigenvalue q_m^2 and

$$v_{i,m} = q_m w_{i,m}, \quad (27)$$

where the square root of q_m^2 is taken, with the positive real part.

So now the amplitudes of the fields in the grating regions are partly dealt with. But in fact we are more interested in the amplitudes of the fields in the regions I and II (see figure 2). They form the diffracted field. We call R_i the amplitudes of the reflected waves and T_i the amplitudes of the transmitted waves. The electric fields have to fit to the fields in the grating region at the interfaces in $z = 0$, $z = d$. Due to the given periodicity in the grating region these fields are written:

$$E_{I,y} = E_{inc,y} + \sum_i R_i \exp[-j(k_{xi}x - k_{I,zi}z)], \quad (28)$$

$$E_{II,y} = \sum_i T_i \exp(-j[k_{xi}x + k_{II,zi}(z - d)]), \quad (29)$$

with the incoming wave

$$E_{inc,y} = \exp[-jk_0 n_I (\sin \theta x + \cos \theta z)]. \quad (30)$$

When one of these fields is put back in equation (10) the relation $k^2 = k_{xi}^2 + k_{\ell,zi}^2$ is revealed. We are looking for an expression for $k_{\ell,zi}$ and therefore we get:

$$k_{\ell,zi} = \pm \sqrt{k_0^2 n_\ell^2 - k_{xi}^2}. \quad (31)$$

This still leaves us with four possibilities. So, let's first consider the case $k_0^2 n_\ell^2 > k_{xi}^2$: That would mean $\pm k_{\ell,zi}$ is real. Filled in the transmitting field (29) we would get a wave with imaginary exponents with an absolute value of 1 everywhere. Also with increasing z this wave doesn't vanish. This is conform with the description of the far field. In order to obtain the downward direction, we take $+k_{\ell,zi}$. It also happens that $k_0^2 n_\ell^2 < k_{xi}^2$ for other indices i , so that $\pm k_{\ell,zi}$ is imaginary. Let's first consider the $+$ -sign. Result would be a transmitted field growing in amplitude with increasing z . As this makes no sense we look at $-k_{\ell,zi}$: this time the transmitted field would approach zero. As this is a more realistic solution we choose for this algebraic sign. All in all we get:

$$k_{\ell,zi} = \begin{cases} k_0 [n_\ell^2 - (k_{xi}/k_0)^2]^{1/2} & k_0^2 n_\ell^2 > k_{xi}^2 \\ -jk_0 [(k_{xi}/k_0)^2 - n_\ell^2]^{1/2} & k_{xi}^2 > k_0^2 n_\ell^2 \end{cases} \quad \ell = I, II. \quad (32)$$

All these fields fulfill the Helmholtz equation (10). If we want to know the magnetic fields we make use of equations (7) and (9).

Now it's time to take into account the interface conditions at $z = 0$ and $z = d$. Continuity at the borders is required for the tangential electric field components \mathbf{E}_y and the magnetic field. Note that, according to (9), continuity of \mathbf{E}_y and \mathbf{H}_x across the interfaces implies continuity of \mathbf{H}_z . We obtain four equations:

$$E_{I,y}(x, 0) = E_{gy}(x, 0), \quad (33)$$

$$H_{I,x}(x, 0) = H_{gx}(x, 0), \quad (34)$$

$$E_{gy}(x, d) = E_{II,y}(x, d), \quad (35)$$

$$H_{gx}(x, d) = H_{II,x}(x, d), \quad (36)$$

and insert the fields. We multiply by $\exp(-jk_{xl}x)$ and integrate over x . Then orthogonality ($\frac{1}{\Lambda} \int_0^\Lambda \exp(-jk_{xl}x) * \exp(-jk_{xi}x)dx = \delta_{il}$, where $\delta_{il} = 1$ for $l = i$ and $\delta_{il} = 0$ for $l \neq i$) is used to eliminate the dependence on x . This leads to the equations

$$\delta_{i0} + R_i = \sum_{m=1}^n w_{i,m} [c_m^+ + c_m^- \exp(-k_0 q_m d)], \quad (37)$$

$$j[n_I \cos \theta \delta_{i0} - (k_{I,zi}/k_0)R_i] = \sum_{m=1}^n v_{i,m} [c_m^+ - c_m^- \exp(-k_0 q_m d)], \quad (38)$$

$$T_i = \sum_{m=1}^n w_{i,m} [c_m^+ \exp(-k_0 q_m d) + c_m^-], \quad (39)$$

$$j(k_{II,zi}/k_0)T_i = \sum_{m=1}^n v_{i,m} [c_m^+ \exp(-k_0 q_m d) - c_m^-]. \quad (40)$$

For purposes of implementation it appears to be convenient to write these in matrix forms:

$$\begin{bmatrix} \delta_{i0} \\ jn_I \cos \theta \delta_{i0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_I \end{bmatrix} [\mathbf{R}] = \begin{bmatrix} \mathbf{W} & \mathbf{WX} \\ \mathbf{V} & -\mathbf{VX} \end{bmatrix} \begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix} \quad (41)$$

and

$$\begin{bmatrix} \mathbf{WX} & \mathbf{W} \\ \mathbf{XV} & -\mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ j\mathbf{Y}_{II} \end{bmatrix} [\mathbf{T}] \quad (42)$$

with \mathbf{I} the identity matrix, \mathbf{Y}_ℓ diagonal matrices with the diagonal elements $(k_{\ell,zi}/k_0)$, \mathbf{W} the matrix with the eigenvectors, \mathbf{X} a diagonal matrix with the diagonal elements $\exp(-k_0 q_m d)$ and $\mathbf{V} = \mathbf{W}\mathbf{Q}$, where \mathbf{Q} contains the eigenvalues q_m in its diagonal entries. It turns out [1] to be numerically advantageous to first calculate \mathbf{c}^\pm by eliminating R_i from equations (37) and (38), or (41) respectively, and T_i from equations (39) and (40), or (42). This results in

$$\begin{bmatrix} j\mathbf{Y}_I \mathbf{W} + \mathbf{V} & j\mathbf{Y}_I \mathbf{W}\mathbf{X} - \mathbf{VX} \\ \mathbf{VX} - j\mathbf{Y}_{II} \mathbf{W}\mathbf{X} & -\mathbf{V} - j\mathbf{Y}_{II} \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix} = \begin{bmatrix} jn_I \cos \theta \delta_{i0} + j\mathbf{Y}_I \delta_{i0} \\ 0 \end{bmatrix}. \quad (43)$$

This linear set of equations is now solvable by Matlab. \mathbf{c}^\pm can be inserted in equations (41) and (42) to receive R_i and T_i . According to the algorithm we have now computed the amplitudes of the reflected and transmitted lightwaves. When these are inserted back in the original ansatz (14) and (15) we obtain the electric diffracted fields. They illustrate where and how the light of the DOE is spreading.

The diffraction efficiency is the fraction of energy that is transported by a certain order relative to the energy of the incoming light. It is calculated for the reflected and transmitted case separately. First the Poynting vector is computed:

$$\mathbf{S} = \frac{1}{2\mu_0} \text{Re}(\mathbf{E} \times \mathbf{B}^*). \quad (44)$$

We are assuming that we are only looking at one order undisturbed by others. This is only applicable in the far field. Using the terms in equations (28), (29), here only the reflected part, related to

index i and equation (6), this leads to

$$\mathbf{S}_{ri} = \frac{1}{2\mu_0\omega} R_i R_i^* \text{Re} \begin{pmatrix} k_{xi} \\ 0 \\ -k_{I,zi} \end{pmatrix} \quad (45)$$

for the reflected field,

$$\mathbf{S}_{ti} = \frac{1}{2\mu_0\omega} T_i T_i^* \text{Re} \begin{pmatrix} k_{xi} \\ 0 \\ k_{II,zi} \end{pmatrix} \quad (46)$$

for the transmitted field and

$$\mathbf{S}_{inc} = \frac{1}{2\mu_0\omega} \text{Re} \begin{pmatrix} k_0 n_I \sin \theta \\ 0 \\ k_0 n_I \cos \theta \end{pmatrix} \quad (47)$$

for the incoming light. Now only the vertical contribution of the reflected and transmitted orders is considered and divided through the vertical contribution of the incoming field to obtain the ratio we are looking for. As for the reflected wave, which travels in the other direction as the incoming and transmitted part of the field, we are taking the absolute value of $\text{Re}(-k_{I,zi})$. Therefore we arrive at the expressions:

$$\text{DE}_{ri} = R_i R_i^* \text{Re} \left(\frac{k_{I,zi}}{k_0 n_I \cos \theta} \right) \quad (48)$$

and

$$\text{DE}_{ti} = T_i T_i^* \text{Re} \left(\frac{k_{II,zi}}{k_0 n_I \cos \theta} \right) \quad (49)$$

for the diffraction efficiencies of the reflected and transmitted waves of order i . Conservation of energy requires that the sum of these should equal 1:

$$\sum_i (\text{DE}_{ri} + \text{DE}_{ti}) = 1. \quad (50)$$

This expression can serve as a consistency check for our computations.

2.2 TM polarization

The implementation of the TM polarization isn't very different from that of the TE case. This time the leading field is the magnetic field H_y . Expressions are the same as in the TE case if not otherwise stated. Again, we start with the Maxwell equations (1) and (2). When filling (3) and (4) into (1) and (2) we get six equations. Now we pick out those which contain the y -component of H :

$$\frac{\partial H_y}{\partial z} = -j\epsilon\omega E_x, \quad (51)$$

$$\frac{\partial H_y}{\partial x} = j\epsilon\omega E_z, \quad (52)$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -j\mu_0\omega H_y. \quad (53)$$

The fields are as follows:

$$H_{I,y} = \exp[-jk_0 n_I (\sin \theta x + \cos \theta z)] + \sum_i R_i \exp[-j(k_{xi}x - k_{I,zi}z)], \quad (54)$$

$$H_{II,y} = \sum_i T_i \exp\{-j[k_{xi}x + k_{II,zi}(z - d)]\}, \quad (55)$$

$$H_{gy} = \sum_i U_{yi}(z) \exp(-jk_{xi}x) \quad (56)$$

and

$$E_{gx} = j\sqrt{\frac{\mu_0}{\epsilon_0}} \sum_i S_{xi}(z) \exp(-jk_{xi}x). \quad (57)$$

The ansatz of the fields (54)-(57) is inserted in equations (51)-(53) and we receive (in matrix form):

$$[\partial^2 \mathbf{U}_y / \partial (z')^2] = [\mathbf{E} \mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_x - \mathbf{E}] [\mathbf{U}_y]. \quad (58)$$

We'll need the eigenvalues of the matrix $[\mathbf{E} \mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_x - \mathbf{E}]$ to solve this system of second order differential equations. The amplitudes of the magnetic and electric fields are given by:

$$U_{yi}(z) = \sum_{m=1}^n w_{i,m} (c_m^+ \exp(-k_0 q_m z) + c_m^- \exp[k_0 q_m (z - d)]) \quad (59)$$

and

$$S_{xi}(z) = \sum_{m=1}^n v_{i,m} (-c_m^+ \exp(-k_0 q_m z) + c_m^- \exp[k_0 q_m (z - d)]), \quad (60)$$

with $w_{i,m}$ and q_m as defined in the TE case, but $v_{i,m}$ are the elements of the matrix $\mathbf{V} = \mathbf{E}^{-1} \mathbf{W} \mathbf{Q}$. This makes the resulting matrix equation slightly different:

$$\begin{bmatrix} j\mathbf{Z}_I \mathbf{W} + \mathbf{V} & j\mathbf{Z}_I \mathbf{W} \mathbf{X} - \mathbf{V} \mathbf{X} \\ \mathbf{V} \mathbf{X} - j\mathbf{Z}_{II} \mathbf{W} \mathbf{X} & -\mathbf{V} - j\mathbf{Z}_{II} \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix} = \begin{bmatrix} j \cos \theta \delta_{i0} / n_I + j\mathbf{Z}_I \delta_{i0} \\ 0 \end{bmatrix}. \quad (61)$$

Here \mathbf{X} is the same as above and \mathbf{Z}_ℓ ($\ell = I, II$) are diagonal matrices with the diagonal entries $k_{I,zi} / k_0 n_\ell^2$.

Also the diffraction efficiency of the transmitted orders changes:

$$\text{DE}_{ti} = T_i T_i^* \text{Re} \left(\frac{k_{II,zi}}{n_{II}^2} \right) / \left(\frac{k_0 \cos \theta}{n_I} \right). \quad (62)$$

2.3 Multilayer systems

The next step was to move on to a system with several layers, as to be seen in figure 3. As before, we focus on a TE configuration first. Fourier expansions are defined separately for each layer. The field of layer ℓ is given by:

$$E_{\ell,gy} = \sum_i S_{\ell,yi}(z) \exp(-jk_{xi}x) \quad (63)$$

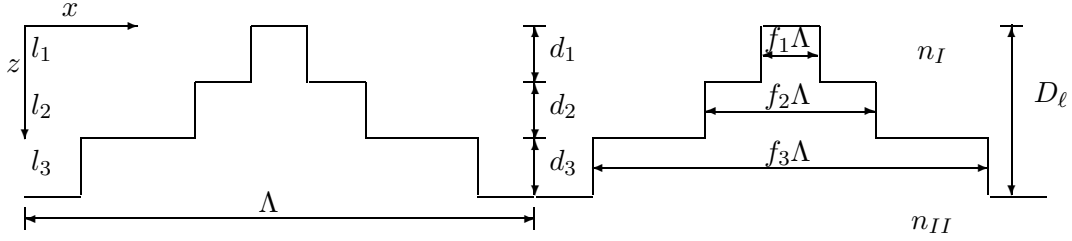


Figure 3: A surface-relief grating with three layers.

with

$$S_{\ell,xi}(z) = \sum_{m=1}^n w_{\ell,i,m} (c_{\ell,m}^+ \exp(-k_0 q_{\ell,m}(z - D_{\ell} + d_{\ell})) + c_{\ell,m}^- \exp[k_0 q_{\ell,m}(z - D_{\ell})]) \quad (64)$$

and $D_{\ell} = \sum_{p=1}^{\ell} d_p$, i.e. the coordinate of the height.

For each layer we've got a matrix \mathbf{E}_{ℓ} , \mathbf{W}_{ℓ} , \mathbf{V}_{ℓ} and \mathbf{X}_{ℓ} . Also the \mathbf{c}_{ℓ}^{\pm} coefficients and the fields of course, have to be calculated for each layer. The relation between the coefficients is derived from the interface conditions between the layers and results in:

$$\begin{bmatrix} \mathbf{W}_{\ell-1} \mathbf{X}_{\ell-1} & \mathbf{W}_{\ell-1} \\ \mathbf{V}_{\ell-1} \mathbf{X}_{\ell-1} & -\mathbf{V}_{\ell-1} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{\ell-1}^+ \\ \mathbf{c}_{\ell-1}^- \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{\ell} & \mathbf{W}_{\ell} \mathbf{X}_{\ell} \\ \mathbf{V}_{\ell} & -\mathbf{V}_{\ell} \mathbf{X}_{\ell} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{\ell}^+ \\ \mathbf{c}_{\ell}^- \end{bmatrix}, \quad (65)$$

so that the transitions are continuous. And the matrix equations at the outmost boundaries become

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0} \cos(\theta) n_I \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_I \end{bmatrix} [\mathbf{R}] = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_1 \mathbf{X}_1 \\ \mathbf{V}_1 & -\mathbf{V}_1 \mathbf{X}_1 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1^+ \\ \mathbf{c}_1^- \end{bmatrix} \quad (66)$$

and

$$\begin{bmatrix} \mathbf{W}_L \mathbf{X}_L & \mathbf{W}_L \\ \mathbf{V}_L \mathbf{X}_L & -\mathbf{V}_L \end{bmatrix} \begin{bmatrix} \mathbf{c}_L^+ \\ \mathbf{c}_L^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ j\mathbf{Y}_{II} \end{bmatrix} [\mathbf{T}]. \quad (67)$$

Combining equations (65) till (67) leads to:

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0} \cos(\theta) n_I \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_I \end{bmatrix} [\mathbf{R}] = \prod_{\ell=1}^L \begin{bmatrix} \mathbf{W}_{\ell} & \mathbf{W}_{\ell} \mathbf{X}_{\ell} \\ \mathbf{V}_{\ell} & -\mathbf{V}_{\ell} \mathbf{X}_{\ell} \end{bmatrix} \begin{bmatrix} \mathbf{W}_{\ell} \mathbf{X}_{\ell} & \mathbf{W}_{\ell} \\ \mathbf{V}_{\ell} \mathbf{X}_{\ell} & -\mathbf{V}_{\ell} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I} \\ j\mathbf{Y}_{II} \end{bmatrix} [\mathbf{T}]. \quad (68)$$

Unfortunately, this last equation isn't stable if solved directly for R and T . The paper suggested two methods implementing the system otherwise. The following one was used: At first the right side of equation (68) (only with the factor L) is changed into

$$\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L \mathbf{X}_L \end{bmatrix} \begin{bmatrix} \mathbf{W}_L \mathbf{X}_L & \mathbf{W}_L \\ \mathbf{V}_L \mathbf{X}_L & -\mathbf{V}_L \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I} \\ j\mathbf{Y}_{II} \end{bmatrix} [\mathbf{T}] = \begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L \mathbf{X}_L \end{bmatrix} \begin{bmatrix} \mathbf{X}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \\ \mathbf{V}_L & -\mathbf{V}_L \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_{L+1} \\ \mathbf{g}_{L+1} \end{bmatrix} [\mathbf{T}], \quad (69)$$

with

$$\begin{bmatrix} \mathbf{f}_{L+1} \\ \mathbf{g}_{L+1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ j\mathbf{Y}_{II} \end{bmatrix} \quad (70)$$

and

$$\begin{bmatrix} \mathbf{W}_L \mathbf{X}_L & \mathbf{W}_L \\ \mathbf{V}_L \mathbf{X}_L & -\mathbf{V}_L \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{X}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \\ \mathbf{V}_L & -\mathbf{V}_L \end{bmatrix}^{-1}. \quad (71)$$

We now state

$$\begin{bmatrix} \mathbf{a}_L \\ \mathbf{b}_L \end{bmatrix} = \begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \\ \mathbf{V}_L & -\mathbf{V}_L \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_{L+1} \\ \mathbf{g}_{L+1} \end{bmatrix}. \quad (72)$$

We substitute that in (69):

$$\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L \mathbf{X}_L \end{bmatrix} \begin{bmatrix} \mathbf{X}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{a}_L \\ \mathbf{b}_L \end{bmatrix} [\mathbf{T}]. \quad (73)$$

Further we introduce another intermediate quantity T_L through $\mathbf{T} = \mathbf{a}_L^{-1} \mathbf{X}_L \mathbf{T}_L$. This is substituted into the last expression which leads to

$$\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L \mathbf{X}_L \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{b}_L \mathbf{a}_L^{-1} \mathbf{X}_L \end{bmatrix} [\mathbf{T}_L] \quad (74)$$

which we set equal to

$$\begin{bmatrix} \mathbf{f}_L \\ \mathbf{g}_L \end{bmatrix} [\mathbf{T}_L]. \quad (75)$$

Now equation (68) becomes

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0} \cos(\theta) n_I \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_I \end{bmatrix} [\mathbf{R}] = \prod_{\ell=1}^{L-1} \begin{bmatrix} \mathbf{W}_\ell & \mathbf{W}_\ell \mathbf{X}_\ell \\ \mathbf{V}_\ell & -\mathbf{V}_\ell \mathbf{X}_\ell \end{bmatrix} \begin{bmatrix} \mathbf{W}_\ell \mathbf{X}_\ell & \mathbf{W}_\ell \\ \mathbf{V}_\ell \mathbf{X}_\ell & -\mathbf{V}_\ell \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_L \\ \mathbf{g}_L \end{bmatrix} [\mathbf{T}_L], \quad (76)$$

i.e. the product on the right-hand side has one factor less. After repeating this analogously for all further inner layers we end up with

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0} \cos(\theta) n_I \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_I \end{bmatrix} [\mathbf{R}] = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{g}_1 \end{bmatrix} [\mathbf{T}_1], \quad (77)$$

which is reformulated:

$$\begin{bmatrix} -\mathbf{I} & \mathbf{f}_1 \\ j\mathbf{Y}_I & \mathbf{g}_1 \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{T}_1 \end{bmatrix} = \begin{bmatrix} \delta_{i0} \\ j\delta_{i0} \cos(\theta) n_I \end{bmatrix} \quad (78)$$

so that \mathbf{R} and \mathbf{T}_1 can now be calculated directly (\mathbf{f}_1 and \mathbf{g}_1 follow from equations (74) and (75)). Finally, \mathbf{T} is computed by

$$\mathbf{T} = \mathbf{a}_L^{-1} \mathbf{X}_L \dots \mathbf{a}_\ell^{-1} \mathbf{X}_\ell \dots \mathbf{a}_1^{-1} \mathbf{X}_1 \mathbf{T}_1. \quad (79)$$

Calculating the electric and magnetic fields turned out to be slightly more difficult than in the one-layer system, since the coefficients c_ℓ^\pm weren't calculated explicitly (as before).

Equation (67) is reformulated:

$$\begin{bmatrix} -\mathbf{W}_L & \mathbf{f}_{L+1} \\ -\mathbf{V}_L & \mathbf{g}_{L+1} \end{bmatrix} \begin{bmatrix} \mathbf{c}_L^- \\ \mathbf{c}_{L+1}^+ \end{bmatrix} = \begin{bmatrix} \mathbf{W}_L \mathbf{X}_L \\ \mathbf{V}_L \mathbf{X}_L \end{bmatrix} \mathbf{c}_L^+ \quad (80)$$

where $\mathbf{c}_{L+1}^+ = \mathbf{T}$, \mathbf{f}_{L+1} and \mathbf{g}_{L+1} are defined as \mathbf{f}_{L+1} and \mathbf{g}_{L+1} in equation (70).

Now we want to determine a relation between the coefficients, let's say $\mathbf{c}_L^- = \mathbf{a}\mathbf{a}_L\mathbf{c}_L^+$. We also assign $\mathbf{c}_{L+1}^+ = \mathbf{b}\mathbf{b}_L\mathbf{c}_L^+$ and substitute this in equation (80) what results in:

$$\begin{bmatrix} \mathbf{a}\mathbf{a}_L \\ \mathbf{b}\mathbf{b}_L \end{bmatrix} = \begin{bmatrix} -\mathbf{W}_L & \mathbf{f}\mathbf{f}_{L+1} \\ \mathbf{V}_L & \mathbf{g}\mathbf{g}_{L+1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L\mathbf{X}_L \end{bmatrix} \quad (81)$$

The general formula for $\mathbf{f}\mathbf{f}_\ell$ and $\mathbf{g}\mathbf{g}_\ell$ is

$$\begin{bmatrix} \mathbf{f}\mathbf{f}_\ell \\ \mathbf{g}\mathbf{g}_\ell \end{bmatrix} \mathbf{c}_\ell^+ = \begin{bmatrix} \mathbf{W}_\ell & \mathbf{W}_\ell\mathbf{X}_\ell \\ \mathbf{V}_\ell & -\mathbf{V}_\ell\mathbf{X}_\ell \end{bmatrix} \begin{bmatrix} \mathbf{c}_\ell^+ \\ \mathbf{c}_\ell^- \end{bmatrix} \quad (82)$$

Now all coefficients $\mathbf{a}\mathbf{a}_\ell$ and $\mathbf{b}\mathbf{b}_\ell$ can be calculated, starting with \mathbf{c}_1^+ all other coefficients \mathbf{c}_ℓ^\pm can be determined. \mathbf{c}_1^+ follows from the reformulated equation (66):

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos(\theta)n_I \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_I \end{bmatrix} [\mathbf{R}] = \begin{bmatrix} \mathbf{f}\mathbf{f}_1 \\ \mathbf{g}\mathbf{g}_1 \end{bmatrix} \mathbf{c}_1^+ \quad (83)$$

The problematic inversion of equation (65) is avoided this way.

The TM polarization for the multiple layer system is very similar to the multi-layered TE polarization. Instead of the eigenvalues of matrix $[\mathbf{K}_x^2 - \mathbf{E}_\ell]$ we take those of $[\mathbf{E}_\ell\mathbf{K}_x\mathbf{E}_\ell^{-1}\mathbf{K}_x - \mathbf{E}_\ell]$ and we use \mathbf{Z}_ℓ instead of \mathbf{Y}_ℓ . Furthermore, the first matrix in equation (66) will become $\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos(\theta)/n_I \end{bmatrix}$.

3 Benchmarking results

A program that implements the theory above has been implemented in Matlab. In order to avoid stability problems some matrix equations had to be rewritten, as was to be seen in the last section.

The input parameters are:

- TE or TM
- Wavelength in free space λ_0
- Angle of incidence θ
- Refractive indices n_I and n_{II}
- Number of layers l (in a multilayer system)
- Number of Fourier harmonics N
- Grating period Λ
- Height of layer(s) d_ℓ
- Duty cycle f_ℓ

As output we get:

- Amplitudes of the reflected and transmitted waves R_i and T_i
- Diffraction efficiencies of i^{th} order, reflected DE_{r_i} and transmitted DE_{t_i}
- Electric and magnetic fields E and H

3.1 TE polarization

This code was tested by a reproduction of figure 2 from [1], the results are shown in figure 4: Left above, the two curves agree quite well, only the dips on the right side don't reach zero like the original graph. The plot next to it on the right is also very similar to the original, the amplitude and frequency look the same. But again the curve doesn't reach 0 at the minima, also it's quite shaky as if there's another frequency present. The lower left plot agrees very much with the TM curve. The last plot, right below, matches very good with the figure in the paper. Note that the lower left plots of the TE (figure 4) and TM polarization (figure 7) are exchanged in [1].

To verify the results of this program the free software 'Optiscan' by the University of Arizona (2005), written by students, was used. It also employed the Rigorous Coupled Wave Theory but the code was not accessible. Also it used a slightly different coordinate system, as a positive θ in my code corresponds to a $-\theta$ in 'Optiscan'. The configurations in this program also couldn't be changed to reproduce figure 4. So two other cases with varied wavelength were chosen. The results for the transmitted and reflected diffraction efficiencies is shown in figure 5. The agreement in all cases was satisfying.

It was also possible to calculate the optical fields, see figure 6 for the electric field. It thoroughly follows the form of the grating (also depicted) and is continuous all over the field as desired.

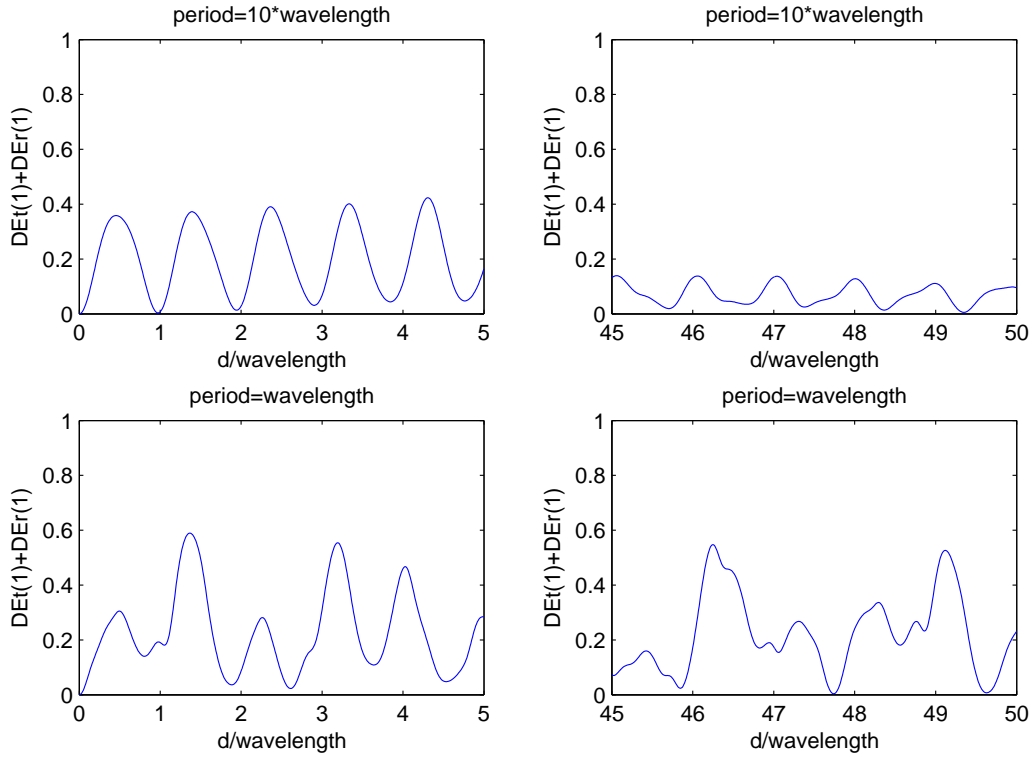
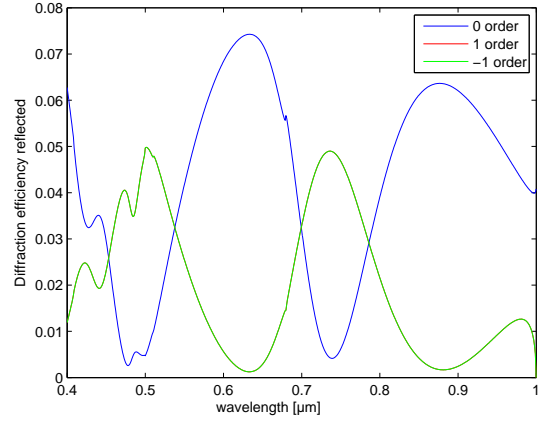
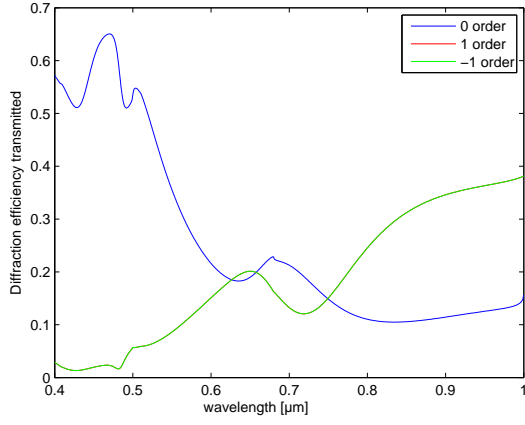
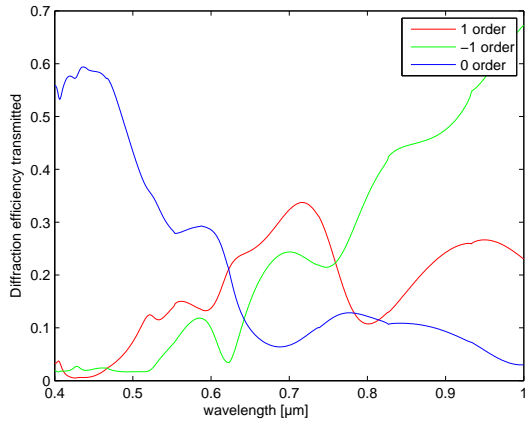


Figure 4: Here $n_1 = 1$, $n_2 = 2.04$, $\theta = 10^\circ$, the duty cycle equals 0.5 and the number of Fourier harmonics is 101 ($N = 50$). The first-order diffraction efficiency (TE) is depicted against the ratio d/λ_0 . In the graphs above $\Lambda = 10\lambda_0$ and below $\Lambda = \lambda_0$. On the left side the low values of the ratio d/λ_0 are depicted, on the right plots its magnitudes reach 50. This is a reproduction of figure 2 in the paper [1].

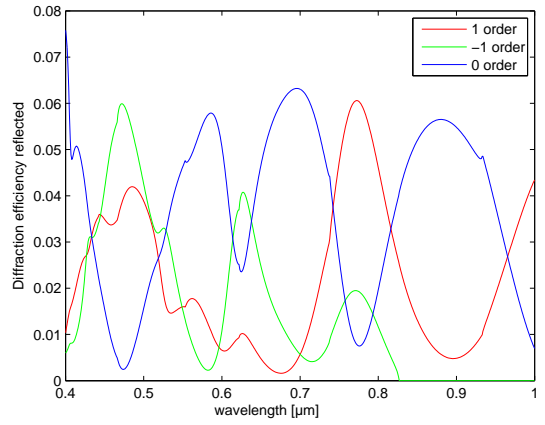


(a)

(b)



(c)



(d)

Figure 5: Here $n_1 = 1$, $n_2 = 2.04$, height = 400nm , $\Lambda = 1\mu\text{m}$, the duty cycle equals 0.5 and the number of Fourier harmonics is 101 ($N = 50$). The wavelength is varied. Above the transmitted and reflected diffraction efficiency at 0 degrees is depicted. The 1st and -1st order are identical in this symmetric case, as it should be. Below the transmitted and reflected diffraction efficiency for 10° for three orders is calculated. The curves in each graph (also in the following figures) are calculated by our code and again by a free RCWA solver by the University of Arizona (Optiscan) in the same plot. All results coincide on the scale of the plots.

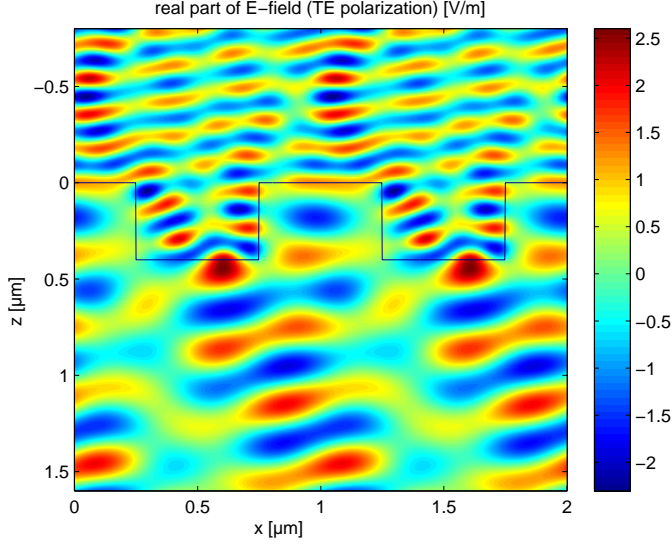
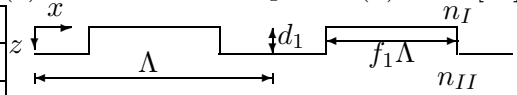


Figure 6: The electric field for the grating with $\lambda_0 = 0.355\mu\text{m}$, $\Lambda = 1\mu\text{m}$, $n_1 = 2.04$, $n_2 = 1$, $\theta = 10^\circ$, duty cycle equals 0.5, the number of Fourier harmonics is 101 ($N = 50$) and height of the layer= $0.4\mu\text{m}$. The light wave enters from above and the grating structure is also plotted.

Other benchmark tests were done with another paper [24]. Here the Finite Element Method was used to analyse binary gratings and design optimization was done by gradient descent. Respectively a two-, three-, four- and five-beam splitter was presented. We'll look at the first two examples, in both cases for perpendicular incidence $\theta = 0^\circ$. The parameters are listed in table 1.

Table 1: The parameters of the two-beam splitter (a) and three-beam splitter (b) from [24].

	n_I	n_{II}	λ_0	Λ	d	f
a)	1.5315	1	$0.633\mu\text{m}$	$2\lambda_0$	$0.734\mu\text{m}$	0.72
b)	1.5315	1	$0.633\mu\text{m}$	$2\lambda_0$	$0.43\mu\text{m}$	0.58



In table 2 the transmitted energy of the desired orders is compared: the first column lists the results of the cited paper, the second column tells what our RCWA program calculated.

These numbers are very similar, but it's striking that the diffraction efficiency of the 1^{th} and -1^{th} order in the FEM calculation is not exactly the same, although the light comes in perpendicularly. Symmetry therefore requires identical diffraction efficiencies for orders ± 1 . Experimenting with the duty cycle revealed that these parameters didn't reach the optimum: with $f = 0.77$ the numbers improved, see table 2. It follows the three-beam splitter. This time there's quite a difference between the two methods (see table 2). Again, the results of the FEM are not symmetric as required.

Table 2: The transmitted diffraction efficiencies of the two- and three-beam splitter calculated by means of the FEM ([24]) and the RCWA.

DEt	a)			b)	
	FEM	RCWA	RCWA ($f = 0.77$)	FEM	RCWA
-1 order	0.435	0.4351	0.4456	0.295	0.3204
0 order	0.023	0.023	0.009	0.279	0.2469
1 order	0.433	0.4351	0.4456	0.293	0.3204

3.2 TM polarization

The results for TM polarization were comparable to those of TE: The graphs were similar (see figure 7). Again the low marks don't reach zero. Our program did not work for a configuration with more than five layers and a large period, due to.

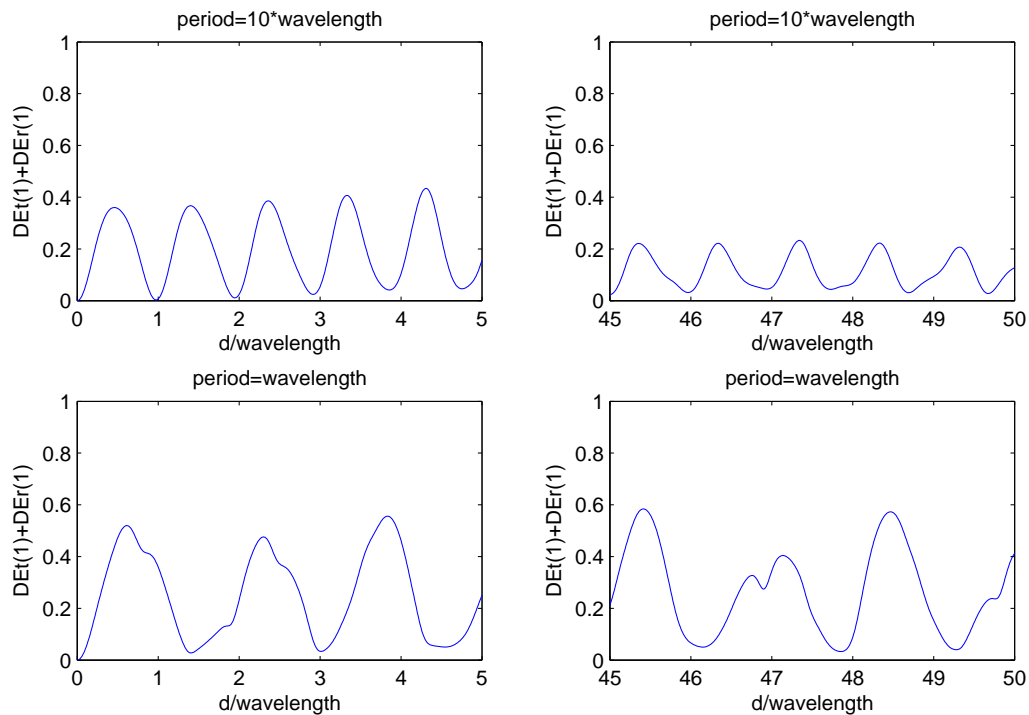
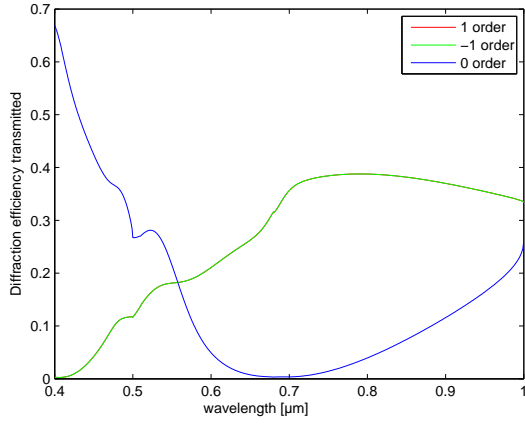
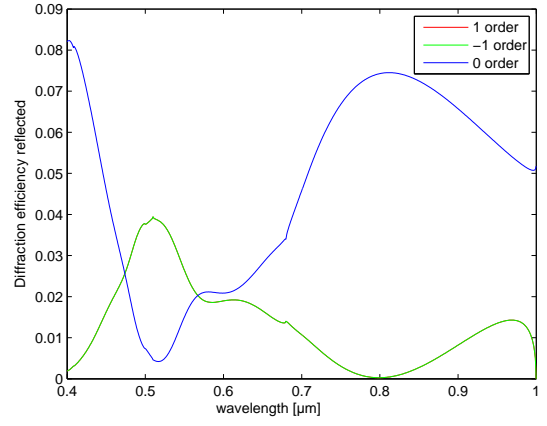


Figure 7: Reproduction of figure 2 from [1], TM polarization. Here $n_I = 1$, $n_{II} = 2.04$, $\theta = 10^\circ$, duty cycle equals 0.5 and the number of Fourier harmonics is 101 ($N = 50$). The first-order diffraction efficiency is depicted.

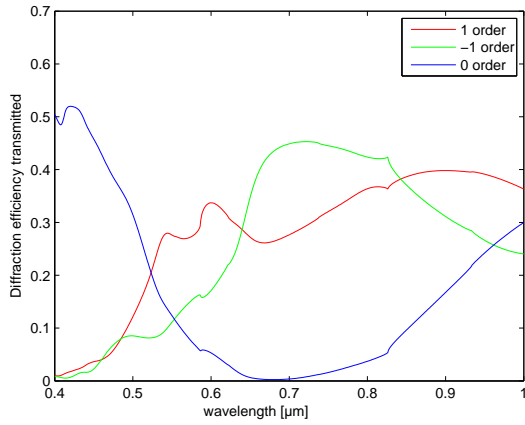
Again, the comparison with 'Optiscan' was satisfying, as the graphs computed by the RCWA coincided when put in the same plot (figure 8).



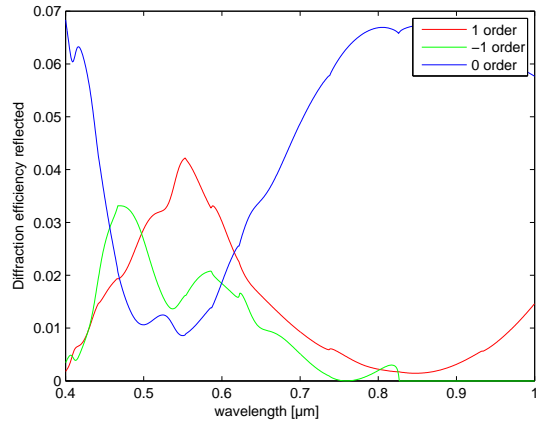
(a)



(b)



(c)



(d)

Figure 8: Here $n_I = 1$, $n_{II} = 2.04$, height = $0.4\mu m$, $\Lambda = 1\mu m$, the duty cycle equals 0.5 and the number of Fourier harmonics is 161 ($N = 80$). The wavelength is varied. The transmitted and reflected diffraction efficiency at 0 degrees with TM polarization is depicted in (a) and (b). The 1st and -1st order are identical because $\theta = 0^\circ$. The transmitted and reflected diffraction efficiency for incidence angle $\theta = 10^\circ$ for three orders is calculated in the lower graphs (c) and (d). They coincide with the 'Optiscan'.

3.3 Multilayer systems

Also a multilayer system has been benchmarked. First, the pictures in [18], figure 2, were reproduced: They treated a 15-layer system, with partly quite wide and deep grooves ($\Lambda = 10\lambda_0$ and/or $d/\lambda_0 = 50$). The grating looks like a sawtooth function decreasing from left to right. Each layer depth is $1/15$ of the whole grating depth D and the stepsize in width per layer is $1/16$ of the grating period (see figure 9). In order to do so, equation (13) has to be changed, due to the no longer symmetric setting of the groove/teeth-ratio within each period. We therefore consider the equation

$$\epsilon(x) = \begin{cases} n_{II}^2 & \text{for } 0 < x < f\Lambda \\ n_I^2 & \text{for } f\Lambda < x < \Lambda \end{cases} \quad (84)$$

and apply Fourier expansion:

$$\tilde{\epsilon}_h = j \frac{(n_I^2 - n_{II}^2)}{2\pi h} (1 - \exp(-jh2\pi f)). \quad (85)$$

Furthermore, $\theta = 10^\circ$, $n_I = 1$, $n_{II} = n_{ridge} = 2.04$ and the groove depth is being varied. The results are to be seen in figure 10. For $\lambda_0 = \Lambda$ both TE and TM polarization are agreeing with the paper. But when Λ goes up to $10\lambda_0$ the TM modulation becomes unstable, due to an ill-conditioned system in equation (79): Also with an increasing number of Fourier harmonics ($N \rightarrow 501$) the curve consists of an uncountable number of peaks, partly reaching 15. This is obviously due to the fact that for a big period ($\Lambda = 10\lambda_0$) and in general the TM polarization more harmonics are needed. In the TE case it's still good coinciding with the paper.

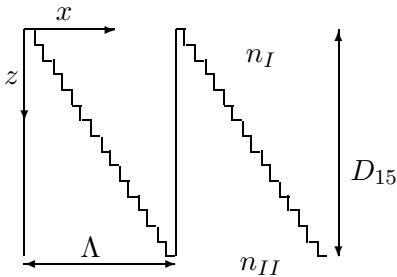


Figure 9: A sawtooth grating with $D_{15} = 1.5\Lambda$.

Again, 'Optiscan' was used to compare the curves, see figure 11. We observed virtually no differences versus our results.

With the present solver for multilayer systems it's also possible now to simulate a structure of finite thickness (see figure 12).

In this case $n_I = n_{II}$. For a simple example, a binary grating with finite thickness, we make use of two layers: one persistent with duty cycle 1 and the other one with the desired grating period. For the results, see figure 13(a,b). A more realistic case is a grating which is only $1/100$ of the whole finite layer system (see figures 13(c,d)). As the grating is etched into a plate this value is more lifelike. The number of peaks is remarkable, for that reason a smaller wavelength range is chosen. It's due to the many reflections inside the continuous layer. 'Optiscan' with these parameters produced the same figure.

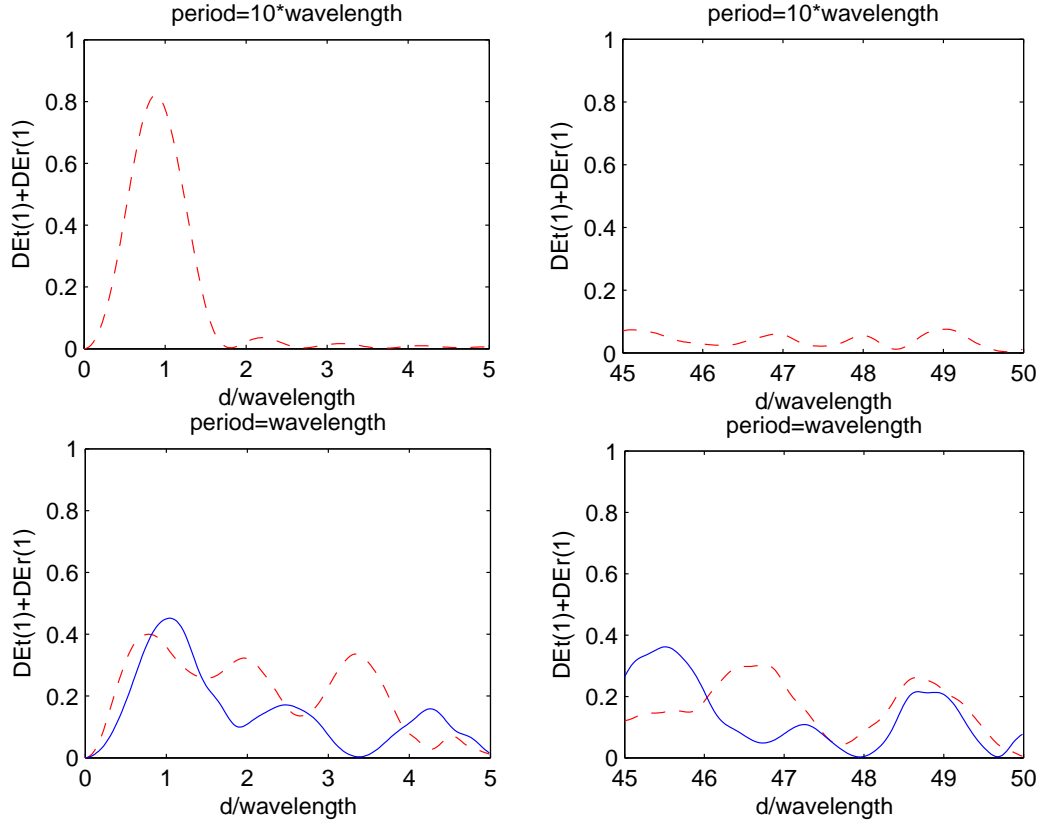
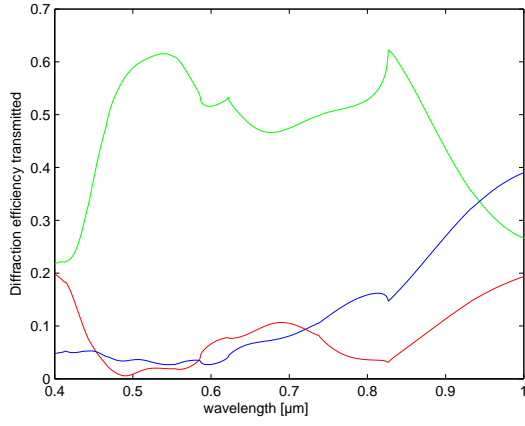
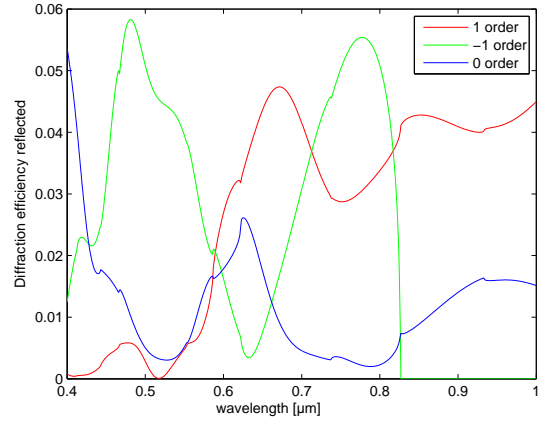


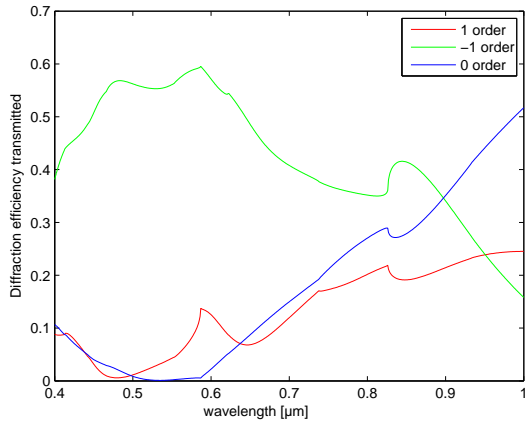
Figure 10: Reproduction of graph from [18], figure 3. The red dashed line represents the TE results and TM comes along with the blue curves. It depicts the first order diffraction efficiency dependence with $n_{II} = 2.04$, $n_I = 1$ at $\theta = 10^\circ$, the grating consists of a 15 layer system. The two pictures above use $\Lambda = 10\lambda_0$, as this is a large system, only the TE polarization produced stable results. In the plots below $\Lambda = \lambda_0$ holds, here also the TM curves are steady. The matching between the plots in the paper and the reproductions is very good.



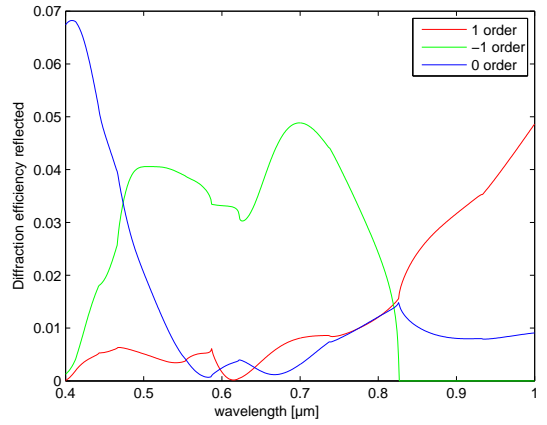
(a)



(b)



(c)



(d)

Figure 11: Here for the TE polarization (a,b) the wavelength is varied from $0.4\mu m$ to $1\mu m$ in a two-layer system. We use $\theta = 10^\circ$, the period is $1\mu m$, $N = 41$ and the height of one layer is $0.2\mu m$. The duty cycles are 0.3 and 0.6 respectively. Here the transmitted and reflected diffraction efficiencies of the -1^{st} (green), 0^{th} (blue) and 1^{st} order (red) are depicted. In the lower pictures (c) and (d), the wavelength for the TM polarization is varied from $0.4\mu m$ to $1\mu m$, $\theta = 10^\circ$. The curves of both programs coincide.

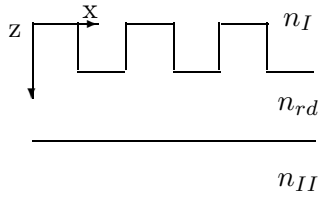


Figure 12: A grating of finite thickness.

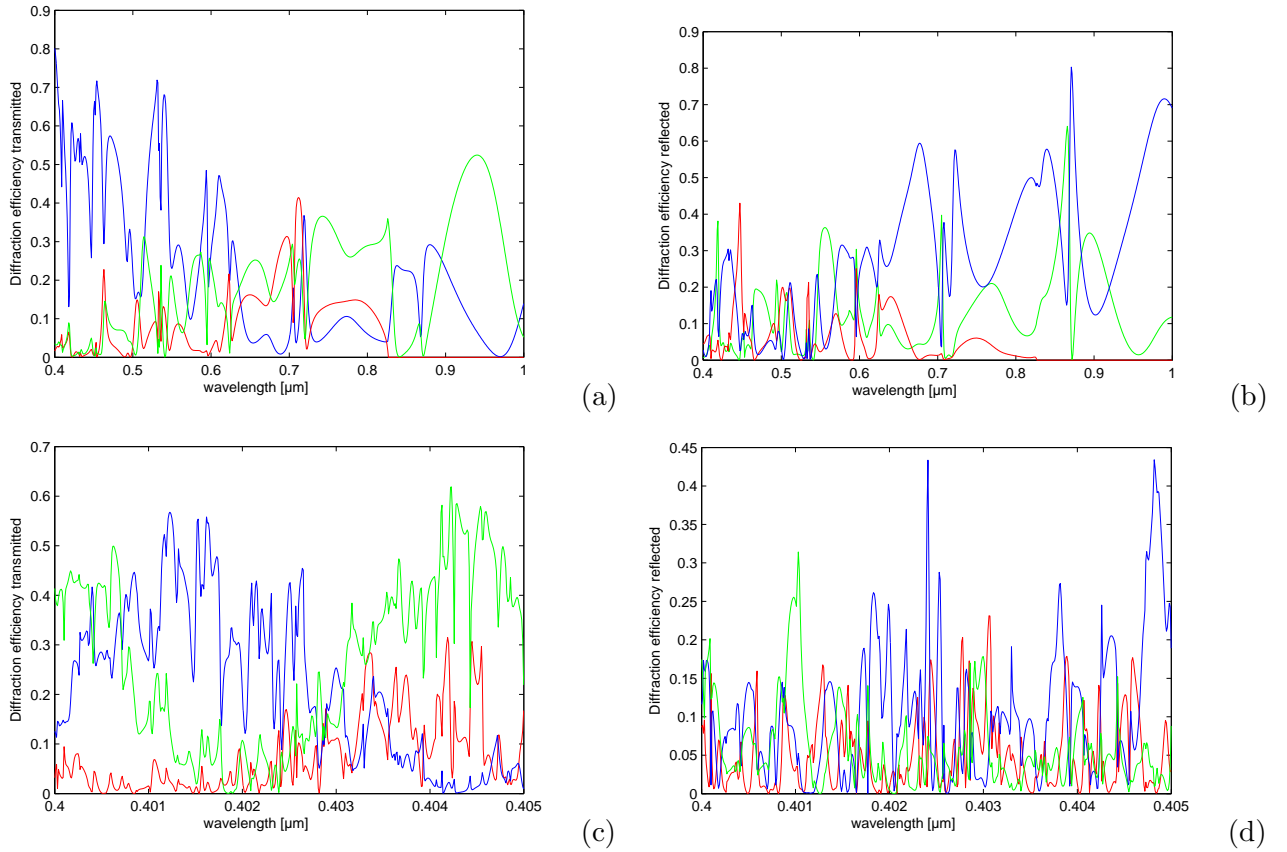


Figure 13: The transmitted (a) and reflected (b) diffraction efficiencies of three orders with a multi layer system with finite thickness: height (of each layer) $= 0.4\mu\text{m}$, $n_I = n_{II} = 1$ and $n_{rd} = 2.04$. The grating period is 0.5 . Here $\Lambda = 0.1\mu\text{m}$ and $\theta = -10^\circ$. The result agreed with that from 'Optiscan'. The transmitted diffraction efficiencies of the -1^{st} (green), 0^{th} (blue) and 1^{st} order (red) are depicted. In (c) and (d) the transmitted and reflected diffraction efficiency of the -1^{st} , 0^{th} and 1^{st} order with another grating of finite thickness: $\Lambda = 1\mu\text{m}$, $\theta = -10^\circ$, height of the grating $= 0.4\mu\text{m}$, height of the substrate $= 40\mu\text{m}$.

4 Design of a three-beam splitter

Now the RCWA solver is to be used to design a three-beam splitter. Therefore we will look for the appropriate choice of parameters. Some will be variable, others stay fixed. A number of $N = 50$ Fourier harmonics appeared to be sufficient for all simulations carried out during optimization. Some first considerations concern the period Λ , as to be seen in the following section.

4.1 Determination of the number of orders

In order to design a three-beam splitter, we have to ask: How can we predict the number of orders a grating transmits? The following rough reasoning shows why that depends on the wavelength λ_0 and the period Λ . We start with Huygens principle when looking at the grating: Here light waves in the slits at $z = 0$ behave like point sources interfering with each other. In the directions where constructive interference takes place, a diffractive order is observed. The order 0 always exists as it relates to those waves from the point sources that go straight downwards. Elementary waves from all equivalent points meet in-phase (figure 14(a)) with the parallel wave from the point source one period further. Hence, these light waves interfere constructively.

But there are also other directions where this happens. The waves propagating under the respective angle are superimposed with the same phase with waves from the equivalent point sources, i.e. the waves are interfering constructively (see figure (14(b))). This occurs when the distance ℓ is a multiple of the wavelength λ , the distance between two wavefronts with equal phase. λ is the wavelength in the material with $\lambda = \frac{\lambda_0}{n_{II}}$. That means

$$\ell = m\lambda \tag{86}$$

with m an integer which indicates the diffraction order. According to the picture one has

$$\sin \alpha = \frac{\ell}{\Lambda}. \tag{87}$$

Both expressions lead to the grating equation

$$\Lambda \sin \alpha = m \frac{\lambda_0}{n_{II}}. \tag{88}$$

This equation may be read as follows: For given wavelength and period, the diffractive order m appears under an angle α .

We also can consider the situation with light coming in under an angle (figure 14(c)). The picture looks similar, but we deal now with two distances we have to take into account: ℓ_1 and ℓ_2 . Incoming waves that pass through adjacent equivalent points differ by l_1 and l_2 in optical path length. When now $l_1 k_I + l_2 k_{II} = m2\pi$ the wavefronts will meet again in phase and interfere constructively. We therefore get the equation

$$\Lambda(n_{II} \sin \alpha + n_I \sin \theta) = m\lambda_0. \tag{89}$$

Now we want to know which values of Λ are allowed, i.e. for which range of periods Λ the grating supports only the fundamental and first diffraction orders (here $\theta = 0^\circ$). For that we have

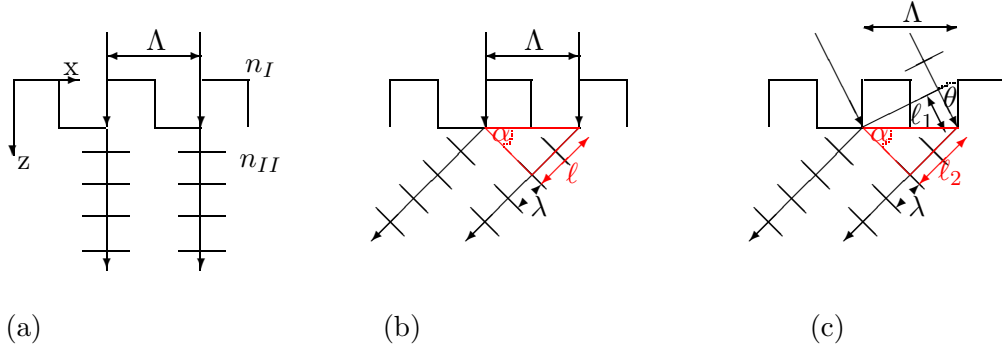


Figure 14: Interference of wavefronts from two equivalent point sources in the periodic grating. (a): the diffractive order 0 and (b) order 2. The wavefronts are interfering constructively. (c) shows interference of wavefronts in a grating with incoming light from above under the angle θ .

to remember that α mustn't reach 90° , otherwise the situation wouldn't make sense any more. So, according to equation (88) the grating supports a diffraction order of 1 and no second order, if

$$2\frac{\lambda_0}{n_{II}} \geq \Lambda > \frac{\lambda_0}{n_{II}}. \quad (90)$$

Also the transmitted electric field according to the RCWA formulation can be used to find out this condition. We consider the equation for the transmitted field (29) and look at $k_{II,zi}$ (equation (32)) for the electromagnetic far field. The corresponding inequation and equation (17) with $\theta = 0^\circ$ lead to the condition:

$$\Lambda > m\frac{\lambda_0}{n_{II}}. \quad (91)$$

This equation tells the number of transmitted orders, depending on the grating period and the wavelength.

4.2 Optimization

The task is now to design a DOE, here a grating that, if illuminated by laser light, transmits only three orders, each containing approximately the same energy. In practice this is the easiest to realise if one concentrates on the orders -1, 0 and 1. As we set $\theta = 0^\circ$ the diffraction efficiency of the orders 1 and -1 are equal. But still the other variables have to be optimized to distribute 1/3 of the transmitted diffraction efficiencies on each of the three orders. This can be expressed by requiring the quantity

$$\text{Err} = (\text{DE}_{t-1} - \frac{1}{3})^2 + (\text{DE}_{t0} - \frac{1}{3})^2 + (\text{DE}_{t1} - \frac{1}{3})^2 \quad (92)$$

to become as small as possible. For every order the difference between the real transmitted energy and the requested one third is calculated, the results are squared and added ¹. Some parameters can be fixed already in advance:

- TE polarization
- incidence angle $\theta = 0^\circ$
- wavelength $\lambda_0 = 0.6328\mu m$ (red laser)
- refractive indices $n_I = 2.04$ and $n_{II} = 1$ (air)

The varied parameters are:

- height d_ℓ of layer ℓ
- number of layers
- duty cycle f_ℓ of layer ℓ

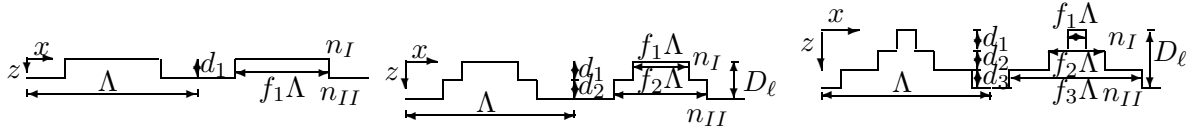
Optimization is realized by first determining the number of layers. For only one layer we therefore are dealing with two parameters, d_1 and f_1 . As for the period, $1.1\lambda, 1.25\lambda, 1.5\lambda, 1.75\lambda$ and 2λ were tried out. We took a stepsize of 0.01 and then calculated Err for all the possible gratings of d_1 within $0 \rightarrow 1\mu m$ and f_1 between 0 and 1. After that, the system with the smallest Err was picked. With a two-layer system, f_2 is added but we keep $d_2 = d_1$, as optimizing with more than three variables results in a computing time of weeks. Also here the possible combinations are calculated and searched for the best values. For practical reasons, the duty cycle of the upper layer had to be smaller than the duty cycle of the lower layer. Due to the direct search method, adding a third layer resulted in a computing time of weeks, as it considers four variables. For this reason a four-layer system was not considered.

First, a one-layer system was optimized. The results are to be seen in table 3a). Following equation (88), the period of $\Lambda = 1.25\lambda$ results in an outcoming angle $\alpha = \pm 53.13^\circ$ of the first modes. It is striking that the teeth of the grating are very narrow. The absolute electric field with a sketch of the grating is depicted in figure 16(a). The landscape (i.e. the function Err) is depicted

¹More generally: the desired value of distribution is subtracted from the diffraction efficiency of the respective order and squared according to the Least Squares method. It is also possible to add a weighting factor to a certain term. For example: a small $\text{Err}' = (\text{DE}_{t-1} - 0.5)^2 + 3(\text{DE}_{t0})^2 + (\text{DE}_{t1} - 0.5)^2$ would deliver a two-beam splitter with emphasis on a suppressed order 0.

Table 3: The optimized gratings and the resulting Err.

	Λ	d_1	d_2	d_3	f_1	f_2	f_3	$DE_{t\pm 1}$	DE_{t0}	Err
a)	1.25λ	$0.89\mu\text{m}$			0.92			31.49%	30.68%	0.0014
b)	1.25λ	$0.67\mu\text{m}$	$0.67\mu\text{m}$		0.62	0.95		32.64%	32.48%	0.00016897
c)	1.25λ	$0.67\mu\text{m}$			0.62			1.11%	86.13%	0.4865
d)	1.25λ	$0.67\mu\text{m}$	$0.68\mu\text{m}$		0.62	0.95		32.82%	32.59%	0.0001088
e)	1.25λ	$0.8\mu\text{m}$	$0.8\mu\text{m}$		0.58	0.79		29.86%	29.55%	0.0039
f)	1.75λ	$0.38\mu\text{m}$	$0.38\mu\text{m}$	$0.38\mu\text{m}$	0.66	0.7	0.74	30.52%	29.94%	0.0027



in figure 15(a). There exist several dips, but the smallest errors are still to be found in the area of the big duty cycles.

When we are switching to a two-layer system, we are keeping the same height for both layers and are therefore dealing with three variables. So we get for every value of the first duty cycle an error landscape as in figure 15(a), but only the configurations where the upper duty cycle is smaller than the lower one are considered. Adding this second layer delivered a better result, see table 3b). The corresponding field is to be seen in figure 16(b). Also, here the teeth of the lower layer are very narrow. The error landscape for $f_1 = 0.62$ is to be seen in figure 15(b). To estimate the influence of the lower layer a similar system with the same values but with only the first layer was considered (table 3c)). Surprisingly, this changed the distribution of the diffraction efficiencies enormously: For the corresponding electric field, see figure 16(c). Apparently, the rods have a significant influence on this grating. This is quite surprising, given their narrow appearance. Now, both duty cycles and only the height of the upper layer were fixed and optimization (over the height of the lower layer) delivered the grating with the specification of table 3d), a slightly better system with even a bit longer lower teeth. The electric field of this grating is displayed in figure 17(a).

Also varying over both heights (still with fixed duty cycles) revealed the same data. So this system was considered to be the best for this project. If one looks for a grating with more robust appearance one should fix a maximum range for the lower layer (0.8). For $\Lambda = 1.25\lambda$ and the same height for both layers this was tested and resulted in the system in table 3e). The result however is slightly worse than configuration (d), for the field see 17(b).

Adding a third layer didn't give better results. For the parameters, see table 3f). The outgoing angle α amounts to 34.85° . The field is to be seen in figure 17(c). At a first glance the regular slope (0.38/0.04) of the duty cycles suggested that a certain form of the teeth could also influence the outcome. Unfortunately, several experiments with this slope on more layers or similar shapes didn't support this theory. It is indicated that this optimized grating is coincidentally the best one for this number of layers.

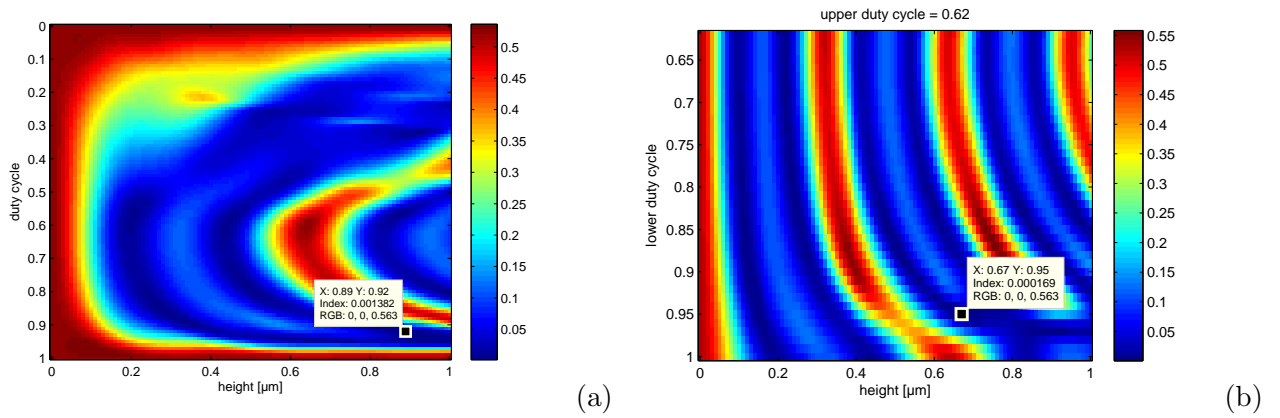


Figure 15: The function Err of a one-layer grating (a) and of a two-layer grating (b), the latter for one specified value of $f_1 = 0.62$.

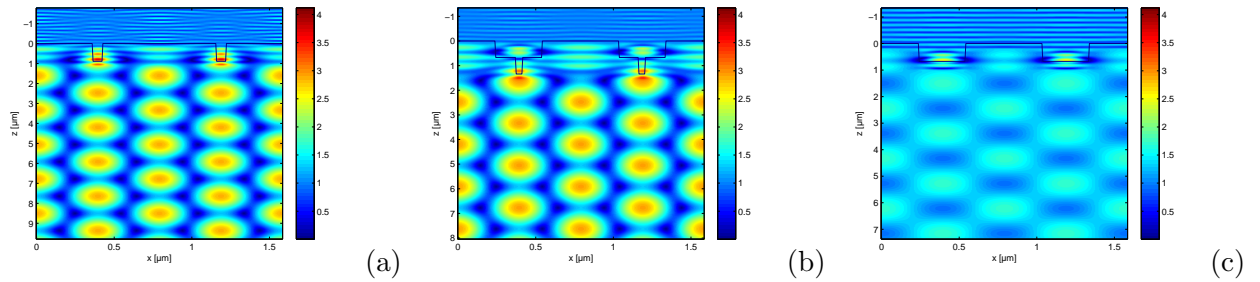


Figure 16: The (absolute) electric fields of the one-layer grating (a), the two-layer grating (b) and the latter without the lower layer (c).

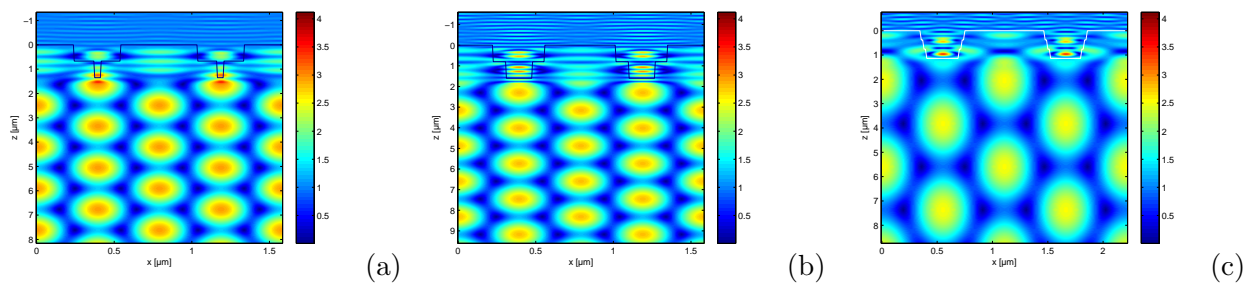


Figure 17: The electric fields of the optimized two-layer system with two different heights (a), a two-layer system with constricted duty cycles (b) and the optimized three-layer system (c) (absolute values).

Also, the gradient descent method was tried out to find a minimum of Err. The heights d_ℓ and duty cycles f_ℓ for three layers were varied, their gradients were calculated to look for the smallest values of Err. But as it seemed the error landscape was so full of local minima, that using this program didn't deliver any feasible results. On the other hand it has been used to verify the optimized gratings already found.

So concludingly, some tests with slightly varying data were performed on grating d). We were mainly interested in the stability of certain parameters: in figure 18 the duty cycle and height of the second layer, the period and the angle of incidence was diversified. This way it was verified that Err at the specified point really reaches a minimum. The graphs also tell about the range in which the parameters can be changed without significant effect on the outcome, i.e. the graphs could provide a basis for an analysis of fabrication tolerances. The lower duty cycle is, as it seems, the most sensitive parameter. Varying the period/wavelength-ratio, on the other hand, has little effect.

We also simulated the same grating for a TM-polarized wave. This resulted in a much worse performance with $DE_{t\pm 1} = 4.9\%$ and $DE_{t0} = 80.53\%$, $Err = 0.3845$. A three-beam splitter performing well for both TE and TM polarization could be optimized by adding the error function (92) for both TE and TM up.

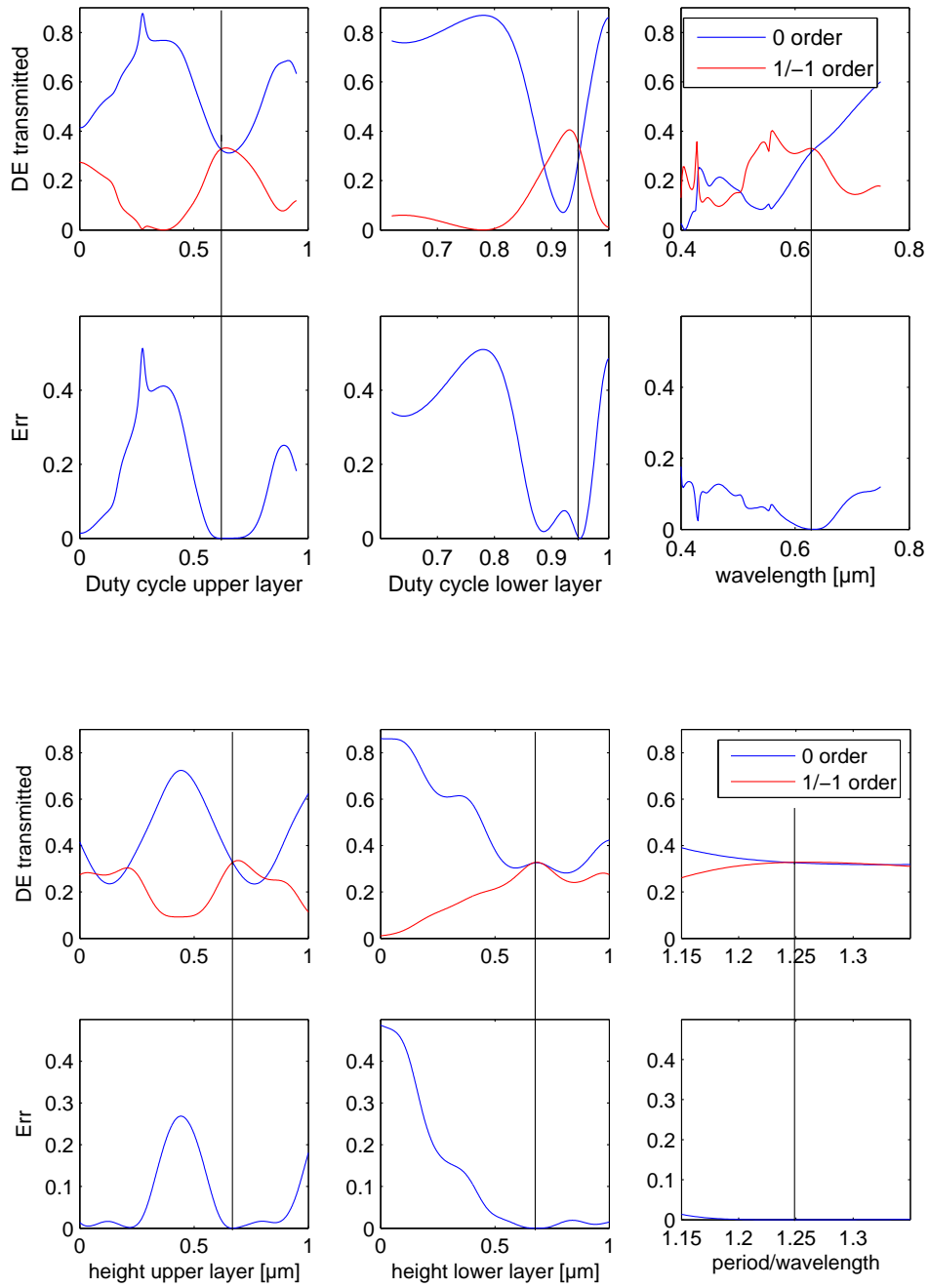


Figure 18: The best system with diversified duty cycles and wavelength (upper graphs) and heights and period (lower graphs). The line marks the original value.

5 Conclusion

The Rigorous Coupled Wave Analysis has been implemented according to the formulation in [1]. Both TE and TM polarization were considered. First, gratings consisting of one layer were examined. Those gratings are periodic in one dimension, while the angle of incidence can also be varied in one dimension. The algorithm calculates the amplitudes of the reflected and transmitted waves. Those can be used to determine the diffraction efficiencies and the electric and magnetic fields. The same is possible for gratings with several layers. The computational effort grows with every added layer since the number of unknowns and matrix inversions increases. The 15 layer system referred to in figure 10 in the TM case with a big period didn't show stable results at all. Five layers needed more Fourier harmonics but could be calculated. Also finite gratings can be analysed.

In principle, although tested so far only for lossless dielectric structures, the technique should work for metallic structures as well, provided that no additional numerical difficulties occur [16], [17]. Note that if we are dealing with a metallic grating, part of the energy is absorbed by the material and therefore (50) wouldn't hold anymore. Also, here a grating of finite thickness should be considered for practical reasons.

By means of this method we designed a two-layer system that divides the energy of a straight incoming laser beam quite homogeneously into three laser rays. It is striking that the second layer is reduced to series of narrow spikes and therefore the system will probably be difficult to produce. Removing the rods or, at least trying to shorten them, resulted in a much worse performance. The same problem is encountered in the one-layer system, it performed best with a very big duty cycle. Results in a constricted, more realistic range were slightly worse. The three-layer grating might be a more practical solution: with rather broad rods it performs worse than the two-layer system but better than the restricted two-layer system.

It surely depends on the application in particular what maximum deviation from the ideal performance is acceptable. This thesis concentrated more on the theory, as these constrained optimizations weren't carried further.

Also including more variables during optimization is an opportunity, such as the period Λ , different heights and more layers. A much more efficient numerical way to optimize would be required. Unfortunately, a straight forward gradient descent method didn't work, as the error landscape seems to be very bumpy and the program got stuck very fast in one of the many local minima. More layers will complicate the structure for producibility, too. Another fast approach would be the direct search method with several increments: we start searching with a rough stepsize, then focus on the area with the lowest Err and search there with a smaller stepsize. This can be repeated as often as necessary.

It would be interesting to know if the parameters of e.g. a two- or four-beam splitter require similar duty cycles and parameters. Possibly these systems wouldn't need such narrow blocks and the producibility wouldn't be a problem at all.

As an extension to the present implementation one could look at an RCWA method with radial coordinates ([10]). In the long run a three-dimensional view on the grating can be used to design not just a simple beam-splitter but also two-dimensional patterns.

References

- [1] 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 of binary gratings. *J. Opt. Soc. Am A*, 12(5): 1068-1076, May 1995.
- [2] http://en.wikipedia.org/wiki/Finite-difference_time-domain_method
- [3] M. G. Moharam and T. K. Gaylord. Rigorous coupled-wave analysis of planar-grating diffraction, *J. Opt. Soc. Am.* 71: 811-818 (1981).
- [4] N. Y. Chang and C. J. Kuo. Algorithm based on rigorous coupled-wave analysis for diffractive optical elements design. *J. Opt. Soc. Am. A*, 18(10): 2491-2501, October 2001.
- [5] Wilson, Raymond G.. Fourier series and optical transform techniques in contemporary optics : an introduction / John Wiley / 1995.
- [6] K. Byun, S. Kim, D. Kim. Design study of highly sensitive nanowire-enhanced surface plasmon resonance biosensors using rigorous coupled wave analysis. *OPTICS EXPRESS*, 13(10): 3737–3742, 16 May 2005.
- [7] E. N. Glytsis and T. K. Gaylord. Rigorous three-dimensional coupled-wave diffraction analysis of single and cascaded anisotropic gratings. *J. Opt. Soc. Am. A*, 4(11): 2061–2080, November 1987.
- [8] Dustin W. Carr, J. P. Sullivan, and T. A. Friedmann. Laterally deformable nanomechanical zeroth-order gratings: anomalous diffraction studied by rigorous coupled-wave analysis. *OPTICS LETTERS*, 28(18): 1636–1638, September 15, 2003.
- [9] M. G. Moharam and T. K. Gaylord. Rigorous coupled-wave analysis of grating diffraction-E-mode polarization and losses. *J. Opt. Soc. Am.*, 73(4): 451–455, April 1983.
- [10] J. M. Jarem. RIGOROUS COUPLED WAVE ANALYSIS OF RADIALY AND AZIMUTHALLY-INHOMOGENEOUS, ELLIPTICAL, CYLINDRICAL SYSTEMS. *Progress In Electromagnetics Research, PIER* 34: 89–115, 2001.
- [11] J. M. Jarem. RIGOROUS COUPLED WAVE ANALYSIS OF BIPOLAR CYLINDRICAL SYSTEMS: SCATTERING FROM INHOMOGENEOUS DIELECTRIC MATERIAL, ECCENTRIC, COMPOSITE CIRCULAR CYLINDERS. *PIER* 43: 181-237, 2003.
- [12] Z. Zylberberg and E. Marorn. Rigorous coupled-wave analysis of pure reflection gratings. *J. Opt. Soc. Am.*, 73 (3): 392–398, March 1983.
- [13] N. Kamiya. Rigorous coupled-wave analysis for practical planar dielectric gratings: 1. Thickness-changed holograms and some characteristics of diffraction efficiency. *APPLIED OPTICS*, 37(25): 5843–5853, 1 September 1998.
- [14] J. M. Jarem and P. P. Banerjee. Rigorous coupled-wave analysis of photorefractive reflection gratings. *J. Opt. Soc. Am. B*, 15(7): 2099–2106, July 1998.

- [15] J.Jarem. A Rigorous Coupled-Wave Analysis and Crossed-Diffraction Grating Analysis of Radiation and Scattering from Three-Dimensional Inhomogeneous Objects. IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, 46(5): 740–741, MAY 1998.
- [16] L. Li and C. W Haggans. Convergence of the coupled-wave method for metallic lamellar diffraction gratings. J. Opt. Soc. Am. A, 10(6): 1184–1189, June 1993.
- [17] G. Granet and B. Guizal. Efficient implementation of the coupled-wave method for metallic lamellar gratings in TM polarization. J. Opt. Soc. Am. A, 13(5): 1019–1023, May 1996.
- [18] M.G. Moharam, D.A. Pommet, E.B. Grann 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.
- [19] S. Sinzinger, J. Jahns. Microoptics. 2nd edition. 2003 WILEY-VCH GmbH & Co. KGaA, Weinheim.
- [20] N. Chateau and J.-P. Hugonin. Algorithm for the rigorous coupled-wave analysis of grating diffraction. J. Opt. Soc. Am. A, 11(4): 1321-1331, April 1994.
- [21] Goodman, Joseph W.. Introduction to Fourier optics / 2nd ed / McGraw- Hill / 1996.
- [22] Okan K. Ersoy. Diffraction, Fourier Optics and Imaging. John Willey & Sons, Inc., 2007.
- [23] S. Peng and G. M. Morris. Efficient implementation of rigorous coupled wave analysis for surface-relief gratings. J. Opt. Soc. Am A, 12(5): 1087-1096, May 1995.
- [24] J. Elschner, G. Schmidt. A rigorous numerical method for the optimal design of binary gratings. WIAS Berlin, 1991.
- [25] Hirsch, Morris W., Smale, Stephen and Devaney, Robert L.. Differential equations, dynamical systems, and an introduction to chaos / 2nd ed / Elsevier Academic Press / 2004.