A rigorous numerical method for the optimal design of binary gratings

J. Elschner and G. Schmidt
WIAS Berlin

1991 Mathematics Subject Classification. 78-05, 78A45, 35J20, 65N30, 49J20.

Keywords. Diffraction problems, Helmholtz equation, transmission problem, strongly elliptic variational formulation, generalized FEM, gradient methods

Abstract

In this paper we describe recent developments in the application of mathematical and computational techniques to the problem of designing binary gratings on top of a multilayer stack in such a way that the propagating modes have a specified intensity or phase pattern for a chosen range of wavelengths or incidence angles. This optimal design problem is solved by a minimization algorithm based on gradient descent, the exact calculation of gradients of certain functionals with respect to the parameters of the grating profile and the thickness of the layers. For the computation of diffraction efficiencies and of the gradients we use a reliable finite element method which originates from variational formulations of the diffraction problems. We provide several numerical examples including polarisation gratings and beam splitters to demonstrate the efficiency of the algorithm.

1 Introduction

The practical application of diffractive optics technology has driven the need for mathematical models and numerical codes both to provide rigorous solutions of the full electromagnetic vector-field equations for complicated grating structures, thus predicting performance given the structure, and to carry out optimal design of new structures.

Periodic gratings can be modeled as quasi–periodic transmission problems for the Helmholtz equation in the whole plane. Special difficulties are associated with the numerical solution of these problems due to the highly oscillatory nature of waves and interfaces. For the evaluation of rigorous solutions for a given structure, the direct diffraction problem, various methods have been proposed. Among the most well known are modal expansion, differential and integral methods (cf. the classical monograph [1] and recent extensions and improvements in e.g. [2–7]). These methods turned out to be efficient for solving the direct diffraction problem for certain classes of grating structures, but it is difficult to find any rigorous treatment of convergence in the literature. Such a convergence analysis can be found in the case of smooth interfaces between different materials for integral equation methods and the analytical continuation method introduced in [8]. In the case of binary structures, whose surface profile is given by a piecewise constant function, the mathematical complexities are amplified by singularities of the solutions. Recently, a new variational approach was proposed (see [9–10] and the references therein), which appears to be well adapted for the analytical and numerical treatment of very general diffraction structures as well as complex materials and allows straightforward extensions to diffraction problems for conical mounting and crossed gratings ([11–12]). In particular, this approach is the basis for the convergence analysis of finite element solution methods, introduced in [13–16].

But it is more important that the variational approach leads to effective formulae for the gradient of cost functionals arising in optimal design problems as shown in [17], [15], such that gradient based minimization methods can be used to find gratings with specified optical functions. There have been a number of papers from the engineering community that are concerned with the optimization of periodic gratings. In some of these papers (e.g. [18–19]) descent methods are considered using approximations of the gradient by simple difference quotients, which, however, are very expensive for a large number of parameters.

In the present paper we apply some mathematical results from [15] to the model problem of designing binary gratings on top of a multilayer stack in such a way that the propagating modes have a specified intensity or phase pattern for a chosen range of wavelengths or incidence angles. First we present the variational formulations of the diffraction problems for TE and TM polarization and give a summary of some existence and uniqueness results. In section 3 we consider a typical optimal design problem, formulate the cost functional and write down the formulae for the gradients with respect to the parameters of the grating profile and the thicknesses of the layers. Then the optimal design problem can be solved by minimization algorithms based on gradient descent. For the computation of diffraction efficiencies and of the gradients we use a reliable numerical method which originates from the variational formulations. This method, which combines a generalized finite element method in the grating structure with Fourier expansions in the multilayer system, is discussed in section 4. Finally we provide some numerical examples to demonstrate the convergence properties of this method for evaluating diffraction efficiencies and gradients.

2 Variational formulation of the direct scattering problem

Consider a binary grating of period d, with height H and transition points t_j at the top of a stack of layers of thickness h_j. The materials are nonmagnetic with the permeability μ_0 and have the dielectric constants ϵ. The coordinate system is chosen such that the diffraction problem is invariant in the x_3 direction and that the x_1 axis is parallel to the layers. Thus the problem is determined by the function $\epsilon(x_1, x_2)$ which is d–periodic in x_1. We assume that the material above the grating profile Γ is homogeneous with $\epsilon = \epsilon^+ > 0$. Below Γ the material may be inhomogeneous and we assume that the function $\epsilon = \epsilon^-$ is piecewise constant corresponding to the different layers and constant for the substrate. Further we suppose that the ϵ^- can be complex valued with $\operatorname{Im}\epsilon^- \geq 0$ and $\operatorname{Re}\epsilon^- > 0$ if $\operatorname{Im}\epsilon^- = 0$.

Assume that an incoming plane wave with time dependence $\exp(-i\omega t)$ is normally incident upon the grating from the top with the angle of incidence $\theta \in (-\pi/2, \pi/2)$. In either case of polarization, one of the fields $\mathbf{E}$ or $\mathbf{H}$ remains parallel to the grooves and is therefore determined by a single scalar quantity $v = v(x_1, x_2)$ (equal to the transverse component of $\mathbf{E}$ in the TE case and to the transverse component of $\mathbf{H}$ in the TM case). The function v satisfies twodimensional Helmholtz equations

$$\Delta v + \omega^2 \mu_0 \epsilon\, v = 0$$

in the regions with constant permittivity, together with the usual outgoing wave condition at infinity. At the material interfaces the solutions are subjected to well known transmission conditions. For TE polarisation the solution and its normal derivative $\partial_n v$ have to cross the interface continuously, whereas in TM polarisation the product $\epsilon^{-1}\partial_n v$ has to be continuous.

The diffraction problems admit variational formulations in a bounded periodic cell. In the following $k = \omega(\mu_0\epsilon)^{1/2}$ denotes the piecewise constant function taking the values k^+ and of the function $k^- = \omega(\mu_0\epsilon^-)^{1/2}$ (k_g, k_1, k_2 and k_3 in figure 2.1), which is chosen such that

$$\operatorname{Re} k^+ > 0\,, \quad \operatorname{Re} k^- > 0\,,\ \operatorname{Im} k^- \geq 0\,. \tag{2.1}$$

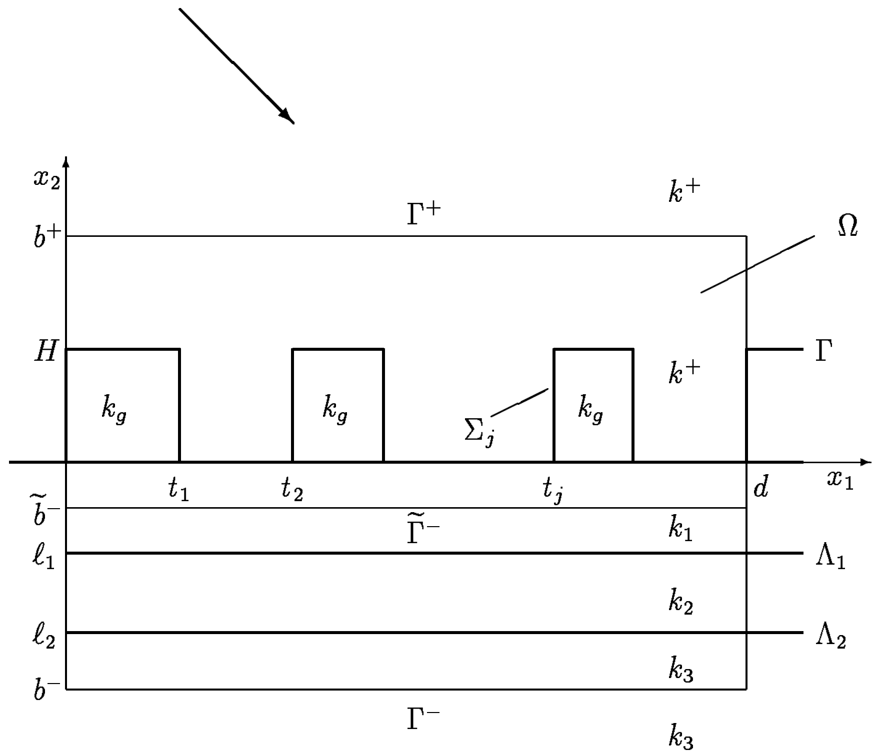

Figure 2.1: Problem geometry

Note that $k = \omega\nu(\mu_0\epsilon_0)^{1/2}$, where ϵ_0 is the permittivity of the vacuum and $\nu = (\epsilon/\epsilon_0)^{1/2}$ denotes the optical index. The incoming wave has the form $v^i = \exp(i\alpha x_1 - i\beta x_2)$, where $\alpha = k^+ \sin\theta$, $\beta = k^+ \cos\theta$. If we introduce two artificial boundaries $\Gamma^\pm = \{x_2 = b^\pm\}$ lying above Γ and below the layer structure, respectively, denote by Ω the rectangle $(0,d)\times(b^-,b^+)$ and define the d-periodic function $u = v\exp(-i\alpha x_1)$, then the diffraction problem for TE polarization is equivalent to the variational equation

$$B_{TE}(u,\varphi) := \int\limits_{\Omega} \nabla_\alpha u \cdot \overline{\nabla_\alpha \varphi} - \int\limits_{\Omega} k^2\, u\, \bar\varphi + \int\limits_{\Gamma^+} (T_\alpha^+ u)\, \bar\varphi + \int\limits_{\Gamma^-} (T_\alpha^- u)\, \bar\varphi$$

$$= -\int\limits_{\Gamma^+} 2i\beta \exp(-i\beta b^+)\, \bar\varphi\,, \quad \forall\varphi\,, \tag{2.2}$$

where $\nabla_\alpha = (\partial_{x_1,\alpha}, \partial_{x_2}) := \nabla + i(\alpha, 0)$. The functions $T_\alpha^\pm u$ are defined on $\Gamma^\pm$ as

$$(T_\alpha^\pm u)(x_1, b^\pm) := -\sum_{n=-\infty}^{\infty} i\beta_n^\pm \hat u_n^\pm \exp(inKx_1)\,, \tag{2.3}$$

where $K = 2\pi/d$ and $\hat u_n^\pm$ denote the Fourier coefficients of $u(x_1, b^\pm)$

$$\hat u_n^\pm = \frac{1}{d}\int\limits_0^d u(x_1, b^\pm)\exp(-inKx_1)\, dx_1\,.$$

The numbers $\beta_n^\pm$ are defined as

$$\beta_n^\pm = \beta_n^\pm(\alpha) := ((k^\pm)^2 - \alpha_n)^{1/2}\,,\ 0 \le \arg\beta_n^\pm < \pi\,,$$

where as usual $\alpha_n = \alpha + nK$ and $k^- = k^-(x_1, b^-)$.

The variational equation (2.2) should be satisfied for all test functions $\varphi \in H^1_p(\Omega)$, that is the function space of all complex–valued functions φ which are d-periodic in x_1 and together with their first–order partial derivatives square integrable in Ω. This variational formulation is very useful, because the transmission and outgoing wave conditions are enforced implicitly and it allows to seek the solution in the function space $H^1_p(\Omega)$, which is natural for second order partial differential equations on non-smooth domains. Here one can apply well established methods for the analysis and numerical solution of the diffraction problems.

Note that any solution of (2.2) satisfies on $\Gamma^\pm$ the boundary conditions

$$\partial_n u|_{\Gamma^+} + T^+_\alpha u|_{\Gamma^+} = -2i\beta \exp(-i\beta b^+)\,, \quad \partial_n u|_{\Gamma^-} + T^-_\alpha u|_{\Gamma^-} = 0\,. \tag{2.4}$$

which implies the Fourier series expansion

$$\begin{aligned} u(x_1, b^+) &= \sum_{n=-\infty}^{\infty} A^+_n \exp(i\beta^+_n b^+) \exp(inKx_1) + \exp(-i\beta b^+)\,, \\ u(x_1, b^-) &= \sum_{n=-\infty}^{\infty} A^-_n \exp(-i\beta^-_n b^-) \exp(inKx_1)\,, \end{aligned} \tag{2.5}$$

Thus the operators $T^\pm_\alpha$ are the Dirichlet–to–Neumann mappings

$$\partial_n u^\pm|_{\Gamma^\pm} = -T^\pm_\alpha u^\pm|_{\Gamma^\pm} \tag{2.6}$$

for functions of the form

$$u(x_1, x_2) = \sum_{n=-\infty}^{\infty} A^\pm_n \exp(\pm i\beta^\pm_n x_2) \exp(inKx_1) \quad |x_2| \geq |b^\pm| \quad x_2 \gtrless b^\pm\,.$$

Similarly the TM diffraction problem can be formulated as follows:

$$\begin{aligned} B_{TM}(u, \varphi) &:= \int_\Omega \frac{1}{k^2} \nabla_\alpha u \cdot \overline{\nabla_\alpha \varphi} - \int_\Omega u\,\bar{\varphi} + \int_{\Gamma^+} \frac{1}{(k^+)^2} (T^+_\alpha u)\,\bar{\varphi} + \int_{\Gamma^-} \frac{1}{(k^-)^2} (T^-_\alpha u)\,\bar{\varphi} \\ &= -\int_{\Gamma^+} \frac{2i\beta}{(k^+)^2} \exp(-i\beta b^+)\,\bar{\varphi}\,, \quad \forall\, \varphi \in H^1_p(\Omega)\,. \end{aligned} \tag{2.7}$$

Based on the variational formulation of the diffraction problems the following existence and uniqueness results can be established ([15], [9]):

1° The TE and TM diffraction problems admit solutions $u \in H^1_p(\Omega)$ for all $\omega > 0$ and θ. These solutions are unique for all but a sequence of countable frequencies ω_j, $\omega_j \to \infty$.

2° For TE polarisation the solution $u(x_1, x_2) = E_{x_3}(x_1, x_2) \exp(-i\alpha x_1)$ has square integrable second–order partial derivatives, $u \in H^2_p(\Omega)$.

3° In the TM case the solution $u(x_1, x_2) = H_{x_3}(x_1, x_2) \exp(-i\alpha x_1)$ may have singularities at the corner points $(t_j, 0)$ and (t_j, H) of the grating. More precisely, near corner points there holds $u = r^\lambda f + g$, where r denotes the distance to the corner point, the exponent $0 < \lambda < 1$ is determined by the optical index of the grating material and f, g are some smoother functions. Hence the partial derivatives of u are only of the form $r^{\lambda-1} f + g$, i.e. the electric field components E_{x_1} and E_{x_2} are strongly singular and the normal derivative of $\partial_n u$ on Γ does not satisfy the Meixner condition $\partial_n u \in L^2(\Gamma)$, in general. This may lead to slow convergence of methods based on truncated modal or Fourier series representations of the field components.

4° Introduce the set of exceptional values (Rayleigh frequencies)

$$\mathcal{R}(\epsilon) = \{(\omega, \theta) : \exists\, n \in \mathbf{Z} \text{ such that } (nK + \omega(\mu\epsilon^+)^{1/2} \sin\theta)^2 = \omega^2 \mu \epsilon^\pm\}\,.$$

If for $(\omega_0, \theta_0) \notin \mathcal{R}(\epsilon)$ one of the diffraction problems is uniquely solvable, then the solution u depends analytically on ω and θ in a neighbourhood of this point.

5° If one of the materials below Γ is absorbing then the TE problem has a unique solution for all frequencies $\omega > 0$.

6° If one of the layer materials is absorbing then the TM problem has a unique solution for all frequencies $\omega > 0$.

7° Let $\epsilon(x) > 0$ for $x \in \Omega$. Suppose that there exists $\tau \in \mathbb{R}$ such that

$$\Big((x_2 + \tau)\frac{\partial \epsilon}{\partial x_2}, v\Big)_{L^2(\Omega)} \geq 0 \quad \text{for all } v \geq 0\,.$$

Then the TE diffraction problem is uniquely solvable for $\omega > 0$.
(This condition is always satisfied if only two materials are present.)

Note that the variational formulation of the diffraction problems and the validity of the corresponding mathematical results are not restricted to binary or other rectangular grating profiles. They remain valid for general piecewise constant functions k satisfying condition (2.1), hence the presented approach is applicable to rather complex grating structures. But here we focus on the case of binary profiles, for which optimal design problems will be considered.

Define the finite sets of indices $P^\pm = \{n \in \mathbf{Z} : \beta_n^\pm \in \mathbb{R}\}$. Then the Rayleigh amplitudes A_n^+ $(n \in P^+)$ resp. A_n^- $(n \in P^-)$, which are called the reflection resp. transmission coefficients, correspond to the propagating modes of u. Note that $P^- = \emptyset$ if $\operatorname{Im} k^-(x_1, b^-) \neq 0$. The reflection and transmission coefficients A_n^+ $(n \in P^+)$ resp. A_n^- $(n \in P^-)$, which correspond to the propagating modes of u, are determined by the Fourier coefficients of u on the artificial boundaries $\Gamma^\pm$

$$A_0^+ = -\exp(-2i\beta b^+) + \exp(-i\beta b^+)\,\hat{u}_0^+\,,\; A_n^\pm = \exp(\mp i\beta_n^\pm b^\pm)\,\hat{u}_n^\pm\,, \quad n \in P^+ \backslash \{0\}\,, n \in P^-$$

The reflected and transmitted efficiencies are defined by

$$e_n^{TE,\pm} = (\beta_n^\pm/\beta)|A_n^\pm|^2\,, \quad e_n^{TM,+} = (\beta_n^+/\beta)|A_n^+|^2\,, \;\; e_n^{TM,-} = (k^+/k^-)^2(\beta_n^-/\beta)|A_n^-|^2\,.$$

3 An optimal design problem

A typical minimization problem occurring in the optimal design of binary gratings on some multilayer system is the following. Assume that the number of transition points and of thin–film layers is fixed and, for given numbers $c_r^{TE,\pm}, c_r^{TM,\pm} \in \{-1, 0, 1\}$, define the functional

$$J(\Xi) = \sum_{r\in P^+} \left(c_r^{TE,+} e_r^{TE,+} + c_r^{TM,+} e_r^{TM,+}\right) + \sum_{r\in P^-} \left(c_r^{TE,-} e_r^{TE,-} + c_r^{TM,-} e_r^{TM,-}\right) . \qquad (3.1)$$

Note that the efficiencies $e_r^\pm$ are functions of the grating profile Γ and the layer interfaces Λ_j. If we fix one transition point t_0 at the origin the efficiencies $e_r^\pm$ are therefore functions of $t_1, ..., t_{m-1}, H, \ell_1, ...\ell_p$ (cf. figure 1). Now the minimization problem reads as follows:

Find transition points $t_1^0, ..., t_{m-1}^0$ and the height H^0 of the binary grating profile Γ^0 as well as thicknesses of the layer structure such that

$$\min_{(t_1,...,t_{m-1},H,\ell_1,...\ell_p)\in K} J(\Xi) = J(\Xi^0)\,, \qquad (3.2)$$

where K is some compact set in the parameter space $\mathbb{R}^{m+p}$ reflecting e.g. natural constraints on the design of the grating and the thin–film layers. Note that the choice $c_r^\pm = -1$ resp. $c_r^\pm = 1$ in (3.1) amounts to maximizing resp. minimizing the efficiency of the corresponding reflected or transmitted propagating mode of order n.

Other minimization problems:

1° If one wants to obtain prescribed values for certain reflection and transmission efficiencies, given by the index sets $I^+ \subset P^+$ and $I^- \subset P^-$, the following smooth functional can be useful

$$\sum_{r\in I^+} \left(|e_r^{TE,+} - c_r^{TE,+}|^2 + |e_r^{TM,+} - c_r^{TM,+}|^2\right)$$
$$+\sum_{r\in I^-} \left(|e_r^{TE,-} - c_r^{TE,-}|^2 + |e_r^{TM,-} - c_r^{TM,-}|^2\right) \to \min$$

2° The optimal design of a grating providing a given phase shift φ between the r-th reflected TE and TM mode can be performed using the functional

$$-e_r^{TE,+} - e_r^{TM,+} + |A_r^{TE,+} - \exp(i\varphi)A_r^{TM,+}|^2 \to \min \qquad (3.3)$$

Obviously many other functionals are possible especially if a corresponding optimization over a range of wavelengths or incidence angles is required.

To find local minima of these functionals, the method of gradient descent or other gradient–type methods can be applied. Thus we must calculate the gradient of $J(\Xi)$, for example, which can be easily expressed in terms of the partial derivatives $D_j A_r^\pm(\Gamma)$ (with respect to the transition points $t_1, \ldots, t_m - 1$, the height H and the layer thicknesses, given by the coordinates ℓ_j) of the reflection and transmission coefficients in both the TE and TM case. Here we propose to use rigorous gradient formulae based on the solution of the direct and its adjoint problem, instead of simple difference quotients which are very expensive to compute for a large number of parameters.

The gradient of $J(\Xi)$ is given by

$$\begin{aligned} D_j J(\Xi) = \sum_{r\in P^+} 2(\beta_r^+/\beta)\Big\{ & c_r^{TE,+}\mathrm{Re}\,(\overline{A_r^{TE,+}(\Xi)} D_j A_r^{TE,+}(\Xi)) \\ & + c_r^{TM,+}\mathrm{Re}\,(\overline{A_r^{TM,+}(\Xi)} D_j A_r^{TM,+}(\Xi))\Big\} \\ + \sum_{r\in P^-} 2(\beta_r^-/\beta)\Big\{ & c_r^{TE,-}\mathrm{Re}\,(\overline{A_r^{TE,-}(\Xi)} D_j A_r^{TE,-}(\Xi)) \\ & + (k^+/k^-)^2 c_r^{TM,-}\mathrm{Re}\,(\overline{A_r^{TM,-}(\Xi)} D_j A_r^{TM,-}(\Xi))\Big\}. \end{aligned} \tag{3.4}$$

Once one has derived explicit formulae for those partial derivatives, it is possible to compute also the gradients for a much more general class of functionals involving the Rayleigh coefficients for a given range of incidence angles or wavelengths.

The formulae for all components of the gradient of $A_r^\pm$ in the TE case take the form:

$$\begin{aligned} D_j A_r^\pm(\Xi) &= (-1)^{j-1}(k_g^2 - (k^+)^2)\int_{\Sigma_j} u\bar{w}_\pm dx_2\,, \quad j = 1,\dots,m-1\,, \\ D_m A_r^\pm(\Xi) &= (k_g^2 - (k^+)^2)\int_{\Sigma_m} u\bar{w}_\pm dx_1\,, \\ D_{m+j} A_r^\pm(\Xi) &= (k_j^2 - k_{j+1}^2)\int_{\Lambda_j} u\bar{w}_\pm dx_1\,, \quad j = 1,\dots,p\,, \end{aligned} \tag{3.5}$$

where u is the solution to the TE diffraction problem (2.2) and the functions $w_\pm$ solve the adjoint TE problems

$$B_{TE}(\varphi, w_\pm) = \frac{\exp(\mp i\beta_r^\pm b^\pm)}{d}\int_{\Gamma^\pm} \varphi\,\exp(-irKx_1)\,dx_1\,, \quad \forall\varphi\in H_p^1(\Omega)\,. \tag{3.6}$$

Here Σ_m is the union of all upper horizontal segments of Γ, whereas Σ_j ($j = 1,\dots,m-1$) denotes the vertical segment at the transition point t_j.

In the TM case the gradient formulae involve the partial derivatives of the solution of direct and adjoint problems at the interfaces. If the optical index of the grating material is such that the solution u satisfies the Meixner condition then

$$\begin{aligned} D_j A_r^\pm(\Xi) &= (-1)^{j-1}(k_g^2 - (k^+)^2)\int_{\Sigma_j} gr(u)\cdot\overline{gr(w_\pm)}\,dx_2\,,\ j = 1,\dots,m-1\,, \\ D_m A_r^\pm(\Xi) &= (k_g^2 - (k^+)^2)\int_{\Sigma_m} gr_H(u)\cdot\overline{gr_H(w_\pm)}\,dx_1\,, \\ D_{m+j} A_r^\pm(\Xi) &= (k_j - k_{j+1}^2)\int_{\Sigma_j} gr_j(u)\cdot\overline{gr_j(w_\pm)}\,dx_2\,,\ j = 1,\dots,p\,, \end{aligned} \tag{3.7}$$

Here, u is the solution of the direct TM problem (2.7), the functions $w_\pm$ solve the adjoint problem

$$B_{TM}(\varphi, w_\pm) = \frac{\exp(\mp i\beta_r^\pm b^\pm)}{d}\int_{\Gamma^\pm} \varphi\,\exp(-irKx_1)\,dx_1\,, \quad \forall\varphi\in H_p^1(\Omega)\,, \tag{3.8}$$

and

$$gr(u) = \frac{1}{k^+ k_g} \left(\frac{k_g}{k^+} \partial_{x_1,\alpha} u \Big|_{\Sigma_j}^+ , \partial_{x_2} u \Big|_{\Sigma_j}^+ \right) = \frac{1}{k^+ k_g} \left(\frac{k^+}{k_g} \partial_{x_1,\alpha} u \Big|_{\Sigma_j}^- , \partial_{x_2} u \Big|_{\Sigma_j}^- \right) ,$$

$$gr_H(u) = \frac{1}{k^+ k_g} \left(\partial_{x_1,\alpha} u \Big|_{\Sigma_m}^+ , \frac{k_g}{k^+} \partial_{x_2} u \Big|_{\Sigma_m}^+ \right) = \frac{1}{k^+ k_g} \left(\partial_{x_1,\alpha} u \Big|_{\Sigma_m}^- , \frac{k^+}{k_g} \partial_{x_2} u \Big|_{\Sigma_m}^- \right) ,$$

$$gr_j(u) = \frac{1}{k_j k_{j+1}} \left(\partial_{x_1,\alpha} u \Big|_{\Lambda_j}^+ , \frac{k_{j+1}}{k_j} \partial_{x_2} u \Big|_{\Lambda_j}^+ \right) = \frac{1}{k_j k_{j+1}} \left(\partial_{x_1,\alpha} u \Big|_{\Lambda_j}^- , \frac{k_j}{k_{j+1}} \partial_{x_2} u \Big|_{\Lambda_j}^- \right) .$$

If the Meixner condition is not fulfilled then the gradient formula has to be modified by an additional term depending on λ.

Concerning the solvability of the adjoint problems (3.6) and (3.8) the same existence and uniqueness results as for the direct problems remain valid. In general the solutions have no physical interpretation. It can be seen that the the solution of the adjoint TE or TM problem solve the corresponding diffraction problem with the complex conjugate wavenumbers $\overline{k}$ and a special radiation condition. For example, for w_- this condition takes the form

$$\begin{aligned} w_-(x_1,x_2) &= \sum_{n=-\infty}^{\infty} A_n^+ \exp(-i\overline{\beta_n^+} x_2) \exp(inKx_1) , \qquad x_2 \ge b^+ , \\ w_-(x_1,x_2) &= \sum_{n=-\infty}^{\infty} A_n^- \exp(i\overline{\beta_n^-} x_2) \exp(inKx_1) + \frac{iC}{4\pi \overline{\beta_r^-}} \exp(-i\overline{\beta_r^-} x_2) \exp(irKx_1) , \\ & \qquad x_2 \le b^- \end{aligned} \tag{3.9}$$

with $C = 1$ for TE and $C = (k^+/k^-)^2$ for TM.

Note that in order to compute all partial derivatives of functionals arising in the optimal design of binary gratings it is sufficient to solve the direct TE and TM diffraction problem and only one corresponding adjoint problem. We demonstrate this for the functional $J(\Xi)$ defined in (3.1).

From formulae (3.4), (3.5) and (3.7) we obtain by linearity that the components of the gradient of $J(\Xi)$ are equal to

$$D_j J(\Xi) = (-1)^{j-1} \mathrm{Re} \left\{ (k_g^2 - (k^+)^2) \Big(\int_{\Sigma_j} u^{TE} \, \overline{w^{TE}} \, dx_2 + \int_{\Sigma_j} gr(u^{TM}) \cdot \overline{gr(w^{TM})} \, dx_2 \Big) \right\}, \quad j = 1, \ldots, m-1,$$

$$D_m J(\Xi) = \mathrm{Re} \left\{ (k_g^2 - (k^+)^2) \Big(\int_{\Sigma_m} u^{TE} \, \overline{w^{TE}} \, dx_1 + \int_{\Sigma_m} gr_H(u^{TM}) \cdot \overline{gr_H(w^{TM})} \, dx_1 \Big) \right\},$$

$$D_{m+j} J(\Xi) = \mathrm{Re} \left\{ (k_j^2 - k_{j+1}^2) \Big(\int_{\Lambda_j} u^{TE} \, \overline{w^{TE}} \, dx_1 + \int_{\Lambda_j} gr_j(u^{TM}) \cdot \overline{gr_j(w^{TM})} \, dx_1 \Big) \right\}, \quad j = 1, \ldots, p,$$

where u^{TE} and u^{TM} are the solutions of the direct TE and TM problems, respectively, and w^{TE}, w^{TM} solve the following adjoint problems

$$\begin{aligned} B_{TE}(\varphi, w^{TE}) = & \sum_{r\in P^+} c_r^{TE,+} \overline{A_r^{TE,+}} \frac{2\beta_r^+ \exp(-i\beta_r^+ b^+)}{d\,\beta} \int_{\Gamma^+} \varphi \exp(-inKx_1)\, dx_1 \\ & + \sum_{r\in P^-} c_r^{TE,-} \overline{A_r^{TE,-}} \frac{2\beta_r^- \exp(i\beta_r^- b^-)}{d\,\beta} \int_{\Gamma^-} \varphi \exp(-inKx_1)\, dx_1 \,, \end{aligned} \tag{3.10}$$

$$\begin{aligned} B_{TM}(\varphi, w^{TM}) = & \sum_{r\in P^+} c_r^{TM,+} \overline{A_r^{TM,+}} \frac{2\beta_r^+ \exp(-i\beta_r^+ b^+)}{d\,\beta} \int_{\Gamma^+} \varphi \exp(-inKx_1)\, dx_1 \\ & + \sum_{r\in P^-} c_r^{TM,-} \overline{A_r^{TM,-}} \frac{2\beta_r^- (k^+)^2 \exp(i\beta_r^- b^-)}{d\,\beta\,(k^-)^2} \int_{\Gamma^-} \varphi \exp(-inKx_1)\, dx_1 \,, \\ & \forall \varphi \in H_p^1(\Omega)\,. \end{aligned} \tag{3.11}$$

Here $A_r^{TE,\pm}$ and $A_r^{TM,\pm}$ denote the Rayleigh amplitudes of u^{TE} and u^{TM}, respectively.

Note that for simple difference approximations of the gradient the number of the direct problems to be solved is at least equal to the number of optimization parameters, whereas the computational costs for solving adjoint and direct problems are the same.

4 Implementation and numerical results

Having described the variational formulation and some basic mathematical properties of the direct diffraction problems as well as the gradient formulae and the variational equations of the adjoint problems, we now consider the numerical solution method of these variational problems.

The proposed method combines a finite element method (FEM) in the grating region, where the solutions are not smooth, with Rayleigh series expansions of the solution within the different layers below the grating.

The FEM is a well–established numerical method for solving boundary value problems for elliptic partial diffraction equations, which is based on the variational approach and the approximation of multivariate functions by piecewise polynomials.

As discussed in Sections 2 and 3 the direct and adjoint problems (2.2), (2.7), (3.10), and (3.11) has the form: find $u \in H_p^1(\Omega)$ satisfying the equations

$$a(u,\varphi) = (f,\varphi)\,, \quad \forall \varphi \in H_p^1(\Omega)\,, \tag{4.1}$$

where $a(u,\varphi)$ is a continuous sesquilinear form (linear in u and antilinear in φ), and (f,φ) stands for a linear and continuous functional on the function space $H_p^1(\Omega)$.

In the FEM the domain Ω is partitioned into a sequence Ω_h of simple subdomains with maximum mesh size h and a family S^h of finite–dimensional subspaces of $H_p^1(\Omega)$ is defined, usually a space of piecewise polynomials subordinate to the corresponding partition. The finite element approximations $u_h \in S^h$ of the solution u of (4.1) are obtained from the equations

$$a(u_h,\varphi_h) = (f,\varphi_h)\,, \quad \forall \varphi_h \in S^h\,. \tag{4.2}$$

If one chooses a basis $\varphi_j,\ j = 1, \dots, N$ in S^h, then (4.2) is equivalent to the system of linear equations

$$\sum_{j=1}^{N} c_j\, a(\varphi_j, \varphi_k) = (f, \varphi_k)\ , \quad j = 1, \dots, N\ , \tag{4.3}$$

for the unknown coefficients of $u_h = \sum_{j=1}^{N} c_j \varphi_j$.

It turns out that all problems under consideration lead to uniquely solvable linear systems of the form (4.3) if h is sufficiently small and that the approximate solutions converge to the corresponding exact solution in the norm of the function space $H_p^1(\Omega)$.

But before using the FEM one can reduce the integration domain Ω by introducing a new artificial boundary $\widetilde{\Gamma}^- = \{x_2 = \widetilde{b}^-\}$ into the first layer, $\ell_1 < \widetilde{b}^- < 0$ (cf. figure 2.1), and new nonlocal boundary operators $\widetilde{T}_\alpha^{TE}$ and $\widetilde{T}_\alpha^{TM}$ which model the layer system below $\widetilde{\Gamma}^-$ together with the radiation condition for $x_2 < b^-$.

Using the Rayleigh expansion in each of the layers, the transmission conditions at the interfaces and the radiation condition one gets by using 2×2 transmission matrices explicit formulae for the numbers γ_n connecting the n-th Fourier coefficient of a solution and its normal derivative on $\widetilde{\Gamma}^-$

$$\partial_n u|_{\widetilde{\Gamma}^-} = - \sum_{n \in \mathbf{Z}} i \gamma_n \hat{u}_n \exp(inKx_1)\ , \tag{4.4}$$

where now

$$\hat{u}_n = \frac{1}{d} \int_0^{2\pi} u(x_1, \widetilde{b}^-) \exp(-inKx_1)\, dx_1\ ,$$

The coefficients γ_n are different for TE and TM polarization, but it can be easily seen that they converge to $\beta_n^1 = (k_1^2 - \alpha_n^2)^{1/2}$, i.e. $|\gamma_n - \beta_n^1| \to 0$ as $|n| \to \infty$. For evaluating these scalars one can use a recursive algorithm which is numerically stable for any number of layers and there is no limit in layer thickness. Matrix algorithms of this type are widely used in other numerical methods for analyzing layered structures (see [7] and the references therein).

Thus if we define nonlocal boundary operators on $\widetilde{\Gamma}^-$

$$\widetilde{T}_\alpha^{TE} u = - \sum_{n \in \mathbf{Z}} i \gamma_n^{TE} \hat{u}_n \exp(inKx_1)\ ,\ \widetilde{T}_\alpha^{TM} u = - \sum_{n \in \mathbf{Z}} i \gamma_n^{TM} \hat{u}_n \exp(inKx_1)\ ,$$

the direct problems (2.2) and (2.7) are equivalent to the variational equations on the smaller rectangle $\widetilde{\Omega} = (0, d) \times (\widetilde{b}^-, b^+)$

$$\begin{aligned} \widetilde{B}_{TE}(u, \varphi) &:= \int_{\widetilde{\Omega}} \nabla_\alpha u \cdot \overline{\nabla_\alpha \varphi} - \int_{\widetilde{\Omega}} k^2\, u\, \bar{\varphi} + \int_{\Gamma^+} (T_\alpha^+ u)\, \bar{\varphi} + \int_{\widetilde{\Gamma}^-} (\widetilde{T}_\alpha^{TE} u)\, \bar{\varphi} \\ &= - \int_{\Gamma^+} 2i\beta \exp(-i\beta b^+)\, \bar{\varphi}\ , \quad \forall \varphi \in H_p^1(\widetilde{\Omega})\ , \end{aligned} \tag{4.5}$$

respectively

$$\widetilde{B}_{TM}(u, \varphi) := \int_{\widetilde{\Omega}} \frac{1}{k^2} \nabla_\alpha u \cdot \overline{\nabla_\alpha \varphi} - \int_{\widetilde{\Omega}} u\, \bar{\varphi} + \int_{\Gamma^+} \frac{1}{(k^+)^2} (T_\alpha^+ u)\, \bar{\varphi} + \int_{\widetilde{\Gamma}^-} \frac{1}{k_1^2} (\widetilde{T}_\alpha^{TM} u)\, \bar{\varphi}$$

$$= -\int_{\Gamma^+} \frac{2i\beta}{(k^+)^2} \exp(-i\beta b^+)\, \bar{\varphi}\,, \quad \forall \varphi \in H^1_p(\widetilde{\Omega})\,. \tag{4.6}$$

The adjoint problems (3.6) and (3.8) are reduced analogously to variational formulations on $\widetilde{\Omega}$, but note that for w_- the right–hand sides changes according to the layer structure.

Due to the simple geometry of binary gratings it is quite natural to choose as finite elements piecewise bilinear functions on a uniform rectangular partition of Ω. The traces of these functions on $\Gamma^\pm$ are piecewise linear functions with uniformly distributed break points. Therefore the computation of the nonlocal terms in the sesquilinear forms

$$\int_{\Gamma^+} (T^+_\alpha \varphi_j)\, \bar{\varphi}_k \, dx_1 \tag{4.7}$$

can be performed very efficiently with an accuracy comparable with the computer precision if the recurrence relations for the Fourier coefficients of spline functions and convergence acceleration methods are used. If the artificial Γ^+ are divided into m subintervals of equal length and the basis of hat functions is used, then (4.7) is a $m \times m$ circulant matrix with the eigenvalues

$$\begin{aligned} \tau_0 &= -i\, d\beta \\ \tau_p &= -2i\, d \left(\frac{\sin(\pi p/m)}{\pi}\right)^4 \sum_{r=-\infty}^{\infty} \frac{\beta^+_{rm+p}}{m(r+p/m)^4}\,, \quad p = 1, \dots, n-1\,. \end{aligned}$$

Thus one has only to expand $\beta^+_{rm+p}/m = \sqrt{(k^+/m)^2 - (\alpha/m + (r+p/m)K)^2}$ with respect to powers of $|r + p/m|$ and to use fast computation of the generalized Zeta function

$$\zeta(x,s) = \sum_{r=0}^{\infty} (r+x)^{-s}\,.$$

Since the scalars γ_n converge very fast to β^1_n, for the computation of the forms

$$\int_{\tilde{\Gamma}^-} (\widetilde{T}^{TE}_\alpha \varphi_j)\, \bar{\varphi}_k \, dx_1\,, \quad \int_{\tilde{\Gamma}^-} (\widetilde{T}^{TM}_\alpha \varphi_j)\, \bar{\varphi}_k \, dx_1\,,$$

one has to compute only few of these coefficients and to apply the summation method mentioned before.

Thus the discretization error of the direct and adjoint problems is mainly determined by the optimal approximation error of the solution with bilinear finite elements. There holds the following convergence results:

1° If the TE problem (2.2) has a unique solution, then for all sufficiently small $h > 0$ the FE discretization of (2.2) and (3.10) are uniquely solvable and the approximate solutions converge to the corresponding exact solution in the norm of $L^2(\Omega)$ with the rate $O(h^2)$.

2° If the TM problem (2.7) has a unique solution, then for all sufficiently small $h > 0$ the FE discretization of (2.7) and (3.11) are uniquely solvable and the approximate solutions converge to the corresponding exact solution with the rate $O(h^{2\lambda})$.

Together with error estimates in the norm of the function space $H_p^1(\Omega)$ it is easy to derive similar estimates for the approximation of the diffraction efficiencies and the gradients of the minimizing functionals.

Hence, in contrast to the optimal convergence rate in the TE case the approximate solutions of the TM problems converge only with the smaller rate $O(h^{2\lambda})$, which is determined by the optical index of the grating material. One way for improving the convergence rate would be the use of special mesh refinements near the corner points of Γ. However, this seems to be not necessary after introducing a generalized FEM (GFEM), which already for rather poor discretizations of the domain $\widetilde{\Omega}$ provided excellent results compared with the usual FEM for both TE and TM modes.

The motivation for generalizing the FEM is due to the well–know fact that the accuracy of FEM for boundary value problems governed by the Helmholtz equation deteriorates with increasing wave number and enlarging domains. Roughly speaking, the evaluation of the sesquilinear form in the interior of $\widetilde{\Omega}$ with the wave number k leads to an approximate solution possessing a different wave number k_h. It turns out that this "phase lag" affects the value of the constants in the error estimates for the FE solutions. Therefore it is necessary to design a FEM with minimal phase lag by modifying the evaluation of the sesquilinear form in the interior of $\widetilde{\Omega}$. For the onedimensional case one can easily define such a method with vanishing phase lag, but in higher dimensions this is impossible. Here we applied and extended the approach of [20] to design a so–called GFEM with minimal pollution ensuring that the wave number of the approximate solution coincides almost with the given k for piecewise uniform rectangular partitions of $\widetilde{\Omega}$ (for details we refer to [15]).

After having solved the linear system corresponding to the FE discretization of the variational equations, the diffraction efficiencies are determined from the Fourier coefficients of the solution on Γ^+ and $\widetilde{\Gamma}^-$. For the computation of the transmission efficiencies and the solutions on the layer interfaces, which appear in the gradient formulae we use a stable recursive algorithms similar to that for evaluating the coefficients γ_n.

The method was used to evaluate the reflection and transmission efficiencies of binary gratings on multilayer systems of different geometries and materials and it turned out to be robust and reliable in both the TE and TM case. Compared with the usual FEM the obtained results were accurate already for rather poor discretizations. In figure 2 we compare the numerical values of some reflection and transmission efficiencies versus the square root n of total number of grid points computed with the usual FEM and the GFEM on quadratic meshes for a simple binary grating with the optical index $\nu = 2.5$ situated on a layer with $\nu = 3.5$. In each case the GFEM results differ already for $n = 40$ only by 2 % from the corresponding values for $n = 200$, whereas the FEM results converge rather slowly to these values.

Furthermore, we compared the results of our method with those obtained with other methods which are known to provide reliable results for binary gratings (e.g. integral equation or modal methods). As an example we give in table 1 the zero order reflection efficiencies of TM polarization for a simple binary grating calculated with different methods. The grating consists of aluminium with the optical index $\nu = 0.47 + 4.8i$ for the given wavelength of 436 nm, the grating period d is equal to 1 μm, the fill factor $f = 0.5$, and the angle of incidence $\theta = 0$.

Table 1 compares the corresponding values of GFEM with an quadratic partitioning

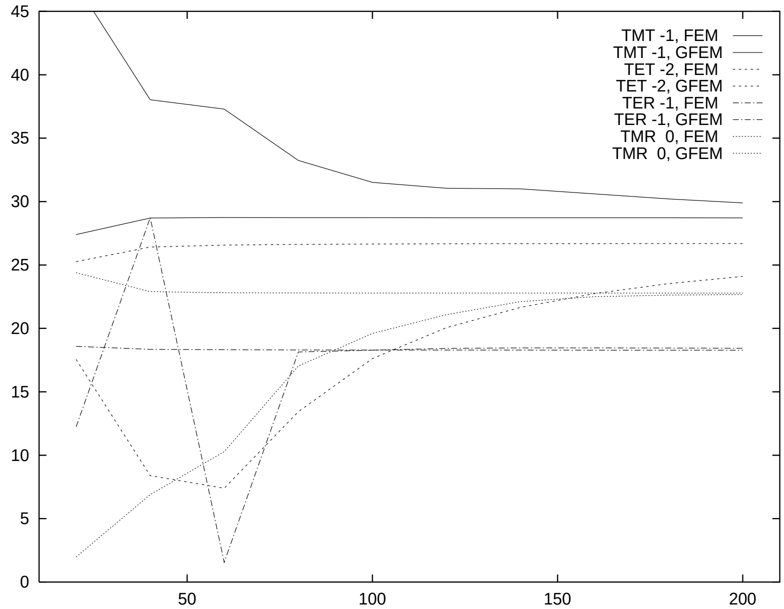

Figure 4.1: Comparison of some efficiencies computed with FEM and GFEM for a simple binary grating with $\nu = 3.5$ versus the square root n of total grid points.

of the rectangular domain with $h_1 = h_2 = 10$ nm for different heights H of the binary structure with the results of three other methods, taken from [21]. These methods are two modal methods, AWG (analytic waveguide method) introduced in [3], [22] and RCWA (rigorous coupled–wave analysis) going back to [2] and essentially improved in recent years (cf. [6]). The third method called IESMP is based on the integral equation method as described in [5] and [21]. Note that GFEM can handle with very general geometries of the diffraction structures and with complex materials, where the implementation of effective solvers for the other methods is very complicated. Moreover, a rigorous convergence analysis for these methods is not known at present.

The GFEM for solving direct and adjoint problems was integrated into a computer program for the study of optimal design problems for binary gratings. By using the standard algorithm of gradient descent local minima of functionals are determined, which characterize desired optical properties. These functionals involve the Rayleigh coefficients of the discrete models on a given partition of the domain Ω for a prescribed range of incidence angles or wavelengths. Of course, the gradients are computed by discretized versions of the formulae given in section 3. Corresponding to the gradients the thicknesses of the layers and the shape of Γ are varied within a class of admissible parameters, which are restricted by certain technological constraints.

Certainly better minimization algorithms exist, for example conjugate gradient methods or methods based on higher order derivative information. The design and analysis of different minimization methods for coated binary gratings will be the topic of future research.

In the following we provide some results of the optimization of a polarisation grating, beam splitters and high reflection mirrors.

The first example concerns the application of metallic subwavelength gratings for polarization devices. Figure 2 shows the results for the optimal design of such a zero order grating that should maximize the reflection of TE polarisation and the transmission of TM polarisation over the range of wavelengths from 450 to 633 nm. The grating period

H/d	AWG	RCWA	IESPM	GFEM
0.1	0.0186	0.0173	0.0190	0.0190
0.2	0.8532	0.8539	0.8529	0.8533
0.3	0.0095	0.0096	0.0100	0.0098
0.4	0.8079	0.8080	0.8095	0.8095
0.5	0.0440	0.0445	0.0465	0.0452
0.6	0.7000	0.7000	0.7068	0.7027
0.7	0.1497	0.1496	0.1511	0.1506
0.8	0.6250	0.6234	0.6277	0.6257
0.9	0.2500	0.2503	0.2503	0.2504
1.0	0.4810	0.4808	0.4840	0.4816

Table 4.1: Comparision of zero order TM efficiency computed with different methods for simple aluminium gratings for normal incidence. Grating parameters are $\lambda = 436$ nm, $d = 1\mu$m, $\nu = 0.47 + 4.8i$ and $f = 0.5$.

is 200 nm, the width of the bar amounts to 60 nm and the height is 150 nm.

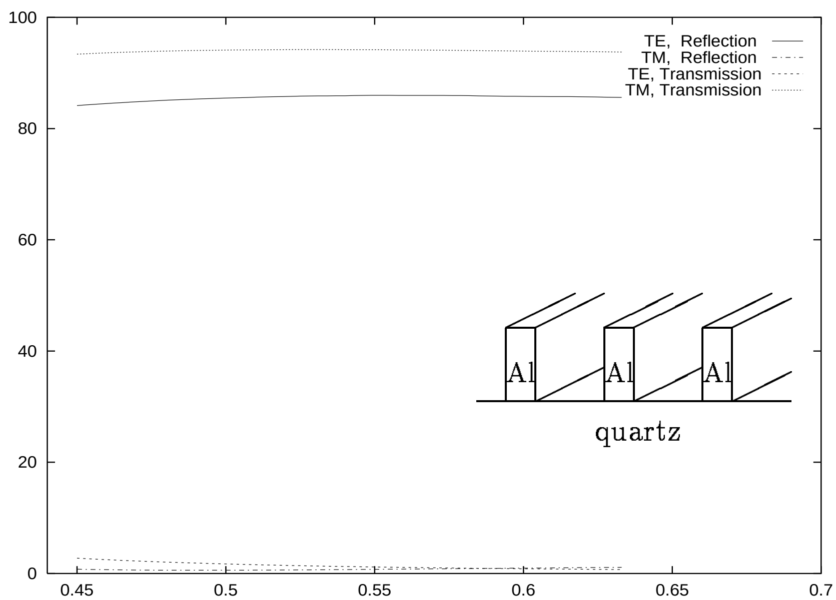

Figure 4.2: Optimal design for a simple polarisation grating for the range of wavelengths from 450 to 633 nm. Grating parameters are $d = 200$ nm, $H = 150$ nm and $f = 0.3$.

Next we provide the optimization results for some beam splitters. The illuminating unpolarized wave with $\lambda = 0.633\mu$m is normally incident from a dielectric medium with refractive index $\nu = 1.5315$. Choosing the period $d = 1.266\mu$m three diffraction orders propagate with angles 0 and $\pm 30^o$. The goal is in

a) to maximize the efficiencies of the orders ± 1

b) to obtain maximal and equal efficiencies of all three orders

by optimizing the height H and the fill factor f of the grating with one groove per period. The results are depicted in figures x, y , the obtained values are

a) $H = 0.734\mu$m, $f = 0.72$, b) $H = 0.43\mu$m, $f = 0.58$.

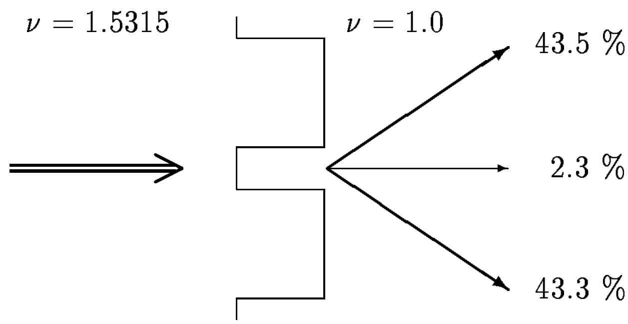

Figure 4.3(a): Optimal design of a 1–to–2 beam splitter. Grating parameters are $\lambda = 0.633\mu$m, $d = 1.266\mu$m, $H = 0.734\mu$m and $f = 0.72$.

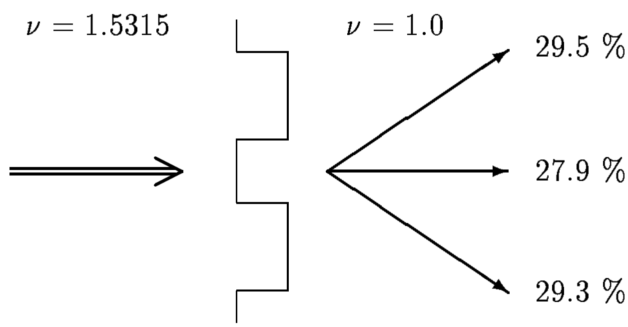

Figure 4.3(b): Optimal design of a 1–to–3 beam splitter. Grating parameters are $\lambda = 0.633\mu$m, $d = 1.266\mu$m, $H = 0.43\mu$m and $f = 0.58$.

For the same parameters as before we seek a one–to–four beam splitter with the diffraction angles $\pm 14.5^o$ and $\pm 30^o$. Choosing the period $d = 2.532\mu$m 9 diffraction orders propagate, the goal is to maximize the efficiencies of the orders ± 1 and ± 2. To obtain a satisfactory solution it is necessary to use two grooves per period. For the optimal solution the height of these grooves is $H = 1.747\mu m$, the scaled transition points are $0., 0.24, 0.38, 0.63$. For the same parameters as before we optimized a one–to–five beam

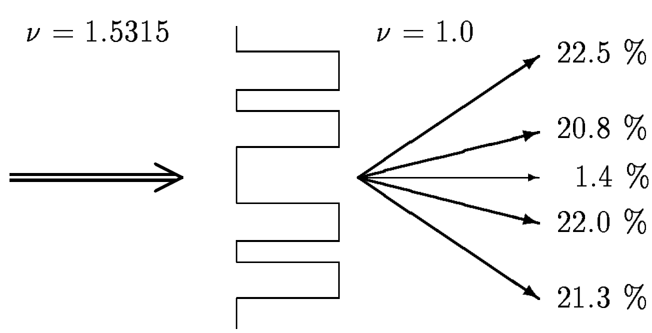

Figure 4.4: Optimal design of a 1–to–4 beam splitter. Grating parameters are $\lambda = 0.633\mu$m, $d = 2.532\mu$m and $H = 1.747\mu$m. The distribution of the transition points is $0., 0.24, 0.38, 0.63$.

splitter with the diffraction angles 0^o, $\pm 20.7^o$ and $\pm 45^o$. The period of the grating is $d = 1.79\mu$m, the height of the optimal grooves is $H = 0.77\mu$m, the scaled transition points are $0, 0.12, 0.36, 0.76$. The next problem concerns the design of a zero–order copper grating ($\nu = 12.7 + 51.1i$) as circular polarizer for CO_2 laser with $\lambda = 10,6\mu$m

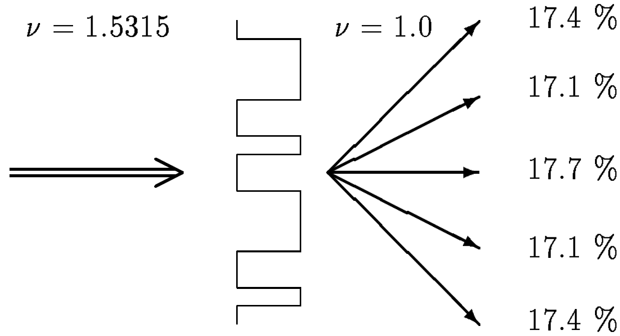

Figure 4.5: Optimal design of a 1–to–5 beam splitter. Grating parameters are $\lambda = 0.633\mu$m, $d = 1.79\mu$m and $H = 0.77\mu$m. The distribution of the transition points is $0., 0.12, 0.36, 0.76$.

θ	TE	TM	phase
29.0	97.50	95.72	90.72
29.2	97.50	95.72	90.58
29.4	97.51	95.72	90.45
29.6	97.51	95.72	90.32
29.8	97.52	95.72	90.18
30.0	97.52	95.72	90.04
30.2	97.53	95.72	89.91
30.4	97.53	95.72	89.77
30.6	97.54	95.72	89.63
30.8	97.54	95.72	89.49
31.0	97.55	95.72	89.35

Table 4.2: Zero order efficiencies and phase difference for circular polarizer. Grating parameters are $\lambda = 10,6\mu$m, $\nu = 12.7 + 51.1i$, $d = 3\mu$m, $H = 1.65\mu$m and f$=$ .24.

such that in the range of incident angles $\theta \in (29^o, 31^o)$ the efficiencies of the reflected TE and TM polarized wave are maximal and the phase difference between them is close to $\pi/2$. Here one has to minimize the functional (3.3) extended over the range of incident angles, which possesses many local minima. One of the reasonable geometries is $d = 3.0\mu$m, $H = 1.65\mu$m and $f = .24$. Table 2 contains the computed values. Finally we consider a high reflection grating on top of a quarter–wave system of 15 layers for the wavelength $\lambda = 1.45\mu$m. The even homogeneous–layer parameters are $\nu = 1.45$ and $h_j = 248$ nm with the odd homogeneous–layer parameters being $\nu = 2.3$ and $h_j = 157$ nm. The substrate is quartz with $\nu = 1.45$. Without any grating structure the reflection efficiency is almost 100 % (99,76 % in normal incidence). The problem is to find a grating surface in an additional quartz layer on the top in order to maximize the TE reflection of order -1 in Littrow mounting for $\theta = 20.4^o$. Correspondingly, the period of the grating is $d = 2.06\mu$m. Optimal values were obtained for the thickness of the additional quartz layer of 866 nm, the binary grating within this layer has the height $H = 804$ nm and the fill factor $f = 0.56$. In that case the efficiency of order -1 amounts $99,42\%$.

5 Conclusion

In this paper we focused on optimal design problems for binary gratings, using exact formulae for the gradients of the cost functionals and a fast and reliable method for the numerical solution of direct and adjoint diffraction problems. The latter method is based on a variational formulation and combines a finite element method in the grating structure with Rayleigh series expansions in the layer system below the grating. This approach is not restricted to binary profiles, but allows the numerical treatment of rather general diffraction structure, together with a rigorous convergence analysis.

We proposed a generalized finite element method (GFEM) with minimal pollution, which provides highly accurate numerical results in the computation of diffraction efficiencies for both the TE and TM mode. In particular, for TM diffraction problems having a mild singularity of the solution, the convergence performance of our method was comparable with that of the rigorous coupled–wave analysis of [6] and the integral equation method of [5]. Moreover, accurate numerical results can be obtained even in the presence of strong singularities of the solution. We expect that the approach can be also extended to the more general case of conical diffraction and biperiodic gratings.

To solve optimal design problems for binary gratings by gradient descent we presented explicit formulae for the gradients with respect to the parameters of the grating profile and the thicknesses of layers. These formulae involve the solutions of direct and adjoint TE and TM problems and reduce considerably the computational costs compared to simple difference approximations of the gradients. The GFEM and the gradient formulae were integrated into a computer program to find the optimal design of binary gratings with desired phase or intensity pattern for a given range of incidence angles or wavelength. Several numerical examples including polarisation gratings and beam splitters successfully demonstrate the efficiency of the algorithm.

Acknowledgments

We are grateful to Bernd Kleemann (Berlin Institute of Optics) and Andreas Pomp (University of Stuttgart) for fruitful discussions concerning direct solvers for diffraction problems and for providing us several interesting optimal design problems.

References

[1] PETIT, R. (ed.), 1980, *Electromagnetic theory of gratings.* Topics in Current Physics, Vol. **22**. (Berlin: Springer-Verlag).

[2] MOHARAM, M. G. and GAYLORD T. K., 1981, Rigorous coupled-wave analysis of planar-grating diffraction. *J. Opt.Soc. Am.* **71**, no. 7, 811–818.

[3] BOTTEN, L. C., CRAIG, M. S., McPHEDRAN, R. C., ADAMS, J. L. and ANDREWARTHA, J.R., 1981, The dielectric lamellar diffraction grating. *Optica Acta* **28**, no. 3, 413–428.

[4] MONTIEL, F. and NEVIÈRE, M., 1994, Differential theory of gratings: extension to deep gratings of arbitrary profile and permittivity through the R–matrix propagation algorithm. *J. Opt.Soc. Am.* A, **11**, 1321–1331.

[5] POMP, A., 1991, The integral method for coated gratings: computational cost. *J. Mod. Optics* **38**, 109–120.

[6] LALANNE, P. and MORRIS, G. M., 1996, Highly improved convergence of the coupled-wave method for TM polarization. *J. Opt.Soc. Am.* A, **13**, no. 4, 779–784.

[7] LI, L., 1996, Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings. J. Opt. Soc. Am. A, **13**, No. 5, 1024–1035.

[8] BRUNO, O. P. and REITICH, F., 1993, Numerical solution of diffraction problems: a method of variation of boundaries. *J. Opt. Soc. Amer.* **A 10**, 1168–1175.

[9] BONNET-BENDHIA, A.-S. and STARLING, F., 1994, Guided waves by electromagnetic gratings and non–uniqueness examples for the diffraction problem. *Math. Meth. Appl. Sci.* **17**, 305–338.

[10] BAO, G., DOBSON, D.C. and COX, J.A., 1995, Mathematical studies in rigorous grating theory, *J. Opt. Soc. Amer.* **A 12**, 1029–1042.

[11] DOBSON, D. C., 1994, A variational method for electromagnetic diffraction in biperiodic structures. *Model. Math. Anal. Numer.* **28**, 419–439.

[12] DOBSON, D.C. and COX, J.A., 1994, Mathematical modeling for diffractive optics. *Proceedings SPIE Diffractive and Miniaturized Optics*, **CR49**, 32–53.

[13] BAO, G., 1995, Finite element approximation of time harmonic waves in periodic structures. *SIAM J. Numer. Anal.* **32**, 1155–1169.

[14] BAO, G., 1996, Numerical analysis of diffraction by periodic structures: TM polarization. *Num. Math.* **75**, 1–16.

[15] ELSCHNER, J. and SCHMIDT, G., 1997, Analysis and numerics for the optimal design of binary diffractive gratings, Preprint No. 323, WIAS Berlin.

[16] PARK, I.H., OH, H.K. and HYUN, S.B., 1997, Photolithography simulation on nonplanar substrate using the finite-element method with absorbing boundary conditions. *Proceedings SPIE Optical Microlithography* X, **3051**, 578–587.

[17] DOBSON, D. C., 1993, Optimal design of periodic antireflective structures for the Helmholtz equation. *European J. Appl. Math.* **4**, 321–340.

[18] NOPONEN, E., TURUNEN, J. and VASARA, A., 1992, Parametric optimization of multilevel diffractive optical elements by electromagnetic theory. Applied Optics **31**, 5910–5912.

[19] KIKUTA, H., OHIRA, Y., IWATA, K, 1996, Subwavelength gratings optimized for broadband quarter-wave plates. *Proceedings SPIE Polarization Analysis and Applications to Device Technology*, **2873**, 218–221.

[20] BABUŠKA, I., IHLENBURG, F., PAIK, E. and SAUTER, S., 1995, A generalized finite element method for solving the Helmholtz equation in two dimensions with minimal pollution, *Comp. Meth. Appl. Mech. Eng.* **128**, 325–359.

[21] DAVIDSON, M., KLEEMANN, B. and BISCHOFF J., 1997, A comparison between rigorous light scattering methods. *Proceedings SPIE Optical Microlithography* X, **3051**, 606–619.

[22] DAVIDSON, M. , MONAHAN, K. and MONTEVERDE, R., 1991, Linearity of coherence probe metrology: simulation and experiment. *Proceedings SPIE Integrated Circuit Metrology, Inspection, and Process Control* V, **1464**, 155–176.