Grating Profile Reconstruction Based on Finite Elements and Optimization Techniques

G.C. Hsiao
Department of Mathematical Sciences, University of Delaware
Newark, DE 19716, U.S.A.
hsiao@math.udel.edu

J. Elschner and A. Rathsfeld
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstr. 39
D-10117 Berlin, Germany
elschner@wias-berlin.de, rathsfeld@wias-berlin.de

February 4, 2003

1991 Mathematics Subject Classification. 35R30, 35J05, 78A46, 78M50.
Keywords. diffraction grating, profile reconstruction, optimization method, conjugate gradient algorithm

Abstract

We consider the inverse diffraction problem to recover a two-dimensional periodic structure from scattered waves measured above and beneath the structure. The task is reformulated in form of an optimization problem including special regularization terms. The solvability and the dependence on the parameter of regularization is analyzed. Numerical results for synthetic data demonstrate the practicability of the inversion algorithm.

1 Introduction

The scattering theory in periodic structures has many applications in micro-optics, where periodic structures are often called diffraction gratings (cf. [18] for an introduction to the direct problem). The treatment of the inverse problem, recovering the periodic structure or the shape of the grating profile from the scattered field, is useful e.g. in quality control and design of diffractive elements with prescribed far field patterns (see [4], [19]).

Various methods for the computation of the grating profile curve of perfectly conducting gratings have been proposed by Ito, Reitich, Arens, Kirsch, Hettlich, Bruckner, Elschner, and Yamamoto ([12, 2, 15, 6, 7]). We follow the technique of [6] (cf. [9, Section 5.4] for the original algorithms applied to scattering obstacles). However, we consider reflection by and transition through gratings described by general material dependent wave number functions, and replace the boundary integral approach of [6] by a finite element algorithm (cf. [1] for a similar finite element optimization of a different functional over a set of transition curves).

To be more precise, we start with a short introduction to the direct problem of diffraction by gratings in Section 2. The TE component of the electric field of the time-harmonic light wave is the solution of a two-dimensional Helmholtz equation over the cross section of the grating device. We recall the variational formulation corresponding to the coupling of differential and boundary integral representations and define the Rayleigh coefficients of the Helmholtz solution. These correspond to the portion of light and the phase shift of the reflected and transmitted modes. In Section 3 we introduce the inverse problem. From measured Rayleigh coefficients for several incidence directions, we wish to reconstruct the grating, i.e. the wave number distribution over the grating cross section. The solution is obtained as the minimizer of an optimization problem, where the objective function consists of three terms. The first is the residual of the Helmholtz equation, the second the deviation of the computed Rayleigh coefficients from the measured data, and the third is a regularization term to cope with the ill-posedness of the inverse problem. We show the existence of minimizers and prove the convergence of these minimizers to the true solution if the regularization parameter tends to zero. Section 4 is devoted to the finite element discretization of the Helmholtz equation and Section 5 to the discretization of the optimization problem. For the solution of the finite dimensional optimization problem we propose the conjugate gradient algorithm of Fletcher, Reeves, Polak, and Ribière. In the last section we present numerical experiments.

Finally, we remark that the proposed treatment of the inverse problem is a first theoretical approach. Due to the severe ill-posedness of the problem the accuracy of the reconstruction cannot be satisfactory for realistic applications. For better approximations the class of admissible gratings must be restricted in accordance with the technical requirements

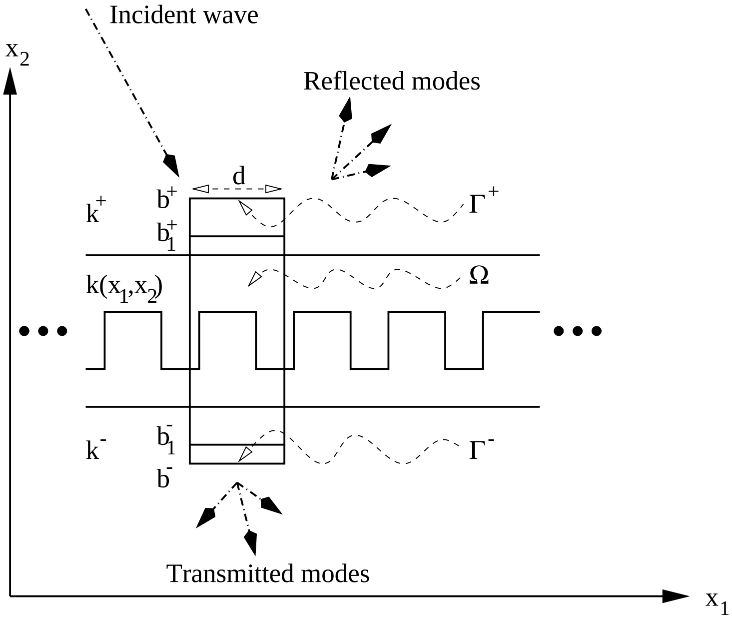

Figure 1: Cross section of grating.

2 The Direct Problem

Consider an ideal optical grating (cf. 1). This is an ideal infinite plate in three–dimensional space covering a half space filled by a substrate material. The plate consists of different materials. Moreover, the materials are disposed in such a manner that the material does not change in one of the two directions parallel to the plane of the plate. With respect to the other direction parallel to the plane the material distribution is supposed to be periodic with period d. The materials are non-magnetic with the permeability μ_0 and have the dielectric constants ε. The coordinate system is chosen such that the x_2 axis is perpendicular to the plane of the grating, such that the material distribution together with the resulting diffraction solution is invariant in the x_3 direction, and such that the x_1 axis is parallel to the plane of the grating. Thus the materials of the problem are determined by the function $\varepsilon(x_1, x_2)$ which is d–periodic in x_1. Due to the invariance with respect to x_3 it is sufficient to consider the electromagnetic fields restricted to the plane spanned by the x_1 and x_2 axes. More precisely, we introduce two artificial boundaries $\Gamma^\pm := \{(x_1, x_2) : \ x_2 = b^\pm\}$ forming the upper and lower bounds of the cross section of the grating structure, respectively, and denote by Ω the rectangle $(0, d) \times (b^-, b^+)$ which covers one period of the cross section. We assume that the material above Γ^+ and below Γ^- is homogeneous with $\varepsilon = \varepsilon^+ > 0$ and $\varepsilon = \varepsilon^-$, respectively. Between Γ^+ and Γ^- the material may be inhomogeneous and we assume that the function ε is piecewise continuous. Further, we introduce the wave number function $k = k(x_1, x_2) := \omega\sqrt{\mu_0\varepsilon}$ and $k^\pm := \omega\sqrt{\mu_0\varepsilon^\pm}$ with ω the angular frequency of the incident light wave. Thus the wave can be described by a time independent factor times $\exp(-\mathbf{i}\omega t)$. We suppose that

$$k^+ > 0\,, \quad \Re\mathrm{e}\, k^- > 0\,, \quad \Im\mathrm{m}\, k^- \geq 0\,, \quad \Re\mathrm{e}\, k(x_1, x_2) > 0\,, \quad \Im\mathrm{m}\, k(x_1, x_2) \geq 0\,. \qquad (2.1)$$

Moreover, we suppose that there exists $b_1^\pm$ with $b^- < b_1^- < b_1^+ < b^+$ such that $k|_{\Omega^\pm} \equiv k^\pm$ for $\Omega^- := (0, d) \times (b^-, b_1^-)$ and $\Omega^+ := (0, d) \times (b_1^+, b^+)$.

Assume that an incoming plane wave is incident in the (x_1, x_2)–plane upon the grating from the top with the angle of incidence $\theta \in (-\pi/2, \pi/2)$. Then the electromagnetic field does not depend on x_3. For simplicity, we restrict to the case of TE polarization, i.e. the electric field $\mathbf{E}$ is supposed to remain parallel to the x_3 axis (to the grooves) and is therefore determined by a single scalar quantity $v = v(x_1, x_2)$ (the transverse component of $\mathbf{E}$). The function v satisfies the two–dimensional Helmholtz equation

$$\Delta v + k^2 v = 0 \tag{2.2}$$

in the regions with constant permittivity. In the infinite regions the usual outgoing wave conditions are required. At the material interfaces the solutions are subjected to the transmission conditions, i.e. the solution v and its normal derivative $\partial_n v$ have to cross the interface continuously.

The diffraction problems admit variational formulations in a bounded periodic cell which were introduced in [20, 5, 4, 13]. The incoming wave has the form $v^i(x_1, x_2) = \exp(\mathbf{i}\alpha x_1 - \mathbf{i}\beta x_2)$, where $\alpha = k^+ \sin\theta$, $\beta = k^+ \cos\theta$. If we define the d-periodic function $u(x_1, x_2) := v(x_1, x_2)\exp(-\mathbf{i}\alpha x_1)$, then the diffraction problem for TE polarization can be transformed to a variational problem for u in the rectangle Ω. Multiplying the differential equation (2.2) by some smooth function, applying Green's formula and taking into account the transmission conditions at the material interfaces and the outgoing wave condition on $\Gamma^\pm$ it can be shown (cf. [20, 4, 10]) that the diffraction problem for TE polarization is equivalent to the variational equation

$$\begin{aligned} \mathcal{B}_{TE}(u, \varphi) &:= \int_\Omega \nabla_\alpha u \cdot \overline{\nabla_\alpha \varphi} - \int_\Omega k^2\, u\, \bar\varphi + \int_{\Gamma^+} (T_\alpha^+ u)\, \bar\varphi + \int_{\Gamma^-} (T_\alpha^- u)\, \bar\varphi \\ &= -\int_{\Gamma^+} 2\mathbf{i}\beta \exp(-\mathbf{i}\beta b^+)\, \bar\varphi\,, \quad \forall \varphi\,, \end{aligned} \tag{2.3}$$

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

$$(T_\alpha^\pm u)(x_1, b^\pm) := -\sum_{n=-\infty}^{\infty} \mathbf{i}\beta_n^\pm \hat u_n^\pm \exp(\mathbf{i}nKx_1)\,, \tag{2.4}$$

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

$$\hat u_n^\pm = \frac{1}{d}\int_0^d u(x_1, b^\pm)\exp(-\mathbf{i}nKx_1)\, dx_1\,.$$

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

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

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

The variational equation (2.3) should be satisfied for all test functions $\varphi \in H^1_{per}(\Omega)$, that is the function space of all complex–valued functions φ which are d-periodic in x_1 and which together with their first–order partial derivatives are square integrable in Ω (cf. [8] for the variational approach to classical elliptic boundary value problems).

The variational formulation (2.3) is very useful, because the transmission and outgoing wave conditions are enforced implicitly and it allows to seek the solution in the function space $H^1_{per}(\Omega)$, 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.

Theorem 2.1 ([10]): *Suppose that k satisfies condition* (2.1). *Then the sesquilinear form $\mathcal{B}_{TE}$ is strongly elliptic over $H^1_{per}(\Omega)$.*

We recall that a bounded sesquilinear form $\mathcal{B}_{TE}(\cdot,\cdot)$ given on the Hilbert space $H^1_{per}(\Omega)$ is called strongly elliptic if there exist a complex number ϕ, $|\phi| = 1$, a constant $c > 0$, and a compact form $\mathcal{Q}(\cdot,\cdot)$ such that

$$\Re\mathrm{e}\, \mathcal{B}_{TE}(\phi\, u, u) \geq c\|u\|^2_X - \mathcal{Q}(u,u)\ , \quad \forall\, u \in H^1_{per}(\Omega)\ .$$

As usual, the sesquilinear form $\mathcal{B}_{TE}$ corresponds to a bounded linear operator B mapping $H^1_{per}(\Omega)$ into its dual $H^1_{per}(\Omega)'$ via $\mathcal{B}_{TE}(u,v) = \langle Bu, v\rangle,\ u,v \in H^1_{per}(\Omega)$. According to the proof of the last theorem (cf. [10]), the bilinear form $\mathcal{B}_{TE}$ splits into the compact form $\mathcal{C}_k(u,v) := -\int_\Omega uv$, into a strongly elliptic form $\mathcal{P}$ with $\mathcal{P}(u,u) \geq c\|u\|^2_{H^1_{per}(\Omega)}$ and constant $c > 0$, and a finite dimensional form $\mathcal{T}$. Correspondingly, we get $B(k,\theta) = P + T + C_k$ with $\langle Pu, u\rangle \geq c\|u\|^2_{H^1_{per}(\Omega)}$, with finite range operator T, and with compact C_k. From this splitting we infer that $B(k,\theta)$ is a Fredholm operator of index zero. Thus the strong ellipticity is the basis to prove the invertibility of operator B under additional conditions.

We write $B = B(k,\theta)$ to indicate the dependence of B on the wave number function k and on the incidence angle θ. The variational equation (2.3) is equivalent to the operator equation $B(k,\theta)u = w$ with $w \in H^1_{per}(\Omega)'$ such that $\langle w, u\rangle$ with $u \in H^1_{per}(\Omega)$ is defined by the right-hand side of (2.3). The operator $B(k,\theta)$ is a second order differential operator. To get an equation with a well conditioned operator acting in the single space $H^1_{per}(\Omega)$ we multiply the equation by the inverse of $\widetilde{B}(\theta) := B(\tilde{k},\theta)$ with a fixed simple wave number function $\tilde{k}$. Thus (2.3) is equivalent to $[\widetilde{B}(\theta)^{-1}B(k,\theta)]u = \widetilde{B}(\theta)^{-1}w$.

The invertibility of the operators $\widetilde{B}(\theta)$ and $B(\tilde{k},\theta)$ will be supposed in the following. Partial results on this are reported e.g. in [20, 4, 10]. Here we only give a stability result with respect to the wave number function.

Theorem 2.2 *Suppose that the squared wave number functions k_n^2 form a weakly convergent sequence in the space $L^2(\Omega)$. (Note that the squared wave number function enters linearly into the scene.) If $B(k_0,\theta)$ is the operator defined with k_0 such that k_0^2 is the weak limit of the k_n^2 and if this $B(k_0,\theta)$ is invertible, then there exist an integer $n_0 > 0$ and a real $c > 0$ such that $\|B(k_n,\theta)u\|_{H^1_{per}(\Omega)'} \geq c\|u\|_{H^1_{per}(\Omega)}$ for any $u \in H^1_{per}(\Omega)$ and $n \geq n_0$. Since the $B(k_n,\theta)$ are Fredholm operators with index zero, the last estimate implies the invertibility of $B(k_n,\theta)$ if $n \geq n_0$.*

Proof: If the theorem were not true, then there is a sequence $\{u_n\} \subset H^1_{per}(\Omega)$ such that $\|u_n\|_{H^1_{per}(\Omega)} = 1$ and $\|B(k_n,\theta)u_n\|_{H^1_{per}(\Omega)'} \to 0$. Then, without loss of generality, we may suppose that u_n tends weakly to u_0 in $H^1_{per}(\Omega)$. Hence, $\|u_n - u_0\|_{L^2(\Omega)} \to 0$. From the weak convergence of u_n we infer the weak convergence in $H^1_{per}(\Omega)'$ of $[P+T]u_n \rightharpoonup [P+T]u_0$. Indeed $Tu_n \to Tu_0$ and $\langle Pu_n, \varphi\rangle = \langle u_n, P'\varphi\rangle \to \langle u_0, P'\varphi\rangle = \langle Pu_0, \varphi\rangle$ for all $\varphi \in H^1_{per}(\Omega)$.

Furthermore, since $\|u_n - u_0\|_{L^p(\Omega)} \to 0$ for any $p < \infty$, we obtain $\|u_n\overline{\varphi} - u_0\overline{\varphi}\|_{L^2(\Omega)} \to 0$ for any $\varphi \in H^1_{per}(\Omega)$. Hence, $\mathcal{C}_k(u_n,\varphi) = -\int_\Omega k_n^2 u_n\overline{\varphi} \to -\int_\Omega k_0^2 u_0\overline{\varphi}$. Together with the weak convergence $[P+T]u_n \rightharpoonup [P+T]u_0$ we have $B(k_n,\theta)u_n \rightharpoonup B(k_0,\theta)u_0$. This implies $B(k_0,\theta)u_0 = 0$ and $u_0 = 0$. Consequently, $C_{k_n}u_n \to 0$ and $Tu_n \to 0$ together with $\|B(k_n,\theta)u_n\|_{H^1_{per}(\Omega)'} \to 0$ yield $Pu_n \to 0$ which contradicts $\langle Pu_n, u_n\rangle \geq c\|u_n\|_{H^1_{per}(\Omega)} = 1$. ∎

Note that any solution of (2.3) satisfies on $\Gamma^\pm$ the boundary conditions

$$\partial_n u|_{\Gamma^+} + T_\alpha^+ u|_{\Gamma^+} = -2\mathbf{i}\beta \exp(-\mathbf{i}\beta b^+)\ , \quad \partial_n u|_{\Gamma^-} + T_\alpha^- u|_{\Gamma^-} = 0\ . \tag{2.5}$$

which implies the Fourier series expansion

$$\begin{aligned} u(x_1, b^+) &= \sum_{n=-\infty}^{\infty} A_n^+ \exp(\mathbf{i}\beta_n^+ b^+) \exp(\mathbf{i}nKx_1) + \exp(-\mathbf{i}\beta b^+)\ , \\ u(x_1, b^-) &= \sum_{n=-\infty}^{\infty} A_n^- \exp(-\mathbf{i}\beta_n^- b^-) \exp(\mathbf{i}nKx_1)\ , \end{aligned} \tag{2.6}$$

for suitable coefficients $A_n^\pm$. It is not hard to see that the operators $T_\alpha^\pm$ are the Dirichlet–to–Neumann mappings

$$\partial_n u^\pm|_{\Gamma^\pm} = -T_\alpha^\pm u^\pm|_{\Gamma^\pm} \tag{2.7}$$

for functions solving the Helmholtz equation with outgoing wave condition for $x_2 \gtrless b^\pm$, i.e. for functions of the form

$$\sum_{n=-\infty}^{\infty} A_n^\pm \exp(\pm\mathbf{i}\beta_n^\pm x_2) \exp(\mathbf{i}nKx_1)\ , \quad x_2 \gtrless b^\pm\ . \tag{2.8}$$

The coefficients $A_n^\pm$ in the expansion (2.8) are called Rayleigh coefficients. The most interesting are those with $n \in \mathcal{U}^\pm$,

$$\mathcal{U}^\pm \ := \ \begin{cases} \emptyset & \text{if } \Im\mathrm{m}\, k^\pm > 0 \\ \left\{ n \in \mathbb{Z} : |n+\alpha| < k^\pm \right\} & \text{if } \Im\mathrm{m}\, k^\pm = 0 \end{cases} .$$

Indeed, these coefficients $A_n^\pm$ describe the magnitude and the phase shift of those terms $A_n^\pm \exp(\mathbf{i}nKx_1)\exp(\pm\mathbf{i}\beta_n^\pm x_2)$ in the representation of $u(x_1,x_2)$ for $x_2 \gtrless b^\pm$, which correspond to propagating plane waves. The terms with $n \notin \mathcal{U}^\pm$ lead to evanescent waves, only. Hence, the $A_n^\pm$ with $n \in \mathcal{U}^\pm$ can be considered to be the far field data of the diffraction problems at optical gratings. The optical efficiencies of the grating are defined by

$$e_n^\pm \ := \ (\beta_n^\pm/\beta)|A_n^\pm|^2,\ (n,\pm) \in \mathcal{U}^* := \left\{(n,+):\ n \in \mathcal{U}^+\right\} \cup \left\{(n,-):\ n \in \mathcal{U}^-\right\}, \tag{2.9}$$

which is the ratio of the energy of the nth propagating mode to the energy of the incident wave.

Restricting the solution $\Omega \ni (x_1,x_2) \mapsto u(k,\theta)(x_1,x_2)$ to $\Gamma^\pm$, we get the Rayleigh coefficients $A_n^\pm$ by computing the Fourier coefficients according to (2.6). The linear operator of restricting u to $\Gamma^\pm$ and of computing the Rayleigh coefficients $A_n^\pm$ will be denoted by $F(\theta)$, i.e.,

$$\begin{aligned} A \ &:= \ \left(A_n^\pm\right)_{(n,\pm)\in\mathcal{U}^*} \ = \ F(\theta)\, u + A^i, \\ A^i \ &:= \ \Big(-\exp(-\mathbf{i}\beta b^+)\delta_{(n,\pm),(0,+)} \Big)_{(n,\pm)\in\mathcal{U}^*}. \end{aligned}$$

If the refractive indices of the cover material above the grating and the substrate material beneath the grating are fixed, then the operator $F(\theta)$ is independent of the function k inside the grating.

3 The Inverse Problem

For the inverse problem, we suppose that the distribution of the material in the grating between the lines $\{(x_1, b_1^+) : 0 < x < d\}$ and $\{(x_1, b_1^-) : 0 < x < d\}$ is unknown. In other words, our task is to determine the unknown function $k(x_1, x_2)$ for $b_1^- < x_2 < b_1^+$. To get this, we illuminate the grating by plane waves $v^i(x_1, x_2) = \exp(\mathbf{i}k^+ \sin\theta x_1 - ik^+ \cos\theta x_2)$ under the incident angles $\theta = \theta_l,\ l = 1, \ldots, L$ and measure the Rayleigh coefficients $A^{\pm}_{meas,n}(\theta_l)$ for $(n, \pm) \in \mathcal{U}^* = \mathcal{U}^*(\theta_l)$ and for each angle $\theta_l,\ l = 1, \ldots, L$. We seek a material distribution and the corresponding wave number function $k(x_1, x_2)$ such that the Rayleigh coefficients $A^{\pm}_n = A^{\pm}_n(k, \theta_l)$, obtained by solving the variational equation (2.3) with respect to $u(x_1, x_2) = u(k, \theta_l)(x_1, x_2)$ and by computing the Fourier coefficients $A^{\pm}_n(k, \theta_l)$ of $u|_{\Gamma^\pm}$ according to (2.8), coincide with the measured data $A^{\pm}_{meas,n}(\theta_l)$ for $(n, \pm) \in \mathcal{U}^*(\theta_l)$ and $l = 1, \ldots, L$. In other words, we seek an unknown squared wave number function k^2 and the corresponding solutions $u(k, \theta_l)$ from $[\widetilde{B}(\theta_1)^{-1}B(k, \theta_l)]u(k, \theta_l) = \widetilde{B}(\theta_1)^{-1}w(\theta_l)$ such that the computed Rayleigh coefficients $A(\theta_l) = F(\theta_l)u(k, \theta_l) + A^i$ coincide with the measured $A_{meas}(\theta_l) := (A^{\pm}_{meas,n}(\theta_l))_{(n,\pm)\in\mathcal{U}^*(\theta_l)}$. Expressing our objective in formulae, we seek k^2 and $u_l = u(k, \theta_l)$ such that

$$
\begin{aligned}
\sum_{l=1}^{L} \left\| \left[\widetilde{B}(\theta_1)^{-1} B(k, \theta_l) \right] u_l - \widetilde{B}(\theta_1)^{-1} w_l \right\|^2_{L^2(\Omega)} &= 0, \\
\sum_{l=1}^{L} \left\| \left[F(\theta_l)\, u_l + A^i \right] - A_{meas}(\theta_l) \right\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))} &= 0.
\end{aligned}
$$

Here $w_l = w(\theta_l)$ stands for the right-hand side functional in (2.3) with θ replaced by θ_l. The symbol $\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))$ denotes the complex Euclidean space of vectors over the index set $\mathcal{U}^*(\theta_l)$.

The operator $F(\theta_l)$ is smoothing and the equation $F(\theta_l)u_l = A(\theta_l) - A^i$ is severely ill-posed. To cope with measurement errors in the values of $A(\theta_l)$ we need a regularization, i.e. we try to find solutions k^2 and u_l such that the left-hand sides of the last two equations are small and that, simultaneously, the solution is relatively smooth. Relatively smooth means that the H^1_{per} Sobolev norms of u_l and the $H^{1/2}_{per}$ Sobolev norm of k^2 do not blow up. This will be helpful also if the solution should not be unique. Finally, we define the non-linear objective functional

$$
\begin{aligned}
\mathcal{F}\Big(k^2, u_1, \ldots, u_L; \gamma\Big) \quad :=& \quad \frac{\displaystyle\sum_{l=1}^{L} \left\| \widetilde{B}(\theta_1)^{-1} B(k, \theta_l) u_l - \widetilde{B}(\theta_1)^{-1} w_l \right\|^2_{L^2(\Omega)}}{\displaystyle\sum_{l=1}^{L} \left\| \widetilde{B}(\theta_1)^{-1} w_l \right\|^2_{L^2(\Omega)}} \\
&+ c_d \frac{\displaystyle\sum_{l=1}^{L} \left\| \left[F(\theta_l)\, u_l + A^i \right] - A_{meas}(\theta_l) \right\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}}{\displaystyle\sum_{l=1}^{L} \left\| A_{meas}(\theta_l) \right\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}} \\
&+ c_v \gamma \left\| k^2 \right\|^2_{H^{1/2}_{per}(\Omega)} + c_s \gamma \sum_{l=1}^{L} \left\| u_l \right\|^2_{H^1_{per}(\Omega)} . \qquad (3.1)
\end{aligned}
$$

Here c_d, c_v, and c_s denote appropriate equilibration constants which are to be determined by numerical experiments. The number γ is a small positive regularization parameter which is to be chosen in dependence on the measurement error. Using the functional F, we finally arrive at the optimization problem

$$\begin{aligned} \mathcal{F}\Big(k^2, u_1, \dots, u_L; \gamma\Big) \quad &\longrightarrow \quad \min \ . \\ k^2 \in H^{1/2}_{per}(\Omega), & \\ u_l \in H^1_{per}(\Omega), \ l = 1, \dots, L & \end{aligned} \tag{3.2}$$

This optimization problem will be discretized and solved numerically in the next sections. For its solvability and its connection to the exact inverse problem, we get the following two theorems.

Theorem 3.1 *For any fixed positive regularization parameter γ, there exists a minimizer $\{k_0^2,\ u_{l,0},\ l = 1, \dots, L\}$ of the optimization problem (3.2).*

Proof: Suppose $\{k_n^2,\ u_{l,n},\ l = 1, \dots, L\}_{n\in\mathbb{N}}$ is a minimizing sequence. Without loss of generality we may suppose $k_n^2 \rightharpoonup k_0$ in $H^{1/2}_{per}(\Omega)$, $k_n^2 \to k_0$ in $L^2(\Omega)$, and $u_{l,n} \rightharpoonup u_{l,0}$ weakly in $H^1_{per}(\Omega)$ since $\|k_n^2\|_{H^{1/2}_{per}(\Omega)}$ and $\|u_{l,n}\|_{H^1_{per}(\Omega)}$ are trivially bounded. Similarly to the proof of Theorem 2.2 we conclude $B(k_n, \theta_l)u_{l,n} \rightharpoonup B(k_0, \theta_l)u_{l,0}$ weakly in $H^1_{per}(\Omega)'$ and thus $\widetilde{B}(\theta_1)^{-1}B(k_n, \theta_l)u_{l,n} \rightharpoonup \widetilde{B}(\theta_1)^{-1}B(k_0, \theta_l)u_{l,0}$ weakly in $H^1_{per}(\Omega)$. Hence, $\widetilde{B}(\theta_1)^{-1}B(k_n, \theta_l)u_{l,n} \to \widetilde{B}(\theta_1)^{-1}B(k_0, \theta_l)u_{l,0}$ strongly in $L^2(\Omega)$. Moreover, $u_{l,n} \rightharpoonup u_{l,0}$ implies that $u_{l,n}|_{\Gamma^\pm} \to u_{l,0}|_{\Gamma^\pm}$ strongly in $L^2(\Gamma^\pm)$ and the strong convergence $F(\theta_l)u_{l,n} \to F(\theta_l)u_{l,0}$. In other words, the first two terms in the objective functionals converge and the limit relations for weakly convergent sequences $\|u_{l,n}\|_{H^1_{per}(\Omega)} \le \liminf \|u_{l,n}\|_{H^1_{per}(\Omega)}$ and $\|k_0^2\|_{H^{1/2}_{per}(\Omega)} \le \liminf \|k_n^2\|_{H^{1/2}_{per}(\Omega)}$ lead us to the upper estimate $\mathcal{F}(k_0^2, u_{1,0}, \dots, u_{L,0}; \gamma) \le \liminf \mathcal{F}(k_n^2, u_{1,n}, \dots, u_{L,n}; \gamma)$. Since $\{k_n^2,\ u_{l,n},\ l = 1, \dots, L\}_{n\in\mathbb{N}}$ is a minimizing sequence, we get that $\mathcal{F}(k_0^2, u_{1,0}, \dots, u_{L,0}; \gamma)$ is the attained minimal value. ∎

Theorem 3.2 *Suppose that, for the give data $A_{meas}(\theta_1), \dots, A_{meas}(\theta_L)$, there exists a wave number function $k_* \in H^{1/2}_{per}(\Omega)$ such that the Rayleigh coefficients corresponding to k_* exactly match the values $A_{meas}(\theta_1), \dots, A_{meas}(\theta_L)$, i.e. $F(\theta_l)u(k_*, \theta_l) + A^i = A_{meas}(\theta_l)$ for the solutions $u(k_*, \theta_l)$ of $B(k_*, \theta_l)u(k_*, \theta_l) = w_l$. Further suppose $0 < \gamma_m \to 0$ and that $\{k_m^2,\ u_{l,m},\ l = 1, \dots, L\}$ is a minimizer of the functional $\mathcal{F}(\dots; \gamma_m)$. Then there exists a $k_0 \in H^{1/2}_{per}(\Omega)$ and a subsequence of $\{k_m^2\}_{m\in\mathbb{N}}$ converging to k_0^2 weakly in $H^{1/2}_{per}(\Omega)$ and strongly in $L^2(\Omega)$. The corresponding solutions $u(k_0, \theta_l)$ of the variational equations (cf. (2.3)) or equivalently of $B(k_0, \theta_l)u(k_0, \theta_l) = w_l$ satisfy $F(\theta_l)u(k_0, \theta_l) + A^i = A_{meas}(\theta_l)$, i.e. their Rayleigh coefficients coincide with the measured data $A_{meas}(\theta_1), \dots, A_{meas}(\theta_L)$.*

Proof: From our assumption on the existence of k_* and from

$$\begin{aligned} c_v\gamma_m \left\|k_m^2\right\|^2_{H^{1/2}_{per}(\Omega)} + c_s\gamma_m \sum_{l=1}^{L} \|u_{l,m}\|^2_{H^1_{per}(\Omega)} &\le \mathcal{F}\Big(k_m^2, u_{1,m}, \dots, u_{L,m}; \gamma_m\Big) \\ &\le \mathcal{F}\Big(k_*^2, u(k_*, \theta_1), \dots, u(k_*, \theta_L); \gamma_m\Big) \\ &= c_v\gamma_m \left\|k_*^2\right\|^2_{H^{1/2}_{per}(\Omega)} + \end{aligned} \tag{3.3}$$

$$c_s \gamma_m \sum_{l=1}^{L} \|u(k_*,\theta_1)\|^2_{H^1_{per}(\Omega)} \longrightarrow 0,$$

we obtain the uniform boundedness of $\|k_m^2\|_{H^{1/2}_{per}(\Omega)}$ and $\|u_{l,m}\|_{H^1_{per}(\Omega)}$. Therefore, we can switch to weakly convergent subsequences. Without loss of generality suppose that k_m^2 and $u_{l,m}$ converge weakly in the corresponding Sobolev spaces. Repeating the arguments of the proof to Theorem 3.1 and using (3.3) lead to

$$\mathcal{F}\Big(k_m^2, u_{1,m}, \dots, u_{L,m}; \gamma_m\Big) \longrightarrow \frac{\sum_{l=1}^{L} \left\| \widetilde{B}(\theta_1)^{-1} B(k_0,\theta_l) u_{l,0} - \widetilde{B}(\theta_1)^{-1} w_l \right\|^2_{L^2(\Omega)}}{\sum_{l=1}^{L} \left\| \widetilde{B}(\theta_1)^{-1} w_l \right\|^2_{L^2(\Omega)}} + c_d \frac{\sum_{l=1}^{L} \left\| \left[F(\theta_l)\, u_{l,0} + A^i \right] - A_{meas}(\theta_l) \right\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}}{\sum_{l=1}^{L} \left\| A_{meas}(\theta_l) \right\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}} = 0\,.$$

The assertions of the theorem follow. ∎

Corollary 3.1 *Suppose the assumptions of the last theorem and, additionally, that the wave number function k_* is the unique solution of the inverse problem, i.e. that the relations $F(\theta_l)u(k,\theta_l) + A^i = A_{meas}(\theta_l)$ and $B(k,\theta_l)u(k,\theta_l) = w_l$ for $k = 1, \dots, L$ imply $k_* = k$. Then the whole sequence $\{k_m^2\}_{m\in\mathbb{N}}$ converges to k_*^2 weakly in $H^{1/2}_{per}(\Omega)$ and strongly in $L^2(\Omega)$.*

Proof: The proof is straightforward since a sequence is convergent if all subsequences have subsequences with a fixed limit. ∎

Remark 3.1 *In general, the uniqueness assumption is hard to verify. For perfectly conducting gratings bounded by a curve of small oscillation represented as a finite Fourier series, uniqueness is proved in [14]. In [13] it has been shown that the knowledge of a finite number of Rayleigh coefficients even for all incident angles is not sufficient to determine the grating. The situation improves slightly if the measurement of Rayleigh coefficients is replaced by the measurement of the field u restricted to the lines $\{(x_1, b_2^+) : 0 < x_1 < d\}$ and $\{(x_1, b_2^-) : 0 < x_1 < d\}$ with $b_2^- < b^- < b^+ < b_2^+$. Note that the differences between the two data types is not so essential if the second data is discretized. Moreover, the theoretical results of this section remain valid for the new kind of measurements. The case of smooth wave number functions depending only on the x_1 variable is treated in [13]. For grating structures corresponding to perfectly conducting gratings bounded by C^2 curves and for the reflected data measured in any direction of incidence, uniqueness is shown in [14]. A fixed incidence direction together with measured data corresponding to a finite number of wave lengths λ is treated in [16]. Gratings consisting of two materials (corresponding to the wave numbers $k^\pm$) separated by a Lipschitz curve and absorbing substrate materials are considered in the subsequent Theorem 3.3. If only local uniqueness in the inverse* *problem is known, then the optimization problem and the numerical methods in Section 5 with suitable initial guess can be used to recover the grating. For local uniqueness, we refer to the local stability results and the papers quoted in [11].*

Corollary 3.2 *Suppose the assumptions of Theorem 3.2 are satisfied. However, consider noisy data $A_{noisy}(\theta_l) \in \ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))$ such that the error to the exactly measured data $A_{meas}(\theta_l)$ satisfies $\|A_{noisy}(\theta_l) - A_{meas}(\theta_l)\|_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))} \le \gamma_m$. Suppose the minimizers are determined for the functional $\mathcal{F}$ with $A_{meas}(\theta_l)$ replaced by $A_{noisy}(\theta_l)$. Then the assertions of Theorem 3.2 remains valid. If, additionally, k_* is the unique solution of the inverse problem, then the assertions of Corollary 3.1 stay in force.*

Proof: The proof is a straightforward modification of that to Theorem 3.2. ∎

Theorem 3.3 *Assume that the graphs $\{(x_1, f_j(x_1)) : \ 0 < x_1 < d\}$ of two different Lipschitz continuous functions f_j $(j = 1, 2)$ cut Ω into an upper region $\{(x_1, x_2) : \ f_j(x_1) < x_2 < b^+\}$ with constant wave number k^+ and a lower region $\{(x_1, x_2) : \ b^- < x_2 < f_j(x_1)\}$ with constant wave number k^- such that $\Re\mathrm{e}\, k^- > 0$ and $\Im\mathrm{m}\, k^- > 0$. For these two gratings and for one planar incident wave ($L = 1$), we assume that u_1 and u_2 are the solutions of the TE problem (cf. (2.2)). Then coincidence of the data $u_1|_{\Gamma^+} = u_2|_{\Gamma^+}$ and $u_1|_{\Gamma^-} = u_2|_{\Gamma^-}$ implies $f_1 = f_2$.*

Remark 3.2 *This generalizes the uniqueness result by Bao [3] for a perfectly reflecting substrate material below the interface.*

Proof: Setting $f(x_1) := \max\{f_1(x_1), f_1(x_1)\}$ and $g(x_1) := \min\{f_1(x_1), f_1(x_1)\}$, we consider the function $u := u_1 - u_2$. Then $u|_{\Gamma^+} = 0$, $u|_{\Gamma^-} = 0$, $\partial_\nu u|_{\Gamma^+} = 0$, and $\partial_\nu u|_{\Gamma^-} = 0$ (cf. (2.5)), which together with the unique continuation theorem implies $u = 0$ in the regions $\{(x_1, x_2) : \ f(x_1) < x_2 < b^+\}$ and $\{(x_1, x_2) : \ b^- < x_2 < g(x_1)\}$. If D (cf. Figure 2) is a simply connected region bounded by the graphs of f and g (and possibly by the vertical lines $\{(x_1, x_2) : \ x_2 = 0\}$ and $\{(x_1, x_2) : \ x_2 = d\}$), then we have $\Delta u_1 + [k^+]^2 u_1 = 0$ and $\Delta u_2 + [k^-]^2 u_2 = 0$ in D, or vice versa. Additionally, we get $u_1 = u_2$ and $\partial_\nu u_1 = \partial_\nu u_2$ on the boundary ∂D of D, where ∂_ν stands for the normal derivative at the boundary points of ∂D. Applying Green's formula, which is justified for $u_j \in H^2(\Omega)$, $j = 1, 2$, we arrive at

$$\begin{aligned} 0 &= \int_D \{u_1 \Delta \overline{u}_1 - \overline{u}_1 \Delta u_1\} = \int_{\partial D} \{u_1 \partial_\nu \overline{u}_1 - \overline{u}_1 \partial_\nu u_1\} = \int_{\partial D} \{u_2 \partial_\nu \overline{u}_2 - \overline{u}_2 \partial_\nu u_2\} \\ &= \int_D \{u_2 \Delta \overline{u}_2 - \overline{u}_2 \Delta u_2\} = 2\mathbf{i}\Im\mathrm{m}\, [k^-]^2 \int_D |u_2|^2 = 0. \end{aligned}$$

Note that for the third equality we have used the quasi-periodicity of the solutions u_j leading to $\{u_j \partial_\nu \overline{u}_j - \overline{u}_j \partial_\nu u_j\} = 0$ over the vertical boundary parts of D. Therefore, $u_2 = 0$ in D. Consequently, $u_2 = 0$ in Ω which is a contradiction to the fact that u_2 is the scattered wave component corresponding to a non-zero incident wave. ∎

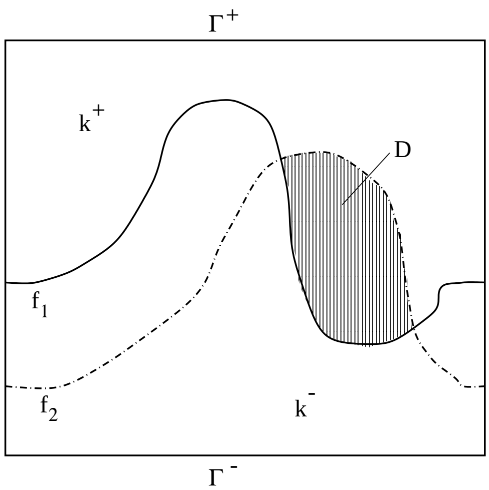

Figure 2: Gratings defined by two graphs.

4 The Finite Element Solution

To define the finite element method, we split domain Ω into the union of triangles such that the diameter of each triangle is less than a prescribed mesh size h and that the triangles have no interior points in common. By S_h^1 we denote the set of all piecewise linear functions subordinate to the partition. Then, the finite element solution u_h of the Helmholtz equation (2.2) in its variational form (2.3) is the unique solution $u_h \in S_h^1$ satisfying

$$\begin{aligned} \mathcal{B}_{TE}(u_h, \varphi_h) &= \int_\Omega \nabla_\alpha u_h \cdot \overline{\nabla_\alpha \varphi_h} - \int_\Omega k^2\, u_h\, \overline{\varphi_h} + \int_{\Gamma^+} (T_\alpha^+ u_h)\, \overline{\varphi_h} + \int_{\Gamma^-} (T_\alpha^- u_h)\, \overline{\varphi_h} \\ &= -\int_{\Gamma^+} 2\mathbf{i}\beta \exp(-\mathbf{i}\beta b^+)\, \overline{\varphi_h}\,, \quad \forall \varphi_h \in S_h^1\,, \end{aligned} \tag{4.1}$$

Clearly, choosing a basis $\{\varphi_{h,j} :\ j = 1, \ldots, N\}$ of S_h^1, the last discrete variational equation is equivalent to an equation in the N dimensional complex Euclidean space $\ell^2_{\mathbb{C}}(N)$, i.e. to the matrix equation $B_h \xi = \eta$ for the unknown coefficients ξ_j of the function u_h, where

$$\begin{aligned} B_h &:= \Big(\mathcal{B}_{TE}(\varphi_{h,j}, \varphi_{h,j'})\Big)_{j',j=1,\ldots,N}, \\ \xi &:= (\xi_j)_{j=1,\ldots,N}\,, \quad u_h(x_1,x_2) \;=\; \sum_{j=1}^N \xi_j\, \varphi_{h,j}(x_1,x_2), \\ \eta &:= (\eta_j)_{j=1,\ldots,N}, \quad \eta_j := -\int_{\Gamma^+} 2\mathbf{i}\beta \exp(-\mathbf{i}\beta b^+)\, \overline{\varphi_{h,j}}\;. \end{aligned}$$

In other words, if the function $k = k(x_1, x_2)$ is given, then we can determine an approximate solution u_h by solving $B_h \xi = \eta$. Clearly, the matrix, the right-hand side, and the solution depend on the angle of incidence θ and on the wave number function $k = k(x_1, x_2)$. To indicate this dependence, we write $u_h = u_h(k, \theta)$ and the matrix equation as $B_h(k, \theta)\xi(k, \theta) = \eta(\theta)$.

Restricting the solution $\Omega \ni (x_1, x_2) \mapsto u_h(k,\theta)(x_1,x_2)$ to $\Gamma^{\pm}$, we get the Rayleigh coefficients $A_n^{\pm}$ by computing the Fourier coefficients according to (2.6). We denote the so obtained approximate values of $A_n^{\pm}$ by $A_{h,n}^{\pm}$ and get the approximate efficiencies by setting $e_{h,n}^{\pm} := (\beta_n^{\pm}/\beta)|A_{h,n}^{\pm}|^2$. The linear operator of restricting u_h to $\Gamma^{\pm}$ and of computing the Rayleigh coefficients $A_{h,n}^{\pm}$ will be denoted by $F_h(\theta)$, i.e.,

$$A_h \;:=\; \left(A_{h,n}^{\pm}\right)_{(n,\pm)\in\mathcal{U}^*} \;=\; F_h(\theta)\,\xi + A^i.$$

If the refractive indices of the cover material above the grating and the substrate material beneath the grating are fixed, then the operator $F_h(\theta)$ is independent of the function k inside of the grating.

Finally, we remark that the linear system of equations $B_h(k,\theta)\xi(k,\theta) = \eta(\theta)$ is the discretization of a second order differential equations. Consequently, the condition number of the finite element matrix $B_h(k,\theta)$ behaves like $\mathcal{O}(h^{-2})$ for h tending to zero. Therefore, a preconditioner is used for the iterative solution of $B_h(k,\theta)\xi(k,\theta) = \eta(\theta)$, i.e. we solve $[\widetilde{B}_h(\theta)^{-1}B_h(k,\theta)]\xi(k,\theta) = \widetilde{B}_h(\theta)^{-1}\eta(\theta)$ instead of $B_h(k,\theta)\xi(k,\theta) = \eta(\theta)$ with a matrix $\widetilde{B}_h(\theta)$ easy to invert and close to $B_h(k,\theta)$. Several preconditioning techniques are possible. In our special case, we can choose $\widetilde{B}_h(\theta)$ e.g. as the finite element matrix $\widetilde{B}_h(\theta) := B_h(\tilde{k},\theta)$, where the wave number function $\tilde{k}(x_1,x_2)$ is equal to k^+ for $x_2 > (b^- + b^+)/2$ and equal to k^- for $x_2 < (b^- + b^+)/2$. If the partition of the finite element method is obtained by dividing the rectangles of a uniform rectangular partition of the rectangle Ω along the diagonals, then $B_h(\tilde{k},\theta)$ is easy to invert. Indeed, if we group the degrees of freedom in clusters according to their x_2 coordinates, then $B_h(\tilde{k},\theta)$ is a triangular block matrix with circular blocks.

5 The Discretized Inverse Problem

For a numerical solution of the inverse problem of Section 3, we switch to the discrete level, i.e. we seek the coefficient vectors $\xi(l)$ of the finite element solutions $u_h(k,\theta_l) = \sum \xi(l)_j \varphi_{h,j}$ from $\widetilde{B}_h(\theta_1)^{-1}B_h(k,\theta_l)\xi(l) \approx \widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l)$ such that the computed Rayleigh coefficients $A_h(\theta_l) = F_h(\theta_l)\xi(l) + A^i$ differ only slightly from the measured $A_{meas}(\theta_l) := (A_{meas,n}^{\pm}(\theta_l))_{(n,\pm)\in\mathcal{U}^*(\theta_l)}$. The unknown squared wave number function k^2 is to be approximated by a function from a discrete space. We fix a partition coarser than that of the finite element method and choose the space S_h^0 as the set of all functions which are piecewise constant subordinate to the fixed partition and which fulfill $k^2(x_1,x_2) = [k^+]^2$ for $0 < x_1 < d$ and $b_1^+ < x_2 < b^+$ as well as $k^2 = [k^-]^2$ for $0 < x_1 < d$ and $b^- < x_2 < b_1^-$. As usual, the corresponding basis of functions equal to one over one triangle of the grid and to zero over the others will be denoted by $\{\chi_{h,j} : \; j = 1,\ldots,M\}$. We can identify the functions $k^2 \in S_h^0$ with the vectors of coefficients $\kappa = (\kappa_j)_{j=1,\ldots,M}$ satisfying $k(x_1,x_2)^2 = \sum_j \kappa_j \chi_{h,j}(x_1,x_2)$. In particular, we write $B_h(\kappa,\theta_l)$ for $B_h(\sum_j \sqrt{\kappa_j}\chi_{h,j},\theta_l)$. Using the discretized and reduced $H_{per}^{1/2}(\Omega)$ norm

$$\|\kappa\|_{V(\Omega)}^2 \;:=\; \sum_j |\kappa_j|^2 w_j + \sum_{\substack{j,j':\\ \text{indices of}\\ \text{neighbours}}} |\kappa_j - \kappa_{j'}|^2 \sqrt{w_j}$$

with w_j the measure of the jth triangle, we define the discrete non-linear objective functional by

$$\begin{aligned}\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big) \;:=\; & \frac{\sum_{l=1}^{L}\Big\|\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)\xi(l)-\widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l)\Big\|^2_{\ell^2_{\mathbb{C}}(N)}}{\sum_{l=1}^{L}\Big\|\widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l)\Big\|^2_{\ell^2_{\mathbb{C}}(N)}}\\ & + c_d\,\frac{\sum_{l=1}^{L}\Big\|\big[F_h(\theta_l)\;\xi(l)+A^i\big]-A_{meas}(\theta_l)\Big\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}}{\sum_{l=1}^{L}\Big\|A_{meas}(\theta_l)\Big\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}}\\ & + c_v\gamma\,\|\kappa\|^2_{V(\Omega)} \;+\; c_s\gamma\sum_{l=1}^{L}\Big\|\sum_{j=1}^{N}\xi(l)_j\varphi_{h,j}\Big\|^2_{H^1_{per}(\Omega)}\,. \end{aligned} \tag{5.1}$$

Here c_d, c_v, and c_s denote appropriate equilibration constants which are to be determined by numerical experiments, and γ is the regularization parameter. Using the functional $\mathcal{F}_h$, we finally arrive at the discrete optimization problem

$$\begin{array}{c}\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big)\qquad\longrightarrow\quad\min\;.\\ \kappa\in\ell^2_{\mathbb{C}}(M),\\ \xi(l)\in\ell^2_{\mathbb{C}}(N),\; l=1,\ldots,L\end{array} \tag{5.2}$$

This will be solved using the following non-linear conjugate gradient algorithm which we prepare by giving formulas for the gradients. Note that the complex variables are treated as couples of real variables.

First we observe that the mapping $k^2\mapsto[B_h(k,\theta_l)-B_h(0,\theta_l)]=(-\int_\Omega k^2\varphi_{h,j}\overline{\varphi_{h,j'}})_{j',j}$ is linear and independent of θ_l. We easily get that

$$\begin{aligned}\nabla_\kappa B_h(\kappa,\theta_l) \;&=\; \nabla_\kappa B_h \;=\; \Big(\Big[\nabla_\kappa B_h\Big]_j\Big)_{j=1,\ldots,M}\,,\\ \Big[\nabla_\kappa B_h\Big]_j \;&:=\; \Big(-\int_\Omega\varphi_{h,i}\overline{\varphi_{h,i'}}\chi_{h,j}\Big)_{i',i=1,\ldots,N}\,,\\ \nabla_\kappa B_h\;\kappa' \;&=\; \sum_{j=1}^{M}\Big[\nabla_\kappa B_h\Big]_j\kappa'_j\;.\end{aligned}$$

If $\langle\cdot,\cdot\rangle_{\ell^2_{\mathbb{C}}(N)}$ stands for the scalar product in the N-dimensional Euclidean space, then the gradient of $\mathcal{F}_h$ is given by

$$\begin{aligned}&\nabla\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big)\Big(\kappa',\xi(1)',\ldots,\xi(L)'\Big)=\\ &\Big(\nabla_\kappa\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big)\,\kappa',\nabla_{\xi(1)}\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big)\,\xi(1)',\\ &\quad\nabla_{\xi(2)}\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big)\xi(2)',\ldots,\nabla_{\xi(L)}\mathcal{F}_h\Big(\kappa,\xi(1),\ldots,\xi(L);\gamma\Big)\xi(L)'\Big),\end{aligned}$$

$$
\begin{aligned}
&\nabla_\kappa \mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)\kappa' = \\
&\Re\mathrm{e}\Big\langle 2\rho_1 \sum_{l=1}^{L} \Big\langle \widetilde{B}_h(\theta_1)^{-1}\Big[B_h(\kappa,\theta_l)\xi(l) - \eta(\theta_l)\Big], \widetilde{B}_h(\theta_1)^{-1}\nabla_\kappa B_h\ \xi(l)\Big\rangle_{\ell^2_{\mathbb{C}}(N)}, \kappa'\Big\rangle_{\ell^2_{\mathbb{C}}(M)}, \\
&\nabla_{\xi(l)} \mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)\xi(l)' = \\
&\Re\mathrm{e}\Big\langle 2\rho_1 \Big[\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)\Big]^* \Big[\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)\xi(l) - \widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l)\Big], \xi(l)'\Big\rangle_{\ell^2_{\mathbb{C}}(N)} + \\
&\Re\mathrm{e}\Big\langle 2\rho_2\, F_h(\theta_l)^*\Big[\Big[F_h(\theta_l)\ \xi(l) + A^i\Big] - A_{meas}(\theta_l)\Big], \xi(l)'\Big\rangle_{\ell^2_{\mathbb{C}}(N)} + \\
&\Re\mathrm{e}\Big\langle 2\, c_s\gamma\ \Big[\sum_{j=1}^{N}\xi(l)_j\varphi_{h,j}\Big], \Big[\sum_{j'=1}^{N}\xi(l)'_{j'}\varphi_{h,j'}\Big]\Big\rangle_{H^1_{per}(\Omega)}, \\
&\rho_1 := \frac{1}{\sum_{l=1}^{L}\Big\|\widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l)\Big\|^2_{\ell^2_{\mathbb{C}}(N)}}, \quad \rho_2 := \frac{c_d}{\sum_{l=1}^{L}\Big\|A_{meas}(\theta_l)\Big\|^2_{\ell^2_{\mathbb{C}}(\mathcal{U}^*(\theta_l))}}.
\end{aligned}
$$

Here $[\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)]^*$ is the adjoint (transposed and complex conjugate) of matrix $[\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)]$ and $F_h(\theta_l)^*$ that of $F_h(\theta_l)$. Treating the gradients as vectors, we arrive at

$$
\begin{aligned}
\nabla\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big) = \Big(&\nabla_\kappa\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big), \nabla_{\xi(1)}\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big), \\
&\nabla_{\xi(2)}\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big), \ldots, \nabla_{\xi(L)}\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)\Big),
\end{aligned}
$$

$$
\begin{aligned}
\nabla_\kappa\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big) &= \Big(\nabla_\kappa\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)_j\Big)_{j=1,\ldots,M} \in \ell^2_{\mathbb{C}}(M), \\
\nabla_\kappa\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)_j &:= 2\rho_1 \times \\
&\sum_{l=1}^{L}\Big\langle \widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)\xi(l) - \widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l), \widetilde{B}_h(\theta_1)^{-1}\Big[\nabla_\kappa B_h\Big]_j\xi(l)\Big\rangle_{\ell^2_{\mathbb{C}}(N)} + \\
&2\, c_v\gamma\Big\{w_j\kappa_j + \sqrt{w_j}\sum_{\substack{j':j,j' \text{ are} \\ \text{indices of} \\ \text{neighbours}}}\big[\kappa_j - \kappa_{j'}\big]\Big\}, \\
\nabla_{\xi(l)}\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big) &= \Big(\nabla_{\xi(l)}\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)_j\Big)_{j=1,\ldots,N} \in \ell^2_{\mathbb{C}}(N), \\
\nabla_{\xi(l)}\mathcal{F}_h\Big(\kappa, \xi(1), \ldots, \xi(L); \gamma\Big)_j &:= 2\rho_1 \times \\
&\Big[\Big[\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)\Big]^*\Big[\widetilde{B}_h(\theta_1)^{-1}B_h(\kappa,\theta_l)\xi(l) - \widetilde{B}_h(\theta_1)^{-1}\eta(\theta_l)\Big]\Big]_j + \\
&2\rho_2\Big[F_h(\theta_l)^*\Big[\Big[F_h(\theta_l)\ \xi(l) + A^i\Big] - A_{meas}(\theta_l)\Big]\Big]_j +
\end{aligned}
$$

$$2\, c_s\gamma \left\langle \left[\sum_{j'=1}^{N} \xi(l)_{j'}\varphi_{h,j'}\right], \varphi_{h,j}\right\rangle_{H^1_{per}(\Omega)} .$$

Now the non-linear conjugate gradient algorithm of Fletcher-Reeves modified by Polak-Ribière (cf. e.g. [17]) takes the form

Conjugate Gradient Algorithm

Given the constant $0 < c_1 = 10^{-3}$*;*
Given the initial guess $\Big(\kappa_0, \xi_0(1), \ldots, \xi_0(L)\Big)$*;*
Evaluate $\mathcal{F}_{h,0} = \mathcal{F}_h\Big(\kappa_0, \xi_0(1), \ldots, \xi_0(L); \gamma\Big), \nabla\mathcal{F}_{h,0} := \nabla\mathcal{F}_h\Big(\kappa_0, \xi_0(1), \ldots, \xi_0(L); \gamma\Big)$*;*
Set the first search direction $p_0 := \Big(\kappa_0^d, \xi_0^d(1), \ldots, \xi_0^d(L)\Big) = -\nabla\mathcal{F}_{h,0}$*, set* $j = 0$*;*
while $\nabla\mathcal{F}_{h,j} = \nabla\mathcal{F}_h\Big(\kappa_j, \xi_j(1), \ldots, \xi_j(L); \gamma\Big) \neq 0$

Compute step size α_j *of the correction* $\Big(\alpha_j\kappa_j^d, \alpha_j\xi_j^d(1), \ldots, \alpha_j\xi_j^d(L)\Big)$ *such that* α_j *is the largest number in* $\{256, 128, 64, 32, 16, \ldots\}$ *with*

$$\mathcal{F}_h\Big(\kappa_j + \alpha_j\kappa_j^d, \xi_j(1) + \alpha_j\xi_j^d(1), \ldots, \xi_j(L) + \alpha_j\xi_j^d(L); \gamma\Big) \leq \mathcal{F}_h\Big(\kappa_j, \xi_j(1), \ldots, \xi_j(L); \gamma\Big) + c_1\alpha_j\nabla\mathcal{F}_{h,j}^T\Big(\kappa_j^d, \xi_j^d(1), \ldots, \xi_j^d(L); \gamma\Big);$$

Set the new iterate solution $\Big(\kappa_{j+1}, \xi_{j+1}(1), \ldots, \xi_{j+1}(L)\Big)$ *to* $\Big(\kappa_j + \alpha_j\kappa_j^d, \xi_j(1) + \alpha_j\xi_j^d(1), \ldots, \xi_j(L) + \alpha_j\xi_j^d(L)\Big)$*;*
Evaluate gradient $\nabla\mathcal{F}_{j+1} := \nabla\mathcal{F}\Big(\kappa_{j+1}, \xi_{j+1}(1), \ldots, \xi_{j+1}(L); \gamma\Big)$*;*
Set $\beta_{j+1} = \max\left\{\dfrac{\nabla\mathcal{F}_{j+1}^T(\nabla\mathcal{F}_{j+1} - \nabla\mathcal{F}_j)}{\|\nabla\mathcal{F}_j\|^2}, 0\right\}$*;*
Set $p_{j+1} := \Big(\kappa_{j+1}^d, \xi_{j+1}^d(1), \ldots, \xi_{j+1}^d(L)\Big) = -\nabla\mathcal{F}_{j+1} + \beta_{j+1}\Big(\kappa_j^d, \xi_j^d(1), \ldots, \xi_j^d(L)\Big)$*;*
Set $j = j + 1$

end(while)

The line search part, i.e. the determination of the α_j can be improved. In fact, instead of changing α_j to half its value, we can take the argument of the minimum of a quadratic interpolation to $\alpha_j \mapsto \mathcal{F}(\kappa_j + \alpha_j\kappa_j^d, \xi_j(1) + \alpha_j\xi_j^d(1), \ldots, \xi_j(L) + \alpha_j\xi_j^d(L); \gamma)$ as the next value for α_j.

Usually, this conjugate gradient method converges to a local minimum of the objective function $\mathcal{F}_h$. The determination of the global minimum for high dimensional optimization is a difficult and expensive problem. Note that a high number of degrees of freedom is required for the finite element method in order to resolve the oscillations of the Helmholtz equation. Even if a fast method for the computation of the global minimum were available, we would have to be careful. Indeed, the global solution of the optimization problem with regularization parameter γ set to zero might be close to a local minimum of the regularized problem ($\gamma > 0$) different from the global minimum. In any case, due to the locality of the conjugate gradient solution, the choice of the initial solution is very important. Fortunately, for our numerical experiments, the choice of the initial guess as

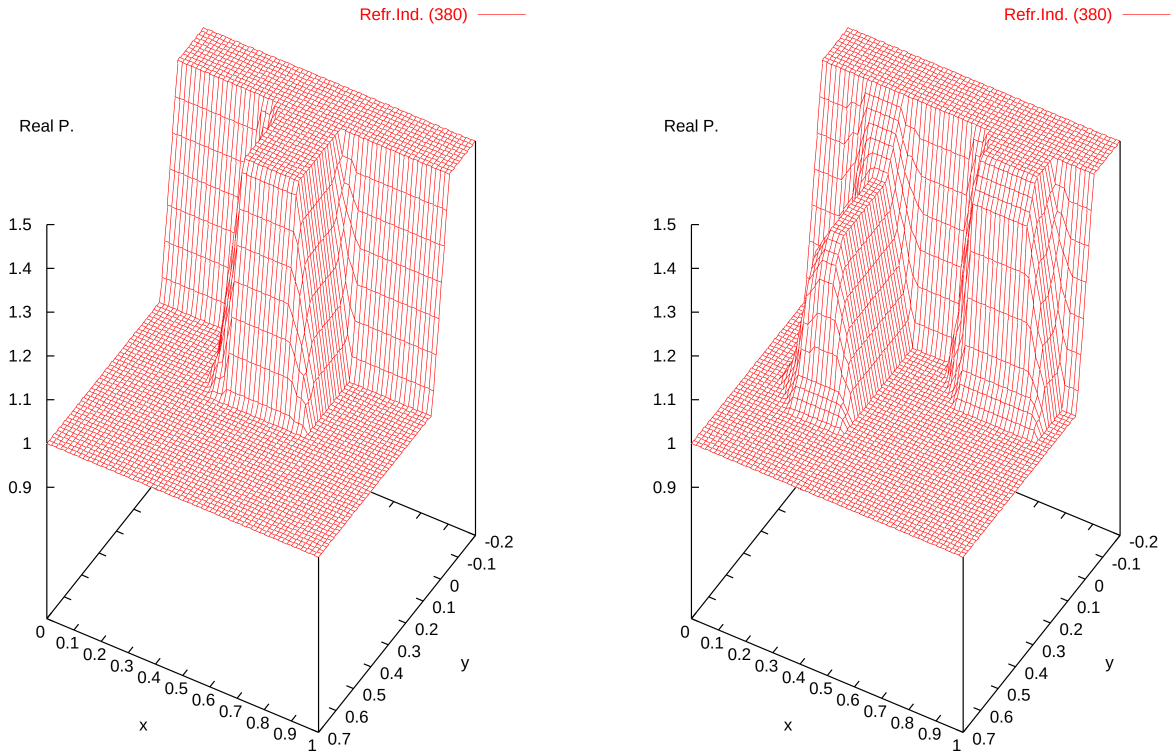

Figure 3: The two gratings: Rectangular & Two Towers

the mean value wave number function and the corresponding solutions of the Helmholtz solutions, i.e.

$$\begin{aligned} \kappa_{0,j} &:= k_0^2,\ j = 1,\ldots,M, \quad k_0 := \frac{k^- + k^+}{2}, \\ \xi_0(l) &:= [B_h(k_0,\theta_l)]^{-1}\eta(\theta_l),\ l = 1,\ldots,L \end{aligned}$$

was satisfactory.

6 The Numerical Experiment

For our numerical tests we consider two gratings. Both are chosen with $b^- = -0.2\,\mu$m, $b_1^- = -0\,\mu$m, $b_1^+ = -0.5\,\mu$m, and $b^+ = 0.7\,\mu$m. The period is $d = 1\,\mu$m. The grating materials are characterized by the refractive index ν which determines the value of the wave number function by the formula $k = \nu d/\lambda$. The wave length of light is $\lambda = 635\,$nm. The cover material over the grating (for $x_2 > b_1^+$) is Air with $\nu = 1$. The index of the substrate material (for $x_2 < b_1^-$) is $\nu = 1.5$. The first grating is rectangular (cf. Figure 3 where a continuous linear interpolation of the piecewise constant function is plotted), i.e. the refractive index is

$$\nu = \nu(x_1,x_2) \quad := \quad \begin{cases} 1.5 & \text{for } \left|x_1 - \frac{1}{2}\right| \le \frac{1}{6} \text{ and } x_2 \le \frac{1}{4} \\ 1.0 & \text{for } \left|x_1 - \frac{1}{2}\right| > \frac{1}{6} \text{ or } x_2 > \frac{1}{4}\,. \end{cases}$$

The second (cf. Figure 3) is a two-tower-grating with

$$\nu = \nu(x_1, x_2) \quad := \quad \begin{cases} 1.5 & \text{for } \left|x_1 - \frac{3}{4}\right| \le \frac{1}{6} \text{ and } x_2 \le \frac{1}{8} \\ 1.35 & \text{for } \left|x_1 - \frac{1}{4}\right| \le \frac{1}{6} \text{ and } x_2 \le \frac{3}{8} \\ 1.5 & \text{for } x_2 < 0 \\ 1.0 & \text{else .} \end{cases}$$

For the two gratings, we have computed the Rayleigh coefficients corresponding to non-evanescent modes under the angles of incidence $\theta_l,\ l = 1, \dots L$.

$$\{\theta_l:\ l = 1, \dots L\} \quad := \quad \begin{cases} \{0\} & \text{if } L = 1 \\ \{-60, 0, 60\} & \text{if } L = 3 \\ \{-60, -40, -20, 0, 20, 40, 60\} & \text{if } L = 7 \\ \{-60, -55, -50, -45, \dots, 45, 50, 55, 60\} & \text{if } L = 25\ . \end{cases}$$

Depending on the angle of incidence, these Rayleigh coefficients are three numbers of the $A_n^+,\ n = -2, -1, 0, 1, 2$ and five numbers of the $A_n^-,\ n = -3, -2, -1, 0, 1, 2, 3$. From these numbers we have to recover the grating by the inverse algorithm described in Section 5.

Since our simulated measurement data should be obtained by a method different from that involved in the inverse algorithm, we have computed the $A_n^\pm$ by a finite element method over a high level non-uniform triangulation. Actually we have employed a standard grid generator and more than 200 000 unknown finite elements. The finite element operator $B_h(k, \theta)$ used for the algorithm of Section 5 is based on a coarse uniform triangulation. More precisely, we split the domain $\Omega = (0, 1) \times (-0.2, 0.7)$ into 40×36 equal squares and divide each square into two triangles by a cut along the diagonal. Taking into account the periodicity, the resulting number of finite elements is 1 600. The unknown wave number function is sought as a function piecewise continuous over the triangulation resulting from halving the squares of a 20×18 uniform rectangular partition. This means 720 triangles in Ω and exactly 400 unknowns for the wave number function corresponding to the triangles falling into the strip $(0, d) \times (b_1^-, b_1^+) = (0, 1) \times (0, 0.5)$.

The constants c_s, c_d, c_v, and γ have to be adapted to the special case at hand. So one should take a typical example with known wave number solution and determine the constants such that the resulting approximation of the wave number function is the closest to the known exact solution. Then the unknown gratings should be recovered using the just obtained constants.

Following this philosophy, we have determined the optimal constants for the first rectangular grating. We have set $c_d = 0.005$ and the other numbers including the number of necessary conjugate gradient iterations are given in Table 1. The plots of the resulting reconstructed wave number functions are shown in Figures 4 and 5. With larger L, i.e. with more measurement data the recovered wave number improves slightly.

Next we have taken the optimal parameters of the rectangular grating and employed them in the algorithm for the two-towers-grating. Figures 6 and 7 show the results of the reconstruction which are close to the exact function (cf. Figure 3).

Acknowledgements. The authors gratefully acknowledge the support of the German Ministry of Education, Research and Technology under Grant No. 03-ELM3B5.

L	$c_v q$	$c_s q$	Iterations
1	0.000 9	0.000 3	2 000
3	0.000 5	0.000 05	8 000
7	0.000 03	0.000 001	25 000
25	0.000 2	0.000 000 01	50 000

Table 1: Constants for the objective functional.

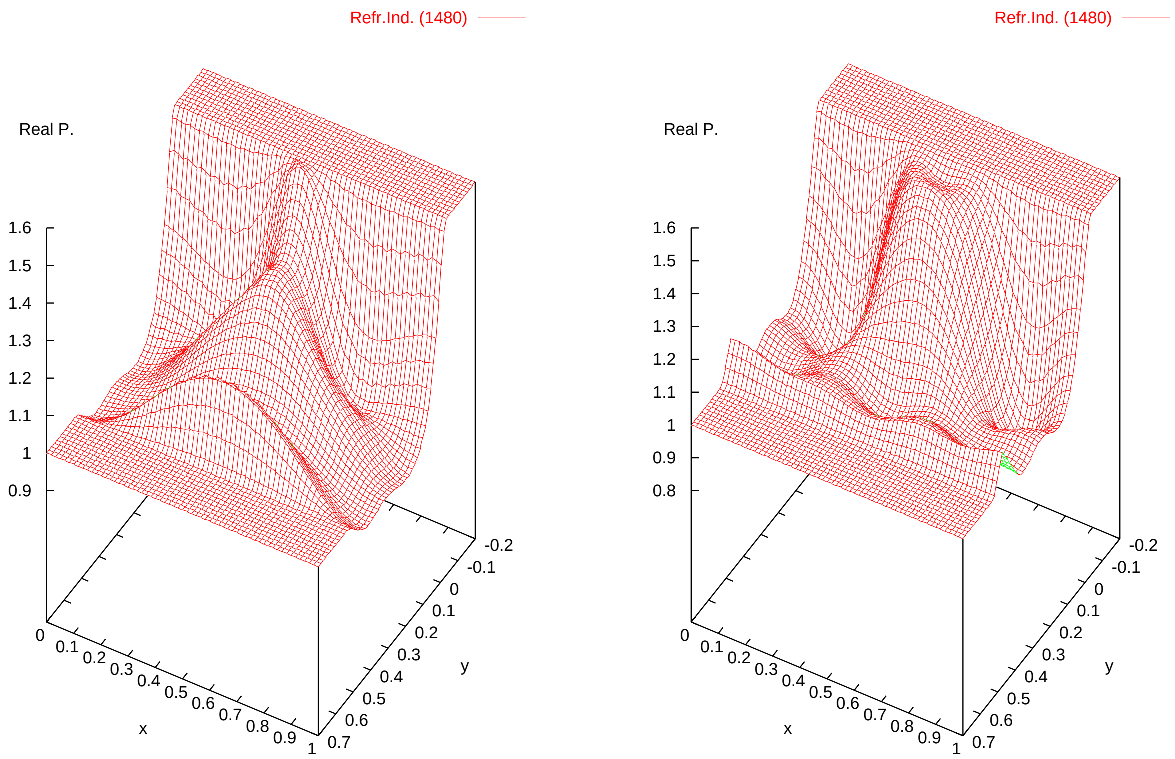

Figure 4: Reconstructed rectangular grating. $L = 1$ and $L = 3$

References

[1] Y. Achdou and O. Pironneau, Optimization of a photocell, Optimal Control Applications & Methods **12**, (1991) 221–246.

[2] T. Arens and A. Kirsch, The factorization method in inverse scattering from periodic structures, Preprint 02/12, Fakultät für Mathematik, Universität Karlsruhe, 2002.

[3] G. Bao, A uniqueness theorem for an inverse problem in periodic diffractive optics, Inverse Problems **10**, (1994) 335–340.

[4] G. Bao, D.C. Dobson, and J.A. Cox, Mathematical studies in rigorous grating theory, J. Opt. Soc. Amer. **A 12**, (1995) 1029–1042.

[5] A.-S. Bonnet-Bendhia and F. Starling, Guided waves by electromagnetic gratings and non–uniqueness examples for the diffraction problem, Math. Meth. Appl. Sci. **17**, (1994) 305–338.

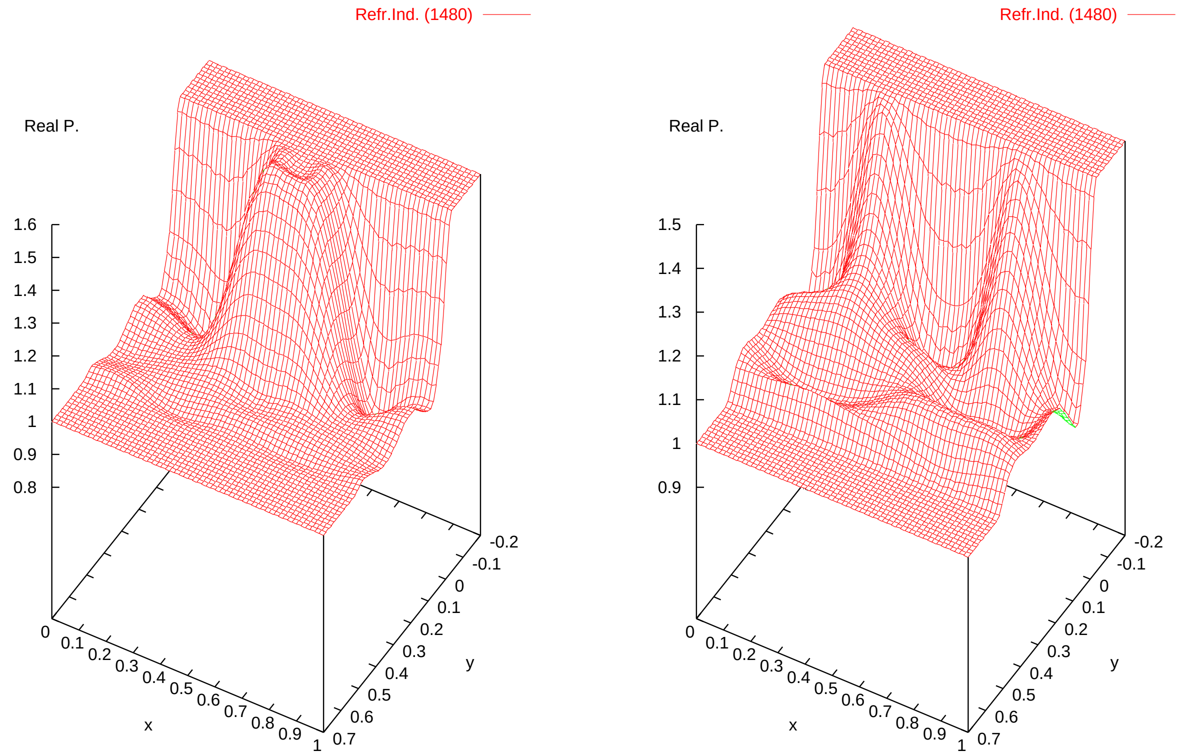

Figure 5: Reconstructed rectangular grating. $L = 7$ and $L = 25$

[6] G. Bruckner, J. Elschner, and M. Yamamoto, An optimization method for grating profile reconstruction, WIAS Preprint 682, Weierstrass Institute for Applied Analysis and Stochastics, Berlin, 2002, to appear in Proceedings of ISAAC 2001.

[7] G. Bruckner and J. Elschner, A two-step algorithm for the reconstruction of perfectly reflecting periodic profiles, WIAS Preprint ***, Weierstrass Institute for Applied Analysis and Stochastics, Berlin, 2002.

[8] P. Ciarlet, The Finite Element Method for Elliptic Problems, North–Holland, Amsterdam, New York, 1978.

[9] D. Colton and R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory, Second Edition, Applied Mathematical Sciences 93, Springer Verlag, Berlin, Heidelberg, 1998.

[10] J. Elschner and G. Schmidt, Diffraction in periodic structures and optimal design of binary gratings. I. Direct problems and gradient formulas, Math. Meth. Appl. Sci. **21**, (1998) 1297–1342.

[11] J. Elschner and G. Schmidt, Inverse scattering for periodic structures: Stability of polygonal interfaces, Inverse Problems **17**, (2001) 1817–1829.

[12] K. Ito and F. Reitich, A high-order perturbation approach to profile reconstruction: I. Perfectly conducting gratings, Inverse Problems **15**, (1999) 1067–1085.

[13] A. Kirsch, An inverse scattering problem for periodic structures, in: R.E. Kleinman, R. Kress, and E. Martensen (eds.), Inverse Scattering and Potential Problems in Mathematical Physics, Methoden und Verfahren der mathematischen Physik, Bd. 40, Lang, Frankfurt am Main, 1995, 75–93.

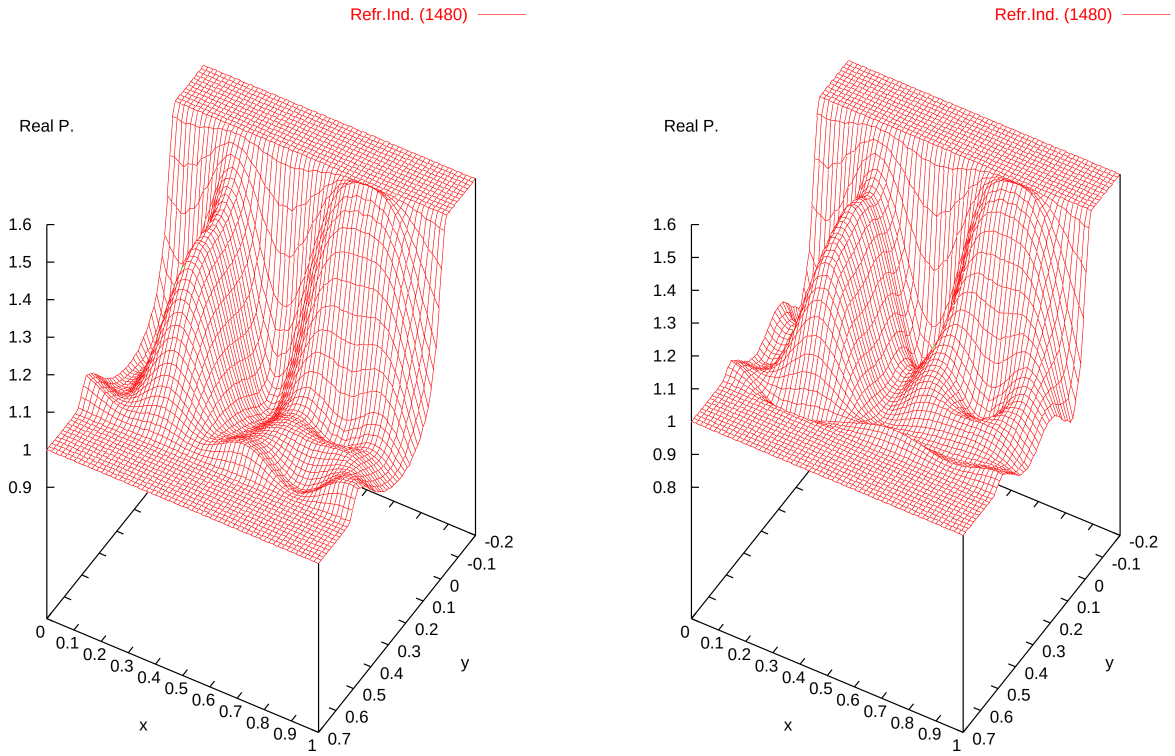

Figure 6: Reconstructed two towers. $L = 1$ and $L = 3$

[14] A. Kirsch, Uniqueness theorems in inverse scattering theory for periodic structures, Inverse Problems **10**, (1994) 145–152.

[15] F. Hettlich, Iterative regularization schemes in inverse scattering by periodic structures, Preprint 01/31, Fakultät für Mathematik, Universität Karlsruhe, 2002.

[16] F. Hettlich and A. Kirsch, Schiffer's theorem in inverse scattering theory for periodic structures, Inverse Problems **13**, (1997) 351–361.

[17] J. Nocedal and S.J. Wright, Numerical Optimization, Springer Series in Operations Research, New York, 1999.

[18] R. Petit (ed.), Electromagnetic Theory of Gratings, Springer, Berlin, 1980.

[19] J. Turunen, F. Wyrowski (eds.), Diffractive Optics for Industrial and Commercial Applications, Akademie Verlag, Berlin, 1997.

[20] H.P. Urbach, Convergence of the Galerkin method for two-dimensional electro-magnetic problems, SIAM J.Numer.Anal. **28**, (1991) 697–710.

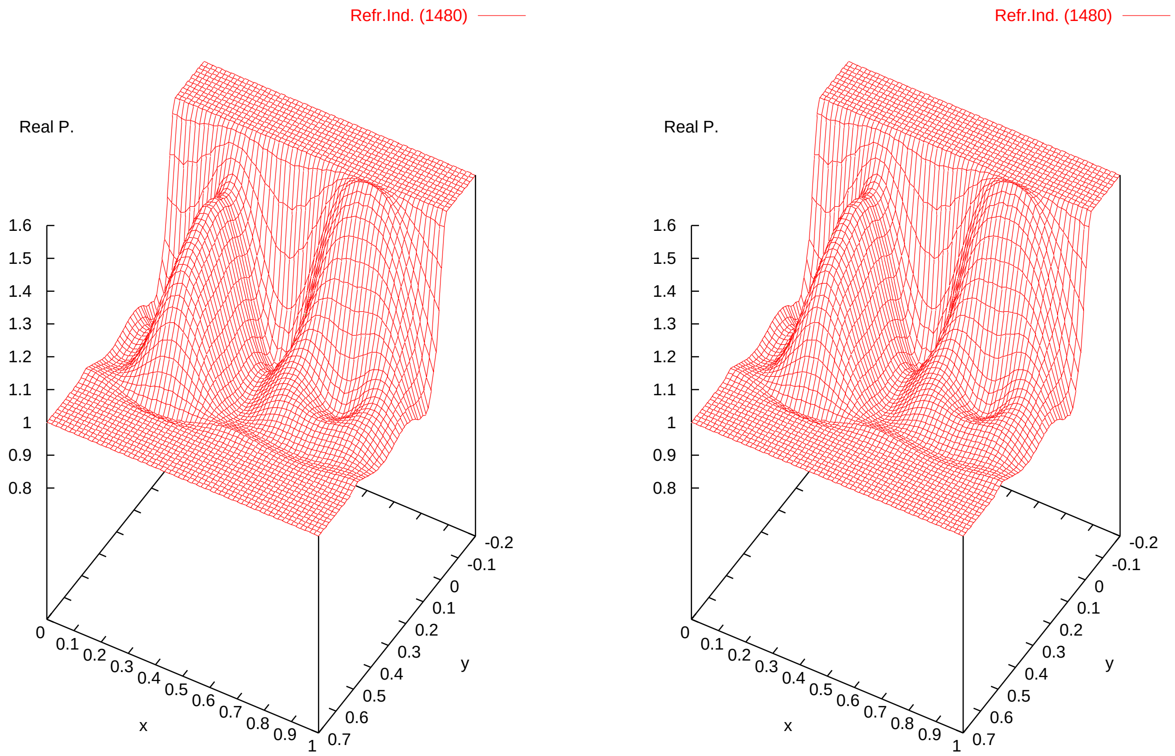

Figure 7: Reconstructed two towers. $L = 7$ and $L = 25$