Modeling and Optimal Design of Diffractive Optical Structures

Gang Bao
Department of Mathematics
University of Florida
Gainesville, FL 32611-2080

David C. Dobson
Department of Mathematics
Texas A&M University
College Station, TX 77843-3368

Abstract

In this paper we describe recent developments in the application of mathematical and computational techniques to problems in diffractive optics, particularly to problems involving spatially periodic structures (gratings). Many of the techniques are also applicable to other scattering problems involving periodic structures. We focus on variational techniques, finite element approximations, inverse and optimal design problems, and nonlinear materials.

1 Introduction

Traditional optical devices have structural features that are much larger than the length of a typical visible light wave, say $0.5\mu m$. One consequence of this great disparity in length scales is that geometrical optics approximations to the underlying electromagnetic equations are generally very accurate. Thus the behavior of traditional optical elements can usually be predicted by ray-tracing methods that have been in use for hundreds of years.

With the advent of high precision micromachining techniques, it is presently practical to fabricate optical devices with complicated structural features on the order of $0.25\mu m$. Such "diffractive optics" devices have great advantages in terms of size and weight, and can often be designed to perform functions unattainable with traditional optical elements. For example, structures with spatially periodic features (diffraction gratings) are used as spectral filters, polarizers, waveguide couplers, etc. The development and application of this new technology increasingly relies on accurate mathematical models and numerical calculations both for the prediction of device behavior and for the determination of "optimal" device designs. In contrast to the case of traditional optical structures, geometrical optics is generally not sufficiently accurate for these diffractive devices. The computational problem is much more challenging, requiring the solution of a full partial differential equation model.

In this survey we describe some recent progress in the application of variational techniques to wave propagation problems associated with spatially periodic diffractive structures. We discuss weak formulations and finite element approximations of the time-harmonic Maxwell's equations in such structures, with both linear and nonlinear optical materials. We also discuss the inverse problem of determining shapes from scattered fields and we describe some recent work aimed at finding optimal shape designs for surface-relief gratings. Although this work has been motivated by diffractive optics applications, many of the techniques presented are applicable to other periodic scattering problems in antennas, acoustics, crystalline materials, etc.

Much of our discussion on the solution of periodic scattering problems is centered around our own work, which follows a variational formulation of the problem. We emphasize at the outset that there are many other important approaches which are applicable to problems in diffractive optics. Of particular significance are the methods of variation of boundaries (MVB) of Bruno and Reitich [22, 23], the coupled-waves method of Gaylord and Moharam [38], and the integral equation approaches of Nédélec [47]. The method of Bruno and Reitich has the advantages of speed, accuracy and conceptual simplicity, and is probably the best known method for the solution of problems involving relatively simple, smooth surface-relief structures. The coupled-waves approach of Gaylord and Moharam is also conceptually simple and is particularly well suited for typical structures manufactured by "mask-etch" production techniques in common use in industry. Various implementations and modifications of this basic method have been in wide use in the engineering community and in industry for the past decade or so. The integral equation approaches of Nédélec are mathematically attractive and of quite general applicability. Our focus on variational methods is due mainly to their flexibility, robustness, and mathematical simplicity. Careful numerical (finite element) implementations of the underlying variational approach, combined with iterative linear systems solvers and preconditioning techinques result in methods which are very competitive with those discussed above for most practical problems. In summary, as is often the case with numerical techniques, there is no one "best" method. Each approach has distinct advantages and disadvantages for different types of problems.

For periodic structures of *finite* extent, the asymptotic matching techniques of Kriegsmann [46] can be applied with any of the above methods.

The reader is referred to Petit [48] for various aspects of the electromagnetic theory of gratings up to the early 1980s, including numerical methods and a comparison of various approaches. Descriptions of several mathematical problems which arise in industrial applications may be found in the books by Friedman [35]–[37].

2 Diffraction in biperiodic structures

The general geometrical configuration for the scattering problems we consider is illustrated in Figure 1. The basic problem is to predict the scattered modes which arise when a linearly polarized electromagnetic plane wave is incident on the biperiodic structure. In this survey we consider only the infinite periodic geometry. The important case of a periodic structure of finite extent can be analyzed using a field-matching technique due

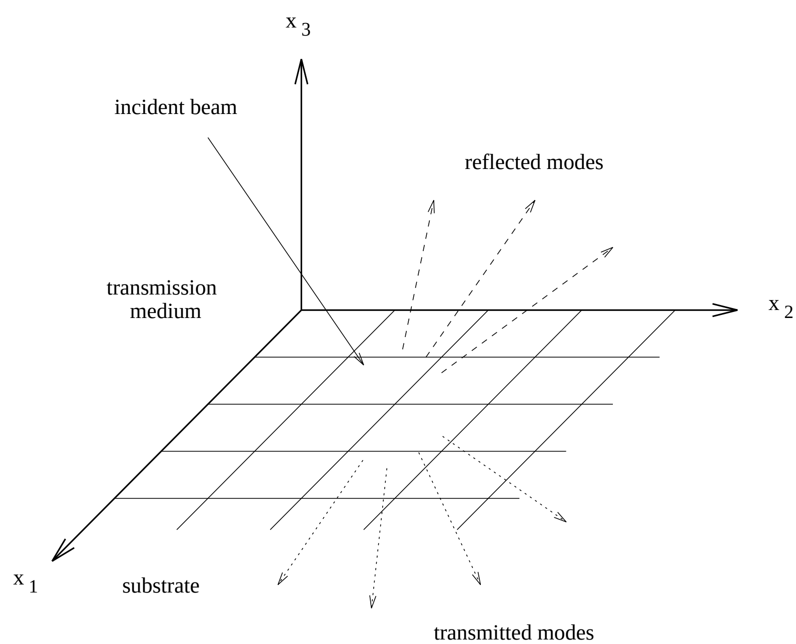

Figure 1: General geometrical configuration. The periodic pattern is assumed to extend infinitely in the (x_1, x_2) plane. Each cell contains some structure, defined by the dielectric coefficient $\epsilon(x_1, x_2, x_3)$.

to Kriegsmann [46], provided the infinite periodic problem can be solved. Thus the infinite periodic geometry represents both a practically useful limiting case, and a "stepping stone" to the analysis of finite structures.

Scattering from infinite periodic structures is a classical problem, dating back to Rayleigh, Floquet and Bloch. A fundamental feature of these problems is that in homogeneous regions, the field can be expanded as an infinite sum of plane waves. In its simplest form, say for a reflected field u of two variables with normal incidence and a 2π-periodic structure, this expansion can be written

$$u(x_1, x_3) = \sum_{n=-\infty}^{\infty} a_n e^{i(nx_1 + \beta_n x_3)}, \tag{2.1}$$

where the a_n are unknown coefficients. This is sometimes called the *Rayleigh expansion.* In the case of a medium with real refractive index $k > 0$, the coefficients β_n in the sum (2.1) are defined by

$$\beta_n = \begin{cases} \sqrt{k^2 - n^2} & \text{if } k \geq |n|, \\ i\sqrt{n^2 - k^2} & \text{if } k < |n|. \end{cases} \tag{2.2}$$

Since β_n is real for at most a *finite* number of indices n, we see from (2.1) that only a finite number of plane waves in the sum propagate into the far field, with the remaining *evanescent* modes decaying exponentially as $x_3 \to \infty$. The number of propagating modes and the direction of propagation for each mode is determined by the frequency of the incident wave, the refractive index of the material, and the period (cell dimension) of the structure. In engineering problems, the usual quantities of interest are the energies in each propagating mode as a proportion of the total incident energy. The ratio of the energy in a given mode to the incident energy is called the *efficiency* of that mode. A typical problem to be solved is: given a description of the structure and the incident beam, find the efficiency in each mode.

2.1 The time-harmonic Maxwell equations

A very accurate model for the underlying light propagation problem is provided by the classical time-harmonic Maxwell equations

$$\nabla \times E - i\omega\mu H = 0, \tag{2.3}$$

$$\nabla \times H + i\omega\epsilon E = 0. \tag{2.4}$$

Here E and H denote the electric and magnetic field vectors, respectively, μ is the magnetic permeability, ϵ is the dielectric coefficient, and ω is the frequency. Most optical structures are made from nonmagnetic materials, so μ is safely assumed to be constant. The physical structure is then completely described by the dielectric coefficient $\epsilon(x)$, $x = (x_1, x_2, x_3)$, which is in general biperiodic as indicated in Figure 1. Thus there are two constants L_1 and L_2, the axis lengths of each cell, such that

$$\epsilon(x_1 + n_1 L_1, x_2 + n_2 L_2, x_3) = \epsilon(x_1, x_2, x_3),$$

for all $x \in \mathbb{R}^3$, and all integers n_1, n_2. The structure is assumed to have finite depth, *i.e.*, there exists a number $b > 0$ such that

$$\begin{aligned} \epsilon(x) &= \epsilon_1, \quad \text{for } x_3 \geq b, \\ \epsilon(x) &= \epsilon_2, \quad \text{for } x_3 \leq -b. \end{aligned}$$

Here ϵ_1 and ϵ_2 are the (constant) dielectric coefficients of the transmission medium and the substrate, respectively. The transmission medium is assumed to be nonconductive, so ϵ_1 is real. The case $Im\ \epsilon_2 > 0$ accounts for a substrate which absorbs energy.

Assume that a plane wave $(E_I, H_I) = (s, p)e^{iq \cdot x}$ is incident on the structure from above. Here $q = (\alpha_1, \alpha_2, -\beta)$ is the incidence vector, and s, p are constant vectors that satisfy

$$s = \frac{1}{\omega\epsilon_1}(p \times q), \; q \cdot q = \omega^2 \epsilon_1 \mu, \; p \cdot q = 0.$$

Given q, one can reduce the original problem of solving Maxwell's equations in all space to a problem involving only one periodic cell, by considering so-called *quasiperiodic* solutions. Define $\alpha = (\alpha_1, \alpha_2, 0)$, and let

$$\begin{aligned} E_\alpha(x) &= e^{-i\alpha \cdot x} E(x), \\ H_\alpha(x) &= e^{-i\alpha \cdot x} H(x). \end{aligned}$$

Assuming that solutions to equations (2.3)–(2.4) are unique, the fields E_α, H_α are spatially periodic in (x_1, x_2), with the same periodicity as $\epsilon(x)$.

Substituting $E = e^{i\alpha\cdot x}E_\alpha$ and $H = e^{i\alpha\cdot x}H_\alpha$ into the Maxwell equations, one obtains a new set of equations for E_α, H_α,

$$\nabla_\alpha \times E_\alpha - i\omega\mu H_\alpha = 0, \tag{2.5}$$

$$\nabla_\alpha \times H_\alpha + i\omega\epsilon E_\alpha = 0, \tag{2.6}$$

where $\nabla_\alpha = \nabla + i\alpha$. Since all quantities in (2.5)–(2.6) are periodic in (x_1, x_2), the equations can be taken to hold on *one* fundamental cell, say $(0, L_1) \times (0, L_2) \times (-\infty, \infty)$, with periodic boundary conditions in the x_1 and x_2 directions. To complete the statement of the problem, a radiation condition is enforced, which states that all reflected and transmitted waves are outgoing and remain bounded as $|x_3| \to \infty$. In a weak sense, equations (2.5)–(2.6) are equivalent to the decoupled system

$$\nabla_\alpha \times (\frac{1}{\epsilon\mu}\nabla_\alpha \times H_\alpha) - \omega^2 H_\alpha = 0, \tag{2.7}$$

$$\nabla_\alpha \times H_\alpha + i\omega\epsilon E_\alpha = 0. \tag{2.8}$$

Let ϵ_c be a constant which satisfies $\inf Re\frac{1}{\epsilon(x)} \geq \frac{3}{4\epsilon_c}$. When considering the variational form of this problem over the Sobolev space $(\mathcal{H}^1)^3$, a natural coercivity condition leads to the problem

$$\nabla_\alpha \times (\frac{1}{\epsilon\mu}\nabla_\alpha \times H_\alpha) - \nabla_\alpha(\frac{1}{\epsilon_c\mu}\nabla_\alpha \cdot H_\alpha) - \omega^2 H_\alpha = 0, \tag{2.9}$$

$$\nabla_\alpha \times H_\alpha + i\omega\epsilon E_\alpha = 0, \tag{2.10}$$

as an alternative to problem (2.7)–(2.8). It can be shown [11, 28, 12] that the system (2.9)–(2.10) admits a unique weak solution, and that the scattering problems (2.7)–(2.8) and (2.9)–(2.10) are actually equivalent, except possibly at a discrete set of parameter values.

2.2 Variational formulation

In the variational formulation, the domain of the equations (2.5)–(2.6) is reduced from the periodic fundamental cell, which has infinite extent in the x_3-direction, to a bounded periodic "box" with finite extent in the x_3-direction. To do so requires the derivation of appropriate "transparent" boundary conditions; this is done by establishing artificial plane boundaries $\{x_3 = b\}$ and $\{x_3 = -b\}$ and matching the Green's function solution and Fourier expansion of the field on these boundaries. Doing so, one obtains nonlocal operators which map the traces of the field components on the artificial boundaries to their derivatives.

Since H_α is periodic in x_1 and x_2, we can expand H_α in a Fourier series:

$$H_\alpha(x) = \sum_{n\in Z^2} H_\alpha^{(n)}(x_3)e^{-i\alpha_n\cdot x}, \tag{2.11}$$

where $n = (n_1, n_2),\ Z = \{0, \pm 1, \pm 2, \ldots\}$,

$$H_\alpha^{(n)}(x_3) = \frac{1}{L_1 L_2} \int_0^{L_1} \int_0^{L_2} H_\alpha(x) e^{i\alpha_n \cdot x} dx_1 dx_2$$

and

$$\alpha_n = (2\pi n_1/L_1, 2\pi n_2/L_2, 0).$$

Let $\Omega_0 = \{x \in \mathbf{R}^3 : -b < x_3 < b\}$, $\Omega_1 = \{x \in \mathbf{R}^3 : x_3 > b\}$, $\Omega_2 = \{x \in \mathbf{R}^3 : x_3 < -b\}$. Denote

$$,_1 = \{x \in \mathbf{R}^3 : x_3 = b\}\ ,\quad ,_2 = \{x \in \mathbf{R}^3 : x_3 = -b\}.$$

Define for $j = 1, 2$ the coefficients

$$\beta_j^n(\alpha) = e^{i\gamma_j^n/2} \left|\omega^2 \epsilon_j \mu - |\alpha_n - \alpha|^2\right|^{1/2}, \quad n \in Z^2, \tag{2.12}$$

where

$$\gamma_j^n = arg(\omega^2 \epsilon_j \mu - |\alpha_n - \alpha|^2),\ \ 0 \le \gamma_j^n < 2\pi. \tag{2.13}$$

We assume that $\omega^2 \epsilon_j \mu \neq |\alpha_n - \alpha|^2$ for all $n \in Z^2$, $j = 1, 2$. This condition excludes "resonance". Note that β_j^n is real at most for finitely many n.

In particular, for real ϵ_2, we have the following equivalent form of (2.12)

$$\beta_j^n(\alpha) = \begin{cases} \sqrt{\omega^2 \epsilon_j \mu - |\alpha_n - \alpha|^2}, & \omega^2 \epsilon_j \mu > |\alpha_n - \alpha|^2\ , \\ i\sqrt{|\alpha_n - \alpha|^2 - \omega^2 \epsilon_j \mu}, & \omega^2 \epsilon_j \mu < |\alpha_n - \alpha|^2\ . \end{cases} \tag{2.14}$$

Observe that inside Ω_1 and Ω_2 the dielectric coefficients ϵ are constants, Maxwell's equations then become

$$(\Delta_\alpha + \omega^2 \epsilon_j \mu) H_\alpha = 0\ , \tag{2.15}$$

where $\Delta_\alpha = \Delta + 2i\alpha \cdot \nabla - |\alpha|^2$.

Since the medium in Ω_j is homogeneous ($\epsilon = \epsilon_j$), the method of separation of variables implies that H_α can be expressed as a sum of plane waves:

$$H_\alpha|_{\Omega_j} = \sum_{n \in Z^2} a_j^n e^{\pm i\beta_j^n(\alpha) x_3 - i\alpha_n \cdot x}, \quad j = 1, 2, \tag{2.16}$$

where the a_j^n are constant (complex) vectors.

Next, since β_j^n is real for at most finitely many n, there are only a finite number of propagating plane waves in the sum (2.16). The remaining waves are exponentially decayed (or unbounded) as $|x_3| \to \infty$. We impose a radiation condition, *i.e.*, H_α is composed of bounded outgoing plane waves in Ω_1 and Ω_2, plus the incident incoming wave in Ω_1.

From (2.11) and (2.16) we deduce

$$H_\alpha^{(n)}(x_3) = \begin{cases} H_\alpha^{(n)}(b) e^{i\beta_1^n(\alpha)(x_3 - b)}, & n \neq 0, \ \text{in } \Omega_1, \\ H_\alpha^{(0)}(b) e^{i\beta_1(x_3 - b)} + pe^{-i\beta_1 x_3} - pe^{i\beta_1(x_3 - 2b)}, & n = 0, \ \text{in } \Omega_1, \\ H_\alpha^{(n)}(-b) e^{-i\beta_2^n(\alpha)(x_3 + b)}, & \text{in } \Omega_2, \end{cases} \tag{2.17}$$

where p is the polarization vector associated with the incident wave H_0. From (2.17) we can then calculate the derivative of $H_\alpha^n(x_3)$ with respect to ν, the unit normal, on Ω_0:

$$\left.\frac{\partial H_\alpha^{(n)}}{\partial \nu}\right|_{\Gamma_j} = \begin{cases} i\beta_1^n(\alpha) H_\alpha^{(n)}(b), & n \neq 0, \quad \text{on} \ ,_1, \\ i\beta_1 H_\alpha^{(0)}(b) - 2i\beta_1 p e^{-i\beta_1 b}, & n = 0, \quad \text{on} \ ,_1, \\ i\beta_2^n(\alpha) H_\alpha^{(n)}(-b)\ , & \qquad\qquad \text{on} \ ,_2. \end{cases} \tag{2.18}$$

Thus from (2.16) and (2.18),

$$\left.\frac{\partial H_\alpha}{\partial \nu}\right|_{\Gamma_1} = \sum_{n\in Z} i\beta_1^n(\alpha) H_\alpha^{(n)}(b) e^{-i\alpha_n \cdot x} - 2i\beta_1 p e^{-i\beta_1 b}, \tag{2.19}$$

$$\left.\frac{\partial H_\alpha}{\partial \nu}\right|_{\Gamma_2} = \sum_{n\in Z} i\beta_2^n(\alpha) H_\alpha^{(n)}(-b) e^{-i\alpha_n \cdot x}\ , \tag{2.20}$$

where the unit vector $\nu = (0, 0, 1)$ on $,_1$ and $(0, 0, -1)$ on $,_2$.

For functions $f \in \mathcal{H}^{\frac{1}{2}}(,_j)^3$, define the operator T_j^α by

$$(T_j^\alpha f)(x_1, x_2) = \sum_{n\in\Lambda} i\beta_j^n(\alpha) f^{(n)} e^{-i\alpha_n \cdot x}, \tag{2.21}$$

where $f^{(n)} = \frac{1}{L_1 L_2}\int_0^{L_1}\int_0^{L_2} f(x) e^{i\alpha_n \cdot x}$, and equality is taken in the sense of distributions.

Similarly, one could derive the following tangential transparent boundary conditions:

$$\nu \times (\nabla_\alpha \times (H_\alpha - H_{I,\alpha})) = B_1(P(H_\alpha - H_{I,\alpha})) \ \text{ on } ,_1, \tag{2.22}$$

$$\nu \times (\nabla_\alpha \times H_\alpha) = B_2(P(H_\alpha)) \ \text{ on } ,_2, \tag{2.23}$$

where the operator B_j is defined by

$$B_j f = -i \sum_{n\in Z^2} \frac{1}{\beta_j^{(n)}} \{(\beta_j^{(n)})^2 (f_1^{(n)}, f_2^{(n)}, 0) + ((\alpha + \alpha_n)\cdot f^{(n)})(\alpha + \alpha_n)\} e^{i\alpha_n \cdot x}, \tag{2.24}$$

where P is the projection onto the plane orthogonal to ν, i.e.,

$$Pf = -\nu \times (\nu \times f).$$

Therefore, the scattering problem can be formulated as follows [12]:

$$\begin{aligned} \nabla_\alpha \times (\frac{1}{\epsilon}\nabla_\alpha \times H_\alpha) - \nabla_\alpha(\frac{1}{\epsilon_c}\nabla_\alpha \cdot H_\alpha) - \omega^2 H_\alpha &= 0 \ \text{ in } \Omega_0\ , \\ \nu \times (\nabla_\alpha \times (H_\alpha - H_{I,\alpha})) &= B_1(P(H_\alpha - H_{I,\alpha})) \ \text{ on } ,_1\ , \\ \nu \times (\nabla_\alpha \times H_\alpha) &= B_2(P(H_\alpha)) \ \text{ on } ,_2\ , \\ (T_1^\alpha - \frac{\partial}{\partial\nu}) H_{\alpha,3} &= 2i\beta_1 p_3 e^{-i\beta_1 b}\ , \ \text{ on } ,_1\ , \\ (T_2^\alpha - \frac{\partial}{\partial\nu}) H_{\alpha,3} &= 0\ , \ \text{ on } ,_2\ . \end{aligned}$$

It follows from some simple vector dentities and an integration by parts that the scattering problem has an equivalent variational form: find $H_\alpha \in \mathcal{H}^1(\Omega_0)^3$ such that

$$B(H_\alpha, \phi) = R(\phi) \ , \quad \text{for all } \phi \in \mathcal{H}^1(\Omega_0)^3, \tag{2.25}$$

where

$$\begin{aligned} B(H,F) \ &= \ \int_\Omega \frac{1}{\epsilon}\nabla_\alpha \times H \cdot \overline{\nabla_\alpha \times F} + \int_\Omega \frac{1}{\epsilon_c}\nabla_\alpha \cdot H \ \overline{\nabla_\alpha \cdot F} + \int_{\Gamma_1} \epsilon_1^{-1} B_1(P(H)) \cdot \overline{F} \\ &\quad + \int_{\Gamma_2} \epsilon_2^{-1} B_2(P(H)) \cdot \overline{F} - \int_{\partial\Omega} \epsilon_c^{-1} \nabla_{\alpha t} \cdot H \ \overline{\nu \cdot F} - \int_{\Gamma_1} \epsilon_c^{-1} T_1(H_3)\overline{F_3} \\ &\quad - \int_{\Gamma_2} \epsilon_c^{-1} T_2(H_3)\overline{F_3} - \omega^2 \int_\Omega H \cdot \overline{F} \end{aligned}$$

and

$$R(F) = \int_{\Gamma_1} (\nu \times \nabla_\alpha \times H_I - B_1 P(H_I)) \cdot \overline{F} + \int_{\Gamma_1} 2i\beta_1 \epsilon_c^{-1} p_3 e^{-i\beta_1 b} \overline{F_3} \ .$$

2.3 Simplifications in two-dimensional geometries

Substantial simplifications of the full Maxwell equations occur when the periodic structure is constant in one direction. This is the case for "classical" diffraction gratings and many other useful structures.

Let us assume that the underlying structure is constant in the x_2 direction, that is, $\epsilon(x_1, x_2, x_3) = \epsilon(x_1, x_3)$. There are three fundamentally different situations which arise, depending upon the direction and polarization of the incident wave:

1. *TE (transverse electric) polarization:* the incidence vector is orthogonal to the x_2-axis and the E field is parallel to x_2,

2. *TM (transverse magnetic) polarization:* the incidence vector is orthogonal to the x_2-axis and the H field is parallel to x_2,

3. *Conical diffraction:* the incidence vector is not orthogonal to x_2.

These situations are essentially listed in order of increasing difficulty. In the first case, Maxwell's equations reduce to a simple scalar Helmholtz equation

$$(\triangle_\alpha + k^2)u_\alpha = 0, \tag{2.26}$$

where $\triangle_\alpha = \partial_{x_1}^2 + \partial_{x_3}^2 + 2i\alpha_1\partial_{x_1} - |\alpha_1|^2$, $k^2 = \omega^2\epsilon\mu$, and u_α represents the component of the E_α field in the x_2 direction. In the second case, Maxwell's equations also reduce to a simple scalar model, this time of the form

$$\nabla_\alpha \cdot (\frac{1}{k^2}\nabla_\alpha u_\alpha) + u_\alpha = 0, \tag{2.27}$$

where u_α now represents the x_2 component of the H_α field. The regularity of solutions u_α of this model is reduced relative to the Helmholtz equation case (2.26). In both cases, transparent boundary conditions that take similar forms as (2.19)(2.20) may be derived. For the third case, the full vector equations generally need to be retained, but the problem can be solved over a "two-dimensional" domain.

In each of the three cases above, a variational formulation can be derived in a manner analogous to that for the full Maxwell equations described in Section 2.2.

2.4 Well-posedness

Here we give a brief summary of recent results and approaches dealing with issues of well-posedness of the grating problem. For Maxwell's equations in a biperiodic structure that separates two homogeneous materials and is piecewise C^2, the existence and uniqueness of the solutions were established in [32] by an integral equation approach. Using jump conditions, the authors reduced the problem to an equivalent system of integral equations and then applied Fredholm theory. In particular, it was shown that there exists a unique solution at all but a countable set of frequencies. The result generalized the earlier work [24] for 1D gratings, see also [47].

Another approach is based on the variational method. It has the advantage for dealing with extremely general diffractive structures and materials. The basic idea is to establish coercivity for the bilinear form of the variational formulation then apply the Lax-Milgram lemma and the Fredholm alternative. Existence and uniqueness results were proved in TE polarization [27], in TM polarization [10], and finally for biperiodic structures [11]. The general result may be stated as:

Theorem 2.1 *For all but a countable sequence of frequencies ω_j, $|\omega_j| \to +\infty$, the diffraction problem has a unique solution.*

Abboud and Nédélec [3] independently developed a variational formulation for Maxwell's equations in a nonperiodic bounded inhomogeneous medium. They were interested in the more general problem where the magnetic permeability is nonconstant. Their approach was further extended by Abboud to the periodic case in [1] and [2].

In general, the result in Theorem 2.1 is the best possible. There are examples which indeed exhibit the existence of singular frequencies (the sequence $\{\omega_j\}$). We refer to [19] where explicit examples are constructed and nonuniqueness is shown at the singular frequencies in the TM case. It was also shown that in general the sequence in Theorem 2.1 is unbounded. It is interesting to note that, in addition to Theorem 2.1, Abboud [1] showed that

Lemma 2.1 *If $Im\ \epsilon_1 > 0$ or $Im\ \epsilon_2 > 0$, then the diffraction problem has a unique solution.*

Thus, for media that are absorbing either in Ω_1 or Ω_2, the diffraction problem always has a unique solution.

Continuous dependence of solutions on material parameters was studied by Kirsch [43] in the two-dimensional TE case with gratings that separate a dielectric medium from a perfectly reflecting medium (conductor). The model equation takes the form

$$(\Delta + k^2)u = 0 , \tag{2.28}$$

$$u|_S = 0 \tag{2.29}$$

together with the radiation condition, where S is the grating profile. It was shown that the solution depends on k and the angle of incidence θ analytically, provided that $(\alpha_n + k\cos\theta)^2 \neq k^2$ for every integer n, where k^2 is real. However, the dependence of solutions on the grating profile turns out to be a bit more complicated. In fact, using a variational approach, it was shown [43] that

$$||u_f - u_g||_{H^1} \leq C||f - g||_{C^1} \tag{2.30}$$

where the constant C is independent of u_f, u_g and f, g. Here u_f and u_g are the solutions of (2.28)–(2.29) generated by grating profiles f and g, respectively, for a fixed incident wave.

3 Convergence of finite element methods

In this section we present some recent convergence results for the finite element method in solving the diffraction problem in various geometries and polarization modes.

For each $h \in (0, 1)$, the domain Ω_0 is discretized with a quasi-uniform mesh of size h. Let $\{S^h : h \in (0, 1)\}$ denote a family of finite dimensional subspaces of $\mathcal{H}^1(\Omega_0)^3$, for example the continuous piecewise linear functions.

We define the finite element approximation of the solution H_α of (2.25) by the following equation: for each $v_h \in S^h$,

$$B(H_h, v_h) = (f, v_h) . \tag{3.1}$$

Thus the problem is reduced to finite dimensions. Solving the resulting matrix equation gives rise to a finite element approximation of the solution.

3.1 Convergence analysis

We analyze two types of errors in the finite element approximation. We consider first the discretization of the continuous problem. The goal is to show that H_h, the solution to (3.1), is a good approximation to H_α. We then analyze truncations of nonlocal boundary operators T_j B_j $(j = 1, 2)$ given by (2.21)(2.24). The fact that these boundary operators are nonlocal follows immediately from the infinite series expansion. In practice, it is essential to obtain error estimates when truncations of these operators take place.

The following well-posedness results for the discretized problem and error estimates have been recently obtained in [11].

Theorem 3.1 *Suppose that (2.25) has a unique solution* $H_\alpha \in \mathcal{H}^1(\Omega_0)^3$ *for each* $f \in (\mathcal{H}^1(\Omega_0)^3)'$*. Then for any given* $\delta > 0$*, there exists* $h_0 = h_0(\delta)$ *such that for* $0 < h < h_0$*,*

$$||H_\alpha - H_h||_0 \leq \delta ||H||_1.$$

Moreover, if $f \in L^2(\Omega_0)^3$*, there exists an* $h_1 = h_1(\delta)$ *such that for all* $0 < h < h_1(\epsilon)$*,*

$$||H_\alpha - H_h||_1 \leq \delta ||f||_0 \; .$$

Since h_0, h_1 are independent of H_α, the estimates are uniform with respect to H_α.

It is also important to establish error estimates when the truncations of the nonlocal boundary operators T_j and B_j for $j = 1, 2$ are made. In [11], we show that the discretized variational problem is well-posed for N (the number of terms used to represent the boundary operators) sufficiently large and h sufficiently small. In other words, the convergence properties of Theorem 3.1 remain valid when N is sufficiently large.

Remark. In general, the convergence results may not be improved. This is essentially due to the fact that the solution $H_\alpha \in \mathcal{H}^1(\Omega_0)^3$ but $\notin \mathcal{H}^{1+\delta}(\Omega_0)^3$ for any $\delta > 0$, without assuming some additional regularity on ϵ.

In the two-dimensional geometries, the situation becomes more interesting. In the TE case, since discontinuous coefficients only occur in the lower order terms, the solution is in $\mathcal{H}^2$. It follows that improved error estimates are possible. In fact, it is shown in [8] that for appropriate approximating subspaces S^h,

$$\begin{aligned} ||u_\alpha - u_h||_1 \quad &\leq Ch \; , \\ ||u_\alpha - u_h||_0 \quad &\leq Ch^2 \; , \end{aligned}$$

where u_α solves the problem (2.26) and u_h is the finite element approximation of u_α. On the other hand, no improved regularity on the solution is available in the TM case. The singularities caused by the discontinuous coefficients in the principle part of the operator can spread more destructively. As a result, the solution is only in $\mathcal{H}^1$. Thus one can only expect similar uniform convergence results as the estimates in the above theorems [10]. Our numerical experiments support that although the finite element method does converge in the TM case, the convergence is slower than in the TE case.

3.2 Numerical experiments

Finite element methods based on the variational approach described above have been implemented. In our implementations, the domain Ω_0 is discretized with a uniform rectangular grid, and the fields are approximated with piecewise bilinear or trilinear elements in two or three dimensions, respectively. In the case of the full Maxwell equations, the solution space corresponding to the variational form (2.25) is simply $\mathcal{H}^1(\Omega_0)^3$, so there is no need to construct special vector finite element spaces as with $\mathcal{H}^1_{div}$ or $\mathcal{H}^1_{curl}$ formulations.

The key disadvantage of the nodal element approach is the possibility of "spurious modes" developing, particularly in problems involving high conductivity materials with

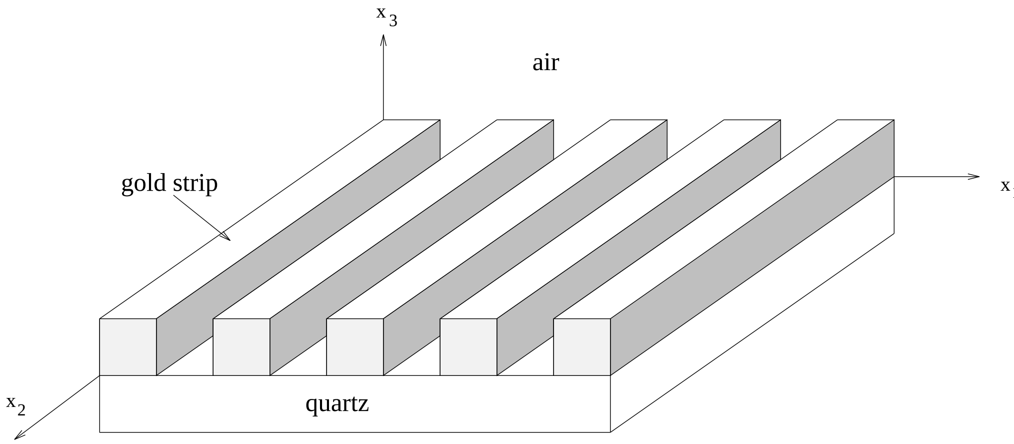

Figure 2: Prototypical diffraction grating. Gold strips are placed on quartz substrate. The structure is assumed to extend infinitely in the (x_1, x_2) plane.

corners. This phenomenon is well-known in the computational electromagnetics community. It is generally understood that "edge element" formulations eliminate the problem of spurious modes, and our own experiments are in agreement with that finding. The nodal approach does however have some advantages in terms of simplicity of implementation, and often gives good results. The question of edge versus nodal elements is not an issue of course in the cases of TE or TM polarization in two dimensions, since the problem is then scalar.

The boundary operators are calculated by truncating the Fourier series representations. The resulting scheme is consistent with the assumptions of the convergence estimates in Theorem 3.2. The discretized linear system is solved using the Orthomin iterative solver. Convergence is robust, but generally slow without preconditioning. Various preconditioners can be applied for greatly improved convergence. Development of effective preconditioners is a topic of current research interest.

These methods have been applied to various practical problems involving modeling diffractive optical structures. As a simple example of a prototypical problem, consider the structure illustrated in Figure 2. The grating period is $2.0\mu m$ and the height and cross-sectional width of each gold strip is $1.0\mu m$. Such structures can be fabricated with microlithographic mask-etch techniques. To study the scattering characteristics of this structure it is necessary to accurately solve Maxwell's equations for many different incoming incidence angles, polarizations, and wavelengths. For incidence vectors which do not lie in the (x_1, x_3) plane, the full vector equations must be solved on a two-dimensional domain (the conical diffraction case described in Section 2.3). Figure 3 shows a cross-section in the (x_1, x_3) plane of the real and imaginary parts of each component $(H_1, H_2, H_3) = H_\alpha$ of the magnetic field vector, for an incoming plane wave at a skewed 30 degree angle. The incoming wave is visible light (wavelength $0.46\mu m$), polarized out of the TM direction. The computation was done on a 84×85 grid and required several minutes on a Sun Sparcstation. In the figure, one can clearly see the penetration of

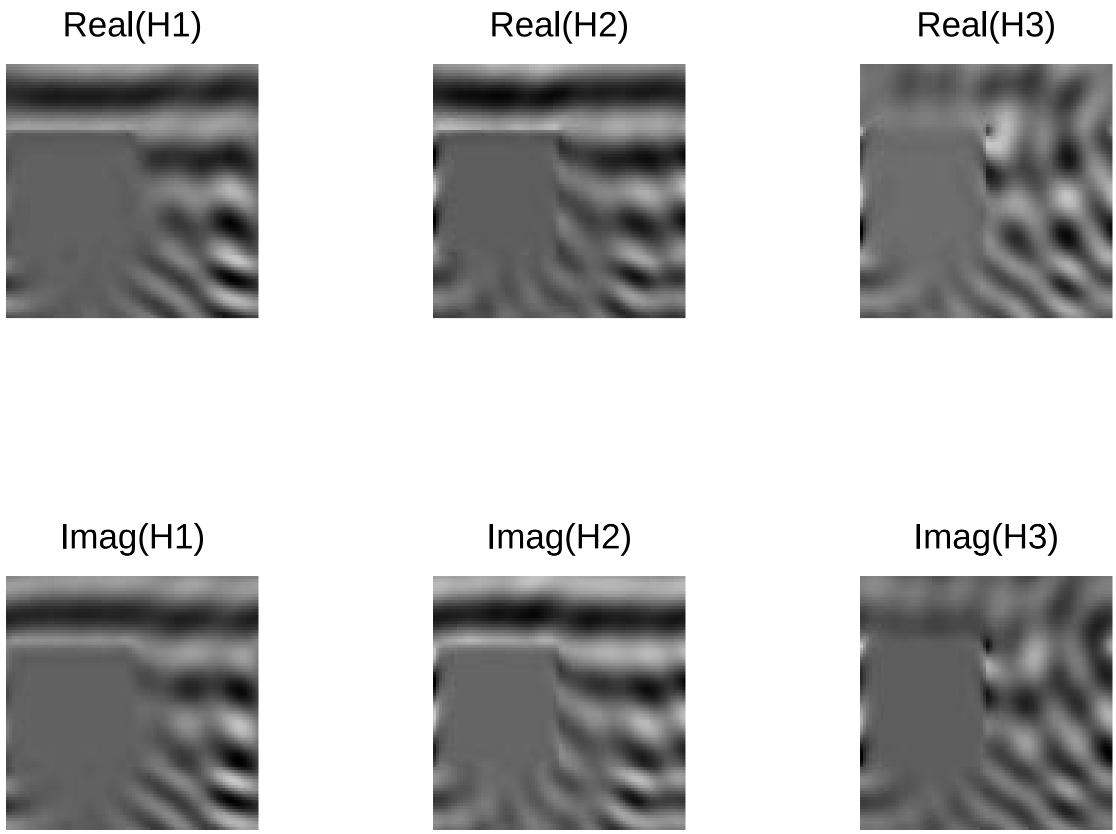

Figure 3: (x_1, x_3) cross-section of the H field through structure. Note the penetration of the field into the gold strip.

the field into the metal strip. The scattering behavior depends quite sensitively on the complex dielectric coefficient of the metal. Modeling the gold strip as a perfect conductor would be very inaccurate here.

Structures of practical interest are often substantially more complicated than the example presented above, involving detailed features and possibly several different materials such as films or coatings within each unit cell. In addition, fully three-dimensional geometries give rise to very large computational problems. We have successfully applied these methods to prototypical three-dimensional problems [28], but such computations are generally quite expensive. Significant improvements in efficiency can be expected with improved preconditioning techniques. The calculations described above used *no preconditioner* for the iterative solver. As the length scale of practical fabrication processes continues to decrease, the demand for more accurate computations on more complicated

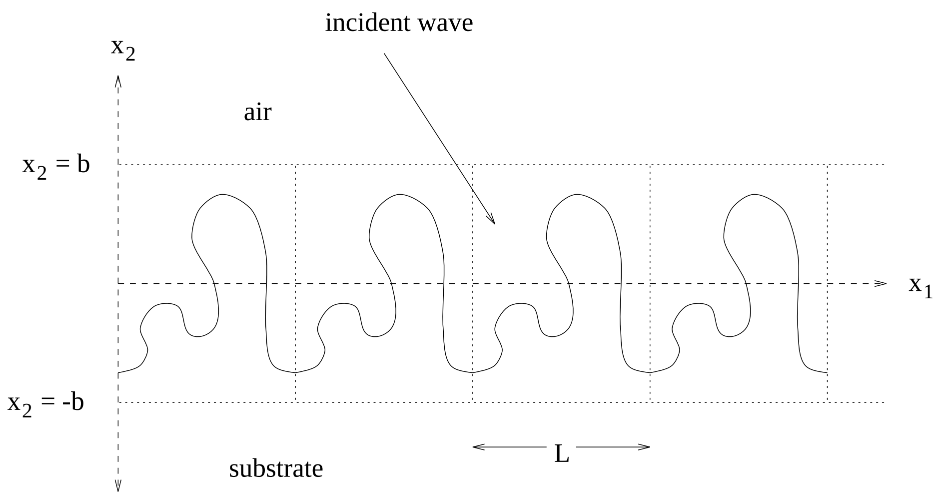

Figure 4: Geometrical configuration for inverse and optimal design problems.

devices will only increase.

4 The inverse problem

Given the incident field and the desired outgoing fields, the inverse problem concerns the determination of the grating profile. Closely related is the optimal design problem of determining a profile which optimizes some performance criteria.

4.1 Determination of a grating profile

We describe some recent progress in the study of inverse diffraction problems and optimal design problems. We discuss only 1D gratings, *i.e.*, surface profiles which separate two homogeneous materials and which are constant in the x_2 direction.

Consider a plane wave incident on a periodic structure from above as shown in Figure 4. The structure separates two regions. In one region, above the periodic structure, the dielectric coefficient ϵ is a fixed constant, so is the index of refraction k. The other region contains a perfectly reflecting material (or conductor). Given the incident field, an inverse diffraction problem is then to determine the periodic structure from the scattered field. Let the incident wave be of the form

$$u_I = e^{i\alpha x_1 - i\beta_1 x_3} \tag{4.1}$$

where $\alpha = k\sin\theta$, $\beta_1 = k\cos\theta$, and $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ is the angle of incidence. As discussed in Section 2.3, the direct model equation has a simple form

$$(\Delta + k^2)u = 0, \tag{4.2}$$

$$u|_S = 0, \tag{4.3}$$

where S is the grating profile. We again seek quasiperiodic solutions to this problem, *i.e.*, solutions u such that $ue^{-i\alpha x_1}$ are L-periodic for every x_3, here L is the period of the grating. Using the quasiperiodicity of the solution and the radiation condition that requires the boundedness of u as x_3 tends to infinity, we arrive at the boundary condition as in Section 2.2

$$\frac{\partial u}{\partial \nu}|_{x_3=b} = B(u|_{x_3=b}) - 2i\beta e^{-i\beta b + i\alpha x_1}, \tag{4.4}$$

where

$$B(f) = \sum_{n \in Z} i\beta_j^n f^{(n)} e^{i(\alpha_n + \alpha)x_1}, \tag{4.5}$$

and α_n and β_n are defined in Section 2.2. The inverse problem is to determine S from the information $u|_{x_3=b}$.

A closely related problem is to determine the grating structure on some nonconductive optical material. In that case, one places optical detectors both above and below the material. Consequently, the measurements consist of information on the reflected wave and transmitted wave. Note that the boundary condition (4.3) should be replaced with a nonlocal boundary condition that is to (4.4). This inverse problem was proposed and studied in [17].

4.2 Uniqueness and stability of the interface

By counting the dimensions of the unknowns and data, it is easy to see that the inverse problem is underdetermined. Thus, in general, properties on uniqueness and stability are very hard, if not impossible, to establish. But because of the important impact of these properties on applications, characterizations of uniqueness and stability are required.

Here, we present a uniqueness result for the inverse problem. Let us assume that for the given incident field u_I, u_1 and u_2 solve the direct problem (4.2-4.4) with respect to $S_1 = \{x_3 = f_1(x_1)\}$ and $S_2 = \{x_3 = f_2(x_1)\}$, respectively. The functions f_1 and f_2 are assumed to be sufficiently smooth, say C^2, and L-periodic. Let $b > \max\{f_1(x_1), f_2(x_1)\}$ be a fixed constant. Denote $T = \max\{f_1(x_1), f_2(x_1)\} - \min\{f_1(x_1), f_2(x_1)\}$.

Theorem 4.1 *Assume that $u_1(x_1, b) = u_2(x_1, b)$. Assume further that one of the following conditions is satisfied:*

(i) *k has a nonzero imaginary part;*

(ii) *k is real and T satisfies $k^2 < 2[T^{-2} + L^{-2}]$.*

Then $f_1(x_1) = f_2(x_1)$.

Thus in the case when k has nonzero imaginary part corresponding to a lossy medium, a global uniqueness result is available [7]. In the case with real k corresponding to a dielectric medium, one can only prove a local uniqueness result, *i.e.*, any two surface profiles are identical if they generate the same diffraction patterns and the area in between the two profiles are sufficiently small. Moreover, the smallness of the area is characterized explicitly in terms of a condition which relates the index of refraction k, the period, and the maximum of the difference in height allowed for the two profiles, see [9] for details.

Uniqueness for the inverse diffraction problem in periodic structure was also studied in [44] for a dielectric medium, where a uniqueness theorem was proved by a approach for the general inverse scattering problem. The main idea was to prove by using many incident waves the denseness of a set of special solutions. For the optical applications we are interested in, one is only allowed to use single or a small number of incident plane waves. See [41] for some additional results along this direction.

More recent results on uniqueness for the inverse diffraction problem in biperiodic structures may be found in [18] and [6].

In applications, it is impossible to make exact measurements. Stability is crucial in the practical reconstruction of profiles since it contains necessary information to determine to what extent the data can be trusted.

Before stating the stability result, let us first introduce some notations. For any two domains D_1 and D_2 in $\mathbf{R}^2$, define $d(D_1, D_2)$ the Hausdorff distance between them by

$$d(D_1, D_2) = \max\{\rho(D_1, D_2),\ \rho(D_2, D_1)\} \tag{4.6}$$

where

$$\rho(D_1, D_2) = \sup_{x\in D_1} \inf_{y\in D_2} |x - y| \,. \tag{4.7}$$

Denote $D = \{x;\ f(x_1) < x_3 < b\}$, and a sequence of domains $D_h = \{x;\ f(x_1) + h\sigma_h(x_1)\mu(x_1) < x_3 < b\}$ for any $0 < h < h_0$, where $\mu(x_1)$ is the normal to $S = \{x_3 = f(x_1)\}$. Assume also that the boundary $S_h = \{x_3 = f(x_1) + h\sigma_h(x_1)\mu(x_1)\}$ is periodic of the same period L. Further, the function σ_h satisfies $|\sigma_h(x_1)| \le C$. Furthermore, for h_0 sufficiently small, the sequence of domains is assumed to satisfy

$$C_1 h \le d(D, D_h) \le C_2 h, \tag{4.8}$$

where C_1 and C_2 are positive constants.

For the fixed incident plane wave u_I, assume that u and u_h solve the scattering problem with respect to periodic structures S and S_h, respectively. Then we have the following stability result.

Theorem 4.2 *Under the above assumptions,*

$$d(D_h, D) \le C||u_h|_{x_3=b} - u|_{x_3=b}||_{\mathcal{H}^{1/2}} \,,$$

where the constant C may depend on the family $\{\sigma_h\}$.

The result indicates that for small h, if the measurements are $O(h)$ close to the true scattered fields in the $\mathcal{H}^{\frac{1}{2}}$ norm, then D_h is $O(h)$ close to D in the Hausdorff distance. This result as well as stability results for other models were proved in [17].

4.3 Relaxed optimal design

The general form of the design problem we consider is: find a diffractive structure which generates a specified output for a given incident beam, or range of incident beams. The "output" usually consists of the far-field intensity pattern of the scattered field. Until recently, most approaches to this problem have depended upon approximations to the underlying PDE model in order to simplify the "profile-to-diffraction pattern" map [34, 39, 40].

In [27] and [29], the variational formulation of the full PDE (2.26) was employed to model the direct problem, and the design problem was solved by a relaxation approach. This approach is described next. The variational approach was introduced in [4, 5] for solving optimal photocell design problems; while the relaxation method for general optimal design problems was first proposed in [45].

We retain the same geometrical configuration described earlier, with a TE wave incident on a structure as shown in Figure 4. For convenience, from now on we assume the period L of the structure is 2π. The periodic curve $S \subset \Omega$ defines the grating profile. The material above S has refractive index k_1 and the material below S has index k_2. To make explicit the dependence of the refractive index on S, define

$$a_S(x) = \begin{cases} k_1^2 & \text{if } x \text{ is above } S, \\ k_2^2 & \text{if } x \text{ is below } S. \end{cases}$$

We then consider the problem

$$(\triangle_\alpha + a_S)u = 0 \quad \text{in } \Omega, \tag{4.9}$$

$$(T_1^\alpha - \frac{\partial}{\partial x_2})u = 2i\beta_1 e^{-i\beta_1 b} \quad \text{on } \Gamma_1, \tag{4.10}$$

$$(T_2^\alpha - \frac{\partial}{\partial x_2})u = 0 \quad \text{on } \Gamma_2. \tag{4.11}$$

With all parameters except the surface profile fixed, there are a fixed, finite number of propagating modes (each of which corresponds to an index n for which the propagation constant β_j^n is real-valued). Define two sets of indices of propagating modes

$$\Lambda_j = \{n \in Z : Im\ (\beta_j^n) = 0\}, \quad j = 1, 2.$$

The set Λ_1 contains the indices of the reflected propagating modes; Λ_2 corresponds to the transmitted modes. The coefficients of each propagating reflected mode are determined by the Fourier components of the trace $u|_{\Gamma_1}$:

$$\begin{aligned} r_n &= u_n(b)e^{-i\beta_1 b} && \text{for } n \neq 0, \quad n \in \Lambda_1, \\ r_0 &= u_0(b)e^{-i\beta_1 b} - const. && \text{for } n = 0, \end{aligned} \tag{4.12}$$

where $u_n(x_3) = \frac{1}{2\pi}\int_0^{2\pi} u(x_1, x_3)e^{-inx_1}dx_1$. Similarly, the coefficients of the propagating transmitted modes are

$$t_m = u_m(-b)e^{-i\beta_2 b} \quad \text{for } m \in \Lambda_2. \tag{4.13}$$

Writing the reflection and transmission coefficients as vectors

$$r = (r_n)_{n\in\Lambda_1}, \quad t = (t_m)_{m\in\Lambda_2},$$

we denote the pair $(r,t) = F$. The coefficients r_n and t_m, and hence F, are functions of the interface profile S. Denote this dependence by $F(a_S)$. Suppose that the desired "output vector" of diffraction coefficients is q, *i.e.* we wish to find S such that $F(a_S) = q$. One plausible way to formulate the optimal design problem is then

$$\min_{a_S\in\mathcal{A}} J(a_S) = \frac{1}{2}\|F(a_S) - q\|_2^2.$$

The choice of the admissible set of coefficients $\mathcal{A}$ is important. To achieve a well-posed optimization problem, there are two general routes. The first is to choose a relatively small admissible set, compact with respect to the topology induced by the map $J(a_S)$ and thus ensuring that the problem has a solution. This has the possible side-effect of introducing "artificial constraints", which could result in sub-optimal designs.

The other route is to begin with a large class of admissible curves, and "relax" the problem, enlarging the admissible set to include appropriate "mixtures" of materials. This can be accomplished as follows. Denote the set of all continuous simple curves contained in the domain Ω by $\mathcal{S}$. We allow any profile $S \in \mathcal{S}$ as admissible. The set of admissible refractive coefficients is then

$$\tilde{\mathcal{A}} = \{a_S : S \in \mathcal{S}\}.$$

We want to find the closure of $\tilde{\mathcal{A}}$ with respect to the functional $J(a_S)$. Consider

$$\mathcal{A} = \{a = k_2^2\gamma + k_1^2(1-\gamma) : \gamma \in L^\infty(\Omega), \quad 0 \le \gamma \le 1, \},$$

which could be described as the set of all mixtures of the two materials.

Under the assumption of low-frequency waves, it can be shown that problem (4.9)–(4.11) has weak solutions for any refractive index $a \in \mathcal{A}$. Furthermore, one can bound $\|u\|_{H^1(\Omega)}$ independent of the particular mixture $a \in \mathcal{A}$ [27]. We can then define for each mixture $a \in \mathcal{A}$ corresponding reflection and transmission vectors $r(a)$, $t(a)$. Using weak convergence arguments, it can be shown that for each $a \in \mathcal{A}$, there exists a sequence $a_n \in \tilde{\mathcal{A}}$ such that $r(a_n) \to r(a)$, and similarly for $t(a)$. In this sense, $\mathcal{A}$ is the closure of $\tilde{\mathcal{A}}$ with respect to $F(a)$.

We then arrive at the "relaxed" formulation of the design problem

$$\min_{a\in\mathcal{A}} J(a) = \frac{1}{2}\|F(a) - q\|_2^2. \tag{4.14}$$

Such problems are studied in [29], where numerical results are also presented.

One could of course generalize and specify a range of incidence angles, or a range of frequencies (or both). One such problem involving the design of antireflective structures was studied in [27], where existence of solutions to a minimization similar to (4.14) is proved, again under the assumption of low-frequency waves. Several examples of optimal antireflective structures are also illustrated.

4.4 Optimal interface shape design

In diffractive optics applications, the relaxed optimal design approach can generate structures which are prohibitively expensive or difficult to fabricate. One remedy is to use total variation constraints on the profile $a \in \mathcal{A}$ in the minimization in an effort to generate "simple" designs [29]. This approach can generate simple designs composed of homogeneous "blocks" of material. Still, this is not the most satisfactory approach in all cases, since the incorporation of intermediate-index materials into the structure can be costly.

The relaxed formulation was obtained by taking a relatively large class $\mathcal{S}$ of admissible interface profiles and passing to the limit as the profiles became oscillatory. Another natural route is to instead restrict the admissible interfaces S to those given by the graph of a bounded function $s(x_1)$. This suggests the problem

$$\min J(s) = \frac{1}{2}\|F(s) - q\|_2^2, \quad \text{subject to: } \|s\|_{L^\infty} \leq b' < b \tag{4.15}$$

where $F(s)$ is the diffraction pattern corresponding to the grating profile given by the graph of s. Without further constraints on s, one would not expect that a solution to (4.15) would exist in general, due to the possibility of oscillatory minimizing sequences. "Smoothness" constraints on s would not be appropriate, since mask-etch type manufacturing processes generally produce profiles with corners. Thus requiring that s is smooth eliminates manufacturable designs from the admissible class.

One convenient quantity which measures oscillations but allows corners is the total variation seminorm

$$TV(s) = \int |s'(x_1)|\, dx_1.$$

Thus we are led to consider the problem

$$\begin{aligned} \min J(s) \quad &= \quad \frac{1}{2}\|F(s) - q\|_2^2, \\ \text{subject to:} \quad & \quad TV(s) \leq M, \\ & \quad \|s\|_{L^\infty} \leq b' < b. \end{aligned} \tag{4.16}$$

The following result is proved in [30].

Theorem 4.3 *For sufficiently low-frequency incident waves, the constrained minimization problem (4.16) admits a solution $s \in BV$.*

A simple numerical approach to the shape design problem is to discretize the interface profile as the graph of a sum of step functions

$$s(x_1) = \sum_{j=1}^{N} s_j \chi_j \tag{4.17}$$

where χ_j is the indicator function in the interval $[(j-1)h, jh)$ and h is the cell width in the x_1 direction. Any s in the form (4.17) with $|s_j| \leq b'$ automatically satisfies $TV(s) \leq 2Nb'$, so one could proceed with the numerical minimization without explicitly enforcing a total variation constraint. Of course, for N large, minimizing sequences could still be "too oscillatory" to be useful in any practical design. In this case $TV(s) \leq M$ should be explicitly enforced.

Consider application of the gradient descent method to find local minima of problem (4.16). For now, assume that the total variation constraint is not enforced. The gradient of $J(s)$ can be found using an adjoint-state calculation as follows.

Let us view $J(s)$ as a map over some subset D of $L^2(0, 2\pi)$. Let δs be a small perturbation to a continuous "background" function s, and consider the linearized response $DJ(s)(\delta s)$ of $J(s)$ to δs. Formally, $DJ(s)(\delta s) = Re\{DF(s)(\delta s) \cdot \overline{(F(s)-q)}\}$, where $DF(s)(\delta s)$ denotes the linearization of F. The components of DF are

$$\begin{aligned} Dr_n(s)(\delta s) &= \frac{e^{-i\beta_1 b}}{2\pi} \int_{\Gamma_1} \delta u \, e^{-inx_1}, \\ Dt_m(s)(\delta s) &= \frac{e^{-i\beta_2 b}}{2\pi} \int_{\Gamma_2} \delta u \, e^{-imx_1}, \end{aligned}$$

where δu solves the linearized problem

$$\begin{aligned} (\triangle_\alpha + a_s)\delta u &= -\delta s \, (k_2^2 - k_1^2)\mu_s u \quad \text{in } \Omega, \\ (T_j - \frac{\partial}{\partial \nu})\delta u &= 0 \qquad \text{on } \Gamma_j, \quad j = 1, 2. \end{aligned}$$

Here μ_s is the measure defined by

$$\int_\Omega f \mu_s = \int_0^{2\pi} f(x_1, s(x_1)) \, dx_1.$$

for functions f on Ω. Taking the domain of DF as $L^2(0, 2\pi)$, the adjoint $DF^*(s)(\cdot)$ is defined by

$$DF(s)(\delta s) \cdot \bar{\delta q} = \int_0^{2\pi} \delta s \cdot \overline{DF^*(s)(\delta q)} \, dx_1,$$

for $\delta q = (\psi, \phi)$ with $\psi = (\psi_n)_{n \in \Lambda_1}$ and $\phi = (\phi_m)_{m \in \Lambda_2}$. Let $w \in H^1(\Omega)$ solve

$$(\triangle_\alpha + a_s)w = 0 \qquad \text{in } \Omega, \tag{4.18}$$

$$(T_1^* - \frac{\partial}{\partial \nu})w = -\frac{e^{i\beta_1 b}}{2\pi} \sum_{n \in \Lambda_1} \psi_n e^{inx_1} \qquad \text{on } \Gamma_1, \tag{4.19}$$

$$(T_2^* - \frac{\partial}{\partial \nu})w = -\frac{e^{i\beta_2 b}}{2\pi} \sum_{m \in \Lambda_2} \phi_m e^{imx_1} \quad \text{on } \Gamma_2, \tag{4.20}$$

where $T_j^* f = -\sum i\bar{\beta}_j^n f_n e^{inx}$. Notice that this adjoint problem for w represents waves propagating *into* Ω. With an integration by parts calculation, one finds that

$$DF(s)(\delta s) \cdot \overline{\delta q} = (k_2^2 - k_1^2) \int_0^{2\pi} \delta s(x_1) \, (\overline{w} u)(x_1, s(x_1)) \, dx_1$$

We then make the identification $\overline{DF^*(s)(\delta q)}(x_1) = (k_2^2 - k_1^2)\overline{w}u(x_1, s(x_1))$, and the gradient of $J(s)$ is given by $G(s) = Re\ \{DF^*(s)(F(s) - q)\}$, or

$$G(s)(x_1) = Re\{(k_2^2 - k_1^2)\overline{w}u\}(x_1, s(x_1)),$$

where w solves (4.18)–(4.20) with $(\psi, \phi) = F(s) - q$. Since both w and u are in H^2 and hence have continuous representatives, $G(s)(x_1)$ is well-defined pointwise. In [30], the differentiability of F is proved under mild assumptions.

A gradient descent step defined by $s_1 = s_0 - tG(s_0)$, $t > 0$ does not necessarily lie in the computational domain Ω. Thus we define a projection operator P by

$$(Pf)(x_1) = \begin{cases} \min\{f(x_1), b'\} & \text{if } f(x_1) \geq 0, \\ \max\{f(x_1), -b'\} & \text{if } f(x_1) < 0, \end{cases}$$

where $b' < b$.

Straightforward gradient descent would then proceed as follows:

1. Choose an initial guess s_0.
2. For $j = 0, \ldots,$convergence, set $s_{j+1} = P(s_j - t_j G(s_j))$ for a suitably chosen step parameter t_j.

In practice, this algorithm is slow but generally effective. Techniques to improve the efficiency of the basic algorithm have been developed for the relaxed design problem [29]. The general idea is to take advantage of the underlying PDE model by viewing it as a constraint and performing inexact solves following infeasible point techniques from constrained optimization. The same ideas can be applied to the shape optimization problem. Further details, as well as a description of the incorporation of the TV constraint into the algorithm can be found in [30].

One useful modification of the problem is to allow the specification of "intensity only" far-field data. Often in applications, the phase of the diffracted orders is irrelevant. Specifying them arbitrarily, as required by the formulation above, may preclude a viable solution. An obvious remedy is to reformulate the cost functional as

$$J(s) = \frac{1}{4} \sum \left(|F_j(s)|^2 - |q_j|^2 \right)^2,$$

where the F_j represent the individual diffracted modes. The general approach described above is easily adapted to this new cost functional.

Numerical experiments in which optimal shapes are obtained for several different diffractive structures can be found in [30, 31]

4.5 A sampling of other approaches

An approach to optimal design of gratings also based on the variational formulation of the problem, but restricted to the case of binary structures, has been formulated by

Elschner and Schmidt [33]. By considering only binary structures, the authors obtain a simpler, finite dimensional minimization problem which avoids many of the technical issues encountered in the more general approach described above.

Simulated annealing and simulated quenching algorithms have been applied by Prather [49] for the optimization of diffractive optical elements of finite extent and with nonperiodic geometry, using boundary element and hybrid finite element–boundary element methods to solve the underlying diffraction problem. Very good final designs were obtained. One advantage of these approaches is that no derivative information on the cost functional is required, so that implementation is relatively simple. However, simulated annealing type algorithms are well known to be very computationally intensive, so that rapid solution of the direct problem becomes even more important than in methods based on derivatives.

A very interesting grating profile reconstruction (or design) method based on the method of variation of boundaries (MVB) has been recently introduced by Ito and Reitich [42]. This method uses a minimization approach, but uses the analytic continuation ideas inherent in MVB to enable a global line search at each step in the minimization, thereby overcoming possible problems with nonconvexity in the cost functional.

4.6 Conclusion

To summarize this section, inverse and optimal design problems in diffractive optics is an interesting and very active area of current research. Methods of practical use for industrial problems have already been developed. New methods are currently being introduced and refined. The area is by no means mature and it is thus difficult at this stage to summarize all important work or to make meaningful comparisons between methods. However the importance of these techniques is large and immediate. The area will continue to grow as methods are adopted by more industrial researchers.

5 Second harmonic generation in gratings

In addition to the linear structures described in the previous sections, variational techniques can be productively applied to study certain nonlinear devices. Here we describe some recent work on the phenomenon of "grating enhanced second-harmonic generation". In physical experiments, a nonlinear optical material is either applied over a (usually metallic) diffraction grating, or a grating structure is etched into the surface of a nonlinear optical crystal. Materials with relatively large second-order nonlinear susceptibility tensor (eg. GaAs, GaSe, InSe) are used. When an intense "pump" beam is applied, a second-harmonic field (at twice the pump frequency) is generated. Thus for example, coherent blue light can be generated from a red pump laser. In most cases, the nonlinear susceptibility is very small, and thus the intensity of the second-harmonic field is weak. It has been found experimentally that the grating structure can significantly enhance the second-harmonic conversion efficiency [50]. The enhancement occurs when the grating is operated near a resonant point, for example, when a diffracted mode is directed parallel to the structure. Presumably the enhancement is due to increased pump field intensity within the nonlinear material.

Previous approaches to this problem relied on either the "undepleted pump approximation", which linearizes the problem and tends to lose accuracy as conversion efficiency increases [21] or the "slowly varying approximation", which tends to be inaccurate in the small structures we consider. Without making these approximations, in the simplest two-dimensional case the model equations take the form

$$\left[\triangle + (\omega k_1)^2\right] u = \chi_1 u^* v, \tag{5.21}$$

$$\left[\triangle + (2\omega k_2)^2\right] v = \chi_2 u^2, \tag{5.22}$$

where u and v are the "pump" and second-harmonic fields, respectively, and χ_1 and χ_2 are components of the second-order nonlinear susceptibility tensors. Using an extension of the approach described in Section 2 this system can be cast in variational form, with transparent boundary conditions similar to the linear case. Estimates involving Schauder's lemma combined with a contraction mapping argument yield the following result.

Theorem 5.1 *Given a fixed nonresonant geometry and fixed incident wave, there exists a constant $\epsilon > 0$, such that if $||\chi_1||_{L^\infty}||\chi_2||_{L^\infty} \leq \epsilon$ then the problem (5.21–5.22), supplemented with appropriate radiation conditions, admits a unique solution $(u, v) \in H^1(\Omega)^2$.*

Thus for small nonlinear susceptibilities (a very physical condition), the problem has a unique solution. This was established first for the case of nonlinear films in [14], then for the case of grating structures in [15].

In the same papers, a numerical approach was employed which combines finite elements with a fixed-point iteration, and several numerical experiments were carried out.

The model described above is valid only for certain types of crystal structures in the nonlinear medium. If the crystal is such that the second harmonic field is generated in TM polarization, or is only excited by TM pump beams, then the model is more complicated and proof of results like Theorem 5.1 is more difficult. Additional complications occur for other polarization modes, and for parameter values near a resonant point. In fact, each of these complications represent cases of great practical interest, see [13] for some recent results. Improved analytical and computational techniques are currently being developed with the hope of tackling these difficulties.

Acknowledgments

Most of this work is a direct consequence of the authors' collaboration with Dr. J. Allen Cox of the Honeywell Technology Center, initiated at the Institute for Mathematics and its Applications at the University of Minnesota. The first author was partially supported by NSF grant number DMS 95-01099 and the NSF University-Industry Cooperative Research Programs grant DMS 97-05139. The second author was supported by AFOSR. Acknowledgement and Disclaimer: Effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0497. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

References

[1] T. Abboud, *Etude mathématique et numérique de quelques problémes de diffraction d'ondes électromagnétiques*, PhD thesis, Ecole Polytechnique, Palaiseau, 1991.

[2] T. Abboud, *Electromagnetic waves in periodic media*, in *Second International Conference on Mathematical and Numerical Aspects of Wave Propagation*, edited by R. Kleinman, T. Angell, D. Colton, F. Santosa, and I. Stakgold, SIAM, Philadelphia, 1–9 (1993).

[3] T. Abboud and J.C. Nédélec, *Electromagnetic waves in an inhomogeneous medium*, J. Math. Anal. Appl., 164 (1992), 40–58.

[4] Y. Achdou, *Numerical optimization of a photocell*, Opt. Comput. Meth. Appl. Mech. Eng. **102**, 89–106 (1993).

[5] Y. Achdou and O. Pironneau, *Optimization of a photocell*, Optimal Control Appl. Meth. **12**, 221–246 (1991).

[6] H. Ammari, *Uniqueness theorems for an inverse problem in a doubly periodic structure*, C. R. Acad. Sci. Paris, t. 320, Série I (1995), 301–306.

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

[8] G. Bao, *Finite elements approximation of time harmonic waves in periodic structures*, SIAM J. Numer. Anal. **32** (1995), 1155–1169.

[9] G. Bao, *An inverse diffraction problem in periodic structures*, Third International Conference on Mathematical and Numerical Aspects of Wave Propagation, ed. by G. Cohen, SIAM, Philadelphia, 694–704, 1995.

[10] G. Bao, *Numerical analysis of diffraction by periodic structures: TM polarization*, Numerische Mathematik, Numer. Math. **75** (1996), 1–16.

[11] G. Bao, *Variational approximation of Maxwell's equations in biperiodic structures*, SIAM J. Appl. Math., **57** (1997), 364–381.

[12] G. Bao, *On the scattering by a biperiodic structure*, preprint.

[13] G. Bao and Y. Chen, *A nonlinear grating problem in diffractive optics*, SIAM J. Math. Anal., Vol. 28, No. 2 (1997), 322–337.

[14] G. Bao and D. Dobson, *Second harmonic generation in nonlinear optical films*, J. Math. Phys. **35** (1994), 1622-1633.

[15] G. Bao and D. Dobson, *Diffractive optics in nonlinear media with periodic structure*, Euro. J. Appl. Math., **6** (1995), 573–590.

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

[17] G. Bao and A. Friedman, *Inverse problems for scattering by periodic structures*, Arch. Rat. Mech. Anal. **132** (1995), 49–72.

[18] G. Bao and Z. Zhou, *Inverse diffraction by a doubly periodic structure*, C.R. Acad. Sci. Paris, t. 324, Série I, (1997), 627–632.

[19] A. 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.

[20] M. Born and E. Wolf, *Principles of Optics,* sixth edition, Pergamon Press, Oxford, 1980.

[21] E. Bringuier and A. Bourdon, *Optical second-harmonic generation in lossy media: Application to GaSe and InSe,* Phys. Rev. B, to appear.

[22] O. Bruno and F. Reitich, *Numerical solution of diffraction problems: a method of variation of boundaries; II. Dielectric gratings, Pade approximants and singularities; III. Doubly-periodic gratings*, J. Opt. Soc. Amer. A., 10 (1993), 1168–1175, 2307–2317, 2551–2562.

[23] O. Bruno and F. Reitich, *Accurate calculation of diffractive grating efficiencies*, in "Mathematics in Smart Structures", edited by H. T. Banks, SPIE Proc. 1919 (1993), 236–247.

[24] X. Chen and A. Friedman, *Maxwell's equations in a periodic structure*, Trans. Amer. Math. Soc., **323** (1991), 465–507.

[25] J. A. Cox, *Inverse and optimal design problems for imaging and diffractive optical systems*, in Inverse Problems and Optimal Design in Industry, ed. H. Engl and J. McLaughlin, B.G. Teubner, Stuttgart, in press.

[26] J. A. Cox and D. Dobson, *Mathematical modeling for diffractive optics*, in "Diffractive and Miniaturized Optics (Critical Reviews)" Sing Lee, ed., SPIE CR-49 (1994), 32–53.

[27] D. Dobson, *Optimal design of periodic antireflective structures for the Helmholtz equation*, Euro. J. Appl. Math., **4** (1993), 321–340.

[28] D. Dobson, *A variational method for electromagnetic diffraction in biperiodic structures*, RAIRO Model. Math. Anal. Numér. **28** (1994), 419–439.

[29] D. Dobson, *Exploiting ill-posedness in the design of diffractive optical structures,* in "Mathematics in Smart Structures", edited by H. T. Banks, SPIE Proc. 1919 (1993), 248–257.

[30] D. Dobson, *Optimal shape design of blazed diffraction gratings,* Appl. Math. Opt., to appear.

[31] D. Dobson, *Controlled scattering of light waves: optimal design of diffractive optics*, in "Control Problems in Industry", edited by I. Lasiecka and B. Morton, Birkhauser, Boston (1995), 17–118.

[32] D. Dobson and A. Friedman, *The time-harmonic Maxwell equations in a doubly periodic structure,* J. Math. Anal. Appl., **166** (1992), 507–528.

[33] J. Elschner and G. Schmidt, *Analysis and numerics for the optimal design of binary diffractive gratings*, Weierstrass Institute for Applied Analysis and Stochastics, Berlin, preprint No. 323 (1997).

[34] M. W. Farn, *New iterative algorithm for the design of phase-only gratings*, in Computer and Optically Generated Holographic Optics, ed. I.N. Cindrich and S. Lee, Proc. SPIE 1555 (1991), 34–42.

[35] A. Friedman, *Mathematics in Industrial Problems*, IMA Volume 16, Springer–Verlag, New York (1988).

[36] A. Friedman, *Mathematics in Industrial Problems, Part 3*, IMA Volume 38, Springer–Verlag, New York (1991).

[37] A. Friedman, *Mathematics in Industrial Problems, Part 7*, Springer–Verlag, New York (1994).

[38] T. K. Gaylord and M. G. Moharam, *Analysis and applications of optical diffraction by gratings*, IEEE Proceedings, Vol. 73 No. 5 (1985), 894-937.

[39] R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of phase from image and diffraction plane pictures*, Optik (1972), 237–246.

[40] Y. Han and C.A. Delisle, *Exact surface relief profile of kinoform lenses from a given phase function,* Proc. SPIE **2152** (1994).

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

[42] K. Ito and F. Reitich, *A high-order perturbation approach to profile reconstruction. I: perfectly conducting gratings*, University of Minnesota, Institute for Mathematics and its Applications preprint series, no. 1518, (1997).

[43] A. Kirsch, *Diffraction by periodic structures*, Proc. of the Lapland Conf. on Inverse Problems, edited by L. Pävarinta and E. Somersalo (1993), Springer-Verlag, 87–102.

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

[45] R. Kohn and G. Strang, *Optimal design and relaxation of variational problems I, II, III*, Comm. Pure Appl. Math. **39**(1986), 113–137, 139–182, 353–377.

[46] G.A. Kriegsmann, *Scattering by acoustically large corrugated planar surfaces*, J. Acoust. Soc. Am. **88** (1990), 492–495.

[47] J. C. Nédélec and F. Starling, *Integral equation methods in a quasi-periodic diffraction problem for the time-harmonic Maxwell's equations,* SIAM J. Math. Anal., **22** (1991), 1679–1701.

[48] R. Petit, ed., *Electromagnetic Theory of Gratings,* Topics in Current Physics, Vol. 22, Springer-Verlag, Heidelberg, 1980.

[49] D. W. Prather, *Analysis and Synthesis of Finite Aperiodic Diffractive Optical Elements Using Rigorous Electromagnetic Models*, Ph.D. Thesis, Department of Electrical Engineering, University of Maryland (1997).

[50] R. Reinisch, M. Nevière, H. Akhouayri, J. Coutaz, D. Maystre and E. Pic, *Grating enhanced second harmonic generation through electromagnetic resonances,* Opt. Eng. **27** (1988), 961–971.