Numerical performance of finite-difference modal methods for the electromagnetic analysis of one-dimensional lamellar gratings
Philippe Lalanne, Jean-Paul Hugonin

To cite this version:

HAL Id: hal-00867708
https://hal-iogs.archives-ouvertes.fr/hal-00867708
Submitted on 30 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Numerical performance of finite-difference modal methods for the electromagnetic analysis of one-dimensional lamellar gratings

Philippe Lalanne and Jean-Paul Hugonin
Laboratoire Charles Fabry de l’Institut d’Optique Théorique et Appliquée, Centre National de la Recherche Scientifique, B.P. 147, 91403 Orsay Cedex, France

Received October 7, 1999; accepted January 10, 2000; revised manuscript received February 16, 2000

The numerical performance of a finite-difference modal method for the analysis of one-dimensional lamellar gratings in a classical mounting is studied. The method is simple and relies on first-order finite difference in the grating to solve the Maxwell differential equations. The finite-difference scheme incorporates three features that accelerate the convergence performance of the method: (1) The discrete permittivity is interpolated at the lamellar boundaries, (2) mesh points are located on the permittivity discontinuities, and (3) a nonuniform sampling with increased resolution is performed near the discontinuities. Although the performance achieved with the present method remains inferior to that achieved with up-to-date grating theories such as rigorous coupled-wave analysis with adaptive spatial resolution, it is found that the present method offers rather good performance for metallic gratings operating in the visible and near-infrared regions of the spectrum, especially for TM polarization. © 2000 Optical Society of America [S0740-3232(00)00606-2]


1. INTRODUCTION

Many rigorous methods exist for analyzing the diffraction by surface-relief gratings, and much work has been done during recent years to improve and generalize rigorous methods. The differential method was associated with the $R$-matrix algorithm to improve its stability for highly conductive metallic gratings, especially for TM polarization (magnetic field vector perpendicular to the grating vector); the integral method was generalized to study diffraction by echelles covered with dielectric layers. The rigorous coupled-wave analysis (RCWA) was improved in a similar manner, and recently its convergence for TM polarization and conical mounts was substantially improved. The method of coordinate transformation was extended to gratings with vertical facets. Finite-difference (FD) methods for solving partial differential equations are also widely used in electromagnetism for solving Maxwell’s equations. The finite-difference time domain (FDTD) method is one example used extensively. FD or finite-element methods are not widely used in grating theory but do require a two-dimensional (2D) mesh for the discretization of a 1D grating. The present method, in contrast, is time harmonic and uses a 1D mesh for the discretization of 1D gratings. It is therefore similar to numerical techniques that are based on FD modal approaches and used in waveguide theories.

The FD modal method is described in Section 2 with special attention devoted to the TM polarization case. The TE polarization case is briefly reported. In Section 3 the convergence rate of the present method is studied for metallic and dielectric lamellar gratings for TE and TM polarizations. It is compared with that achieved by the classical RCWA and by the enhanced version of the RCWA recently reported. Limitations of the present method for analyzing dielectric gratings are emphasized. Section 4 contains several comments on the present method. Section 5 concludes the paper.

2. FINITE-DIFFERENCE MODAL METHOD

Let us consider a 1D lamellar grating along the $x$ axis with a relative permittivity profile $\varepsilon(x)$; see Fig. 1. The $z$ axis is perpendicular to the grating boundaries. The diffraction problem is invariant in the $y$ direction. Magnetic effects are not considered in this paper, and the constant $\mu_0$ denotes the permeability of the periodic structure. $\varepsilon_0$ is the permittivity of the vacuum. The grating period is denoted by $\Lambda$, and the modulus $K$ of the grating vector is equal to $2\pi/\Lambda$. An incident plane wave of frequency $\omega$ and wavelength $\lambda$ in the vacuum makes an angle $\theta$ with the $z$ direction in a nonconical mounting.
Fig. 1. Parameter definition for the classical grating diffraction problems considered in this paper.

We denote the modulus of the wave vector of the incident wave by \( k_0 \), where \( k_0 = 2 \pi / \lambda \) in the vacuum. A temporal dependence in \( e^{j \omega t} \) of the wave is assumed. In the following, the FD modal method is first described for TM polarization. The TE polarization case is then briefly outlined.

**A. TM Polarization**

The incident normalized magnetic field is given by

\[
\Psi_{\text{inc}} = \exp[-j k_0 n_1 \sin(\theta) x + \cos(\theta) z)].
\]  

(1)

The Rayleigh expansions for the magnetic field in the incident medium and in the substrate are given by

\[
\Psi_1 = \Psi_{\text{inc}} + \sum_i R_i \exp[-j(k_x i x - k_{1,zi} z)],
\]

(2)

\[
\Psi_3 = \sum_i T_i \exp[-j(k_x i x + k_{3,zi} (z - h))],
\]

(3)

respectively. In Eqs. (2) and (3), \( R_i \) and \( T_i \) are the backward- and forward-diffracted amplitudes, \( k_{1,zi} \) is equal to \( k_0 n_1 \sin \theta - ik_x \), and \( k_{p,zi} \) (\( p = 1 \) or \( 3 \)) is defined by \( k_{p,zi}^2 + k_{p,zi}^2 = n_0^2 k_0^2 \), with \( \text{Re}(k_{p,zi}) + \text{Im}(k_{p,zi}) > 0 \). The magnetic field \( \Psi(x) \) in the grating region satisfies the Helmholtz equation

\[
\frac{\partial^2 \Psi}{\partial z^2} + \frac{\partial}{\partial x} \left( \frac{1}{\epsilon} \frac{\partial \Psi}{\partial x} \right) + k_0^2 \epsilon \Psi = 0.
\]  

(4)

To solve this differential equation, we use a FD scheme in the \( x \) direction and then calculate analytically the propagation in the \( z \) direction. The first step in applying a FD method is to select a discrete set of values of \( x \) (the discrete points) inside one grating period, the \( x_i \)'s, \( i = 1, ..., N \), represented with crosses in Fig. 2. The magnetic field becomes a function of the discrete index \( i \) and can be noted as a vector \( \Psi \) with \( N \) components, the \( \Psi'_i \)'s in Fig. 2. The FD expression for the first derivative along the \( x \) direction is computed at the discrete locations \( x_i' \). The \( x_i' \)'s are represented by circles in Fig. 2. An interlaced grid similar to the Yee's space lattice\(^{25}\) used in FDTD methods is thus considered. In its discretized form, Eq. (4) is written as

\[
\frac{\partial^2 \Psi}{\partial z^2} + E \Psi = 0,
\]  

(5)

where

\[
E = \epsilon_1 (D_2 \epsilon_2^{-1} D_1 + k_0^2 I).
\]  

(6)

In Eq. (6) the \( N \times N \) matrix \( E \) is tridiagonal with nonnull upper right and lower left coefficients, \( I \) is the identity matrix, and \( \epsilon_1 \) and \( \epsilon_2 \) are two diagonal matrices that result from the relative permittivity profile. Because this profile is piecewise constant, a naive discretization of the relative permittivity by simply assigning to the \( \epsilon_1 \) and \( \epsilon_2 \) coefficients the values of the relative permittivity at the discrete points is a poor representation along the boundaries inside the grating region. Instead, we use an interpolation scheme that locally averages the permittivity or its inverse according to the following rules:

\[
\epsilon_1(i, i) = \frac{1}{\langle 1/\epsilon(x) \rangle_{[x_{i-1}, x_i]}}
\]

(7a)

\[
\epsilon_2(i, i) = \langle \epsilon(x) \rangle_{[x_{i-1}, x_i]}
\]

(7b)

In Eqs. (7a) and (7b) the brackets indicate the intervals over which the averaging has to be performed. The use of this interpolation scheme has a drastic impact on the convergence performance of the present method and is justified in Appendix A. The differential operator \( D_1 \) and \( D_2 \) are given by

\[
D_1 = \begin{bmatrix}
\epsilon_1^{-1}(1) & & & & & \\
& \epsilon_1^{-1}(2) & & & & \\
& & \ddots & & & \\
& & & \epsilon_1^{-1}(N) & & \\
\end{bmatrix}
\]  

(8a)

\[
D_2 = \begin{bmatrix}
1 & & & & & \\
-\alpha & 1 & & & & \\
& -1 & 1 & & & \\
& & \ddots & & & \\
& & & 1 & -\alpha & \\
& & & & -1 & 1
\end{bmatrix}
\]  

(8b)
The substrate at the mesh points. We rather prefer (see electromagnetic field in the incident medium and that in the grating region: \( \Psi(x + \Lambda, z) = \Psi(x, z) \exp(-jk_0n_1 \sin \theta) \). The vectors \( h_1 \) and \( h_2 \) are defined by \( h_1(i) = x_{i+1} - x_i \) and \( h_2(i) = x_i - x_{i-1} \).

The magnetic field \( \Psi \) in the grating region is given by

\[
\Psi(z) = \sum_{m=1}^{N} w_m \{ c_m^+ \exp(-\lambda_h^m z) + c_m^{-} \exp[\lambda_m(z-h)] \},
\]

(9)

where \( W_m \) and \( \lambda_m \) are the vector of the eigenvector matrix \( W \) and the positive square root of the eigenvalues of the matrix \( E \), respectively. The \( c_m^+ \) and \( c_m^- \) coefficients, like the \( R_i \) and \( T_i \), coefficients, are unknowns.

All the unknown coefficients have to be evaluated by matching the tangential field components \( \Psi \) and \( \xi \) at the grating interfaces, \( z = 0 \) and \( z = h \). For this purpose, we first note that the fields in the incident medium and in the substrate are expressed in the Rayleigh basis, whereas in the grating region, the field is just known at some discrete locations, the mesh points \( x_i \)'s. A natural choice here would be to discretize the electromagnetic field in the incident medium and that in the substrate at the mesh points. We rather prefer (see Section 4) to interpret the discrete \( \Psi_i \)'s values computed at the \( x_i \)'s points in the grating region as the coefficients of the magnetic field expansion in a set of rectangle functions. For this purpose, we define the function \( \text{rect}_i(x) \), \( i = 1, \ldots, N \), as

\[
\text{rect}_i(x) = \begin{cases} 
1 & \text{if } x_i < x < x_{i+1} \\
0 & \text{otherwise}
\end{cases}
\]

(10)

By matching the boundary conditions, we obtain at the top interface \( z = 0 \),

\[
\exp(-jk_0n_1 \sin \theta x) + \sum_i R_i \exp(-jk_{xi}x) = \sum_{p=1}^{N} \sum_{m=1}^{N} w_{mp} \text{rect}_p(x) [ c_m^+ \exp(-\lambda_m h) + c_m^- \exp[\lambda_m(z-h)] ] ,
\]

(11a)

\[
-\frac{j k_0 \cos \theta}{n_1} \exp(-jk_0n_1 \sin \theta x) + \sum_i \frac{j k_{xi}}{\epsilon_1} R_i \exp(-jk_{xi}x) = \sum_{p=1}^{N} \sum_{m=1}^{N} \lambda_m v_{mp} \text{rect}_p(x) [-c_m^+ + c_m^- \exp(-\lambda_m h)] ,
\]

(11b)

and at the substrate interface \( z = h \),

\[
\sum_i T_i \exp(-jk_{xi}x) = \sum_{p=1}^{N} \sum_{m=1}^{N} w_{mp} \text{rect}_p(x) [ c_m^+ \exp(-\lambda_m h) + c_m^- ] ,
\]

(12a)

\[
\sum_i -\frac{j k_{3,zi}}{\epsilon_3} T_i \exp(-jk_{zi}x) = \sum_{p=1}^{N} \sum_{m=1}^{N} \lambda_m v_{mp} \text{rect}_p(x) [-c_m^+ \exp(-\lambda_m h) + c_m^- ] .
\]

(12b)

In Eqs. (11) and (12) the sign \( = \) is used to specify the equality between two functions of the variable \( x \) that are expressed in different function expansions. \( w_{mp} \) is the \( p \)th component of vector \( W_m \) and \( v_{mp} \) is the \( p \)th component of vector \( V_m \). \( V_m = \epsilon_1 \frac{\partial}{\partial x} W_m Q \) with \( Q \) being a diagonal matrix with the diagonal element \( \lambda_m \). Hereafter, an equal number of Rayleigh orders and of point locations is considered. Thus in Eqs. (11) and (12) the summation over the index \( i \) runs from \((-N-1)/2 \) to \((N-1)/2 \) for odd \( N \) values or from \(-N/2 \) to \( N/2 - 1 \) for even \( N \) values.

Equations (11) and (12) constitute a system of equations in known function expansions with unknown expansion coefficients \( R_i \), \( T_i \), \( c_m^+ \), and \( c_m^- \). To solve this system, we use the method of moments. In this method, a projection basis is first chosen. Then both sides of the series-expansion equations are projected on the projection basis. Finally, the linear system of equations is solved with standard numerical techniques. As the projection basis, we chose the set of plane waves of the Rayleigh expansion. This choice is commented on in Section 4. It is easily found that \( \text{rect}(x) = \sum_{m=1}^{N} p_{i,m} \exp(-jk_{xm}x) \), with

\[
p_{i,m} = \exp(jk_{xm}x_i) \{ \exp(jk_{xm}(x_{i+1} - x_i)) - \exp(jk_{xm}(x_i - x_{i-1})) \} / (jk_{xm} \Lambda).
\]

(13)

In a matrix format, Eqs. (11) and (12) can be written as

\[
\begin{bmatrix} 
\delta_{i,0} \\
\frac{j \delta_{i,0} k_0 \cos \theta}{n_1} \\
\sum_{p=1}^{N} \sum_{m=1}^{N} \lambda_m v_{mp} \text{rect}_p(x) \end{bmatrix} \begin{bmatrix} 
\frac{1}{-jZ_4} \\
\frac{1}{-jZ_4} \\
\frac{1}{-jZ_4} \\
\end{bmatrix} = \begin{bmatrix} 
\text{PW} \\
\text{PV} \\
\text{PVX} \\
\end{bmatrix} \begin{bmatrix} 
1 \\
T_i \\
R_i \\
\end{bmatrix}.
\]

(14)

\[
\begin{bmatrix} 
\frac{1}{Z_4} \frac{1}{Z_4} \frac{1}{Z_4} \frac{1}{Z_4} \\
\text{PW} \\
\text{PW} \\
\text{PW} \\
\end{bmatrix} = \begin{bmatrix} 
\begin{bmatrix} 
\text{PW} \\
\text{PW} \\
\text{PVX} \\
\end{bmatrix} \end{bmatrix} \begin{bmatrix} 
1 \\
T_i \\
R_i \\
\end{bmatrix}.
\]

(15)

where \( \text{Z}_4 \), \( \text{Z}_4 \), and \( \text{Z}_4 \) are diagonal matrices with the diagonal elements \( \exp(-\lambda_h h) \), \( k_{1,1}/\epsilon_1 \), and \( k_{3,1}/\epsilon_3 \), respectively. \( P \) is the matrix formed by the \( p_{i,m} \) coefficients. Equations (14) and (15) are solved numerically. One may first analytically eliminate \( R_i \) and \( T_i \), then solve the resulting set of equations for the \( c_m^+ \) and \( c_m^- \) coefficients,
and, finally, substitute the $c_n^+$ and $c_n^-$ coefficients back into Eqs. (14) and (15). The diffraction efficiencies are

$$DE_{ri} = R_i R_i^* \text{Re}(k_{1z1}/k_0 n_1 \cos \theta),$$

$$DE_{ti} = T_i T_i^* \text{Re}(k_{3z1}/k_0 n_3 \cos \theta).$$

The field inside the grating region can be computed at the mesh points with Eq. (9).

### B. TE Polarization

We briefly outline here the few modifications necessary to implement the TE polarization case. Equations (1)–(3) remain valid, but now, $\Psi$ represent the electric field. Discretizing the Helmholtz equation ($\partial^2 \Psi/\partial z^2 + k_0^2 e \Psi = 0$) leads to

$$\frac{\partial^2 \Psi}{\partial z^2} + E \Psi = 0,$$

where $E = D_2 D_1 + k_0^2 e$. (16)

In Eq. (16), $e_3$ is the diagonal matrix with diagonal coefficients

$$e_3(i,i) = \langle \epsilon(x) \rangle_{[x_{i-1}, x_i]}.$$  (17)

For TE polarization, the tangential field components are $\Psi$ and ($\partial \Psi/\partial z$). Equations (14) and (15) become

$$\begin{bmatrix} \delta_{i,0} \cr j \delta_{i,0} k_0 n_1 \cos \theta \end{bmatrix} + \begin{bmatrix} 1 \cr -j Y_1 \end{bmatrix} R = \begin{bmatrix} PW & PWX \\
PV & -PVX \end{bmatrix} \begin{bmatrix} c^+ \\
c^- \end{bmatrix},$$

(18)

$$\begin{bmatrix} 1 \\
j Y_3 \end{bmatrix} T \begin{bmatrix} PWX \\
PVX \end{bmatrix} \begin{bmatrix} c^+ \\
c^- \end{bmatrix},$$

(19)

where $Y_1$ and $Y_3$ are diagonal matrices with the diagonal elements $k_{1z1}$ and $k_{3z1}$, respectively. $V$ is the matrix formed by the vectors $V_m$, $V_m = W_m Q$, $W$, $Q$, and $X$ are as defined previously. The diffraction efficiencies are

$$DE_{ri} = R_i R_i^* \text{Re}(k_{1z1}/k_0 n_1 \cos \theta),$$

$$DE_{ti} = T_i T_i^* \text{Re}(k_{3z1}/k_0 n_3 \cos \theta).$$

### 3. NUMERICAL RESULTS

In this section the convergence performance of the present method is studied for TM and TE polarizations. For the sake of comparison, the convergence speed achieved is systematically compared with that achieved with the RCWA. Our RCWA implementation for TM polarization is described in Refs. 7 and 8. As will be shown, the relative performance with respect to the RCWA strongly depends on the grating materials. Metallic lamellar gratings operating in the visible and in the infrared region of the spectrum are first considered in Subsection 3.A. Dielectric gratings for which relatively bad convergence performance are observed are then considered in Subsection 3.B. In all calculations, the number of retained Rayleigh orders is equal to the number $N$ of point locations.

#### A. Metallic Gratings

In general, with FD methods, the convergence performance depends on the set of selected discrete points. We first consider a uniform sampling and a metallic grating on an aluminum substrate in a Littrow mounting for operation with CO2 lasers. This example was previously considered in the literature by several researchers concerned with the convergence performance of the RCWA. The diffraction parameters are given in the caption of Fig. 3. Only the minus-first and zeroth reflected orders are propagating. For a perfectly uniform discretization, Figs. 3(a) and 3(b) show the error (defined as the computed diffraction efficiency minus the exact value) of the zeroth order as a function of $N$ for TM and TE polarizations, respectively. In these figures, pluses are obtained with the present method and circles with the RCWA. This convention is used throughout the paper.

---

**Fig. 3.** Uniform discretization: computational error for the reflected zeroth order of a metallic grating: (a) TM, (b) TE. The values of the parameters are: $\theta = 30^\circ$, $f = 0.5$, $n_1 = n_1^* = 1$, $n_3 = n = 0.22 + 6.71i$, and $\lambda = \Lambda = h = 1.0 \mu m$. The exact value of the reflected zeroth order is 84.848% for TM and 73.428% for TE. Pluses, present method; circles, RCWA.
Although perfectly regular grids are often preferred for ease of programming, nonuniform discretizations with a minute sampling in localized regions near sharp discontinuities or edges is desirable to model the local field phenomena accurately. For a given accuracy, nonuniform discretizations result in reduced matrix sizes (see Chapter 11 of Ref. 12 for FDTD methods or, in a more related context, Ref. 22, for instance.

For the numerical results reported in this subsection, the set of discrete points is generated as follows. First, a discretization step \( h_g = \Lambda / N' \) is computed, with \( N' \) being an arbitrary number of discrete points. Then the following step-by-step procedure, illustrated in Fig. 4, is performed:

(a) A cross (circle) point is associated with every left (right) profile-transition location; see Fig. 4(a).

(b) For each point of Fig. 4(a), \( 2P \) discrete points \( (p = 1, \ldots, P) \) are generated symmetrically around the discontinuity. Figure 4(b) illustrates the case \( P = 2 \). Let us consider the left transition location, a similar operation being performed for the right transition location. The distance \( D_p \) between two successive crosses is \( h_g (P + 1 - p)^r \), where \( r \) is a real number, typically, 0.5, 1, or 2. After the crosses are generated, circles are placed equidistantly between them. Note that this discretization procedure results in increased resolution near boundaries.

(c) The last step consists in generating a uniform sampling in all ridge or groove intervals denoted by \( T \) in Fig. 4(c). Every \( T \) interval is discretized by a set of equidistant discrete points with alternate crosses and circles. The distance between two successive crosses or circles is chosen to be as close as possible to the discretization step \( h_g = \Lambda / N' \). We end up with Fig. 4(c). Note that this discretization procedure results in \( N \) discrete crosses or circles and that the value of \( N \) slightly differs from the initial value \( N' \). In the following, \( P \) is equal to 3.

We first apply this procedure to the metallic grating previously considered for uniform discretization. Results for the reflected zeroth-order diffraction efficiency are shown in Tables 1 and 2 for TM and TE polarizations, respectively. The first column is for the grating of Fig. 3 (TM Polarization) and for the grating of Fig. 3 (TE Polarization), respectively. The third columns correspond to numerical values computed with the nonuniform discretization, and \( r = 0.5 \). For the sake of comparison, the results obtained with a uniform discretization in Fig. 3 are given in the second columns and those obtained with the RCWA are in the fourth columns. The values in the last columns are taken from the second column of Table 1 in Ref. 24, where an improved convergence rate is achieved with the RCWA by adaptive spatial resolution. The comparison of the numerical values in columns 2 and 3 shows that the nonuniform discretization accelerates the convergence speed, especially for small \( N \) values for both TE and TM polarization. Moreover, by comparison of columns 3, 4, and 5, we conclude that the numerical performance of the present method for TM polarization is between those of the RCWA and the RCWA with adaptive spatial resolution and is worse than that of the RCWA for TE polarization.

We now consider metal with larger conductivity. The example is taken from Ref. 27, where the computation of the near-field pattern with the RCWA was studied for TM polarization and where the convergence performance of the RCWA was tested for highly conductive metals. A lamellar grating in a Littrow mount with only the zeroth and minus-first orders propagating is considered for \( n = n_0 = 1 - 40i \). The diffraction parameters are given in the caption of Fig. 5. For these very specific values, a resonance effect that is related to the presence of a leaky wave at the grating surface is observed for TM polarization. In general, the diffraction efficiencies quickly vary near the resonance, and grating resonances are particularly suitable for studying the robustness of any numerical methods.\(^{27}\) Figures 5(a) and 5(b) correspond to

---

**Table 1. Reflected Zero-Order Efficiencies for Various Truncation Orders \( N \) and for the Grating of Fig. 3 (TM Polarization)**

<table>
<thead>
<tr>
<th>( N )</th>
<th>Uniform Sampling</th>
<th>Nonuniform Sampling</th>
<th>RCWA</th>
<th>From Ref. 24</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>0.78284</td>
<td>0.83704</td>
<td>0.83485</td>
<td>0.84519</td>
</tr>
<tr>
<td>17</td>
<td>0.81855</td>
<td>0.84305</td>
<td>0.83383</td>
<td>0.84663</td>
</tr>
<tr>
<td>21</td>
<td>0.83300</td>
<td>0.84592</td>
<td>0.84211</td>
<td>0.84831</td>
</tr>
<tr>
<td>25</td>
<td>0.83980</td>
<td>0.84717</td>
<td>0.83960</td>
<td>0.84862</td>
</tr>
<tr>
<td>29</td>
<td>0.84339</td>
<td>0.84774</td>
<td>0.84241</td>
<td>0.84850</td>
</tr>
<tr>
<td>41</td>
<td>0.84712</td>
<td>0.84829</td>
<td>0.84425</td>
<td>0.84848</td>
</tr>
<tr>
<td>61</td>
<td>0.84825</td>
<td>0.84844</td>
<td>0.84579</td>
<td>0.84847</td>
</tr>
<tr>
<td>81</td>
<td>0.84844</td>
<td>0.84847</td>
<td>0.84677</td>
<td>0.84848</td>
</tr>
</tbody>
</table>

**Table 2. Reflected Zero-Order Efficiencies for Various Truncation Orders \( N \) and for the Grating of Fig. 3 (TE Polarization)**

<table>
<thead>
<tr>
<th>( N )</th>
<th>Uniform Sampling</th>
<th>Nonuniform Sampling</th>
<th>RCWA</th>
<th>From Ref. 24</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>0.23815</td>
<td>0.44096</td>
<td>0.81979</td>
<td>0.72760</td>
</tr>
<tr>
<td>17</td>
<td>0.47834</td>
<td>0.59064</td>
<td>0.78196</td>
<td>0.73441</td>
</tr>
<tr>
<td>21</td>
<td>0.59588</td>
<td>0.65601</td>
<td>0.76227</td>
<td>0.73422</td>
</tr>
<tr>
<td>25</td>
<td>0.65193</td>
<td>0.68540</td>
<td>0.75181</td>
<td>0.73427</td>
</tr>
<tr>
<td>29</td>
<td>0.68109</td>
<td>0.70082</td>
<td>0.73857</td>
<td>0.73428</td>
</tr>
<tr>
<td>41</td>
<td>0.71425</td>
<td>0.71962</td>
<td>0.73857</td>
<td>0.73428</td>
</tr>
<tr>
<td>61</td>
<td>0.72711</td>
<td>0.72821</td>
<td>0.73561</td>
<td>0.73428</td>
</tr>
<tr>
<td>81</td>
<td>0.73063</td>
<td>0.73097</td>
<td>0.73485</td>
<td>0.73428</td>
</tr>
</tbody>
</table>
TM and TE polarizations, respectively, and are obtained for $r = 1$. For this highly conductive case, the present method achieves better convergence performance than the RCWA for TE and for TM polarizations.

We have tested the FD method for several other diffraction problems by considering several metals, incident wavelengths, fill factors, grating depths, and grating periods and also by considering multigroove profiles. The results are not reported here for the sake of brevity, but several observations are worth mentioning. In general, situations for which the performance of the present method compares favorably with the performance of the RCWA correspond to highly conductive metals, especially for TM polarization. As the metal conductivity is lowered, the RCWA becomes progressively superior. We also observed that for a given metal, the convergence rate of the present method, relative to the RCWA, decreases as larger and larger air grooves (typically larger than $\sim 2$ or $3\lambda$) are considered. We also observed that the grating period (for a fixed air-groove width) and the grating depth have nearly no impact on the performance of the present method.

### B. Limitation: Dielectric Gratings

The above observations can be interpreted if one considers that the nearly-uniform discretization procedure provided by step (c) is not suited for sampling regions where, for example, the electromagnetic field is slowly varying, large air-groove regions. This limitation results in slower convergence performance and in increased computational loads. This bad efficiency is especially stringent when dielectric gratings with large periods and with large grooves or ridges are considered. Figure 6 illustrates our purpose for a dielectric grating with a $2\pi$-period and with a $\sim 1.5\lambda$-large air groove illuminated under TM polarization. It is clearly seen that the convergence performance of the present method ($r = 0.5$) compares very badly with that of the RCWA.

Although a better performance can be reasonably expected for a more appropriate discretization, we believe that the bad performance reported in Fig. 6 is intrinsic to the FD method. Even with a refined discretization, the performance of the present method is likely not to meet the performance of the RCWA. In fact, it is conceivable that a Fourier basis, especially with respect to low Fourier harmonics, is much more appropriate than a sampling on discrete point locations to represent slowly varying functions.

![Fig. 5. Nonuniform discretization ($r = 1$). Minus-first reflected order for the grating considered in Ref. 27. The values of the parameters are $\theta = \arcsin(\lambda/2\Lambda)$, $f = 0.57$, $n_1 = n' = 1$, $n_3 = n = 1 + 40i$, $\lambda = 1 \mu m$, $\Lambda = 1.2361 \mu m$, and $h = 0.4\Lambda$. (a) TM polarization, (b) TE polarization. Pulses, present method; circles, RCWA.](image)

![Fig. 6. TM polarization and nonuniform discretization ($r = 0.5$). Computational error for the transmitted minus-first order of a dielectric grating. The dashed curve represents the deviation to the energy conservation obtained with the present method. The values of the parameters are $\theta = 30^\circ$, $f = 0.234$, $n_1 = n' = 1$, $n_3 = 1.5$, $n = 2.3$, $\lambda = 1 \mu m$, $\Lambda = 2 \mu m$, and $h = 1 \mu m$. The exact value for the diffraction efficiency is 51.062%. Pluses, present method; circles, RCWA.](image)
4. REMARKS

A. Convergence Performance and Discretization

The convergence performance of the present method is highly improved by the use of the interpolation rules justified in Appendix A. With a majority rule that simply consists in assigning to the coefficients of matrices $e_1$ and $e_2$ the values of the permittivity at the discrete point locations $e_1(i, i) = e(x_i)$ or $e_2(i, i) = e(x'_i)$, an oscillatory convergence with large oscillation amplitude is observed.

Discretizing at the boundaries as suggested by step (a) in Subsection 3(A) is known to be advisable with FD methods because it results in an accurate localization of discontinuities. Perhaps, it is less well known that the alternating discretization (one-half of the discontinuities are located on the $x_i$ points, the other half on the $x'_i$ points) proposed in step (a) results in better convergence rates than the standard discretization approaches that consists in locating all boundaries either with $x_i$ points or with $x'_i$ points. Step (b) results in increased resolution near boundaries and accelerates the convergence speed, especially for small $N$ values and for metallic gratings: see Tables 1 and 2, for instance. Step (c) consists in generating a discretization grid as uniform as possible. This discretization procedure is not optimal; enhanced convergence performances are likely to be achieved with refined discretization relying on physical and mathematical considerations; see Chapter 11 in Ref. 12. Such considerations are out of the scope of this work.

B. Boundary Conditions

In Section 2 the method of moments is used for matching the boundary conditions after the electromagnetic fields in the grating region are expressed by an expansion composed of rectangle functions centered at the discrete points. A more natural choice consists in calculating the electromagnetic fields in the incident and in the substrate at the discrete $x_i$ points and then in matching the boundary conditions at the discrete points. This natural and simple approach was first considered, and nearly identical convergence rates were obtained.

However, when a stack of lamellar gratings with identical periods is considered or when a continuous-profile grating (a blazed grating, for example) is sliced for computational purposes, the natural approach is problematic. This is due to the fact that each slice requires a specific discrete set of $x_i$ points. This drawback does not hold with the present method. Since the electromagnetic field quantities in the substrate, in the incident medium, and in all the slices are all expressed in the Rayleigh basis, standard procedures like the $S$- or $R$-matrix methods may be applied.

C. Computational Effort: Sparsity

The present method is twofold. FD’s are first applied to discretize Maxwell’s equations inside the grating region, and a system of differential equations is obtained. Once the eigenvalues and eigenvectors of this system are found, the boundary conditions at the grating interfaces are matched to compute the diffraction efficiencies. This is achieved by solving a linear system of equations.

The numerical implementation is very similar to that of the RCWA. Generally, except for the additional $P$ matrix involved when the boundary conditions at the grating interfaces are solved with the method of moments, the numerical complexities of the RCWA and the FD modal method for solving the boundary conditions are identical. The eigenproblem formulations, however, are rather different. Whereas the RCWA requires solution of a full matrix that has no symmetry in general, the $Q$ matrix of the FD modal approach is sparse. It is a tridiagonal matrix with nonzero upper right and lower left coefficients. Hence a significant enhancement in computational efficiency and a reduction in computer memory requirement can be achieved by use of an appropriate eigenvalue software package.

D. Special Case of Lossless Metals

In practice, to access the ultimate performance of many designs with metallic gratings operating in the visible or near-infrared region of the spectrum, it is relevant to consider absorption-free metals, i.e., metals with a purely real negative permittivity. The RCWA method for TM polarization fails in this case. The reason comes from the fact that the inverse rule cannot be applied. More specifically, neither condition (a) nor condition (b) in theorem 3 of Ref. 29 is satisfied for a lossless metal with a purely real negative permittivity. Figure 7 illustrates our purpose. The same diffraction geometry as in Fig. 3 is considered, except that the metal refractive index is $6.7i$ ($n = n_3 = 6.71i$). Clearly, the present method shown with pluses does not present this drawback.

E. Lossless Dielectric Gratings and Energy Conservation

Conservation of energy for a lossless grating stipulates that the sum of all the diffraction efficiencies is one. As opposed to the RCWA, for instance, the proposed method does not automatically verify this conservation law. Consequently, an energy balance criterion may be used to access the accuracy of an individual diffraction efficiency. In Fig. 6 the dashed curve represents the sum of all the diffraction efficiencies.

![Fig. 7](image-url)
diffraction efficiencies minus one. It is noticeable that this error is approximately ten times smaller than the error for the diffraction efficiency of the reflected minus-first order. The same conclusion holds when considering the relative error. We conclude that the energy balance criterion cannot be reliably used to access the accuracy of the computed diffracted orders.

5. CONCLUSION

We have presented a first-order finite-difference modal method for the analysis of 1D lamellar gratings in a classical mounting. The FD scheme incorporates three features that accelerate the convergence performance of the method: (1) the discrete permittivity is interpolated at the lamellar boundaries, (2) mesh points are located on the permittivity discontinuities, and (3) a nonuniform sampling with increased resolution is performed near the discontinuities. The convergence performance of the present method was compared with the RCWA, also named the Fourier modal method. We found that the present method is much inferior to the RCWA for dielectric gratings. In contrast, for metallic gratings operating in the near-infrared regions of the spectrum, the present method compares favorably with the RCWA, especially for TM polarization. In fact, our numerical results indicate that the higher the conductivity, the better the convergence of the present method with respect to the RCWA.

The main conclusion of this work is that, although they are now applied in the visible for metallic guides. They are now applied in the near-infrared regions of the spectrum, the present method is much inferior to the RCWA for dielectric waveguides. In the view of the results obtained in this paper, this appears questionable.

APPENDIX A: INTERPOLATION AT A BOUNDARY

This appendix provides a justification for the interpolation rules of Eqs. (7a), (7b), and (17). Only the TM polarization case is considered for the sake of conciseness. We consider a planar interface that separates two media with relative permittivities $\varepsilon_1$ and $\varepsilon_2$; see Fig. 8. In the two homogeneous media, the electromagnetic fields can be expanded in a plane-wave basis. We consider one of the Fourier components of the spectrum and denote by $q$ the $z$ component of its wave vector. In medium 1, the $y$ component of the magnetic field vector can be written as

$$\Psi_1(x, z) = \exp(-j q z) [a_1 \exp(-j r x) + \beta_1 \exp(j r x)].$$

Similarly, in medium 2, we have

$$\Psi_2(x, z) = \exp(-j q z) [a_2 \exp(-j r' x) + \beta_2 \exp(j r' x)],$$

with

$$r^2 + q^2 = \varepsilon_1 k^2, \quad r'^2 + q^2 = \varepsilon_2 k^2.$$

For $\Psi_1$ and $\Psi_2$ to be solutions of Maxwell’s equations, $\Psi$ and $(1/\varepsilon_0)\partial \Psi/\partial x$ have to be continuous at the interface. Thus the complex amplitudes $a_1$, $a_2$, $\beta_1$, and $\beta_2$ verify

$$a_1 + \beta_1 = a_2 + \beta_2, \quad \frac{r}{\varepsilon_1} (a_1 - \beta_1) = \frac{r'}{\varepsilon_2} (a_2 - \beta_2).$$

The three Maxwell’s curl equations for TM waves are

$$j \omega \varepsilon_0 E_z = \partial_x \Psi,$$

$$j \omega \varepsilon_0 E_x = -\partial_z \Psi,$$

$$j \omega \mu_0 \Psi' = \partial_z E_z - \partial_x E_x.$$

According to Eq. (A5), the $z$ components of the electric field in media 1 and 2 are

$$E_{z1} = -\frac{r}{\omega \varepsilon_0 \varepsilon_1} \exp(-j q z) [a_1 \exp(-j r x) - \beta_1 \exp(j r x)],$$

$$E_{z2} = -\frac{r'}{\omega \varepsilon_0 \varepsilon_2} \exp(-j q z) [a_2 \exp(-j r' x) - \beta_2 \exp(j r' x)],$$

respectively. Our objective is to determine the discrete values of the relative permittivity to minimize discretization errors. We proceed first with Eq. (A5). Its discretized version is

$$j \omega \varepsilon_i^z E_{z,i} = \frac{1}{h_i^z} (\Psi_{i+1} - \Psi_i),$$

where $h_i = x_{i+1} - x_i$ and $\varepsilon_i^z$ is the discrete value (at abscissa $x_i^z$) of $\varepsilon(x)$. The superscript $z$ indicates that the discretized permittivity is related to the $z$ component of the electric field. Clearly, in FD methods, the discrete
functions differ from the exact solution; the difference vanishes when the discretization step tends toward zero. Consequently, in Eq. (A10), the discrete field quantities $E_{z,i}$, $\Psi_{i+1}$, and $\Psi_i$ differ from the exact values $E_{z}(x_i)$, $\Psi_{i+1}(x_{i+1})$, and $\Psi_i(x_i)$. How to determine the value of $\epsilon_i^2$ so that the discrete field quantities best approximate the exact values? A reasonably good solution to that problem is obtained by replacing in Eq. (A10) the discrete field quantities by their exact values, then by expanding Eq. (A10) in a power series of the discretization step, and finally by selecting the value of $\epsilon_i^2$ that nullifies the lower-order term of the power series. Doing so, we obtain for the power-series expansion

$$r' \epsilon_i^2 (\beta_2 - \alpha_2) h_1 = (\beta_2 + \alpha_2 - \beta_1 - \alpha_1)$$

$$+ r'(1 - f_1)(\beta_2 - \alpha_2)$$

$$+ r f_1(\beta_1 - \alpha_1) h + O(h_2^2),$$  \hspace{1cm} (A11)

where $f_1$ is defined as in Fig. 8. Taking into account the boundary conditions (A4), Eq. (A11) is satisfied up to the first order in $h_1$, provided that $\epsilon_i^2$ is the average relative permittivity

$$\epsilon_i^2 = f_1 \epsilon_1 + (1 - f_1) \epsilon_2.$$  \hspace{1cm} (A12)

Equation (A12) holds for any value of $r$ (or $q$) up to the first power of $h_1$.

We now proceed with the interpolation of the discrete value of the relative permittivity related to the $x$ component of the electric field. From Eqs. (A6) and (A7) and noting that $\partial_x \Psi = -jq \Psi$, we first obtain $\epsilon(k^2 \Psi + j \omega \epsilon_0 \partial_x E_z) = q^2 \Psi$ and, after discretization,

$$\epsilon_i^2 \left[ \frac{k^2 \Psi_i}{h_2} + \frac{j \omega \epsilon_0}{h_2} (E_{z,i} - E_{z,i-1}) \right] = q^2 \Psi_i,$$  \hspace{1cm} (A13)

where $\epsilon_i^2$ is the discrete value of $\epsilon$ at the discrete point $x_i$ and $h_2 = x_i - x_{i-1}$. Incorporating terms up to the order 2, we obtain

$$\epsilon_i^2 \left[ \frac{r}{\epsilon_1} (\alpha_1 - \beta_1) - \frac{r'}{\epsilon_2} (\alpha_2 - \beta_2) + h_2 \left( \frac{k^2 (\alpha_2 + \beta_2)}{\epsilon_2} \right) \right. $$

$$- \left. \frac{r^2}{\epsilon_1} (\alpha_1 + \beta_1) f_2 - \frac{r'^2}{\epsilon_2} (\alpha_2 + \beta_2) (1 - f_2) \right] = q^2 (\alpha_2 + \beta_2) h_2 + O(h_2^2).$$  \hspace{1cm} (A14)

The zeroth-order term in the left-hand side of Eq. (A14) is null; see Eq. (A4). Expressing $r$ and $r'$ as a function of $q$ and $k$ according to Eq. (2) and using the continuity of $\Psi$, Eq. (A14) reduces to the weighted average of the inverse of the relative permittivities

$$\epsilon_i^2 = \frac{1}{f_2 \epsilon_1 + (1 - f_2) / \epsilon_2},$$  \hspace{1cm} (A15)

which holds for any value of $q$ up to the first power of $h_2$.

The tensorial interpolation rules of Eqs. (A12) and (A15) hold for any value of the $z$ component of the wave vector. Since any field distribution can be expanded in a plane-wave basis, it is concluded that they allow for the field to satisfy locally the boundary conditions at an interface up to the first order of the discretization step.

Equations (A12) and (A15) have a simple interpretation: For electric field components parallel to the interface, the relative permittivity has to be interpolated with a standard weighted average, and for electric field components normal to the interface, with an inverse weighted average. Although not mentioned above for the sake of clarity, a similar interpolation scheme holds for magnetic material interfaces. Finally, note that, although much effort has been carried out to take into account boundaries in FD methods accurately, the tensorial interpolation rules of Eqs. (A12) and (A15) are not widely used. For instance, in FDTD methods, related but different average rules exist for the permittivity and the permeability when a staircase approximations of curved interfaces are generated on uniform grids. In FD modal approaches used for waveguide computations,\textsuperscript{21–23} more complex FD expressions obtained from the second-order derivative (Helmholtz) Maxwell’s equations have been studied.\textsuperscript{30–32}

In contrast, the interpolation rules of Eqs. (A12) and (A15) are simple to implement and verify the first-order derivative curl equations. They were previously used for photonic bandgap computation in Ref. 33, where their use was briefly and weakly justified through effective medium theory considerations. We believe that the justification provided above is more relevant. Moreover, the numerical results obtained by the authors of Ref. 33 for multidimensional periodic structures indicate that the present interpolation rules may be efficiently used for crossed-grating diffraction problems.

**ACKNOWLEDGMENTS**

We are grateful to Violaine Louvet of the Laboratoire de Mathématique in Orsay and to Maria Huhtala of the Helsinki University of Technology in Finland. This research was supported in part by the Program Micro-systèmes of the Ministère de l’Éducation Nationale et de la Recherche Technologique under contract 99N37/0002.

**REFERENCES**


28. See Chap. 10 in Ref. 12, for instance.