Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach

M. G. Moharam, Drew A. Pommet, and Eric B. Grann

Center for Research and Education in Optics and Lasers, University of Central Florida, Orlando, Florida 32826

T. K. Gaylord

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Received September 20, 1994; revised manuscript received December 5, 1994; accepted December 6, 1994

An enhanced, numerically stable transmittance matrix approach is developed and is applied to the implementation of the rigorous coupled-wave analysis for surface-relief and multilevel gratings. The enhanced approach is shown to produce numerically stable results for excessively deep multilevel surface-relief dielectric gratings. The nature of the numerical instability for the classic transmission matrix approach in the presence of evanescent fields is determined. The finite precision of the numerical representation on digital computers results in insufficient accuracy in numerically representing the elements produced by inverting an ill-conditioned transmission matrix. These inaccuracies will result in numerical instability in the calculations for successive field matching between the layers. The new technique that we present anticipates and preempts these potential numerical problems. In addition to the full-solution approach whereby all the reflected and the transmitted amplitudes are calculated, a simpler, more efficient formulation is proposed for cases in which only the reflected amplitudes (or the transmitted amplitudes) are required. Incorporating this enhanced approach into the implementation of the rigorous coupled-wave analysis, we obtain numerically stable and convergent results for excessively deep (50 wavelengths), 16-level, asymmetric binary gratings. Calculated results are presented for both TE and TM polarization and for conical diffraction.

1. INTRODUCTION

The rigorous coupled-wave analysis (RCWA) has been widely used for the analysis and the design of diffractive structures. It is an exact solution of Maxwell's equations for the electromagnetic diffraction by grating structures. It is a relatively straightforward, noniterative, deterministic technique. The accuracy of the solution obtained depends solely on the number of terms retained in the space-harmonic expansions of the fields, with conservation of energy always being satisfied. The RCWA was developed initially for the investigation of holographic gratings. Later it was extended to surface-relief and multilevel grating structures. It has been successfully applied to transmission and reflection planar dielectric and/or absorption holographic gratings, arbitrarily profiled dielectric and/or metallic surface-relief gratings, multiplexed holographic gratings, two-dimensional surface-relief gratings, and anisotropic gratings, for both planar and conical diffraction.[1–9]

In the technique to analyze the diffraction from surface-relief grating structures, the grating is divided into a large number of sufficiently thin planar grating slabs to approximate the grating profile to an arbitrary degree of accuracy. The electromagnetic fields in each grating layer are determined by the coupled-wave approach (or by another approach, such as the modal approach). The electromagnetic boundary conditions (continuity of the tangential electric- and magnetic-field components) are then applied in sequence at the interfaces among the output region, the individual grating layers, and finally the input region to yield the reflected and the transmitted diffracted field amplitudes and the diffraction efficiencies. Whereas the RCWA formulation is numerically stable for single-layer binary or planar gratings, the particular approach used in the successive electromagnetic field matching at the boundaries between the individual grating layers may introduce numerical instability. For example, the standard transmittance matrix approach for propagating the field matching is known to produce numerical instability because of the presence of evanescent fields, even in the case of simple uniform multilayer systems.[10]

Several approaches have been proposed to produce stable implementation for the surface-relief grating analysis problem.[3,11–14] Moharam and Gaylord[3] obtained numerically stable RCWA calculations of the diffraction characteristics for surface-relief dielectric gratings to a grating depth of as many as four wavelengths, for TE polarization, by a sequential Gaussian elimination scheme. Pai and Awada[11] introduced layer transmission matrices and interface reflection and transmission matrices to derive the solution for the RCWA in terms of a multiple-reflection series that appears to be stable for dielectric gratings to a grating depth of as many as four wavelengths, for TE polarization. Han *et al.*[12] employed a shoot-back method to obtain a stable numerical solution for the two-dimensional RCWA of dielectric surface-relief

0740-3232/95/051077-10$06.00

gratings, to a grating depth of one wavelength. Li[13] used the R-matrix propagation algorithm to propagate the field through the layers in the modal approach to obtain stable results for deep dielectric and metallic one-dimensional gratings in the conical mount. Chateau and Hugonin[14] presented an algorithm for the RCWA and obtained numerically stable results for deep one-dimensional (one- and two-level) binary gratings and holographic gratings, for TE polarization.

It is generally well known that evanescent fields may cause numerical problems in determining the electromagnetic propagation characteristics in uniform multilayer systems.[10] The numerical instability, which of course is not inherent in the reflection–transmission problem, is introduced when the transmittance (characteristic) matrix or other approaches are used for the successive matching of the tangential electromagnetic fields at the interface between the layers. Several techniques have been proposed and have been successfully implemented to yield a numerically stable solution for this uniform multilayer problem, including the R-matrix approach presented by Schwartz and DeSandre,[15] the scattering matrix approach proposed by Ko and Sambles,[16] and the impedance formalism developed by Cousins and Gottschalk.[17] However, these approaches are not suitable for the RCWA for surface-relief gratings. The RCWA formulation involves a Fourier expansion of the fields, in each grating layer, into a set of coupled space harmonics. These field harmonics, in any one layer, are phase matched to the field harmonics at all the other grating layers. This method is compatible with the transmittance matrix approach, whereby both the electric and the magnetic fields, rather than the ratio between the two fields, are matched at the boundary between the layers.

In this paper an enhanced numerically stable transmittance matrix approach is developed. First, we examine the source and the nature of the numerical instability in the transmittance matrix approach to develop a clear understanding of the problem. We achieve this by considering the situation of reflection and transmission from a stack of uniform homogeneous layers. It is shown that the evanescent fields, which are associated with total internal reflection or losses, result in ill-conditioned transmittance matrices. The inversion of this ill-conditioned matrix results in a matrix with numerically large diagonal elements that cannot be represented with sufficient numerical accuracy because of the finite precision of digital computers. This numerical inaccuracy will produce numerical instability in the calculated transmitted- and reflected-field amplitudes. We develop an enhanced, numerically stable transmittance approach to eliminate the numerical instability by anticipating and preempting these potential numerical problems. The source and the nature of the numerical instability observed in the diffraction analysis of surface-relief and multilevel gratings analysis are shown to be identical to those observed in the reflection–transmission problem presented in the stack of uniform homogeneous layers. Hence the proposed enhanced technique is extended to the diffraction from surface-relief dielectric gratings. A formulation for the implementation of the RCWA for surface-relief dielectric gratings incorporating the developed enhanced transmittance approach is presented. In addition to a full-solution approach that permits the calculation of all the diffraction-reflected and -transmitted amplitudes, we present a simpler efficient approach that allows only the diffracted reflected (or the diffracted) amplitudes to be calculated. This second approach is suitable for analyzing mainly transmissive or mainly reflective grating structures, for which only the forward or the backward diffraction must be determined. It is shown that the implementation is numerically stable, and calculated results are presented for excessively deep and multilevel asymmetric surface-relief gratings. The calculated diffraction efficiencies converge to the correct value with an increasing number of space harmonics. The proposed technique is applicable to other problems that involve wave propagation through cascaded layer systems. Specifically, it is applicable to slanted periodic structures (within the grating layer) in which the eigenvalues are in positive and negative pairs.

2. REFLECTION AND TRANSMISSION FROM A STACK OF UNIFORM HOMOGENEOUS LAYERS

Consider the reflection and the transmission of a TE-polarized plane wave of free-space wavelength λ_0, incident at angle θ, on L uniform layers of refractive indices $n_1, \ldots, n_L$ and thickness $d_1, \ldots, d_L$, as shown in Fig. 1 (generalization to TM polarization is straightforward and is discussed below). The normalized electric field (in the y direction) for the input and the output regions and in each of the layers may be written as

$$
\begin{aligned}
E_0 &= [\exp(-jk_{I,z}z) + R\,\exp(jk_{I,z}z)]\exp(-jk_x x), \qquad z \le 0,\\
E_\ell &= \{P_\ell\,\exp[-k_0\gamma_\ell(z - D_{\ell-1})] \\
&\quad + Q_\ell\,\exp[k_0\gamma_\ell(z - D_\ell)]\}\exp(-jk_x x), \qquad D_{\ell-1} \le z \le D_\ell,\\
E_t &= T\,\exp\{-j[k_x x + k_{\mathrm{II},z}(z - D_L)]\}, \qquad z \ge D_L, \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
k_x &= k_0 n_{\mathrm{I}}\,\sin(\theta),\\
k_{\mathrm{I},z} &= k_0 n_{\mathrm{I}}\,\cos(\theta),\\
k_{\mathrm{II},z} &= k_0[n_{\mathrm{II}}{}^2 - n_{\mathrm{I}}{}^2\,\sin^2(\theta)]^{1/2},\\
\gamma_\ell &= j[n_\ell{}^2 - n_{\mathrm{I}}{}^2\,\sin^2(\theta)]^{1/2}, \qquad \ell = 1, \ldots, L,\\
D_\ell &= \sum_{p=1}^{\ell} d_p, \qquad (2)
\end{aligned}
$$

where R and T are the reflected and the transmitted amplitudes of the electric fields, P and Q are the field amplitudes in the uniform slab, $k_0 = 2\pi/\lambda_0$ is the wave-vector magnitude, and n_{I} and n_{II} are the refractive indices of the input and the output regions. Note that, for a lossless region, γ_ℓ is purely imaginary when $n_\ell > n_{\mathrm{I}}\,\sin(\theta)$ and is purely real when $n_\ell < n_{\mathrm{I}}\,\sin(\theta)$ (total internal reflection, i.e., evanescent fields in the dielectric layer). For a lossy region (complex refractive index) γ_ℓ is complex. We obtain the reflected and the transmitted amplitudes by matching the tangential electric- and magnetic-field components at the boundaries between the layers. At

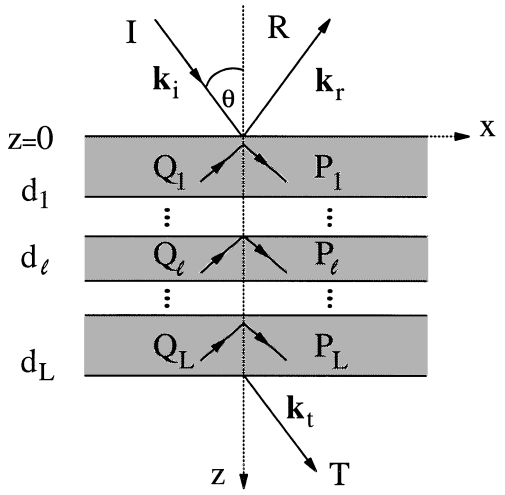

Fig. 1. Geometry for the reflection and the transmission from a stack of uniform homogeneous layers.

the boundary between the input region and the first layer ($z = 0$)

$$1 + R = P_1 + Q_1 \exp(-k_0\gamma_1 d_1),$$
$$j(k_{\mathrm{I},z}/k_0)(1 - R) = \gamma_1[P_1 - Q_1 \exp(-k_0\gamma_1 d_1)]; \quad (3)$$

at the boundary between the $(\ell - 1)$st and the ℓth layers ($z = D_{\ell-1}$)

$$P_{\ell-1} \exp(-k_0\gamma_{\ell-1}d_{\ell-1}) + Q_{\ell-1} = P_\ell + Q_\ell \exp(-k_0\gamma_\ell d_\ell),$$
$$\gamma_{\ell-1}[P_{\ell-1} \exp(-k_0\gamma_{\ell-1}d_{\ell-1}) - Q_{\ell-1}] = \gamma_\ell[P_\ell - Q_\ell \exp(-k_0\gamma_\ell d_\ell)]; \quad (4)$$

and at the boundary between the last layer and the output region ($z = D_L$)

$$P_L \exp(-k_0\gamma_L d_L) + Q_L = T,$$
$$\gamma_L[P_L \exp(-k_0\gamma_L d_L) - Q_L] = j(k_{\mathrm{II},z}/k_0)T. \quad (5)$$

Note that matching the electric and the magnetic fields at each interface produces two equations; therefore, for a structure with L layers, $2(L + 1)$ equations would be generated. This entire system of equations may be solved simultaneously for R and T by standard techniques for the solution of a system of equations (e.g., LU or **QR** decomposition) without any numerical instability for any number of layers, any layer thickness, or any number of refractive indices. The numerically inaccurate representation (including numerical overflow) of these elements, for large positive (γd), has been preempted by the normalization ($z - D$) of the exponential term involving the positive value of γ in Eq. (1). Unfortunately, for a large number of layers, the entire system of equations may be too large for an efficient numerical computation. Methods have been used to improve the computational efficiency. However, these techniques may introduce additional numerical instability, as is shown below.

Typically, Eqs. (3)–(5) are solved by a transmittance matrix approach to reduce the size of the system of simultaneous equations. In the transmittance matrix approach Eqs. (5) are used to determine the field amplitudes P_L and Q_L in terms of the transmitted coefficient T:

$$\begin{bmatrix} P_L \\ Q_L \end{bmatrix} = \begin{bmatrix} \exp(-k_0\gamma_L d_L) & 1 \\ \gamma_L \exp(-k_0\gamma_L d_L) & -\gamma_L \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ j(k_{\mathrm{II},z}/k_0) \end{bmatrix} T, \quad (6)$$

which is then substituted into Eqs. (4) for the field amplitudes P_{L-1} and Q_{L-1}:

$$\begin{bmatrix} P_{L-1} \\ Q_{L-1} \end{bmatrix} = \begin{bmatrix} \exp(-k_0\gamma_{L-1}d_{L-1}) & 1 \\ \gamma_{L-1} \exp(-k_0\gamma_{L-1}d_{L-1}) & -\gamma_{L-1} \end{bmatrix}^{-1} \times \begin{bmatrix} 1 & \exp(-k_0\gamma_L d_L) \\ \gamma_L & -\gamma_L \exp(-k_0\gamma_L d_L) \end{bmatrix} \times \begin{bmatrix} \exp(-k_0\gamma_L d_L) & 1 \\ \gamma_L \exp(-k_0\gamma_L d_L) & -\gamma_L \end{bmatrix}^{-1} \times \begin{bmatrix} 1 \\ j(k_{\mathrm{II},z}/k_0) \end{bmatrix} T. \quad (7)$$

Repeating this procedure for the remaining layers, we obtain

$$\begin{bmatrix} 1 \\ j(k_{\mathrm{I},z}/k_0) \end{bmatrix} + \begin{bmatrix} 1 \\ -j(k_{\mathrm{I},z}/k_0) \end{bmatrix} R = \prod_{\ell=1}^{L} \begin{bmatrix} 1 & \exp(-k_0\gamma_\ell d_\ell) \\ \gamma_\ell & -\gamma_\ell \exp(-k_0\gamma_\ell d_\ell) \end{bmatrix} \times \begin{bmatrix} \exp(-k_0\gamma_\ell d_\ell) & 1 \\ \gamma_\ell \exp(-k_0\gamma_\ell d_\ell) & -\gamma_\ell \end{bmatrix}^{-1} \times \begin{bmatrix} 1 \\ j(k_{\mathrm{II},z}/k_0) \end{bmatrix} T. \quad (8)$$

Note that Eq. (8) includes one matrix inversion for each layer. In any one layer, if the quantity ($\gamma_\ell d_\ell$) is large and positive (lossy media or evanescent wave), all the elements in the left-hand column of the matrix will be virtually zero. The inversion of an ill-conditioned matrix will produce some very large elements that cannot be represented with sufficient numerical accuracy because of the finite precision of the numerical representation in the digital computer (truncation error), resulting in numerical instability. Numerical overflow is an extreme case of this numerical inaccuracy. Therefore, although the reflection–transmission problem is fundamentally stable, the transmittance matrix approach may produce erroneous (or no) results for lossy media or total internal reflection situations, even for this simple multilayer (or single-layer) system.

The formulation for the TM-polarization case is similar to those of Eqs. (2)–(8); however, γ_ℓ is replaced by γ_ℓ/n_ℓ^2 (except in the exponent); $k_{\mathrm{I},z}$ is replaced by $k_{\mathrm{I},z}/n_{\mathrm{I}}^2$; and $k_{\mathrm{II},z}$ is replaced by $k_{\mathrm{II},z}/n_{\mathrm{II}}^2$. Now R and T are the reflected and the transmitted magnetic-field components (normal to the plane of incidence in the y direction). The numerical instability observed for TE polarization is also encountered in TM-polarization case.

3. NUMERICALLY STABLE ENHANCED TRANSMITTANCE APPROACH

To preempt the numerical instability associated with the inversion of the matrix, consider the last factor in the product ($\ell = L$) of Eq. (8):

$$\begin{bmatrix} 1 & \exp(-k_0\gamma_L d_L) \\ \gamma_L & -\gamma_L \exp(-k_0\gamma_L d_L) \end{bmatrix} \begin{bmatrix} \exp(-k_0\gamma_L d_L) & 1 \\ \gamma_L \exp(-k_0\gamma_L d_L) & -\gamma_L \end{bmatrix}^{-1} \times \begin{bmatrix} f_{L+1} \\ g_{L+1} \end{bmatrix} = \begin{bmatrix} 1 & \exp(-k_0\gamma_L d_L) \\ \gamma_L & -\gamma_L \exp(-k_0\gamma_L d_L) \end{bmatrix} \times \begin{bmatrix} \exp(-k_0\gamma_L d_L) & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ \gamma_L & -\gamma_L \end{bmatrix}^{-1} \begin{bmatrix} f_{L+1} \\ g_{L+1} \end{bmatrix}, \quad (9)$$

where $f_{L+1} = 1$ and $g_{L+1} = jk_{\mathrm{II},z}/k_0$. Note that the matrix to be inverted (with possible numerical problems) has been rewritten as the product of two matrices. The matrix on the right (in the product) is inverted without numerical instability. The right-hand side of Eq. (9) is then rewritten as

$$\begin{bmatrix} 1 & \exp(-k_0\gamma_L d_L) \\ \gamma_L & -\gamma_L \exp(-k_0\gamma_L d_L) \end{bmatrix} \begin{bmatrix} \exp(-k_0\gamma_L d_L) & 0 \\ 0 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} a_L \\ b_L \end{bmatrix} T, \quad (10)$$

where

$$\begin{bmatrix} a_L \\ b_L \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \gamma_L & -\gamma_L \end{bmatrix}^{-1} \begin{bmatrix} f_{L+1} \\ g_{L+1} \end{bmatrix}.$$

The remaining matrix to be inverted is ill conditioned when the diagonal element $\exp(-k_0\gamma_L d_L)$ is very small. Clearly, for this simple (2×2) diagonal matrix we can perform the inversion analytically to obtain the inverted matrix accurately. However, the inverted matrix will still have one very large element that cannot be represented with sufficient numerical accuracy because of numerical truncation in the finite precision of the numerical representation. This numerical inaccuracy will result in numerical instability in the calculation of the reflected and the transmitted amplitudes. We can avoid this problem by making the substitution $T = \exp(-k_0\gamma_L d_L)T_L$. This last factor in Eq. (6) is reduced to

$$\begin{aligned} \begin{bmatrix} f_L \\ g_L \end{bmatrix} T_L &= \begin{bmatrix} 1 & \exp(-k_0\gamma_L d_L) \\ \gamma_L & -\gamma_L \exp(-k_0\gamma_L d_L) \end{bmatrix} \times \begin{bmatrix} a_L \\ b_L \exp(-k_0\gamma_L d_L) \end{bmatrix} T_L \\ &= \begin{bmatrix} a_L + b_L \exp(-2k_0\gamma_L d_L) \\ \gamma_L[a_L - b_L \exp(-2k_0\gamma_L d_L)] \end{bmatrix} T_L. \end{aligned} \quad (11)$$

Repeating the process of Eq. (9) for the other layers, we obtain

$$\begin{bmatrix} 1 \\ j(k_{\mathrm{I},z}/k_0) \end{bmatrix} + \begin{bmatrix} 1 \\ -j(k_{\mathrm{I},z}/k_0) \end{bmatrix} R = \begin{bmatrix} f_1 \\ g_1 \end{bmatrix} T_1, \quad (12)$$

where the relationship between T and T_1 is given by

$$T = \exp(-k_0\gamma_L d_L) \cdots \exp(-k_0\gamma_\ell d_\ell) \cdots \exp(-k_0\gamma_1 d_1) T_1 \quad (13)$$

and a_ℓ, b_ℓ, f_ℓ, and g_ℓ are obtained by the process described by expressions (10) and (11). Equation (12) is easily solved for R and T (through T_1) without any numerical instability. This enhanced transmittance matrix approach is an efficient and numerically stable method. The above formulation can be obtained by means of a much simpler manipulation because it involves simple (2×2) scalar matrices.

4. RIGOROUS COUPLED-WAVE ANALYSIS FORMULATION FOR MULTILEVEL AND SURFACE-RELIEF GRATINGS

In this section the enhanced transmittance approach is extended to multilevel binary and surface-relief gratings. The surface-relief grating diffraction problem is depicted in Fig. 2. A linearly polarized electromagnetic wave is obliquely incident at an arbitrary angle of incidence θ upon a multilevel or surface-relief dielectric or lossy grating. The grating period Λ is, in general, composed of several regions with differing refractive indices. The grating is bound by two different media with refractive indices n_{I} and n_{II}. In the formulation presented here, without any loss of generality, the normal to the boundary is in the z direction, and the grating vector is in the x direction. In the grating region the periodic relative permittivity associated with the surface-relief or the multilevel grating is expandable in a Fourier series as

$$\varepsilon(x, z) = \sum_h \varepsilon_h(z) \exp[\, j(2\pi h/\Lambda)], \quad (14)$$

where $\varepsilon_h(z)$ is the hth Fourier component of the relative permittivity in the grating region and is a function of the depth. It is, in general, complex for lossy or nonsymmetric dielectric gratings. By dividing the grating into a large number of thin planar grating slabs perpendicular to the normal to the boundary (z axis), as shown in Fig. 2, we can then apply the RCWA, which was first developed for planar (including binary) gratings, to each slab grating. If the individual planar gratings are sufficiently thin, then any grating profile can be analyzed to an arbitrary degree of accuracy. The Fourier components of the grating permittivity are constant within each thin grating layer. The periodic relative permittivity in the ℓth layer, of thickness d_ℓ, is then given by

$$\varepsilon_\ell(x) = \sum_h \varepsilon_{\ell,h} \exp[\, j(2\pi h/\Lambda)], \quad D_\ell - d_\ell < z < D_\ell = \sum_{p=1}^{\ell} d_p. \quad (15)$$

The surface-relief grating diffraction problem is solved in a sequence of steps. First, the coupled-wave equations are constructed and are solved for the electromagnetic fields in each grating layer. Second, the electromagnetic boundary conditions (continuity of the tangential electric- and magnetic-field components) are applied between the input region and the first grating layer, then between

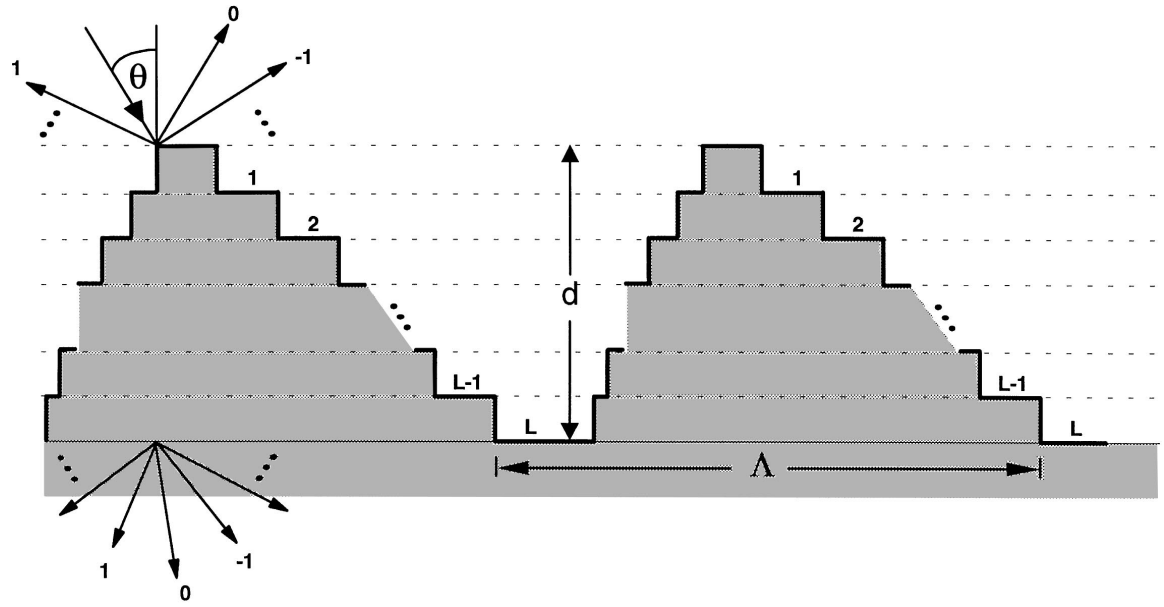

Fig. 2. Geometry for the surface-relief grating diffraction problem analyzed herein.

the first and the second grating layers, and so forth, and finally between the last grating layer and the output region. Third, the resulting array of boundary-condition equations is solved for the reflected and the transmitted diffracted field amplitudes, and the diffraction efficiencies are determined.

The detailed formulation for the efficient and stable implementation for the RCWA for planar (including binary) gratings was presented elsewhere[18] and is only reviewed here. For completeness the formulation for the TM polarization is presented (extension to TE polarization or to conical mount is straightforward). The incident reflected diffracted and transmitted diffracted fields are given by

$$H_{\mathrm{I},y} = \exp\{-jk_0n_0[\sin(\theta)x + \cos(\theta)z]\} + \sum_i R_i \exp[-j(k_{xi}x - k_{\mathrm{I},zi}z)],$$
$$H_{\mathrm{II},y} = \sum_i T_i \exp\{-j[k_{xi}x + k_{\mathrm{II},zi}(D_L - z)]\}. \quad (16)$$

R_i and T_i are the normalized magnetic-field amplitude of the ith backward-diffracted (reflected) and forward-diffracted (transmitted) waves, respectively. The quantity k_{xi} is determined from the Floquet condition and is

$$k_{x_i} = k_0[n_{\mathrm{I}} \sin(\theta) - i(\lambda_0/\Lambda)],$$
$$k_{\ell,zi} = (k_0{}^2 n_\ell{}^2 - k_{xi}{}^2)^{1/2}, \qquad \ell = \mathrm{I}, \mathrm{II}. \quad (17)$$

Note that $k_{\mathrm{I},zi}$ and $k_{\mathrm{II},zi}$ are either positive real or negative imaginary. The diffraction efficiencies are defined as

$$\mathrm{DE}_{ri} = R_iR_i{}^* \,\mathrm{Re}[k_{\mathrm{I},zi}/k_0n_{\mathrm{I}} \cos(\theta)],$$
$$\mathrm{DE}_{ti} = T_iT_i{}^* \,\mathrm{Re}\left(\frac{k_{\mathrm{II},2zi}}{n_{\mathrm{II}}{}^2}\right)\Bigg/\left[\frac{k_0 \cos(\theta)}{n_{\mathrm{I}}}\right]. \quad (18)$$

The tangential magnetic and electric fields in the ℓth grating layer are expressed in a Fourier expansion of the form

$$H_{\ell,gy} = \sum_i U_{\ell,yi}(z)\exp(-jk_{xi}x),$$
$$E_{\ell,gx} = j(\mu_0/\varepsilon_o)^{1/2}\sum_i S_{\ell,xi}(z)\exp(-jk_{xi}x). \quad (19)$$

ε_0 and μ_0 are the permittivity and the permeability of free space. $U_{\ell,yi}(z)$ and $S_{\ell,xi}(z)$ are the normalized amplitudes of the ith space-harmonic fields such that the electromagnetic fields satisfy Maxwell's equations in the grating layer:

$$\partial H_{\ell,gy}/\partial z = -j\omega\varepsilon_0\varepsilon_\ell(x)E_{\ell,gx},$$
$$\partial E_{\ell,gx}/\partial z = -j\omega\mu_0H_{\ell,gy} + (\partial E_{\ell,gz}/\partial x), \quad (20)$$

where ω is the angular frequency of the incident wave. Substituting the expansions of Eqs. (15) and (19) into Maxwell's equations, we obtain

$$[\partial^2\mathbf{U}_{\ell,\mathbf{y}}/\partial(\mathbf{z}')^2] = [\mathbf{E}_\ell][\mathbf{B}_\ell][\mathbf{U}_{\ell,\mathbf{y}}], \quad (21)$$

where

$$\mathbf{B}_\ell = \mathbf{K}_\mathbf{x}\mathbf{E}_\ell{}^{-1}\mathbf{K}_\mathbf{x} - \mathbf{I},$$

$\mathbf{z}' = k_0z$, $\mathbf{I}$ is the unity matrix, $\mathbf{K}_\mathbf{x}$ is a diagonal matrix with diagonal elements k_{xi}/k_0, and $\mathbf{E}_\boldsymbol{\ell}$ is the matrix with its i,pth element is equal to $\varepsilon_{\ell,i-p}$ defined in Eqs. (15). We solve the coupled-wave equations (21) by calculating the eigenvalues and the eigenvectors associated with the $(n \times n)$ matrix $\mathbf{E}_\boldsymbol{\ell}\mathbf{B}_\boldsymbol{\ell}$, where n is the number of harmonics retained in the field expansions. The tangential electric field is related to the tangential magnetic field by

$$[\mathbf{S}_{\ell,\mathbf{x}}] = [\mathbf{E}_\ell]^{-1}[\partial\mathbf{U}_{\ell,\mathbf{x}}/\partial(\mathbf{z}')]. \quad (22)$$

The space harmonics of the tangential magnetic and electric fields in the ℓth grating layer, which are represented in terms of the calculated eigenvalues and eigenvectors, are given by

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

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

$$D_\ell - d_\ell < z < D_\ell = \sum_{p=1}^{\ell} d_p, \quad (23)$$

where $w_{\ell,i,m}$ and $q_{\ell,m}$ are the elements of the eigenvector matrix $\mathbf{W}_\ell$ and the square root (with positive real part) of the eigenvalues of the matrix $\mathbf{E}_\ell\mathbf{B}_\ell$, respectively. The quantities $v_{\ell,i,m}$ are the elements of the product matrix $\mathbf{V}_\ell = \mathbf{E}_\ell^{-1}\mathbf{W}_\ell\mathbf{Q}_\ell$, where $\mathbf{Q}_\ell$ is a diagonal matrix with the diagonal elements $q_{\ell,m}$. $c_{\ell,m}$ are unknown constants to be determined.

As in the case of the uniform homogeneous layer, we obtain the reflected and the transmitted diffracted amplitudes by matching the tangential electromagnetic fields at the boundaries between the grating layers. At the boundary between the input region and the first grating layer ($z = 0$)

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos(\theta)/n_{\mathrm{I}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Z}_{\mathrm{I}} \end{bmatrix}\mathbf{R} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_1\mathbf{X}_1 \\ \mathbf{V}_1 & -\mathbf{V}_1\mathbf{X}_1 \end{bmatrix}\begin{bmatrix} \mathbf{c}_1^{+} \\ \mathbf{c}_1^{-} \end{bmatrix}, \quad (24)$$

at the boundary between the $\ell - 1$ and the ℓ grating layers ($z = D_{\ell-1}$)

$$\begin{bmatrix} \mathbf{W}_{\ell-1}\mathbf{X}_{\ell-1} & \mathbf{W}_{\ell-1} \\ \mathbf{V}_{\ell-1}\mathbf{X}_{\ell-1} & -\mathbf{V}_{\ell-1} \end{bmatrix}\begin{bmatrix} \mathbf{c}_{\ell-1}^{+} \\ \mathbf{c}_{\ell-1}^{-} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\ell & \mathbf{W}_\ell\mathbf{X}_\ell \\ \mathbf{V}_\ell & -\mathbf{V}_\ell\mathbf{X}_\ell \end{bmatrix}\begin{bmatrix} \mathbf{c}_\ell^{+} \\ \mathbf{c}_\ell^{-} \end{bmatrix}, \quad (25)$$

and at the boundary between the last grating layer and the output region ($z = D_L$)

$$\begin{bmatrix} \mathbf{W}_L\mathbf{X}_L & \mathbf{W}_L \\ \mathbf{V}_L\mathbf{X}_L & -\mathbf{V}_L \end{bmatrix}\begin{bmatrix} \mathbf{c}_L^{+} \\ \mathbf{c}_L^{-} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ j\mathbf{Z}_{\mathrm{II}} \end{bmatrix}\mathbf{T}, \quad (26)$$

where $\mathbf{Z}_{\mathrm{I}}$ and $\mathbf{Z}_{\mathrm{II}}$ are diagonal matrices with the diagonal elements $k_{\mathrm{I},zi}/(n_{\mathrm{I}}^2 k_0)$ and $k_{\mathrm{II},zi}/(n_{\mathrm{II}}^2 k_0)$, respectively, and $\mathbf{X}_\ell$ is a diagonal matrix with the diagonal elements $\exp(-q_{\ell,m}d_\ell)$.

For surface-relief structures divided into L grating layers, Eqs. (24)–(26) would be a $2n(L + 1)$ system of equations, where n is the number of field harmonics retained in the formulation. In the brute-force approach all $2n(L + 1)$ equations are solved simultaneously for R_i and T_i by standard techniques for the solution of the system of equations (e.g., LU or **QR** decomposition). In this approach no numerical instability will be encountered for any number of layers, any layer thickness, or any number of refractive indices, even for large positive values of γd. For large L the entire system of equations is too large for practical numerical computation. Other methods must be introduced to increase the efficiency of the computations. Unfortunately, incorporating these methods may introduce additional numerical instability.

The standard transmittance matrix approach, described for a stack of uniform layers, can easily be extended to multilevel binary or surface-relief gratings. Of course, this extension will be accompanied by the numerical instability discussed in Section 3. These numerical instabilities are even more pronounced here because the eigenvalues associated with the coupled-wave equations are, in general, complex with positive real parts. In addition, the magnitudes of the largest eigenvalues will increase as the number of space harmonics in the field expansion increases, and numerical instability will be encountered for dielectric gratings when even a moderate number of field harmonics are retained. Following the procedure for the reflection or the transmittance of a stack of uniform layers, we may rewrite Eqs. (24)–(26) as

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos(\theta)/n_{\mathrm{I}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Z}_{\mathrm{I}} \end{bmatrix}\mathbf{R} = \prod_{\ell=1}^{L}\begin{bmatrix} \mathbf{W}_\ell & \mathbf{W}_\ell\mathbf{X}_\ell \\ \mathbf{V}_\ell & -\mathbf{V}_\ell\mathbf{X}_\ell \end{bmatrix}\begin{bmatrix} \mathbf{W}_\ell\mathbf{X}_\ell & \mathbf{W}_\ell \\ \mathbf{V}_\ell\mathbf{X}_\ell & -\mathbf{V}_\ell \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{I} \\ j\mathbf{Z}_{\mathrm{II}} \end{bmatrix}\mathbf{T}. \quad (27)$$

Again, Eq. (27), which is analogous to Eq. (8), includes one matrix inversion for each grating layer. For any one grating layer, if one or more of the quantities $\gamma_{\ell,m}d_\ell$ is large and positive real, the associated exponential $\exp(-q_{\ell,m}d_\ell)$ in the diagonal matrix $\mathbf{X}_\ell$ will be very small and possibly numerically zero. As a result, all the elements in one or more columns will be virtually zero. The inversion of this almost singular matrix will produce either numerical failure or erroneous results because of truncation errors. Even if the matrix can be analytically inverted, some of the elements of the inverted matrix cannot be represented with sufficient numerical accuracy because of numerical truncation in the finite-precision numerical representation, which results in numerical instability. Reducing the individual grating layer thickness by increasing the number of layers may temporarily relieve this condition by improving the numerical accuracy (reducing the magnitude of the exponent). However, extremely thin grating layers (a very large number of layers) must be used because some of the eigenvalues will be large when a relatively moderate to large number of space harmonics are retained. The largest possible real part of the exponent ($k_0 q d$) must be less than 36 for a computer with machine δ of 10^{-16} (δ is the smallest number that can be added to unity without a loss of precision). In addition to a significant increase in computational time, other forms of numerical round-off and truncation errors will be introduced.

5. NUMERICALLY STABLE ENHANCED TRANSMITTANCE APPROACH FOR SURFACE-RELIEF GRATINGS

Formulations for two numerically stable approaches are presented. The first approach permits the calculation of all the diffraction-reflected and -transmitted amplitudes. The second approach allows only the diffracted transmitted or the diffracted reflected amplitude to be calculated. This second, simpler and more efficient, approach is suitable for analyzing mainly transmissive or mainly reflective grating structures, for which only the forward or the backward diffraction must be determined.

6. FULL-SOLUTION APPROACH

In this section the full-solution approach is a numerically stable approach whereby all the transmitted and the reflected amplitudes are calculated. The approach to preempt the numerical instability associated with the matrix inversion is as follows. Consider the last factor ($\ell = L$) in Eq. (27):

$$\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L\mathbf{X}_L \end{bmatrix}\begin{bmatrix} \mathbf{W}_L\mathbf{X}_L & \mathbf{W}_L \\ \mathbf{V}_L\mathbf{X}_L & -\mathbf{V}_L \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{f}_{L+1} \\ \mathbf{g}_{L+1} \end{bmatrix}\mathbf{T}$$
$$= \begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L\mathbf{X}_L \end{bmatrix}\begin{bmatrix} \mathbf{X}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L \\ \mathbf{V}_L & -\mathbf{V}_L \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{f}_{L+1} \\ \mathbf{g}_{L+1} \end{bmatrix}\mathbf{T}, \tag{28}$$

where $\mathbf{f}_{L+1} = \mathbf{I}$ and $\mathbf{g}_{L+1} = j\mathbf{Z}_{\mathrm{II}}$. The matrix to be inverted has been rewritten as the product of two matrices. The matrix on the right in the product is well conditioned, and its inversion its numerically stable. However, it is advisable to use the single-value decomposition technique when numerically degenerate eigenvalues and eigenvectors are suspected (e.g., a large number of harmonics and a large number of equations). The right-hand side of Eq. (28) then reduces to

$$\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L\mathbf{X}_L \end{bmatrix}\begin{bmatrix} \mathbf{X}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{a}_L \\ \mathbf{b}_L \end{bmatrix}\mathbf{T}, \tag{29}$$

where

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

As in the case of a stack of uniform layers, the remaining matrix to be inverted is ill conditioned when the elements of the diagonal matrix $\mathbf{X}_L$ are very small. Clearly, for this diagonal matrix we can perform the inversion analytically to obtain the inverted matrix accurately. However, as discussed above, the diagonal elements cannot be represented with sufficient numerical accuracy because of the finite precision and the truncation errors. We avoid this problem by introducing the substitution $\mathbf{T} = \mathbf{a}_L{}^{-1}\mathbf{X}_L\mathbf{T}_L$, and the last factor ($\ell = L$) in Eq. (27) is reduced to

$$\begin{bmatrix} \mathbf{f}_L \\ \mathbf{g}_L \end{bmatrix}\mathbf{T}_L = \begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L\mathbf{X}_L \end{bmatrix}\begin{bmatrix} \mathbf{I} \\ \mathbf{b}_L\mathbf{a}_L{}^{-1}\mathbf{X}_L \end{bmatrix}\mathbf{T}_L$$
$$= \begin{bmatrix} \mathbf{W}_L(\mathbf{I} + \mathbf{X}_L\mathbf{b}_L\mathbf{a}_L{}^{-1}\mathbf{X}_L) \\ \mathbf{V}_L(\mathbf{I} - \mathbf{X}_L\mathbf{b}_L\mathbf{a}_L{}^{-1}\mathbf{X}_L) \end{bmatrix}\mathbf{T}_L. \tag{30}$$

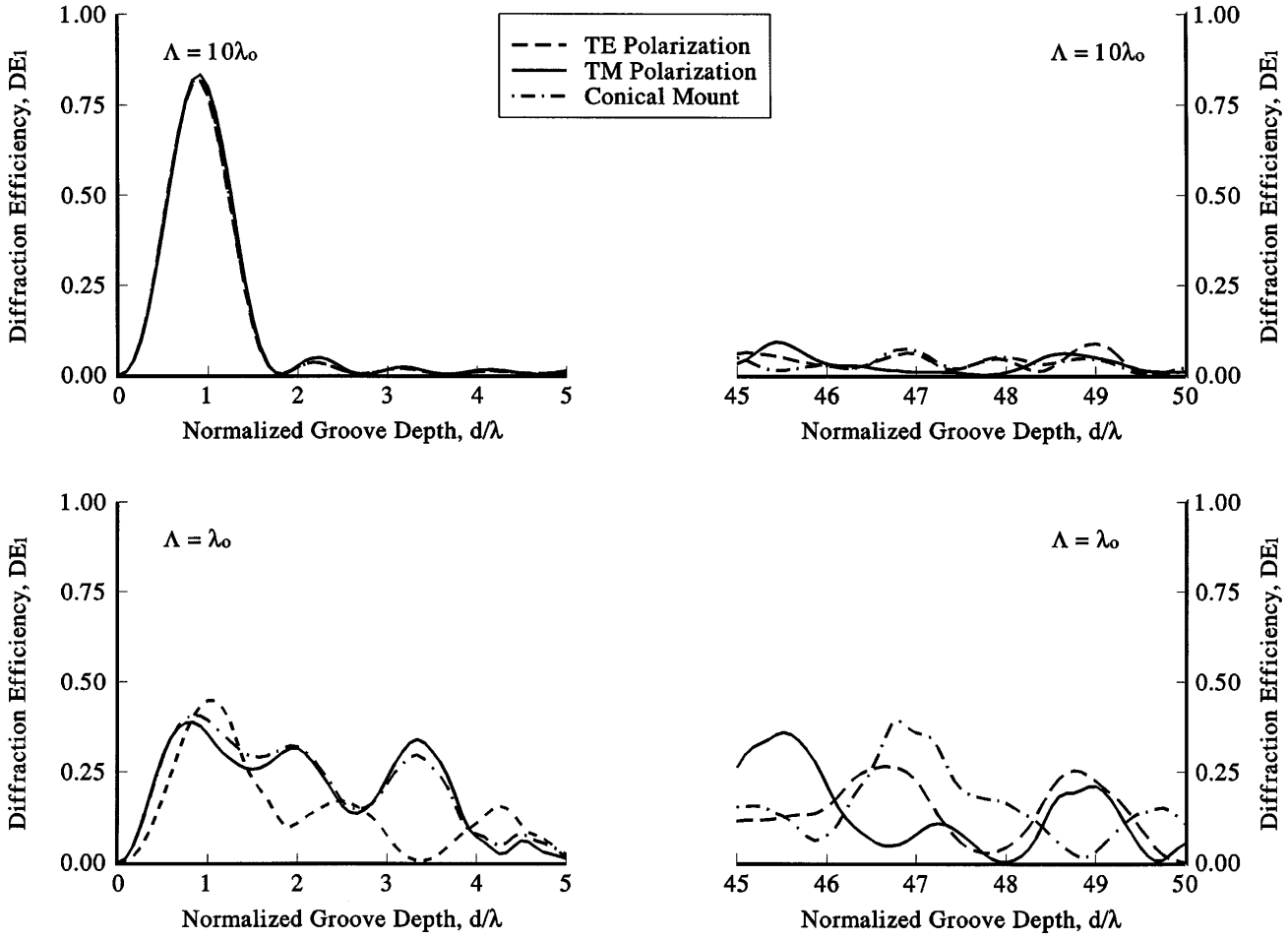

Fig. 3. Diffraction efficiency dependency on the normalized grating of a 16-level (15 layers) asymmetric binary dielectric grating ($n_g = n_{\mathrm{II}} = 2.04$, $n_{\mathrm{I}} = 1$). The angle of incidence is $\theta = 10°$. TE-polarization, TM-polarization, and conical-mount results are shown for two grating periods of 1 and 10 wavelengths, respectively.

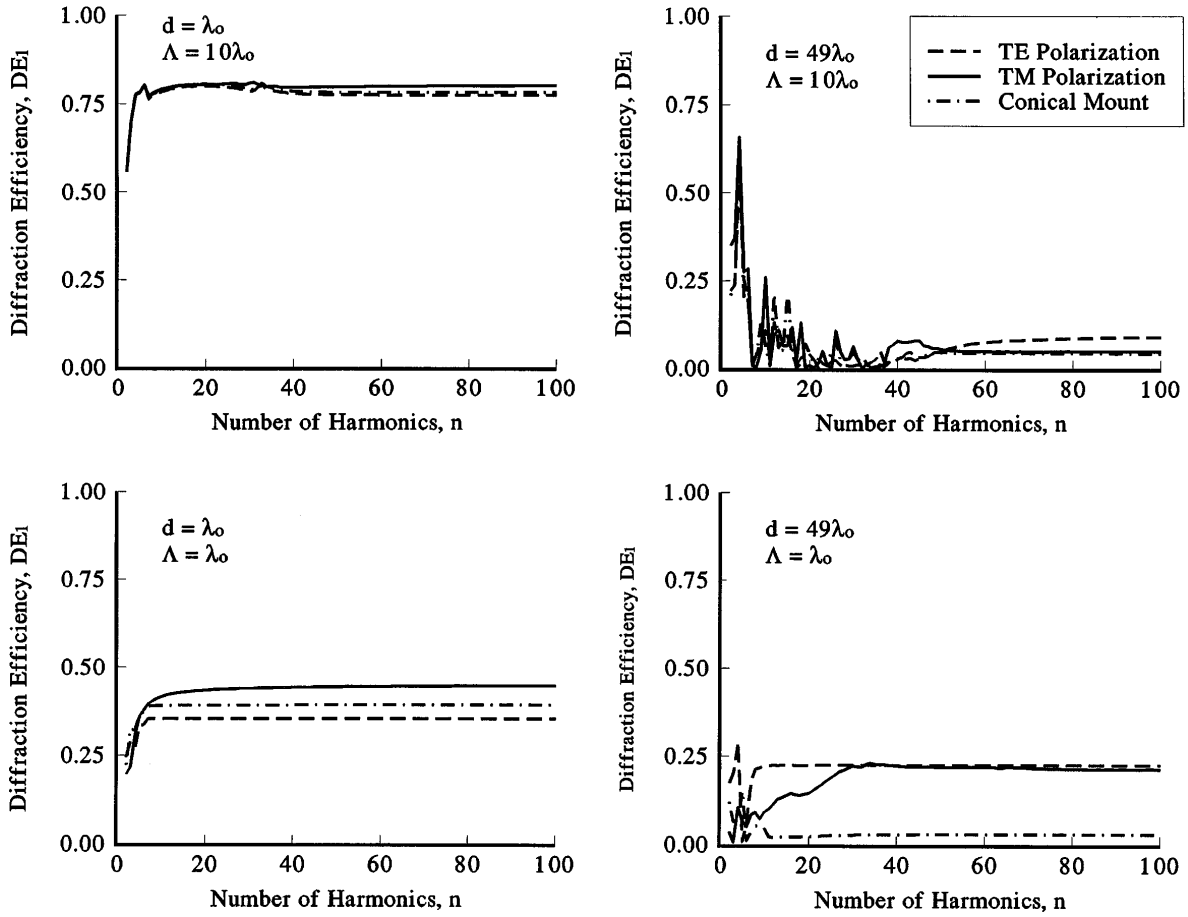

Fig. 4. Diffraction efficiency dependency on the number of space harmonics for the grating shown in Fig. 3 for two values of the grating depth (1 and 49 wavelengths, respectively) and for two grating periods of 1 and 10 wavelengths, respectively).

Repeating the process for all the layer(s), we obtain an equation of the form

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos(\theta)/n_{\mathrm{I}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Z}_{\mathrm{I}} \end{bmatrix} \mathbf{R} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{g}_1 \end{bmatrix} \mathbf{T}_1, \tag{31}$$

where

$$\mathbf{T} = \mathbf{a}_L{}^{-1}\mathbf{X}_L \cdots \mathbf{a}_\ell{}^{-1}\mathbf{X}_\ell \cdots \mathbf{a}_1{}^{-1}\mathbf{X}_1\mathbf{T}_1. \tag{32}$$

The quantities $\mathbf{a}_\ell$, $\mathbf{b}_\ell$, $\mathbf{f}_\ell$, and $\mathbf{g}_\ell$ are obtained by a process similar to expressions (29) and (30). Equation (31) is easily solved for R_i and T_i without numerical instability for any number of layers. The single-value decomposition technique should be considered in inverting the matrix $\mathbf{a}$ to avoid numerical difficulties because of round-off errors when a large number of layers and a large number of harmonics are used.

7. PARTIAL-SOLUTION APPROACH

In this section we present a second, simpler and more efficient, numerically stable transmittance approach whereby either the diffracted reflected or the diffracted transmitted amplitudes (but not both) are determined. The following formulation is for calculating the reflected amplitudes. Equation (26) (field matching at $L + 1$ boundary) may be rewritten as

$$\begin{bmatrix} -\mathbf{W}_L & \mathbf{f}_{L+1} \\ \mathbf{V}_L & \mathbf{g}_{L+1} \end{bmatrix} \begin{bmatrix} \mathbf{c}_L{}^- \\ \mathbf{c}_{L+1}{}^+ \end{bmatrix} = \begin{bmatrix} \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L\mathbf{X}_L \end{bmatrix} \mathbf{C}_L{}^+, \tag{33}$$

where $\mathbf{C}_{L+1}{}^+ \equiv \mathbf{T}$, $\mathbf{f}_{L+1} = \mathbf{I}$, and $\mathbf{g}_{L+1} = j\mathbf{Z}_{\mathrm{II}}$. From Eq. (33) the relationship between and $\mathbf{C}_L{}^-$ and $\mathbf{C}_L{}^+$ is given by

$$\mathbf{C}_L{}^- = \mathbf{a}_L\mathbf{C}_L{}^+, \tag{34}$$

where

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

The matrix inversion is numerically stable because the matrix is well conditioned. The matrix $\mathbf{b}_L$ is not used. From the latter equation the fields just inside the L layer at the $\ell = L$ boundary, given by Eq. (26), are

$$\begin{bmatrix} \mathbf{W}_L & \mathbf{W}_L\mathbf{X}_L \\ \mathbf{V}_L & -\mathbf{V}_L\mathbf{X}_L \end{bmatrix} \begin{bmatrix} \mathbf{c}_L{}^+ \\ \mathbf{c}_L{}^- \end{bmatrix} = \begin{bmatrix} \mathbf{f}_L \\ \mathbf{g}_L \end{bmatrix} \mathbf{c}_L{}^+, \tag{35}$$

where

$$\begin{bmatrix} \mathbf{f}_L \\ \mathbf{g}_L \end{bmatrix} = \begin{bmatrix} \mathbf{W}_L(\mathbf{I} + \mathbf{X}_L\mathbf{a}_L) \\ \mathbf{V}_L(\mathbf{I} - \mathbf{X}_L\mathbf{a}_L) \end{bmatrix}.$$

From Eq. (25) the fields just inside the $L - 1$ layer at the $\ell = L$ boundary are then given by

$$\begin{bmatrix} -\mathbf{W}_{L-1} & \mathbf{f}_L \\ \mathbf{V}_{L-1} & \mathbf{g}_L \end{bmatrix} \begin{bmatrix} \mathbf{c}_{L-1}{}^- \\ \mathbf{c}_L{}^+ \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{L-1} & \mathbf{X}_{L-1} \\ \mathbf{V}_{L-1} & \mathbf{X}_{L-1} \end{bmatrix} \mathbf{c}_{L-1}{}^+. \quad (36)$$

Repeating the process [Eqs. (34)–(36)] for the remaining number of layers, we obtain

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos(\theta)/n_{\mathrm{I}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Z}_{\mathrm{I}} \end{bmatrix} \mathbf{R} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{g}_1 \end{bmatrix} \mathbf{C}_1{}^+, \quad (37)$$

where the quantities $\mathbf{f}_1$ and $\mathbf{g}_1$ are defined in Eq. (35). Equation (37) is easily solved for R_i without numerical stability. This simpler approach is much more efficient than the full-solution approach, with approximately half as many matrix multiplications. In addition, only one $(2n \times 2n)$ matrix inversion is performed (efficient inversion by partition may be used because only the top half of the inverted matrix is required). However, only the reflected amplitudes are obtained. The formulation for calculating the transmitted amplitudes, whereby the matching process starts at the input side with fields propagating toward the substrate, is very similar.

8. NUMERICAL STABILITY AND CONVERGENCE

To illustrate the stability of the technique, in Fig. 3 the diffraction efficiency of the first diffracted order is plotted versus the normalized grating depth (the ratio of the full grating depth to the light wavelength) for a 16-level asymmetric dielectric grating ($n = 2.04$ and $n_{\mathrm{I}} = n_{\mathrm{II}} = 1.0$) up to excessive depths of 50 wavelengths for two grating periods of 1 and 10 wavelengths, respectively. The asymmetric grating is a sawtoothlike 15-layer stairstep with a step width of 1/16 of the grating period and a layer depth of 1/15 of the total depth of the structure. The diffraction efficiency is shown for TE and TM polarization and for conical diffraction (azimuth angle, $\phi = 30°$; polarization angle between the incident electric field and the plane of incidence, $\psi = 45°$). A sufficient number of terms are retained in the space-harmonic expansions to ensure accuracy to four places past the decimal (convergence of the diffraction efficiency with an increasing number of space harmonics is shown in Fig. 4). Conservation of energy has always been satisfied, to within one part in 10^{10}, even for extremely deep gratings. Conservation of energy is defined by

$$\sum_i (\mathrm{DE}_{ri} + \mathrm{DE}_{ti}) = 1 \quad (38)$$

and is a necessary criterion for the numerical stability (nonfailure) of the algorithm. It is important to bear in mind that satisfying the conservation-of-power condition does not ensure the accuracy of the diffraction efficiency for each diffracted order. The accuracy of the individual diffraction efficiencies depends on the number of space harmonics retained in the field expansions.

Figure 4 illustrates the convergence of the diffraction efficiency of the 16-level, asymmetric dielectric grating shown in Fig. 3 as the number of field harmonics is increased. Results are given for two grating depths (1 and 49 wavelengths, respectively) and for two grating periods (1 and 10 wavelengths, respectively) for both TE and TM polarization and for conical mounting. It is clear that, in all cases, the diffraction efficiency converges to the proper values when a sufficient number of harmonics are included in the formulation. Note that the TE polarization requires fewer harmonics than are required by the conical diffraction and by the TM polarization. Also, more harmonics are required for deeper gratings and for gratings with larger grating periods.

9. SUMMARY AND CONCLUSIONS

An enhanced, numerically stable transmittance matrix approach was developed and was applied to the implementation of the rigorous coupled-wave analysis for surface-relief and multilevel gratings. This new approach was shown to produce numerically stable results for excessively deep multilevel surface-relief dielectric gratings. Two numerically stable approaches were presented. In the first approach all the diffraction-reflected and -transmitted amplitudes are calculated. In the second approach only the diffraction-transmitted or the diffraction-reflected amplitude is calculated. This second, simpler and more efficient, approach is suitable for analyzing mostly transmissive or mostly reflective grating structures, for which only the forward or the backward diffraction is required.

We developed the enhanced transmittance approach by first examining the reflection and the transmission from a stack of uniform homogeneous layers to determine the nature and the source of possible numerical instability. The presence of evanescent fields results in an ill-conditioned transmittance matrix. Inverting this T matrix results in a situation in which the elements of the inverted matrix cannot be represented with sufficient numerical accuracy because of truncation errors associated with the finite precision of the numerical representation of digital computers. These inaccuracies will result in numerical instability in the calculations for successive field matching between the layers.

The new technique anticipates and preempts these potential numerical problems. It was found that the nature of the numerical instability observed in the diffraction analysis of surface-relief and multilevel gratings analysis is identical to that of the stack of uniform layers (reflection–transmission problem). The new approach, when incorporated into the implementation of the rigorous coupled-wave analysis, was shown to produce numerically stable and convergent results for excessively deep (50 wavelengths) 16-level, asymmetric binary gratings. Calculated results were presented for TE and TM polarization and for conical diffraction. The calculated diffraction efficiencies converge to the appropriate value, with an increasing number of space harmonics. The proposed technique is also applicable to other problems involving wave propagation through cascaded layer systems. Specifically, it is applicable to slanted periodic structures (within the grating layer) in which the eigenvalues are not in positive and negative pairs.

REFERENCES

1. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," J. Opt. Soc. Am. **71,** 811–818 (1981).

2. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar grating diffraction—E-mode polarization and losses," J. Opt. Soc. Am. **73,** 451–455 (1983).
3. M. G. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings," J. Opt. Soc. Am. **72,** 1385–1392 (1982).
4. M. G. Moharam and T. K. Gaylord, "Three-dimensional vector coupled-wave analysis of planar grating diffraction," J. Opt. Soc. Am. **73,** 1105–1112 (1983).
5. W. E. Baird, M. G. Moharam, and T. K. Gaylord, "Diffraction characteristics of planar absorption gratings," Appl. Phys. B **32,** 15–20 (1983).
6. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of metallic surface-relief gratings," J. Opt. Soc. Am. A **3,** 1780–1796 (1986).
7. M. G. Moharam, "Diffraction analysis of multiplexed holographic gratings," in *Digest of Topical Meeting on Holography* (Optical Society of America, Washington, D.C., 1986), pp. 100–103.
8. M. G. Moharam, "Coupled-wave analysis of two-dimensional gratings," in *Holographic Optics: Design and Applications,* I. Cindrich, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **883,** 8–11 (1988).
9. E. N. Glytsis and T. K. Gaylord, "Rigorous three-dimensional coupled-wave diffraction analysis of single and cascaded anisotropic gratings," J. Opt. Soc. Am. A **4,** 2061–2080 (1987).
10. L. M. Brekhovshikh, *Waves in Layered Media* (Academic, New York, 1960).
11. D. M. Pai and K. A. Awada, "Analysis of dielectric gratings of arbitrary profiles and thickness," J. Opt. Soc. Am. A **8,** 755–762 (1991).
12. S. T. Han, Y.-L. Tsao, R. M. Walser, and M. F. Becker, "Electromagnetic scattering of two-dimensional surface-relief dielectric gratings," Appl. Opt. **31,** 2343–2352 (1992).
13. L. Li, "Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity," J. Opt. Soc. Am. A **10,** 2581–2591 (1993).
14. N. Chateau and J.-P. Hugonin, "Algorithm for the rigorous coupled-wave analysis of grating diffraction," J. Opt. Soc. Am. A **11,** 1321–1331 (1994).
15. C. Schwartz and L. F. DeSandre, "New calculational approach for multilayer stacks," Appl. Opt. **26,** 3140–3144 (1987).
16. D. Y. K. Ko and J. R. Sambles, "Scattering matrix method for propagation of radiation in stratified media: attenuated reflection studies of liquid crystals," J. Opt. Soc. Am. A **5,** 1863–1866 (1988).
17. A. K. Cousins and S. C. Gottschalk, "Application of the impedance formalism to diffraction grating with multiple coating layers," Appl. Opt. **29,** 4268–4271 (1990).
18. M. G. Moharam, E. B. Grann, D. A. Pommet, and T. K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings," J. Opt. Soc. Am. A **12,** (1995).