Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings

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

Center for Research and Education in Optics and Lasers, University of Central Florida, Orlando, Florida 32816

T. K. Gaylord

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

Received August 24, 1994; accepted October 12, 1994; revised manuscript received November 7, 1994

The rigorous coupled-wave analysis technique for describing the diffraction of electromagnetic waves by periodic grating structures is reviewed. Formulations for a stable and efficient numerical implementation of the analysis technique are presented for one-dimensional binary gratings for both TE and TM polarization and for the general case of conical diffraction. It is shown that by exploitation of the symmetry of the diffraction problem a very efficient formulation, with up to an order-of-magnitude improvement in the numerical efficiency, is produced. The rigorous coupled-wave analysis is shown to be inherently stable. The sources of potential numerical problems associated with underflow and overflow, inherent in digital calculations, are presented. A formulation that anticipates and preempts these instability problems is presented. The calculated diffraction efficiencies for dielectric gratings are shown to converge to the correct value with an increasing number of space harmonics over a wide range of parameters, including very deep gratings. The effect of the number of harmonics on the convergence of the diffraction efficiencies is investigated. More field harmonics are shown to be required for the convergence of gratings with larger grating periods, deeper gratings, TM polarization, and conical diffraction.

1. INTRODUCTION

Over the past 10 years the rigorous coupled-wave analysis (RCWA) has been the most widely used method for the accurate analysis of the diffraction of electromagnetic waves by periodic structures. It has been used successfully and accurately to analyze both holographic and surface-relief grating structures. It has been formulated to analyze transmission and reflection planar dielectric–absorption holographic gratings, arbitrary profiled dielectric–metallic surface-relief gratings, multiplexed holographic gratings, two-dimensional surface-relief gratings, and anisotropic gratings for both planar and conical diffraction.[1–9]

The RCWA is a relatively straightforward technique for obtaining the exact solution of Maxwell's equations for the electromagnetic diffraction by grating structures. It is a noniterative, deterministic technique utilizing a state-variable method that converges to the proper solution without inherent numerical instabilities. The accuracy of the solution obtained depends solely on the number of terms in the field space-harmonic expansion, with conservation of energy always being satisfied.

Our purpose in this paper is to present a detailed review of the RCWA and to provide a step-by-step guide for its efficient and stable implementation. A simple compact formulation for the efficient and stable numerical implementation of the RCWA for one-dimensional, rectangular-groove binary surface-relief dielectric gratings is presented. Formulations for TE and TM polarization and for the conical-diffraction configuration are included. It is shown that a very efficient formulation, with up to an order-of-magnitude improvement in the numerical efficiency, can be achieved by exploitation of the symmetry of the diffraction problem. The technique is shown to be fundamentally stable. The criteria for numerical stability are (1) energy conservation and (2) convergence to the proper solution with an increasing number of field harmonics for all the grating and the incident-wave parameters. Potential numerical difficulties can be preempted by proper formulation and normalization. Specifically, the nonpropagating evanescent space harmonics in the grating region must be properly handled in the numerical implementation. The effect of the number of terms in the field space-harmonic expansion on the convergence of the diffraction efficiency is investigated. It is shown that for dielectric gratings, even very deep gratings, the calculated diffraction efficiencies always converge to the correct value as the number of space harmonics increases. As expected, more field space harmonics are required for the convergence of gratings with larger grating periods, deeper gratings, TM polarization, and conical diffraction.

2. FORMULATION

The general three-dimensional binary grating diffraction problem is depicted in Fig. 1. A linearly polarized electromagnetic wave is obliquely incident at an arbitrary angle of incidence θ and at an azimuthal angle ϕ upon a binary 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 for-

0740-3232/95/041068-09$06.00

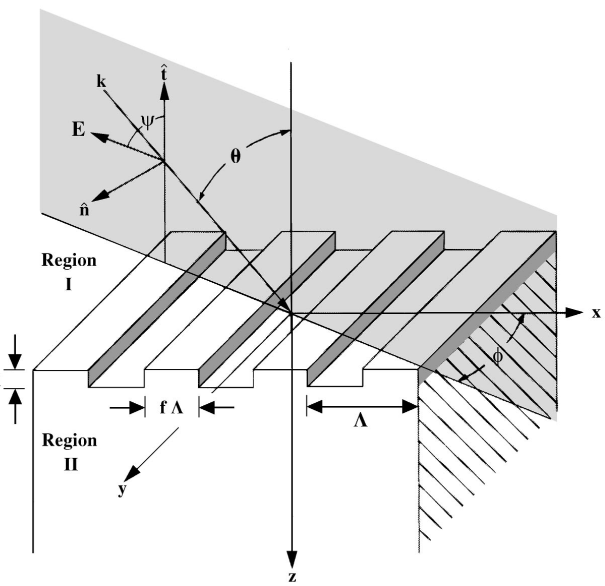

Fig. 1. Geometry for the binary rectangular-groove grating diffraction problem analyzed herein.

mulation 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 ($0 < z < d$) the periodic relative permittivity is expandable in a Fourier series of the form

$$\varepsilon(x) = \sum_h \varepsilon_h \exp\left(j\,\frac{2\pi h}{\Lambda}\right), \tag{1}$$

where ε_h is the hth Fourier component of the relative permittivity in the grating region, which is complex for lossy or nonsymmetric dielectric gratings. For simple grating structures with alternating regions of refractive indices n_{rd} (ridge) and n_{gr} (groove) the Fourier harmonics are given by

$$\varepsilon_0 = n_{\mathrm{rd}}{}^2 f + n_{\mathrm{gr}}{}^2(1-f), \qquad \varepsilon_h = (n_{\mathrm{rd}}{}^2 - n_{\mathrm{gr}}{}^2)\frac{\sin(\pi h f)}{\pi h}, \tag{2}$$

where f is the fraction of the grating period occupied by the region of index n_{rd} and ε_0 is the average value of the relative permittivity, not the permittivity of free space.

The general approach for solving the exact electromagnetic-boundary-value problem associated with the diffraction grating is to find solutions that satisfy Maxwell's equations in each of the three (input, grating, and output) regions and then match the tangential electric- and magnetic-field components at the two boundaries. For the case of planar diffraction ($\phi = 0$) the incident polarization may be decomposed into a TE- and a TM-polarization problem, which are handled independently. Here all the forward- and the backward-diffracted orders lie in the same plane (the plane of incidence, the x–z plane). For the general three-dimensional problem ($\phi \neq 0$), or conical diffraction, the wave vectors of the diffracted orders lie on the surface of a cone, and the perpendicular and the parallel components of the electric and the magnetic fields are coupled and must be obtained simultaneously. The three cases are considered separately in Sections 3–5.

3. PLANAR DIFFRACTION: TE POLARIZATION

The incident normalized electric field that is normal to the plane of incidence is given by

$$E_{\mathrm{inc},y} = \exp[-jk_0 n_{\mathrm{I}}(\sin\theta\, x + \cos\theta\, z)], \tag{3}$$

where $k_0 = 2\pi/\lambda_0$ and λ_0 is the wavelength of the light in free space. The normalized solutions in region I ($0 < z$) and in region II ($z > d$) are given by

$$E_{\mathrm{I},y} = E_{\mathrm{inc},y} + \sum_i R_i \exp[-j(k_{xi}x - k_{\mathrm{I},zi}z)], \tag{4}$$

$$E_{\mathrm{II},y} = \sum_i T_i \exp\{-j[k_{xi}x - k_{\mathrm{II},zi}(z-d)]\}, \tag{5}$$

where k_{xi} is determined from the Floquet condition and is given by

$$k_{xi} = k_0[n_{\mathrm{I}} \sin\theta - i(\lambda_0/\Lambda)] \tag{6}$$

and where

$$k_{\ell,zi} = \begin{cases} +k_0[n_\ell{}^2 - (k_{xi}/k_0)^2]^{1/2} & k_0 n_\ell > k_{xi} \\ -jk_0[(k_{xi}/k_0) - n_\ell{}^2]^{1/2} & k_{xi} > k_0 n_\ell \end{cases}, \quad \ell = \mathrm{I}, \mathrm{II}. \tag{7}$$

R_i is the normalized electric-field amplitude of the ith backward-diffracted (reflected) wave in region I. T_i is the normalized electric-field amplitude of the forward-diffracted (transmitted) wave in region II. The magnetic fields in regions I and II may be obtained from Maxwell's equation

$$H = \left(\frac{j}{\omega\mu}\right)\nabla \times E, \tag{8}$$

where μ is the permeability of the region and ω is the angular optical frequency.

In the grating region ($0 < z < d$) the tangential electric (y-component) and magnetic (x–component) fields may be expressed with a Fourier expansion in terms of the space-harmonic fields as

$$E_{gy} = \sum_i S_{yi}(z)\exp(-jk_{xi}x), \tag{9}$$

$$H_{gx} = -j\left(\frac{\epsilon_0}{\mu_0}\right)^{1/2} \sum_i U_{xi}(z)\exp(-jk_{xi}x), \tag{10}$$

where ϵ_0 is the permittivity of free space. $S_{yi}(z)$ and $U_{xi}(z)$ are the normalized amplitudes of the ith space-harmonic fields such that E_{gy} and H_{gx} satisfy Maxwell's equation in the grating region, i.e.,

$$\frac{\partial E_{gy}}{\partial z} = j\omega\mu_0 H_{gx}, \tag{11}$$

$$\frac{\partial H_{gx}}{\partial_z} = j\omega\epsilon_0\varepsilon(x)E_{gy} + \frac{\partial H_{gz}}{\partial_x}. \tag{12}$$

Substituting Eqs. (9) and (10) into Eqs. (11) and (12) and eliminating H_{gz}, we obtain the coupled-wave equations

$$\frac{\partial S_{yi}}{\partial z} = k_0 U_{xi},$$

$$\frac{\partial U_{xi}}{\partial z} = \left(\frac{k_{xi}{}^2}{k_0}\right)S_{yi} - k_0 \sum_p \varepsilon_{(i-p)} S_{yp}, \tag{13}$$

or, in matrix form,

$$\begin{bmatrix} \partial \mathbf{S}_y/\partial(z') \\ \partial \mathbf{U}_x/\partial(z') \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{S}_y \\ \mathbf{U}_x \end{bmatrix}, \tag{14}$$

which may be reduced to

$$[\partial^2 \mathbf{S}_y/\partial(z')^2] = [\mathbf{A}][\mathbf{S}_y], \tag{15}$$

where $z' = k_0 z$ and

$$\mathbf{A} = \mathbf{K}_x{}^2 - \mathbf{E}, \tag{16}$$

where $\mathbf{E}$ is the matrix formed by the permittivity harmonic components, with the i, p element being equal to $\varepsilon_{(i-p)}$; $\mathbf{K}_x$ is a diagonal matrix, with the i, i element being equal to k_{xi}/k_0; and $\mathbf{I}$ is the identity matrix. Note that $\mathbf{A}$, $\mathbf{K}_x$, and $\mathbf{E}$ are $(n \times n)$ matrices, where n is the number of space harmonics retained in the field expansion, with the ith row of the matrix corresponding to the ith space harmonic. The $(2n \times 2n)$ matrix in Eq. (14) thus becomes an $(n \times n)$ matrix in Eq. (15).

We solve the set of the coupled-wave equations by calculating the eigenvalues and the eigenvectors associated with the matrix $\mathbf{A}$. The simplification step taken from Eq. (14) to Eq. (15) effectively reduces the overall computational time of the eigenvalue problem by a factor of 8. Moreover, for symmetric gratings, the matrix $\mathbf{A}$ is symmetric for dielectric or Hermitian for lossy binary gratings. Hence a significant enhancement in the computational efficiency and a reduction in the computer memory requirement can be achieved by use of an appropriate eigenvalue software package. The space harmonics of the tangential electric and magnetic fields in the grating region are then given by

$$S_{yi}(z) = \sum_{m=1}^{n} w_{i,m}\{c_m{}^+ \exp(-k_0 q_m z) + c_m{}^- \exp[k_0 q_m (z-d)]\}, \tag{17}$$

$$U_{xi}(z) = \sum_{m=1}^{n} v_{i,m}\{-c_m{}^+ \exp(-k_0 q_m z) + c_m{}^- \exp[k_0 q_m (z-d)]\}, \tag{18}$$

where $w_{i,m}$ and q_m are the elements of the eigenvector matrix $\mathbf{W}$ and the positive square root of the eigenvalues of the matrix $\mathbf{A}$, respectively. The quantity $v_{i,m} = q_m w_{i,m}$ is the i, m element of the matrix $\mathbf{V} = \mathbf{WQ}$, where $\mathbf{Q}$ is a diagonal matrix with the elements q_m. The quantities $c_m{}^+$ and $c_m{}^-$ are unknown constants to be determined from the boundary conditions. Note that the exponential terms involving the positive square root of the eigenvalues are normalized to prevent possible numerical overflow, as is shown below.

We calculate the amplitudes of the diffracted fields R_i and T_i (together with $c_m{}^+$ and $c_m{}^-$) by matching the tangential electric- and magnetic-field components at the two boundaries. At the input boundary ($z = 0$)

$$\delta_{i0} + R_i = \sum_{m=1}^{n} w_{i,m}[c_m{}^+ + c_m{}^- \exp(-k_0 q_m d)], \tag{19}$$

$$j[n_\mathrm{I} \cos\theta\ \delta_{i0} - (k_{\mathrm{I},zi}/k_0)R_i] = \sum_{m=1}^{n} v_{i,m}[c_m{}^+ - c_m{}^- \exp(-k_0 q_m d)], \tag{20}$$

or, in matrix form,

$$\begin{bmatrix} \delta_{i0} \\ jn_\mathrm{I}\cos\theta\ \delta_{i0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y}_\mathrm{I} \end{bmatrix}[\mathbf{R}] = \begin{bmatrix} \mathbf{W} & \mathbf{WX} \\ \mathbf{V} & -\mathbf{VX} \end{bmatrix} \begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix}, \tag{21}$$

and at $z = d$

$$\sum_{m=1}^{n} w_{i,m}[c_m{}^+ \exp(-k_0 q_m d) + c_m{}^-] = T_i, \tag{22}$$

$$\sum_{m=1}^{n} v_{i,m}[c_m{}^+ \exp(-k_0 q_m d) - c_m{}^-] = j(k_{\mathrm{II},zi}/k_0)T_i, \tag{23}$$

or, in matrix form,

$$\begin{bmatrix} \mathbf{WX} & \mathbf{W} \\ \mathbf{VX} & -\mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ j\mathbf{Y}_\mathrm{II} \end{bmatrix}[\mathbf{T}], \tag{24}$$

where $\delta_{i0} = 1$ for $i = 0$ and $\delta_{i0} = 0$ for $i \neq 0$ and $\mathbf{X}$, $\mathbf{Y}_\mathrm{I}$, and $\mathbf{Y}_\mathrm{II}$ are diagonal matrices with the diagonal elements $\exp(-k_0 q_m d)$, $(k_{\mathrm{I},zi}/k_0)$, and $(k_{\mathrm{II},zi}/k_0)$, respectively. Equations (21) and (24) are solved simultaneously for the forward- and backward-diffracted amplitudes T_i and R_i. Numerical overflow is successfully preempted by the normalization process; i.e., at both boundaries [Eqs. (19)–(23)] the arguments of the exponential are always negative. One may significantly improve numerical efficiency by eliminating R_i from Eqs. (19) and (20) and T_i from Eqs. (22) and (23), solving the resulting set of equations for the $c_m{}^+$ coefficients, and then substituting these coefficients back into Eqs. (21) and (24) to calculate R_i and T_i. However, attempts to solve Eq. (24) for $c_m{}^+$ and $c_m{}^-$ in terms of T_i and then substitute for $c_m{}^+$ and $c_m{}^-$ in Eq. (21) to determine T_i and R_i will probably cause numerical errors. This is due to possible zero columns on the left-hand sides of Eqs. (21) and (24), which result from very small terms in the diagonal matrix $\mathbf{X}$ when some of the generally complex eigenvalues have a large positive real part. The diffraction efficiencies are defined as

$$\mathrm{DE}_{ri} = R_i R_i{}^* \,\mathrm{Re}\left(\frac{k_{\mathrm{I},zi}}{k_0 n_\mathrm{I} \cos\theta}\right),$$

$$\mathrm{DE}_{ti} = T_i T_i{}^* \,\mathrm{Re}\left(\frac{k_{\mathrm{II},zi}}{k_0 n_\mathrm{I} \cos\theta}\right). \tag{25}$$

The sum of the reflected and the transmitted diffraction efficiencies given by Eq. (25) must be unity for lossless gratings. This sum is independent of the number of space harmonics retained in the field expansion, which determines the accuracy of the individual diffracted orders.

4. PLANAR DIFFRACTION: TM POLARIZATION

The incident normalized magnetic field is normal to the plane of incidence and may be written as

$$H_{\mathrm{inc},y} = \exp[-jk_0 n_\mathrm{I}(\sin\theta\ x + \cos\theta\ z)]. \tag{26}$$

The normalized solutions in region I ($0 < z$) and region II ($z > d$) are given, respectively, by

$$H_{\mathrm{I},y} = H_{\mathrm{inc},y} + \sum_i R_i \exp[-j(k_{xi}x - k_{\mathrm{I},zi}z)], \quad (27)$$

$$H_{\mathrm{II},y} = \sum_i T_i \exp\{-j[k_{xi}x + k_{\mathrm{II},zi}(z-d)]\}, \quad (28)$$

where k_{xi}, $k_{\mathrm{I},zi}$, and $k_{\mathrm{II},zi}$ are defined as in Eqs. (6) and (7).

R_i is the normalized magnetic-field amplitude of the ith backward-diffracted (reflected) wave in region I. T_i is the normalized magnetic-field amplitude of the forward-diffracted (transmitted) wave in region II. The magnetic-field vectors in the two regions can be obtained from Maxwell's equation

$$\mathbf{E} = \left(\frac{-j}{\omega\epsilon_0 n^2}\right)\nabla \times \mathbf{H}. \quad (29)$$

In the modulated region ($0 < z < d$) the tangential magnetic (y-component) and electric (x-component) fields may be expressed as a Fourier expansion:

$$H_{gy} = \sum_i U_{yi}(z)\exp(-jk_{xi}x), \quad (30)$$

$$E_{gx} = j\left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} \sum_i S_{xi}(z)\exp(-jk_{xi}x), \quad (31)$$

where $U_{yi}(z)$ and $S_{xi}(z)$ are the normalized amplitudes of the ith space-harmonic fields such that H_{gy} and E_{gx} satisfy Maxwell's equation in the grating region, i.e.,

$$\frac{\partial H_{gy}}{\partial z} = -j\omega\epsilon_0\varepsilon(x)E_{gx}, \quad (32)$$

$$\frac{\partial E_{gx}}{\partial z} = -j\omega\mu_0 H_{gy} + \frac{\partial E_{gx}}{\partial x}. \quad (33)$$

Substituting Eqs. (30) and (31) into Eqs. (32) and (33) and eliminating H_{gz}, we find that the set of coupled-wave equations, in matrix form, is

$$\begin{bmatrix} \partial\mathbf{U}_y/\partial(z') \\ \partial\mathbf{S}_x/\partial(z') \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{E} \\ \mathbf{B} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{U}_y \\ \mathbf{S}_x \end{bmatrix}, \quad (34)$$

which may be reduced to

$$[\partial^2\mathbf{U}_y/\partial(z')^2] = [\mathbf{EB}][\mathbf{U}_y], \quad (35)$$

where

$$\mathbf{B} = \mathbf{K}_x\mathbf{E}^{-1}\mathbf{K}_x - \mathbf{I}, \quad (36)$$

with $\mathbf{E}$ and $\mathbf{K}_x$ being defined as in Eq. (16). As in the TE case, the above set of coupled-wave equations is solved by calculation of the eigenvalues and the eigenvectors associated with the $(n \times n)$ matrix $\mathbf{EB}$, where n is the number of harmonics retained in the field expansion. The $(2n \times 2n)$ matrix in Eq. (34) is reduced to an $(n \times n)$ matrix in Eq. (35), thus reducing the overall computational time of the eigenvalue problem by a factor of 8. The space harmonics of the tangential magnetic and electric fields are then given by

$$U_{yi}(z) = \sum_{m=1}^{n} w_{i,m}\{c_m{}^+ \exp(-k_0q_mz) + c_m{}^- \exp[k_0q_m(z-d)]\}, \quad (37)$$

$$S_{xi}(z) = \sum_{m=1}^{n} v_{i,m}\{-c_m{}^+ \exp(-k_0q_mz) + c_m{}^- \exp[k_0q_m(z-d)]\}, \quad (38)$$

where $w_{i,m}$ and q_m are the elements of the eigenvector matrix $\mathbf{W}$ and the positive square root of the eigenvalues of the matrix $\mathbf{EB}$, respectively. The quantities $v_{i,m}$ are the elements of the product matrix $\mathbf{V} = \mathbf{E}^{-1}\mathbf{WQ}$, with $\mathbf{Q}$ being a diagonal matrix with the diagonal elements q_m. The quantities $c_m{}^+$ and $c_m{}^-$ are unknown constants to be determined from the boundary conditions. Again, note that the exponential terms involving the positive square root of the eigenvalues are normalized so that potential numerical overflow is preempted.

As in the TE-polarization case, one calculates the amplitudes of the diffracted fields R_i and T_i (together with $c_m{}^-$ and $c_m{}^+$) by matching the tangential field components at the two boundaries. In matrix form the set of equations for tangential field matching at the input boundary ($z = 0$) is

$$\delta_{i0} + R_i = \sum_{m=1}^{n} w_{i,m}[c_m{}^+ + c_m{}^- \exp(-k_0q_md)], \quad (39)$$

$$j\left[\left(\frac{\cos\theta}{n_{\mathrm{I}}}\right)\delta_{i0} - \left(\frac{k_{\mathrm{I},zi}}{k_0{n_{\mathrm{I}}}^2}\right)R_i\right] = \sum_{m=1}^{n} v_{i,m}[c_m{}^+ - c_m{}^- \exp(-k_0q_md)], \quad (40)$$

or, in matrix form,

$$\begin{bmatrix} \delta_{i0} \\ j\delta_{i0}\cos\theta/n_{\mathrm{I}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -\mathbf{Z}_{\mathrm{I}} \end{bmatrix}[\mathbf{R}] = \begin{bmatrix} \mathbf{W} & \mathbf{WX} \\ \mathbf{V} & -\mathbf{VX} \end{bmatrix}\begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix}, \quad (41)$$

and at $z = d$

$$\sum_{m=1}^{n} w_{i,m}[c_m{}^+ \exp(-k_0q_md) + c_m{}^-] = T_i, \quad (42)$$

$$\sum_{m=1}^{n} v_{i,m}[c_m{}^+ \exp(-k_0q_md) + c_m{}^-] = j\left(\frac{k_{\mathrm{I},zi}}{k_0{n_{\mathrm{II}}}^2}\right)T_i, \quad (43)$$

or, in matrix form,

$$\begin{bmatrix} \mathbf{WX} & \mathbf{W} \\ \mathbf{VX} & -\mathbf{V} \end{bmatrix}\begin{bmatrix} \mathbf{c}^+ \\ \mathbf{c}^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ j\mathbf{Z}_{\mathrm{II}} \end{bmatrix}[\mathbf{T}], \quad (44)$$

where $\mathbf{X}$ is as defined previously and $\mathbf{Z}_{\mathrm{I}}$ and $\mathbf{Z}_{\mathrm{II}}$ are diagonal matrices with the diagonal elements $(k_{\mathrm{I},zi}/k_0{n_{\mathrm{I}}}^2)$ and $(k_{\mathrm{II},zi}/k_0{n_{\mathrm{II}}}^2)$, respectively.

Equations (41) and (44) are solved simultaneously for the forward- and the backward-diffracted amplitudes T_i and R_i. As in the TE-polarization case, one may significantly improve numerical efficiency by analytically eliminating R_i and T_i from Eqs. (41) and (44), solving the resulting set of equations for the $c_m{}^+$ coefficients, and then substituting the $c_m{}^+$ coefficients back into Eqs. (41) and (44) for R_i and T_i. However, as in the TE case, numerical problems will occur if one attempts to solve Eq. (44) for $c_m{}^+$ and $c_m{}^-$ in terms of T_i and then substitute into Eq. (41) to find T_i and R_i. The diffraction efficiencies are defined as

$$\mathrm{DE}_{ri} = R_i R_i^* \ \mathrm{Re}(k_{\mathrm{I},zi}/k_0 n_\mathrm{I} \cos\theta),$$
$$\mathrm{DE}_{ti} = T_i T_i^* \ \mathrm{Re}\left(\frac{k_{\mathrm{II},zi}}{n_{\mathrm{II}}^2}\right) \Bigg/ \left(\frac{k_0 \cos\theta}{n_\mathrm{I}}\right). \tag{45}$$

5. CONICAL DIFFRACTION

In region I the incident normalized electric-field vector is

$$\mathbf{E}_{\mathrm{inc}} = \mathbf{u} \exp[-jk_0 n_\mathrm{I}(\sin\theta\cos\phi\ x + \sin\theta\sin\phi\ y + \cos\theta\ z)], \tag{46}$$

where

$$\mathbf{u} = (\cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi)\hat{x} + (\cos\psi\cos\theta\cos\phi - \sin\psi\cos\phi)\hat{y} - \cos\psi\sin\theta\ \hat{z}, \tag{47}$$

where ψ is the angle between the electric-field vector and the plane of incidence. For $\psi = 0°$ and $\psi = 90°$ the magnetic and the electric fields, respectively, are perpendicular to the plane of incidence.

The normalized solutions in region I $(0 < z)$ and in region II $(z > d)$ are given by

$$\mathbf{E}_\mathrm{I} = \mathbf{E}_{\mathrm{inc}} + \sum_i \mathbf{R}_i \exp[-j(k_{xi}x + k_y y - k_{\mathrm{I},zi}z)], \tag{48}$$

$$\mathbf{E}_{\mathrm{II}} = \sum_i \mathbf{T}_i \exp\{-j[k_{xi}x + k_y y + k_{\mathrm{II},zi}(z-d)]\}, \tag{49}$$

where

$$k_{xi} = k_0[n_\mathrm{I}\sin\theta\cos\phi - i(\lambda_0/\Lambda)], \tag{50}$$

$$k_y = k_0 n_\mathrm{I}\sin\theta\sin\phi, \tag{51}$$

$$k_{\ell,zi} = \begin{cases} +[(k_0 n_\ell)^2 - k_{xi}^2 - k_y^2]^{1/2} & (k_{xi}^2 + k_y^2) < k_0 n_\ell \\ -j[k_{xi}^2 + k_y^2 - (k_0 n_\ell)^2]^{1/2} & (k_{xi}^2 + k_y^2)^{1/2} > k_0 n_\ell \end{cases}, \quad \ell = \mathrm{I}, \mathrm{II}. \tag{52}$$

$\mathbf{R}_i$ is the normalized vector electric-field amplitude of the ith backward-diffracted (reflected) wave in region I. $\mathbf{T}_i$ is the normalized electric-field vector amplitude of the forward-diffracted (transmitted) wave in region II. The magnetic-field vectors in region I and II can be obtained from Maxwell's equation (6). Note that the output plane of diffraction for the ith propagating diffraction order has an inclination angle given by

$$\varphi_i = \tan^{-1}(k_y/k_{xi}). \tag{53}$$

In the modulated region $(0 < z < d)$ the electric and the magnetic vector fields, $\mathbf{E}_g$ and $\mathbf{H}_g$, respectively, may be expressed as the Fourier expansion in terms of the space-harmonic fields as

$$\mathbf{E}_g = \sum_i [S_{xi}(z)\mathbf{x} + S_{yi}(z)\mathbf{y} + S_{zi}(z)\mathbf{z}]\exp[-j(k_{xi}x + k_{yy})], \tag{54}$$

$$\mathbf{H}_g = -j\left(\frac{\epsilon_0}{\mu_0}\right)^{1/2} \sum_i [U_{xi}(z)\mathbf{x} + U_{yi}(z)\mathbf{y} + U_{zi}(z)\mathbf{z}] \times \exp[-j(k_{xi}x + k_{yy})]. \tag{55}$$

$S_i(z)$ and $U_i(z)$ are the normalized vector amplitudes of the ith space-harmonic fields such that $\mathbf{E}_g$ and $\mathbf{H}_g$ satisfy Maxwell's equations in the grating region:

$$\nabla \times \mathbf{E}_g = -j\omega\mu_0\mathbf{H}_g,$$
$$\nabla \times \mathbf{H}_g = j\omega\epsilon_0\varepsilon(x)\mathbf{E}_g. \tag{56}$$

Substituting Eqs. (54) and (55) into Eqs. (56) and eliminating the normal components of the field (H_{gz} and E_{gz}), we obtain the set of coupled-wave equations in a matrix form:

$$\begin{bmatrix} \partial\mathbf{S}_y/\partial(z') \\ \partial\mathbf{S}_x/\partial(z') \\ \partial\mathbf{U}_y/\partial(z') \\ \partial\mathbf{U}_x/\partial(z') \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{K}_y\mathbf{E}^{-1}\mathbf{K}_x & \mathbf{I} - \mathbf{K}_y\mathbf{E}^{-1}\mathbf{K}_y \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_x\mathbf{E}^{-1}\mathbf{K}_x - \mathbf{I} & -\mathbf{K}_x\mathbf{E}^{-1}\mathbf{K}_y \\ \mathbf{K}_x\mathbf{K}_y & \mathbf{E} - \mathbf{K}_y^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_x^2 - \mathbf{E} & -\mathbf{K}_x\mathbf{K}_y & \mathbf{0} & \mathbf{0} \end{bmatrix} \times \begin{bmatrix} \mathbf{S}_y \\ \mathbf{S}_x \\ \mathbf{U}_y \\ \mathbf{U}_x \end{bmatrix}, \tag{57}$$

where $\mathbf{K}_y$ is a diagonal matrix with the elements (k_y/k_0) and $\mathbf{E}$ and $\mathbf{K}_x$ are as previously defined. Equation (57), a $(4n \times 4n)$ matrix, is reduced to either of the following two $(2n \times 2n)$ matrices:

$$\begin{bmatrix} \partial^2\mathbf{S}_y/\partial(z')^2 \\ \partial^2\mathbf{S}_x/\partial(z')^2 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_x^2 + \mathbf{DE} & \mathbf{K}_y[\mathbf{E}^{-1}\mathbf{K}_x\mathbf{E} - \mathbf{K}_x] \\ \mathbf{K}_x[\mathbf{E}^{-1}\mathbf{K}_y\mathbf{E} - \mathbf{K}_y] & \mathbf{K}_y^2 + \mathbf{BE} \end{bmatrix} \begin{bmatrix} \mathbf{S}_y \\ \mathbf{S}_x \end{bmatrix}, \tag{58}$$

$$\begin{bmatrix} \partial^2\mathbf{U}_y/\partial(z')^2 \\ \partial^2\mathbf{U}_x/\partial(z')^2 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_y^2 + \mathbf{EB} & [\mathbf{K}_x - \mathbf{EK}_x\mathbf{E}^{-1}]\mathbf{K}_y \\ [\mathbf{K}_y - \mathbf{EK}_y\mathbf{E}^{-1}]\mathbf{K}_x & \mathbf{K}_x^2 + \mathbf{ED} \end{bmatrix} \begin{bmatrix} \mathbf{U}_y \\ \mathbf{U}_x \end{bmatrix}. \tag{59}$$

The submatrix $\mathbf{B}$ is defined as in Eq. (36), and the submatrix $\mathbf{D} = \mathbf{K}_y\mathbf{E}^{-1}\mathbf{K}_y - \mathbf{I}$.

When the diagonal matrix $\mathbf{K}_y$ is a simple unity matrix (multiplied by a constant), as in the case for conical mount for a one-dimensional grating, further simplification is possible, and Eqs. (58) and (59) are reduced to two $(n \times n)$ matrices, respectively, of the form

$$[\partial^2\mathbf{U}_x/\partial(z')^2] = [k_y^2\mathbf{I} + \mathbf{A}][\mathbf{U}_x],$$
$$[\partial^2\mathbf{S}_x/\partial(z')^2] = [k_y^2\mathbf{I} + \mathbf{BE}][\mathbf{S}_x]. \tag{60}$$

The submatrix $\mathbf{A}$ is defined in Eq. (16). As in the TE or the TM case, one solves the above set of coupled-wave equations by calculating the eigenvalues and the eigenvectors associated with two $(n \times n)$ matrices, where n is the number of harmonics retained in the field expansion. The reduction of Eq. (57) to Eqs. (60) reduces the eigenvalues' and the eigenvectors' computational time by a factor of 32. For the two-dimensional grating diffraction problem, either Eq. (58) or Eq. (59) may be used to determine the eigenvalues and the eigenvectors for an improvement in the numerical efficiency of a factor of 4.

The space harmonics of the tangential magnetic and electric fields are given by

$$U_{xi}(z) = \sum_{m=1}^{n} w_{1,i,m}\{-c_{1,m}{}^{+} \exp(-k_0 w q_{1,m} z) + c_{1,m}{}^{-} \exp[k_0 q_{1,m}(z-d)]\}, \tag{61}$$

$$S_{xi}(z) = \sum_{m=1}^{n} w_{2,i,m}\{c_{2,m}{}^{+} \exp(-k_0 q_{2,m} z) + c_{2,m}{}^{-} \exp[k_0 q_{2,m}(z-d)]\}, \tag{62}$$

$$S_{yi}(z) = \sum_{m=1}^{n} v_{11,i,m}\{c_{1,m}{}^{+} \exp(-k_0 q_{1,m} z) + c_{1,m}{}^{-} \exp[k_0 q_{1,m}(z-d)]\} + \sum_{m=1}^{n} v_{12,i,m}\{c_{2,m}{}^{+} \exp(-k_0 q_{2,m} z) + c_{2,m}{}^{-} \exp[k_0 q_{2,m}(z-d)]\}, \tag{63}$$

$$U_{yi}(z) = \sum_{m=1}^{n} v_{11,i,m}\{c_{1,m}{}^{+} \exp(-k_0 q_{1,m} z) + c_{1,m}{}^{-} \exp[k_0 q_{1,m}(z-d)]\} + \sum_{m=1}^{n} v_{22,i,m}\{c_{2,m}{}^{+} \exp(-k_0 q_{2,m} z) + c_{2,m}{}^{-} \exp[k_0 q_{2,m}(z-d)]\}, \tag{64}$$

where $w_{1,i,m}$ and $q_{1,m}$ are the elements of the eigenvector matrix $\mathbf{W}_1$ and the positive square root of the eigenvalues of the matrix $[k_y{}^2\mathbf{I} + \mathbf{A}]$, respectively. The quantities $w_{2,i,m}$ and $q_{2,m}$ are the elements of the eigenvectors matrix $\mathbf{W}_2$ and the positive square root of the eigenvalues of the matrix $[k_y{}^2\mathbf{I} + \mathbf{BE}]$, respectively. The quantities $v_{11,i,m}$, $v_{12,i,m}$, $v_{21,i,m}$, and $v_{22,i,m}$ are the elements of the matrices $\mathbf{V}_{11}$, $\mathbf{V}_{12}$, $\mathbf{V}_{21}$, and $\mathbf{V}_{22}$ and are given by

$$\begin{aligned}
\mathbf{V}_{11} &= \mathbf{A}^{-1}\mathbf{W}_1\mathbf{Q}_1, \\
\mathbf{V}_{12} &= (k_y/k_0)\mathbf{A}^{-1}\mathbf{K}_x\mathbf{W}_2, \\
\mathbf{V}_{21} &= (k_y/k_0)\mathbf{B}^{-1}\mathbf{K}_x\mathbf{E}^{-1}\mathbf{W}_1, \\
\mathbf{V}_{22} &= \mathbf{B}^{-1}\mathbf{W}_2\mathbf{Q}_2,
\end{aligned} \tag{65}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are diagonal matrices with the diagonal elements $q_{1,m}$ and $q_{2,m}$, respectively. The quantities $c_{1,m}{}^{+}$, $c_{1,m}{}^{-}$, $c_{2,m}{}^{+}$, and $c_{2,m}{}^{-}$, are unknown constants, to be determined from the boundary conditions. Again note that the exponential terms involving the positive square root of the eigenvalues are normalized, so that numerical overflow is preempted.

As in the TE- and the TM-polarization cases, one calculates the amplitudes of the diffracted fields $\mathbf{R}_i$ and $\mathbf{T}_i$ (together with $c_m{}^{-}$ and $c_m{}^{+}$) by matching the tangential field components (rotated into the diffraction plane) at the two boundaries. At the input boundary ($z = 0$)

$$\sin\psi\delta_{i0} + R_{s,i} = \cos\varphi_i S_{yi}(0) - \sin\varphi_i S_{xi}(0), \tag{66}$$

$$j[\sin\psi\, n_{\mathrm{I}} \cos\theta - (k_{\mathrm{I},zi}/k_0)R_{s,i}] = -[\cos\varphi_i\, U_{xi}(0) + \sin\varphi_i\, U_{yi}(0)], \tag{67}$$

$$\cos\psi\cos\theta - j[k_{\mathrm{I},zi}/(k_0 n_{\mathrm{I}}{}^2)]R_{p,i} = \cos\varphi_i\, S_{xi}(0) + \sin\varphi_i\, S_{yi}(0), \tag{68}$$

$$-jn_{\mathrm{I}}\cos\psi + R_{p,i} = -[\cos\varphi_i\, U_{yi}(0) - \sin\varphi_i\, U_{xi}(0)], \tag{69}$$

where $R_{s,i}$ and $R_{p,i}$ are the components of the amplitude of the electric- and the magnetic-field vectors normal to the diffraction plane given by Eq. (53). They may be considered the TE and the TM components of the reflected diffracted field and are defined by

$$\begin{aligned}
R_{s,i} &= \cos\varphi_i\, R_{yi} - \sin\varphi_i\, R_{xi}, \\
R_{p,i} &= (-j/k_0)[\cos\varphi_i(ik_{\mathrm{I},zi}R_{xi} - k_{xi}R_{zi}) - \sin\varphi_i(k_y R_{zi} + k_{\mathrm{I},zi}R_{yii})].
\end{aligned} \tag{70}$$

Equations (66)–(69) may be rewritten in matrix form as

$$\begin{bmatrix} \sin\psi\,\delta_{i0} \\ j\sin\psi\, n_{\mathrm{I}}\cos\theta\,\delta_{i0} \\ -j\cos\psi\, n_{\mathrm{I}}\,\delta_{i0} \\ \cos\psi\cos\theta\,\delta_{i0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -j\mathbf{Y}_{\mathrm{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -j\mathbf{Z}_{\mathrm{I}} \end{bmatrix} \begin{bmatrix} \mathbf{R}_s \\ \mathbf{R}_p \end{bmatrix} = \begin{bmatrix} \mathbf{V}_{ss} & \mathbf{V}_{sp} & \mathbf{V}_{ss}\mathbf{X}_1 & \mathbf{V}_{sp}\mathbf{X}_2 \\ \mathbf{W}_{ss} & \mathbf{W}_{sp} & -\mathbf{W}_{ss}\mathbf{X}_1 & -\mathbf{W}_{sp}\mathbf{X}_2 \\ \mathbf{W}_{ps} & \mathbf{W}_{pp} & -\mathbf{W}_{ps}\mathbf{X}_1 & -\mathbf{W}_{pp}\mathbf{X}_2 \\ \mathbf{V}_{ps} & \mathbf{V}_{pp} & \mathbf{V}_{ps}\mathbf{X}_1 & \mathbf{V}_{pp}\mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1{}^{+} \\ \mathbf{c}_1{}^{-} \\ \mathbf{c}_2{}^{+} \\ \mathbf{c}_2{}^{-} \end{bmatrix}, \tag{71}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are diagonal matrices with the diagonal elements $\exp(-k_0 q_{1,m} d)$ and $\exp(-k_0 q_{2,m} d)$, respectively, and

$$\begin{aligned}
\mathbf{V}_{ss} &= \mathbf{F}_c\mathbf{V}_{11}, & \mathbf{W}_{pp} &= \mathbf{F}_c\mathbf{V}_{22}, \\
\mathbf{W}_{ss} &= \mathbf{F}_c\mathbf{W}_1 + \mathbf{F}_s\mathbf{V}_{21}, & \mathbf{V}_{pp} &= \mathbf{F}_c\mathbf{W}_2 + \mathbf{F}_s\mathbf{V}_{12}, \\
\mathbf{V}_{sp} &= \mathbf{F}_c\mathbf{V}_{12} - \mathbf{F}_c\mathbf{W}_2, & \mathbf{W}_{ps} &= \mathbf{F}_c\mathbf{V}_{21} - \mathbf{F}_s\mathbf{W}_1, \\
\mathbf{W}_{sp} &= \mathbf{F}_s\mathbf{V}_{22}, & \mathbf{V}_{ps} &= \mathbf{F}_s\mathbf{V}_{11},
\end{aligned} \tag{72}$$

with $\mathbf{F}_c$ and $\mathbf{F}_s$ being diagonal matrices with the diagonal elements $\cos\varphi_i$ and $\sin\varphi_i$, respectively. At $z = d$

$$\cos\varphi_i\, S_{yi}(d) - \sin\varphi_i\, S_{xi}(d) = T_{s,i}, \tag{73}$$

$$-[\cos\varphi_i\, U_{xi}(d) + \sin\varphi_i\, U_{yi}(d)] = j(k_{\mathrm{I},zi}/k_0)T_{s,i}, \tag{74}$$

$$-[\cos\varphi_i\, U_{yi}(d) + \sin\varphi_i\, U_{xi}(d)] = T_{p,i}, \tag{75}$$

$$\cos\varphi_i\, S_{xi}(d) + \sin\varphi_i\, S_{yi}(d) = j(k_{\mathrm{I},zi}/k_0 n_{\mathrm{I}}{}^2)T_{p,i}, \tag{76}$$

where $T_{s,i}$ and $T_{p,i}$ are the components of the amplitude of the electric- and the magnetic-field vectors normal to the diffraction plane given by Eq. (53). They may be considered the TE and the TM components of the transmitted diffracted field and are defined by

$$\begin{aligned} T_{s,i} &= \cos \varphi_i\, T_{yi} - \sin \varphi_i\, T_{xi}\,, \\ T_{p,i} &= (-j/k_0)[\cos \varphi_i(k_{\mathrm{II},zi}T_{xi} - k_{xi}T_{zi}) \\ &\quad - \sin \varphi_i(-k_{\mathrm{II},zi}T_{yi} + k_y T_{zi})]\,. \end{aligned} \tag{77}$$

Equations (73)–(76) may be rewritten in matrix form as

$$\begin{bmatrix} \mathbf{V}_{ss}\mathbf{X}_1 & \mathbf{V}_{sp}\mathbf{X}_2 & \mathbf{V}_{ss} & \mathbf{V}_{sp} \\ \mathbf{W}_{ss}\mathbf{X}_1 & \mathbf{W}_{sp}\mathbf{X}_2 & -\mathbf{W}_{ss} & -\mathbf{W}_{sp} \\ \mathbf{W}_{ps}\mathbf{X}_1 & \mathbf{W}_{pp}\mathbf{X}_2 & -\mathbf{W}_{ps} & -\mathbf{W}_{pp} \\ \mathbf{V}_{ps}\mathbf{X}_1 & \mathbf{V}_{pp}\mathbf{X}_2 & \mathbf{V}_{ps} & \mathbf{V}_{pp} \end{bmatrix} \begin{bmatrix} \mathbf{c}_1{}^+ \\ \mathbf{c}_1{}^- \\ \mathbf{c}_2{}^+ \\ \mathbf{c}_2{}^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ j\mathbf{Y}_{\mathrm{II}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & j\mathbf{Z}_{\mathrm{II}} \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix}. \tag{78}$$

Equations (71) and (78) are solved simultaneously for the forward- and the backward- diffracted amplitudes. As in the TE- or the TM-polarization case, one may significantly improve numerical efficiency by first analytically eliminating $\mathbf{R}_s$, $\mathbf{T}_s$, $\mathbf{R}_p$, and $\mathbf{T}_p$ from Eqs. (71) and (78). One then solves the resulting set of equations for the $c_m{}^{\pm}$ coefficients and then substitutes the $c_m{}^{\pm}$ coefficients back into the original equations for $\mathbf{R}_s$, $\mathbf{T}_s$, $\mathbf{R}_p$, and $\mathbf{T}_p$. Numerical problems will be encountered if one attempts to solve Eq. (78) for $c_m{}^+$ and $c_m{}^-$ in terms of $\mathbf{T}_s$ and $\mathbf{T}_p$ and then substitute the coefficients back into Eq. (71) for $\mathbf{R}_s$ and $\mathbf{R}_p$. The diffraction efficiencies are defined as

$$\begin{aligned} \mathrm{DE}_{ri} &= |R_{s,i}|^2 \operatorname{Re}\left(\frac{k_{\mathrm{I},zi}}{k_0 n_\mathrm{I} \cos\theta}\right) + |R_{p,i}|^2 \operatorname{Re}\left(\frac{k_{\mathrm{I},zi}/n_\mathrm{I}{}^2}{k_0 n_\mathrm{I} \cos\theta}\right), \\ \mathrm{DE}_{ti} &= |T_{s,i}|^2 \operatorname{Re}\left(\frac{k_{\mathrm{II},zi}}{k_0 n_\mathrm{I} \cos\theta}\right) + |T_{p,i}|^2 \operatorname{Re}\left(\frac{k_{\mathrm{II},zi}/n_\mathrm{II}{}^2}{k_0 n_\mathrm{I} \cos\theta}\right). \end{aligned} \tag{79}$$

6. NUMERICAL STABILITY

The criteria for numerical stability are (1) energy conservation and (2) convergence to the proper solution with an increasing number of field harmonics for all the grating and the incident-wave parameters.

7. CONSERVATION OF ENERGY

Conservation of energy for a lossless grating is defined as

$$\sum_i (\mathrm{DE}_{ri} + \mathrm{DE}_{ti}) = 1\,. \tag{80}$$

This condition should be achieved, to an accuracy of at least 1 part in 10^{10}, regardless of the number of terms

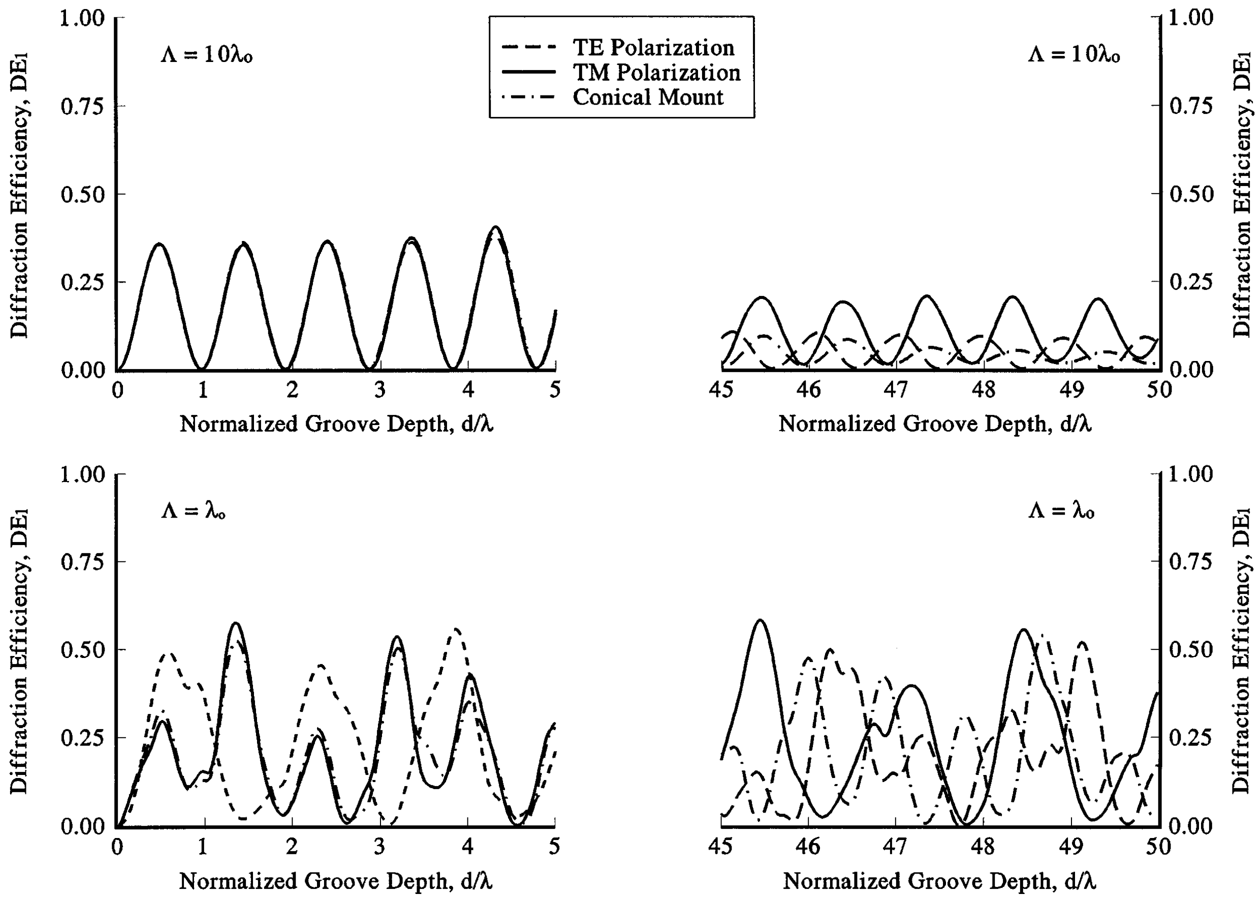

Fig. 2. Diffraction-efficiency dependence on the normalized grating depth of a binary dielectric grating ($n_{\mathrm{II}} = n_{\mathrm{rd}} = 2.04$, $n_{\mathrm{I}} = 1$) for TE polarization, TM polarization, and conical mount ($\phi = 30°$ and $\psi = 45°$) at $\theta = 10°$.

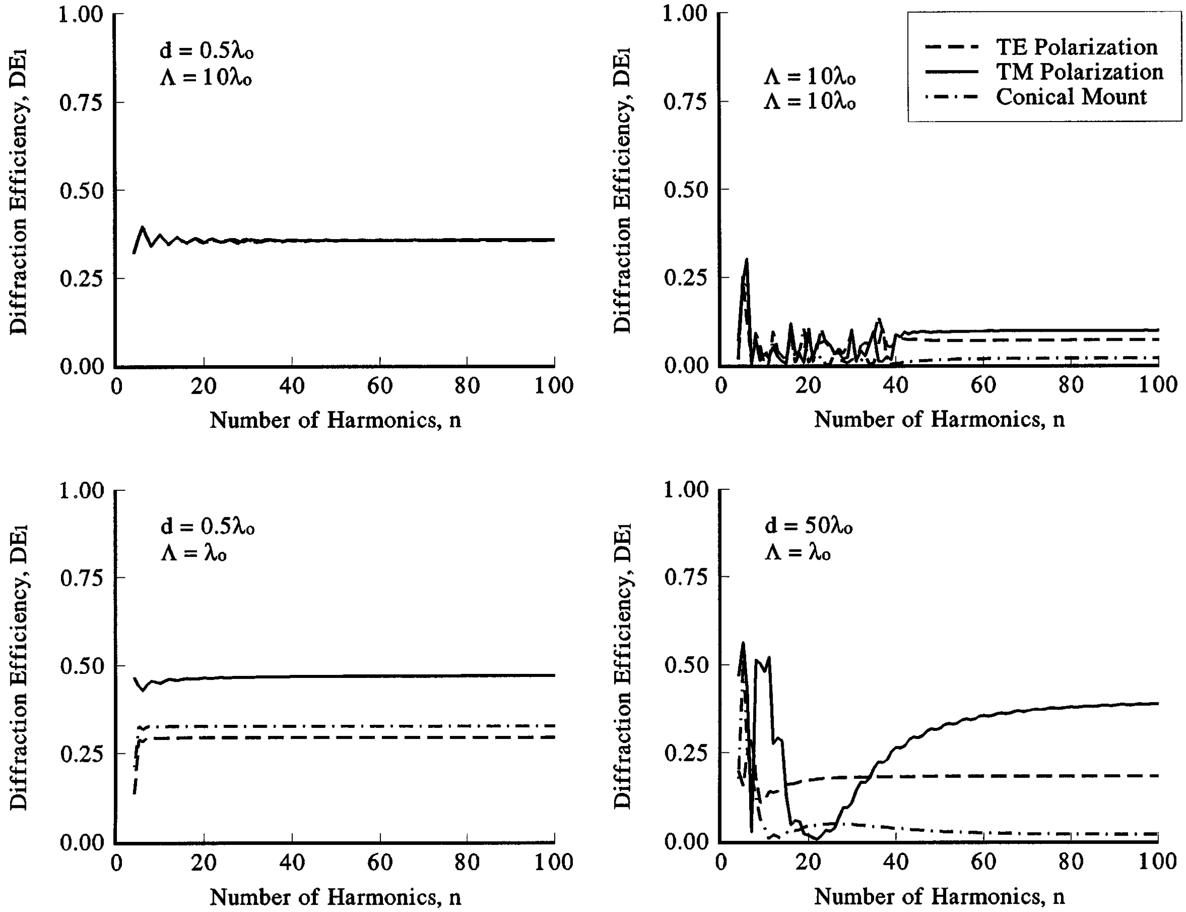

Fig. 3. Diffraction-efficiency dependence on the number of space harmonics for a binary dielectric grating ($n_{\mathrm{II}} = n_{\mathrm{rd}} = 2.04$) for TE polarization, TM polarization, and conical mount ($\phi = 30°$ and $\psi = 45°$) at $\theta = 10°$.

in the field-expansion series that are retained in the formulation. This condition is necessary but not sufficient for the success (nonfailure) of the numerical algorithm. Conservation of power does not ensure the accuracy of the diffraction efficiency for each diffracted order. The individual diffraction efficiency depends on the number of space harmonics retained in the field expansion, which is discussed in Section 8.

The primary source of potential numerical instabilities is complex eigenvalues with a large, positive real part. Without appropriate normalization of the terms involving $\exp(k_0 q_m d)$, a numerical overflow will occur. This unrecoverable numerical instability is preempted by the normalization utilized in Eqs. (17) and (18), (37) and (38), and (61)–(64). With this simple normalization the formulation will involve terms in $\exp(-k_0 q_m d)$ that could be yet another source of numerical instabilities. These terms may result in several zero columns in the matrices on the left-hand sides of Eqs. (21), (23), (41), (43), (71), and (78). Attempts to invert these matrices will result in either a numerical failure, because of numerical overflow, or erroneous results, because of large round-off errors. This numerical instability can easily be avoided by simultaneous solution of Eqs. (21) and (23) (in the TE-polarization case). One may achieve solutions that are more efficient and still stable by eliminating R_i and T_i from Eqs. (21) and (23), solving the two resulting sets of equations simultaneously for the $c_m^{\pm}$ coefficients, and then substituting these coefficients back into the original equations to calculate R_i and T_i. This method is applicable to TM polarization with Eqs. (41) and (43) and to the conical-diffraction case with Eqs. (71) and (78). Solutions for these sets of equations are always numerically stable, and even standard LU decomposition routines are normally sufficient for obtaining stable, accurate solutions. We have never encountered any numerical-instability problems in solving these sets of equations, even for very deep gratings. For extremely large matrices (retaining large number of harmonics) in which round-off errors might cause potential numerical problems, a **QR** decomposition routine may be used at the cost of some numerical efficiency.

It is important to note that the above method is not practical or suitable for removing the numerical instabilities in multilevel binary and surface-relief grating problems. This is due to the extremely large size of the system of equations, which will require prohibitive computational resources for implementation. A technique for removing these numerical instabilities is presented in a companion paper.[10]

To illustrate the stability of the present technique the diffraction efficiency of the first diffracted order is plotted versus the normalized grating depth (with respect to the light wavelength) for a dielectric binary grating

($n_{\mathrm{II}} = n_{rd} = 2.04$) up to extreme depths (see Fig. 2). The diffraction efficiency is shown for both TE and TM polarization and for conical diffraction ($\phi = 30°$ and $\psi = 45°$) for two values of the grating period. A sufficient number of terms are retained in the space-harmonics expansion to ensure accuracy to four places past the decimal. Conservation of energy is always achieved, to within 1 part in 10^{10}, even for extremely deep gratings. This is independent of the number of terms retained in the field expansion.

8. CONVERGENCE

As discussed in Section 6, the stable and efficient implementation of the RCWA, as described above, will always converge to yield the diffracted field amplitudes. The accuracy of the solution depends solely on the number of terms retained in the expansion of the space-harmonic fields in the grating region. The effects of incident polarization, including conical mounting diffraction, grating-period-to-wavelength ratios, and grating depth, on the number of field harmonics needed for the convergence of the diffraction efficiency is investigated. The convergence of the diffraction efficiency of a dielectric grating ($n_{\mathrm{II}} = 2.04$) as the number of field harmonics is increased is shown in Fig. 3. Results are shown for two normalized grating depths (1 and 50) and two normalized grating periods (1 and 10) for both TE and TM polarization and for the conical mounting diffraction ($\phi = 30°$ and $\psi = 45°$). 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 TE polarization requires fewer harmonics than does conical-diffraction TM polarization. Also, more harmonics are required for deeper gratings and for gratings with larger grating periods.

The numerical convergence investigation presented here is for a dielectric binary rectangular-groove grating. However, it was shown previously that the RCWA converges to the proper solution for metallic binary gratings.[11] Convergence for incident TE polarization is relatively efficient, requiring a small number of field harmonics. However, a significantly larger number of field harmonics are required for the case of TM polarization, and convergence is very slow.

9. SUMMARY

The RCWA technique for describing the diffraction of electromagnetic waves by periodic grating structures was reviewed. A detailed, step-by-step formulation for a stable and efficient numerical implementation of this analysis technique was presented for one-dimensional binary gratings for both TE and TM polarization and for the general case of conical diffraction. It was shown that a very efficient formulation, with up to an order-of-magnitude improvement in the numerical efficiency, can be achieved by exploitation of the symmetry of the diffraction problem. It was shown that the technique is inherently stable and that energy conservation and convergence to the proper solution with an increasing number of field harmonics are achieved with all the grating and the incident-wave parameters. Potential numerical difficulties can be preempted by proper formulation and normalization. Specifically, the nonpropagating evanescent space harmonics in the grating region must be properly handled in the numerical implementation, and the potential numerical underflow and overflow problems inherent in digital calculations must be anticipated and preempted. The effect of the number of harmonics on the convergence was investigated, and the calculated diffraction efficiencies for dielectric gratings were shown to converge to the correct value in each case as the number of space harmonics in the series expansion of the electromagnetic fields was increased. As expected, more field harmonics are required for the convergence of gratings with larger grating periods, deeper gratings, TM polarization, and conical diffraction.

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. M. G. Moharam, D. A. Pommet, E. B. Grann, and T. K. Gaylord, "Stable implementation of the rigorous coupled-wave analysis for surface-relief dielectric gratings: enhanced transmittance matrix approach," J. Opt. Soc. Am. A **12,** 1077–1086 (1995).
11. L. Li and C. W. Haggans, "Convergence of the coupled-wave method for metallic lamellar diffraction gratings," J. Opt. Soc. Am. A **10,** 1184–1187 (1993).