Differential theory: application to highly conducting gratings

Evgeny Popov

Institut Fresnel, Case 161, Unité Mixte de Recherche Associée au Centre National de la Recherche Scientifique (UMR 6133), Faculté des Sciences et Techniques de St.-Jérôme, Avenue Escadrille Normandie Niémen, 13397 Marseille Cedex 20, France

Boris Chernov

Optical and Quantum Radiophysics Department, State University of Telecommunications, St. Petersburg 191186, Russia

Michel Nevière and Nicolas Bonod

Institut Fresnel, Case 161, Unité Mixte de Recherche Associée au Centre National de la Recherche Scientifique (UMR 6133), Faculté des Sciences et Techniques de St.-Jérôme, Avenue Escadrille Normandie Niémen, 13397 Marseille Cedex 20, France

Received June 23, 2003; revised manuscript received September 12, 2003; accepted September 24, 2003

The recently developed fast Fourier factorization method, which has greatly improved the application range of the differential theory of gratings, suffers from numerical instability when applied to metallic gratings with very low losses. This occurs when the real part of the refractive index is small, in particular, smaller than 0.1–0.2, for example, when silver and gold gratings are analyzed in the infrared region. This failure can be attributed to Li's "inverse rule" [L. Li, J. Opt. Soc. Am. A **13**, 1870 (1996)] as shown by studying the condition number of matrices that have to be inverted. Two ways of overcoming the difficulty are explored: first, an additional truncation of the matrices containing the coefficients of the differential system, which reduces the numerical problems in some cases, and second, an introduction of lossier material inside the bulk, thus leaving only a thin layer of the highly conducting metal. If the layer is sufficiently thick, this does not change the optical properties of the system but significantly improves the convergence of the differential theory, including the rigorous coupled-wave method, for various types of grating profiles. © 2004 Optical Society of America

OCIS codes: 050.0050, 050.1940, 050.1950, 050.1960, 050.2770, 260.1960, 260.2110, 650.2770.

1. INTRODUCTION

The differential theory of gratings,[1–4] which was developed in the 1970s, experienced two important breakthroughs during the 1990s. The first consisted of the introduction of the S-matrix propagation algorithm,[5,6] which eliminated the numerical contamination during the integration process that was related to the presence of growing exponential functions. The second was the development of the fast Fourier factorization (FFF) method,[6–8] which greatly increased the speed of convergence of the truncated Fourier series of the field diffracted under arbitrary polarization conditions. The FFF method relies on factorization rules that were stated by Li.[9] The "direct rule" allows computation of the Fourier components of the product of a discontinuous function by a continuous one, and the "inverse rule" applies to the product of complementary discontinuous functions, i.e., functions whose product is continuous.

Numerical results[6,7] have shown the FFF method to be efficient with highly reflecting metals used in the visible. However, problems have recently been discovered with use of silver and gold gratings in the near-infrared region. A low rate of convergence of the series of the field was found when the refractive index n_1 of the grating material reached the value $0.1 + i10$. This is not an academic problem, since metallic gratings in the near infrared region have great potential for use in dense wavelength demultiplexing.

To illustrate this, in Fig. 1 we present the convergence of a sinusoidal metallic grating with period d equal to the groove depth $h = 0.5\,\mu$m, used in TM polarization at 45° incidence from air, with wavelength equal to 0.6328 μm. The convergence parameter N determines the number of Fourier components (equal to $2N + 1$) of the field, kept in the calculations. The real part of the refractive index of the material is progressively lowered from 0.3 to 0 starting with $n_1 = 0.3 + i9$ (thin solid curve in Fig. 1), then $n_1 = 0.1 + i10$ (thick solid curve), and finishing with a metal having an artificially chosen index $n_1 = 0 + i10$ (dashed curve). We recall that the electromagnetic losses are proportional to the imaginary part of the permittivity $\epsilon \equiv \epsilon_0\epsilon_r = \epsilon_0 n^2$, equal to the product of the vacuum permitivity ϵ_0 and the relative permittivity ϵ_r of the substance. In the last example the metal is lossless. As can be observed, even in the first example, presenting a relatively rapid convergence at $N = 13$, the results start to deteriorate for $N > 20$. The case then worsens with a decrease of $\mathrm{Re}(n_1)$.

1084-7529/2004/020199-08$15.00

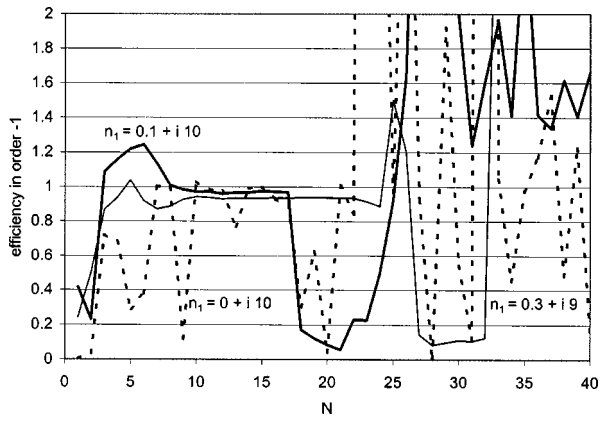

Fig. 1. Convergence with respect to the truncation parameter N of the -1st-order diffraction efficiency of a metallic sinusoidal grating with period d equal to the groove depth $h = 0.5\,\mu$m in TM polarized light with wavelength 0.6328 μm, incident from air at an angle of 45°. Calculations were made for three different complex refractive indices of the substrate with values shown in the figure.

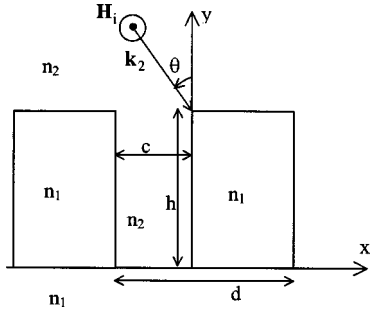

Fig. 2. Schematical representation of a lamellar grating.

In contrast, no such problems exist in TE polarization. Together with the fact that the numerical problems increase with a decrease of the real part of the grating material and are the worst when $\mathrm{Re}(n_1) = 0$, this suggests that the reason for the problem could lie in the validity of the inverse rule[9] applied to a value of permittivity ϵ_r that goes across the grating surface from the value of 1 in air to an almost real negative value in the metal. As shown by Li, this may cause the corresponding Toeplitz matrix to become almost singular.

To clarify the problem, we first choose to work on a lamellar grating profile illuminated in TM polarization for which various methods are available to check the results, namely, the modal method,[10–12] the rigorous coupled-wave (RCW) method,[13–16] and the differential theory.[6] We first show that the value $n_1 = 0 + i10$ produces bad condition numbers[17] for the matrix, which must be inverted when the inverse rule has to be used. Two ways to eliminate the numerical instabilities are then tried. The first one consists of truncating the matrix of the coefficients of the differential system to be integrated. The second one represents a physical solution and consists of replacing the bulk region of the highly conducting material with a not-so-highly-conducting metal. If an optically thick (typically 20–30 nm) surface layer of the original metal is preserved, this procedure does not modify the diffracted field significantly, and it removes the (quasi-) singularity of the corresponding Toeplitz matrix, which appears during application of the inverse rule. The advantage of the second method is that it can be generalized to an arbitrary profile, in particular for sinusoidal and triangular gratings, and can be extended to crossed gratings and three-dimensional periodic structures.

2. THE DIFFRACTION PROBLEM

Figure 2 shows a lamellar grating that will be studied in what follows. Its surface separates the space into two domains, with refractive indices $n_j\,(j = 1, 2)$. The superstrate is air ($n_2 = 1$), and the substrate is a good reflector (gold, silver, etc.). The groove period d, equal to the groove depth h, is equal to 0.5 μm. The grating is illuminated at angle of incidence θ, wavelength λ, and circular frequency ω, with an incident wave vector $\mathbf{k}_2$ lying in the cross-section plane xOy. The incident magnetic field is parallel to the z axis (TM polarization). We introduce the column vector $[E_x]$, which contains the pseudo-Fourier components of E_x, and similar notation for $[H_z]$, and F will denote the column vector containing the two column vectors $[E_x]$ and $[H_z]$. The FFF method leads to the following propagation equations[6]:

$$\frac{\mathrm{d}F}{\mathrm{d}y} = \mathbf{M}F,$$

where

$$\mathbf{M} = i\begin{bmatrix} 0 & \dfrac{\alpha}{\omega}[\![\epsilon]\!]^{-1}\alpha - \omega\mu \\ -\omega\left[\!\left[\dfrac{1}{\epsilon}\right]\!\right]^{-1} & 0 \end{bmatrix}. \tag{1}$$

The double square brackets represent a Toeplitz matrix formed by the corresponding Fourier components of the enclosed function, α is a diagonal matrix with elements α_n given by $\alpha_n = k_2 \sin\theta + n2\pi/d$, and n is integer.

Matrix $\mathbf{M}$ contains four submatrices $M_{i,j}\,(i, j = 1, 2)$, which are $(2N + 1) \times (2N + 1)$ matrices. For a lamellar profile, $\mathbf{M}$ is y independent, and the solution $F(y)$ can then be expressed in terms of the eigenvalues ρ_m of matrix $\mathbf{M}$, which represents the essence of the RCW method. If we designate by $\varphi(y)$ the diagonal matrix with elements $\exp(\rho_m y)\delta_{nm}$ and by $\mathbf{V}$ the square matrix made with the eigenvectors of $\mathbf{M}$ put in the same sequence as the sequence of eigenvalues on the diagonal of φ, it is well-known[6] that

$$F(y) = V\varphi(y)V^{-1}F(0). \tag{2}$$

The eigenvalue equation is

$$\det(\mathbf{M} - R\mathbf{I}) = 0, \tag{3}$$

where $R_{n,m} = \rho_m\delta_{n,m}$, and $\mathbf{I}$ is the unity matrix. To decrease the time calculation, it is judicious to resolve the following system:

$$M_{1,2}M_{2,1} - R^2\mathbf{I} = 0. \tag{4}$$

Introducing the matrix $\mathbf{U} = M_{1,2}M_{2,1}$, we obtain, using Eq. (1),

$$\mathbf{U} = \omega^2\mu\left[\!\left[\frac{1}{\epsilon}\right]\!\right]^{-1}\left(\mathbf{I} - \frac{\alpha}{\omega^2\mu}[\![\epsilon]\!]^{-1}\alpha\right). \tag{5}$$

Matrix $\mathbf{U}$ is a $(2N+1)\times(2N+1)$ matrix, and a numerical solution of the eigenvalue problem gives $(2N+1)$ eigenvalues τ_m and $(2N+1)$ eigenvectors $[H_z]$. Owing to the form of matrix $\mathbf{M}$ in Eq. (1), its square matrix has $4N+2$ eigenvalues, which are 2×2 degenerated and equal to τ_m. Since the eigenvalues of $\mathbf{M}^2$ are the square of the eigenvalues of $\mathbf{M}$, the link between τ_m and ρ_m is $\tau_m = {\rho_m}^2$. Thus the $4N+2$ eigenvalues ρ_m of $\mathbf{M}$ are divided into couples having opposite signs and obtained from the $2N+1$ eigenvalues τ_m through the relation $\rho_{m_{1,2}} = \pm\sqrt{\tau_m}$.

In conical incidence, matrix $\mathbf{M}$ contains 16 submatrices (half of which are null) and thus is a $4(2N+1)\times 4(2N+1)$ matrix.[6] For gratings slanted by an angle δ, the blocks of the new matrix $\tilde{\mathbf{M}}$ are simply related to those of matrix $\mathbf{M}$ by[6]

$$\tilde{M}_{i,j} = M_{i,j} + \mathbf{L}, \tag{6}$$

where L is a diagonal matrix with elements equal to $-in(2\pi/d)\tan\delta$.

To demonstrate the numerical problems that arise, we choose as a good reflector gold, with refractive index $0 + i10$, and study the dependence of the -1st-order efficiency as a function of the groove width c. Figure 3(a) shows that strong fluctuations still remain when the field is described by 31 Fourier components ($N = 15$). It must be pointed out that these anomalies have no physical character; they are just numerical errors, and their position as a function of the aspect ratio c varies when the truncation parameter N is changed. This behavior should be compared with the results obtained by using

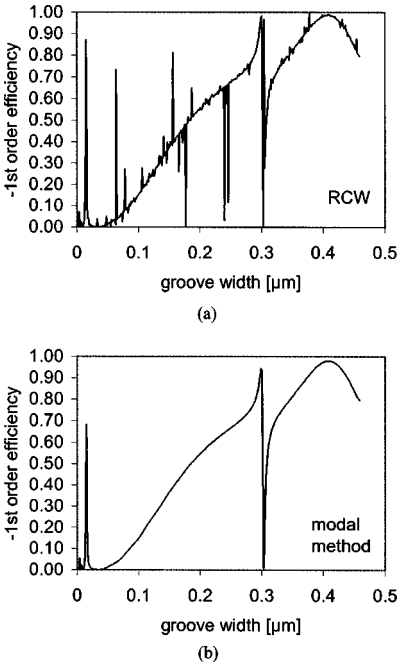

Fig. 3. -1st-order efficiency of a lamellar grating with optical index of the material $n_1 = 0 + i10$ as function of the groove width c for $N = 15$. $d = h = 0.5\,\mu$m, $\lambda = 0.6328\,\mu$m, $\theta = 30$ deg. (a) RCW method, (b) rigorous modal method.

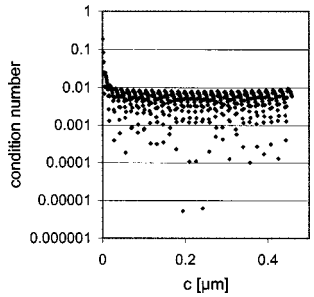

Fig. 4. Condition number of $[\![\epsilon]\!]$ as a function of the groove width c. All the parameters are the same as in Fig. 3.

the rigorous modal method,[10,11] which does not use Fourier decomposition of the field and thus of ϵ [Fig. 3(b)] and gives a smooth dependence except for two values of c where real anomalies occur.

Let us recall that $[\![\epsilon]\!]$ is the Toeplitz matrix constructed with the Fourier components of ϵ; i.e., the (n, m) entry of $[\![\epsilon]\!]$ is ϵ_{n-m}. When $n_1 = 0 + i10$, the relative permittivity of the metal becomes real negative: $\epsilon_{1,r} = {n_1}^2 = -100$. As a result, the Toeplitz matrix $[\![\epsilon]\!]$ that appears in Eq. (1) involves the Fourier components of a real periodic function oscillating between ϵ_0 and $-100\epsilon_0$, and its inverse oscillates between $1/\epsilon_0$ and $-1(100\epsilon_0)$. This behavior violates the condition that ensures the validity of the inverse rule[9] that was used to get Eqs. (1) and (5). The study of the condition number of $[\![\epsilon]\!]$ and $[\![1/\epsilon]\!]$ shows that it becomes less than 10^{-3} for many values of the groove width and can sometimes go as low as 10^{-5}, as presented in Fig. 4. The results are that this matrix is quasi-singular and that its inversion produces errors in the integration process of Eq. (1); this causes the appearance of many thin anomalies in Fig. 3(a), which are merely numerical artifacts. The aim of the study is to get rid of these artifacts.

3. ADDITIONAL TRUNCATION OF THE BLOCK MATRICES APPEARING IN THE DIFFERENTIAL SET OF EQUATIONS

A. Description of the Numerical Procedure

In the limit of infinite number of Fourier components ($N \to \infty$), the difference between the matrices $[\![\epsilon]\!]$ and $[\![1/\epsilon]\!]^{-1}$ vanishes. However, for a limited N, the two truncated matrices differ from each other. As mentioned by Li,[9] the difference is more noticeable close to the two extremities of the main diagonals of the matrices. This fact led to the idea of supplementary truncation of the two matrices $[\![\epsilon]\!]^{-1}$ and $[\![1/\epsilon]\!]^{-1}$ that exist in Eq. (1) after their calculation for a given truncation parameter by eliminating the regions lying close to the extremities of their main diagonals.

To this end, it was necessary to introduce a two-step truncation process. First, we used a larger-than-N truncation order $\mathbf{M} = N + \Delta$ when calculating the matrices $[\![\epsilon]\!]$, $[\![1/\epsilon]\!]$, $[\![\epsilon]\!]^{-1}$, $[\![1/\epsilon]\!]^{-1}$, and all submatrices in matrices $\mathbf{M}$ and $\mathbf{U}$. Then we made a supplementary truncation of all submatrices, reducing their size to the order $(2N+1)\times(2N+1)$. This process is schematically presented in Figs. 5–7. For submatrices formed by multiplication of two matrices, this process results in the multiplication of

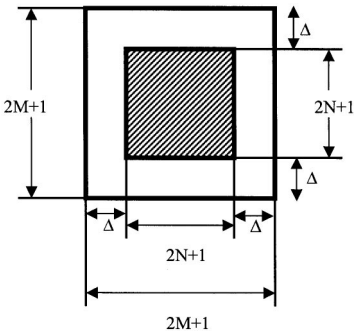

Fig. 5. Matrix truncation parameters M, Δ, and N.

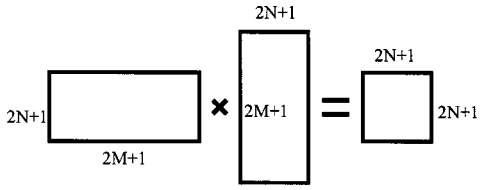

Fig. 6. Two-step truncation method applied to the multiplication of two matrices.

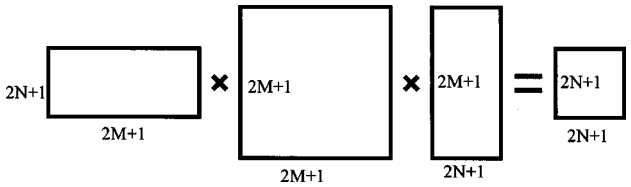

Fig. 7. Two-step truncation method applied to the multiplication of three matrices.

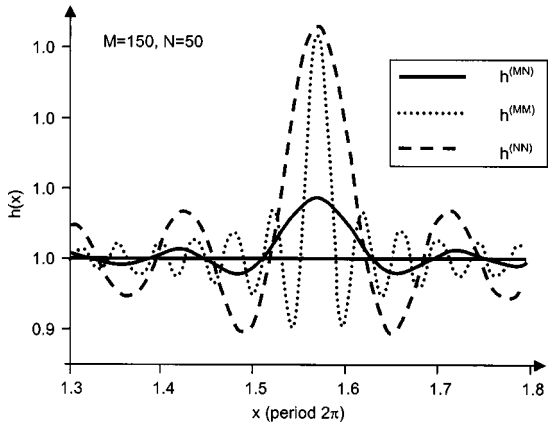

Fig. 8. Effect of the truncation type on the reconstruction of a continuous function obtained by the product of two discontinuous ones. In this example, $N = 50$ and $M = 150$.

two rectangular matrices with orders $(2N + 1) \times (2M + 1)$ and $(2M + 1) \times (2N + 1)$ (Fig. 6). For submatrices formed by multiplication of three matrices, this process is equivalent to multiplication matrices of orders $(2N + 1) \times (2M + 1)$, $(2M + 1) \times (2M + 1)$, and $(2M + 1) \times (2N + 1)$ (Fig. 7). In some cases, this kind of truncation may cause a smoothing effect. For example, following Ref. 9, we choose two functions $f(x)$ and $g(x)$, periodic with period 2π, given by

$$f(x) = \begin{cases} a, & |x| < \dfrac{\pi}{2} \\ \dfrac{a}{2}, & \dfrac{\pi}{2} < |x| \leqslant \pi \end{cases}, \tag{7}$$

and $g(x) = 1/f(x)$. It is obvious that the discontinuities of $f(x)$ and $g(x)$ are pairwise complementary in the sense that their product $h(x)$ is continuous, since $h(x) = f(x)g(x) = 1, \forall x$.

Figure 8 shows, in the neighborhood of the point of discontinuity $x = \pi/2$, the partial sums $h^{(MM)}(x)$, $h^{(MN)}(x)$, and $h^{(NN)}(x)$, defined below in Eqs. (8)–(11). The partial sums represent the function $h(x) = 1$ reconstructed by summing the function's truncated Fourier series (i.e., by a truncated inverse Fourier transform) and using the Fourier components $h_n{}^{(M)}$ obtained as a truncated factorized product of f_n and g_n. This product is computed by using the truncated finite Laurent rule (or direct rule):

$$h_n{}^{(M)} = \sum_{m=-M}^{M} f_{n-m} g_m . \tag{8}$$

The reconstructed function (partial sum) $h(x)$ then depends on two truncation parameters: first, the sum limits M in Eq. (8) and second, the number of Fourier components in the partial sum, Eqs. (9)–(11). We thus can define three different partial sums:

$$h^{(MM)}(x) = \sum_{n=-M}^{M} h_n{}^{(M)} \exp(inx), \tag{9}$$

$$h^{(MN)}(x) = \sum_{n=-N}^{N} h_n{}^{(M)} \exp(inx), \tag{10}$$

$$h^{(NN)}(x) = \sum_{n=-N}^{N} h_n{}^{(N)} \exp(inx). \tag{11}$$

As can be observed in Fig. 8 with $N = 50$ and $M = 150$, the Gibbs phenomenon due to the truncation and the discontinuity of $f(x)$ and $g(x)$ is partially smoothed out when the combination (MN) is used, i.e., when the reconstruction is made with N components in the sum of Eq. (10), a number significantly smaller than the number M in the factorization rule of Eq. (8).

B. Numerical Results for a Slanted Grating in Conical Mounting

First, we consider an arbitrarily chosen lamellar grating etched on a substrate with complex refractive index $0.01 + i10$. The geometry of the grating diffraction problem in conical mountings is depicted in Fig. 9 for the particular case of a surface-relief slanted lamellar diffraction grating. A linearly polarized plane wave is incident at angle of incidence $\theta = 30°$ and at azimuthal angle $\delta' = 30°$. The incident medium is air. The electric-field vector is in the plane of incidence, so that angle $\psi = 0$. The grooves of the lamellar grating are slanted at angle $\delta = -45°$, and the fill-in ratio (the duty cycle) of the grating is 0.5. The grating has a thickness (i.e., the entire

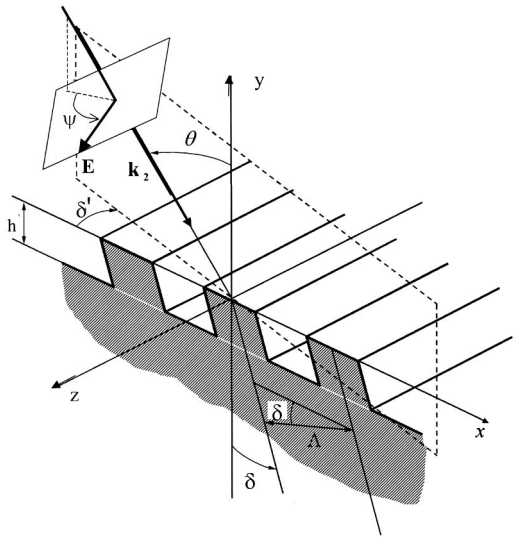

Fig. 9. Schematic representation of a slanted grating illuminated in conical mounting.

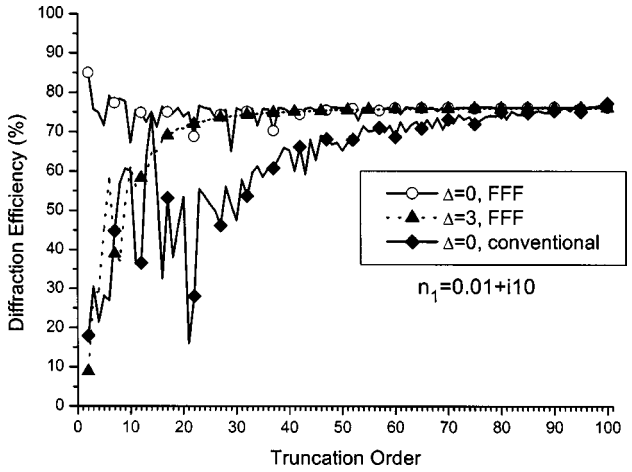

Fig. 10. Improvement of convergence as a function of truncation parameter with the use of the FFF method compared with the conventional formulation. Shown is the effect of the truncation parameter Δ on the fluctuations with the FFF method.

groove depth) $h = 0.3\,\mu$m. The incident (vacuum) wavelength is $\lambda = 0.8\,\mu$m, and the grating period is $d = 1\,\mu$m ($\Lambda = 0.7071\,\mu$m).

Figures 10 and 11 show the convergence of the diffraction efficiencies as N is increased. Figure 10 compares the convergence of the −1st-order diffraction efficiencies computed by the RCW method on the basis of the FFF formulation with $\Delta = 0$ (circles) and $\Delta = 3$ (triangles), and on the basis of the conventional formulation (without FFF) and with $\Delta = 0$ (diamonds) when N is increased from 2 to 100. The results of the conventional formulation oscillate and converge very slowly. The efficiencies computed with the FFF-based formulation converge remarkably quickly as N increases. However, when $\Delta = 0$, there are several strong fluctuations close to $N = 37$ and $N = 52$. Use of the preliminary truncation with $\Delta = 3$ helps to eliminate these fluctuations. The preliminary truncation method does not work for the conventional formulation of the eigenvalue problem of Eq. (3), because there is no inversion of the Toeplitz matrices in the conventional formulation, in contrast to Eq. (1). However, the conventional formulation already suffers from slow convergence, as is obvious from Fig. 10.

Figure 11 illustrates the effect of a preliminary truncation for two values of $\Delta = 3$ and 5. As can be observed, a further increase of Δ does not influence the convergence rate. We also tried other values of Δ, but increasing Δ further did not give as good results.

C. Numerical Results for an Unslanted Grating

To investigate the effect of the double truncation on fluctuations of the calculated diffraction efficiency for unslanted (in Fig. 9, $\delta = 0$) metallic lamellar gratings, we consider a lamellar grating etched on substrate with complex refractive index $0 + i10$. A linearly polarized plane wave is incident at angle of incidence $\theta = 30°$ and at azimuthal angle $\delta' = 0$ (nonconical mounting). As before, the incident medium is air. The electric-field vector is in the plane of incidence ($\psi = 0$, TM polarization). The fill-in ratio of the grating is 0.5. The grating has a groove depth of 0.5 μm. The incident (vacuum) wavelength is $\lambda = 0.6328\,\mu$m, and the grating period is 0.5 μm. This case is interesting, because it permits the use of two formulations of the eigenvalue problem, Eq. (3) or Eq. (4).

The general formulation requires a matrix **M** with dimensions $2(2N + 1) \times 2(2N + 1)$, and the formulation in Eq. (4) requires a $(2N + 1) \times (2N + 1)$ **U** matrix. For ordinary one-step truncation ($\Delta = 0$) the two formulations give identical numerical results; however, the computer code for solving Eq. (4) works much faster.

Figure 12 shows the convergence of the −1st-order diffraction efficiencies and illustrates the effect of preliminary truncation for $\Delta = 3$. Without preliminary truncation, $\Delta = 0$, there are two strong fluctuations for $N = 13$ and $N = 29$. Using preliminary truncation ($\Delta = 3$) eliminates these fluctuations but reduces the convergence rate. This reduction is much more pronounced when Eq. (4) is used instead of Eq. (3), although the latter requires longer computation times for the same value of N.

All results presented in Figs. 10–12 correspond to the gratings with fill-in ratio 0.5. Attempts to extend the

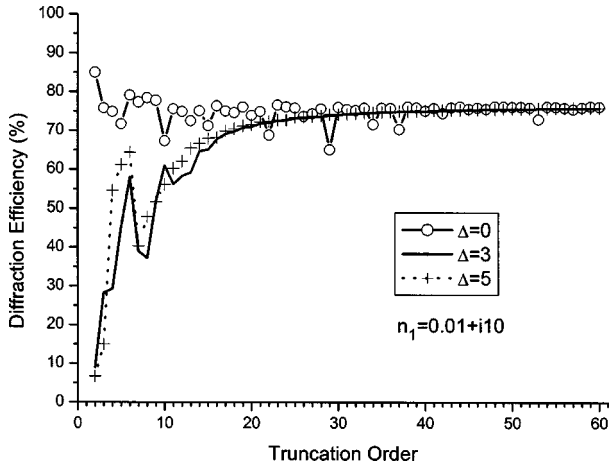

Fig. 11. Influence of the Δ parameter on the elimination of the fluctuations when the truncation number N is increased.

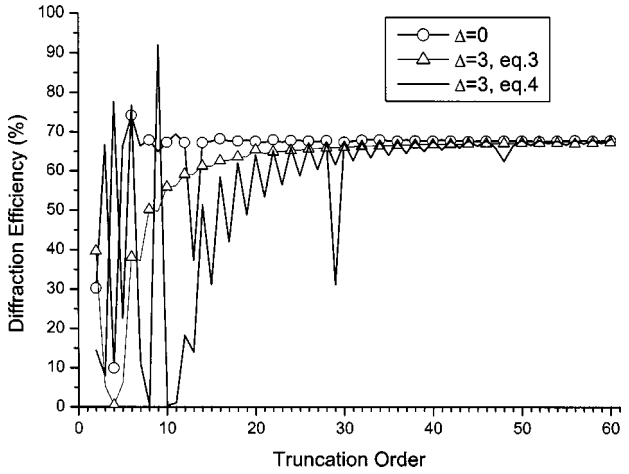

Fig. 12. Performance of the two-step truncation method when it is applied to the eigenvalue problems stated by Eqs. (3) and (4).

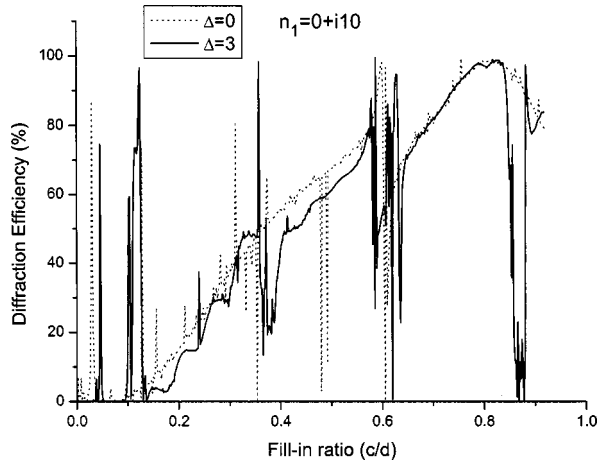

Fig. 13. Performance of the two-step truncation method as a function of the fill-in-ratio grating parameter.

preliminary truncation method on gratings with arbitrary fill-in ratio have not given satisfactory results. Figure 13 shows the dependence of diffraction efficiency on fill-in ratio at $N = 15$, the remaining parameters being the same as in Fig. 3. Figure 13 shows that the preliminary truncation method with $\Delta = 3$ eliminates some of fluctuations that appear when $\Delta = 0$, in particular in the neighborhood of $c/d = 0.5$. However, many other oscillations still remain, in comparison with the curve in Fig. 3(b).

4. PHYSICAL SOLUTION

Since the problem seems to be due to the poor behavior of ϵ and $1/\epsilon$, which jump from real positive to almost real negative values across the grating surface, we can introduce inside the grating material an artificial region in which n_1 is a complex number far from a pure imaginary one. This means that we introduce a bulk region with a much greater real part, typically $n_m = 2 + i10$. On the surface of this bulk, a layer of the original grating material with refractive index n_1 will remain (see Fig. 14, which presents an example of a lamellar grating). When $n_1 = 0.1 + i10$, the attenuation depth is typically less than 10 nm, so provided that the layer thickness is ~30–40 nm (or less), the field inside the artificial bulk will be close to zero; i.e., the incident wave will not experience the existence of the bulk material. It means that inside the artificial domain we should be allowed to put any material we want without modifying the diffraction problem significantly. However, the functions ϵ and $1/\epsilon$ will not so strongly violate the condition that ensures the validity of the inverse rule, so the corresponding Toeplitz matrices are expected to behave much better. On the other hand, although the incident wave could not experience the change in the bulk index, it is important not to change its index too much in comparison with the metal layer index, because the Fourier components of ϵ and $1/\epsilon$ must not differ too much from those in the initial diffraction problem; otherwise, a large value of N would be required for a proper description of ϵ and $1/\epsilon$ and would affect the convergence rate.

A. Lamellar Profile

To prove that it is the jump of ϵ and $1/\epsilon$ from positive to negative values that is at the origin of the problem, we choose the worst case from a numerical point of view with $n_1 = 0 + i10$, i.e., zeroing the real part of the refractive index. The groove and incident angle parameters are the same as in Fig. 3. The thickness e of the layer is equal to 20 nm. Figure 15 shows the -1st-order efficiency as a function of c. Two cases are shown: (a) with $n_m = n_1 = 0 + i10$ and (b) $n_m = 2 + i10$. As can be observed, artifacts similar to those in Fig. 3(a) appear in case (a), the difference being due to the change in N, whereas the curve in Fig. 15(b) is free of them. Moreover, for case (b), changing N to 15 instead of 20 led to a curve practically unchanged, a result not shown here. Figure 16 presents the convergence of the sum of efficiencies (order 0 plus order -1) as a function of N for several cases. Case (a) is identical to case (a) in Fig. 15, with $n_m = n_1 = 0 + i10$, and to case (b), too, with $n_1 = 0 + i10$ and $n_m = 2 + i10$, whereas for case (c) it is assumed that the entire grating is lossy with $n_m = n_1 = 2 + i10$. One can observe that introducing a lossy bulk material [case (b)] removes the numerical fluctuations and that the convergence rate in that case is identical to the convergence rate when the grating material is overall equal to $2 + i10$ [case (c)], the latter being characterized by absorption losses (total diffracted efficiency smaller than 100%).

Case (d) treats a metallic grating with dielectric bulk material, i.e., $n_1 = 0 + i10$ and $n_m = 2 + i0$, and illus-

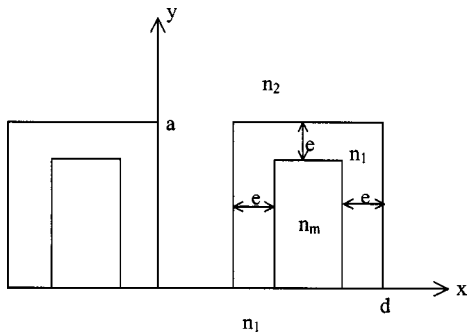

Fig. 14. Schematical representation of the lamellar grating with artificially changed bulk material.

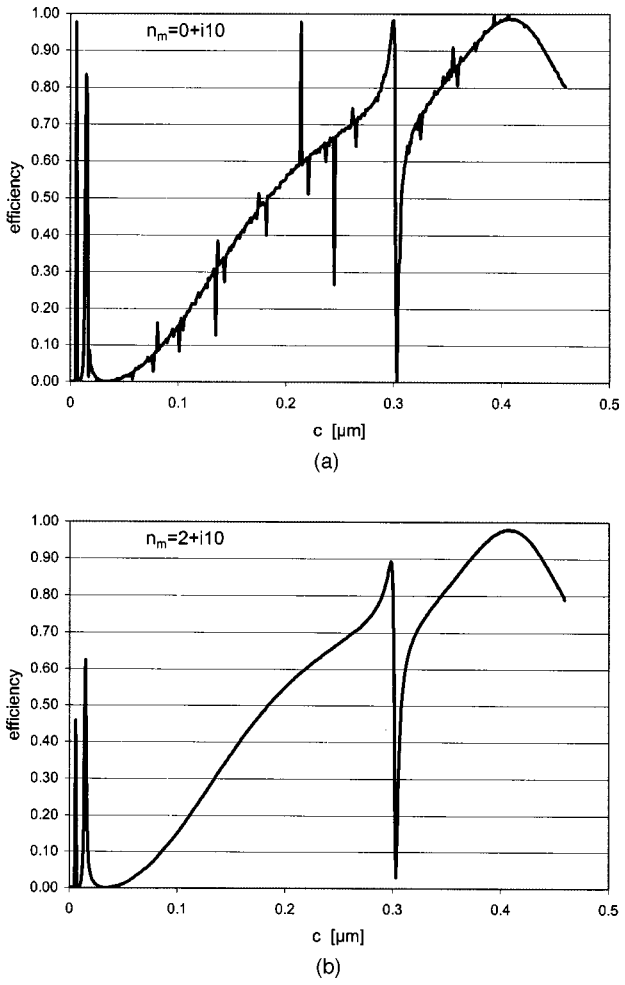

Fig. 15. (a) same as in Fig. 3(a), but with $N = 20$ instead of 15; (b) same as (a) but with a bulk material with losses.

trates the discussion at the end of the previous paragraph that problems of convergence will exist when the bulk index is quite different from the layer index.

From Figs. 15 and 16, it is established that the physical solution that avoids the jumps of ϵ and $1/\epsilon$ from negative to positive real values produces stable results as soon as the layer thickness is twice the attenuation length, whatever the aspect ratio may be.

B. Sinusoidal Profile

In contrast to lamellar profiles, sinusoidal gratings can be seen as having an aspect ratio that varies with the ordinate. Having suppressed the numerical artifacts for any aspect ratio of lamellar gratings, we are tempted to try the physical solution on arbitrary profiles. Figure 17 shows the results for the grating with a sinusoidal profile with parameters given in Fig. 1 (dashed curve). Again, we take the most difficult case with $n_1 = 0 + i10$, which exhibits the greatest oscillations in Fig. 1. If we change the index of the volume material to $n_m = 3 + i10$ and preserve only a 40-nm-thick layer with index $n_1 = 0 + i10$, the convergence is drastically improved. Here it was necessary to increase the layer thickness, measured in the vertical direction; otherwise, the layer becomes too thin at the steep slopes (the modulation depth $h/d = 100\%$), and absorption losses due to the bulk material manifest themselves.

5. CONCLUSION

The numerical problem encountered by the FFF method when applied to highly conducting gratings and due to the limitations of the inverse rule can be overcome in some cases by a suitable additional truncation of the matrices involved in the calculations. However, the improvement is not guaranteed for every aspect ratio of the grating.

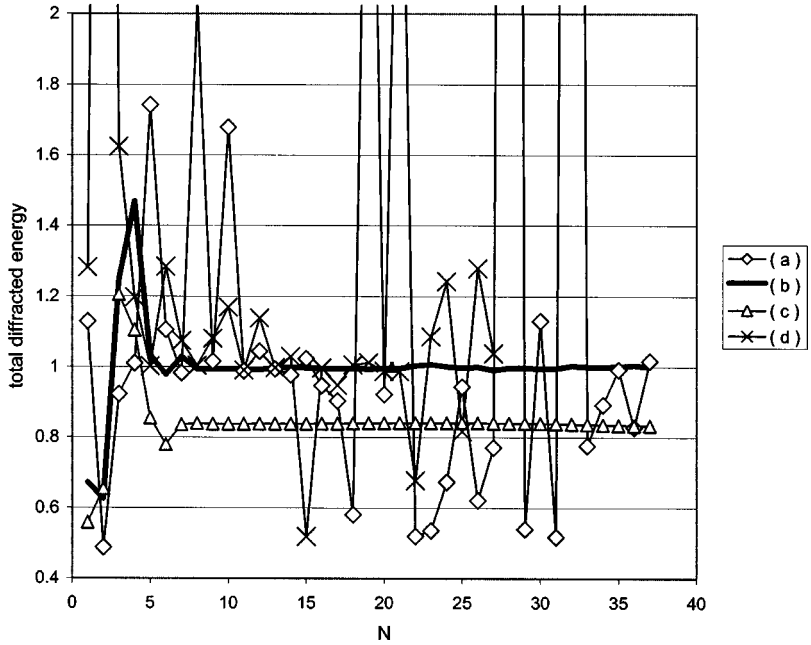

Fig. 16. Convergence with respect to N of the total diffracted energy (sum of efficiencies) for the grating presented in Figs. 3 and 15 with $c/d = 0.5$. (a) $n_m = n_1 = 0 + i10$; (b) $n_1 = 0 + i10$ and $n_m = 2 + i10$; (c) $n_m = n_1 = 2 + i10$; (d) $n_1 = 0 + i10$ and $n_m = 2 + i0$.

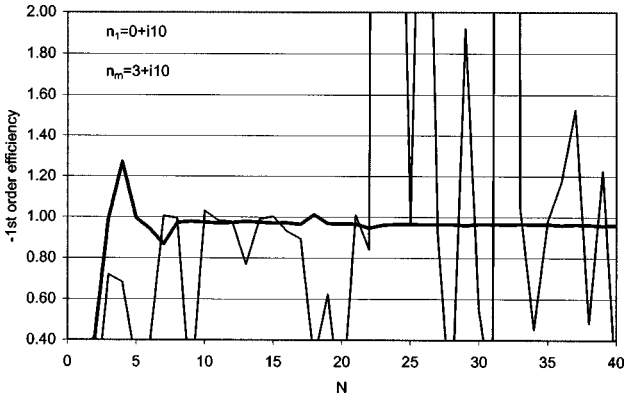

Fig. 17. Convergence of the −1st-order efficiency of a sinusoidal grating when the truncation parameter N increases. The grating is the same as in Fig. 1. The bulk region with an index n_m is covered with a 40-nm-thick layer with index n_1. Thick curve, $n_m = 3 + i10$ and $n_1 = 0 + i10$; thin curve, $n_m = n_1 = 0 + i10$ (the same as for the dashed curve in Fig. 1).

The physical solution, which consists of artificially introducing lossier material into the bulk of the highly conducting metal, leads to a fast convergence of the efficiency with respect to the number of Fourier components of the field, without changing the results significantly. This approach is applicable to most grating materials and profiles. Moreover, it can be easily extended to two- and three-dimensional grating structures, for which the convergence problems are particularly important because of limitations of memory and computation time.

The e-mail address of E. Popov is e.popov@fresnel.fr.

REFERENCES

1. M. Nevière, G. Cerutti-Maori, and M. Cadilhac, "Sur une nouvelle méthode de résolution du problème de la diffraction d'une onde plane par un réseau infiniment conducteur," Opt. Commun. **3**, 48–52 (1971).
2. M. Nevière, R. Petit, and M. Cadilhac, "About the theory of optical grating coupler-waveguide systems," Opt. Commun. **8**, 113–117 (1973).
3. M. Nevière, P. Vincent, R. Petit, and M. Cadilhac, "Systematic study of resonances of holographic thin film couplers," Opt. Commun. **9**, 48–53 (1973).
4. M. Nevière, P. Vincent, and R. Petit, "Sur la théorie du réseau conducteur et ses applications à l'optique," Nouv. Rev. Opt. **5**, 65–77 (1974).
5. L. Li, "Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings," J. Opt. Soc. Am. A **13**, 1024–1035 (1996).
6. M. Nevière and E. Popov, *Light Propagation in Periodic Media: Differential Theory and Design* (Marcel Dekker, New York, 2003).
7. E. Popov and M. Nevière, "Grating theory: New equations in Fourier space leading to fast converging results for TM polarization," J. Opt. Soc. Am. A **17**, 1773–1784 (2000).
8. E. Popov and M. Nevière, "Maxwell equations in Fourier space: Fast converging formulation for diffraction by arbitrary shaped, periodic, anisotropic media," J. Opt. Soc. Am. A **17**, 1773 (2001).
9. L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," J. Opt. Soc. Am. A **13**, 1870–1876 (1996).
10. C. Botten, M. S. Craig, R. C. McPhedran, J. L. Adams, and J. R. Andrewartha, "The dielectric lamellar diffraction grating," Opt. Acta **28**, 413–428 (1981).
11. C. Botten, M. S. Craig, R. C. McPhedran, J. L. Adams, and J. R. Andrewartha, "The finitely conducting lamellar diffraction grating," Opt. Acta **28**, 1087–1102 (1981).
12. J. R. Andrewartha, G. H. Derrick, and R. C. McPhedran, "A general modal theory for reflection gratings," Opt. Acta **28**, 1501–1516 (1981).
13. S. T. Peng, T. Tamir, and H. Bertoni, "Theory of periodic dielectric waveguides," IEEE Trans. Microwave Theory Tech. **MTT-23**, 123–133 (1975).
14. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," J. Opt. Soc. Am. **71**, 811–818 (1981).
15. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of dielectric surface-relief gratings," J. Opt. Soc. Am. **72**, 1385–1392 (1982).
16. M. G. Moharam, D. A. Pommet, E. B. Grann, and T. K. Gaylord, "Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmission matrix approach," J. Opt. Soc. Am. A **12**, 1077–1086 (1995).
17. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2nd ed. (Cambridge U. Press, Cambridge, UK, 1992).