OPTICA ACTA, 1986, VOL. 33, NO. 5, 593–605

Convergence of Rayleigh–Fourier method and rigorous differential method for relief diffraction gratings

E. POPOV and L. MASHEV

Institute of Solid State Physics, Bulgarian Academy of Sciences, boul. Lenin 72, Sofia 1784, Bulgaria

(*Received 26 June 1985*)

Abstract. The Rayleigh–Fourier method and the method of Chandezon *et al.* for the calculation of diffraction efficiency of sinusoidal metal, dielectric and coated dielectric gratings are compared with respect to the threshold truncation value and the thickness of the coating layer. For shallow gratings the convergence of the two methods is practically one and the same. However, for deeper gratings the method of Chandezon *et al.* is more powerful. It is shown that for coated dielectric gratings the thickness of the layer imposes a limit on the truncation value, the limitations being weaker for the method of Chandezon *et al.*

1. Introduction

During the last fifteen years considerable interest has been shown in the problem of light diffraction from relief gratings. An extensive review on this subject is given in references [1–4]. Today, rigorous electromagnetic theories based on the integral [5–7] or the differential [8] formalism are the most trustworthy. These two approaches, numerically implemented in a computer code, give accurate results for arbitrary gratings in a large spectral range and through them the most complete comparison is made with other theories and experimental data [1].

Another widely used approach is the Rayleigh method [9], especially its Rayleigh–Fourier (RF) modification. Although the RF method is non-rigorous in a mathematical sense [10], Wirgin [11, 12] shows that it gives accurate results for sinusoidal gratings with half groove depth to period ratio h/d five times greater than the theoretical limit $h/d = 0{\cdot}072$ [10]. The diffraction efficiency curves, calculated by the abovementioned methods for two-layer gratings have been studied extensively in the literature. However, less attention has been devoted to more complicated structures, for example, three-layer [13, 14], multilayered [15] dielectric gratings and multicoated metallic gratings [16, 17], where some new effects occur. In particular, total reflection and selectivity of the zeroth order for three-layer gratings, quite narrow anomalies for multilayered ones and reduction of the absorption from the metal for multicoated gratings are observed. In this paper a detailed comparison between the calculations of the diffraction efficiency curves for sinusoidal perfectly conducting, metallic, dielectric and coated dielectric gratings, using the method of Chandezon *et al.* [18] (C method) and the RF method is carried out. The choice of the C method is not accidental. First, it is a rigorous vector electromagnetic method, suitable for the treatment of multicoated gratings, and second, it gives accurate results for very deep gratings with $h/d \sim 3$ [18].

It is shown that for bare dielectric or metallic gratings the simpler RF method is preferable at low and moderate values of h/d. However, for deeper gratings the C method is more powerful and its convergence rate is much faster. An interesting consequence is that in the case of coated gratings the thickness of the middle layer introduces a further limitation on the value of h/d which can be achieved for both methods.

2. Formulation of the problem

A plane monochromatic wave with time dependence $\exp(i\omega t)$ is incident at an angle θ on a three-layer grating (figure 1). Region 3 is filled with a metal or a dielectric, the middle layer is a dielectric with a thickness t, and the upper layer is air. The corrugation is one and the same for the two boundaries. Since the convergence of the RF method decreases rapidly for non-sinusoidal gratings, we consider a sinusoidal groove profile with a period d and a depth $2h$.

The Floquet theorem states that for $y \to \pm\infty$ the field can be represented as a sum of propagating plane waves (diffraction orders). The electromagnetic field must satisfy the Maxwell equations in the three media, the boundary conditions at the two corrugated boundaries and the out-going wave conditions when $y \to \pm\infty$. A brief review of the main mathematical aspects of the RF and C methods are given below in order to compare their mathematical base and to introduce some notations to be referred to.

2.1. *RF method*

This method has been discussed in detail by Maystre [2] and Wirgin [11]. We have modified it to work for coated gratings too, as follows. In each medium 1, 2 and 3 (figure 1), the electromagnetic field can be represented in the form

$$F_z^j(x,y) = \sum_m [a_m^{j+} \exp(ih_{m,j}^{RF} y) + a_m^{j-} \exp(-ih_{m,j}^{RF} y)] \exp(i\beta_m x), \tag{1}$$

where $\beta_m = \beta_0 + mK$, $K = 2\pi/d$, $\beta_0 = n_1 \sin\theta$, $m = 0, \pm 1, \pm 2, \ldots$ and

$$h_{m,j}^{RF} = k(n_j^2 - \beta_m^2)^{1/2}, \quad k = 2\pi/\lambda. \tag{2}$$

$a_m^{j\pm}$ are the amplitudes of the field components and F_z^j is the z component of the electric or magnetic field vector for TE or TM polarization, respectively.

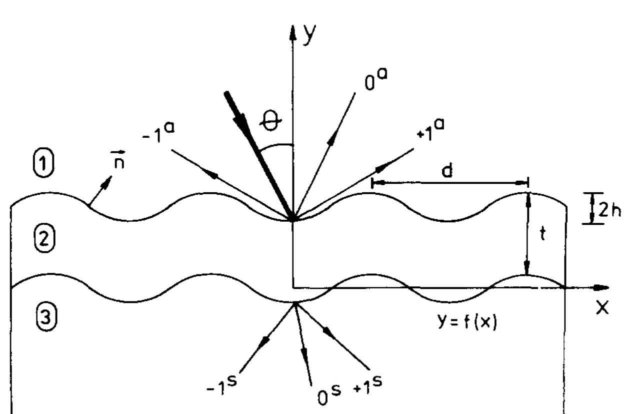

Figure 1. Schematic representation of the configuration under consideration.

The boundary conditions are

$$\left.\begin{aligned}
F_z^1(x, y=f(x)+t) &= F_z^2(x, y=f(x)+t),\\
F_z^2(x, y=f(x)) &= F_z^3(x, y=f(x)),\\
P^1 \left.\frac{\partial F_z^1(x,y)}{\partial \mathbf{n}(x,y)}\right|_{y=f(x)+t} &= P^2 \left.\frac{\partial F_z^2(x,y)}{\partial \mathbf{n}(x,y)}\right|_{y=f(x)+t},\\
P^2 \left.\frac{\partial F_z^2(x,y)}{\partial \mathbf{n}(x,y)}\right|_{y=f(x)} &= P^3 \left.\frac{\partial F_z^3(x,y)}{\partial \mathbf{n}(x,y)}\right|_{y=f(x)},
\end{aligned}\right\} \quad (3)$$

where $P^j = 1$ for the TE case, and $P^j = n_j^{-2}$ for the TM case.

Substituting (1) into (3), the RF method requires the development of $\exp(\pm i h_{m,j}^{\mathrm{RF}} f(x))$ in Fourier series. The resulting equations can be written in a matrix form:

$$\left.\begin{aligned}
M_{\mathrm{R}}^1 \xi^1 &= M_{\mathrm{R}}^2 \phi^2(t) \xi^2,\\
M_{\mathrm{R}}^2 \xi^2 &= M_{\mathrm{R}}^3 \xi^3,
\end{aligned}\right\} \quad (4)$$

where ξ^j is a column vector:

$$\xi^j = \begin{bmatrix} \vdots \\ a_m^{j+} \\ \vdots \\ a_m^{j-} \\ \vdots \end{bmatrix}.$$

M_{R}^j is a square matrix of a double infinite order:

$$M_{\mathrm{R}}^j = \begin{bmatrix} A_{\mathrm{R}}^j & B_{\mathrm{R}}^j \\ C_{\mathrm{R}}^j & D_{\mathrm{R}}^j \end{bmatrix},$$

where

$$\left.\begin{aligned}
A_{m,p}^j &= \frac{1}{d}\int_0^d \exp\left[i h_{p,j}^{\mathrm{RF}} f(x) + i(m-p)Kx\right] \mathrm{d}x,\\
B_{m,p}^j &= A_{p,m}^j,\\
C_{m,p}^j &= h_{p,j}^{\mathrm{RF}} A_{m,p}^j + \beta_p \sum_q F_q A_{m-q,p}^j,\\
D_{m,p}^j &= -h_{p,j}^{\mathrm{RF}} B_{m,p}^j + \beta_p \sum_q F_q B_{m-q,p}^j,
\end{aligned}\right\} \quad (5)$$

F_q is the qth Fourier component of the derivative of the corrugation function $f(x)$ and ϕ^j is a diagonal matrix:

$$\phi^j = \begin{bmatrix} \phi^{j+} & 0 \\ 0 & \phi^{j-} \end{bmatrix}, \quad (6)$$

with $\phi^{j\pm}_{m,p}(y)=\delta_{m,p}\exp(\pm ih^{\mathrm{RF}}_{m,j}y)$ and $\delta_{m,p}$ is a Kronecker symbol. Using (4) a system of linear algebraic equations can be obtained for the unknown amplitudes:

$$M^3_{\mathrm{R}}\xi^3=M^2_{\mathrm{R}}\phi^2(t)(M^2_{\mathrm{R}})^{-1}M^1_{\mathrm{R}}\xi^1. \tag{7}$$

From the out-going wave conditions

$$a^{1+}_m=0, \quad a^{1-}_m=\delta_{m,0}, \tag{8}$$

applied to the system (7) we obtain that the number of the equations in (7) is equal to the number of the unknown amplitudes.

2.2. *C method*

The basic idea of this method lies in the introduction of a new coordinate system $\tilde{S}(\tilde{x},\tilde{y},\tilde{z})$ with

$$\tilde{x}=x, \quad \tilde{y}=y-f(x), \quad \tilde{z}=z. \tag{9}$$

In $\tilde{S}$ the boundary conditions are simplified because of the plane boundaries, however, the form of the Maxwell equations becomes much more complicated a system of second-order partial differential equations with non-constant coefficients. The Fourier transformation applied to this system results in an infinite system of ordinary differential equations. Similar to the RF method, the field in each medium can be represented as a sum of exponentials

$$\exp(\pm ih^{\mathrm{C}}_{m,j}\tilde{y}). \tag{10}$$

However, now $h^{\mathrm{C}}_{m,j}$ are the eigenvalues of the characteristic matrix T of the infinite system of ordinary differential equations. In the matrix form the solution can be written as follows:

$$F^j_z(\tilde{x},\tilde{y})=\boldsymbol{\eta}(\tilde{x})M^j_{\mathrm{C}}\phi^j(\tilde{y})\xi^j, \tag{11}$$

where $\boldsymbol{\eta}$ is a row-vector with components $\exp(i\beta_m\tilde{x})$; M^j_{C} is the matrix of eigenvectors of T and the other members are the same as in the Rayleigh–Fourier method. The boundary conditions in $\tilde{S}$ lead to (7), substituting M^j_{R} with M^j_{C}.

3. Numerical problems

An essential difference between the two methods is that in the RF method the field is searched for in a form of a plane-wave expansion with preliminary stated exponents (2) (non-rigorous treatment in the grating region), while in the C method at first the exponentials (10) are calculated and after that the expansion of the field into these exponentials is performed (rigorous treatment).

In fact, the two methods would be equivalent, if the infinite system (7) with M_{R} or M_{C} can be solved [18, 19]. For a numerical treatment, however, a truncation of the matrices M_{R} and M_{C} is necessary to be made up to a given order $M=2N+1$ with $-N\leqslant m,p,q\leqslant N$.

When the corrugation depth tends to zero, the coordinate system $\tilde{S}$ tends towards the unperturbed coordinate system S thus $h^{\mathrm{C}}_{m,j}\rightarrow h^{\mathrm{RF}}_{m,j}$. Therefore up to a given small value of the groove depth the two methods must give not only identical results, but would have the same interstitial steps, too. This fact is of great important for testing the computer codes.

As the corrugation depth is increased the two methods must have different convergence rates and different limitations. There are no theoretical arguments to predict the tendencies with respect to the restrictions and the accuracy of the two methods when a specific system is considered.

In the following three sections a comparison between the convergence rates of the two methods and the diffraction efficiences defined as

$$\eta_m^j = |a_m^j|^2 n_j \cos\theta_m^j/(n_1 \cos\theta), \qquad (12)$$

is made for metal, dielectric and coated dielectric gratings. We have developed two computer codes, based on the RF and C methods respectively. The calculations have been performed on a computer with a wordlength of 32 bits. The eigenvalue and the eigenvector problems were solved by a standard QR method (see, for example, [20]) for complex non-symmetrical matrices, using EISPACK. The matrix inversion, necessary in the solution of (7) both for the RF and C methods, was performed by the Gauss–Jordan scheme.

The normalization of the eigenvectors in the C method leads to well determined non-singular matrices. For the RF method in the case of perfectly conducting gratings the inner product in (5) results in numerical instability, due to the truncation of the small matrix elements in M_R in comparison with the large ones in a floating-point computer arithmetic. This requires a special renormalization of the matrix elements (5).

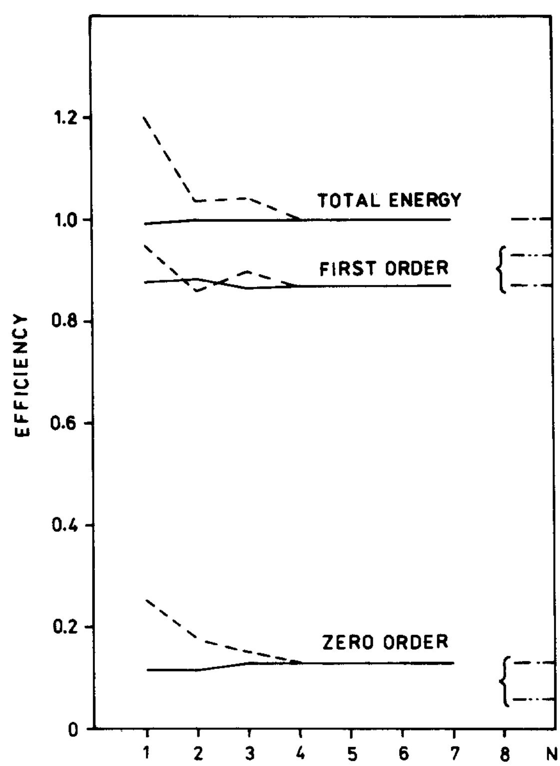

Figure 2. Diffraction efficiency curves and total diffracted energy of a perfectly conducting grating with $d = 15\,\mu$m, $2h = 6{\cdot}83\,\mu$m as a function of a truncation parameter N at constant deviation of 8·9° between −1 backscattered order and TE polarized incident wave. Full curve, C method; broken curve, RF method. In the right-hand side the results of Kalhor and Neureuther (—··—··—) and Wirgin (—·—·—) are taken from figure 9 of [11].

4. Numerical results

4.1. *Metal gratings*

The diffraction efficiencies and the total diffracted energy have been calculated for a sinusoidal perfectly conducting grating using both computer codes. The calculated values for different N are presented in figure 2 together with the set of results obtained by Kalhor and Neureuther and by Wirgin, taken from figure 9 of [11]. It is worth noting that our results for the RF method coincide with those of Wirgin with the same convergence rate $N=4$, while the stabilized values of the C method occur (within 0·1 per cent relative error) for $N=3$.

To test the possibilities of the two computer codes, as a second example we have considered a normally incident plane wave on a perfectly conducting grating for the two fundamental cases of polarization (figure 3). The calculations have been done by the C method. Wirgin (figure 8 of [11]) gives the convergence rate of the RF method for the same type of grating (right-hand side in figure 3) whose stabilized values coincide with the results of the C method. A quite important difference is that the threshold of the RF method is obtained for $N=15$, while the same value for the C method in $N=7$.

Further on the computer codes were generalized to work for real metal gratings taking into account the finite conductivity of the metal. Figure 4 shows the efficiency of an aluminium grating ($n_2=1{\cdot}23+i6{\cdot}95$) with a period $d=0{\cdot}8\,\mu$m and ratio $h/d=0{\cdot}3$. The results of the RF method are highly oscillatory for small N and the threshold is reached for $N=10$. The C method has a much faster convergence rate ($N=5$).

4.2. *Dielectric gratings*

Let us consider a normally incident TE polarized wave on a lossless dielectric grating with a period $d=0{\cdot}8\,\mu$m, half groove depth $h=0{\cdot}173\,\mu$m and refractive index of the lower medium $n_2=2{\cdot}3$. The upper medium is air ($n_1=1$). At these conditions three reflected (0^a, $\pm 1^a$) and five transmitted (0^s; $\pm 1^s$, $\pm 2^s$) diffraction orders are propagating. A convergence of the diffraction efficiency as a function of the truncation parameter N is given in figure 5 only for the transmitted orders for the sake of clarity. Both the RF and C methods have a convergence value of $N=6$. The same results for TE and TM polarizations at the same conditions as in figure except for $h=0{\cdot}24\,\mu$m (i.e. $h/d=0{\cdot}3$) are displayed in figure 6. It is interesting to notice that the stabilized value $N=10$ for the RF method is the same as for the perfectly conducting grating with $h/d=0{\cdot}3$ [11]. However, now the threshold value of the C method is smaller ($N=7$). Reducing the refractive index of the lower medium, the number of the transmitted diffraction orders is decreased. The same grating as in figure 6 with $n_2=1{\cdot}5$ supports two orders less, $\pm 2^s$ orders are passing off. The convergence for the two methods is better, but again the threshold value of N for the C method is smaller than for the RF method (figure 7).

4.3. *Coated dielectric gratings*

In this section two gratings of the type presented in figure 1 are considered. The first one is a grating on the glass substrate ($n_3=1{\cdot}5$) with a period $d=0{\cdot}3\,\mu$m and ratio $h/d=0{\cdot}17$ covered with a dielectric layer ($n_2=2{\cdot}3$). The dependence of the total diffracted energy on the layer thickness t is given in figure 8. Due to the moderately low h/d ratio and that in the air and the substrate only the zeroth diffracted orders exist, the convergence is achieved for $N=3$.

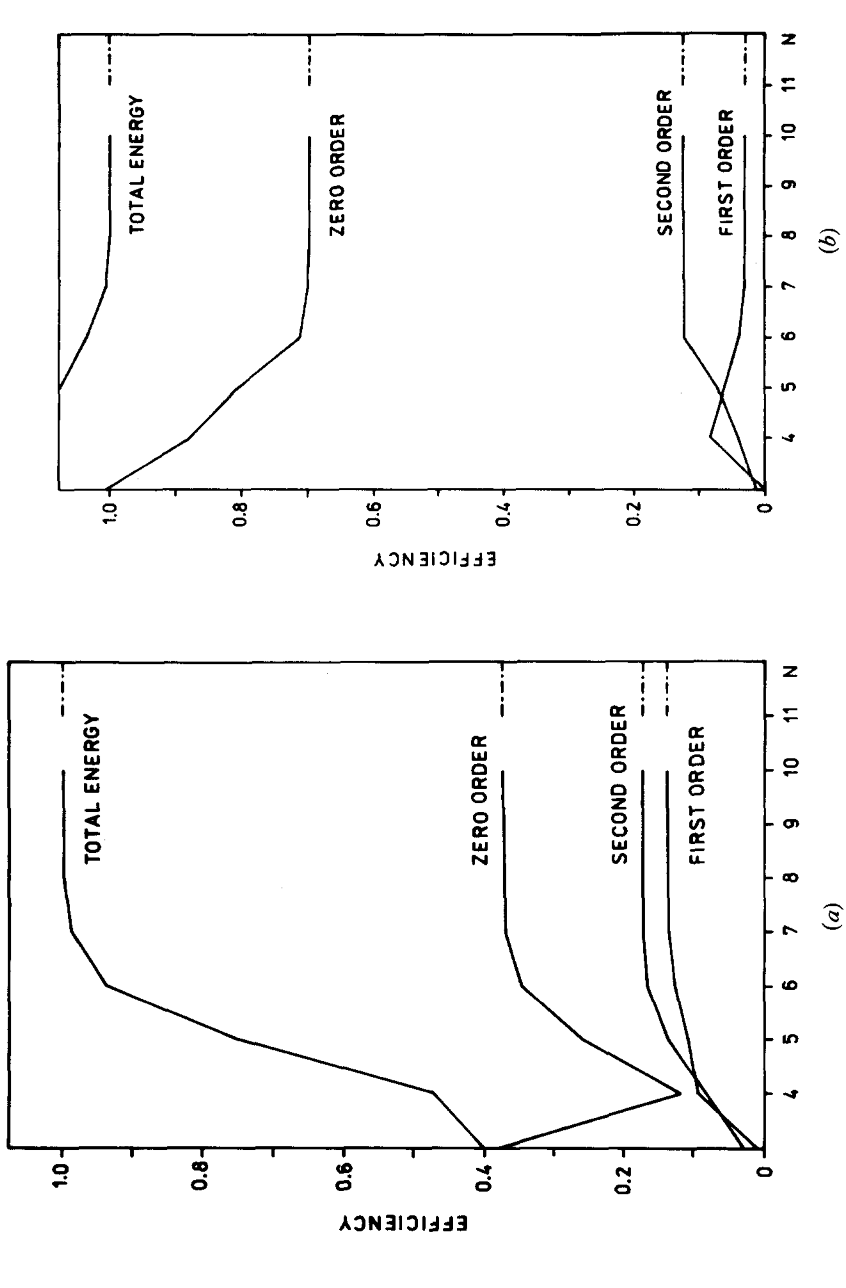

Figure 3. Convergence of the efficiencies for a sinusoidal perfectly conducting grating (C method). $\lambda = 0{\cdot}546\,\mu m$, $d = 1{\cdot}25\,\mu m$, $h/d = 0{\cdot}35$. The data represented by (— · — · —) are the results of Wirgin. (*a*) shows TE polarization and (*b*) TM polarization.

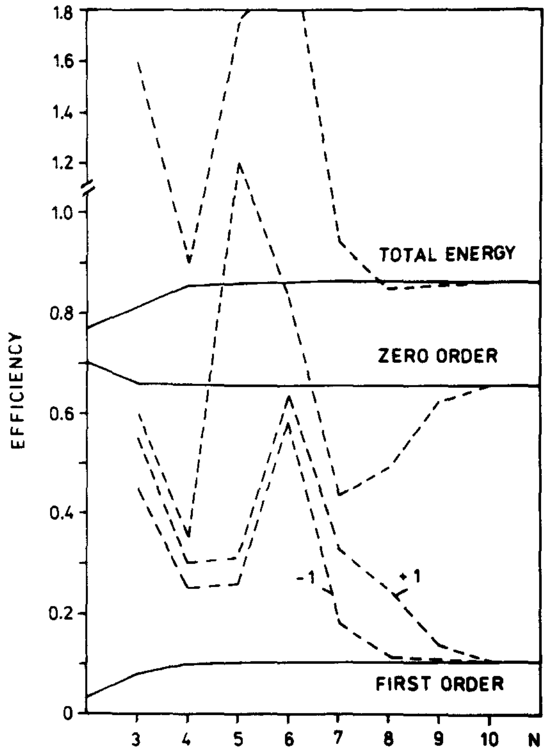

Figure 4. Diffraction efficiency of an aluminium grating with $d = 0{\cdot}8\,\mu m$, $h/d = 0{\cdot}3$ as a function of a truncation parameter N. Normally incident TE polarized wave with $\lambda = 0{\cdot}6328\,\mu m$. The data are (——) C method results and (– – –) RF method results.

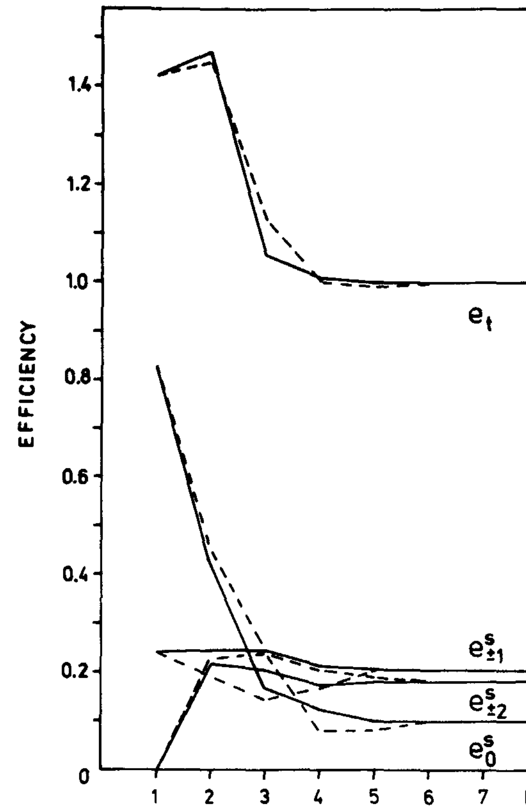

Figure 5. Approximation sequences for a bare dielectric grating with $n_2 = 2{\cdot}3$, calculated by the C method (full curve) and by the RF method (broken curve) at normal incidence. $\lambda = 0{\cdot}6328\,\mu m$, TE polarization, $d = 0{\cdot}8\,\mu m$, $h = 0{\cdot}173\,\mu m$.

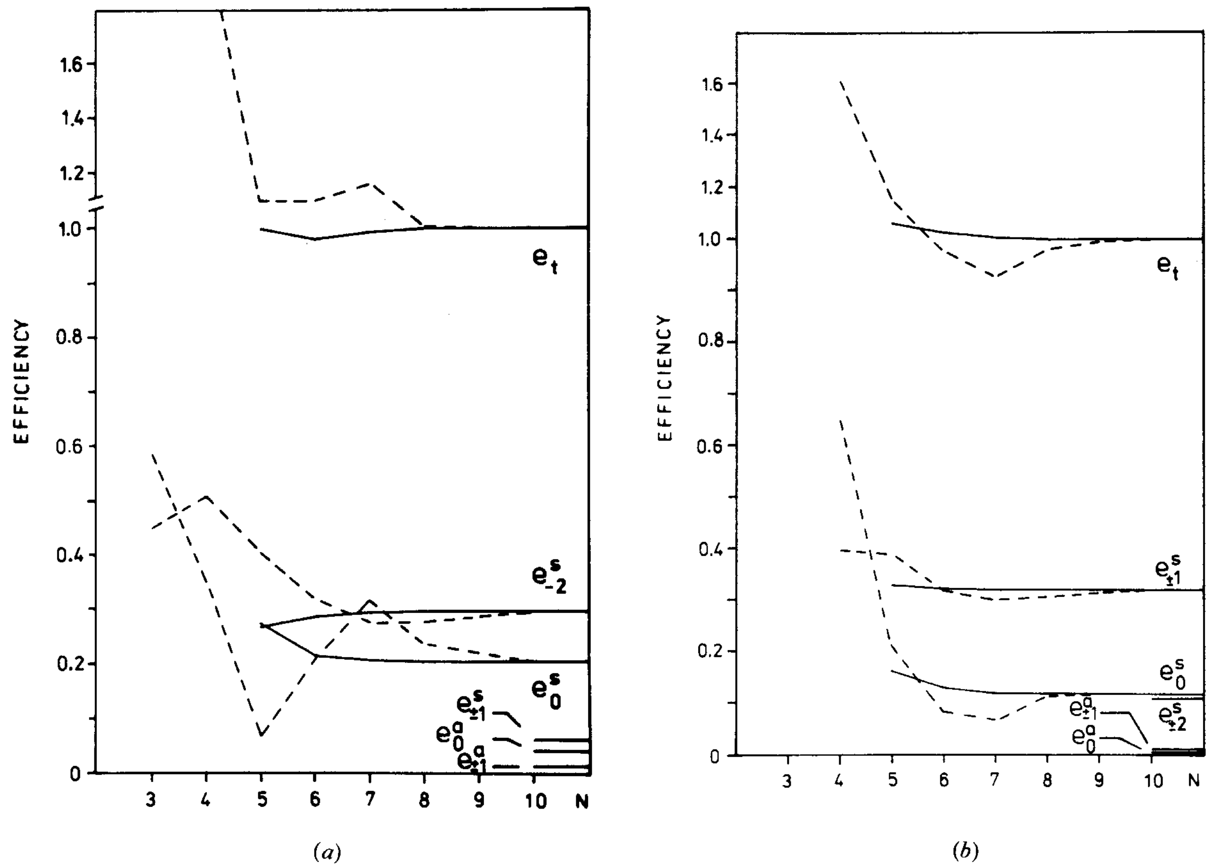

Figure 6. As figure 5 except for $h = 0{\cdot}24\,\mu m$, for (*a*) TE polarization and (*b*) TM polarization.

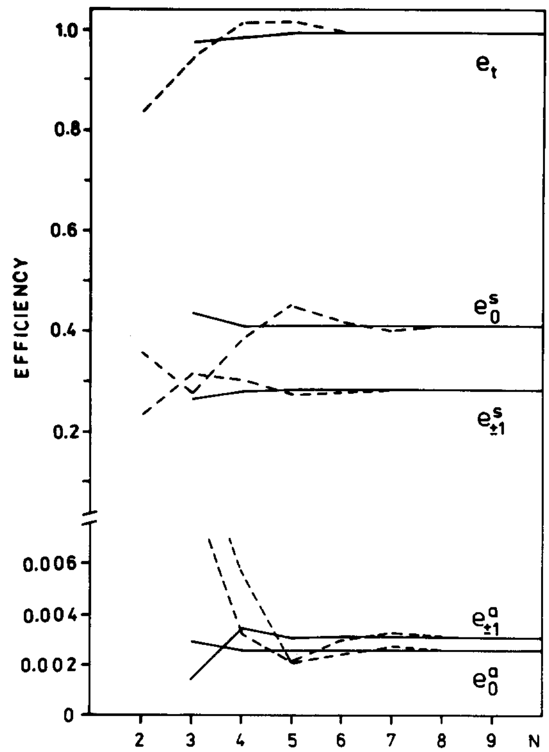

Figure 7. As figure 5 except for $n_2 = 1{\cdot}5$.

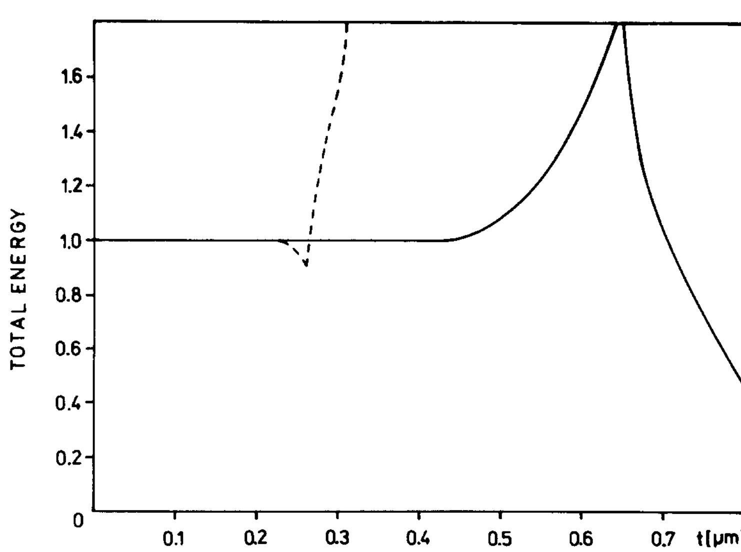

Figure 8. Total diffracted energy of coated dielectric grating as a function of the layer thickness at normal incidence. $n_1 = 1$, $n_2 = 2{\cdot}3$, $n_3 = 1{\cdot}5$, $h/d = 0{\cdot}167$, $\lambda = 0{\cdot}6328\,\mu$m, $d = 0{\cdot}3\,\mu$m.

Above some critical layer thickness, different for the two methods, the total energy diverges. Surprisingly, with increasing the N, the critical thickness is reduced. The same tendency appears in figure 9 for a symmetrical structure. TE polarized light is incident normally on the grating with period $d = 0{\cdot}37\,\mu m$ and $h/d = 0{\cdot}22$. Although the rate of convergence $N = 3$ is one and the same for both methods, the critical thickness for RF method is $t_R = 0{\cdot}175\,\mu m$, while for the C method this value is four times greater ($t_C = 0{\cdot}625\,\mu m$). Again both t_R and t_C rapidly decrease when N is increased.

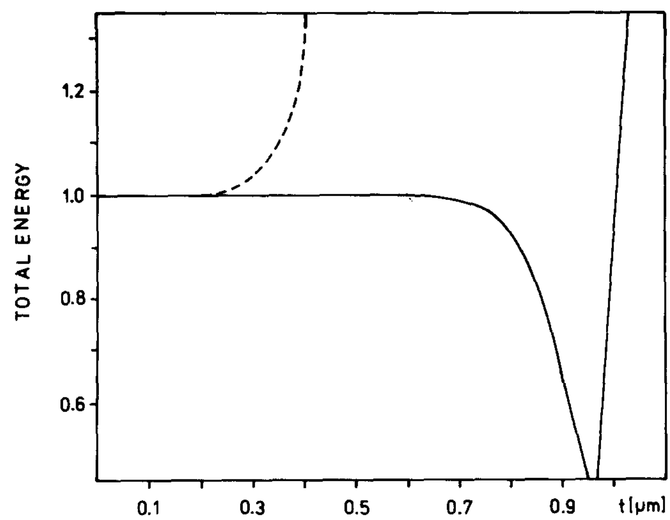

Figure 9. As figure 8 except for $n_3 = 1$, $d = 0{\cdot}37\,\mu m$, $h/d = 0{\cdot}216$.

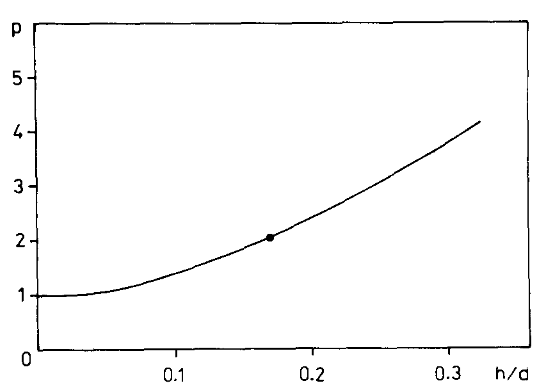

Figure 10. Ratio $p = H_N^{RF}/H_N^C$ depending on the ratio h/d.

5. Discussion

In [11] Wirgin has shown that the Rayleigh approximation in its RF form has great advantages and overcomes the theoretically predicted limit of $h/d \leqslant 0{\cdot}072$. Our results support this conclusion. Figure 3 extends the possibility of RF method up to ratio $h/d = 0{\cdot}35$ for metal gratings and in figure 6 the results are given up to $h/d = 0{\cdot}3$ for dielectric gratings. However, for moderate and high ratio $h/d > 0{\cdot}25$ the convergence of the C method is better.

Increasing the groove depth the difference between the exponential factors $h^{RF}_{m,j}$ and $h^{C}_{m,j}$ becomes larger. Since the exponents in the C method are calculated from the truncated equations, they would be 'better' ones in comparison with those of the RF method, therefore a better convergence is ensured. This provides for shorter computation times (and, if necessary, less memory requirements), which usually compensates the sophistication of the computer code. In our opinion the simpler RF method is more useful for gratings with ratio h/d up to 0·25, while for deeper gratings a more complicated rigorous method must be used. For example, the C method is a rigorous one and does not require such sophisticated and time consuming computer codes, as for instance the integral methods.

The application of the C method even for shallow gratings is preferable in comparison with the RF method when coated gratings are investigated. The divergence of the efficiency upon some critical thickness of the layer is due to the finite computer wordlength. The existence of such a difficulty was mentioned by Chandezon *et al.* [18], but no explanation was given.

For the propagating waves the modulus of the exponent in (6) is less than or equal to unity. The evanescent waves have complex propagation factors thus in (6) real exponential members are included. After matrix multiplication, if the real exponents are large enough, the members of order of unity would be truncated leading to a divergence of the calculations.

The values of $h^{RF}_{m,j}$ and $h^{C}_{m,j}$, $-N \leqslant m \leqslant N$, with $N = 3$ for the middle layer in the case of figure 8 are presented in table 1. The computer wordlength of 32 bits means

Table 1.

	$h^{C}_{m,2}/k$		$h^{RF}_{m,2}/k$	
m	real	imaginary	real	imaginary
0	2·2966	0	2·3	0
1	0·9171	0	0·9169	0
2	0·9987	2·8787	0	3·5365
3	0·9980	2·8791	0	5·8952

Table 2.

N	H^{C}_{N}	H^{RF}_{N}
2	1·843	3·537
3	2·879	5·895
4	3·985	8·118
5	5·078	10·293
6	6·173	12·445

that in the floating point arithmetic 23 bits are reserved for the mantissa of a given number. If

$$\max_m |\operatorname{Im}(h_m)|/k = H_N, \quad -N \leqslant m \leqslant N, \tag{13}$$

then the maximum thickness t_{max} for which the calculations would converge can be given as

$$2\pi t_{max} H_N / \lambda < \ln(2^{23}). \tag{14}$$

The calculated values of t_{max} from (14) using table 1 are $t_{max}^{RF} = 0{\cdot}27\ \mu m$ and $t_{max}^{C} = 0{\cdot}56\ \mu m$, while the critical thicknesses from figure 9 are 0·175 and 0·625 μm respectively.

In table 2 the values of H_N^{RF} and H_N^{C} are represented for different N. As it can be seen the ratio $H_N^{RF}/H_N^{C} \approx 2$ is independent of N. From the fact that $H_N^{RF} > H_N^{C}$ it directly follows that $t_{max}^{RF} < t_{max}^{C}$ even if the threshold values of N for the two methods are equal. The calculated values of the ratio H_N^{RF}/H_N^{C} for the system of figure 8 are displayed in figure 10; the same tendency of independence of the ratio on N is obtained. With a solid circle the position of the results from table 2 is depicted. Taking into account the more rapid convergence of the C method, the ratio t_{max}^{C}/t_{max}^{RF} would be even greater for deeper gratings. Therefore the C method enables to investigate coated gratings with a layer thickness at least four times greater than with the RF method (for $h/d = 0{\cdot}32$). Of course, using double precision, the maximum values of the thickness are doubled but the ratio t_{max}^{RF}/t_{max}^{C} remains the same.

References

[1] PETIT, R., (ed.), 1980, *Electromagnetic Theory of Gratings* (Berlin: Springer).
[2] MAYSTRE, D., 1984, *Progress in Optics*, vol. 21, edited by E. Wolf (Amsterdam: North-Holland).
[3] PETIT, R., 1975, *Nouv. Rev. Optics*, **6**, 129.
[4] MILLAR, R. F., 1973, *Radio Sci.*, **8**, 785.
[5] MAYSTRE, D., 1972, *Optics Commun.*, **6**, 50.
[6] MAYSTRE, D., 1973, *Optics Commun.*, **8**, 216.
[7] MAYSTRE, D., 1978, *J. opt. Soc. Am.*, **68**, 490.
[8] NEVIERE, M., VINCENT, P., and PETIT, R., 1974, *Nouv. Rev. Optics*, **5**, 65.
[9] RAYLEIGH, LORD, 1907, *Proc. R. Soc.* A, **79**, 399.
[10] PETIT, R., and CADILHAC, M., 1966, *C. r. hebd. Seanc. Acad. Sci., Paris*, B, **262**, 468.
[11] WIRGIN, A., 1980, *Optica Acta*, **27**, 1671.
[12] WIRGIN, A., 1981, *Optica Acta*, **28**, 1377.
[13] MASHEV, L., and POPOV, E., 1985, *Optics Commun.*, **55**, 377.
[14] POPOV, E., MASHEV, L., and MAYSTRE, D., 1986, *Optica Acta*, **33**, 607.
[15] MASHEV, L., and POPOV, E., 1984, *Optics Commun.*, **51**, 131.
[16] MAYSTRE, D., 1978, *Optics Commun.*, **26**, 127.
[17] MAYSTRE, D., LAUDE, J. P., GACOIN, P., LEPERE, D., and PRIOU, J. P., 1980, *Appl. Optics*, **19**, 3099.
[18] CHANDEZON, J., DUPUIS, M. T., CORNET, G., and MAYSTRE, D., 1982, *J. opt. Soc. Am.*, **72**, 839.
[19] HUGONIN, J. P., PETIT, R., and CADILHAC, M., 1981, *J. opt. Soc. Am.*, **71**, 593.
[20] PARLETT, B. N., 1980, *The Symmetric Eigenvalue Problem* (New York: Prentice Hall).