OPTICA ACTA, 1981, VOL. 28, NO. 3, 413–428

The dielectric lamellar diffraction grating

I. C. BOTTEN and M. S. CRAIG
School of Mathematical Sciences, The New South Wales Institute of Technology, P.O. Box 123, Broadway, New South Wales 2007, Australia

R. C. McPHEDRAN
Department of Theoretical Physics, The University of Sydney, New South Wales 2006, Australia

J. L. ADAMS and J. R. ANDREWARTHA
Department of Physics, The University of Tasmania, G.P.O. Box 252C, Hobart, Tasmania 7001, Australia

(Received 20 December 1979; revision received 3 June 1980)

Abstract. A rigorous modal theory describing the diffraction properties of a dielectric lamellar grating is presented. The numerical implementation is shown to be suited to modelling the behaviour of high refractive index gratings. This suggests that an approach of this type may be successfully applied to the problem of lossy metallic lamellar gratings.

1. Introduction

The problem of diffraction by a perfectly conducting lamellar grating has long attracted interest [1–4]. The simplicity of the grating profile enabled the specification of modal expansions incorporating explicitly the boundary conditions.

Recently Knop [5] has discussed the problem of diffraction by a dielectric lamellar grating. His interest arose from the suggested use of surface-relief phase-gratings for storing pictorial information [6].

Knop's method relied upon the solution of an eigenvalue problem derived from the Helmholtz equation and involved the use of a Fourier series to describe the discontinuous function equal to the square of the refractive index in the dielectric, and one in air. The differential technique of Nevière *et al.* [7] encountered serious numerical difficulties associated with the use of such series for large refractive index transitions. This was one motive which prompted us to seek an alternative method of solution, with the potential of being generalized to treat a lamellar grating made of a metal with a refractive index complex and large in modulus.

The treatment proposed here for the dielectric lamellar grating is a generalization of the modal treatments appropriate in the infinite conductivity case. The form of the modes was to some extent suggested by functions occurring in the treatment of the asymmetric slab waveguide [8]. We find the method developed here for the dielectric grating to have all the elegance of the classical treatment for perfectly conducting lamellar gratings.

2. Theory

2.1. *Notation and method*

Initially we shall consider the diffraction of a P-polarized plane wave of free space wavelength λ, incident upon the lossless structure (shown in figure 1) at some angle ϕ. The grating extends from $x=-\infty$ to $x=+\infty$ and is perfectly periodic of period d. The incident electric field is defined to be

$$\mathbf{E}^{i}(x,y)=\exp\left[i(\alpha_0 x-\chi_0(y-h/2))\right]\hat{\mathbf{z}}, \tag{1}$$

where $\alpha_0=k_0\sin\phi$, $\chi_0=k_0\cos\phi$ and $k_0=2\pi/\lambda$. (In this paper the temporal dependence of $\exp(-i\omega t)$ has been suppressed from the specification of all field quantities.) Because of the periodicity of the structure all plane wave fields have the same phase shift across a period as the incident field (i.e. all fields are pseudo-periodic.)

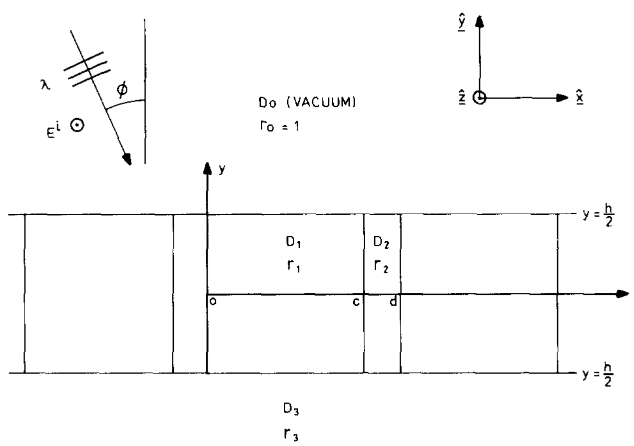


Figure 1. The geometry of the diffraction problem.

The total electric field $\mathbf{E}$ has the same polarization as the incident field and in each region D_i of figure 1 obeys the Helmholtz equations

$$(\nabla^2+k_i^2)\mathbf{E}=0 \quad \text{in region } D_i,$$

where

$$k_i=k_0 r_i \quad \text{for } i=0,1,2 \text{ or } 3,$$

and r_i is the refractive index of medium D_i.

(Because of the polarization-independent nature of the problem we can replace all vector electric fields by scalar fields representing their single non-zero (i.e. z) component.)

In regions D_0 and D_3 the scattered fields are expressed as series of outward-going plane waves (Rayleigh expansions [9]). Thus we write the total electric fields in these regions as

$$E(x,y)=\sum_{p=-\infty}^{\infty}\left[\exp(-i\chi_0(y-h/2))\delta_{p0}+R_p\exp(i\chi_p(y-h/2))\right]\frac{1}{\sqrt{d}}\exp(i\alpha_p x) \quad \text{in } D_0, \tag{2}$$

$$E(x,y)=\sum_{p=-\infty}^{\infty}\left[T_p\exp(-i\eta_p(y+h/2))\right]\frac{1}{\sqrt{d}}\exp(i\alpha_p x) \quad \text{in } D_3, \tag{3}$$

where

$$\alpha_p = \frac{2\pi}{d} p + \alpha_0, \qquad (p \text{ an integer}),$$

$$\begin{aligned} \chi_p &= \sqrt{(k_0^2 - \alpha_p^2)}, && \text{for } |\alpha_p| \leqslant k_0, \\ &= i\sqrt{(\alpha_p^2 - k_0^2)}, && \text{for } |\alpha_p| > k_0, \\ \eta_p &= \sqrt{(k_3^2 - \alpha_p^2)}, && \text{for } |\alpha_p| \leqslant k_3, \\ &= i\sqrt{(\alpha_p^2 - k_3^2)}, && \text{for } |\alpha_p| > k_3, \end{aligned}$$

and δ_{p0} is the Kronecker delta. The sets $\{R_p\}$ and $\{T_p\}$ are the reflection and transmission plane wave coefficients.

The field in the region $D_1 \cup D_2$ is expanded in terms of a modal (eigenfunction) expansion, each term of which analytically obeys the appropriate Helmholtz equations, the necessary continuity conditions relating fields on the left- and right-hand sides of the interfaces:

$$\left. \begin{aligned} E(c^-, y) = E(c^+, y); \quad \frac{\partial E}{\partial x}(c^-, y) = \frac{\partial E}{\partial x}(c^+, y), \\ E(d^-, y) = E(d^+, y); \quad \frac{\partial E}{\partial x}(d^-, y) = \frac{\partial E}{\partial x}(d^+, y), \end{aligned} \right\} \tag{4a}$$

and the pseudo-periodicity conditions

$$\tau E(0^+, y) = E(d^+, y); \quad \tau \frac{\partial E}{\partial x}(0^+, y) = \frac{\partial E}{\partial x}(d^+, y) \tag{4b}$$

with $\tau = \exp(i\alpha_0 d)$, $\alpha_0 d$ being the phase shift of the incident field across one period.

Our approach to the problem differs substantially from the method of Knop [5] and Burckhardt [10]. The individual eigenmodes $\{e_n\}$ used by these authors are a sum of terms possessing the same continuity and pseudo-periodicity properties as the field. The Helmholtz equation

$$\frac{\partial^2 e_n}{\partial x^2} + \frac{\partial^2 e_n}{\partial y^2} + k_0^2 r^2(x) e_n = 0$$

(where $r^2(x)$ is the square of the refractive index profile of the grating and is expanded in a Fourier series) is then transformed into an infinite set of homogeneous linear equations (i.e. an eigenvalue problem). This system must naturally be truncated to permit a numerical solution but it would appear that the dimension of this matrix is highly dependent upon the size of the refractive index discontinuity (i.e. the number of terms that have to be considered in the slowly convergent Fourier series for $r^2(x)$).

It is clear that this treatment is of a sufficiently general form to encompass structures having periodically modulated refractive index profiles of arbitrary form. However, it is our opinion that the method is best suited to continuous refractive index profiles but is less well suited to the consideration of profiles having discontinuities.

In contrast, our analysis is specific to this problem in that it only handles step function refractive index profiles. However, this permits us to describe the eigenmodes and hence the field in a manner which is more elegant analytically and less involved numerically.

Our eigenmodes each involve only a single term (cf. the series of Knop's approach) and the resulting eigenvalue problem necessitates only the solution of a single transcendental equation. (Note that at no stage in the calculation of our modes is any truncation of infinite series required.) It may also be shown that the lossless nature of the structure gives rise to orthogonal modal functions, thereby facilitating the application of the method of moments [11] to solve the diffraction problem.

2.2. *The eigenvalue problem*

2.2.1. *Separation of variables*

We wish to solve the Helmholtz equation of the form

$$\frac{\partial^2 E}{\partial x^2}+\frac{\partial^2 E}{\partial y^2}+k_0^2 r^2(x)=0, \quad 0\leqslant x\leqslant d,$$

where

$$r(x)=\begin{cases} r_1, & 0<x<c, \\ r_2, & c<x<d, \end{cases}$$

and E satisfies the conditions given in equations (4). Since the geometry of the problem permits a separation of variables, i.e.

$$E(x,y)=u(x)v(y),$$

we obtain for u

$$\left.\begin{aligned} u''+k_1^2u=\mu^2u, &\quad 0<x<c, \\ u''+k_2^2u=\mu^2u, &\quad c<x<d, \end{aligned}\right\} \tag{5}$$

and for v

$$v''+\mu^2v=0. \tag{6}$$

Equations (5) can be rewritten in the more concise form

$$u''+\zeta^2S(x-c)u=-\beta^2u, \tag{7}$$

where

$$\left.\begin{aligned} \zeta^2&=k_2^2-k_1^2, \\ \beta^2&=k_1^2-\mu^2 \end{aligned}\right\} \tag{8}$$

and

$$S(x)=\begin{cases} 0, & \text{for } x<0, \\ 1, & \text{for } x\geqslant 0. \end{cases}$$

The boundary conditions imply that u and u' are continuous at $x=c$ and pseudo-periodic so that

$$\tau u(0^+)=u(d^-), \tag{9}$$

$$\tau u'(0^+)=u'(d^-), \tag{10}$$

where $\tau=\exp(i\alpha_0 d)$.

Equations (7–10) define an eigenvalue problem, and as we shall see in §2.2.2 solutions exist only for a countably infinite discrete set of real eigenvalues (β^2).

Let θ, ψ be two linearly independent solutions of equation (7) which are continuous and continuously differentiable at $x=c$ such that

$$\left.\begin{aligned} \theta(0)=1, \quad \psi(0)=0, \\ \theta'(0)=0, \quad \psi'(0)=1. \end{aligned}\right\} \tag{11}$$

Then

$$\theta=\begin{cases} \cos(\beta x), & 0\leqslant x\leqslant c, \\ \cos(\beta c)\cos[\gamma(x-c)]-\dfrac{\beta}{\gamma}\sin(\beta c)\sin[\gamma(x-c)], & c\leqslant x\leqslant d \end{cases} \tag{12}$$

and

$$\psi=\begin{cases} \dfrac{1}{\beta}\sin(\beta x), & 0\leqslant x\leqslant c, \\ \dfrac{1}{\beta}\sin(\beta c)\cos[\gamma(x-c)]+\dfrac{1}{\gamma}\cos(\beta c)\sin[\gamma(x-c)], & c\leqslant x\leqslant d, \end{cases} \tag{13}$$

where $\gamma^2=\beta^2+\zeta^2$. (14)

Since θ, ψ satisfy equation (7), the wronskian

$$W(\theta,\psi)=\theta\psi'-\theta'\psi$$

is constant and so from equations (11)

$$\theta\psi'-\theta'\psi=1. \tag{15}$$

We now set

$$u(x)=A\theta(x)+B\psi(x) \tag{16}$$

and apply the pseudo-periodicity constraints (9) and (10). Thus,

$$\tau A=A\theta(d)+B\psi(d),$$
$$\tau B=A\theta'(d)+B\psi'(d).$$

For this homogeneous set of linear equations to have a non-trivial solution we require that

$$[(\theta(d)-\tau)][\psi'(d)-\tau]-\theta'(d)\psi(d)=0.$$

Using equation (15) this becomes

$$\tau[\theta(d)+\psi'(d)]=1+\tau^2,$$

that is

$$\theta(d)+\psi'(d)=2\cos(\alpha_0 d), \tag{17}$$

using the definition of τ. On expanding equation (17) explicitly (using equations (12)) we obtain the eigenvalue equation

$$\cos(\beta c)\cos(\gamma g)-\frac{1}{2}\frac{(\beta^2+\gamma^2)}{\beta\gamma}\sin(\beta c)\sin(\gamma g)=\cos(\alpha_0 d), \tag{18}$$

where $g=d-c$. We choose β to be that root of β^2 having a positive real part.

For any eigenvalue β satisfying (14) and (18) then

$$\frac{A}{B}=\frac{\psi(d)}{\tau-\theta(d)}=\frac{\tau-\psi'(d)}{\theta'(d)}.$$

Also the solution of equation (6) may be written as

$$v(y)=a\sin(\mu y)+b\cos(\mu y),$$

where μ is specified in equation (8). Thus the separable solution has the form

$$E(x,y)=\left(\frac{A}{B}\theta(x)+\psi(x)\right)(a\sin(\mu y)+b\cos(\mu y)).$$

Note that when $\zeta=0$, then $\beta=\gamma$ and so equation (18) degenerates to

$$\cos(\beta d)=\cos(\alpha_0 d),$$

that is

$$\beta=\frac{2\pi p}{d}+\alpha_0$$

for any integer p, which is simply the classical grating equation.

2.2.2. *Properties of the eigenvalue problem*

In this section we shall discuss

(i) the self-adjoint nature of the boundary value problem specified in equations (7)–(10);

(ii) the distribution of the eigenvalues; and

(iii) the completeness of the eigenfunctions.

(i) *Self-adjointness*

We define the operator

$$L=\left[\frac{d^2}{dx^2}+\zeta^2 S(x-c)\right]$$

and note that for any continuously differentiable functions f, g obeying the boundary conditions (9) and (10)

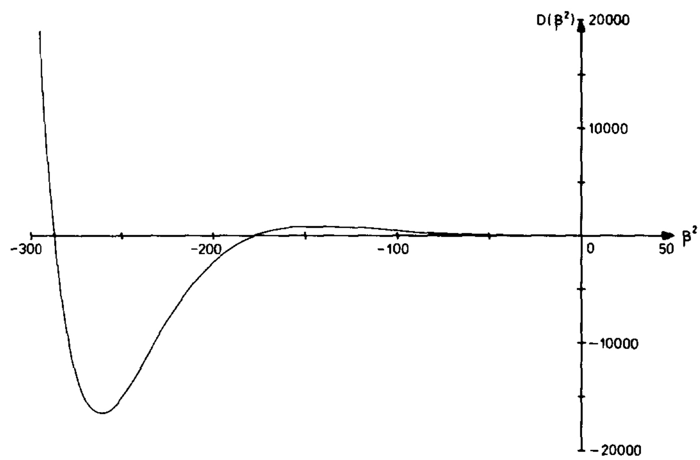
$$\begin{aligned}\int_0^d(\bar{g}Lf-f\overline{Lg})\,dx&=\int_0^d(\bar{g}[f''+\zeta^2S(x-c)f]-f[\bar{g}''+\bar{\zeta}^2S(x-c)\bar{g}])\,dx,\\&=[\bar{g}f'-f\bar{g}']_0^d \quad \text{(since } \zeta^2 \text{ is real)},\\&=0 \quad \text{(by the boundary conditions)}.\end{aligned}$$

In the above, the bar denotes complex conjugation.

Hence L is a self-adjoint operator and consequently has real eigenvalues (β^2), and eigenfunctions corresponding to distinct eigenvalues are orthogonal.

(ii) *Distribution of eigenvalues*

It may be shown [12] that $D(\beta^2) = \theta(d) + \psi'(d)$, taken as a function of β^2 always has behaviour of the type illustrated in figure 2. Firstly $D(\beta^2)$ is continuous. Secondly, $D(\beta^2) > 2$ for $\beta^2 <$ some number Ω. (It may be shown that $\Omega \geqslant -\zeta^2$.) For $\beta^2 > \Omega$, $D(\beta^2)$ oscillates infinitely often; its maxima always being greater than 2 and its minima always being less than -2. $D(\beta^2)$ is always monotonic between values of β^2 where $D(\beta^2) = -2$ and $D(\beta^2) = 2$. Hence the eigenvalue equation (18) always has infinitely many discrete solutions $\beta_1^2, \beta_2^2, \beta_3^2, \ldots$ with $\beta_n^2 \to \infty$ as $n \to \infty$. It may also be shown that provided $\zeta^2 \neq 0$ there is only one eigenfunction $u_m(x)$ corresponding to any eigenvalue β_m^2 (except possibly at certain isolated points if $\tau = \pm 1$).



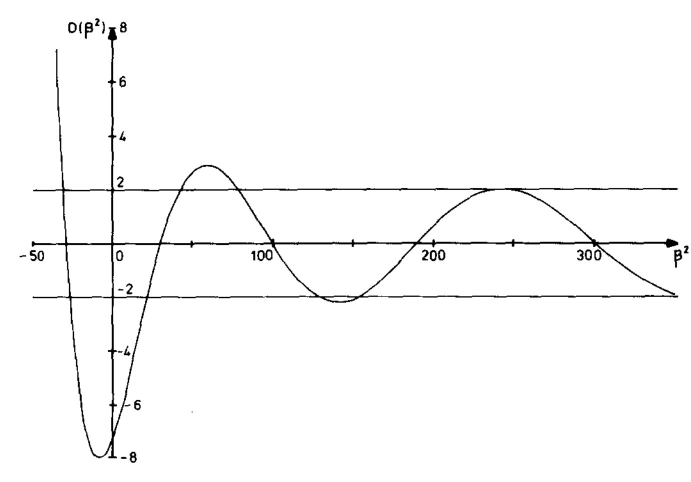


Figure 2. Plot of the function $D(\beta^2)$. The eigenvalues correspond to points where $D(\beta^2) = 2\cos\alpha_0 d$. In the first case shown, at least three eigenvalues occur for $\beta^2 < 0$. In the second case only one eigenvalue occurs for $\beta^2 < 0$ while the first four eigenvalues for $\beta^2 > 0$ are shown.

(iii) *Completeness*

It can further be shown [13] that the countably infinite set of eigenfunctions $\{u_m\}$ forms a complete set. That is to say, any function that is continuous, differentiable and obeys the pseudo-periodicity condition may be expressed as a linear combination of the $\{u_m\}$, i.e.

$$z(x)=\sum_m z_m u_m(x),$$

where

$$z_m=\int_0^d z(x)\bar{u}_m(x)\,dx$$

such that the $\{u_m\}$ are normalized according to

$$\int_0^d |u_m|^2\,dx=1.$$

Hence it follows that the general solution for the total electric field E can be expressed in the form

$$E(x,y)=\sum_m (a_m \sin(\mu_m y)+b_m \cos(\mu_m y))u_m(x),$$

which will now be matched to the plane wave fields at the boundaries $y=\pm h/2$.

2.3. *Method of moments*

With the form of the fields in all regions having been prescribed, we can now solve the diffraction problem by applying the method of moments. The continuity of the electric fields across the interface at $y=h/2$ may be expressed by

$$\sum_p [R_p+\delta_{p0}]\frac{1}{\sqrt{d}}\exp(i\alpha_p x)=\sum_m (a_m^*+b_m^*)u_m(x), \quad \text{for} \quad 0\leqslant x\leqslant d, \tag{19}$$

from which we obtain a system of equations of the form

$$a_n^*+b_n^*=\sum_p (R_p+\delta_{p0})\bar{J}_{pn}, \quad \forall n \tag{20}$$

by multiplying equation (19) by $\bar{u}_n$ and integrating over the interval $[0,d]$. In the above

$$a_n^*=a_n \sin(\mu_n h/2),$$

$$b_n^*=b_n \cos(\mu_n h/2),$$

and

$$J_{pn}=\frac{1}{\sqrt{d}}\int_0^d \exp(-i\alpha_p x)u_n(x)\,dx.$$

The system of equations (20) may be cast in the more concise form†

$$a^*+b^*=J^{\mathrm{H}}(R+F), \tag{21}$$

† In equation (21), and in what follows, the superscript 'H' denotes the hermitian conjugate of a matrix.

where $a^*=[a_n^*]$, $b^*=[b_n^*]$, $J=[J_{pn}]$, $R=[R_p]$ and F is a vector whose entries are $F_p=\delta_{p0}$. Similarly, continuity of E at $y=-h/2$ gives

$$-a^*+b^*=J^{\mathrm{H}}T, \tag{22}$$

where

$$T=[T_p].$$

Applying the continuity of $\partial E/\partial y$ at $y=h/2$ we have

$$\sum_p i\chi_p[R_p-\delta_{p0}]\frac{1}{\sqrt{d}}\exp(i\alpha_p x)=\sum_m(D_{1m}a_m^*+D_{2m}b_m^*)u_m(x), \quad \text{for} \quad 0\leqslant x\leqslant d, \tag{23}$$

where

$$\begin{aligned} D_{1m}&=\mu_m\cot(\mu_m h/2),\\ D_{2m}&=-\mu_m\tan(\mu_m h/2). \end{aligned} \tag{24}$$

From this we derive a system of equations,

$$i\chi_q[R_q-\delta_{q0}]=\sum_m(D_{1m}a_m^*+D_{2m}b_m^*)J_{qm},$$

by multiplying equation (23) by $1/\sqrt{d}\exp(-i\alpha_q x)$ and integrating over $[0,d]$. Expressing this in matrix notation reduces the above equation to

$$R=F-i\chi^{-1}J(D_1a^*+D_2b^*), \tag{25}$$

where

$$\begin{aligned} \chi&=\operatorname{diag}(\chi_p),\\ D_1&=\operatorname{diag}(D_{1m}),\\ D_2&=\operatorname{diag}(D_{2m}). \end{aligned}$$

Similarly at $y=-h/2$ we derive

$$T=i\eta^{-1}J(D_1a^*-D_2b^*), \tag{26}$$

where $\eta=\operatorname{diag}(\eta_p)$.

Substituting equations (25) and (26) into (21) and (22) we arrive at a coupled pair of linear equations in the $\{a_m^*\}$ and $\{b_m^*\}$

$$\begin{bmatrix} iJ^{\mathrm{H}}\chi^{-1}JD_1+I & iJ^{\mathrm{H}}\chi^{-1}JD_2+I \\ -iJ^{\mathrm{H}}\eta^{-1}JD_1-I & iJ^{\mathrm{H}}\eta^{-1}JD_2+I \end{bmatrix}\begin{bmatrix} a^* \\ b^* \end{bmatrix}=2\begin{bmatrix} J^{\mathrm{H}}F \\ 0 \end{bmatrix}, \tag{27}$$

which is solved by standard elimination techniques. In equation (27), I denotes the identity matrix. Reconstruction of the reflected and transmitted amplitudes takes place using equations (25) and (26). Propagating order efficiencies are then given by

$$\rho_p^{\mathrm{R}}=\frac{\chi_p}{\chi_0}|R_p|^2$$

and

$$\rho_p^{\mathrm{T}} = \frac{\eta_p}{\chi_0} |T_p|^2.$$

2.4. *Differences in the formulation of the S-polarization problem*

The modifications arise because of the different forms taken by the boundary conditions for this polarization. For this problem, the magnetic field vector is aligned with the z-axis and we express the boundary conditions in terms of the scalar magnetic field H as follows: both H and $1/r^2\ \partial H/\partial n$ are continuous at each interface and are also pseudo-periodic. Furthermore, the appropriate form of the wave equation is

$$\frac{\partial}{\partial x}\left[\frac{1}{k^2}\frac{\partial H}{\partial x}\right]+\frac{\partial}{\partial y}\left[\frac{1}{k^2}\frac{\partial H}{\partial y}\right]+H=0, \tag{28}$$

where $k^2 = k_0^2 r^2(x)$.

This may be written in the form

$$\frac{\partial^2 H}{\partial x^2}+\frac{\partial^2 H}{\partial y^2}+k^2 H=\frac{1}{k^2}\frac{\partial H}{\partial x}\zeta^2\delta(x-c),$$

a non-homogeneous Helmholtz equation with the distributive term on the right-hand side representing the discontinuity in the normal derivative of H at $x=c$.

We shall now outline the essential changes to the eigenvalue problem. Again, we assume a separable solution

$$H(x,y)=u(x)v(y)$$

of the wave equation and obtain

$$k^2\left[\frac{1}{k^2}u'\right]'+\zeta^2 S(x-c)u=-\beta^2 u, \tag{29}$$

$$v''+\mu^2 v=0, \tag{30}$$

where

$$\mu^2=k_1^2-\beta^2.$$

The boundary conditions imply that

$$\begin{aligned} u(c^-)&=u(c^+); \quad \frac{1}{r_1^2}u'(c^-)=\frac{1}{r_2^2}u'(c^+)\\ (d^-)&=u(d^+); \quad \frac{1}{r_2^2}u'(d^-)=\frac{1}{r_1^2}u'(d^+)\end{aligned} \tag{31}$$

with pseudo-periodicity giving

$$\tau u(0^+)=u(d^+); \quad \frac{1}{r_1^2}\tau u'(0^+)=\frac{1}{r_1^2}u'(d^+). \tag{32}$$

Now defining θ, ψ as the two linearly independent solutions of (29) obeying the same initial conditions as given in equation (11) and having the properties of u given in equation (31) we have that

$$\theta(x)=\begin{cases}\cos(\beta x), & \text{for } 0\leqslant x\leqslant c\\[2ex] \cos(\beta c)\cos[\gamma(x-c)]-\dfrac{r_2^2}{r_1^2}\dfrac{\beta}{\gamma}\sin(\beta c)\sin[\gamma(x-c)], & \text{for } c\leqslant x\leqslant d\end{cases} \tag{33}$$

and

$$\psi(x)=\begin{cases}\dfrac{1}{\beta}\sin(\beta x) & \text{for } 0\leqslant x\leqslant c\\[2ex] \dfrac{1}{\beta}\sin(\beta c)\cos[\gamma(x-c)]+\dfrac{r_2^2}{r_1^2}\dfrac{1}{\gamma}\cos(\beta c)\sin[\gamma(x-c)] & \text{for } c\leqslant x\leqslant d\end{cases} \tag{34}$$

where $\gamma^2=\beta^2+\zeta^2$.

From the differential equation (29) and the initial conditions on θ and ψ (equation (11)) it may be shown that the appropriate wronskian is

$$W(\theta,\psi)=\frac{1}{r^2(x)}(\theta\psi'-\theta'\psi)=\frac{1}{r_1^2}. \tag{35}$$

Using the basis functions θ,ψ we now construct u in the form

$$u(x)=A\theta(x)+B\psi(x).$$

From the boundary conditions (32) the following homogeneous set of equations

$$\tau A=A\theta(d^-)+B\psi(d^-)$$

$$\frac{1}{r_1^2}\tau B=\frac{1}{r_2^2}[A\theta'(d^-)+B\psi'(d^-)]$$

are obtained, and for these to have non-trivial solutions we require that

$$[\theta(d^-)-\tau]\left[\frac{1}{r_2^2}\psi'(d^-)-\frac{1}{r_1^2}\tau\right]-\psi(d^-)\frac{1}{r_2^2}\theta'(d^-)=0,$$

or more simply

$$\theta(d^-)+\frac{r_1^2}{r_2^2}\psi'(d^-)=2\cos(\alpha_0 d), \tag{36}$$

using the result of equation (35). Equation (36) may be expressed in the more explicit form

$$\cos(\beta c)\cos(\gamma g)-\frac{1}{2}\left(\frac{r_2^2}{r_1^2}\frac{\beta}{\gamma}+\frac{r_1^2}{r_2^2}\frac{\gamma}{\beta}\right)\sin(\beta c)\sin(\gamma g)=\cos(\alpha_0 d) \tag{37}$$

—the S-polarization eigenvalue equation which has properties similar to those mentioned in part (ii) of § 2.2.2. Thus in region $D_1\cup D_2$ the field H is expanded in terms of functions of the form

$$u(x)v(y)=\left[\frac{A}{B}\theta(x)+\psi(x)\right][a\sin(\mu y)+b\cos(\mu y)]$$

where

$$\frac{A}{B}=\frac{\psi(d^-)}{\tau-\theta(d^-)}=\frac{\tau r_2^2/r_1^2-\psi'(d^-)}{\theta'(d^-)}.$$

We now demonstrate the self-adjoint nature of the differential operator L defined by

$$Lu=k^2\left[\frac{1}{k^2}u'\right]'+\zeta^2 S(x-c)u.$$

Consider the following integral for any functions f, g obeying the same boundary conditions as u:

$$\begin{aligned}
&\int_0^d \frac{1}{k^2(x)}[\bar{g}Lf-f\overline{Lg}]\,dx\\
&\qquad=\int_0^d\left\{\left[\frac{1}{k^2}f'\right]'\bar{g}+\frac{\zeta^2}{k^2}S(x-c)f\bar{g}\right\}-\left\{\left[\frac{1}{k^2}\bar{g}'\right]'f+\frac{\bar{\zeta}^2}{k^2}S(x-c)\bar{g}f\right\}dx,\\
&\qquad=\left[\frac{1}{k^2}f'\bar{g}-\frac{1}{k^2}f\bar{g}'\right]_0^d \quad \text{(since } \zeta^2 \text{ is real)},\\
&\qquad=0 \qquad\qquad \text{(by the boundary conditions)}.
\end{aligned}$$

Thus L is a self-adjoint operator for the inner product defined with respect to the weight distribution $1/k^2(x)$ and hence L has real eigenvalues (β^2). Its eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the above weight distribution.

It may also be shown by an extension of methods given in [13] that the eigenfunctions form a complete set and so we expand H in region $D_1 \cup D_2$ in the series

$$H(x,y)=\sum_m [a_m \sin(\mu_m y)+b_m \cos(\mu_m y)]u_m(x).$$

The solution of the diffraction problem then proceeds by the method of moments with the fields in regions D_0 and D_3 once again being expressed in terms of the Rayleigh expansions (equations (2) and (3)). The analysis of § 2.3 is unaltered except for the redefinition of the inner product J_{pn} which now becomes

$$J_{pn}=\frac{1}{\sqrt{d}}\int_0^d \frac{1}{r^2(x)}\exp(-i\alpha_p x)u_n(x)\,dx,$$

and the replacement of η by η/r_3^2.

3. Verifications and applications of the theory

The formalisms for P- and S-polarizations were numerically tested using the reciprocity theorem and also the symmetry properties discussed in [14]. The above criteria were satisfied to high accuracy. In addition, the results were compared with the graphs given by Knop [5], again with good agreement being observed.

It was decided to test the accuracy of the theory in the region of strong diffraction anomalies, not only for weakly reflecting materials (such as glass) but also for the case of a strongly reflecting material (having a refractive index of 5·0). The results of this study are shown in figures 3–8. The convergence of results with the increasing number of modes and the numerical accuracy of field matching were all satisfied to much better than 1 per cent for all points on the curves. The criterion of conservation of energy supplies no convergence information at all since it may be shown that it holds analytically within the formalism.

Figures 3 and 4 illustrate the behaviour of a glass refraction grating used with normally incident light. Both the total energy transmission and the energy carried by the zeroth transmitted order are shown. It is interesting that the grating is more strongly diffracting (i.e. disperses more energy) for P-polarization than for S-polarization. Also the diffraction anomalies are much stronger for the former polarization. The P-polarization anomalies are dark bands of the type first observed for reflection gratings by Palmer [15]. It will be noted that for refraction gratings, Wood anomalies occur at differing wavelengths in air and in the substrate. For the pth order in air the Rayleigh wavelength is

$$\lambda_R = \begin{cases} d(+1-\sin(\phi))/p, & \text{for all } p>0, \\ d(+1+\sin(\phi))/(-p), & \text{for all } p<0, \end{cases} \tag{38}$$

while for the pth order in the dielectric substrate of refractive index r_3 it is

$$\lambda'_R = \begin{cases} d(+r_3-\sin(\phi))/p, & \text{for all } p>0, \\ d(+r_3+\sin(\phi))/(-p), & \text{for all } p<0. \end{cases} \tag{39}$$

For very long wavelengths the energy transmitted in both polarizations tends to the geometrical optics limit of 96 per cent.

For the case of a transmission grating in glass (see figures 5 and 6) the diffraction anomalies in the two polarizations take the form of dark bands. The S-polarization anomaly is the sharper of the two, as is generally the case with diffraction gratings. At long wavelengths the transmittance tends towards unity, since the incident wave no longer 'resolves' the grating structure.

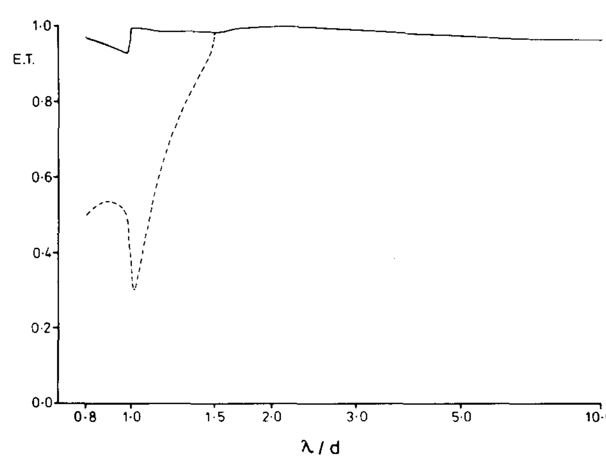


Figure 3. The total transmitted energy (solid line) and the energy transmitted by the zeroth order (broken line) are shown as a function of normalized wavelength (λ/d) for normally incident light ($\phi=0°$). The grating parameters are $c/d=0{\cdot}6$, $h/d=0{\cdot}4$, $r_1=1{\cdot}0$, $r_2=r_3=1{\cdot}5$. P-polarized light.

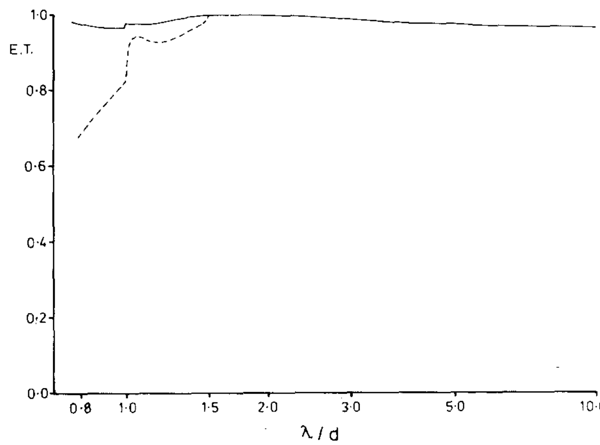


Figure 4. As for figure 3, but with S-polarized light.

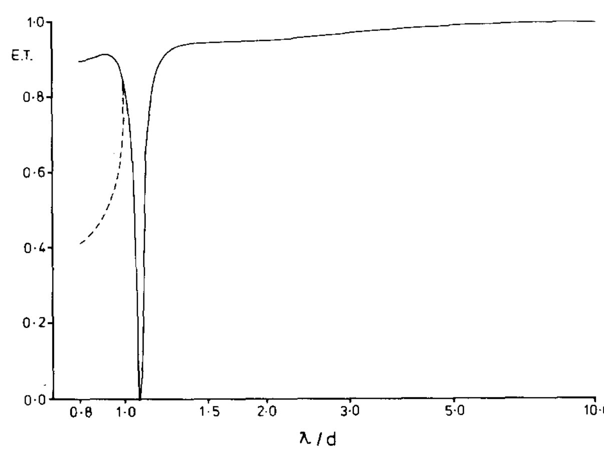


Figure 5. As for figure 3, but with $r_2 = 1{\cdot}5$, $r_3 = 1{\cdot}0$.

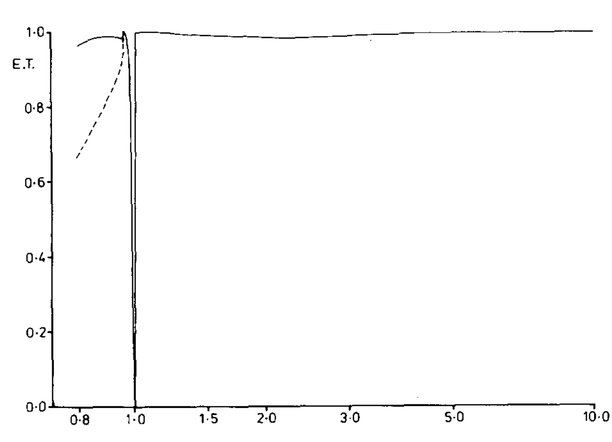


Figure 6. As for figure 4, but with $r_2 = 1{\cdot}5$, $r_3 = 1{\cdot}0$.

The behaviour of the high refractive index transmission grating of figures 7 and 8 is very much more complicated than that of the low refractive index structures of figures 3–6. The total transmitted energy in both polarizations undergoes a multiplicity of sharp resonance anomalies. We attribute these anomalies to the excitation of 'leaky' resonant cavity modes of the type discussed by Hessel and Oliner [16]. It is to be noted that the consequence of increasing the refractive index of the dielectric is an increase in the reflectance of the side walls of the leaky cavity and thus in the Q-factor of its resonances. Also because of the large value of r_2, the cavity can support a large number of 'propagating' modes (i.e. modes for which μ_m is real) and these resonate in the vicinity of their cut-off wavelengths. Further work devoted to understanding the detailed behaviour of these resonance anomalies is warranted.

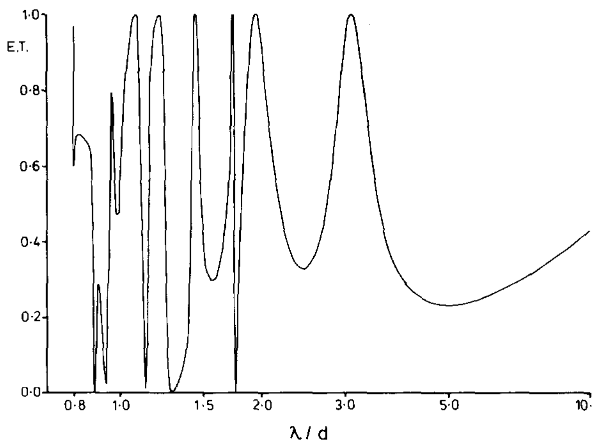


Figure 7. As for figure 3, but with $r_2 = 5{\cdot}0$, $r_3 = 1{\cdot}0$. (To simplify the figure, the zeroth-order energy transmission curve has been omitted.)

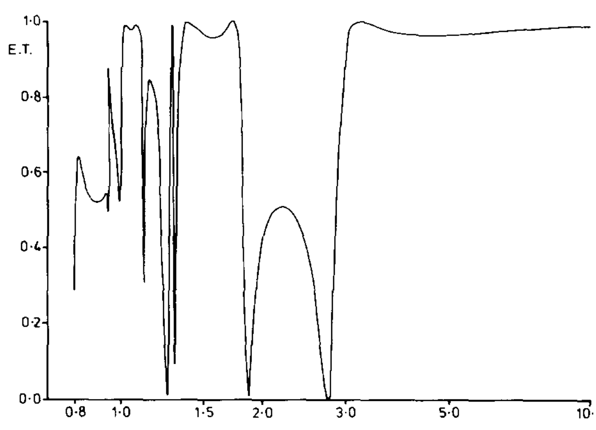


Figure 8. As for figure 7, but with S-polarized light.

4. Conclusions

The theoretical method outlined here represents the generalization of the classical modal formalisms for perfectly conducting lamellar gratings. It has been developed as part of a process designed to lead to modal methods for lossy lamellar gratings and integral methods for bimetallic gratings (see [17]). Work has already commenced on these extensions and it is hoped to report on these at an early stage.

Acknowledgments

This work was undertaken while two of the authors (J. L. A. and J. R. A.) were recipients of Commonwealth of Australia Post-Graduate Research Fellowships. R. McPhedran acknowledges financial assistance provided by the Australian Research Grants Committee, and the Science Foundation for Physics within the University of Sydney for the provision of facilities.

On présente une théorie modale rigoureuse décrivant les propriétés de diffraction d'un réseau diélectrique lamellaire. On montre que la méthode numérique utilisée convient pour modéliser le comportement de réseaux à haut indice de réfraction. Ceçi suggère qu'une telle approche peut être appliquée avec succès au problème des réseaux métalliques lammellaires présentant des pertes.

Es wird eine streng formale Theorie der Beugungseigenschaftern eines dielektrischen Lamellengitters präsentiert. Die numerische Durchführung erweist sich als geeignet, das Verhalten von Gittern hoher Brechungsindizes modellmäßig zu beschreiben. Dies deutet an, daß eine Näherung dieser Art auf das Problem verlustbehafteter metallischer Lamellengitter erfolgreich angewandt werden könnte.

References

[1] Deriugin, L. N., 1960, *Radiotekhnika*, **15**, 15.
[2] Deriugin, L. N., 1960, *Radiotekhnika*, **15**, 9.
[3] Wirgin, A., and Deleuil, R., 1969, *J. opt. Soc. Am.*, **59**, 1348.
[4] Maystre, D., and Petit, R., 1972, *Optics Commun.*, **5**, 90.
[5] Knop, K., 1978, *J. opt. Soc. Am.*, **68**, 1206.
[6] Knop, K., 1976, *Optics Commun.*, **18**, 298.
[7] Nevière, M., Vincent, P., and Petit, R., 1974, *Nouv. Revue Optique*, **5**, 65.
[8] Marcuse, D., 1974, *Theory of Dielectric Optical Waveguides* (London: Academic Press).
[9] Lord Rayleigh, 1907, *Proc. R. Soc.* A, **79**, 399.
[10] Burckhardt, C. B., 1966, *J. opt. Soc. Am.*, **56**, 1502.
[11] Harrington, R. F., 1968, *Field Computation by Moment Methods* (London: Collier–Macmillan).
[12] Eastham, M. S. P., 1973, *The Spectral Theory of Periodic Differential Equations* (Edinburgh: Scottish Academic Press).
[13] Coddington, E. A., and Levinson, N., 1955, *Theory of Ordinary Differential Equations* (London: McGraw-Hill).
[14] Botten, L. C., Adams, J. L., McPhedran, R. C., and Derrick, G. H., 1980, *J. Optics, Paris*, **11**, 43.
[15] Palmer, C. H., 1952, *J. opt. Soc. Am.*, **42**, 269.
[16] Hessel, A., and Oliner, A. A., 1965, *Appl. Optics*, **4**, 1275.
[17] Botten, L. C., 1980, *J. Optics, Paris*, **11**, 161.