Optical response of a thin film with arbitrary deterministic roughness of the interfaces

Shu Wang* and P. Halevi†
Centro de Investigación en Física, Universidad de Sonora, Apartado Postal 5-088, Hermosillo, Sonora, 83190 Mexico
(Received 5 May 1992)

We present a theoretical study of the optical response of a thin film with arbitrary, deterministic roughness of the interfaces (in one dimension). The three layers of the film are characterized each by a spatially nondispersive dielectric constant and magnetic permeability. The incident light may have TE(s) or TM(p) polarization. Using the Rayleigh hypothesis, we derive an integral matrix equation which relates the reflected fields, the transmitted fields, and the fields inside the thin film to the incident wave. This equation is applied to the special case of periodic corrugation, leading to a solution in terms of an infinite set of linear equations for the amplitudes of the diffracted partial waves in the three media. For aperiodic roughness the numerical solution is still given by a similar set of equations. For periodic (aperiodic) films the solution involves Fourier coefficients (Fourier integrals) of functions related to the roughness profiles. We also derive the secular equations for the polariton eigenmodes of periodic and aperiodic films.

I. INTRODUCTION

The first studies of wave propagation in periodically corrugated waveguides involved optically transparent dielectrics. Dabby, Kesteubaum, and Pack[1] calculated the normal modes assuming identical bounding media (symmetric configuration) and corrugation at both interfaces. On the other hand, Yariv and co-workers[2,3] did the calculation for dissimilar bounding media (asymmetric configuration) with only one interface being corrugated. In these works,[1-3] numerical solutions were derived for specific profiles of the corrugation. Rigorous calculations of propagation modes for arbitrary profiles have been developed by Neviere and co-workers.[4-6] Two reviews by Maystre[7,8] are available on this subject. It should be noted that waveguide modes in corrugated, semiconductor thin-film structures found important applications such as distributed-feedback lasers.[9,10]

The waves that propagate in the above-mentioned waveguides are essentially volume (bulk) modes inside the plate or thin film; usually, they decay exponentially away from the interfaces in the bounding media. In metallic thin films, propagation of surface modes is also possible. By definition, these decay exponentially away from the interfaces inside the thin film, as well as outside. The first experiments on surface-plasmon polaritons in periodically corrugated (silver) films were performed by Pockrand and co-workers.[11-13] It was observed[11] that, with increased height of the corrugation, the phase velocity of the polariton decreases and its damping increases. In Ref. 12, attenuated total-reflection (ATR) spectroscopy was used to excite the two modes each associated with one of the interfaces. These become coupled as a consequence of the periodic corrugation, resulting in mode repulsion[12] in the vicinity of points of interaction of the dispersion curves in the empty-lattice approximation. A convincing confirmation of this effect was provided by an experiment due to Gruhlke, Holland, and Hall.[14] Here, molecules on one side of the film decay by exciting the surface-plasmon mode corresponding to the nearby interface. This mode, via the mode cross-coupling, excites the surface plasmon associated with the other interface, which then radiates by the usual grating-coupling mechanism. In other situations, however, the cross-coupling may be too weak to be observable, as pointed out by Weber and Mills.[15] These authors discard the possibility that cross-coupling played a role in an experiment by Brueck *et al.*,[16] whose purpose actually was to enhance the quantum efficiency of an internal photoemission detector.

The dispersion of polaritons in the immediate neighborhood of the coupling region was investigated analytically by Halevi and Mata-Méndez.[17] The outcome depends on the sign of a coupling constant B, given by a complicated formula. For $B>0$, the interaction results in mode repulsion or in an energy gap ("minigap") depending on whether the unperturbed modes, at their point of intersection, have slopes of the same sign or of different signs, respectively. The possibility $B<0$ leads to a momentum gap or to a simultaneous energy and momentum gap. Unfortunately, a negative value of B has not been found by numerical calculations.

Further experiments, on free-standing silver films corrugated on both sides, were performed by Inagaki *et al.*[18,19] Using a photoacoustic technique, they measured the propagation constant and the resonance half-width as a function of film thickness (at a given frequency). It was found[19] that a considerable discrepancy exists between the experimental values and those calculated from the dispersion relation of the planar (smooth) film. The differences were an order of magnitude greater than the estimated contribution of the corrugation. The authors[19] tentatively attributed the discrepancy to imperfections in the corrugation profile. However, recently, Brudny and Depine[20] calculated the scattering, due to surface plasmons, of a sinusoidal grating on which there was superimposed a statistically varying roughness (which represented the defects of a real grating). These authors found that the dominant peaks of the scattered light were largely unaffected by the random component of the roughness, that is, the corresponding scattering angles were still given by the surface-plasmon polariton dispersion relation for the smooth (plane) surface. Mata-Méndez and Halevi[21] showed, on the basis of a perturbational calculation, that the propagation constant and the attenuation constant for a thin film are both proportional to the corrugation height squared, just as in the case of surfaces. However, the proportionality constants were not evaluated numerically. It seems that the discrepancy reported in Ref. 19 is still open to interpretation.

Recently, Chen and Simon[22] studied the ATR line shape for a silver film, with sawtooth form of corrugation on both sides, in the symmetric configuration (quartz on both sides). The measured and calculated reflectance, as a function of the angle of incidence, were found to be in very good agreement, with the exception of the thinnest film (140 Å thick). Using the same grating structure, Simon and Chen[23] studied optical second-harmonic generation.

A series of papers[24–26] employed the Rayleigh hypothesis in order to investigate the effects of corrugation on the coupled surface plasmons of a thin film. Farias, Maradudin, and Celli assumed identical profiles at the two interfaces, as well as identical boundary media, and took the retardationless limit. On the other hand, Auto, Farias, and Maradudin[25] allowed for retardation and considered the asymmetric configuration with only one interface being corrugated. The dispersion curves, plotted in the first Brillouin zone, include parts that have been "folded in" from other Brillouin zones. In the same geometry, Cavalcante, Farias, and Maradudin[26] calculated the (specular) reflectance, which was dominated by first-order diffraction. It should be mentioned that Chen and Simon[22] noticed disagreement with some of these results.

Up to this point the discussion was limited to periodically corrugated thin films. On the other hand, in optics aperiodic scatterers ("obstacles") are of considerable interest. This topic is related to the inverse scattering problem, namely, the determination of roughness profiles on the basis of the scattering pattern. Another application, in the field of integrated optics, concerns coupling devices. Aperiodic, but deterministically rough, waveguides and stratified media were studied numerically by Hugonin and Petit.[27,28] However, the obstacles are assumed to be localized, and of heights smaller than one wavelength. A more recent work, experimental as well as theoretical, is that of Greffet and Ladan.[29] Good agreement was found for a system of two "ribs" on a germanium substrate.

As far as we are aware, ours is the first calculation of optical response of a thin film with arbitrary, deterministic roughness of both interfaces (in one direction parallel to the film). Thus, the roughness profile may be periodic or aperiodic (obstacle), identical or different at the two interfaces. Our method of calculation resembles that of Toigo *et al.*[30] for a deterministically rough surface; this was extended to a spatially dispersive medium by Wang, Barrera, and Mochán.[31] The three layers are each characterized by their dielectric constant and magnetic permeability. The incident light may have TE(s) or TM(p) polarization.

In Sec. II we use the Rayleigh hypothesis—our only approximation, in principle—in order to derive the "central equation," Eq. (21). This is an integral matrix equation that relates the reflected fields, the transmitted fields, and the fields inside the thin film to the incident wave. The general formalism is applied to the special case of periodic corrugations in Sec. III. A numerical method of solution for aperiodic roughness profiles is given in Sec. IV. In Secs. III and IV we also derive the secular equations for the eigenmodes of the periodic and aperiodic films, respectively.

In a future publication, we shall present a perturbative (in the corrugation height), analytic solution for the optical response of rough films.[32] Preliminary numerical results compare favorably with experimental and calculated reflectivity curves in Ref. 22.

II. BASIC FORMALISM

Let us consider a thin film of thickness D and dielectric constant ϵ_2 sandwiched between a substrate of dielectric constant ϵ_3 and a superstrate of dielectric constant ϵ_1. The corresponding magnetic permeabilities are μ_1, μ_2, and μ_3. The two interfaces have arbitrary, deterministic roughness in one dimension (x); the z axis is chosen to be perpendicular to the thin film, and therefore, the y axis is parallel to the grooves of the roughness. The upper and lower interfaces are defined by the profile functions $f(x)$ and $g(x)$, namely,

$$z = D + hf(x) , \tag{1}$$

and

$$z = hg(x) , \tag{2}$$

where h is a measure of the height of the surface roughness. Clearly, the "teeth" of the two interfaces should not run into each other, the condition for which is

$$D + hf(x) > hg(x) .$$

It may be convenient to choose the direction of the x axis, the point $z=0$, and the thickness D in such a way that $f(x)$ and $g(x)$, integrated from $-\infty$ to ∞, give vanishing results. Other choices, however, are possible. All three media comprising the thin film are taken to be isotropic, homogeneous, linear, and local. (See Fig. 1.)

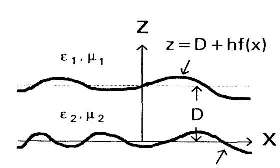

FIG. 1. Geometry considered in this work. $f(x)$ and $g(x)$ are arbitrary deterministic functions.

We assume that a monochromatic plane wave is incident from medium 1 at the interface given by Eq. (1). The plane of incidence is chosen to be the xz plane, and the angle of incidence is θ. The wave vector of the incident wave may be written as

$$\hat{x}Q_0+\hat{z}\beta_0=(\omega/c)(\epsilon_1\mu_1)^{1/2}(\hat{x}\sin\theta+\hat{z}\cos\theta) ,$$

where ω is the circular frequency. We treat simultaneously the cases of TE or s polarization and TM or p polarization of the incident wave. Its amplitude I represents the electric field E_y in the case of s polarization and it is the magnetic field H_y for p polarization. That is, I is the amplitude of the y component of the electromagnetic field for both polarizations. Then the incident wave (y component) is

$$U_i(x,z)=Ie^{i(Q_0x+\beta_0z)}, \quad z\geq D+hf(x) . \tag{3}$$

Because of the surface roughness, the parallel component Q_0 of the wave vector is not conserved. Then the reflected and transmitted fields and the field inside the thin film all suffer diffraction. They are a superposition of plane waves with all values, in general, of the wave-vector component Q parallel to the surface. Correspondingly, the perpendicular components of the wave vectors in the three media are

$$\beta_j=\left[\epsilon_j\mu_j\frac{\omega^2}{c^2}-Q^2\right]^{1/2}, \quad j=1,2,3 .$$

The y component of the reflected field, the field in the thin film, and the transmitted field are postulated to be given, respectively, by the following expressions for their Fourier integrals:

$$U_r(x,z)=\int_{-\infty}^{\infty}dQ\,R(Q)e^{i(Qx-\beta_1z)}, \quad z\geq D+hf(x) , \tag{4}$$

$$U_f(x,z)=\int_{-\infty}^{\infty}dQ\,[A(Q)e^{i\beta_2z}+B(Q)e^{-i\beta_2z}]e^{iQx} ,$$
$$hg(x)\leq z\leq D+hf(x) , \tag{5}$$

$$U_t(x,z)=\int_{-\infty}^{\infty}dQ\,T(Q)e^{i(Qx+\beta_3z)}, \quad z\leq hg(x) . \tag{6}$$

Equations (4)–(6) are exact solutions of the Helmholtz equation only outside the selvedge region. We assume, however, that they may be "continued" into the selvedge region, presumably with an acceptably small error. This, indeed, is the essence of the Rayleigh hypothesis. The quantities that are assumed to be given are the ϵ_j, μ_j, D, h, $f(x)$, $g(x)$, ω, θ, and I. Our task is to determine the field amplitudes $R(Q)$, $A(Q)$, $B(Q)$, and $T(Q)$.

At every point of the interfaces 1/2 and 2/3 the parallel components of **E** and **H** must be continuous. Because $U(x,z)$ is chosen to be the electric (magnetic) field for transverse-electric (magnetic) polarization, then, clearly, $U(x,z)$ is continuous for both polarizations. There is a simple device for handling the continuity of the other parallel component. Namely, one uses the fact that $(1/\nu_j)(\partial U/\partial n)$ is also continuous; here, $\partial U/\partial n$ is the normal derivative and $\nu_j=1$ for s polarization and $\nu_j=\epsilon_j$ for p polarization. Then, the boundary conditions at the two interfaces read

$$(U_i+U_r)|_{z=hf(x)+D}=U_f|_{z=hf(x)+D} , \tag{7}$$

$$\frac{1}{\nu_1}\frac{\partial}{\partial n}(U_i+U_r)|_{z=hf(x)+D}=\frac{1}{\nu_2}\frac{\partial}{\partial n}U_f|_{z=hf(x)+D} , \tag{8}$$

$$U_f|_{z=hg(x)}=U_t|_{z=hg(x)} , \tag{9}$$

$$\frac{1}{\nu_2}\frac{\partial}{\partial n}U_f|_{z=hg(x)}=\frac{1}{\nu_3}\frac{\partial}{\partial n}U_t|_{z=hg(x)} , \tag{10}$$

where

$$\frac{\partial}{\partial n}\psi\equiv\hat{\mathbf{n}}\cdot\nabla\psi , \tag{11}$$

and $\hat{\mathbf{n}}$ is a unit vector perpendicular to the interfaces at an arbitrary point. Thus,

$$\frac{\partial}{\partial n}e^{i\{Qx\pm\beta_jz[f(x)]\}}=F_j^{\pm}e^{i\{Qx\pm\beta_jz[f(x)]\}} , \tag{12}$$

where

$$F_j^{\pm}=\frac{i[-Q\dot{f}(x)h\pm\beta_j]}{\{1+[\dot{f}(x)h]^2\}^{1/2}} , \tag{13}$$

and similarly,

$$\frac{\partial}{\partial n}e^{i\{Qx\pm\beta_jz[g(x)]\}}=G_j^{\pm}e^{i\{Qx\pm\beta_jz[g(x)]\}} , \tag{14}$$

where

$$G_j^{\pm}=\frac{i[-Q\dot{g}(x)h\pm\beta_j]}{\{1+[\dot{g}(x)h]^2\}^{1/2}} . \tag{15}$$

It is understood that, for $j=0$, $Q=Q_0$, and the overdots on $f(x)$ and $g(x)$ denote derivatives.

Using Eqs. (3)–(6), (12), and (14), Eqs. (7)–(10) can be written as

$$Ie^{iQ_0x+i\beta_0f(x)h+i\beta_0D}+\int_{-\infty}^{\infty}dQ\,R(Q)e^{iQx-i\beta_1f(x)h-i\beta_1D}=\int_{-\infty}^{\infty}dQ\,[A(Q)e^{i\beta_2f(x)h+i\beta_2D}+B(Q)e^{-i\beta_2f(x)h-i\beta_2D}]e^{iQx} , \tag{16}$$

$$\frac{1}{\nu_1}IF_0^{+}e^{iQ_0x+i\beta_0f(x)h+i\beta_0D}+\frac{1}{\nu_1}\int_{-\infty}^{\infty}dQ\,R(Q)F_1^{-}e^{iQx-i\beta_1f(x)h-i\beta_1D}$$
$$=\frac{1}{\nu_2}\int_{-\infty}^{\infty}dQ\{A(Q)F_2^{+}e^{i\beta_2f(x)h+i\beta_2D}+B(Q)F_2^{-}e^{-i\beta_2f(x)h-i\beta_2D}\}e^{iQx} , \tag{17}$$

$$\int_{-\infty}^{\infty} dQ\,[A(Q)e^{i\beta_2 g(x)h}+B(Q)e^{-i\beta_2 g(x)h}]e^{iQx}=\int_{-\infty}^{\infty} dQ\,T(Q)e^{iQx+i\beta_3 g(x)h}\,, \tag{18}$$

$$\frac{1}{\nu_2}\int_{-\infty}^{\infty} dQ\,[A(Q)G_2^+e^{i\beta_2 g(x)h}+B(Q)G_2^-e^{-i\beta_2 g(x)h}]e^{iQx}=\frac{1}{\nu_3}\int_{-\infty}^{\infty} dQ\,T(Q)G_3^+e^{iQx+i\beta_3 g(x)h}\,. \tag{19}$$

Now we can write the continuity conditions [(16)–(19)] in a matrix form

$$\int_{-\infty}^{\infty} dQ \begin{bmatrix} e^{-i\beta_1(fh+D)} & -e^{i\beta_2(fh+D)} & -e^{-i\beta_2(fh+D)} & 0 \\ \frac{1}{\nu_1}F_1^-e^{-i\beta_1(fh+D)} & -\frac{1}{\nu_2}F_2^+e^{i\beta_2(fh+D)} & -\frac{1}{\nu_2}F_2^-e^{-i\beta_2(fh+D)} & 0 \\ 0 & e^{i\beta_2 gh} & e^{-i\beta_2 gh} & -e^{i\beta_3 gh} \\ 0 & \frac{1}{\nu_2}G_2^+e^{i\beta_2 gh} & \frac{1}{\nu_2}G_2^-e^{-i\beta_2 gh} & -\frac{1}{\nu_3}G_3^+e^{i\beta_3 gh} \end{bmatrix}$$

$$\times \begin{bmatrix} R(Q) \\ A(Q) \\ B(Q) \\ T(Q) \end{bmatrix} e^{iQx} = \begin{bmatrix} -Ie^{i\beta_0(fh+D)} \\ -\frac{1}{\nu_1}IF_0^+e^{i\beta_0(fh+D)} \\ 0 \\ 0 \end{bmatrix} e^{iQ_0x}\,, \tag{20}$$

where $f\equiv f(x)$ and $g\equiv g(x)$. If we multiply the second row by $-i$ times the denominator of $F_j^{\pm}$, and the fourth row by $-i$ times the denominator of $G_j^{\pm}$, then we can write the above equation in a simple form:

$$\int_{-\infty}^{\infty} dQ\,\mathcal{M}(Q,x)\mathcal{R}(Q)e^{iQx}=\mathcal{I}(Q_0,x)e^{iQ_0x}\,, \tag{21}$$

where

$$\mathcal{M}(Q,x)=\begin{bmatrix} e^{-i\beta_1(fh+D)} & -e^{i\beta_2(fh+D)} & -e^{-i\beta_2(fh+D)} & 0 \\ \frac{-Q\dot{f}h-\beta_1}{\nu_1}e^{-i\beta_1(fh+D)} & \frac{Q\dot{f}h-\beta_2}{\nu_2}e^{i\beta_2(fh+D)} & \frac{Q\dot{f}h+\beta_2}{\nu_2}e^{-i\beta_2(fh+D)} & 0 \\ 0 & e^{i\beta_2 gh} & e^{-i\beta_2 gh} & -e^{i\beta_3 gh} \\ 0 & \frac{-Q\dot{g}h+\beta_2}{\nu_2}e^{i\beta_2 gh} & \frac{-Q\dot{g}h-\beta_2}{\nu_2}e^{-i\beta_2 gh} & \frac{Q\dot{g}h-\beta_3}{\nu_3}e^{i\beta_3 gh} \end{bmatrix}, \tag{22}$$

$$\mathcal{R}(Q)=\begin{bmatrix} R(Q) \\ A(Q) \\ B(Q) \\ T(Q) \end{bmatrix}, \tag{23}$$

and

$$\mathcal{I}(Q_0,x)=\begin{bmatrix} -1 \\ \frac{Q_0\dot{f}h-\beta_0}{\nu_1} \\ 0 \\ 0 \end{bmatrix} Ie^{i\beta_0(fh+D)}\,. \tag{24}$$

The last four equations summarize our results. In principle, it is necessary to invert the integral equation (21) and express $\mathcal{R}(Q)$ in terms of the known matrices $\mathcal{M}(Q,x)$, Eq. (22), and $\mathcal{I}(Q_0,x)$, Eq. (24). The vector $\mathcal{R}(Q)$ gives the Fourier transforms of the four relevant field amplitudes [see Eq. (23)] and, through Eqs. (4)–(6), determines the complete optical response of our rough film. These calculations are valid for light of either TE or TM polarization and for arbitrary roughness profiles. In fact, the Rayleigh hypothesis is the only limitation on the generality of Eqs. (21)–(24). We shall refer to Eq. (21) as the central equation.

III. PERIODIC ROUGHNESS

In this section we assume that the profile functions $f(x)$ and $g(x)$ are both periodic with the same period d. These functions are not necessarily the same, although this is an important special case, corresponding to numerous experimental configurations. Another particular geometry that conserves the overall periodicity is the case where one of the interfaces is plane; then either $f(x)$ or $g(x)$ vanishes identically.

For a periodic system, the Bloch theorem must be satisfied, that is, apart from the wave factor $\exp(iQ_0x)$, the electromagnetic fields have the same periodicity as the corrugations of the thin film (for every value of the coordinate z). This periodic part may be expanded in a Fourier series. Then, the transverse (y) component of the reflected field is written as

$$U_r(x,z)=\sum_{-\infty}^{\infty} R_n e^{i[Q_n x-\beta_1(Q_n)z]}, \tag{25}$$

$$Q_n=Q_0+2\pi n/d, \quad n=\ldots,-1,0,1,\ldots, \tag{26}$$

$$\beta_j(Q_n)=\left[\epsilon_j\mu_j\frac{\omega^2}{c^2}-Q_n^2\right]^{1/2}, \quad j=1,2,3 . \tag{27}$$

On comparing Eqs. (25) and (4), we see that $R(Q)$ must be given by the expression

$$R(Q)=\sum_{-\infty}^{\infty} R_n(Q_n)\delta(Q-Q_n) . \tag{28}$$

Analogous formulas hold for the fields inside the film and for the transmitted fields. In our matrix formalism we write

$$\mathcal{R}(Q)=\sum_{-\infty}^{\infty} \mathcal{R}_n(Q_n)\delta(Q-Q_n) . \tag{29}$$

The column matrix $\mathcal{R}(Q)$ has been defined in Eq. (23) and

$$\mathcal{R}_n(Q_n)=\begin{bmatrix} R_n(Q_n) \\ A_n(Q_n) \\ B_n(Q_n) \\ T_n(Q_n) \end{bmatrix} . \tag{30}$$

Equation (29) indicates that now only discrete values of the wave vector Q are permitted; these Q_n are given by Eq. (26). The thin film reflects light only at angles θ_n such that

$$\sin\theta_n=\frac{Q_n}{(\omega/c)(\epsilon_1\mu_1)^{1/2}}=\sin\theta+\frac{2\pi c}{\omega(\epsilon_1\mu_1)^{1/2}d}n . \tag{31}$$

For $n=0$ specular reflection is obtained. For a given value of $n\,(\neq 0)$ the reflection corresponds to order n of diffraction.

Let us substitute Eq. (29) in our central equation, Eq. (21). We get

$$\sum_{-\infty}^{\infty} \mathcal{M}(Q_n,x)\mathcal{R}_n(Q_n)e^{iQ_n x}=\mathcal{I}(Q_0,x)e^{iQ_0 x} . \tag{32}$$

Here we are summing an infinite number of column matrices. If we multiply every one of the four rows of Eq. (32) by $\exp(-iQ_m x)$ and integrate over the period d, the result is

$$\sum_{n=-\infty}^{\infty} \hat{\mathcal{M}}(Q_n,m-n)\mathcal{R}_n(Q_n)=\hat{\mathcal{I}}(Q_0,m) , \quad m=\ldots,-1,0,1,\ldots . \tag{33}$$

Every element of the matrices $\hat{\mathcal{M}}$ and $\hat{\mathcal{I}}$ is the Fourier coefficient of the corresponding matrix element of $\mathcal{M}$ and $\mathcal{I}$:

$$\hat{\mathcal{M}}_{ij}(Q_n,m)=\frac{1}{d}\int_0^d dx\, \mathcal{M}_{ij}(Q_n,x)e^{-i2\pi mx/d} , \quad i,j=1,2,3,4 , \tag{34}$$

$$\hat{\mathcal{I}}_i(Q_0,m)=\frac{1}{d}\int_0^d dx\, \mathcal{I}_i(Q_0,x)e^{-i2\pi mx/d} , \quad i=1,2,3,4 . \tag{35}$$

A glance at Eqs. (22) and (24) reveals that the right-hand sides of Eqs. (34) and (35) involve only the integrals

$$J_f(\pm\beta_i,m)\equiv\frac{1}{d}\int_0^d dx\, e^{\pm i\beta_i hf(x)-i2\pi mx/d} , \tag{36}$$

$$\tilde{J}_f(\pm\beta_i,m)\equiv\frac{1}{d}\int_0^d dx\, \dot{f}(x)e^{\pm i\beta_i hf(x)-i2\pi mx/d} , \tag{37}$$

and similar integrals with $g(x)$ replacing $f(x)$. Using Eqs. (22), (24), and (34)–(37), we get the formulas in the Appendix. If $f(x)$ is composed entirely of linear segments, then the integration in Eq. (36) is immediate.

If m is allowed to take all values, then the complete set of Eqs. (33) may be written as

$$\begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \hat{\mathcal{M}}(Q_{-1},0) & \hat{\mathcal{M}}(Q_0,-1) & \hat{\mathcal{M}}(Q_1,-2) & \cdots \\ \cdots & \hat{\mathcal{M}}(Q_{-1},1) & \hat{\mathcal{M}}(Q_0,0) & \hat{\mathcal{M}}(Q_1,-1) & \cdots \\ \cdots & \hat{\mathcal{M}}(Q_{-1},2) & \hat{\mathcal{M}}(Q_0,1) & \hat{\mathcal{M}}(Q_1,0) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \cdots \\ \mathcal{R}_{-1} \\ \mathcal{R}_0 \\ \mathcal{R}_1 \\ \cdots \end{bmatrix} = \begin{bmatrix} \cdots \\ \hat{\mathcal{I}}(Q_0,-1) \\ \hat{\mathcal{I}}(Q_0,0) \\ \hat{\mathcal{I}}(Q_0,1) \\ \cdots \end{bmatrix} . \tag{38}$$

The three infinite matrices in this equation are partitioned matrices, with the elements defined by the submatrices (A1), (A2), and (30). We may also rewrite Eq. (33) in the explicit form

$$\sum_{n=-\infty}^{\infty}\sum_{j=1}^{4} \hat{\mathcal{M}}_{ij}(Q_n,m-n)\mathcal{R}_{nj}(Q_n)=\hat{\mathcal{I}}_i(Q_0,m) , \quad m=\ldots,-1,0,1,\ldots , \quad i=1,2,3,4 . \tag{39}$$

By Eq. (30) $\mathcal{R}_{n1}=R_n$, $\mathcal{R}_{n2}=A_n$, $\mathcal{R}_{n3}=B_n$, and $\mathcal{R}_{n4}=T_n$.

Of course, in practice the number of equations (39) must be finite, as must be the number of terms in the summation over n. If we choose $|m|,|n|\le N$, where N is a sufficiently large positive integer, then there are $4(2N+1)$ equations for the $4(2N+1)$ unknowns $\mathcal{R}_{nj}$:

$$R_{-N},A_{-N},B_{-N},T_{-N},\ldots,\quad R_0,A_0,B_0,T_0,\ldots,$$
$$R_N,A_N,B_N,T_N\ .$$

Thus, we see that the solution of Eq. (39) yields the complete optical response of the corrugated film: not just the reflected-field amplitudes, but also the amplitudes of the fields inside the film and those of the transmitted field. Note that R_n is the amplitude of the transverse field reflected (or diffracted) at the angle θ_n given by Eq. (31).

In the special case $\epsilon_1=\epsilon_3$, $\mu_i=1$ (nonmagnetic media), $f(x)=g(x)$, and $\nu_i=\epsilon_i$ (p-polarized light), our results should be equivalent to those of Ref. 22. Also, if $\mu_i=1$, $f(x)=0$, and $\nu_i=\epsilon_i$ then our formulas should reduce to those in Ref. 26. Unfortunately, in both cases the proofs do not seem to be straightforward.

The condition for self-sustained electromagnetic oscillations is that the amplitude of the incident field vanishes. By Eq. (A2), if $I=0$ then the column matrix $\hat{\mathcal{I}}$ also vanishes. On the other hand, the reflectivity amplitudes $\mathcal{R}_{nj}$ should not be all equal to zero, that is, the response matrix $\mathcal{R}$ must not be a null matrix. Then Eq. (38) gives the secular equation for the eigenmodes of the thin film,

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \hat{\mathcal{M}}(Q_{-1},0) & \hat{\mathcal{M}}(Q_0,-1) & \hat{\mathcal{M}}(Q_1,-2) & \cdots \\ \cdots & \hat{\mathcal{M}}(Q_{-1},1) & \hat{\mathcal{M}}(Q_0,0) & \hat{\mathcal{M}}(Q_1,-1) & \cdots \\ \cdots & \hat{\mathcal{M}}(Q_{-1},2) & \hat{\mathcal{M}}(Q_0,1) & \hat{\mathcal{M}}(Q_1,0) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}=0\ . \tag{40}$$

An equivalent result has been derived previously.[33] There, however, the elements of the determinant are 2×2 submatrices, rather than the 4×4 submatrices in Eq. (40). Also, for $g(x)=0$ and $\mu_i=1$, Eq. (40) should reduce to Eq. (13) of Ref. 21, and to Eq. (2.11) of Ref. 25 if, in addition, $\nu_i=\epsilon_i$.

It is instructive to consider the case of a smooth film, $f(x)=g(x)=0$. Then Eq. (36) gives $J_f(\pm\beta_i,m)=\delta_{m,0}$ and $\tilde{J}_f(\pm\beta_i,m)=0$. We see that the determinant of Eq. (40) becomes diagonal in the submatrices $\hat{\mathcal{M}}(Q_n,0)$. The determinant of such a matrix is simply equal to the product of the determinants of the submatrices, that is, Eq. (40) becomes

$$\cdots\times|\hat{\mathcal{M}}(Q_{-1},0)|\times|\hat{\mathcal{M}}(Q_0,0)|\times|\hat{\mathcal{M}}(Q_1,0)|\times\cdots=0\ ,$$

or

$$|\hat{\mathcal{M}}(Q_n,0)|=0,\quad n=\ldots,-1,0,1,\ldots\ .$$

If we substitute $m=0$ in Eq. (A1), this matrix greatly simplifies. Its determinant is readily found to be

$$|\hat{\mathcal{M}}(Q_n,0)|=\left[\frac{\beta_1}{\nu_1}+\frac{\beta_2}{\nu_2}\right]\left[\frac{\beta_2}{\nu_2}+\frac{\beta_3}{\nu_3}\right]e^{i(\beta_2-\beta_1)D}+\left[\frac{\beta_1}{\nu_1}-\frac{\beta_2}{\nu_2}\right]\left[\frac{\beta_2}{\nu_2}-\frac{\beta_3}{\nu_3}\right]e^{-i(\beta_2+\beta_1)D}\ .$$

Setting this equal to zero, we have

$$\left[\frac{\beta_1(Q_n)}{\nu_1}+\frac{\beta_2(Q_n)}{\nu_2}\right]\left[\frac{\beta_2(Q_n)}{\nu_2}+\frac{\beta_3(Q_n)}{\nu_3}\right]e^{i\beta_2(Q_n)D}+\left[\frac{\beta_1(Q_n)}{\nu_1}-\frac{\beta_2(Q_n)}{\nu_2}\right]\left[\frac{\beta_2(Q_n)}{\nu_2}-\frac{\beta_3(Q_n)}{\nu_3}\right]e^{-i\beta_2(Q_n)D}=0\ ,$$
$$n=\ldots,-1,0,1,\ldots\ . \tag{41}$$

For $n=0$, this is just the well-known dispersion relation of polaritons in a smooth (planar) thin film, for either TE ($\nu_i=1$) or TM ($\nu_i=\epsilon_i$) polarization. If $n\neq0$, then the dispersion curves for $n=0$ are displaced along the Q_0 axis by all the reciprocal-lattice vectors $(2\pi/d)n$.[17] This corresponds to the "empty-lattice approximation" in the band structure of periodic solids. The effect of the periodicity is kept, although the limit $h\to0$ for the height of the corrugation is taken.

Intersection of polariton dispersion branches may occur outside of Brillouin-zone boundaries as well as at these boundaries (see Fig. 1 of Ref. 17). If $f(x)$ and/or $g(x)$ is permitted to be finite, then Eq. (40) will lead to mode repulsion or to gaps in the $\omega(Q_0)$ spectrum.

IV. APERIODIC ROUGHNESS

Returning to our central equation, we multiply both sides of Eq. (21) by $\exp(-iQ'x)$ and integrate over the x axis. We get

$$\int_{-\infty}^{\infty}dQ\,\hat{\mathcal{M}}(Q,Q'-Q)\mathcal{R}(Q)=\hat{\mathcal{I}}(Q_0,Q'-Q_0)\ , \tag{42}$$

where $\hat{\mathcal{M}}$ and $\hat{\mathcal{I}}$ are now the inverse Fourier integrals of $\mathcal{M}$ and $\mathcal{I}$:

$$\hat{\mathcal{M}}(Q,Q')=\int_{-\infty}^{\infty}dx\,\mathcal{M}(Q,x)e^{-iQ'x}\ , \tag{43}$$

$$\hat{\mathcal{I}}(Q_0,Q')=\int_{-\infty}^{\infty}dx\,\mathcal{I}(Q_0,x)e^{-iQ'x}\ . \tag{44}$$

In analogy to Eqs. (36) and (37), the independent integrals on the right-hand sides of Eqs. (43) and (44) are

$$J_f(\pm\beta_i,Q')=\int_{-\infty}^{\infty}dx\, e^{\pm i\beta_i h f(x)-iQ'x}\ , \tag{45}$$

$$\tilde{J}_f(\pm\beta_i,Q')=\int_{-\infty}^{\infty}dx\, \dot{f}(x)e^{\pm i\beta_i h f(x)-iQ'x}\ , \tag{46}$$

and a similar one for $g(x)$. The matrices $\hat{\mathcal{M}}(Q,Q')$ and $\hat{\mathcal{I}}(Q_0,Q')$ are quite similar to Eqs. (A1) and (A2). It is not difficult to see that they are obtained from these equations by the replacements

$$\begin{aligned} &Q_n\to Q,\quad 2\pi m/d\to Q'\ ,\\ &J_{f,g}(\pm\beta_i,m)\to J_{f,g}(\pm\beta_i,Q')\ ,\\ &\tilde{J}_{f,g}(\pm\beta_i,m)\to\tilde{J}_{f,g}(\pm\beta_i,Q')\ . \end{aligned} \tag{47}$$

As in the case of periodic roughness, the integral (45) may be solved analytically if $f(x)$ is constituted from linear segments. Assuming that $\hat{\mathcal{M}}(Q,Q')$ and $\hat{\mathcal{I}}(Q,Q')$ have been calculated—analytically or numerically—we proceed to solve Eq. (42).

We divide the Q axis into a set of convenient intervals $\Delta Q_{-N},\ldots,\Delta Q_0,\ldots,\Delta Q_N$, which are not necessarily equal. [The choice of the ΔQ_n in practice would depend on the profiles $f(x)$ and $g(x)$.] The wave vectors Q and Q' in Eq. (42) are allowed to assume only discrete values $Q_{-N},\ldots,Q_0,\ldots,Q_N$, where Q_n is chosen to be the midpoint of the interval ΔQ_n. Then we rewrite Eq. (42) for $Q'=Q_{-N},\ldots,Q_0,\ldots,Q_N$, and replace the integration by summation over the intervals ΔQ_n:

$$\sum_{n=-N}^{N}\Delta Q_n\hat{\mathcal{M}}(Q_n,Q_m-Q_n)\mathcal{R}(Q_n)=\hat{\mathcal{I}}(Q_0,Q_m-Q_0)\ ,\qquad m=-N,\ldots,0,\ldots,N\ . \tag{48}$$

These matrix equations are similar to Eqs. (33) for the periodic film. We also write down the explicit form [see Eq. (39)]

$$\sum_{n=-N}^{N}\Delta Q_n\sum_{j=1}^{4}\hat{\mathcal{M}}_{ij}(Q_n,Q_m-Q_n)\mathcal{R}_j(Q_n)=\hat{\mathcal{I}}_i(Q_0,Q_m-Q_0)\ ,\qquad m=-N,\ldots,0,\ldots,N;\ \ i=1,2,3,4\ . \tag{49}$$

Here we have $4(2N+1)$ equations to calculate the $4(2N+1)$ amplitudes $R_j(Q_n)$. We remind the reader that $\mathcal{R}_1(Q_n)\equiv R(Q_n)$ is the amplitude of the y component of the electromagnetic field that is reflected (diffracted) at an angle θ_n given by the first equality of Eq. (31). Similarly, $\mathcal{R}_4(Q_n)\equiv T(Q_n)$ is the corresponding transmitted field. If all the intervals ΔQ_n are chosen to be equal to some ΔQ, then

$$Q_n=Q_0+n\Delta Q,\quad n=-N,\ldots,0,\ldots,N\ . \tag{50}$$

This is similar to Eq. (26); however, $2\pi/d$ is replaced by ΔQ. Of course, Eq. (26) gives all the *possible* directions of diffraction (for a periodic film), while Eq. (50) specifies certain *chosen* directions of diffraction (for the aperiodic film). Now Eq. (49) simplifies to read

$$\sum_{n=-N}^{N}\sum_{j=1}^{4}\hat{\mathcal{M}}_{ij}[Q_n,(m-n)\Delta Q]\mathcal{R}_j(Q_n)=\hat{\mathcal{I}}_i(Q_0,m\Delta Q)/\Delta Q\ ,\qquad m=-N,\ldots,0,\ldots,N;\ \ i=1,2,3,4\ . \tag{51}$$

The matrix elements $\hat{\mathcal{M}}_{ij}(Q_n,m\Delta Q)$ and $\hat{\mathcal{I}}_i(Q_0,m\Delta Q)$ are obtained from Eqs. (A1) and (A2) by replacing $J_{f,g}(\pm\beta_i,m)$ and $\tilde{J}_{f,g}(\pm\beta_i,m)$ with $J_{f,g}(\pm\beta_i,m\Delta Q)$ and $\tilde{J}_{f,g}(\pm\beta_i,m\Delta Q)$, the latter quantities being defined by Eqs. (45) and (46). We see that by choosing equal intervals ΔQ, the solution equation (51) becomes very similar to Eq. (39) for a periodic film. The major difference lies in the replacement of the Fourier coefficients of $\exp[\pm i\beta_i hf(x)]$, Eqs. (36) and (37), by the inverse Fourier integral, Eqs. (45) and (46), of the same function.

In practice, the light is diffracted in all directions in a continuous fashion. The method described here serves as an algorithm for numerical calculations of the optical response of thin films with aperiodic roughness and it becomes exact in the limit $\Delta Q_n\to 0$ and $N\to\infty$.

From Eq. (48) the secular equation for the electromagnetic modes of the system is

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \Delta Q_{-1}\hat{\mathcal{M}}(Q_{-1},0) & \Delta Q_0\hat{\mathcal{M}}(Q_0,Q_{-1}-Q_0) & \Delta Q_1\hat{\mathcal{M}}(Q_1,Q_{-1}-Q_1) & \cdots\\ \cdots & \Delta Q_{-1}\hat{\mathcal{M}}(Q_{-1},Q_0-Q_{-1}) & \Delta Q_0\hat{\mathcal{M}}(Q_0,0) & \Delta Q_1\hat{\mathcal{M}}(Q_1,Q_0-Q_1) & \cdots\\ \cdots & \Delta Q_{-1}\hat{\mathcal{M}}(Q_{-1},Q_1-Q_{-1}) & \Delta Q_0\hat{\mathcal{M}}(Q_0,Q_1-Q_0) & \Delta Q_1\hat{\mathcal{M}}(Q_1,0) & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}=0\ . \tag{52}$$

If one chooses $\Delta Q_n=\Delta Q$, then Eq. (52) assumes the same form as Eq. (40), except that the elements $\hat{\mathcal{M}}(Q_n,m)$ are replaced by $\hat{\mathcal{M}}(Q_n,m\Delta Q)$.

ACKNOWLEDGMENTS

We acknowledge useful comments by Raul García-Llamas. S.W. appreciates support by CONACyT and by DGICSA. P.H. thanks CONACyT for support under Grant No. 0746-E9109 and for support from the Universidad de Sonora.

APPENDIX

The matrices $\hat{\mathcal{M}}$ and $\hat{\mathcal{I}}$ which appear in Eq. (33) are

$$\hat{\mathcal{M}}(Q_n,m)=\begin{bmatrix} e^{-i\beta_1 D}J_f(-\beta_1,m) & -e^{i\beta_2 D}J_f(+\beta_2,m) & -e^{-i\beta_2 D}J_f(-\beta_2,m) & 0 \\ -e^{-i\beta_1 D}K_{f1}(-\beta_1,m) & e^{i\beta_2 D}K_{f2}(+\beta_2,m) & e^{-i\beta_2 D}K_{f2}(-\beta_2,m) & 0 \\ 0 & J_g(+\beta_2,m) & J_g(-\beta_2,m) & -J_g(+\beta_3,m) \\ 0 & -K_{g2}(+\beta_2,m) & -K_{g2}(-\beta_2,m) & K_{g3}(+\beta_3,m) \end{bmatrix}, \tag{A1}$$

$$\hat{\mathcal{I}}(Q_0,m)=\begin{bmatrix} -1 \\ \dfrac{Q_0 h}{\nu_1}\tilde{J}_f(\beta_0,m)-\dfrac{\beta_0}{\nu_1}J_f(\beta_0,m) \\ 0 \\ 0 \end{bmatrix} I e^{i\beta_0 D}, \tag{A2}$$

where $\beta_i=\beta_i(Q_n)$,

$$K_{fj}(\pm\beta_j,m)=\frac{Q_n h}{\nu_j}\tilde{J}_f(\pm\beta_j,m)-\frac{\pm\beta_j}{\nu_j}J_f(\pm\beta_j,m)\ ,\quad K_{gj}(\pm\beta_j,m)=\frac{Q_n h}{\nu_j}\tilde{J}_g(\pm\beta_j,m)-\frac{\pm\beta_j}{\nu_j}J_g(\pm\beta_j,m)\ .$$

---

*Present address: Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 2681, Ensenada, Baja California, 22800 México.

†Permanent address: Instituto de Física, Universidad Autónoma de Puebla, Apartado Postal J-48, Puebla, 72570 México.

[1]F. Dabby, A. Kesteubaum, and U. Pack, Opt. Commun. **6**, 125 (1972).

[2]K. Sakuda and A. Yariv, Opt. Commun. **8**, 1 (1973).

[3]H. Stoll and A. Yariv, Opt. Commun. **8**, 5 (1973).

[4]M. Neviere, P. Vincent, and R. Petit, Nouv. Rev. Opt. **5**, 65 (1974).

[5]M. Neviere, D. Maystre, and P. Vincent, J. Opt. (Paris) **8**, 231 (1977).

[6]P. Vincent and M. Neviere, Appl. Phys. **20**, 345 (1979).

[7]D. Maystre, in *Electromagnetic Theory of Gratings,* edited by R. Petit, Topics in Current Physics Vol. 22 (Springer-Verlag, Berlin, 1980), p. 63.

[8]D. Maystre, in *Electromagnetic Surface Modes,* edited by A. D. Boardman (Wiley, New York, 1982), p. 661.

[9]P. K. Tien, Rev. Mod. Phys. **49**, 361 (1977).

[10]*Integrated Optics,* edited by T. Tamir, Topics in Applied Physics Vol. 7 (Springer-Verlag, Berlin, 1979).

[11]I. Pockrand, Phys. Lett. **49A**, 259 (1974).

[12]I. Pockrand, Opt. Commun. **13**, 311 (1975).

[13]I. Pockrand and H. Raether, Opt. Commun. **18**, 395 (1976).

[14]R. W. Gruhlke, W. R. Holland, and D. G. Hall, Phys. Rev. Lett. **56**, 2838 (1986).

[15]M. G. Weber and D. L. Mills, Phys. Rev. B **32**, 5057 (1985).

[16]S. R. J. Brueck, V. Diadiuk, T. Jones, and W. Lenth, Appl. Phys. Lett. **46**, 915 (1985).

[17]P. Halevi and O. Mata-Méndez, Phys. Rev. B **39**, 5694 (1989).

[18]T. Inagaki, M. Motosuga, E. T. Arakawa, and J. P. Goudonnet, Phys. Rev. B **31**, 2548 (1985).

[19]T. Inagaki, M. Motosuga, E. T. Arakawa, and J. P. Goudonnet, Phys. Rev. B **32**, 6238 (1985).

[20]V. L. Brudny and R. A. Depine, Opt. Commun. **82**, 420 (1991).

[21]O. Mata-Méndez and P. Halevi, Phys. Rev. B **36**, 1007 (1987).

[22]Z. Chen and H. J. Simon, J. Opt. Soc. Am. B **5**, 1396 (1988).

[23]H. J. Simon and Z. Chen, Phys. Rev. B **39**, 3077 (1989).

[24]G. A. Farias, A. A. Maradudin, and V. Celli, Surf. Sci. **129**, 9 (1983).

[25]M. M. Auto, G. A. Farias, and A. A. Maradudin, Surf. Sci. **167**, 57 (1986).

[26]M. G. Cavalcante, G. A. Farias, and A. A. Maradudin, J. Opt. Soc. Am. B **4**, 1372 (1987).

[27]J. P. Hugonin and R. Petit, Opt. Commun. **22**, 221 (1977).

[28]J. P. Hugonin and R. Petit, J. Opt. Soc. Am. **71**, 664 (1981).

[29]J.-J. Greffet and F.-R. Ladan, J. Opt. Soc. Am. **8**, 1261 (1991).

[30]F. Toigo, A. Marvin, V. Celli, and N. R. Hill, Phys. Rev. B **15**, 5618 (1977).

[31]Shu Wang, R. G. Barrera, and W. L. Mochán, Phys. Rev. B **40**, 1571 (1989).

[32]Shu Wang and P. Halevi, J. Phys. Condens. Matter (to be published).

[33]C. López-Carrillo and P. Halevi (unpublished).