PHYSICAL REVIEW B VOLUME 23, NUMBER 10 15 MAY 1981

Surface polaritons on large-amplitude gratings

Bernardo Laks,* D. L. Mills, and A. A. Maradudin
Department of Physics, University of California, Irvine, California 92717
(Received 15 September 1980)

We present numerical studies of the dispersion relation for surface polaritons that propagate on the surface of a model free-electron metal upon which a periodic grating has been impressed. The calculations are based on the integral equation formulation of the boundary-value problem, as discussed earlier by Toigo, Marvin, Hill, Celli, and others. We employ the "extinction theorem" scheme, which yields convergent results for values of the ratio ζ_0/a far in excess of 0.072, where the Rayleigh-Fano method fails when applied to the scattering of a scalar wave from a hard wall with profile given by $\zeta_0\cos(2\pi x/a)$. We give results for the dispersion relation for two grating profiles, the sinusoidal and the symmetric sawtooth forms.

I. INTRODUCTION

In recent years, there has been considerable theoretical interest in the description of optical interactions on surfaces that are not perfectly flat. One may consider rough surfaces, where the profile has random character, or the case where a periodic grating of some profile has been impressed on the surface. These questions are directly relevant to a diverse variety of experiments underway at a number of laboratories. For example, the optical constants of materials are often deduced from reflectivity studies, but these are carried out on real surfaces that are surely not smooth. Recent experiments[1] explore the influence of roughness deliberately introduced onto the surface, and characterize this in quantitative terms. A number of intriguing physical phenomena, such as the "giant" Raman signal produced by molecules adsorbed on metal surfaces,[2] and light emission by tunnel junctions,[3] require the presence of roughness[4,5] to be observed. Finally, grating couplers allow one to excite surface polaritons[6] on surfaces which support these modes.

The theoretical studies of the influence of roughness or periodic gratings on optical interactions at surfaces confine their attention to the limit where the deviation from flatness is small. Then perturbation theoretic methods may be employed to describe its effect. Two principal methods have been employed in these studies. One is a Green's-function approach,[7] and the second matches wave fields across the true boundary by means of an approximation scheme valid in the limit of small amplitude perturbations.[8] Both methods yield identical results.[9]

This paper presents a study of the dispersion relation of surface polaritons on a metal surface upon which a perfectly periodic diffraction grating resides. Our purpose is to explore the dispersion relation within the framework of a method that does not assume that the depth ζ_0 of the diffraction-grating grooves is small. If a is the spatial period of the grating, then the method we use here gives accurate, stable, and convergent results for values of the ratio ζ_0/a as large as 0.6. The significance of this remark will be evident from comments below.

We proceed by expressing the boundary-value problem in the form of a set of integral equations. An excellent discussion of properties of the same integral equations was presented earlier by Toigo, Marvin, Celli, and Hill.[10] These authors show how the small-amplitude-perturbation theoretic results cited earlier may be obtained, and their Eqs. (3.11′) and (3.12′), displayed in a form suitable for calculation of the reflectivity rather than the surface-polariton dispersion relation, are identical save for notation to the equations that form the basis of our work. We have also tried one other scheme, described briefly below, without success at the time of this writing. Before we turn to a description of our calculations, some introductory remarks may prove useful.

While rather little attention has been devoted to the theoretical study of optical interactions on corrugated surfaces outside the small-amplitude limit, there is considerable literature on a closely related and conceptually simpler problem. This is the scattering of a scalar wave field $\psi(\tilde{x},t)$ from a corrugated hard wall. This problem has been attacked by a variety of methods. From a physical point of view it may be utilized to describe the scattering of atoms from crystal surfaces. When an attractive van der Waals potential outside the wall is introduced, the model provides a quantitative description of atom-surface scattering in impressive agreement with experimental data.[11,12]

The simplest approach to the finite-amplitude corrugated-wall problem is provided by the Rayleigh-Fano method, discussed for the problem of present interest by Toigo *et al.*[10] In the me-

thod, the scattering problem is approached by considering $\psi(\vec{x},t)$ to consist everywhere of an incoming plane wave, and, in addition, a superposition of scattered waves which for open channels are outgoing waves, and which decay exponentially as one moves away from the surface for closed channels. It is not clear that such a function provides a complete description of the problem, and for this reason the Rayleigh-Fano form of $\psi(\vec{x},t)$ leads to convergent results only for a limited range of parameters. To see this, let the x_3 axis be normal to the average surface, with the x_1 and x_2 axes in a plane parallel to the average surface. Suppose further that the surface has the form of a sinusoidal grating, with the surface profile specified by the relation $x_3=\zeta(x_1)$, and $\zeta(x_1)=\zeta_0\cos(2\pi x_1/a)$. Then the Rayleigh-Fano method clearly provides a proper form for the wave function outside the selvedge region where $x_3>\zeta_0$, but for $-\zeta_0\leq x_3\leq+\zeta_0$, one may expect $\psi(\vec{x},t)$ to consist of scattered waves with both incoming and outgoing character.

When the Rayleigh-Fano form of the wave function is used as the basis for numerical calculations, convergent and well behaved results are obtained only if the ratio $(\zeta_0/a)<0.072$.[12] When the method converges, the uniqueness theorem ensures that one has found the proper solution to the problem. However, for $(\zeta_0/a)>0.072$, in the numerical analysis, one encounters ill-conditioned matrices, and the method fails.

The "extinction-theorem" approach we used here does not require the Rayleigh-Fano ansatz for the electromagnetic fields in the near vicinity of the surface, as Toigo *et al.* have emphasized.[10] Of course, there is no guarantee ill-conditioned matrices will not be encountered,[13] or that the method will converge rapidly enough that very large matrices are not required. As remarked earlier, we have found convergence for values of (ζ_0/a) as large as 0.6, a value of the grating amplitude nearly an order of magnitude larger than where the Rayleigh-Fano method breaks down in the scalar-wave problem. Our criterion for convergence is that the dispersion relation be given to four-figure accuracy; to achieve this, the largest matrix we have encountered is a 52×52 matrix. One may find the determinant of such a matrix and search for zeros, or invert it with a modest investment of computer time.

The discussion in this paper is confined entirely to the study of the surface-polariton-dispersion relation. We have studies of the reflectivity of our model surface underway. When completed, these will be reported elsewhere.

The outline of the paper is as follows. In Sec. II, we present a derivation of the integral equations that form the basis for our analysis, and we cast them into the particular form we shall use. As remarked earlier, these same equations have been discussed by other authors, but the derivation is brief and we present it for completeness. Section III describes our results, and comments briefly on one other approach we have tried to implement, without success.

II. THEORY

In this section we obtain the dispersion relation for a p-polarized surface polariton propagating in the x_1 direction over a grating whose surface is defined by the equation $x_3=\zeta(x_1)$. We assume that the region $x_3<\zeta(x_1)$ is occupied by a dielectric medium, characterized by an isotropic, frequency-dependent dielectric constant $\epsilon(\omega)$; the region $x_3>\zeta(x_1)$ is vacuum.

In this geometry it is particularly convenient to work with the single, nonvanishing component of the magnetic field, which is directed along the x_2 axis, and is a function of the spatial coordinates x_1 and x_3, as well as of the time. If we write this component of the electromagnetic field in the forms

$$H_2(\vec{x},t)=\begin{cases}H_2^{>}(x_1x_3|\omega)\,e^{-i\omega t}, & x_3>\zeta(x_1) \quad (2.1a)\\ H_2^{<}(x_1x_3|\omega)e^{-i\omega t}, & x_3<\zeta(x_1) \quad (2.1b)\end{cases}$$

the amplitude functions $H_2^{\gtrless}(x_1x_3|\omega)$ are the solutions of the equations

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}+\frac{\omega^2}{c^2}\right)H_2^{>}(x_1x_3|\omega)=0, \quad x_3>\zeta(x_1) \quad (2.2a)$$

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}+\epsilon(\omega)\frac{\omega^2}{c^2}\right)H_2^{<}(x_1x_3|\omega)=0, \quad x_3<\zeta(x_1). \quad (2.2b)$$

The boundary conditions satisfied by $H_2^{\gtrless}(x_1x_3|\omega)$ are

$$H_2^{<}(x_1x_3|\omega)\big|_{x_3=\zeta(x_1)}=H_2^{>}(x_1x_3|\omega)\big|_{x_3=\zeta(x_1)}, \quad (2.3a)$$

$$\frac{1}{\epsilon(\omega)}\frac{\partial}{\partial n_+}H_2^{<}(x_1x_3|\omega)\big|_{x_3=\zeta(x_1)}=\frac{\partial}{\partial n_+}H_2^{>}(x_1x_3|\omega)\big|_{x_3=\zeta(x_1)}. \quad (2.3b)$$

In Eq. (2.3b), $(\partial/\partial n_+)$ denotes differentiation along the unit vector normal to the surface at each point, which is assumed to be directed from the dielectric into the vacuum:

$$\frac{\partial}{\partial n_+}=\left[1+\left(\frac{d\zeta}{dx_1}\right)^2\right]^{-1/2}\left(-\frac{d\zeta}{dx_1}\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_3}\right). \quad (2.4)$$

In what follows we will also have an occasion to use the derivative $(\partial/\partial n_-)$ that denotes differentia-

tion along the unit vector normal to the surface at each point, which is directed from the vacuum into the dielectric. The relation between these two normal derivatives is clearly

$$\frac{\partial}{\partial n_-} = -\frac{\partial}{\partial n_+} . \tag{2.5}$$

The starting point for the derivation of an exact dispersion relation for surface polaritons on a grating is Green's theorem. If $u(\vec{x})$ and $v(\vec{x})$ are arbitrary scalar fields defined in the volume V bounded by the closed surface Σ this theorem states that

$$\int_V (u\nabla^2 v - v\nabla^2 u)d^3x = \int_\Sigma \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right)dS , \tag{2.6}$$

where $(\partial/\partial n)$ is the normal derivative at the surface Σ, directed outward from inside the volume V.

We consider first the case that the volume V is the vacuum above the dielectric, $x_3 > \zeta(x_1)$. In this case the surface Σ consists of two parts: the surface S of the dielectric defined by the equation $x_3 = \zeta(x_1)$, and a hemisphere in the upper half-space, $S^{(+\infty)}$, whose radius is allowed to become infinite. We now introduce the Green's function $G_0(x_1x_3|x_1'x_3')$ that satisfies the equation

$$\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_3^2} + \frac{\omega^2}{c^2}\right)G_0(x_1x_3|x_1'x_3') = -4\pi\delta(x_1 - x_1')\delta(x_3 - x_3') , \tag{2.7}$$

and outgoing-wave or exponentially decaying boundary conditions at infinity. We multiply Eq. (2.7) from the left by $H_2^>(x_1x_3|\omega)$ and subtract from the resulting equation the equation that is obtained by multiplying Eq. (2.2a) from the left by $G_0(x_1x_3|x_1'x_3')$. When the result is integrated over the volume V and Eq. (2.6) is used, we obtain the pair of integral equations

$$\frac{1}{4\pi}\int_S \left(G_0(x_1x_3|x_1'x_3')\frac{\partial}{\partial n'}H_2^>(x_1'x_3'|\omega) - \frac{\partial}{\partial n'}G_0(x_1x_3|x_1'x_3')H_2^>(x_1'x_3'|\omega)\right)ds_1' = \begin{cases} H_2^>(x_1x_3|\omega), & x_3 > \zeta(x_1) \quad (2.8a) \\ 0, & x_3 < \zeta(x_1) \quad (2.8b) \end{cases}$$

where ds_1' is the element of path length along the surface S. Equation (2.8b) expresses the extinction theorem: The field and its normal derivative on the surface act as sources that "extinguish" the field in the dielectric medium. Since we seek a function $H_2^>(x_1x_3|\omega)$ that vanishes as $x_3 \to \infty$ there is no contribution to the integral over the surface from the integration over the surface of the hemisphere of infinite radius, $S^{(+\infty)}$.

In writing Eqs. (2.8) we have used the symmetry of $G_0(x_1x_3|x_1'x_3')$ expressed by

$$G_0(x_1x_3|x_1'x_3') = G_0(x_1'x_3'|x_1x_3) . \tag{2.9}$$

This symmetry property follows from general conditions, and is displayed explicitly in the Fourier integral representation of $G_0(x_1x_3|x_1'x_3')$ that we will use here,

$$G_0(x_1x_3|x_1'x_3') = \int_{-\infty}^{\infty} dq \, \frac{1}{\alpha(q\omega)} \exp[iq(x_1 - x_1') - \alpha(q\omega)|x_3 - x_3'|] . \tag{2.10}$$

The function $\alpha(q\omega)$ appearing in Eq. (2.10) is defined by

$$\alpha(q\omega) = \begin{cases} \left(q^2 - \frac{\omega^2}{c^2}\right)^{1/2}, & q^2 > \frac{\omega^2}{c^2} \quad (2.11a) \\ -i\left(\frac{\omega^2}{c^2} - q^2\right)^{1/2}, & q^2 < \frac{\omega^2}{c^2} . \quad (2.11b) \end{cases}$$

We now turn to the case that the volume V is the dielectric medium $x_3 < \zeta(x_1)$. In this case the surface Σ consists of the surface S of the dielectric medium and a hemisphere of infinite radius in the lower half-space $S^{(-\infty)}$. Here we introduce the Green's function $G_\epsilon(x_1x_3|x_1'x_3')$ that satisfies the equation

$$\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_3^2} + \epsilon(\omega)\frac{\omega^2}{c^2}\right)G_\epsilon(x_1x_3|x_1'x_3') = -4\pi\delta(x_1x_1')\delta(x_3 - x_3') \tag{2.12}$$

and outgoing-wave or decaying exponential boundary conditions at infinity. A Fourier integral representation of this Green's function is

$$G_\epsilon(x_1x_3|x_1'x_3') = \int_{-\infty}^{\infty} dq\, \frac{1}{\beta(q\omega)} \exp[iq(x_1 - x_1') - \beta(q\omega)|x_3x_3'|], \tag{2.13}$$

where

$$\beta(q\omega) = \left(q^2 - \epsilon(\omega)\frac{\omega^2}{c^2}\right)^{1/2} \quad \mathrm{Im}\,\beta(q\omega) < 0. \tag{2.14}$$

When we apply Green's theorem to the volume occupied by the dielectric medium, we obtain the pair of integral equations

$$\frac{1}{4\pi}\int_S \left(G_\epsilon(x_1x_3|x_1'x_3')\frac{\partial}{\partial n_+'}H_2^<(x_1'x_3'|\omega) - \frac{\partial}{\partial n_+'}G_\epsilon(x_1x_3|x_1'x_3')H_2^<(x_1'x_3'|\omega)\right)ds_1' = \begin{cases} 0, & x_3 > \zeta(x_1) \\ H_2^<(x_1x_3|\omega), & x_3 < \zeta(x_1). \end{cases} \tag{2.15a, 2.15b}$$

Equation (2.15) expresses the extinction theorem in this case. If we take into account the continuity conditions (2.3), as well as Eq. (2.5), we can rewrite Eqs. (2.15) in the form

$$-\frac{1}{4\pi}\int_S \left(\epsilon(\omega)G_\epsilon(x_1x_3|x_1'x_3')\frac{\partial}{\partial n_-'}H_2^>(x_1'x_3'|\omega) - \frac{\partial}{\partial n_-'}G_\epsilon(x_1x_3|x_1'x_3')H_2^>(x_1'x_3'|\omega)\right)ds_1' = \begin{cases} 0, & x_3 > \zeta(x_1) \\ H_2^<(x_1x_2|\omega), & x_3 < \zeta(x_1). \end{cases} \tag{2.16a, 2.16b}$$

We can convert the integrals on the right-hand sides of Eqs. (2.8) and (2.16) into integrals over the surface $x_3' = 0$ by noting that

$$ds_1' = \left[1 + \left(\frac{d\zeta(x_1')}{dx_1}\right)^2\right]^{1/2} dx_1', \tag{2.17}$$

setting $x_3' = \zeta(x_1')$ in the integrands, and using Eqs. (2.4)–(2.5). In this way we obtain for the pair of equations representing the extinction theorem, Eqs. (2.8b) and (2.16a),

$$0 = \frac{1}{4\pi}\int dx_1'\left[G_0(x_1x_3|x_1'x_3')L(x_1'|\omega) - \left(\frac{d\zeta}{dx_1'}\frac{\partial}{\partial x_1'} - \frac{\partial}{\partial x_3'}\right)G_0(x_1x_3|x_1'x_3')H(x_1'|\omega)\right]_{x_3' = \zeta(x_1')}, \quad x_3 < \zeta(x_1) \tag{2.18a}$$

$$0 = -\frac{1}{4\pi}\int dx_1'\left[\epsilon(\omega)G_\epsilon(x_1x_3|x_1'x_3')L(x_1'|\omega) - \left(\frac{d\zeta}{dx_1'}\frac{\partial}{\partial x_1'} - \frac{\partial}{\partial x_3'}\right)G_\epsilon(x_1x_3|x_1'x_3')H(x_1'|\omega)\right]_{x_3' = \zeta(x_1')}, \quad x_3 > \zeta(x_1) \tag{2.18b}$$

where we have introduced the definitions

$$H(x_1|\omega) = H_2^>(x_1x_3|\omega)\big|_{x_3 = \zeta(x_1)}, \tag{2.19a}$$

$$L(x_1|\omega) = \left[1 + \left(\frac{d\zeta(x_1)}{dx_1}\right)^2\right]^{1/2}\frac{\partial}{\partial n_-}H_2^>(x_1x_3|\omega)\big|_{x_3 = \zeta(x_1)}. \tag{2.19b}$$

None of the results obtained thus far has depended on the periodicity of the surface profile function $\zeta(x_1)$. However, because of this periodicity the magnetic field amplitude functions $H_2^<(x_1x_3|\omega)$ must possess the Bloch property

$$H_2^<(x_1 + a, x_3|\omega) = e^{ika}H_2^<(x_1x_3|\omega), \tag{2.20}$$

where k is the wave vector of the surface polariton. It follows that each of the functions $H(x_1|\omega)$ and $L(x_1|\omega)$ possesses this property as well:

$$H(x_1+a|\omega)=e^{ika}H(x_1|\omega), \tag{2.21a}$$

$$H(x_1+a|\omega)=e^{ika}L(x_1|\omega). \tag{2.21b}$$

Equation (2.18a) holds for all $x_3<\zeta(x_1)$. We obtain a simple form of this equation if we require that it be satisfied for any $x_3<\zeta_{\min}$. This is because the modulus sign in the exponent of the representation for $G_0(x_1x_3|x_1'x_3')$ given by Eq. (2.10) can be removed for such values of x_3. For the same reason we will require that Eq. (2.18b) be satisfied for any $x_3>\zeta_{\max}$. We next substitute the representations (2.10) and (2.13) into Eqs. (2.18), and use the periodicity of $\zeta(x_1)$, as well as the properties (2.21), to obtain convenient equations for $H(x_1|\omega)$ and $L(x_1|\omega)$,

$$\frac{1}{a}\int_{-a/2}^{a/2}dx_1e^{-ik_mx_1-\alpha_m(k\omega)\zeta(x_1)}\{L(x_1|\omega)-[\alpha_m(k\omega)-ik_m\zeta'(x_1)]H(x_1|\omega)\}=0, \tag{2.22a}$$

$$\frac{1}{a}\int_{-a/2}^{a/2}dx_1e^{-ik_mx_1+\beta_m(k\omega)\zeta(x_1)}\{\epsilon(\omega)L(x_1|\omega)-[\beta_m(k\omega)+ik_m\zeta'(x_1)]H(x_1|\omega)\}=0, \tag{2.22b}$$

where

$$k_m\equiv k+\frac{2\pi m}{a}, \tag{2.23a}$$

$$\alpha_m(k\omega)\equiv\alpha(k_m\omega), \tag{2.23b}$$

$$\beta_m(k\omega)=\beta(k_m\omega), \tag{2.23c}$$

and $m=0,\pm1,\pm2,\pm3,\ldots$. Equations (2.22) are to be solved for each of these values of m.

To solve these equations we expand $H(x_1|\omega)$ and $L(x_1|\omega)$ in Fourier series according to

$$H(x_1|\omega)=\sum_{n=-\infty}^{\infty}e^{ik_nx_1}\hat{H}_n(k\omega), \tag{2.24a}$$

$$L(x_1|\omega)=\sum_{n=-\infty}^{\infty}e^{ik_nx_1}\hat{L}_n(k\omega). \tag{2.24b}$$

Note that these expansions retain the Bloch property expressed by Eqs. (2.21). When the expansions (2.24) are substituted into Eqs. (2.22), and integrations by parts are used to simplify the terms containing $\zeta'(x_1)$, the Fourier coefficients $[\hat{H}_n(k\omega)]$ and $[\hat{L}_n(k\omega)]$ are found to obey the simple equations[10]

$$\sum_{n=-\infty}^{\infty}I_{m-n}^{(m)}(k\omega)\left(\frac{\omega^2/c^2-k_mk_n}{\alpha_m(k\omega)}\hat{H}_n(k\omega)+\hat{L}_n(k\omega)\right)=0, \tag{2.25a}$$

$$\sum_{n=-\infty}^{\infty}J_{m-n}^{(m)}(k\omega)\left(\frac{\epsilon(\omega)\omega^2/c^2-k_mk_n}{\epsilon(\omega)\beta_m(k\omega)}\hat{H}_n(k\omega)-\hat{L}_n(k\omega)\right)=0, \tag{2.25b}$$

where

$$I_n^{(m)}(k\omega)=\frac{1}{a}\int_{-a/2}^{a/2}dx_1e^{-i2\pi n/ax_1}e^{-\alpha_m(k\omega)\zeta(x_1)}, \tag{2.26a}$$

$$J_n^{(m)}(k\omega)=\frac{1}{a}\int_{-a/2}^{a/2}dx_1e^{-i2\pi n/ax_1}e^{\beta_m(k\omega)\zeta(x_1)}. \tag{2.26b}$$

Equations (2.25) constitute a doubly infinite set of homogeneous linear equations for the coefficients $[\hat{H}_n(k\omega)]$ and $[\hat{L}_n(k\omega)]$. The solvability condition for this set of equations, the vanishing of the determinant of the coefficients, yields the dispersion relation for surface polaritons on a grating.

In the next section we describe the numerical determination of this dispersion relation for two choices of the surface profile function. The first

is the sinusoidal profile,

$$\zeta(x_1)=\zeta_0\cos\frac{2\pi x_1}{a}; \tag{2.27}$$

the second is the symmetric sawtooth profile,

$$\zeta(x_1)=\begin{cases} h+\dfrac{4h}{a}x_1, & -\dfrac{a}{2}\le x_1\le 0 \\ h-\dfrac{4h}{a}x_1, & 0\le x_1\le\dfrac{a}{2}\,. \end{cases} \tag{2.28}$$

We conclude this section by giving the expressions for the Fourier coefficients $I_n^{(m)}(k\omega)$ and $J_n^{(m)}(k\omega)$ for each of these profiles. For the sinusoidal profile, we have

$$I_n^{(m)}(k\omega)=(-1)^n I_n(\zeta_0\alpha_m(k\omega)), \tag{2.29}$$

$$I_n^{(m)}(k\omega)=I_n(\zeta_0\beta_m(k\omega)), \tag{2.30}$$

where $I_n(x)$ is a modified Bessel function. For the sawtooth form, one finds

$$I_n^{(m)}(k\omega)=\begin{cases} \dfrac{4h\alpha_m(k\omega)}{\pi^2n^2+4h^2\alpha_m^2(k\omega)}\sinh[h\alpha_m(k\omega)], & n \text{ even} \\ \dfrac{4h\alpha_m(k\omega)}{\pi^2n^2+4h^2\alpha_m^2(k\omega)}\cosh[h\alpha_m(k\omega)], & n \text{ odd} \end{cases} \tag{2.31a, 2.31b}$$

$$J_n^{(m)}(k\omega)=\begin{cases} \dfrac{4h\beta_m(k\omega)}{\pi^2n^2+4h^2\beta_m^2(k\omega)}\sinh[h\beta_m(k\omega)], & n \text{ even} \\ \dfrac{4h\beta_m(k\omega)}{\pi^2n^2+4h^2\beta_m^2(k\omega)}\cosh[h\beta_m(k\omega)], & n \text{ odd}\,. \end{cases} \tag{2.32a, 2.32b}$$

III. DISCUSSION OF RESULTS

We have studied the dispersion relation of surface polaritons on grating structures of both sinusoidal and sawtooth profiles by searching for the zeros of the determinant formed from Eqs. (2.25a) and (2.25b). We have truncated the sum on the index n in these equations by allowing n to run from $-N+1$ to $+N$ and we are thus led to a $2N\times 2N$ determinant. In the calculations, N is chosen sufficiently large that convergence to four figures is realized. Finally, the calculations all employ the nearly-free-electron form for the dielectric constant $\epsilon(\omega)$,

$$\epsilon(\omega)=1-\frac{\omega_p^2}{\omega^2}\,, \tag{3.1}$$

with the electron plasma frequency ω_p chosen equal to 2 eV. Before we present the numerical results, some general comments are in order.

It is useful to frame our discussion in language similar to that employed in the analysis of the electronic band structure of a one-dimensional crystal lattice. We begin with the dispersion relation for the flat surface (the "empty lattice" of the band-structure discussion), introduce the appropriate Brillouin zone, and translate portions of the dispersion curve back into the first Brillouin zone. After this is done for the flat-surface dispersion curve, the influence of the grating may be appreciated in qualitative terms. We shall see that there are features in this procedure unique to the surface-polariton problem.

For the flat surface, the surface-polariton-dispersion relation is given by the well known expression

$$\frac{c^2k^2}{\omega^2}=\frac{\epsilon(\omega)}{\epsilon(\omega)+1}\,, \tag{3.2}$$

where k is its wave vector parallel to the surface. The mode exists only in the frequency regime where $\epsilon(\omega)<-1$, and for our model this constraint is obeyed for $0<\omega<\omega_{sp}=\omega_p/\sqrt{2}$. As $k\to 0$, the dispersion curve approaches the light line $\omega=ck$, and as $k\to\infty$, $\omega\to\omega_{sp}$, where $\epsilon(\omega)\equiv -1$.

In Fig. 1(a), the solid line with cross-hatched portions is a sketch of the dispersion relation of the surface polariton on the flat surface. If the grating has spatial period a, then as in the discussion of the problem of the electronic band structure, we may obtain a complete description of the dispersion relation by confining our attention entirely to the first Brillouin zone, i.e., to values of k that lie between $-\pi/a$ and $+\pi/a$. One begins with the dispersion curve illustrated in Fig. 1(a), with k ranging from $-\infty$ to $+\infty$, then translates the portion outside the first Brillouin zone back into it by moving various segments by integral multiples of $2\pi/a$. For example, the piece of the dispersion curve that lies between π/a and $3\pi/a$ is translated to the left by $2\pi/a$, that between $3\pi/a$ and $5\pi/a$ is translated in the same direction by $4\pi/a$, and so on. When this construction is completed, we have a multiband dispersion relation in the first Brillouin zone and of course, the resulting figure contains the same information as the plot in the extended-zone scheme displayed in Fig. 1(a).

There are two features of the construction just described unique to the surface-polariton problem. First of all, in the extended-zone scheme, the surface-polariton-dispersion curve (in the region $k>0$) lies entirely to the right of the light line $\omega=ck$, as illustrated in Fig. 1(a). The electric fields in the vacuum above the metal surface have the x_3 variation $\exp[-\alpha(k,\omega)x_3]$, where $\alpha(k,\omega)$

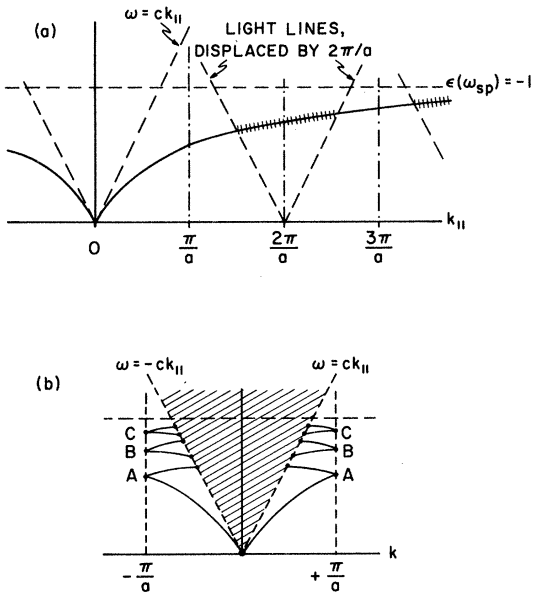

FIG. 1. (a) Sketch of the flat-surface dispersion curve, in the extended-zone scheme. As discussed in the text, the cross-hatched portion becomes unstable with respect to radiation into the vacuum above the crystal, when the periodic grating is introduced. (b) The nonradiative portions of the flat-surface dispersion curve have been translated back into the first Brillouin zone, to form the reduced-zone scheme appropriate to the present problem.

$=(k^2-\omega^2/c^2)^{1/2}$. Thus, as is well known, the fields remain "bound" to the interface only if $k > \omega/c$, so this is a necessary requirement for the surface polariton to remain bound to the surface. Now, after the various portions of the dispersion curve are translated back into the first Brillouin zone, through the construction described in the previous paragraph, the cross-hatched portions lie *between* the light line $\omega = +ck$, and that with $\omega = -ck$. This is the radiative region of the ω-k plane where, as just noted, surface polaritons cannot exist.

What happens is the following: For the flat surface, of course, the folding of the cross-hatched portion of the dispersion curve back into the first Brillouin zone is merely a relabeling procedure, and the electric fields associated with each mode decay to zero exponentially as x_3 increases, since the wave vector parallel to the surface of these modes is $k^{(n)} = k + 2\pi n/a$ with k in the first zone. Even if $|k| < \omega/c$, we have $|k^{(n)}| > \omega/c$. Now as we turn on the periodic potential by increasing the grating amplitude, the periodic potential mixes into our wave with wave vector $k^{(n)}$ modes with wave vector $k^{(m)} = k + 2\pi m/a$, with all possible values of m included (if the grating amplitude is not small). Most particularly, if $k^{(n)} = k + 2\pi n/a$, then the periodic potential adds a component with the wave vector k itself. If $|k| < \omega/c$, the mode of wave vector k does not have fields bound to the surface, but is a radiative mode. Thus, in the presence of the periodic grating, the surface polaritons on the cross-hatched regions of the dispersion curve do not remain true normal modes of the system, but acquire a finite lifetime. The periodic grating induces them to radiate their energy into the vacuum above the crystal. The surface polaritons that lie outside the periodically repeated light cone in Fig. 1(a) remain true normal modes fully bound to the surface.

In this paper, we study the dispersion relation of the true normal modes of the system. That is, in the determinant described in the first paragraph of this section, we assume the wave vector k is real, and we examine the determinant for zeros on the real axis of the frequency plane. In the reduced-zone scheme, all such normal modes lie to the right of the light line, between $k = \omega/c$ and the boundary $k = \pi/a$ of the first Brillouin zone. In our numerical work, as the wave vector k is decreased toward ω/c, to eventually cross the light line, we do lose the solutions abruptly as the light line is crossed.

A common means of exploring the surface-polariton-dispersion curve experimentally is to study the reflectivity of a surface upon which a grating has been ruled. If the surface is illuminated with p-polarized light that makes an angle θ with respect to the surface normal, then the wave vector parallel to the surface of the incident photon is $k = \omega \sin\theta/c < \omega/c$. The surface modes excited by the incident photon have wave vectors $k_{\parallel} + 2\pi n/a$. These necessarily lie on the *cross-hatched* portion of the dispersion curve in the extended-zone scheme. Thus, the true normal modes of the system explored here are, in fact, inaccessible to study by the reflectivity technique. We have underway theoretical analyses of the reflectivity of crystals with finite amplitude gratings impressed, to see to what extent the "radiative leak" contributes to the width of the reflectivity dips produced by coupling to surface polaritons on the cross-hatched portion of the dispersion curve.

Note that while the normal modes studied here may not be excited by a photon incident from the vacuum, above a metal surface upon which a diffraction grating has been ruled, one could couple to them via the attenuated-total-reflection (ATR) technique. That is, by placing a prism above the surface of a metal upon which a grating has been ruled, with an air gap in between (the Otto configuration), one may have access to the modes studied here.

In Fig. 1(b), we sketch the picture one obtains

after translating the flat-surface dispersion curve back into the first Brillouin zone, and terminating the various pieces that cross the light lines, to enter the radiative region. We see that one obtains a hierarchy of surface wave bands that crowd ever closer to ω_{sp}. In the numerical work, we have confined our attention to only two or three subbands, and we have not attempted to sort through the large number of modes that lies very close to ω_{sp}. We shall have particular interest in the gaps that open up at points such as A, B, or C, where two pieces of the dispersion curve meet at the zone boundary.

We now turn to our numerical results. We begin with the case of the sinusoidal grating. In all calculations discussed here, the plasma frequency ω_p in Eq. (3.1) has been chosen equal to 2 eV, as remarked earlier.

In Fig. 2, in the reduced zone, we present some of the segments of the dispersion curves for propagation on a sinusoidal grating with spatial period 500 Å, and with height $\zeta_0 = 50$ Å. Here the boundary at the first Brillouin zone lies at a wave vector large compared to ω/c so, at least at the zone boundary, retardation effects are unimportant.

As one can see from Fig. 2, we find two branches of the surface-polariton-dispersion curve split off quite far from the asymptotic frequency $\omega_{sp} = \omega_p/\sqrt{2}$ approached by the flat-surface dispersion curve as the wave vector $k \to \infty$. The remaining solutions lie very close to ω_{sp} for all wave vectors we have explored. We show in the drawing only the two segments split well off from ω_{sp}, since it is difficult to show more than the one pair on the graph.

It is most interesting to compare the results in Fig. 2 with the predictions of the small-amplitude-perturbation theoretic treatments put forth earlier. Most particularly, Mills[14] has presented a detailed discussion of the propagation of surface polaritons on periodic gratings within perturbation theory with emphasis on the form of the dispersion relation near the Brillouin-zone boundary.

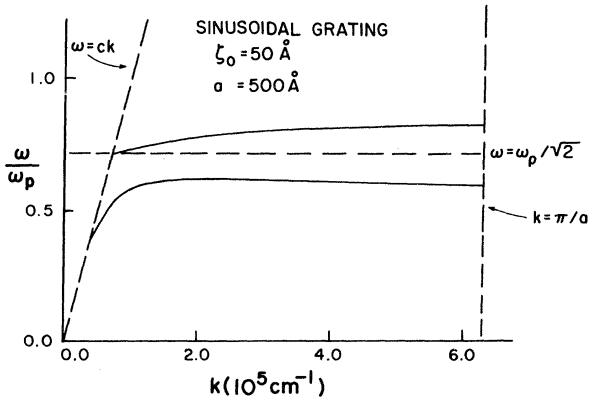

FIG. 2. Surface-polariton branches for propagation on a sinusoidal grating with spatial period a = 500 Å and height ζ_0 = 50 Å. As discussed in the text, the plasma frequency ω_p has been chosen equal to 2 eV.

From perturbation theory applied to the sinusoidal grating, in the extended-zone scheme displayed in Fig. 1(a), we expect the largest gap to open up at $k = \pm\pi/a$, since the wave at $+\pi/a$ and that at $-\pi/a$ are directly mixed by the perturbation. Mills' theory applied to the present geometry and model gives for this principal gap, in the limit retardation is ignored, the result

$$\frac{G}{\omega_p} = \frac{\pi}{\sqrt{2}}\frac{\zeta_0}{a} . \tag{3.3}$$

Here G is the magnitude of the gap which, as always in degenerate first-order perturbation theory, is linear in the amplitude of the perturbation. The gaps at $\pm 2\pi/a$, $\pm 3\pi/a$, etc., require successively higher orders of perturbation theory to generate, and we do not have simple formulas in hand for these.

If we interpret the large gap in Fig. 2 as that opened up at π/a by the grating, then the results of the figure agree qualitatively with our expectations from perturbation theory, since all other zone-boundary gaps are quite small, and the shift of the associated dispersion curve from the flat-surface curve is very small. The magnitude of the principal gap at the zone boundary in Fig. 2 is large, but its magnitude is described very well by the perturbation theory. For $\zeta_0 = 0.1a$, Eq. (3.3) predicts $G = 0.222\omega_p$, while the full calculation gives a result $G = 0.21\omega_p$ that differs by only a few percent from the perturbation theoretic expression.

Figures 3 and 4 display dispersion curves calculated again for $a = 500$ Å, but now $\zeta_0 = 100$ Å and $\zeta_0 = 300$ Å, respectively. For the case $\zeta_0 = 300$ Å, $\zeta_0/a = 0.6$, a value nearly an order of magnitude larger than the point where the Rayleigh criterion breaks down. In the calculations for $\zeta_0 = 300$ Å, results for the dispersion curve accurate to four figures were obtained by using 42×42 matrices formed from Eqs. (2.25a) and (2.25b).

Clearly, simple perturbation theory cannot be used to describe the dispersion curves in Fig. 4 since we now have a substantial number of large gaps. If we begin with Fig. 3 and compare the predicted and calculated value of G, the perturbation theory predicts $G = 0.444\omega_p$ while the calculated value of the principal gap is $0.37\omega_p$. The perturbation theory still works rather well, but

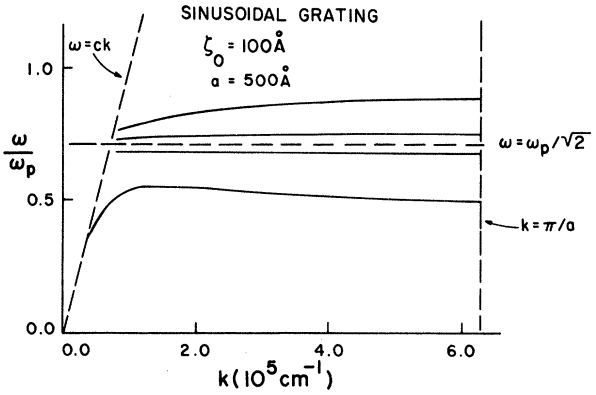

FIG. 3. Surface-polariton branches for propagation on a sinusoidal grating with spatial period $a = 500$ Å and height $\zeta_0 = 100$ Å. Again, ω_p has been chosen equal to 2 eV.

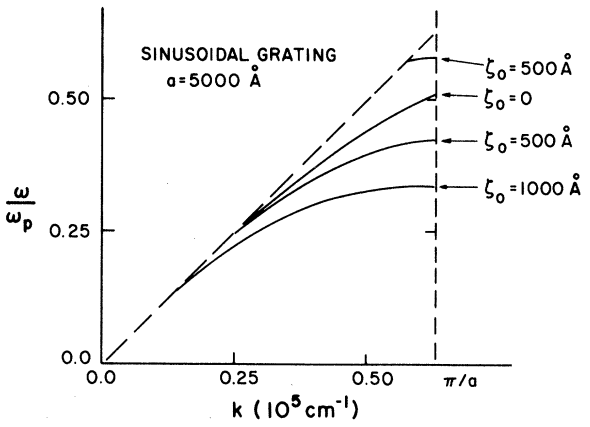

FIG. 5. For the case where the grating period is $a = 5000$ Å, and several values of ζ_0, the dispersion curves in the nonradiative region are given. The curve labeled $\zeta_0 = 0$ is that portion of the flat-surface dispersion curve which lies in the first Brillouin zone; for $\zeta_0 = 500$ Å we show the principal gap that opens up at $k = \pi/a$, and two branches of the dispersion curve, while for $\zeta_0 = 1000$ Å the upper branch is pushed well into the radiative region.

by the time we have $\zeta_0 = 0.6a$, it breaks down badly. Note that in all the examples discussed, the perturbation theory overestimates the magnitude of the gap.

Figure 5 shows dispersion curves calculated for propagation on a sinusoidal grating with spatial period $a = 5000$ Å. Now the Brillouin-zone boundary lies in a regime of wave vector where retardation effects are important. We show in the figure that portion of the flat-surface dispersion curve that lies within the first Brillouin zone. This is the curve labeled $\zeta_0 = 0$. The curves labeled $\zeta_0 = 500$ Å shows the principal gap that opens up at $k = \pi/a$, along with the two related pieces of the dispersion curve which, in the extended-zone scheme, emanate from $k = \pi/a$. By the time $\zeta_0 = 1000$ Å, evidently that upper piece of this part of the dispersion curve is driven into the radiative region, so we find no evidence of an upper branch.

We next turn to the case of the sawtooth grating. For a grating with period $a = 500$ Å, we show in Figs. 6 and 7 calculations of the dispersion curves for $h = 50$ Å and $h = 100$ Å. In each case, convergence was less good than for the sinusoidal grating. To obtain convergence when $h = 100$ Å, we had to use a 52×52 matrix, while, as indicated earlier for the sinusoidal grating, very good results were obtained with a 42×42 matrix for the case where $\zeta_0 = 300$ Å, which describes a much deeper structure.

In Fig. 6, for which $h/a = 0.1$, we see the principal gap is smaller than in the case of the sinu-

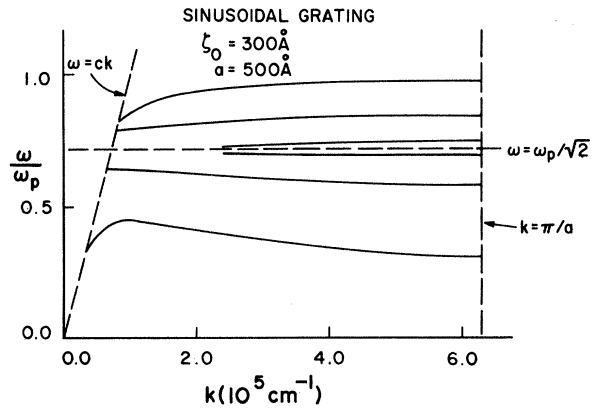

FIG. 4. Surface-polariton branches for propagation on a sinusoidal grating with spatial period $a = 500$ Å and height $\zeta_0 = 300$ Å. The plasma frequency ω_p has been chosen equal to 2 eV.

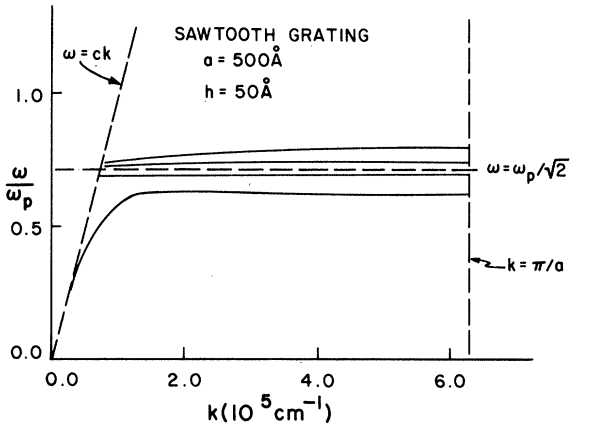

FIG. 6. Calculations of the dispersion curve for propagation on a sawtooth grating with spatial period $a = 500$ Å and height $h = 50$ Å.

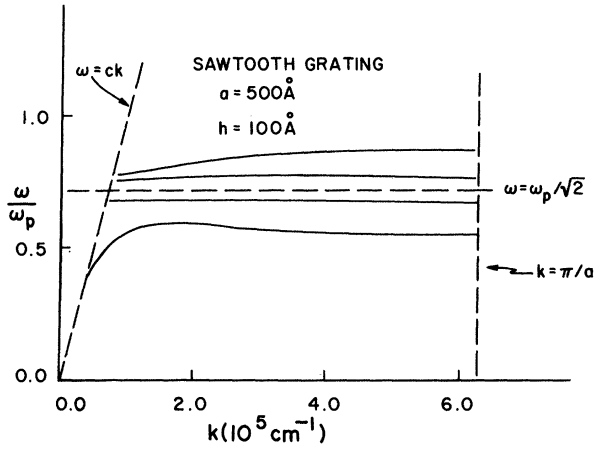

FIG. 7. Calculations of the dispersion curve for propagation on a sawtooth grating with a =500 Å and height h=100 Å.

soidal grating, and at the zone boundary we have a second gap of appreciable magnitude. Since the perturbation here is not sinusoidal, the surface profile function $\zeta(x_1)=x_3$ has a whole spectrum of Fourier components, and to first order in h, one finds gaps not only at $k=\pi/a$ as earlier, but at the other zone boundaries as well. For the sawtooth grating, application of perturbation theory[14] leads to a gap G_n at the zone boundary at $k=k_n=n\pi/a$ given by

$$\frac{G_n}{\omega_p}=\frac{2\sqrt{2}\,h}{\pi a n}[1-(-1)^n]\ . \tag{3.4}$$

Thus, perturbation theory predicts a gap at $k=\pi/a$ given by $(n=1)$

$$\frac{G_1}{\omega_p}=\frac{4\sqrt{2}\,h}{\pi a}=0.180\,, \tag{3.5a}$$

in excellent agreement with the results in Fig. 6. Perturbation theory to lowest order predicts no gap at $k=2\pi/a$, so we identify the smaller gap in the figure as that which opens up at $3\pi/a$. For this we have

$$\frac{G_3}{\omega_p}=\frac{4\sqrt{2}\,h}{3\pi a}=0.060\,, \tag{3.5b}$$

again in remarkably good agreement with the result in Fig. 6. For $h=100$ Å, as in the case of the sinusoidal grating, perturbation theory underestimates the magnitude of the gap by roughly 10%.

The above discussion provides us with a notion of the computational labor to obtain convergent results from Eqs. (2.25a) and (2.25b), and also on the limits of the perturbation theory. We have confined our attention to only the gaps at the zone boundary because the form of the perturbation theory in Ref. 14 provides results for the dispersion curve only in the near vicinity of the Brillouin-zone boundary, where (as in electron band theory) a form for the dispersion curve may be deduced from perturbed fields first order in the amplitude of the surface perturbation. As discussed in Ref. 14, an extension of the method is required to produce the form of the grating-induced shift of the surface-polariton-frequency lowest order in the grating amplitude. Such a form may be found, in implicit form, in the paper by Toigo *et al.*,[10] but the result is not as simple to apply as the elementary formulas for the gaps quoted above. We conclude with some brief remarks on another approach we have attempted to use, without success so far.

In the theory of the scattering of a scalar wave from a corrugated hard wall, Garcia and Cabrera[15] proceeded very differently from the way we have here. We have attempted to apply their approach to the present problem, but we have encountered difficulties.

In their approach, as it is applied here, one begins with Eq. (2.8a) and allows x_3 to approach $\zeta(x_1)$ from above. A second integral equation independent of this follows from Eq. (2.16b) if x_3 is allowed to approach $\zeta(x_1)$ from below. We then have two integral equations for the value of the magnetic fields $H_2^>(x_1x_3\,|\,\omega)$ and $H_2^<(x_1x_3\,|\,\omega)$ and their derivatives on the boundary. It is straightforward to arrange these so they become integral equations for the two functions $L(x_1\,|\,\omega)$ and $H(x_1\,|\,\omega)$. Finally, through use of the Bloch form of the solutions [Eqs. (2.21a) and (2.21b)], these may be reduced to integral equations on x_1' from $-a/2$ to $+a/2$. Instead of matching the fields above and below the selvedge region $[x_3>\zeta(x_1)$, $x_3<\zeta(x_1)$, respectively], we are matching them on the actual boundary surface.

These equations are difficult to solve by the Fourier-transform method, because now one must retain the absolute value $|x_3-x_3'|$ in the exponent in Eq. (2.13); the kernel of the integral equation does not decompose into a product of a function of x_3 and one of x_3'. Garcia and Cabrera proceed by working directly in coordinate space. The interval $-a/2\leq x_1\leq +a/2$ is divided into N points, so the integral equation is approximated by a matrix relation. One encounters a Fourier-transformed Green's function

$$G_0^{(k)}(x_1x_3\,|\,x_1'x_3')=\sum_{m=-\infty}^{+\infty}e^{ikma}G_0(x_1x_3\,|\,x_1'+ma,x_3')\,, \tag{3.6}$$

its companion $G_{\epsilon}^{(k)}(x_1x_3|x_1'x_3')$, and similar transforms for the Green's-function derivatives.

A tricky point in the method of Garcia and Cabrera is the following. The function $G_0^{(k)}(x_1x_3|x_1'x_3')$ is singular when $x_1' \to x_1$, $x_3' \to x_3$, though it may be represented by a properly convergent series when $x_1 \neq x_1'$, $x_3 \neq x_3'$. After the interval is divided into N segments, Garcia and Cabrera replace the singular entry

$$G_0^{(k)}(x_1\zeta(x_1)|x_1\zeta(x_1))$$

by the average of

$$G_0^{(k)}(x_1\zeta(x_1)|x_1'\zeta(x_1'))$$

over an interval of width a/N centered on x_1. The resulting average diverges as $\ln(N)$ in the limit $N \to \infty$.

In the corrugated-hard-wall problem, the basic integral equation involves only the Green's function $G_0(x_1x_3|x_1'x_3')$ and not its derivative with respect to x_3'. Our problem involves both the Green's function and its derivative. When the Garcia and Cabrera procedure is applied to the derivative terms, transformed as in Eq. (3.6), the averaging procedure just outlined produces an entry in the matrix that diverges not simply as $\ln(N)$, but as N itself. Thus, for large N, one encounters determinants very awkward to handle numerically. These have large diagonal elements; the off-diagonal elements are small (as are those that scale as $\ln N$), but in the end, the contribution of the off-diagonal elements is comparable to that from the diagonal elements.

We were very concerned that under the circumstances just outlined, one could not evaluate the resulting determinant accurately, with a reasonable amount of computer time. In the end, we assured ourselves that there was no difficulty, for matrices in the 50×50 range. At the same time, we could not find physically meaningful zeros of the determinant in the relevant range of frequency, even for modest values of ζ/a. It is possible that we chose N too small, but we found it hard to acquire confidence in the method, since it is difficult to see how the dispersion curve emerges from the resulting matrix, even for the perfectly flat surface. In contrast to this, as $\zeta(x_1)$ becomes small, one easily sees how the flat-surface dispersion relation emerges from the Fourier transform method, and it is easy to generate corrections to this in the limit of small $\zeta(x_1)$.

We abandoned the Garcia and Cabrera approach after we found the excellent convergence with the Fourier-transform method. It may be that for very large values of $\zeta(x_1)$, with a fixed, the Fourier-transform method will converge slowly, or not at all. For the sawtooth profile,[1] convergence was slower than for the sinusoidal profile, as remarked above. There may thus be a regime, possibly approached in our calculations for the sawtooth profile, where the method of Garcia and Cabrera becomes preferred, and as a consequence we believe this approach merits further study. We have calculations underway of the reflectivity of the model surfaces explored here, and these will be reported in a subsequent publication.

ACKNOWLEDGMENTS

This research has been supported by the Air Force Office of Scientific Research, through Contract No. F49620-C-0019, and one of us (B. L.) is partially supported by a fellowship from CNPQ (Conselhd Nacional De Desenvolvimento Cientifico e Technologico, Brazil).

---

*Present address: Instituto de Fisica, Unicamp, 13100, Campinas SP Brazil.

[1] A detailed experimental study of the influence of surface roughness on the reflectivity Ag films, along with a comparison to small-amplitude theories, has been presented by S. O. Sari, D. K. Cohen, and K. D. Scherkoske, Phys. Rev. B 21, 2162 (1980). This paper cites a number of earlier studies.

[2] See D. L. Jeanmaire and R. P. Van Duyne, J. Electroanal. Chem. 84, 1 (1977).

[3] John Lambe and S. L. McCarthy, Phys. Rev. Lett. 37, 923 (1976); S. L. McCarthy and J. L. Lambe, Appl. Phys. Lett. 30, 427 (1977); 33, 858 (1978); Arnold Adams, J. C. Wups, and P. K. Hansma, Phys. Rev. Lett. 42, 912 (1979); J. R. Kirtley, T. N. Theis, and J. C. Tsang, Appl. Phys. Lett. 37, 445 (1980).

[4] A discussion of the importance of surface roughness in the "giant" Raman effect has been presented by C. Y. Chen, E. Burstein, and S. Lundqvist, Solid State Commun. 32, 63 (1979).

[5] Bernardo Laks and D. L. Mills, Phys. Rev. B 20, 4962 (1979); 21, 5175 (1980); 22, 5723 (1980).

[6] U. Fano, J. Opt. Soc. Am. 31, 213 (1941).

[7] A. A. Maradudin and D. L. Mills, Phys. Rev. B 11, 1392 (1975); D. L. Mills and A. A. Maradudin, *ibid.* 12, 2934 (1975).

[8] A. Marvin, F. Toigo, and V. Celli, Phys. Rev. B 11, 2777 (1975).

[9] G. S. Agarwal, Phys. Rev. B 14, 846 (1976).

[10] F. Toigo, A. Marvin, V. Celli, and N. R. Hill, Phys. Rev. B 15, 5618 (1975).

[11] H. Chow and E. D. Thompson, Surf. Sci. 59, 225 (1976); N. Garcia, William E. Carlos, Milton W. Cole, and V. Celli, Phys. Rev. B 21, 1636 (1980).

[12]R. Petit and M. Cadilhac, C. R. Acad. Sci. B 262, 468 (1966); R. F. Millar, Proc. Cambridge Philos. Soc. 69, 175 (1971); 69, 217 (1971); R. Neviere and M. Cadilhac, Opt. Commun. 2, 235 (1970); 3, 379 (1971).

[13]A discussion of ill-conditioned matrices, and their appearance in the Rayleigh-Fano approximation to the problem of present interest has been given by N. Garcia *et al.*, Phys. Rev. B 18, 5184 (1978).

[14]D. L. Mills, Phys. Rev. B 15, 3097 (1977). We recently found a typographical error in Eq. (5.27) of this paper. The factor of $|\epsilon|^2$ in the denominator of the second term on the right-hand side should be replaced by $|\epsilon|$.

[15]N. Garcia and N. Cabrera, Phys. Rev. B 18, 576 (1978).