PHYSICAL REVIEW B VOLUME 21, NUMBER 10 15 MAY 1980

Analysis of second-harmonic generation at metal surfaces

J. E. Sipe,* V. C. Y. So, M. Fukui,† and G. I. Stegeman
Department of Physics, Erindale College, University of Toronto, Toronto, Ontario M5S 1A7, Canada
(Received 1 November 1979)

We discuss the second-harmonic generation of light at metal surfaces within the hydrodynamic theory of the electron gas; expressions for the phenomenological parameters a and b of Rudnick and Stern are presented, and the possibility of a resonance in a at optical or near-uv frequencies is discussed. A recent plasmon-enhanced experiment of Simon *et al.* is analyzed, and the use of such experiments to determine a and b is considered; new experiments are proposed to aid in such a determination.

I. INTRODUCTION

There has been a recent renewal of interest in the second-harmonic generation (SHG) of light at metal surfaces, both because SHG can be greatly enhanced by coupling with surface plasmons,[1-3] and because the same nonlinear source terms that are responsible for SHG lead to the nonlinear mixing of traveling surface plasmons[4] and to the second-harmonic generation of surface plasmons and polaritons.[5,6] These sources consist of a bulk current density which extends about a skin depth into the metal, and two "surface" current densities—one normal and one tangential to the surface—which extends only a few Fermi wavelengths into the metal. At optical frequencies these latter current densities radiate essentially as a dipole sheet at the surface. Their theoretical and experimental determination is of some interest, since they depend on the dynamics of electrons in the thin layer at the surface where the equilibrium electron density drops to zero from its bulk value.[7] Further, their magnitude is affected by the presence of adsorbed molecules,[8] and thus their observation can be expected to eventually aid in understanding the changes that occur in the electron dynamics at surfaces undergoing chemi- and physisorption. In this paper we present a theoretical treatment of these surface current densities at clean metal surfaces using the hydrodynamic theory[9] to describe the electron dynamics. We then analyze the results of a recent enhanced-SHG experiment[3] to determine to what extent such experiments are sensitive to the surface current densities, and close by suggesting a set of experiments which we feel would aid in determining these current densities. For a more general discussion of the enhancement of nonlinear optical phenomena in metals using attenuated total reflection (ATR) geometries, of which the experiments discussed herein are examples, we refer to the work of Chen and Burstein.[10]

The nonlinear surface current densities of interest have been previously investigated theoretically at a number of levels. The early work of Jha[11-13] and Bloembergen *et al.*[14] was based on the Sommerfeld free-electron model (see also Ref. 15), which leads to a linear constitutive relation of the form

$$\vec{D}(\vec{r},\omega)=\epsilon(\vec{r},\omega)\vec{E}(\vec{r},\omega)\,, \tag{1.1}$$

and a nonlinear source term of the form

$$\vec{P}^{NL}(\vec{r},2\omega)=\alpha(\vec{r},\omega)\vec{E}(\vec{r},\omega)\times\vec{H}(\vec{r},\omega) + \beta(\vec{r},\omega)\vec{E}(\vec{r},\omega)[\vec{\nabla}\cdot\vec{E}(\vec{r},\omega)]\,. \tag{1.2}$$

However, the free-electron model is clearly physically unreasonable within a few Fermi wavelengths of the surface, where in the absence of Drude damping it in fact predicts that the linear electric field diverges (see Ref. 9 and Sec. II of this paper). The limitations of the free-electron model, and the resulting ambiguities that can appear when it is applied to discuss SHG, were pointed out by Rudnick and Stern.[7] They gave a discussion of the physical nature of SHG at metal surfaces, and presented phenomenological arguments to estimate the size of the surface current densities, obtaining expressions involving phenomenological constants a and b of order unity. In addition, they presented an RPA calculation of the densities which, however, neglected the breaking of inversion symmetry at the surface. More general RPA expressions for the current densities have been presented (see, e.g., Bower[16]), but their evaluation is difficult and has not been attempted, and it is even more difficult to see just on what properties of the surface they depend. Now the hydrodynamic model of electron dynamics is certainly very crude (Griffin and Kranz[17]), but has the advantage of simplicity: It is fairly easy to see on what parameters the results depend, even without a detailed solution. The model has recently been used to discuss the behavior of linear fields at surfaces,[9,18] and the conclusions obtained were generally borne out by RPA calculations.[19,20]

It is in such a spirit of preliminary investigation, and as a preamble to more sophisiticated calculations, that we employ the hydrodynamic model here. We find that Rudnick and Stern's phenomenological parameter b is predicted to be -1, independent of frequency (as is implicit in the work of Rudnick and Stern,[21] but which is apparently not generally well known). However, the parameter a is found to depend on the "effective plasma frequency" of the nonlinear surface current density, and could exhibit a resonance at optical or near-uv frequencies.

Turning then to the experiments, we analyze the striking results of plasmon-enhanced SHG from a silver film plated on α quartz at 1.06 μm,[3] and point out the difficulties in determining the parameters from the existing data for either of the interfaces (rutile-metal and quartz-metal) present. We suggest a simple extension of this experiment which should make such an evaluation possible. We then consider the determination of the parameters a and b for the interface of more fundamental interest—metal-vacuum—and suggest an enhanced-SHG experiment more sensitive to these parameters than the experiments which have been performed, and which is feasible with existing equipment. Finally, we indicate the limitations of this work and the directions for future work which we feel would be fruitful. A preliminary account of some of our results has been previously published.[22]

II. THE HYDRODYNAMIC THEORY OF SHG AT METAL SURFACES

The use and limitations of the hydrodynamic theory to give a qualitative description of the dynamics of electrons near a surface have been dealt with by a number of authors[9,17,23-25]; here we begin by simply giving the basic equations of that theory in our notation. The total charge and current densities are

$$\begin{aligned}\rho(\vec{r},t)&=\rho^{+}(\vec{r})+\rho^{-}(\vec{r},t),\\ \vec{j}(\vec{r},t)&=-en(\vec{r},t)\vec{v}(\vec{r},t),\end{aligned}\tag{2.1}$$

where the positive and negative charge densities are given by

$$\begin{aligned}\rho^{+}(\vec{r})&=en_{+}(\vec{r}),\\ \rho^{-}(\vec{r},t)&=-en(\vec{r},t),\end{aligned}\tag{2.2}$$

and $-e$, $n_{+}(\vec{r})$, and $n(\vec{r},t)$ are, respectively, the charge on an electron, the positive ion number density, and the electron number density; $\vec{v}(\vec{r},t)$ is the electron velocity field. The densities (2.1) satisfy the equation of continuity,

$$\vec{\nabla}\cdot\vec{j}+\dot{\rho}=0,\tag{2.3}$$

and Euler's equation for the electron fluid,

$$mn[\partial\vec{v}/\partial t+(\vec{v}\cdot\vec{\nabla})\vec{v}]=-en\vec{E}-(en/c)\vec{v}\times\vec{B}-\vec{\nabla}p,\tag{2.4}$$

where m is the electron mass and $p(r,t)$ is the "quantum pressure"[9]; in the Thomas-Fermi theory,

$$p(\vec{r},t)=\zeta[n(\vec{r},t)]^{5/3},\tag{2.5}$$

where

$$\zeta=(3\pi^2)^{2/3}\hbar^2/5m.\tag{2.6}$$

The electric and magnetic fields $\vec{E}$ and $\vec{B}$ appearing in Eq. (2.4) satisfy the Maxwell equations which, when a polarization potential $\vec{P}(\vec{r},t)$ is introduced by virtue of Eq. (2.3),

$$\vec{j}=\dot{\vec{P}},\quad \rho=-\vec{\nabla}\cdot\vec{P},\tag{2.7}$$

take a form familiar from dielectric theory,

$$\begin{aligned}&\vec{\nabla}\cdot\vec{E}=-4\pi\vec{\nabla}\cdot\vec{P},\quad \vec{\nabla}\cdot\vec{B}=0,\\ &c\vec{\nabla}\times\vec{B}-\dot{\vec{E}}=4\pi\dot{\vec{P}},\quad c\vec{\nabla}\times\vec{E}+\dot{\vec{B}}=0.\end{aligned}\tag{2.8}$$

To calculate the SHG, we now expand all fields in the usual way,[14,21]

$$\begin{aligned}n(\vec{r},t)&=n_0(\vec{r})+n_1(\vec{r},t)+n_2(\vec{r},t)+\cdots,\\ \vec{E}(\vec{r},t)&=\vec{E}_0(\vec{r})+\vec{E}_1(\vec{r},t)+\vec{E}_2(\vec{r},t)+\cdots,\\ \vec{v}(\vec{r},t)&=\vec{v}_1(\vec{r},t)+\vec{v}_2(\vec{r},t)+\cdots,\\ \vec{B}(\vec{r},t)&=\vec{B}_1(\vec{r},t)+\vec{B}_2(\vec{r},t)+\cdots.\end{aligned}\tag{2.9}$$

We find that the zeroth-order equations resulting from Eqs. (2.4), (2.7), and (2.8) give the usual Thomas-Fermi equations for the equilibrium electron density $n_0(\vec{r})$, when Eqs. (2.5) and (2.6) are used. The first-order equation derived from Eqs. (2.4) and (2.7) may be written in the form

$$\ddot{\vec{P}}_1(\vec{r},t)-\vec{L}(\vec{r})\cdot\vec{P}_1(\vec{r},t)=[e^2n_0(\vec{r})/m]\vec{E}_1(\vec{r},t),\tag{2.10}$$

where

$$\vec{L}=-(5/9m)\zeta n_0^{-1/3}(\vec{\nabla}n_0)\vec{\nabla}+(5/3m)\zeta n_0^{2/3}\vec{\nabla}\vec{\nabla},\tag{2.11}$$

and the second-order equation reduces to

$$\ddot{\vec{P}}_2(\vec{r},t)-\vec{L}(\vec{r})\cdot\vec{P}_2(\vec{r},t)=[e^2n_0(\vec{r})/m]\vec{E}_2(\vec{r},t)+\vec{S}_f(\vec{r},t)+\vec{S}_p(\vec{r},t),\tag{2.12}$$

where the source terms $\vec{S}_f(\vec{r},t)$ and $\vec{S}_p(\vec{r},t)$ are given by

$$\vec{S}_f = en_0(\vec{v}_1\cdot\vec{\nabla})\vec{v}_1 - e\vec{v}\dot{n}_1 + (e^2/m)n_1\vec{E}_1 + (e^2n_0/mc)\vec{v}_1\times\vec{B}_1 \qquad (2.13)$$

and

$$\vec{S}_p = (5e/9m)\zeta\vec{\nabla}(n_1^2 n_0^{-1/3}) . \qquad (2.14)$$

The first two terms in Eq. (2.13) are the purely convective sources that always appear in an expansion solution of Euler's equation; the second two terms are cross terms between matter and electromagnetic fields specific to the Lorentz force appearing on the right-hand side of Eq. (2.4); the term (2.14) is the new second-order source due to the presence of the quantum pressure. Of course, the terms appearing in Eq. (2.13) take on different values than they would if the quantum pressure were neglected in solving Eq. (2.10).

We consider first the solution of Eq. (2.10) and the first-order Maxwell equations. This has been discussed in some detail,[9,18,25] in connection with the theory of multipole surface plasmons, for the special case to which we now restrict ourselves. A jellium model is assumed for the background positive-ion density, leading to an equilibrium electron density $n_0(\vec{r}) = n_0(z)$, where there is a surface parallel to the $z=0$ plane in the neighborhood of which $n_0(z)$ drops from its bulk value to zero over a distance on the order of the Fermi wavelength λ_F. As discussed in Ref. 25, it is possible to take advantage of the fact that deep within the metal it is permissible to neglect $\vec{L}(\vec{r})$ in Eq. (2.10) at optical and near-uv frequencies, since it leads only to corrections of order $(\lambda_F/\lambda)^2 \ll 1$, where λ is the wavelength of light *in vacuo*. We place the $z=0$ plane so that from $-\infty < z \leq 0$ (the vector $\hat{z}$ is taken to point towards the vacuum), $\vec{L}(\vec{r})$ may be neglected; this region we refer to as the bulk. The equilibrium electron density then drops to ~0 at $z=l$, where $l \sim \lambda_F$; we refer to the region $0 < z \lesssim l$, where $\vec{L}(\vec{r})$ may not be neglected, as the selvedge. We shall see that, as in Ref. 25, our results are independent of the exact location of the $z=0$ plane, as long as $l \ll \lambda$.

Writing all first-order fields as

$$f_1(\vec{r},t) = f_1(\vec{r})e^{-i\omega t} + \text{c.c.} = 2\,\text{Re}[f_1(\vec{r})e^{-i\omega t}] , \qquad (2.15)$$

where we take

$$f_1(\vec{r}) = f_1(z)e^{i(k_x x + k_y y)} , \qquad (2.16)$$

and $\vec{k} = (k_x, k_y, 0)$ is the wave vector parallel to the surface, in the bulk we find that Eq. (2.10) reduces to

$$\vec{P}_1(\vec{r}) = (-e\bar{n}_0/m\omega^2)\vec{E}_1(\vec{r}) \equiv \bar{\chi}\vec{E}_1(\vec{r}) \quad (z<0) , \qquad (2.17)$$

where $\bar{n}_0$ is the bulk value of $n_0(z)$. Equation (2.17) leads to the usual dielectric constant

$$\bar{\epsilon} = 1 + 4\pi\bar{\chi} = 1 - \bar{\omega}_p^2/\omega^2 , \qquad (2.18)$$

where $\bar{\omega}_p$,

$$\bar{\omega}_p^2 = 4\pi e^2\bar{n}_0/m , \qquad (2.19)$$

is the bulk plasma frequency of the metal. In the selvedge region, we treat Eq. (2.10) by writing

$$\vec{E}_1(z) = \vec{E}_{1b} + \vec{E}_{1s}(z) , \qquad (2.20)$$

where $\vec{E}_{1b}$ is the sum of the incident field and the linear field from currents in the bulk; over the selvedge region $\vec{E}_{1b}$ may be taken as uniform, since $l \ll \lambda$. The field $\vec{E}_{1s}(z)$ is the field in the selvedge due to currents in the selvedge. It is given by

$$\vec{E}_{1s}(z) = \int_0^l \vec{G}(z-z')\cdot\vec{P}_1(z')dz' , \qquad (2.21)$$

where $\vec{G}(z-z')$ is the tensor Green function of Eqs. (2.8), (2.15), and (2.16), given, e.g., in Sipe.[25] To lowest order in l/λ we may put

$$\vec{G}(z-z') \simeq -4\pi\hat{z}\hat{z}\,\delta(z-z') \qquad (2.22)$$

in Eq. (2.21). To the same order in l/λ we then find that for points in the selvedge, Eq. (2.10) reduces to

$$P_1^k(z) = [-e^2n_0(z)/m\omega^2]E_{1b}^k , \qquad L_0(\omega)P_1^s(z) = [e^2n_0(z)/m]E_{1b}^s , \qquad (2.23)$$

where the operator $L_0(\omega)$ is given by

$$L_0(\omega) = \omega_p^2(z) - \omega^2 + \frac{5}{9m}\zeta n_0^{-1/3}(z)\frac{dn_0(z)}{dz}\frac{d}{dz} + \frac{5}{3m}\zeta n_0^{2/3}(z)\frac{d^2}{dz^2} . \qquad (2.24)$$

$\omega_p(z)$ is the "local plasma frequency,"

$$\omega_p^2(z) = 4\pi e^2 n_0(z)/m , \qquad (2.25)$$

and we have restricted ourselves to p-polarized light and written

$$\vec{P}_1(z) = P_1^s(z)\hat{z} + P_1^k(z)\hat{k} , \qquad (2.26)$$

etc., where $\hat{k} = \vec{k}/|\vec{k}|$. The second of Eqs. (2.23) and Eq. (2.24) illustrate the unphysical nature of the assumption of a "local" theory ($\zeta = 0$) with a smooth density profile $n_0(z)$, $P_1^s(z)$ in this instance diverges at z such that $\omega_p(z) = \omega$. Although this divergence may formally be removed by introducing a Drude collision lifetime in Eq. (2.4), the large values of $P_1^s(z)$ that still result when $\omega_p(z) \sim \omega$ are clearly unphysical, as discussed by Eguiluz and Quinn.[9]

To decouple Eqs. (2.7) and (2.23) we now again invoke the inequality $l/\lambda \ll 1$ to neglect the electric field in the bulk due to currents in the selvedge. Then the bulk fields may be found by solving the Maxwell equations with dielectric constant

$$\epsilon(z) = \begin{cases} 1 - \bar{\omega}_p^2/\omega^2 & z < 0, \\ 1 & z > 0, \end{cases} \tag{2.27}$$

and the field in the selvedge may finally be determined from Eqs. (2.23). Since E_{1b} is the incident field plus the field *from* currents in the bulk (rather than the field *in* the bulk), we have

$$\begin{aligned} E_{1b}^k &= E_1^k(z=0^-), \\ E_{1b}^s &= \bar{\epsilon}^{-1} E_1^s(z=0^-). \end{aligned} \tag{2.28}$$

The approximation (2.22) and the neglect of the field generated in the selvedge on the bulk results in the neglect of the effects of the details of the charge-current distribution in the selvedge on the surface plasmon frequencies. This introduces no serious error for $l \ll \lambda$, as long as $L_0^{-1}(\omega)$ has no poles near frequencies of interest (Refs. 9 and 18 where, as in Ref. 25, an approximate expression for L_0 was adopted). Such poles signal the presence of multipole surface plasmons, and, if they appear, either a numerical solution[9] of Eqs. (2.10) or a bulk-selvedge coupling theory[25] applied to Eqs. (2.17) and (2.23) must be used to obtain the linear fields. Except for our comments in the concluding remarks, we shall neglect the possible existence of multipole surface plasmons in this work. We note, however, that the existence of the lowest multipole surface plasmon has been both predicted in RPA calculations[20] and seen experimentally[26] in metallic systems covered with metal overlayers. This lowest multipole surface plasmon, however, is predicted to exist in such layered systems even within a "local" response theory.[27]

Returning to our problem at hand, we now consider the second-order equation (2.12). We write the second-harmonic component of all second-order fields as

$$\begin{aligned} f_2(\vec{r},t) &= f_2(\vec{r})e^{-i\Omega t} + \text{c.c.}, \\ f_2(\vec{r}) &= f_2(z)e^{i\vec{K}\cdot\vec{r}}, \end{aligned} \tag{2.29}$$

where $\Omega = 2\omega$ and $\vec{K} = 2\vec{k}$, and then find that in the bulk, where $n_1 \equiv e^{-1}\vec{\nabla}\cdot\vec{P}_1 = 0$ and $n_0(z) = \bar{n}_0$, Eq. (2.12) reduces to

$$\vec{P}_2(\vec{r}) = \bar{\chi}\vec{E}_2(\vec{r}) - \frac{e}{8m\omega^2}\left(\frac{\epsilon-1}{4\pi}\right)\vec{\nabla}[\vec{E}_1(\vec{r})\cdot\vec{E}_1(\vec{r})], \tag{2.30}$$

where in (and only in) Eq. (2.30),

$$\bar{\chi} \equiv -e^2\bar{n}_0/m\Omega^2. \tag{2.31}$$

The source term in Eq. (2.30) is the usual expression for the SHG source term in a uniform electron gas (cf., e.g., Ref. 15, and references cited therein). We note that it is not sufficient to keep only the Lorentz term proportional to $\vec{v}_1 \times \vec{B}_1$ in Eq. (2.13); in an experiment involving two waves in the metal, as we discuss in Sec. III, the term proportional to $(\vec{v}_1\cdot\vec{\nabla})\vec{v}_1$ is nonzero. Both of these terms are included in Eq. (2.30). Turning to the selvedge region, we proceed as in the final analysis of the linear fields and neglect all but the lowest-order terms in l/λ. We find that the field $\vec{P}_2(z)$ in the selvedge responsible for the generation of significant second-harmonic fields outside the selvedge is given by

$$\begin{aligned} -\Omega^2 P_2^k(z) &= S^k(z), \\ L_0(\Omega)P_2^s(z) &= S^s(z), \end{aligned} \tag{2.32}$$

where

$$\begin{aligned} S^k &= \frac{2e}{m}E_{1b}^k\frac{\partial P_1^s}{\partial z}, \\ S^s &= \frac{e}{m}E_{1b}^s\frac{\partial P_1^s}{\partial z} + \frac{\partial V}{\partial z}, \end{aligned} \tag{2.33}$$

and

$$V = \frac{5e}{9m}\zeta n_1^2 n_0^{-1/3} - \left(\frac{\omega^2}{en_0} + \frac{2\pi e}{m}\right)(P_1^s)^2. \tag{2.34}$$

A final simplification is possible because $(l/\lambda) \ll 1$. Only the dipole moment per unit area, $\vec{Q}$, of the selvedge need be considered,[7,25]

$$\vec{Q}e^{i\vec{K}\cdot\vec{r}} = \int_0^l \vec{P}_2(z)e^{i\vec{K}\cdot\vec{r}}dz. \tag{2.35}$$

The second-harmonic source terms in the selvedge radiate as a dipole sheet placed *outside* the bulk metal, as is clear from our development here. All the shielding of the currents in the selvedge by the motion of charges in the selvedge is taken into account in Eqs. (2.32).

The first of Eqs. (2.32) may easily be solved, since the tangential component of the linear electric field is essentially uniform over the selvedge [Eqs. (2.20)–(2.23)]. We find

$$\begin{aligned} Q^k &= (e/2m\omega^2)E_{1b}^k P_1^s(z=0^-) \\ &= \frac{e}{2m\omega^2}\left(\frac{\epsilon-1}{4\pi}\right)E_1^k(z=0^-)E_1^s(z=0^-), \end{aligned} \tag{2.36}$$

since $P_1^s(z \gtrsim 0) = 0$, $E_{1b}^k = E_1^k(z=0^-)$, and we have used Eqs. (2.7) and (2.18); the details of the linear charge density in the selvedge do not affect Q^k in our present model; only the total charge confined

to the surface is important.[21] This is, however, not the case for Q^s; to determine that component we would have to solve the second of Eqs. (2.23) for $P_1^s(z)$ with a given $n_0(z)$, and then solve the second of Eqs. (2.32). We do not consider this worthwhile because of the crudeness of the hydrodynamic model (cf. Rudnick and Stern[7]); rather, we only wish to obtain an order-of-magnitude estimate on the size of $\vec{Q}$, and some insight into the parameters on which $\vec{Q}$ might depend. To this end, we define ω_0 according to

$$\int_0^l L_0(\Omega)P_2^s(z)dz \equiv (\omega_0^2 - \Omega^2)Q^s , \tag{2.37}$$

where the $P_2^s(z)$ in Eqs. (2.37) is the exact solution of Eqs. (2.23) and (2.32). Comparing Eqs. (2.24), (2.32), and (2.35), we see that ω_0 can be interpreted as the "effective plasma frequency" of the second-harmonic currents in the selvedge. If $P_2^s(z)$ existed predominantly in the region where $n_0/\bar{n}_0 \ll 1$, we would have $\omega_0 \simeq 0$; if on the other hand it existed predominantly in the region where $n_0/\bar{n}_0 \simeq 1$, we would have $\omega_0 > \bar{\omega}_p$, the inequality appearing because of the effect of the pressure term in Eq. (2.4) in increasing the plasma frequency as the wavelength of the charge oscillation is decreased. We may confidently expect that a full solution of the hydrodynamic model would lead to an ω_0 in the range $0 \leq \omega_0 \leq 2\bar{\omega}_p$. Using Eqs. (2.17)–(2.19), (2.35), and (2.37) we finally obtain

$$Q^s = -\frac{e}{m(\omega_0^2 - \Omega^2)}\left(\frac{\epsilon - 1}{4\pi}\right)\left(\frac{\epsilon + 3}{2}\right)[E_1^s(z = 0^-)]^2 , \tag{2.38}$$

which completes our analysis, as far as we shall carry it, of the second-harmonic source terms within the hydrodynamic model.

We close this section by comparing our results with those of Rudnick and Stern,[7] who estimated $\vec{Q}$ from phenomenological arguments. They argued that the components should be given by

$$\begin{aligned} Q^k &= \frac{e^3 n_0 b}{2m^2\omega^4} E_1^k(z = 0^-)E_1^s(z = 0^-) , \\ Q^s &= \frac{e\bar{\omega}_p^2 a}{16\pi m\omega^4}[E_1^s(z = 0^-)]^2 , \end{aligned} \tag{2.39}$$

where a and b are to be of order unity. We have expressed Rudnick and Stern's results in our notation (note in particular that they denote the charge on an electron by e, while we denote it by $-e$). Comparing Eqs. (2.36), (2.38), and (2.39) we see that the expressions of Rudnick and Stern agree with ours if we set

$$\begin{aligned} b &= -1 , \\ a &= -2\left(\frac{\bar{\omega}_p^2 - \Omega^2}{\omega_0^2 - \Omega^2}\right) . \end{aligned} \tag{2.40}$$

Thus we at least partly confirm the arguments of Rudnick and Stern; if $\omega_0 \gtrsim \bar{\omega}_p$ and $\Omega \ll \bar{\omega}_p$, both a and b should be of order unity. But if $\omega_0 \simeq \Omega \equiv 2\omega$, Eq. (2.38) can lead to a very large $|a|$. The phenomenological arguments of Rudnick and Stern[7] employed the zero-frequency response of the electron gas, and thus did not show such a resonance. However, the resonance denominator in Eq. (2.38) should not be taken too seriously, since the ω_0 defined by Eq. (2.37) may depend on Ω, and the hydrodynamic model itself may give an oversimplified description of the electron dynamics near the surface. Nonetheless, it is not unreasonable to expect a resonance in Q^s associated with an "effective plasma frequency" on general physical grounds, since the part of the second-harmonic currents responsible for Q^s involves essentially longitudinal oscillations of charge density perpendicular to the surface. Q^k on the other hand, which results simply because the charge density accumulated at the surface is driven in second order parallel to the surface, should be free from such a resonance. The damping of this longitudinal oscillation might of course be very large, or it might occur sufficiently near $\bar{\omega}_p$ so that, at optical and near-uv frequencies where the theory developed here is valid, it could not be observed. But we believe that this investigation indicates that further theoretical work on the dispersion of a, within more realistic models of the electron dynamics, and especially considering the effect to adsorbed species, would be worthwhile. The experimental possibilities for determining a are discussed in Sec. III; the effects of surface roughness, which we have completely neglected in our treatment in this section, are mentioned in Sec. IV.

III. SHG IN GEOMETRIES OF INTEREST

We now turn to an analysis of some experiments involving SHG from metal surfaces. The source terms are those discussed in Sec. II. In the bulk we have a source polarization

$$\vec{P}_{NL}(\vec{r}) = -\frac{e}{8m\omega^2}\left(\frac{\epsilon - 1}{4\pi}\right)\vec{\nabla}[\vec{E}_1(\vec{r}) \cdot \vec{E}_1(\vec{r})] \tag{3.1}$$

[cf. Eq. (2.30)], while at a metal-vacuum selvedge we have an effective current density

$$\vec{J}_{NL}(\vec{r}) = -i(2\omega)\vec{Q}e^{i\vec{K}\cdot\vec{r}}\delta(z - z_0^{(\pm)}) \tag{3.2}$$

[cf. Eqs. (2.7) and (2.35)], where the interface is at $z = z_0$, and

$$Q^k = (\mp)\frac{eb}{2m\omega^2}\left(\frac{\epsilon-1}{4\pi}\right)E_1^x(z_0^{(\mp)})E_1^k(z_0^{(\mp)}),$$
$$Q^x = (\mp)\frac{ea}{4m\omega^2}\left(\frac{\epsilon-1}{4\pi}\right)[E_1^x(z_0^{(\mp)})]^2, \tag{3.3}$$

where, for convenience in comparing with earlier work, we write $\vec{Q}$ in terms of a and b; however, we shall set $b=-1$ [Eq. (2.40)] in the following discussions unless otherwise specified. The signs in Eqs. (3.2)–(3.3) refer to the orientation of the surface: If $\hat{z}$ points from metal to vacuum, as in Sec. II, the upper sign should be used, while if $\hat{z}$ points from vacuum to metal the lower sign is appropriate. This guarantees, in Eq. (3.2), that the current sheet is placed outside the "bulk metal" [see comment following Eq. (2.35)] and it corrects Eq. (3.3) when the geometry for which Eqs. (2.36), (2.38) and (2.39) were derived is inverted. The field $\vec{E}_1$ in Eqs. (3.1) and (3.3) is the linear electric field, which may be approximated using the usual macroscopic Maxwell theory with a step discontinuity in the dielectric constant at the surface. The fields in Eq. (3.3) are then to be evaluated on the metal side of the interface, as indicated by the $(\mp)$. Finally, we note that since Eq. (3.1) is a bulk term it should be evaluated *up to but not including* the contribution from the discontinuity in the approximate $\vec{E}_1$ over the step discontinuity in the dielectric constant. The inclusion of that contribution would incorrectly treat effects already correctly described by Eq. (3.3).

Once the sources (3.1)–(3.3) are specified, the generated second-harmonic fields can be calculated either by a straightforward application of the macroscopic Maxwell equations or, more easily, by using a simple extension of the theory of transfer matrices which we give in the Appendix. To describe the propagation of these fields through the metal, we use not the dielectric constant (2.18) which was derived from the simple theory of Eq. (2.4), but rather the observed dielectric constants at ω and 2ω. The ability to do this, which results from our division of the response problem into "bulk" and "selvedge" parts, is crucial; especially in plasmon-enhanced-SHG experiments, the observed SHG is very sensitive to the linear properties of the metal, as pointed out by Simon *et al.*,[1] and as is clear from our discussion below.

Perhaps the most striking of the enhanced-SHG experiments is that of Simon *et al.*,[3] where light at 1.06 μm is coupled through a rutile prism into a metal film bounded by a quartz crystal (Kretschmann geometry). SHG occurs in the quartz, bulk metal, and at both the rutile-metal and metal-quartz selvedges. It varies in intensity over about four orders of magnitude, as the incident angle is scanned over a few degrees, due to enhancement by passing through first- and second-harmonic surface plasmon resonances at the metal-quartz interface. As in other enhanced-SHG experiments, the surfaces here are not the simple metal-vacuum surfaces of fundamental interest, but rather more complicated metal-dielectric surfaces. To treat these surfaces within the framework of the theory of Rudnick and Stern[7] and Sec. II, we assume they may be described by putting a vacuum gap between a metal-vacuum selvedge, treated according to Sec. II, and a vacuum-dielectric surface, treated by normal macroscopic electrodynamics; the thickness of this vacuum gap is then allowed to vanish. That is, the system is treated as shown in Fig. 1, with the interfaces at $z_0=0$ and d. We turn later in this section to experiments directly sensitive to metal-vacuum selvedges.

Calculating the SHG in terms of the parameters of the bulk and those of the selvedge current sheets, we find

$$I(2\omega) = \mathcal{R}[I(\omega)]^2, \tag{3.4}$$

where

$$I(\omega) = 2(c/4\pi)n_1|E_{1+}^{\omega}|^2$$

and

$$I(2\omega) = 2(c/4\pi)N_1|E_{1-}^{2\omega}|^2$$

are the intensities of the upward-propagating incident field and the downward-propagating second-harmonic field, respectively (see Fig. 1). E_{1+}^{ω} and $E_{1-}^{2\omega}$ are the amplitudes of these fields, and n_1 and N_1 are, respectively, the refractive indices of ru-

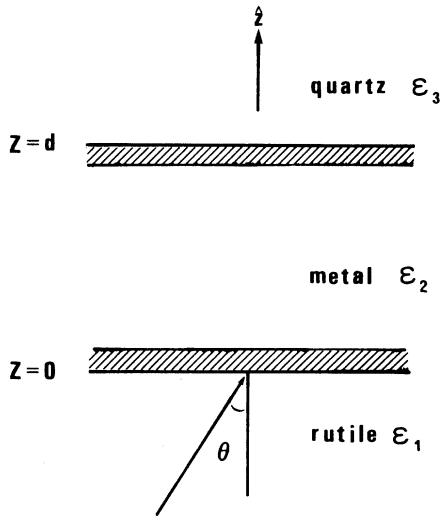

FIG. 1. The model used for the rutile-metal and metal-quartz surfaces. The shaded areas indicate the selvedge current sheets, which are formally placed in vacuum gaps.

tile at ω and 2ω. The conversion factor for the reflected SHG, $\mathcal{R}$, is conveniently written in the form

$$\mathcal{R}=\frac{2\pi}{c}\frac{N_1}{n_1^2}|A|^2, \tag{3.5}$$

where A is the ratio $E_1^{2\omega}/(E_{1+}^{\omega})^2$,

$$A=A_q+A_m+A_0+A_d, \tag{3.6}$$

the contributions indicated in Eq. (3.6) being those from the quartz, the bulk metal, and the selvedges at $z=0$ and d, respectively. We find

$$A_q=\frac{-4\pi\tilde{\Omega}^2 d_{11}\{K[e^x)^2-(e^k)^2]-2W_3e^xe^k\}}{W_3\tilde{\Omega}_3(2w_3+W_3)} \times T_{21}e^{iW_2d}T_{32}e^{2iw_3d}M, \tag{3.7}$$

for the orientation of the quartz crystal specified by Simon *et al.*,[3] where d_{11} is the nonlinear susceptibility (Zernike and Midwinter[28]), and e^x and e^k are the ratios of the amplitudes of the components of the incident linear field in the quartz to E_{1+}^{ω},

$$\begin{Bmatrix} e^x \\ e^k \end{Bmatrix}=\begin{Bmatrix} k\tilde{\omega}_3^{-1} \\ -w_3\tilde{\omega}_3^{-1} \end{Bmatrix} t_{23}t_{12}e^{i(w_2-w_3)d}m. \tag{3.8}$$

Here

$$m=(1-r_{23}r_{21}e^{2iw_2d})^{-1}, \qquad M=(1-R_{23}R_{21}e^{2iW_2d})^{-1} \tag{3.9}$$

are the modification factors which account for multiple reflections in the metal film at ω and 2ω. We have set $\tilde{\omega}=\omega/c$,

$$\tilde{\omega}_i=\tilde{\omega}[\epsilon_i(\omega)]^{1/2}, \qquad w_i=(\tilde{\omega}_i^2-k^2)^{1/2}, \tag{3.10}$$

and the Fresnel coefficients at ω for reflection and transmission, r_{ij} and t_{ij}, are given by

$$r_{ij}=\frac{w_i\epsilon_j-w_j\epsilon_i}{w_i\epsilon_j+w_j\epsilon_i}, \qquad t_{ij}=\frac{2(\epsilon_i\epsilon_j)^{1/2}w_i}{w_i\epsilon_j+w_j\epsilon_i}, \tag{3.11}$$

for our case of p polarization. The corresponding capital letters indicate these quantities evaluated at $\Omega=2\omega$ and $K=2k$, and the subscripts 0, 1, 2, and 3 indicate vacuum, rutile, metal, and quartz, respectively. From the bulk metal we find

$$A_m=T_{21}M\left(\frac{ek(\epsilon_2-1)}{2imc^2W_2\tilde{\Omega}_2\tilde{\omega}_2^2}\right)(t_{12}m)^2\{e^{i(W_2+2w_2)d}(1+R_{23})[k^2(1+r_{23})^2+w_2^2(1-r_{23})^2] -(1+R_{23}e^{2iW_2d})[k^2(1+r_{23}e^{2iw_2d})^2+w_2^2(1-r_{23}e^{2iw_2d})^2]\}, \tag{3.12}$$

from the rutile-metal selvedge current sheet we find

$$A_0(E_{1+}^{\omega})^2/2\pi i\tilde{\Omega}=T_{01}(1-R_{02}R_{01})^{-1}[KW_0^{-1}(1+R_{02})Q_0^x+(1-R_{02})Q_0^k] +T_{21}R_{23}e^{2iW_2d}MT_{02}(1-R_{02}R_{01})^{-1}[KW_0^{-1}(1+R_{01})Q_0^x-(1-R_{01})Q_0^k], \tag{3.13}$$

and from the metal-quartz selvedge current sheet we find

$$A_d(E_{1+}^{\omega})^2/2\pi i\tilde{\Omega}=T_{21}Me^{iW_2d}T_{02}(1-R_{03}R_{02})^{-1} \times[KW_0^{-1}(1+R_{03})Q_d^x+(1-R_{03})Q_d^k], \tag{3.14}$$

where $\vec{Q}_{0,d}$ are the dipole moments per unit area of the current sheets at $z=0$ and d, Eqs. (3.3). The linear field in the metal used to evaluate them is

$$\vec{E}(z)=(\hat{p}_{2+}e^{iw_2z}+r_{23}e^{2iw_2d}\hat{p}_{2-}e^{-iw_2z})t_{12}mE_{1+}^{\omega}, \tag{3.15}$$

where

$$\hat{p}_{i\pm}=(k\hat{z}\mp w_i\hat{k})/\tilde{\omega}_i. \tag{3.16}$$

To calculate these quantities we use the optical constants given in Table I. The recently determined value of the dielectric constant of silver at 0.53 μm that we adopt is in good agreement with the values of Otter,[34] Dujardin and Theye,[31] Johnson and Christy,[35] and Weber and McCarthy.[36] However, there is serious disagreement concerning the dielectric constant of silver at 1.06 μm. The values of Johnson and Christy[35] and Adams *et al.*[37] are rather different from that given in Table I, but the latter value is chosen because the cited investigators did more extensive experiments in the infrared. At least part of the disagreement is likely to be due to the variation in the optical properties of a film with the conditions and rate of deposition[31]; future investigations of SHG should probably be preceded by experiments, such as those of Lafair *et al.*,[38] to determine the optical constants of the film under study.

In Fig. 2 we plot $\mathcal{R}_q=(2\pi N_1/cn_1^2)|A_q|^2$ and $\mathcal{R}_m=(2\pi N_1/cn_1^2)|A_m|^2$, the conversion factors that

TABLE I. Optical constants.

	ω (1.06 μm)	2ω (0.53 μm)
Rutile (Ref. 29)		
refractive index	2.479	2.671
Silver ($d = 550$ Å)		
dielectric constant	$-67.03 + i2.44$ (Refs. 30, 31)	$-11.9 + i0.33$ (Ref. 32)
Quartz (Ref. 33)		
refractive index	1.536	1.542
nonlinear susceptibility:		
$d_{11} = 0.39 \times 10^{-12}$ (m/volt)		
$= 0.93 \times 10^{-9}$ (esu, cgs)		

would be observed if SHG only in the quartz and only in the bulk metal, respectively, were present; we also plot $\mathcal{R}_{qm} = (2\pi N_1/cn_1^2)|A_q + A_m|^2$. The factor $\mathcal{R}_q$ is subject to spectacular enhancement in the neighborhoods of the surface plasmon resonance (SPR) angle and of the second-harmonic surface plasmon resonance (SHSPR) angle for the metal-quartz interface (Simon *et al.*[3]). As the SPR is approached the linear field in the metal near $z = d$ increases, while that near $z = 0$ decreases; this leads to the structure observed in $\mathcal{R}_m$. The dip in $\mathcal{R}_m$ near the SHSPR has been explained qualitatively by Simon *et al.*[3] It should be noted that when those authors refer to SHG at, e.g., the rutile-metal interface, they apparently refer to both that due to the selvedge current at $z = 0$ and that due to the bulk metal near $z = 0$. Finally, we mention that the structure in $\mathcal{R}_q$ at angles just short of the SPR angle is due to changes in the nature of fields in the quartz. For incident angles in the rutile greater than 38.28° the incident field is evanescent in the quartz, while for incident angles in the rutile greater than 38.46° the second-harmonic field generated in the quartz is evanescent.

Comparing $\mathcal{R}_{qm}$ with the experimental data of Simon *et al.*[3] (Fig. 3), we see that the plot in Fig. 2 reproduces the peak and valley positions, but that the ratios of intensities at different angles, and the width of the main SPR peak, do not agree with the experimental results. We now turn to the SHG generated at the two selvedges; in Fig. 4 we plot $\mathcal{R}_0 = (2\pi N_1/cn_1^2)|A_0|^2$ and $\mathcal{R}_d = (2\pi N_1/cn_1^2)|A_d|^2$ for $a = 0$ (and $b = -1$) at both selvedges. The factor $\mathcal{R}_d$ goes through maxima at the SPR and SHSPR much like $\mathcal{R}_q$, but is generally much smaller than $\mathcal{R}_0$ since, except at the SPR, the linear field at $z = 0$ is much larger than that at $z = d$. The dip in $\mathcal{R}_0$ near the SPR angle is due to the decrease in the

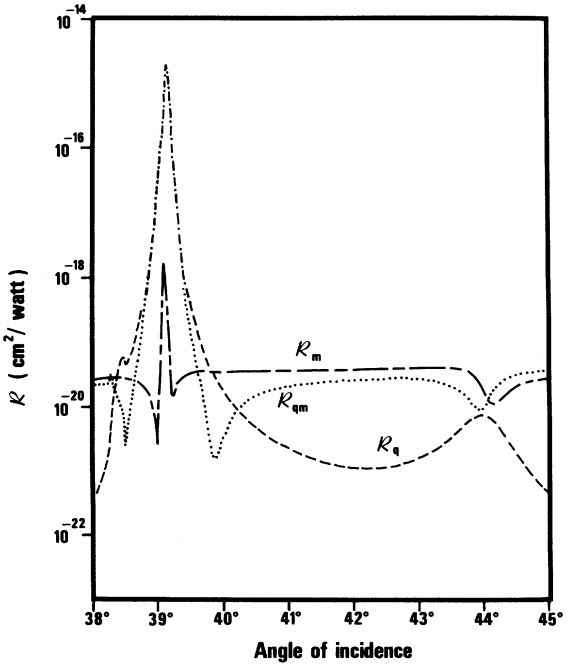

FIG. 2. The second-harmonic reflection coefficients $\mathcal{R}_q$, $\mathcal{R}_m$, and $\mathcal{R}_{qm}$.

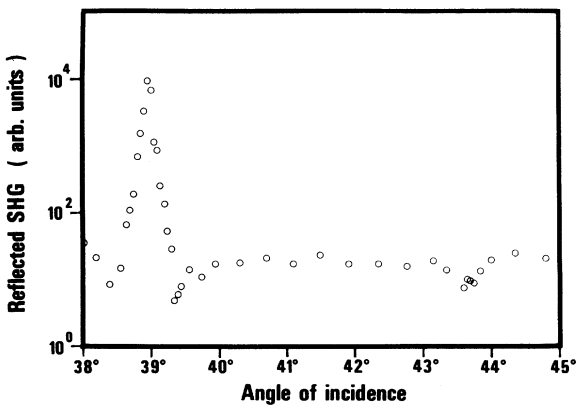

FIG. 3. The experimental reflected SHG of Simon *et al.* (Ref. 3).

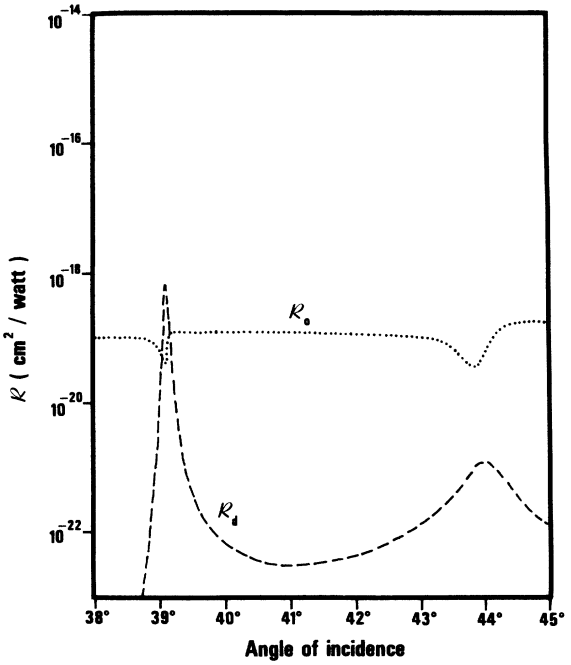

FIG. 4. The second-harmonic reflection coefficients $\mathcal{R}_0$ and $\mathcal{R}_d$.

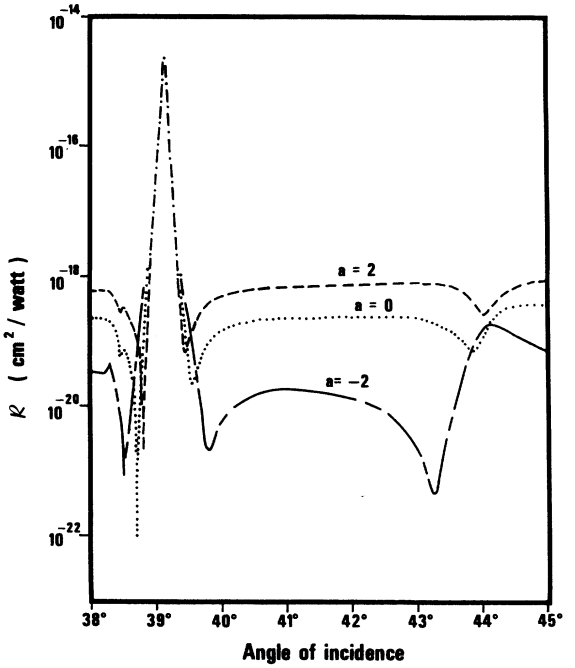

FIG. 5. Total second-harmonic reflection coefficients for $a=-2$, 0, and +2.

linear field at $z=0$ there, while the dip near the SHSPR angle is similar in nature to the corresponding dip in $\mathcal{R}_m$. We graph curves for total factors $\mathcal{R}$ in Fig. 5, for values of $a=-2$, 0, and +2. As is clear from Fig. 4, these curves are essentially independent of the parameters of the metal-quartz selvedge current. All of them give a much better agreement with the experimentally observed width of the SPR peak than does the $\mathcal{R}_{qm}$ factor of Fig. 2, indicating that Eqs. (3.3), with $b=-1$ and $|a|$ of order unity or smaller, predict the correct order of magnitude of the selvedge SHG at 1.06 μm. But a close fit with the experimental data is difficult to obtain, even if we do not worry about the dips right before and after the SPR peak, the first of which is very narrow and both of which, resulting from destructive interference between the second-harmonic fields generated in the quartz and at the rutile-metal selvedge, are very sensitive to the values chosen for the optical constants. Looking just at the ratio of the SPR peak intensity to the plateau intensity around $\theta=42°$, and at the shape of the dip at the SHSPR angle, we find that a better fit of the first with the experimental data is achieved by an $0 \leq a < -2$, while comparing the second with the data would lead one to adopt an $a \simeq 0$. Larger values of $|a|$ lead to plots of $\mathcal{R}$ with either one, the other, or both of these features in serious disagreement with the data. However, we note that the use of the first of these features to compare theory with experiment may not be advisable since the angular width of the peak is very narrow; i.e., the last order of magnitude of the peak height is confined to a width of only 0.1°. Thus, the observation of this feature is complicated by difficulties in resolution and beam divergence. Further, changing the values adopted for the optical constants by not unreasonable amounts leads to changes in the predicted peak height of about an order of magnitude. These problems somewhat obviate the advantage of using the quartz signal near the SPR as a calibration for the metal SHG.

To consider in more detail the dependence of SHG on a and b, we look at the more traditional SHG experiment[14] shown in Fig. 6; a beam is in-

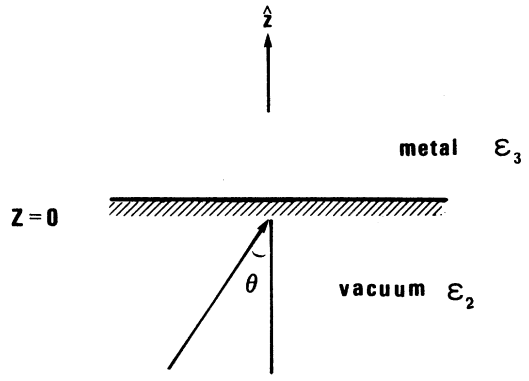

FIG. 6. The traditional SHG experiment.

cident on a vacuum-metal interface. The conversion factor for SHG is here given by $\mathcal{R} = (2\pi/c)|A_{vm}|^2$, where, using Eqs. (3.1)–(3.3), we find

$$A_{vm} = \frac{4\pi i \tilde{\Omega} k t_{23}^2}{W_2 \delta_3 + W_3} f, \tag{3.17}$$

where

$$f = \left(\frac{k^2}{\tilde{\omega}^2}\frac{\delta_3}{\epsilon_3} a - \frac{w_3 W_3}{\tilde{\omega}^2 \epsilon_3} b + \frac{1}{2}\right)\sigma \tag{3.18}$$

and

$$\sigma = \frac{e}{2m\omega^2}\left(\frac{\epsilon_3 - 1}{4\pi}\right), \tag{3.19}$$

and we here use the subscripts (2,3) to denote vacuum, metal. For $\epsilon_3 \simeq 1 - \omega_p^2/\omega^2 \ll -1$, Eq. (3.18) reduces to

$$f \simeq \left[\tfrac{1}{4}(k^2/\tilde{\omega}^2)a - b + \tfrac{1}{2}\right], \tag{3.20}$$

where the $\frac{1}{2}$ in both Eqs. (3.18) and (3.20) is from the SHG in the bulk metal. The relative sizes of the terms in Eq. (3.20) can easily be understood. Compare, for example, the contributions from Q^ϵ and Q^k: Q^k is driven to a much greater amplitude than Q^ϵ because the linear field in the metal has a much larger k component than a z component; however, Q^ϵ radiates much more efficiently than Q^k because the field from Q^ϵ radiated towards the bulk and reflected from it is essentially in phase with the field radiated directly to $-\infty$, while the corresponding fields from Q^k add essentially destructively. The net result is that the contributions are of the same order of magnitude, but that from Q^ϵ varies as k^2 compared to that from Q^k, since both the z component of the linear field, and the effectiveness with which Q^ϵ radiates, vary linearly with k for ϵ_3, $\delta_3 \ll -1$. The fact that both Q^ϵ and Q^k require a z component of the linear field [Eq. (3.3)] leads to the k in Eq. (3.17), and to the familiar $\sin^2\theta$ dependence for SHG at small θ. Similar arguments can be made to compare the SHG from Q^ϵ and from the bulk metal.

In Fig. 7 we use Eqs. (3.17) and (3.18) to plot the predicted $\mathcal{R}$'s from silver at 1.06 μm for $a = -1$, 0, and +1 (all with $b = -1$); the proportionately larger contribution of Q^ϵ at larger k is apparent. Since, comparing the $a = -1$ and $a = +1$ curves, $\mathcal{R}(40°)/\mathcal{R}(\text{peak})$ differs by about 40% of that ratio, careful measurements of reflected SHG in the traditional geometry of Fig. 6 (although they would be at admittedly low intensities), could be used to determine a. The advantage of such an experiment is that its interpretation and dependence on optical constants is rather more straightforward than an experiment such as that shown in Fig. 1. A number of experiments of this

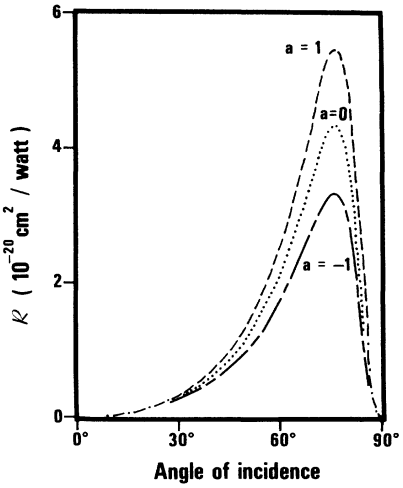

FIG. 7. SH reflection coefficients for the traditional experiment, for $a = -1$, 0, and +1.

type have been performed (see, e.g., Bloembergen *et al.*[14]), but at higher frequencies than considered here, where the theory is more complicated (see the discussion in Sec. IV).

Because of the form of Eqs. (3.18) and (3.20), we can expect even more dramatic dependence on a for experiments in which incident fields can be directed with larger wave vectors parallel to the surface. Consider the experiment shown in Fig. 8, where we assume there is no appreciable SHG in the glass. The conversion factor $\mathcal{R} = (2\pi N_1/cn_1^2)|A|^2$ is given by

$$A = (e^{iw_2 d} t_{12} m)^2 T_{21} e^{iW_2 d} M A_{vm}, \tag{3.21}$$

where m and M are given by Eqs. (3.9) and A_{vm}

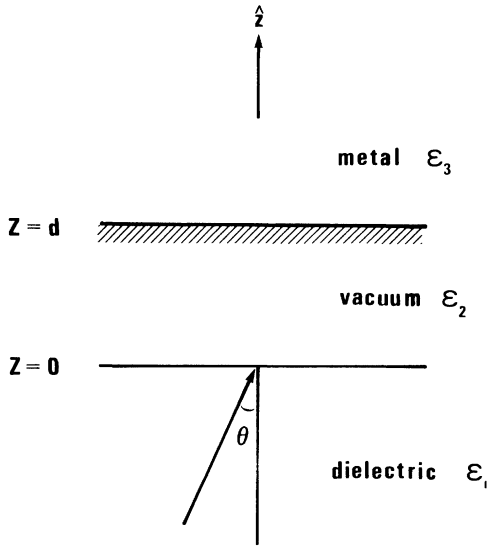

FIG. 8. SHG experiment to reach large values of k.

by Eq. (3.17). Look first at the limit $d=0$; if the dielectric is rutile we then obtain SHG from the bulk metal and from the rutile-metal selvedge present in Fig. 1. We plot the $\mathcal{R}$'s which result for silver metal and an incident field at 1.06 μm in Fig. 9. The predicted signal is both more intense than that shown in Fig. 7, and there is a larger difference between the curves with different values of a, the peaks in fact occurring at different angles of incidence. Because of the comparatively large values of $\mathcal{R}$, this experiment should be feasible; in fact, it would be performed by a simple extension of the experiment of Simon *et al.*[3] to measure the SHG at higher incident angles, since the results of that latter experiment are essentially independent of the parameters characterizing the quartz-metal selvedge. In Fig. 10 we plot the values of $\mathcal{R}$ predicted for the geometry shown in Fig. 1, with higher incident angles than shown in Fig. 5. Looking at Figs. 5 and 10 from, say 40° to 85°, it is clear that different values of a lead to quite different curve shapes. Observations of the SHG through these angles, rather than around the SPR, would be more appropriate in an attempt to determine the value of a for the rutile-metal selvedge, as would observations of the SHG from a simple rutile-metal interface. The dip in the curve for $a=-2$ in Fig. 10 occurs when the second-harmonic field from the rutile-metal selvedge, which is rapidly growing with increasing angle, reaches the same magnitude as the field from the quartz and metal, and adds *destructively*; it has nothing to do with plasmon enhancement [cf. Eq. (3.20), which, however, strictly applies to only the much simpler geometry previously discussed]. A similar dip is predicted for the experiment of Fig. 9, if a is sufficiently large in

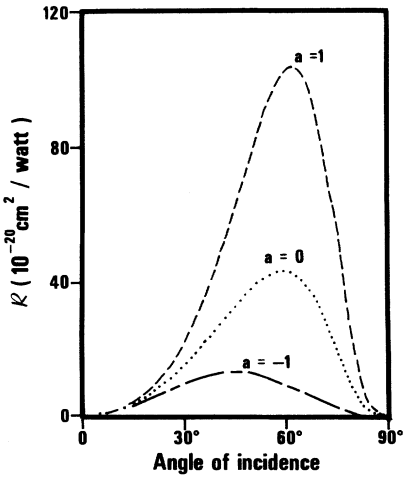

FIG. 9. SH reflection coefficients for a rutile-metal surface.

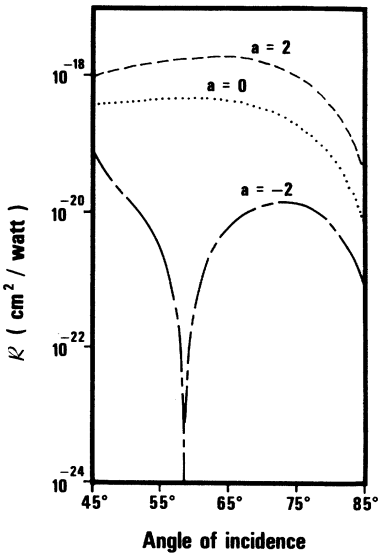

FIG. 10. SH reflection coefficients for the experiment of Simon *et al.* (Ref. 3) at larger angles.

magnitude and negative. If a is in this range, its determination by these experiments would be very simple indeed.

However, the fundamental surface of interest in the SHG from metals is the metal-vacuum surface, rather than a metal-dielectric surface. An enhanced-SHG experiment particularly sensitive to the metal-vacuum selvedge would be useful in determining the coefficients a and b of interest. Experiments similar to that shown in Fig. 1, but with medium 3 being air, have been performed; however, the plots in Figs. 2 and 4 suggest, and calculations confirm, that even within 0.1° of the SPR, the SHG from the rutile-metal interface is not insignificant, and that for the particular geometry of Fig. 1 the bulk metal SHG seems to dominate even at the SPR. As a simple alternative we suggest (and see Ref. 10) the experiment shown in Fig. 8 with a gap of about $d=1$ μm (Otto geometry). In Fig. 11 we plot predicted $\mathcal{R}$'s for a glass with $n_1=N_1=1.40$, and silver metal subject to an incident field at 1.06 μm. The enhancement (cf. Fig. 7) is associated with the SPR at the vacuum-metal interface and with the vacuum light line, but is complicated by the fact that the gap is on the order of a wavelength; at larger gaps well-defined SPR peaks are predicted at higher angles, but they are smaller in intensity and very narrow. Enhancement due to the SHSPR at the vacuum-metal interface also occurs, but at a much lower intensity, since at those greater values of k the evanescent linear fields in the gap are more confined to the glass-vacuum surface, and much weaker SHG results. The predicted values of $\mathcal{R}$

shown in Fig. 11 are fairly sensitive to d, but that distance could be determined from linear experiments; the predictions for smaller, but perhaps not experimentally reproducible gaps are also dramatic. With respect to determining a, we note that the curves, when normalized to their peak intensities, are very similar in shape, the ratios $\mathcal{R}(42°)/\mathcal{R}(\text{peak})$ differing by only about 7% between $a=-1$ and $+1$. However, the signal could be compared with a reference crystal, or the experiment could be made self-calibrating by adding a crystal without inversion symmetry above the metal. If the SPR angle for the vacuum-metal interface were fairly far removed from the SPR angle for the crystal-metal interface, the SHG at the crystal-metal selvedge could be neglected, and the background SHG from the crystal would vary little with incident angle and could be used to calibrate the experiment. In any case, the intensities predicted in Fig. 11 are much higher than in the traditional (Fig. 5) experiment, and since the peaks are fairly broad and the difference between curves of different a fairly substantial, we feel experiments of this sort could be of some aid in understanding SHG at the metal-vacuum surface. The enhancement factors present in more complicated ATR experiments, which could also be used to probe the SHG at a metal-vacuum surface, are discussed in a general way by Chen and Burstein.[10]

IV. CONCLUDING REMARKS

To summarize, in this work we have discussed the SHG from metal surfaces within the hydrodynamic model by dividing the metal into "bulk" and "selvedge" regions. We have shown that Rudnick and Stern's phenomenological parameter b is predicted to be -1; the parameter a may show a resonance (though perhaps either highly damped or

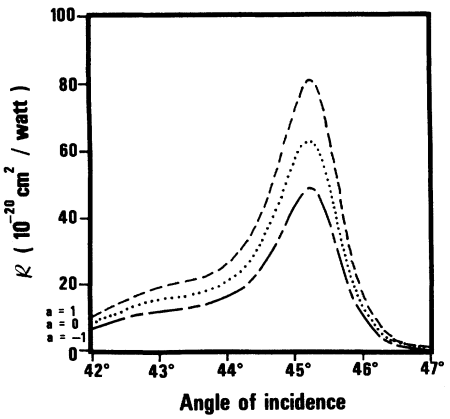

FIG. 11. SH reflection coefficients for the experiment of Fig. 8 (see text).

near the bulk plasma frequency), associated with the effective plasma frequency of the second-harmonic surface currents. Analyzing the plasmon-enhanced-SHG results of Simon *et al.*[3] at 1.06 μm we showed that, taking into account the difficulty of the experiments, the results are fairly well described by $b=-1$ and $|a|$ of order unity or less. Using the fact that SHG experiments are proportionately more sensitive to a at larger values of $(k/\tilde{\omega})$, we suggested an SHG experiment on a rutile-metal interface, and an enhanced-SHG experiment on a vacuum-metal interface, that should aid in determining that parameter.

Throughout we have neglected surface roughness, which can be expected to change the value of b (and of course that of a) from the smooth surface value $(b=-1)$ corresponding to a surface which perfectly reflects electrons impinging on it from the bulk. Rudnick and Stern[7] have discussed the effect surface roughness can be expected to have on b, and there has been much recent work on the effect of surface roughness on the *linear* properties of an interface (Maradudin and Mills[39]). Nonetheless, a theory for the effect of surface roughness on SHG is still lacking. We note, though, that the experiments we have suggested to determine a are, in fact, sensitive to the ratio $a/(-b+\frac{1}{2})$ [Eq. (3.20)]. Thus, even if b is not known, some information can be gained. Ideally, of course, b could be determined from experimental results at low $(k/\tilde{\omega})$ where the contribution from Q^{ϵ} is negligible; this would require a reference crystal, or a self-calibrating experiment of the type discussed at the end of Sec. III.

Since the application of the hydrodynamic theory indicates the possibility of a resonance in a, we believe that a closer look at the RPA expressions for SHG, and some actual calculations, would be justified. As pointed out by a number of authors (cf., e.g., Refs. 7, 15–17, 21), the hydrodynamic theory cannot be relied on to give a quantitative description of electron behavior near surfaces. A more satisfactory approach would be to calculate a and b following the analysis of, e.g., Bower[16]; this would begin with the solution for the linear fields recently presented by Feibelman[40] (and see Refs. 41 and 42). Only within such a more realistic theory could a calculation of a resonance in a be considered physically meaningful; the effect of adsorbed species on a and b would be particularly interesting. Of course, more detailed calculations of even the bulk SHG would be required for a serious experimental and theoretical investigation into the dispersion of a (and b), since it is only in the infrared, where interband transitions can be neglected, that the hydrodynamic model forms a good approximation of elec-

tron dynamics even in the bulk of metals such as silver.[43] Another very interesting topic for future theoretical and experimental investigation is the effect of the presence of multipole surface plasmons[18,20,25] on the SHG at metal surfaces. In addition to the possibility of the nonlinear source terms coupling the plasmon branches, where the strength of the coupling would depend on the surface charge distributions of both branches, it is possible that the SHG from a multipole surface plasmon would be considerably stronger than that from the usual monopole surface plasmon, since the charge density in the former undergoes drastic variation in the selvedge region.[9]

ACKNOWLEDGMENTS

We wish to acknowledge support of the National Science and Engineering Research Council, and helpful conversations with Y. R. Shen and F. DeMartini. One of us (J.E.S.) acknowledges financial assistance from the Killam Program of the Canada Council.

APPENDIX

We here present an extension of the theory of transfer matrices[44] which simplifies the solution of the driven Maxwell equations for systems of the geometry shown in Fig. 1, with an arbitrary number of layers.

For the source of polarization

$$\vec{P}(\vec{r})=\vec{P}(z)e^{i(k_x x+k_y y)} \tag{A1}$$

(for the problems in the text $\vec{k}$ would be replaced by $\vec{K}$, etc.) a particular solution of the Maxwell equations,

$$\begin{aligned}&\vec{\nabla}\cdot(\epsilon_i\vec{E})=-4\pi\vec{\nabla}\cdot\vec{P}\,,\\&\vec{\nabla}\cdot\vec{B}=0\,,\\&\vec{\nabla}\times\vec{B}+i\tilde{\omega}\epsilon_i\vec{E}=-4\pi i\tilde{\omega}\vec{P}\,,\\&\vec{\nabla}\times\vec{E}-i\tilde{\omega}\vec{B}=0\,,\end{aligned} \tag{A2}$$

is given by

$$\vec{E}(z)=\int\tilde{G}(z-z')\cdot\vec{P}(z')dz'\,, \tag{A3}$$

where

$$\tilde{G}(z-z')=2\pi i\tilde{\omega}^2w_i^{-1}[(\hat{s}\hat{s}+\hat{p}_{i+}\hat{p}_{i+})e^{iw_i(z-z')}\Theta(z-z')+(\hat{s}\hat{s}+\hat{p}_{i-}\hat{p}_{i-})e^{-iw_i(z-z')}\Theta(z'-z)]-4\pi\epsilon_i^{-1}\hat{z}\hat{z}\delta(z-z')\,, \tag{A4}$$

$\hat{s}=\hat{k}\times\hat{z}$, $\Theta(z)$ is the usual step function, $d\Theta(z)/dz=\delta(z)$, and w_i and $\hat{p}_{i\pm}$ are given by Eqs. (3.10) and (3.16). The Green function (A4) can be derived simply and directly from Eqs. (A2), without the introduction of potentials or auxiliary fields, following Sipe.[25] The p-polarized solution of the *homogeneous* form of Eqs. (A2), is, of course,

$$\vec{E}(z)=E_{i+}\hat{p}_{i+}e^{iw_iz}+E_{i-}\hat{p}_{i-}e^{-iw_iz}\,, \tag{A5}$$

which we write in the form of a vector,

$$e_i(z)=\begin{bmatrix}E_{i+}e^{iw_iz}\\E_{i-}e^{-iw_iz}\end{bmatrix}\,. \tag{A6}$$

The reflection and transmission of a field (A5) at a discontinuity in the dielectric constant is easily handled by applying transfer matrices to the vector (A6).

Now return to Eqs. (A1) and (A2), and consider a source $\vec{P}(z)=\vec{P}(z_0)\delta(z-z_0)$. We are interested in the p-polarized fields which are radiated, after perhaps reflections and transmissions at interfaces, towards $z=\pm\infty$; the field resulting from the δ function in Eq. (A4) is not of interest. Clearly for $z\neq z_0$ the total field is of the form (A5) and (A6), but with different coefficients $E_{i\pm}$ as $z\gtrless z_0$. Combining the particular solution (A3) with an arbitrary homogeneous solution, we find

$$e_i(z_0^+)=s_i(z_0)+e_i(z_0^-)\,, \tag{A7}$$

where

$$s_i(z_0)=\begin{bmatrix}s_{i+}(z_0)\\-s_{i-}(z_0)\end{bmatrix} \tag{A8}$$

and

$$s_{i\pm}(z_0)=2\pi i\tilde{\omega}^2w_i^{-1}\hat{p}_{i\pm}\cdot\vec{P}(z_0)\,. \tag{A9}$$

Using Eq. (A7) and the appropriate transfer matrices, along with the boundary condition of outgoing waves at $z=\pm\infty$, the radiated fields are easily found by straightforward algebra; the total fields radiated from the source (A1) are then found by superposition.

*Killam Research Associate (1979–80).
†Permanent Address: The University of Tokushima, Tokushima, Japan.

[1]H. J. Simon, D. E. Mitchell, and J. G. Watson, Phys. Rev. Lett. 33, 1531 (1974).
[2]H. J. Simon, D. E. Mitchell, and J. G. Watson, Opt. Commun. 13, 294 (1975).
[3]H. J. Simon, R. E. Brenner, and J. G. Rako, Opt.

Commun. 23, 245 (1977).
[4]M. Fukui, V. C. Y. So, J. E. Sipe, and G. I. Stegeman, J. Phys. Chem. Solids 40, 523 (1979).
[5]D. L. Mills, Solid State Commun. 24, 669 (1977).
[6]M. Fukui, V. So, and G. Stegeman, Phys. Rev. B 18, 2484 (1978).
[7]J. Rudnick and E. A. Stern, Phys. Rev. B 4, 4274 (1971).
[8]F. Brown and M. Matsuoka, Phys. Rev. 185, 985 (1969).
[9]A. Eguiluz and J. J. Quinn, Phys. Rev. B 14, 1347 (1976).
[10]Y. J. Chen and E. Burstein, Nuovo Cimento 39B, 807 (1977).
[11]S. S. Jha, Phys. Rev. 140, A2020 (1965).
[12]S. S. Jha, Phys. Rev. Lett. 15, 412 (1965).
[13]S. S. Jha, Phys. Rev. 145, 500 (1966).
[14]N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, Phys. Rev. 174, 813 (1968).
[15]C. S. Wang, J. M. Chen, and J. R. Bower, Opt. Commun. 8, 275 (1973).
[16]J. R. Bower, Phys. Rev. B 14, 2427 (1976).
[17]A. Griffin and H. Kranz, Phys. Rev. B 15, 5068 (1977).
[18]A. Eguiluz, S. C. Ying, and J. J. Quinn, Phys. Rev. B 11, 2118 (1975).
[19]J. E. Inglesfield and E. Wikborg, Solid State Commun. 14, 661 (1974).
[20]J. E. Inglesfield and E. Wikborg, J. Phys. F 5, 1706 (1975).
[21]J. Rudnick and E. A. Stern, in *Polaritons*, edited by E. Burstein and F. DeMartini (Plenum, New York, 1974).
[22]Solid State Commun., in press.
[23]J. Crowell and R. H. Ritchie, J. Opt. Soc. Am. 60, 794 (1970).
[24]A. J. Bennett, Phys. Rev. B 1, 203 (1970).
[25]J. E. Sipe, Surf. Sci. 84, 75 (1979).
[26]T. Lopez-Rios, F. Abelès, and G. Vuye, J. Phys. (Paris) 39, 645 (1978).
[27]E. N. Economou and K. L. Ngai, Adv. Chem. Phys. 27, 265 (1974).
[28]F. Zernike and J. E. Midwinter, *Applied Nonlinear Optics* (Wiley, New York, 1973).
[29]*American Institute of Physics Handbook* (McGraw-Hill, New York, 1972).
[30]A. G. Mathewson, H. Aronsson, and L. G. Bernland, J. Phys. F 2, L39 (1972).
[31]M. Dujardin and M. Thèye, J. Phys. Chem. Solids 32, 2033 (1971).
[32]T. Hollstein, U. Kreibig, and F. Leis, Phys. Status Solidi 82b, 545 (1977).
[33]*Handbook of Lasers* (Chemical Rubber, Cleveland, 1971).
[34]M. Otter, Z. Phys. 161, 163 (1961).
[35]P. B. Johnson and R. W. Christy, Phys. Rev. B 6, 4370 (1972).
[36]W. H. Weber and S. L. McCarthy, Phys. Rev. B 12, 5643 (1975)
[37]J. R. Adams, J. R. Zeidler, and N. M. Bashara, Opt. Commun. 15, 115 (1975).
[38]J. Lafait, F. Abelès, M. L. Thèye, and G. Vuye, J. Phys. F 8, 1597 (1978).
[39]A. A. Maradudin and D. L. Mills, Phys. Rev. B 11, 1392 (1975).
[40]P. J. Feibelman, Phys. Rev. B 12, 1319 (1975); 14, 762 (1976).
[41]A. Bagchi, Phys. Rev. B 15, 3060 (1977).
[42]G. Mukhopadhyay and S. Lundqvist, Phys. Scr. 17, 69 (1978).
[43]S. S. Jha and C. S. Warke, Phys. Rev. 153, 751 (1967).
[44]O. S. Heavens, *Optical Properties of Thin Solid Films* (Dover, New York, 1965).