Second-harmonic generation on simple metal surfaces: A quantum-mechanical point of view

C. D. Hu

Department of Physics, National Taiwan University, Taipei, Taiwan 107 64, Republic of China

(Received 9 March 1989)

Second-harmonic generation was studied with quantum mechanics. The explicit form of the nonlinear current was derived for the high-frequency regime. The efficiency of second-harmonic generation was also discussed. A nonlinear surface-plasmon mode was found.

I. INTRODUCTION

The second-harmonic generation of light in media was first found by Franken and co-workers.¹ The works by Armstrong *et al.*² and of Bloembergen and Pershan³ started serious theoretical study on second-harmonic generation on surfaces. Early works, summarized in a paper by Bloembergen, Chang, Jha, and Lee⁴ and updated by Shen⁵ (who also included many related topics), were analogous to classical electrodynamics in linear media. More recently, a hydrodynamic model has been developed,⁶⁻⁸ which has the virtue of simplicity and indeed provides more insight to this problem; most importantly, it gives a direct link between second-harmonic generation and media properties such as the electron density and surface-potential barrier. The problem of retardation was also studied within this model.⁹ The first attempt of applying quantum mechanics and taking into account surface electronic behavior was given in an important paper by Rudnick and Stern (RS).¹⁰ Now time-dependent density-functional theory^{11,12} was also used to treat this problem in the non-(time-)retarded case. Some unexpected results were obtained. For example, parameter a defined by RS is far from unity,^{8,11,12} the original estimation.

The problem at hand is still an open field. On the one hand, second-harmonic generation is sensitive to surface preparation for metals and semiconductors;¹³ therefore surface needs to be studied microscopically with use of quantum mechanics. On the other hand, complex calculations could blur the physical picture which, in our opinion, is still not well understood. In this paper, we apply quantum mechanics to second-harmonic generation. For tractability, we studied a simple metal surface which can be described adequately by jellium model. General formalism is given in Sec. II. Nonlinear current density is analyzed in Sec. III. In Sec. IV, we present correct boundary conditions and evaluate efficiency of second-harmonic generation. Section V is a summary.

II. GENERAL FORMALISM

We give a brief description of our system. The metal occupies the $z < 0$ half-space. It will be described by a jellium model. A beam of light with angular frequency ω and momentum $\hbar\mathbf{k}$ incidents on the flat metal surface at $z = 0$. It is to be denoted by its vector potential $\mathbf{A}_0(\mathbf{r}, \omega)$.

We chose the incident plane to be the x - z plane so that $\mathbf{k} = (q, 0, k_0)$ and

$$\mathbf{A}_0(\mathbf{r}, \omega) = \mathbf{A}_{0,\omega} e^{i(qx - k_0 z)} . \quad (1)$$

Of course

$$q^2 + k_0^2 = \omega^2/c^2 , \quad (2)$$

where c is the light speed. The incident light can generate fundamental and second-harmonic modes. The reflected and transmitted waves for the former are represented by $\mathbf{A}_R(\mathbf{r}, \omega)$ and $\mathbf{A}_1(\mathbf{r}, \omega)$, respectively, while the latter is represented by $\mathbf{A}_2(\mathbf{r}, 2\omega)$. Here we used the gauge

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} , \quad (3)$$

$$\mathbf{B} = \nabla \times \mathbf{A} , \quad (4)$$

and

$$\nabla \phi = 0 , \quad (5)$$

where ϕ is the scalar potential. We have to note here that the quantities in Eqs. (3)–(5) are all frequency dependent. In other words they are ac fields. The potential from passive positive ions in metal is not included. What are included is the field of light and that from screening electrons. The current and density induced by ac fields should satisfy continuity equation. Far away from the surface we have

$$\mathbf{A}_1(\mathbf{r}, \omega) = \mathbf{A}_{1,\omega} e^{i(qx - k_1 z)} \quad (6)$$

and

$$\mathbf{A}_R(\mathbf{r}, \omega) = \mathbf{A}_{R,\omega} e^{i(qx + k_0 z)} , \quad (7)$$

where

$$k_1 = [\omega^2 \epsilon(\omega)/c^2 - q^2]^{1/2} \quad (8)$$

and $\epsilon(\omega)$ is the dielectric function. We assumed the metal surface to be flat. Hence momenta parallel to surface are conserved. The x dependence of $\mathbf{A}_2(\mathbf{r}, 2\omega)$ is of the form e^{2iqx} . The z dependence is more complicated. As we shall see

$$\mathbf{A}_2(\mathbf{r}, 2\omega) = \mathbf{a}_r(2\omega)e^{2i(qx + k_0z)} \quad \text{in vacuum,} \quad (9)$$

$$\begin{aligned} \mathbf{A}_2(\mathbf{r}, 2\omega) &= \mathbf{a}_b(2\omega)e^{2i(qx - k_2z)} \\ &+ \mathbf{a}'_b(2\omega)e^{2i(qx - k_1z)} \quad \text{in bulk,} \end{aligned} \quad (10)$$

where

$$k_2 = \left[\frac{\omega^2}{c^2} \epsilon(2\omega) - q^2 \right]^{1/2} \quad (11)$$

and $\mathbf{a}_r$, $\mathbf{a}_b$, and $\mathbf{a}'_b$ are proportional to $|\mathbf{A}_0|^2$. Finally, we remarked that although we said the metal surface was at $z=0$, actually there is a fuzzy region at $|z| \sim 0$ where surface properties dominate. Let us assume the surface healing length to be l . Typically, the length of l is a few Å and much less than the wavelength of incident light. Thus we can divide our space into three regions: $z > l/2$, $l/2 > z > -l/2$, and $z < -l/2$, where, respectively, vacuum, surface, and bulk properties dominate. Equations (7) and (9) are valid in the first region and Eqs. (6) and (10) are valid in the third region. In the surface region, both $\mathbf{A}_1(\mathbf{r}, \omega)$ and $\mathbf{A}_2(\mathbf{r}, 2\omega)$ become very complicated and can only be described correctly with microscopic theory.

When a light beam illuminates nonlinear media, it generates linear and nonlinear currents which in turn give off linear and nonlinear radiation. One had better use quantum mechanics to study the first-half process though Chizmeshya and Zaremba⁸ had shown that the hydrodynamic-model results agreed reasonably well with those of time-dependent density-functional method.¹¹ For the second-half process, Maxwell equations are adequate. We have

$$\nabla \times \nabla \times \mathbf{A} + \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = \frac{4\pi}{c} \mathbf{j}. \quad (12)$$

For the second harmonic part,

$$\begin{aligned} \nabla \times \nabla \times \mathbf{A}_2(\mathbf{r}, 2\omega) - \frac{4\omega^2}{c^2} \mathbf{A}_2(\mathbf{r}, 2\omega) \\ = \frac{4\pi}{c} [\mathbf{j}_{2L}(\mathbf{r}, 2\omega) + \mathbf{j}_{NL}(\mathbf{r}, 2\omega)], \end{aligned} \quad (13)$$

where

$$\mathbf{j}_{2L}(\mathbf{r}, 2\omega) = \frac{2i\omega}{c} \int \sigma(\mathbf{r}, \mathbf{r}', 2\omega) \mathbf{A}_2(\mathbf{r}', 2\omega) d^3\mathbf{r}'. \quad (14)$$

Here $\sigma(\mathbf{r}, \mathbf{r}', 2\omega)$ is the conductivity tensor. It is the linear response function of current. $\mathbf{j}_{NL}$, on the other hand, is related to both linear and nonlinear properties of the media. If the form of $\mathbf{j}_{NL}$ is known (that of σ is known), then we can solve for $\mathbf{A}_2(\mathbf{r}, 2\omega)$ using Eq. (13).

$$\langle \mathbf{J}(\mathbf{r}, t) \rangle = \frac{i}{\hbar} \int_{-\infty}^t dt' \langle [H_{\text{int}}^{(H)}(t'), \mathbf{J}^{(H)}(\mathbf{r}, t)] \rangle - \frac{1}{\hbar^2} \int_{-\infty}^t dt' \int_{-\infty}^{t'} dt'' \langle [H_{\text{int}}^{(H)}(t''), [H_{\text{int}}^{(H)}(t'), \mathbf{J}^{(H)}(\mathbf{r}, t)]] \rangle, \quad (19)$$

where superscripts (H) denote the operators are in Heisenberg picture of H_0 , $\langle \cdots \rangle$ means expectation value over the ground-state wave function, and

$$\mathbf{J}(\mathbf{r}) = \frac{e}{2m} \left[\psi^\dagger(\mathbf{r}) \mathbf{p} \psi(\mathbf{r}) - [\mathbf{p} \psi^\dagger(\mathbf{r})] \psi(\mathbf{r}) - \frac{2e}{c} \psi^\dagger(\mathbf{r}) \mathbf{A}(\mathbf{r}, t) \psi(\mathbf{r}) \right] \quad (20)$$

To find $\mathbf{j}_{NL}(\mathbf{r}, 2\omega)$ and $\mathbf{j}_{2L}(\mathbf{r}, 2\omega)$, i.e., to study the first-half process, we started with the Hamiltonian

$$H = H_{e-i} + H_{e-e} + H_{e-p}, \quad (15)$$

where H_{e-i} is the interaction energy between electrons and ions which were smeared out to form passive uniform positive background in jellium model, H_{e-e} is that between electrons, and H_{e-p} is the kinetic energy of electrons and their interaction with light waves:

$$H_{e-p} = \frac{1}{2m} \int \psi^\dagger(\mathbf{r}) \left[\mathbf{p} - \frac{e}{c} \mathbf{A}_0(\mathbf{r}, t) \right]^2 \psi(\mathbf{r}) d^3r. \quad (16)$$

Here e and m are electron charge and mass, respectively, and $\psi^\dagger(\mathbf{r})$ and $\psi(\mathbf{r})$ are creation and annihilation operators of electrons. In principle, we can use the density-functional method¹⁴ to solve for the ground state of metal surfaces¹⁵ and treat the fields of incident light as perturbation. However, the matter becomes complicated because each perturbed term involves infinite numbers of terms with different orders of H_{e-e} . One approximation is to replace $\mathbf{A}_0(\mathbf{r}, t)$ in Eq. (16) by $\mathbf{A}(\mathbf{r}, t)$, the screened vector potential which has linear and nonlinear parts and treat electrons as independent particles. This has the same spirit as the Bohn-Pines random-phase approximation (RPA).¹⁶ In RPA, each electron feels a screened perturbation. The screening field is the average of all other electrons' fields. Aside from this mean field, electrons do not interact with each other except for obeying Pauli's principle. We followed this approach. This way, what we shall calculate is the analogue of the irreducible response function in RPA:

$$H \approx \int \psi^\dagger(\mathbf{r}) \frac{\mathbf{p}^2}{2m} \psi(\mathbf{r}) d^3r + \int V_{KS}(\mathbf{r}) \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) d^3r + H_{\text{int}} \quad (17)$$

where

$$\begin{aligned} H_{\text{int}} &= \frac{1}{2m} \int \psi^\dagger(\mathbf{r}) \left[-\frac{e}{c} \mathbf{p} \cdot \mathbf{A}(\mathbf{r}, t) - \frac{e}{c} \mathbf{A}(\mathbf{r}, t) \cdot \mathbf{p} \right] \psi(\mathbf{r}) d^3r \\ &+ \frac{1}{2m} \int \psi^\dagger(\mathbf{r}) \frac{e^2}{c^2} \mathbf{A}^2(\mathbf{r}, t) \psi(\mathbf{r}) d^3r, \end{aligned} \quad (18)$$

where $V_{KS}(\mathbf{r})$ is the Kohn-Sham potential under the local-density approximation.¹⁷ Now the first two terms of Eq. (17) combined gives the ground-state Hamiltonian H_0 while H_{int} is perturbation. The linear and nonlinear current can be found with time-dependent perturbation theory:¹⁸

is the current operator. We define

$$H_1(\mathbf{r}, \omega) \equiv \frac{1}{2m} \psi^\dagger(\mathbf{r}) \left[-\frac{e}{c} \mathbf{p} \cdot \mathbf{A}_1(\mathbf{r}, \omega) - \frac{e}{c} \mathbf{A}_1(\mathbf{r}, \omega) \cdot \mathbf{p} \right] \psi(\mathbf{r}), \quad (21a)$$

$$H_2(\mathbf{r}, 2\omega) \equiv \frac{1}{2m} \psi^\dagger(\mathbf{r}) \left[-\frac{e}{c} \mathbf{p} \cdot \mathbf{A}_2(\mathbf{r}, 2\omega) - \frac{e}{c} \mathbf{A}_2(\mathbf{r}, 2\omega) \cdot \mathbf{p} \right] \psi(\mathbf{r}), \quad (21b)$$

$$H_3(\mathbf{r}, 2\omega) \equiv \frac{1}{2m} \psi^\dagger(\mathbf{r}) \frac{e^2}{c^2} [\mathbf{A}_1(\mathbf{r}, \omega)]^2 \psi(\mathbf{r}), \quad (22)$$

$$\mathbf{J}_p(\mathbf{r}) \equiv \frac{e}{2m} \{ \psi^\dagger(\mathbf{r}) \mathbf{p} \psi(\mathbf{r}) - [\mathbf{p} \psi^\dagger(\mathbf{r})] \psi(\mathbf{r}) \}, \quad (23)$$

$$\mathbf{J}_A(\mathbf{r}, \omega) \equiv -\frac{e^2}{mc} \psi^\dagger(\mathbf{r}) \mathbf{A}_1(\mathbf{r}, \omega) \psi(\mathbf{r}). \quad (24)$$

Using Eq. (19) and creating the Fourier transform, we found

$$\langle \mathbf{J}(\mathbf{r}, \omega) \rangle = \mathbf{j}_1(\mathbf{r}, \omega) + \mathbf{j}_{2L}(\mathbf{r}, 2\omega) + \mathbf{j}_{NL}(\mathbf{r}, 2\omega), \quad (25)$$

where

$$\mathbf{j}_1(\mathbf{r}, \omega) = -\int d^3r' \left[\frac{[H_1(\mathbf{r}', \omega)]_{0n} [\mathbf{J}_p(\mathbf{r})]_{n0}}{\hbar\omega + E_{n0} + i\eta} - \frac{[\mathbf{J}_p(\mathbf{r})]_{0n} [H_1(\mathbf{r}', \omega)]_{n0}}{\hbar\omega - E_{n0} + i\eta} \right] - \frac{e^2}{mc} \mathbf{A}_1(\mathbf{r}, \omega) n_0(\mathbf{r}), \quad (26)$$

$$\mathbf{j}_{2L}(\mathbf{r}, \omega) = -\int d^3r' \left[\frac{[H_2(\mathbf{r}', 2\omega)]_{0n} [\mathbf{J}_p(\mathbf{r})]_{n0}}{2\hbar\omega + E_{n0} + i\eta} - \frac{[\mathbf{J}_p(\mathbf{r})]_{0n} [H_2(\mathbf{r}', 2\omega)]_{n0}}{2\hbar\omega - E_{n0} + i\eta} \right] - \frac{e^2}{mc} \mathbf{A}_2(\mathbf{r}, 2\omega) n_0(\mathbf{r}), \quad (27)$$

$$\begin{aligned} \mathbf{j}_{NL}(\mathbf{r}, 2\omega) = & -\int d^3r' \left[\frac{[H_1(\mathbf{r}', \omega)]_{0n} [\mathbf{J}_A(\mathbf{r}, \omega)]_{n0}}{\hbar\omega + E_{n0} + i\eta} - \frac{[\mathbf{J}_A(\mathbf{r}, \omega)]_{0n} [H_1(\mathbf{r}', \omega)]_{n0}}{\hbar\omega - E_{n0} + i\eta} \right] \\ & -\int d^3r' \left[\frac{[H_3(\mathbf{r}', 2\omega)]_{0n} [\mathbf{J}_p(\mathbf{r})]_{n0}}{2\hbar\omega + E_{n0} + i\eta} - \frac{[\mathbf{J}_p(\mathbf{r})]_{0n} [H_3(\mathbf{r}', 2\omega)]_{n0}}{2\hbar\omega - E_{n0} + i\eta} \right] \\ & + \frac{1}{2} \int d^3r' \int d^3r'' \left[\frac{[H_1(\mathbf{r}'', \omega)]_{0n} [H_1(\mathbf{r}', \omega)]_{nm} [\mathbf{J}_p(\mathbf{r})]_{m0}}{(\hbar\omega - E_{nm} + i\eta)(\hbar\omega + E_{n0} + i\eta)} - \frac{[H_1(\mathbf{r}'', \omega)]_{0n} [\mathbf{J}_p(\mathbf{r})]_{nm} [H_1(\mathbf{r}', \omega)]_{m0}}{(\hbar\omega - E_{m0} + i\eta)(\hbar\omega + E_{n0} + i\eta)} \right. \\ & \left. - \frac{[H_1(\mathbf{r}', \omega)]_{0n} [\mathbf{J}_p(\mathbf{r})]_{nm} [H_1(\mathbf{r}'', \omega)]_{m0}}{(\hbar\omega - E_{m0} + i\eta)(\hbar\omega + E_{n0} + i\eta)} + \frac{[\mathbf{J}_p(\mathbf{r})]_{0n} [H_1(\mathbf{r}', \omega)]_{nm} [H_1(\mathbf{r}'', \omega)]_{m0}}{(\hbar\omega - E_{nm} + i\eta)(\hbar\omega - E_{m0} + i\eta)} \right], \quad (28) \end{aligned}$$

where subscripts 0 and n (m) denote the ground state and intermediate states, respectively. E_{mn} is their energy difference. η is a positive infinitesimal number. The sign of $i\eta$ has to do with the fact that $\langle \mathbf{J}(\mathbf{r}, \omega) \rangle$ is a retarded function¹⁸. Finally $n_0(\mathbf{r})$ is the ground-state electron density:

$$n_0(\mathbf{r}) = \langle \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) \rangle. \quad (29)$$

In principle, once the expressions of $\mathbf{j}_{2L}(\mathbf{r}, 2\omega)$ and $\mathbf{j}_{NL}(\mathbf{r}, 2\omega)$ were found, they can be substituted into Eq. (13) to solve for $\mathbf{A}_2(\mathbf{r}, 2\omega)$ and hence other physical quantities such as reflectance. However, due to the presence of surfaces, Eqs. (27) and (28) become very complicated except for in the high-frequency regime, where they can be reduced to simple forms.

III. CURRENT DENSITY

As we have already indicated, $\mathbf{j}_{2L}(\mathbf{r}, 2\omega)$ is related only to a linear-response function of media while $\mathbf{j}_{NL}(\mathbf{r}, 2\omega)$ is

related to both linear- and nonlinear-response functions. However, the nonlinear part [the third integral of Eq. (28)] was often neglected in most of the previous works.^{3,4} It is based on the argument that its ratio to the first two integrals is $(\hbar^2 k q / 2m) / \hbar\omega$, where $(\hbar^2 k q)$ comes from two additional momentum operators. Since the ratio is of order v_F / c where v_F , the Fermi velocity, is much smaller than light speed, the third integral is very small. However, it has to be pointed out that momenta are not conserved in the direction normal to the surface. A momentum operator in that direction may very well give a factor $\hbar/l$. Now consider $(\hbar^2 / 2ml^2) / \hbar\omega$. Both numerator and denominator are of order 1 eV if the red light is used as external source. Therefore, the third integral cannot be neglected except for the high-frequency region. Consequently, the phenomenological constants β , γ , and δ defined by Bloembergen *et al.*⁴ cannot be used for analyzing experimental data unless ω is at least of order ω_p , the bulk plasma frequency. It also gives the contribution of forms $\partial_i A_j \partial_l A_k$ and $\partial_i A_j \partial_l A_k \partial_m A_n$. Therefore, it is

not clear what effect this integral will have on a and b defined by RS. Since the z derivatives clearly are surface sensitive, one would not expect the resulting a to be a universal constant as in the hydrodynamic model.⁹ It is also contradictory to what Guyot-Sionnest and Shen¹⁹ claimed. They concluded that nonlinear polarization and hence current density consists of a series of multiple terms. The series can be terminated at the terms with only the first-order derivative with respect to space coordinates. Apparently, the term which we were discussing has derivatives of third order. We note that in their Eq. (10) the integral does have contribution from the surface because $S_z(z, 2\omega)$ (in their notation) has a peak at the surface. Therefore we concluded that multiple terms of higher-order derivatives do contribute to current and polarization and their contribution is surface related.

Now we study $j_{NL}(\mathbf{r}, \omega)$ in the high-frequency region. In this case we can apply two simplifications: the third integral in Eq. (28) can be ignored and E_{nm} in the denominators of the rest can be neglected. Using the closure relation we find

$$j_{NL}(\mathbf{r}, 2\omega) \approx -\frac{1}{\hbar\omega} \int d^3r' \{ \langle [H_1(\mathbf{r}, \omega), \mathbf{J}_A(\mathbf{r}, \omega)] \rangle + \langle [H_3(\mathbf{r}', 2\omega), \mathbf{J}_p(\mathbf{r})] \rangle / 2 \} . \quad (30)$$

The commutators can be worked out easily to give

$$j_{NL}(\mathbf{r}, 2\omega) \approx -\frac{ie^3}{m^2c^2\omega} \mathbf{A}_1(\mathbf{r}, \omega) \nabla \cdot [\mathbf{A}_1(\mathbf{r}, \omega) n_0(\mathbf{r})] - \frac{ie^3 n_0(\mathbf{r})}{4m^2c^2\omega} \nabla A_1^2(\mathbf{r}, \omega) . \quad (31)$$

The difference between Eq. (31) and the form derived by the classical model^{3,4} is the presence of $n_0(\mathbf{r})$ which is position dependent. Its effect will emerge as we analyze the current density carefully.

$j_{NL}(\mathbf{r}, 2\omega)$ has both volume and surface parts. Since $n_0(\mathbf{r})$ is constant and $\nabla \cdot \mathbf{A}_1 = 0$ deep inside the bulk ($z < -l/2$), only the second term on the right-hand side of Eq. (31) exists in the entire volume. Thus we write

$$j_{NL}^{(v)}(\mathbf{r}, 2\omega) = -\frac{ie^3 \bar{n}}{4m^2c^2\omega} \nabla A_1^2(\mathbf{r}, \omega) = -\frac{ie^3 \bar{n}}{4m^2c^2\omega} A_{1,\omega}^2 \nabla e^{2i(qx - k_1 z)} , \quad (32)$$

where Eq. (6) has been used for the second step. $\bar{n}$ is the electron density deep inside the bulk. The other part is concentrated on the surface. It is due to the fact that $\mathbf{A}_1(\mathbf{r}, \omega)$ and $n_0(\mathbf{r})$ vary rapidly (have a jump classically) near the surface. Consider the following:

$$j_{NL,z}^{(s)}(\mathbf{r}, 2\omega) = -\frac{3ie^3}{4m^2c^2\omega} \frac{\partial}{\partial z} [n_0(\mathbf{r}) A_{1,z}^2(\mathbf{r}, \omega)] - \frac{ie^3}{4m^2c^2\omega} A_{1,z}^2(\mathbf{r}, \omega) \frac{\partial n_0(z)}{\partial z} , \quad (33)$$

$$j_{NL,i}^{(s)}(\mathbf{r}, 2\omega) = -\frac{ie^3}{m^2c^2\omega} A_{1,i}(\mathbf{r}, \omega) \frac{\partial}{\partial z} [n_0(\mathbf{r}) A_{1,z}(\mathbf{r}, \omega)] , \quad (34)$$

where superscripts (v) and (s) denote volume and surface, respectively, and subscript i denotes a component parallel to surface. We have also used the fact that $A_{1,i}(\mathbf{r}, \omega)$ is continuous across the surface. What affects the reflected and transmitted second harmonics are the total surface current which is given as the following:

$$j_{NL,z}^{(s),t}(2\omega) = \frac{3ie^3}{4m^2c^2\omega} \bar{n} A_{1,z}^2 e^{2iqx} - \frac{ie^3}{4m^2c^2\omega} \int_{-l/2}^{l/2} A_{1,z}^2(\mathbf{r}, \omega) \frac{\partial n_0(z)}{\partial z} dz = \frac{3ie^3}{4m^2c^2\omega} \bar{n} A_{1,z}^2 e^{2iqx} - \frac{ie^3}{4m^2c^2\omega} \int_{-\infty}^{\infty} A_{1,z}^2(\mathbf{r}, \omega) \frac{\partial n_0(z)}{\partial z} dz , \quad (35)$$

$$j_{NL,i}^{(s),t}(2\omega) = \frac{ie^3 \bar{n}}{m^2c^2\omega} A_{1,i} A_{1,z} e^{2iqx} \quad (36)$$

where we have integrated $j_{NL}^{(s)}$ from $-l/2$ to $l/2$ and used parameter a of RS.¹⁹ $A_{1,z}$ and $A_{1,i}$ are components of $\mathbf{A}_{1,\omega}$ with ω dependence suppressed. The second step in Eq. (35) was reached with the help of the property of $\partial n_0(z)/\partial z$. It is nontrivial only in the region of $-l/2 < z < l/2$. The quantum-mechanical approach has several advantages over the classical one. First, since $A_{1,z}(\mathbf{r}, 2\omega)$ and $n_0(\mathbf{r})$ are treated microscopically, the ambiguity of $E_z \partial E_z / \partial z$ pointed out by RS does not occur. Secondly, the huge factor $(\epsilon^2 - 1)$ of $j_{NL}^{(s),t}$ given by the classical model⁴ disappears because $n_0(r) = 0$ in vacuum and we do not have to resort to the simplified model of RS to achieve it. Most importantly, we have a term [the integral of Eq. (35)] that the classical model fails to give. It is related to the microscopic structure of the surface. It was also absent in RS's results. They used a simplified model for the surface. Hence they could not account for microscopically electron behavior at the surface. It has been pointed out by Guyot-Sionnest and Shen²⁰ that there is such a term due to the abrupt variation of a nonlinear response function at the surface. But here its form was derived explicitly for high frequencies using quantum mechanics. Finally, following RS, we can use a parameter a to express $j_{NL,z}^{(s),t}(2\omega)$ in the form of $(iae^3 \bar{n} / 2m^2c^2\omega) A_{1,z}^2 e^{2iqx}$. Due to the presence of the integral in Eq. (35) a is not a constant. It depends on ω and also on bulk electron density and surface structure. However, we believe the variation due to the integral should not change a 's order of magnitude. We can also make an estimate of a in the high-frequency limit. Using the relations $\mathbf{D} = \epsilon \mathbf{E}$ and $\epsilon(\omega) = 1 - \omega_p^2 / \omega^2$, we found the discontinuity of $A_{1,z}$ at the surface is inversely proportional to ω^2 and hence can be neglected. Thus the integral in Eq. (35) can be done easily and a is equal to uni-

ty. For the above reasons, we think a is a slow-varying function of ω at least for high frequencies. Thus, from now on we shall use the form $(iae^3\bar{n}/2m^2c^2\omega)A_{1,z}^2e^{2iqx}$ for $j_{NL,z}^{(s)t}(2\omega)$.

IV. BOUNDARY CONDITIONS AND EFFICIENCY OF SECOND-HARMONIC GENERATION

Knowing the form of current density, we can solve for nonlinear vector potentials using Eq. (13). However, to find the efficiency of second-harmonic generation, we need only reflected or transmitted vector potentials far away from the surface. They can be derived using boundary conditions instead of solving Eq. (13) for the entire space. The solution of Eq. (13) far away from the surface is²¹

$$\mathbf{A}_2(\mathbf{r}, 2\omega) = \begin{cases} \mathbf{a}_v(2\omega)e^{2i(qx+k_0z)}, & z \geq l/2 \\ \mathbf{a}_b(2\omega)e^{2i(qx-k_2z)} + \frac{\pi e^3\bar{n}}{4m^2c\omega^2\epsilon(2\omega)} A_{1,\omega}^2 \nabla e^{2i(qx-k_1z)}, & z \leq -l/2 \end{cases} \quad (37)$$

See also Eqs. (9) and (10). From now on the argument 2ω of $\mathbf{a}_v(2\omega)$ and $\mathbf{a}_b(2\omega)$ will be dropped.

The boundary conditions can be reduced from Maxwell's equations.²² First, we found that B_z is continuous across the surface. Using Eqs. (4), (37), and (38), we have

$$a_{b,y} = a_{v,y}. \quad (39)$$

Now consider a loop on the x - z plane with two sides at $z = -l/2$ and $z = l/2$ and the other two sides at $x = l_1$

and $x = l_2$ with $l_2 - l_1 \gg l$. Integrating $\mathbf{H}$ around the loop and neglecting terms of order l , we found

$$-2i(k_0a_{v,y} + k_2a_{b,y}) = \frac{4\pi e^3\bar{n}}{m^2c^2\omega} A_{1,y} A_{1,z}. \quad (40)$$

If the loop is on the y - z plane, we have

$$2i(k_0a_{v,x} - qa_{v,z}) + 2i(k_2a_{b,x} + qa_{b,z}) = -\frac{4\pi e^3\bar{n}}{m^2c^3\omega} A_{1,x} A_{1,z}. \quad (41)$$

If instead, $\mathbf{E}$ is integrated around the loop on the x - z plane, the E_z contribution on the short sides cannot be ignored. The surface current in the z direction builds a dipole layer peaked at the surface. Its distribution is a Dirac delta function at the surface in the classical model. Using relations $\mathbf{P} = \mathbf{j}$, $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ and the fact that D_z is continuous we found there is a peak for E_z at the surface due to $j_{NL,z}^{(s)}$. Taking this into account, we found

$$a_{v,x} - a_{b,x} + \frac{\pi e^3\bar{n}q}{2m^2c\omega^3\epsilon(2\omega)} A_{1,\omega}^2 - \frac{4\pi qc}{2i\omega^2} \frac{ie^3\bar{n}a}{2m^2c^2\omega} A_{1,z}^2 = 0. \quad (42)$$

Finally if the loop is on the y - z plane for $\mathbf{E}$, we get the same result as Eq. (39) because $j_{NL,z}^{(s)t}$ does not vary with y .

Equations (39)–(42) combined with

$$qa_{v,x} + k_0a_{v,z} = 0, \quad (43)$$

$$qa_{b,x} - k_2a_{b,z} = 0 \quad (44)$$

determined six unknowns from components of $\mathbf{a}_v$ and $\mathbf{a}_b$:

$$a_{v,x} = - \left[(k_2 + q^2/k_2) \left(\frac{\pi e^3\bar{n}q}{2m^2c\omega^3\epsilon(2\omega)} A_{1,\omega}^2 - \frac{\pi e^3\bar{n}qa}{m^2c\omega^3} A_{1,z}^2 \right) + \frac{2\pi e^3\bar{n}}{m^2c^3\omega} (A_{1,x} A_{1,z}) \right] [(k_0 + q^2/k_0) + (k_2 + q^2/k_2)]^{-1}, \quad (45)$$

$$a_{v,y} = -\frac{2\pi e^3\bar{n}}{m^2c^3\omega(k_0 + k_2)} A_{1,y} A_{1,z}. \quad (46)$$

Using relations $\nabla \cdot \mathbf{A}_0(\mathbf{r}, \omega) = 0$, $\nabla \cdot \mathbf{A}_1(\mathbf{r}, \omega) = 0$ for $|z| \geq l/2$, and

$$A_{1,z} = \frac{2}{\epsilon(\omega) + k_1/k_0} A_{0,z}, \quad (47)$$

we found the efficiency of p -polarized incident light to be

$$R = \frac{4a_v^2}{A_0^4} = \left[\frac{4\pi e^3\bar{n}q^2k_0^2}{m^2\omega^4} \right]^2 \times \left[\epsilon(2\omega) \left(\frac{q(1 + k_1^2/q^2)}{2\epsilon(2\omega)} - aq \right) + \frac{2k_1k_2}{q} \right]^2 \times [k_0\epsilon(\omega) + k_1]^{-4} [k_0\epsilon(2\omega) + k_2]^{-2}. \quad (48)$$

Note that in Eqs. (45) and (48) the only quantity related to detailed surface structure is a . Coming back to Eq. (35) we found that a has both contributions from surface and bulk. Since they have the same behavior, it seems to agree with what Guyot-Sionnest and Shen claimed,²⁰ i.e., surface and bulk contributions are not separable. However, we are not so pessimistic. The form of nonlinear current in Eqs. (31)–(36) is valid only in the high-frequency regime. In this case a is related only to the linear response function of the surface. For low frequencies, the third integral of Eq. (28) would be significant. It could have different behavior as far as q , k_0 , and k_1 are concerned. Therefore, it is possible that nonlinear contributions from surface and bulk are separable.

The denominator of $a_{v,x}$ is interesting. Under certain conditions it may vanish.²³ It indicates that there is a nonlinear normal model. From Eqs. (2) and (11) we found it to be

$$k_2 + k_0\epsilon(2\omega) = 0. \quad (49)$$

FIG. 1. Dispersion of the second-harmonic and fundamental modes of the surface plasmon and the photon are shown in solid, dashed, and dotted lines, respectively.

This is analogous to the condition of surface plasmon of fundamental fields.²⁴ Clearly, the surface plasmon is excited nonlinearly by absorbing incident light. If $\epsilon(2\omega) = 1 - \omega_p^2/4\omega^2$ is used, we find that for small k_0 and ω , the solution is

$$\omega \approx cq \left[1 - \frac{2c^2 q^2}{\omega_p^2} \right] \quad (50)$$

and for large q and ω , it is

$$\omega \approx \omega_p / 2\sqrt{2} . \quad (51)$$

The second-harmonic surface-plasmon dispersion is plotted in Fig. 1 in solid lines. Note that for Eq. (49) to be satisfied, both k_0 and k_2 should be imaginary. Thus, the surface plasmon of second harmonics can be seen in attenuated total reflection experiments.

The denominator of efficiency could vanish in still another way. That is

$$\epsilon(\omega) + k_1/k_0 = 0 . \quad (52)$$

It comes from fundamental fields [see, for example, Eq. (47)]. When incident light is coupled to the surface plasmon, the transmitted (and also reflected) vector potential becomes very large. It generates large nonlinear current which, in turn, gives off second harmonics of large amplitude. The solution of Eq. (52) is given in Fig. 1 in long-dashed lines. Therefore, there are two ways to produce second harmonics of high intensity. One can arrange incident light so that it is coupled to the surface plasmon in order to enhance the linear vector potential. Or, the surface plasmon can be excited nonlinearly to magnify second harmonics.

V. SUMMARY

We studied second-harmonic generation with quantum mechanics. It is found that ambiguity and unphysical factors in the classical model are eliminated. The nonlinear current is, in general, related to nonlinear properties of media. Only in the high-frequency regime can we reduce nonlinear current density to the simple form which is widely used. In this regime, quantum mechanics give a term which was left out in previous classical calculations. By analyzing surface current density we obtained valid boundary conditions. We then evaluated the efficiency of second-harmonic generation and found a nonlinear surface-plasmon mode. From this work, the picture of second-harmonic generation in the high-frequency regime is clear. However, a lot more is left to be done, e.g., second harmonics in the low- and intermediate-frequency regime, those in the presence of absorbed atoms, surface roughness, and grating, and relations between various theoretical approaches such as the hydrodynamical model and time-dependent density functional theory. Past experiments focused on the angular dependence. We feel that the frequency dependence is equally important for understanding second-harmonic generation.

ACKNOWLEDGMENTS

This work is supported in part by the National Science Council of the Republic of China under contract No. NSC78-0208-M002-51.

- ¹P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinrich, *Phys. Rev. Lett.* **7**, 118 (1961).
- ²J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, *Phys. Rev.* **127**, 1918 (1962).
- ³N. Bloembergen and P. S. Pershan, *Phys. Rev.* **128**, 606 (1962).
- ⁴N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, *Phys. Rev.* **174**, 813 (1968).
- ⁵Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984).
- ⁶J. E. Sipe, V. C. Y. So, M. Fukui, and G. I. Stegeman, *Phys. Rev. B* **21**, 4389 (1980).
- ⁷M. Corvi and W. L. Schaich, *Phys. Rev. B* **33**, 3688 (1986).
- ⁸A. Chizmeshya and E. Zaremba, *Phys. Rev. B* **37**, 2805 (1988).

- ⁹W. L. Schaich and A. Liebsch, *Phys. Rev. B* **37**, 6187 (1988).
- ¹⁰Joseph Rudnick and E. A. Stern, *Phys. Rev. B* **4**, 4274 (1971). It will be referred to as RS in text.
- ¹¹M. Weber and A. Liebsch, *Phys. Rev. B* **35**, 4711 (1987); **36**, 6411 (1987).
- ¹²A. Liebsch, *Phys. Rev. Lett.* **61**, 1233 (1988).
- ¹³For a review, see G. L. Richmond, J. M. Robinson, and V. L. Shannon, *Prog. Surf. Sci.* **28**(1), 1 (1988).
- ¹⁴P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- ¹⁵N. D. Lang and W. Kohn, *Phys. Rev. B* **1**, 4555 (1970).
- ¹⁶D. Pines and P. Nozières, *The Theory of Quantum Liquid* (Benjamin, New York, 1966), Vol. 1.
- ¹⁷W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).

- ¹⁸For details, see C. D. Hu and E. Zaremba, Phys. Rev. B **37**, 9268 (1988).
- ¹⁹Due to our choice of metal position (negative half-space), our $j_{\text{Nf}}^{(s)}(2\omega)$ has a different sign. However, the current density, boundary conditions and hence other measurable physical quantities such as $\mathbf{a}_v$ and $\mathbf{a}_b$ are not affected.
- ²⁰P. Guyot-Sionnest and Y. R. Shen, Phys. Rev. B **38**, 7985 (1988).
- ²¹S. S. Jha and C. S. Warke, Phys. Rev. **153**, 751 (1967).
- ²²C. C. Wang, Phys. Rev. **178**, 1457 (1969).
- ²³D. L. Mills, Solid State Commun. **24**, 669 (1977). Mills was the first who noticed that it might be equal to zero. However, he did not point out its relation with any particular mode.
- ²⁴See, for example, F. Forstmann and R. R. Gerhardts, *Metal Optics Near the Plasma Frequency*, Vol. 109 of *Springer Tracts in Modern Physics*, edited by G. Höhler and E. A. Niekisch (Springer-Verlag, Berlin, 1986).