Theory of prism-coupled light emission from tunnel junctions

S. Ushioda,* J. E. Rutledge, and R. M. Pierce
Department of Physics, University of California, Irvine, California 92717
(Received 9 September 1985; revised manuscript received 30 June 1986)

We have developed a theory of light emission from smooth tunnel junctions through a coupler prism. This method of light emission does not depend on the surface roughness of the junction for the emission mechanism by surface-plasmon polaritons, and allows direct emission by the fast mode. The theory follows the Green's-function formalism developed by Laks and Mills and uses the same assumed form of current fluctuations in the junction. The prediction of this theory for the angle dependence of the emission intensity agrees well with the experimental result, but the calculated emission spectra at several angles do not agree with the experimental results. From the fact that the observed spectra show a well-defined peak in the frequency region in which the slow mode is very strongly excited, we conclude that an important process that produces the observed spectra involves the conversion of the slow mode to the fast mode mediated by residual surface roughness.

I. INTRODUCTION

Since the first discovery of light emission from metal-oxide-metal (*M-O-M*) tunnel junctions by Lambe and McCarthy,[1] many experimental and theoretical investigations into the emission mechanism of these tunnel junctions have been undertaken.[2] Through these efforts a general picture of the emission mechanism has emerged. According to this picture the tunneling current excites surface-plasmon polaritons (SPP) of the junction structure. SPP are the electromagnetic normal modes of the layered *M-O-M* structure and do not couple to external free photons, if the surface is perfectly smooth, i.e., they are nonradiative and no light can be emitted by SPP. This absence of coupling arises from the fact that the wave vector of SPP for any given frequency is greater than that of the free photon of the same frequency. Thus the wave vectors cannot be matched between an SPP and a free photon. However, when the junction surface is roughened,[1] has small metal particles lying on it,[3] or is corrugated by a grating,[4] the wave-vector conservation condition along the surface is relaxed by the absence of translational symmetry and SPP can now radiate free photons. Thus roughening of the junction surface or other means of breaking the wave-vector conservation at the surface has been considered an essential part of the SPP emission process from tunnel junctions.

Recently, we have introduced a new geometry for *M-O-M* junctions that allows light emission without the presence of surface roughness or corrugation.[5] This is accomplished by placing a junction on the flat surface of a hemicylindrical prism with refractive index n, as sketched in the inset of Fig. 1. This configuration consists of five layers and four interfaces. The SPP for this geometry has three branches in the visible range: the slow mode (SM), prism-Al mode (PR-Al), and the fast mode (FM). The dispersion curves for these branches are plotted in Fig. 1. We notice that for this geometry the fast mode lies to the left-hand side of the prism light line (PR) which is also drawn in Fig. 1. This means that for some emission angle θ (measured from the surface normal) the wave vector parallel to the surface can be matched between SPP and light in the prism. At a given frequency ω, the wave-vector matching condition along the surface is given by

$$k_{\|\mathrm{SPP}}(\omega)=n\omega\sin\theta/c\ . \tag{1}$$

Since $k_{\|\mathrm{SPP}}$ is less than $n\omega/c$, there is real θ that satisfies Eq. (1), and θ is the emission angle of light in the prism. Thus we see that wave-vector conservation along the surface can be accomplished using a prism of an appropriate refractive index without breaking the wave-vector conservation by surface roughness or corrugation. A similar configuration consisting of a prism and a metal film has been used in the method of attenuated total reflection (ATR) to excite SPP by injection of light from the prism.[6] This configuration in the field of ATR spectroscopy is

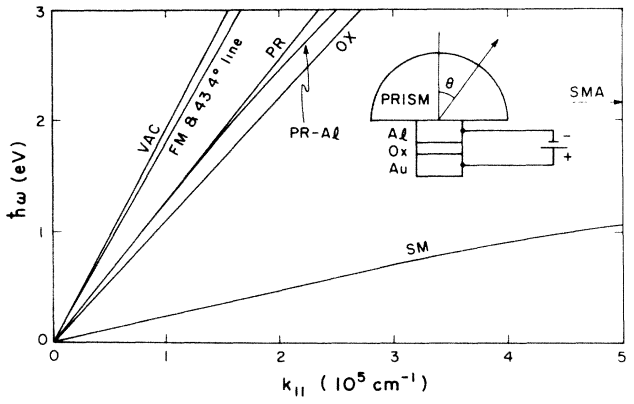


FIG. 1. Geometry of a tunnel junction in contact with a hemicylindrical coupler prism, and the dispersion curves of surface-plasmon polaritons in this layered structure. VAC: vacuum light line; FM: fast-mode dispersion curve; PR: prism light line; PR-Al: prism-aluminum mode dispersion curve; OX: oxide light line; SM: slow-mode dispersion curve; SMA: asymptote for slow-mode dispersion curve.

called "Kretschmann configuration."[7] Light emission through a prism coupler is the reverse process to the excitation of SPP by light injection. Indeed, we have demonstrated that light emission by this method does occur from *M-O-M* tunnel junctions without the intermediary of surface roughness or corrugation.[5]

In the case of roughened junctions, all three modes can radiate in principle, and there have been discussions as to which mode is contributing most strongly to the emitted spectra.[2] On the other hand, for prism-coupled emission, there is no question about which mode radiates, i.e., only the fast mode radiates with p polarization, apart from very weak emission due to residual surface roughness.

Surface roughness is a difficult property to quantify experimentally. Thus so far it has not been possible to make detailed comparisons between experimental emission spectra and theoretical ones which are calculated on the basis of some assumed distribution of surface roughness. With the prism-coupled emission method, there is no experimental parameter that is undetermined. Thus this method presents an opportunity to make a rigorous test of theories and that is what we wish to attempt in this paper. Since the emission part of the process is simple and straightforward, we can focus on the details of the excitation mechanism of SPP modes, i.e., the coupling mechanism between currents and SPP. Kirtley *et al.*[8] have pointed out the possibility that SPP are excited by hot-electron injection into a metal rather than by current fluctuations as assumed by all theories so far. The prism-coupler method will be useful in answering this kind of questions without complications caused by uncertainties introduced by surface roughness.

The main purpose of this paper is to present a theory of prism-coupled light emission from tunnel junctions and to compare the predictions of the theory with our experimental results. To do this we will follow the theoretical framework previously presented by Laks and Mills[9] for roughened junctions, and adapt their Green's-function method to the smooth junction geometry of interest here.

II. THEORY

We will consider a five-layer—four-interface geometry depicted in Fig. 2. The layers numbered 0 through 4 correspond to prism, aluminum, aluminum oxide, gold, and vacuum, respectively, in the experimental configuration. The dielectric constant of each layer is designated by ϵ_n, and the layer boundaries are at $z=-b$, 0, d, and $d+a$. This geometry is different from the one considered by Laks and Mills in two regards: first, we have placed the prism layer in contact with the Al layer, making the Al layer thickness finite; second, our interfaces are assumed to be perfectly smooth. Thus the infinitesimal translational symmetry along the surface is maintained. The objective of the theory is to obtain the electric field in the prism, $E_\mu(\mathbf{x},\omega)$, given the current density in the tunnel junction, $J_\mu(\mathbf{x},t)$, and to calculate the emitted power in the prism. Since we follow the derivation by Laks and Mills closely, we will only sketch the outline of the calculation.

Let us assume that the μth component of the current density in the junction at position $\mathbf{x}$ and time t is given by

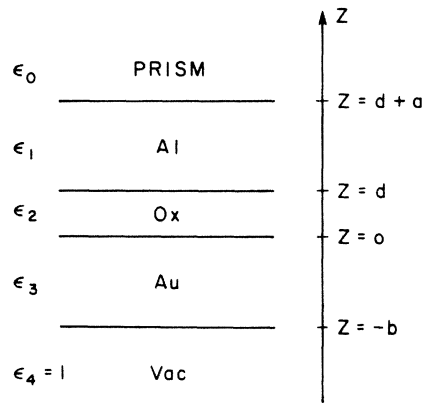


FIG. 2. Definition of symbols for a five-layered metal-dielectric structure.

$$J_\mu(\mathbf{x},t)=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}J_\mu(\mathbf{x};\omega)e^{-i\omega t}\,, \tag{2}$$

where ω is the frequency of the Fourier component. Then the μth component of the electric field at frequency ω and position $\mathbf{x}$ is given by

$$E_\mu(\mathbf{x},\omega)=\frac{i\omega}{c^2}\sum_\lambda\int d^3x'\,D_{\mu\lambda}(\mathbf{x},\mathbf{x}';\omega)J_\lambda(\mathbf{x}';\omega)\,, \tag{3}$$

where $D_{\mu\lambda}(\mathbf{x},\mathbf{x}';\omega)$ is the Green's-function matrix which is a solution of the Maxwell equation

$$\sum_\mu\left[\frac{\omega^2}{c^2}\epsilon(\mathbf{x},\omega)\delta_{\lambda\mu}-\frac{\partial^2}{\partial x_\lambda\partial x_\mu}+\delta_{\lambda\mu}\nabla^2\right]D_{\mu\nu}(\mathbf{x},\mathbf{x}';\omega)=4\pi\delta_{\lambda\nu}(\mathbf{x}-\mathbf{x}')\,. \tag{4}$$

$\epsilon(\mathbf{x},\omega)$ takes on values given by ϵ_n ($n=0,1,2,3,4$) depending on the value of z. Since we have translational symmetry along the surface, we can express $D_{\mu\nu}(\mathbf{x},\mathbf{x}';\omega)$ in terms of its two-dimensional Fourier transform $d_{\mu\nu}(\mathbf{k}_{||},\omega;z,z')$,

$$D_{\mu\nu}(\mathbf{x},\mathbf{x}';\omega)=\int\frac{d^2k_{||}}{(2\pi)^2}e^{i\mathbf{k}_{||}\cdot(\mathbf{x}_{||}-\mathbf{x}'_{||})}d_{\mu\nu}(\mathbf{k}_{||},\omega\,|\,z,z')\,, \tag{5}$$

where $\mathbf{k}_{||}$, $\mathbf{x}_{||}$, and $\mathbf{x}'_{||}$ are the components of the respective vectors parallel to the surface. From the boundary condition at $z\to\infty$ one can write $d_{\mu\nu}(\mathbf{k}_{||},\omega;z,z')$ in the form

$$d_{\mu\nu}(\mathbf{k}_{||},\omega\,|\,z,z')=e^{ik_0z}\epsilon_{\mu\nu}(\mathbf{k}_{||},\omega\,|\,z')\,, \tag{6}$$

where

$$k_0=\left[\frac{n^2(\omega+i\eta)^2}{c^2}-k_{||}^2\right]^{1/2}. \tag{7}$$

η is a positive infinitesimal and n is the real refractive index of the prism $n=(\epsilon_0)^{1/2}$. We take the square root in Eq. (7) so that $\mathrm{Im}k_0\geq 0$ to ensure finite $E(\mathbf{x},\omega)$ at $z\to\infty$.

By substituting Eq. (5) and Eq. (6) into Eq. (3), we obtain

$$E_\mu(\mathbf{x},\omega)=\frac{i\omega}{c^2}\sum_\lambda\int d^3x'\int\frac{d^2k_{||}}{(2\pi)^2}e^{i\mathbf{k}_{||}\cdot(\mathbf{x}_{||}-\mathbf{x}'_{||})}e^{ik_0z}\times\epsilon_{\mu\lambda}(\mathbf{k}_{||},\omega\,|\,z')J_\lambda(\mathbf{x}';\omega)\,. \tag{8}$$

To effect the integration in Eq. (8), we let

$$\mathbf{x}=|\mathbf{x}|(\hat{\mathbf{x}}\sin\theta_0\cos\phi_0+\hat{\mathbf{y}}\sin\theta_0\sin\phi_0+\hat{\mathbf{z}}\cos\theta_0)\ , \tag{9}$$

where (θ_0,ϕ_0) is the direction of observation, and $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$ are the unit vectors in the respective direction. We also let

$$\mathbf{k}_{||}=\frac{n\omega}{c}(\hat{\mathbf{x}}\sin\theta\cos\phi+\hat{\mathbf{y}}\sin\theta\sin\phi)\ . \tag{10}$$

Then

$$|\mathbf{k}_{||}|^2=\left[\frac{n\omega}{c}\right]^2\sin^2\theta \tag{11}$$

and

$$k_0=\left[\frac{n\omega}{c}\right]\cos\theta\ . \tag{12}$$

The radiative part of the contributions to the electric field comes only from the part of the integral for which k_0 is real and positive, i.e., $k_{||}\le n\omega/c$. Thus the appropriate integration limits are $0\le\theta\le\pi/2$ and $0\le\phi\le 2\pi$. When we use the method of stationary phase to carry out the integration, we obtain the radiative part of the electric field $E_\mu^{(R)}$,

$$E_\mu^{(R)}(\mathbf{x},\omega)=\frac{in\omega^2\cos\theta_0}{4\pi c^3|\mathbf{x}|}\exp\left[\frac{in\omega}{c}|\mathbf{x}|\right]\times\sum_\lambda\int d^3x'\,e^{-i\mathbf{k}_{||}^{(0)}\cdot\mathbf{x}_{||}}\times\epsilon_{\mu\lambda}(\mathbf{k}_{||}^{(0)},\omega;z')J_\lambda(\mathbf{x}';\omega)\ , \tag{13}$$

where

$$\mathbf{k}_{||}^{(0)}=\frac{n\omega}{c}(\hat{\mathbf{x}}\sin\theta_0\cos\phi_0+\hat{\mathbf{y}}\sin\theta_0\sin\phi_0)\ . \tag{14}$$

The time-averaged power delivered per unit solid angle per unit frequency range by this electric field is given by the Poynting vector as

$$P(\mathbf{k}_{||}^{(0)},\omega)\equiv\left\langle\frac{d^3W}{d\Omega\,d\omega\,dt}\right\rangle=|\mathbf{x}|^2\frac{c}{8\pi}\sum_\mu\langle E_\mu^{(R)*}(\mathbf{x},\omega)E_\mu^{(R)}(\mathbf{x},\omega)\rangle\ . \tag{15}$$

By substituting Eq. (13) into Eq. (15), we have

$$P(\mathbf{k}_{||}^{(0)},\omega)=\frac{n^2\omega^4}{4\pi^2c^6}\sum_{\lambda,\lambda'}\int d^3x'\int d^3x''\int d^2Q_{||}\,e^{i(\mathbf{k}_{||}^{(0)}-\mathbf{Q}_{||})\cdot(\mathbf{x}''-\mathbf{x}')}\epsilon_{\mu\lambda'}^*(\mathbf{k}_{||}^{(0)},\omega;z'')\epsilon_{\mu\lambda}(\mathbf{k}_{||}^{(0)},\omega;z')J_{\lambda\lambda'}(\mathbf{Q}_{||},\omega;z',z'')\ , \tag{16}$$

where we took advantage of the translational symmetry in the surface plane and defined

$$\langle J_{\lambda'}^*(\mathbf{x}'',\omega)J_\lambda(\mathbf{x}',\omega)\rangle\equiv\int d^2Q_{||}\,e^{-i\mathbf{Q}_{||}\cdot(\mathbf{x}''-\mathbf{x}')}J_{\lambda\lambda'}(\mathbf{Q}_{||},\omega;z'',z')\ . \tag{17}$$

$\mathbf{Q}_{||}$ is a wave vector parallel to the surface. By carrying out the integration in the plane, we obtain

$$P(\mathbf{k}_{||}^{(0)},\omega)=\frac{An^2\omega^4\cos^2\theta_0}{8\pi c^6}\sum_{\lambda,\lambda'}\sum_\mu\int dz'\int dz''\,\epsilon_{\mu\lambda'}^*(\mathbf{k}_{||}^{(0)},\omega;z'')\epsilon_{\mu\lambda}(\mathbf{k}_{||}^{(0)},\omega;z')J_{\lambda\lambda'}(\mathbf{k}_{||}^{(0)},\omega;z',z'')\ , \tag{18}$$

where $A=\int dx^2$ is the area of the junction. Now we introduce a simplifying assumption that only the $J_{zz}(\mathbf{k}_{||},\omega;z,z')$ component of the current fluctuation is nonzero. Then

$$P(\mathbf{k}_{||}^{(0)},\omega)=\frac{An^2\omega^4\cos^2\theta_0}{8\pi c^6}\sum_\mu\int dz'\int dz''\,\epsilon_{\mu z}^*(\mathbf{k}_{||}^{(0)},\omega;z')\epsilon_{\mu z}(\mathbf{k}_{||}^{(0)},\omega;z'')J_{zz}(\mathbf{k}_{||}^{(0)},\omega;z',z'')\ . \tag{19}$$

Now we need the explicit forms of $\epsilon_{\mu z}(\mathbf{k}_{||},\omega;z)$ for our geometry. They can be obtained by following the treatment in the Appendix of Mills and Maradudin[10] and noting that for large z above the junction area $z-z'\gg 0$ and $z-z''\gg 0$. After some algebra we arrive at the expression

$$P(\mathbf{k}_{||}^{(0)},\omega)=\frac{\pi A\omega^2\sin^2\theta_0}{2c^4|F_-(\mathbf{k}_{||}^{(0)},\omega)|^2}\int dz'\int dz''\,E_z^{<*}(z')E_z^{<}(z'')J_{zz}(\mathbf{k}_{||}^{(0)},\omega;z',z'')\ . \tag{20}$$

In Eq. (20), $E_z^<(z)$ for different regions is defined by

$$E_z^<(z)=F_+e^{ik_{0z}}+F_-e^{-ik_{0z}}\ \text{(prism)}\ , \tag{21}$$

$$E_z^<(z)=G_+e^{ik_{1z}}+G_-e^{-ik_{1z}}\ \text{(Al)}\ , \tag{22}$$

$$E_z^<(z)=H_+e^{ik_{2z}}+H_-e^{-ik_{2z}}\ \text{(oxide)}\ , \tag{23}$$

$$E_z^<(z)=J_+e^{ik_{3z}}+J_-e^{-ik_{3z}}\ \text{(Au)}\ , \tag{24}$$

$$E_z^<(z)=e^{ik_{4z}}\ \text{(vacuum)}\ , \tag{25}$$

where

$$k_i=(\epsilon_i\omega^2/c^2-k_{||}^2)^{1/2}\ . \tag{26}$$

In taking the square root in Eq. (27), we choose the root so that $\mathrm{Im}(k_i)\le 0$ for $i\neq 0$. The amplitudes of the electric field, $F_\pm$, $G_\pm$, $H_\pm$, and $J_\pm$, are determined by fitting the Maxwell boundary conditions at four interfaces. Their explicit expressions are given below:

$$J_+ = \frac{1}{2}\left[\frac{\epsilon_4}{\epsilon_3}+\frac{k_4}{k_3}\right]e^{i(k_3-k_4)b}, \tag{27}$$

$$J_- = \frac{1}{2}\left[\frac{\epsilon_4}{\epsilon_3}-\frac{k_4}{k_3}\right]e^{-i(k_3+k_4)b}, \tag{28}$$

$$H_+ = \frac{1}{2}\left[\frac{\epsilon_3}{\epsilon_2}+\frac{k_3}{k_2}\right]J_+ + \frac{1}{2}\left[\frac{\epsilon_3}{\epsilon_2}-\frac{k_3}{k_2}\right]J_- , \tag{29}$$

$$H_- = \frac{1}{2}\left[\frac{\epsilon_3}{\epsilon_2}-\frac{k_3}{k_2}\right]J_+ + \frac{1}{2}\left[\frac{\epsilon_3}{\epsilon_2}+\frac{k_3}{k_2}\right]J_- , \tag{30}$$

$$G_+ = \frac{1}{2}\left[\frac{\epsilon_2}{\epsilon_1}+\frac{k_2}{k_1}\right]H_+ e^{i(k_2-k_1)d} + \frac{1}{2}\left[\frac{\epsilon_2}{\epsilon_1}-\frac{k_2}{k_1}\right]H_- d^{-i(k_2+k_1)d}, \tag{31}$$

$$G_- = \frac{1}{2}\left[\frac{\epsilon_2}{\epsilon_1}-\frac{k_2}{k_1}\right]H_+ e^{i(k_1+k_2)d} + \frac{1}{2}\left[\frac{\epsilon_2}{\epsilon_1}+\frac{k_2}{k_1}\right]H_- e^{-i(k_2-k_1)d}, \tag{32}$$

$$F_+ = \frac{1}{2}\left[\frac{\epsilon_1}{\epsilon_0}+\frac{k_1}{k_0}\right]G_+ e^{i(k_1-k_0)(d+a)} + \frac{1}{2}\left[\frac{\epsilon_1}{\epsilon_0}-\frac{k_1}{k_0}\right]G_- e^{-i(k_0+k_1)(d+a)}, \tag{33}$$

$$F_- = \frac{1}{2}\left[\frac{\epsilon_1}{\epsilon_0}-\frac{k_1}{k_0}\right]G_+ e^{i(k_0+k_1)(d+a)} + \frac{1}{2}\left[\frac{\epsilon_1}{\epsilon_0}+\frac{k_1}{k_0}\right]G_- e^{-i(k_1-k_0)(d+a)}. \tag{34}$$

Following Laks and Mills again let us assume the form of the current fluctuation,

$$J_{zz}(Q_{||},\omega \mid z',z'') = \frac{eI_0(1-\hbar\omega/eV_0)}{2\pi^2 A} \times \frac{\Delta(z',z'')}{(1+Q_{||}^2\xi_0^2)^{3/2}}, \tag{35}$$

where ξ_0 is the coherence length of the current fluctuation in the x-y plane, e is the electron charge, I_0 is the total current in the junction of area A, V_0 is the bias voltage across the junction, and we set $\Delta_{zz'}=1$ in the oxide layer and zero outside. Then we can integrate Eq. (20) analytically and obtain

$$P(\mathbf{k}_{||},\omega) = \frac{\omega^2\sin^2\theta}{\pi c^3}\frac{eI_0(1-\hbar\omega/eV_0)}{(1+k_{||}^2\xi_0^2)^{3/2}} \times \left|\frac{(H_+ e^{ik_2d/2}+H_- e^{-ik_2d/2})\sin(k_2d/2)}{k_2F_-(\mathbf{k}_{||},\omega)}\right|. \tag{36}$$

In arriving at Eq. (36), we have dropped the subscript and superscript 0 from θ_0 and $\mathbf{k}_{||}^{(0)}$, respectively. This is the final expression we will use to calculate the emission spectra as a function of the emission angle θ. $P(\mathbf{k}_{||},\omega)$ can also be considered a function of (θ,ω) because of the relation $k_{||}=n\omega\sin\theta/c$. The radiative part of $P(\mathbf{k}_{||},\omega)$ is limited to the region where $k_{||}\le n\omega/c$.

The dispersion relation for the SPP of the layered structure is given by the condition that $E_z^<(z)$ remains finite at $z\to+\infty$, that is, from Eq. (21),

$$F_-(\mathbf{k}_{||},\omega)=0 . \tag{37}$$

Thus we see that $P(\mathbf{k}_{||},\omega)$ is peaked along the dispersion curves of SPP where $F_-(\mathbf{k}_{||},\omega)$ is small. If we assume ω to be real, the value of $k_{||}$ that satisfies Eq. (37) is complex. Thus for real $k_{||}=(n\omega/c)\sin\theta$, $|F_-(k_{||},\omega)|$ remains finite but becomes small when $k_{||}$ (real) is close to the true zero of Eq. (37) in the complex $k_{||}$ plane.

When we measure the emission spectrum along a fixed direction of observation θ, we measure the variation in the height of $P(\mathbf{k}_{||},\omega)$ along a line defined by $k_{||}=(n\omega/c)\sin\theta$ in the $(\mathbf{k}_{||},\omega)$ space. We will call this line in the $(\mathbf{k}_{||},\omega)$ space an "observation line." When the observation line crosses a dispersion curve, we will see a peak at the crossing frequency. On the other hand, if a dispersion curve and the observation line are essentially parallel, we expect a broad emission spectrum. One example of such a case is shown in Fig. 1 where the observation line for $\theta=43.4°$ overlaps the dispersion curve for the fast mode. The best way to visualize the predicted spectrum for different angles is to plot a contour map of $P(\mathbf{k}_{||},\omega)$ as we show in Sec. III.

Once the dispersion curve is found by solving Eq. (37), the electric field pattern for each mode can be found from Eqs. (21)—(26) as a function of $(k_{||},\omega)$. Since SPP is p polarized, $E_y(z)=0$ everywhere, and $E_x(z)$ is simply related to $E_z(z)$ through one of the Maxwell equations. Thus Eqs. (21)—(26) are sufficient to determine the complete electric field pattern. Throughout our calculation we take ω to be pure real and allow $k_{||}$ to be complex. The imaginary part of $k_{||}$ then represents spatial decay of SPP along the interface.

III. NUMERICAL RESULTS

Equation (36) contains two simplifying assumptions. First, we assumed that the only nonzero element of $J_{\mu\nu}(\mathbf{k}_{||},\omega;z,z')$ is $J_{zz}(\mathbf{k}_{||},\omega;z,z')$. Second, we assumed that $\Delta_{zz'}=1$ inside the oxide layer and zero outside. The first assumption will affect the relative strengths of excitation of the three SPP modes, because the three modes have strong electric fields in different directions which can couple to other components of $J_{\mu\nu}(\mathbf{k}_{||},\omega;z,z')$. The second assumption can be made less restrictive. For instance, we can let $\Delta_{zz'}=1$ everywhere in the junction and still carry out the calculation analytically. However, it will take a major theoretical effort to obtain its functional form and value from first principles. In what follows, we will make calculations using the above assumptions, and see how well the simplified theory can describe the experimental data. There is no other assumption or adjustable parame-

ter in the theory, and numerical calculations can be carried out using only known parameters.

The calculations were made for Al-oxide-Au junctions using the value of $\epsilon_2=3.1$ for aluminum oxide and interpolated values of the dielectric function for Au obtained from the table of complex index of refraction given by Johnson and Christy.[11] The dielectric function for Al was approximated by the free-electron form

$$\epsilon(\omega)=1-\omega_p^2/\omega(\omega+i/\tau_c)\ , \tag{38}$$

where we used the values of $\omega_p=12.7$ eV and $\tau_c=5.12\times10^{-15}$ sec reported by Ehrenreich, Philipp, and Segall.[12]

In the late stage of this work, we realized that better fits of the emission spectra can be achieved by using a more realistic set of data for the frequency dependence of the dielectric function for Al. Thus for Figs. 8(a), 8(b), 8(c), and 10, we used the dielectric function data compiled by Ordal *et al.*[13] None of the other figures are affected significantly by this change of the dielectric function for Al.

For a working prototype we used the thicknesses of the Al, oxide, and Au layers of 200, 25, and 200 Å, respectively. We found that changes in the thickness by a factor of 2 or 3 around these values do not affect the general features of the calculated emission spectra or the angle dependence significantly.

Figure 3 shows an expanded version of the dispersion curves for the three modes: the slow mode, the prism-Al mode, and the fast mode. These curves were found from Eq. (37) assuming ω to be real and by looking for complex $k_{||}$ that satisfy the equation. In this figure we plotted only the real part of $k_{||}$. The complex roots were found by using a numerical complex root finding routine. We notice that the dashed part of the dispersion curve for the fast mode moves to the left of the vacuum light line. In the dashed region $\mathrm{Im}k_{||}>\mathrm{Re}k_{||}$ and this part of the curve does not represent a true propagating mode. Of course, there cannot be a mode that propagates faster than the speed of light.

Figure 4 is a plot of $E_x(z)$ as a function of z for the fast mode and the prism-Al mode. The thicknesses of the layers are those of our prototype model: Al:200 Å, Ox:25 Å, and Au:200 Å. We note that $E_x(z)$ of the fast mode is oscillatory in the prism; this means that the fast mode is radiative in the prism as expected from our earlier discussion. $E_x(z)$ decays exponentially on the vacuum side, meaning that the fast mode is nonradiative into vacuum. In contrast to the fast-mode behavior, the prism-Al mode is nonradiative both in the prism and vacuum, showing exponential decay on both sides.

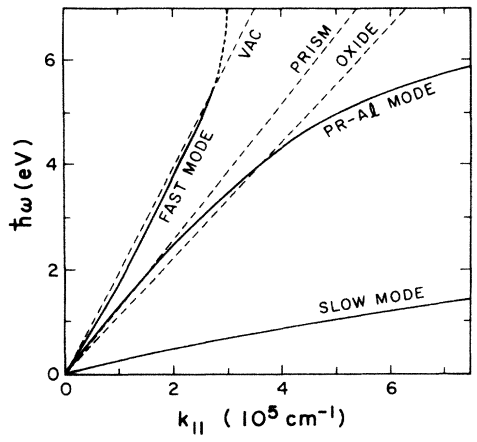


FIG. 3. Expanded view of the dispersion curves of a model layered structure consisting of a glass prism (BK-7), Al (200 Å), aluminum oxide (25 Å), Au (200 Å), and vacuum.

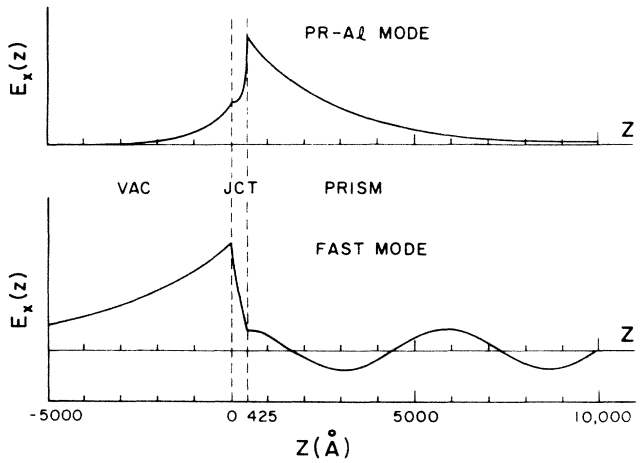


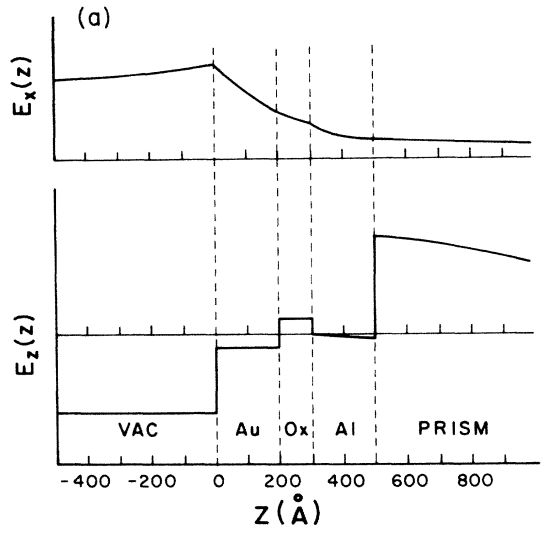
FIG. 4. Electric field patterns of the fast mode and the prism-Al mode for the same model structure as in Fig. 3.

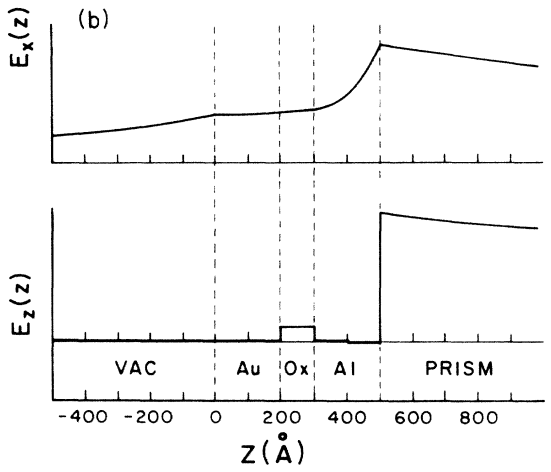
Figures 5(a), 5(b), and 5(c) show both the x and z components of the electric field in the tunnel junction for the three modes. For these figures the thickness of the oxide layer is increased to 100 Å to show the details of the field patterns inside the junction. We note that the strongest field in the z direction appears across the oxide layer for the slow mode. This is the reason why the slow mode couples most efficiently to the J_{zz} component of current fluctuations.

If the dielectric constants of the metals were pure real and negative, one expects $E_z(z)$ to change signs across the metal-dielectric interface [because of the condition that $D_z(z)$ be continuous]. However, we notice that in some cases $E_z(z)$ does not change signs, e.g., at the vacuum-Au interface for the fast mode. This occurs because the imaginary part of the dielectric function for metals plays a very strong role in fitting the boundary conditions.

The angle dependence of the emitted intensity at 5850 Å (2.12 eV) is plotted in Fig. 6. This curve was obtained by fixing ω and varying θ in Eq. (36). Both the peak position ($\theta=43.4°$) and the peak width compare well with the experimental result reproduced in Fig. 7 (see Ref. 5). The experimental result shows weak s-polarized emission, but the theory predicts no s-polarized emission. We believe that the observed s-polarized emission results from the residual surface roughness of the junction.[5]

Figures 8(a), 8(b), and 8(c) show comparisons between the observed emission spectra and the calculated spectra. The dots represent the observed spectra for three emission angles reproduced from Ref. 5, and the solid curves were





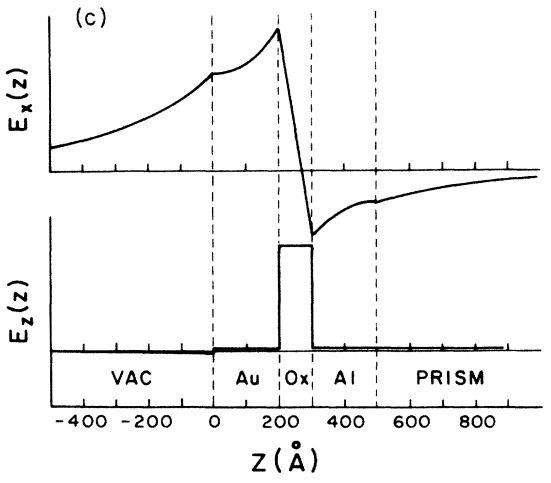


FIG. 5. (a) Electric field pattern of the fast mode showing both $E_x(z)$ and $E_z(z)$. The layer thicknesses are Al (200 Å), oxide (100 Å), and Au (200 Å). Note the expanded thickness for the oxide layer. (b) Electric field pattern of the prism-Al mode for the same layer thicknesses as in (a). (c) Electric field pattern of the slow mode for the same layer thicknesses as in (a).

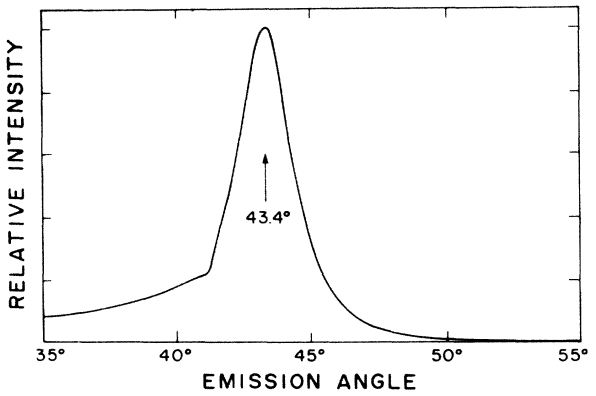


FIG. 6. Calculated angle dependence of the emission intensity at 5850 Å for p-polarized light. The theory predicts no emission with s polarization.

calculated from Eq. (36) with the parameters specified in the figure captions. The data points are corrected for the wavelength dependent sensitivity of the detection system (gratings and photomultiplier). The theory curves were produced by convoluting emission curves for a range of angles, using a triangular pass function corresponding to the acceptance angle of the spectrometer system (approximately 3° FWHM). Normalization between the theory curves and the experimental data is arbitrary, but a single constant ratio was used for all three data sets. Thus the relative intensities in the three figures for both data points and theory curves reflect the real situation.

We see that the cutoff wavelength on the short-wavelength end is correctly reproduced by the theory, but the shape of the spectra is not well represented by the calculated curves. On the other hand, the relative overall in-

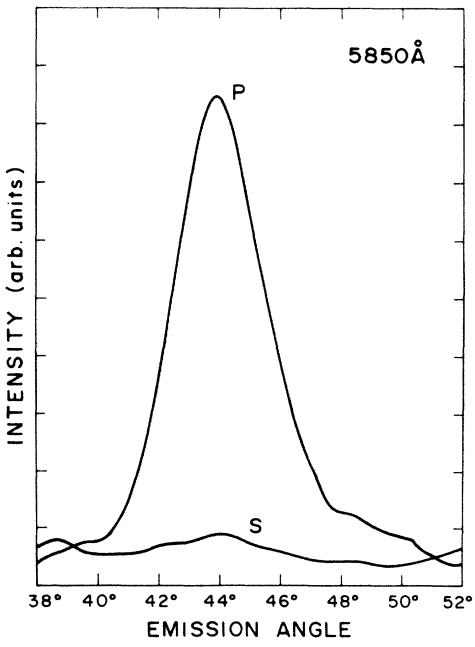


FIG. 7. Measured angle dependence of the emission intensity at 5850 Å from Ref. 5.

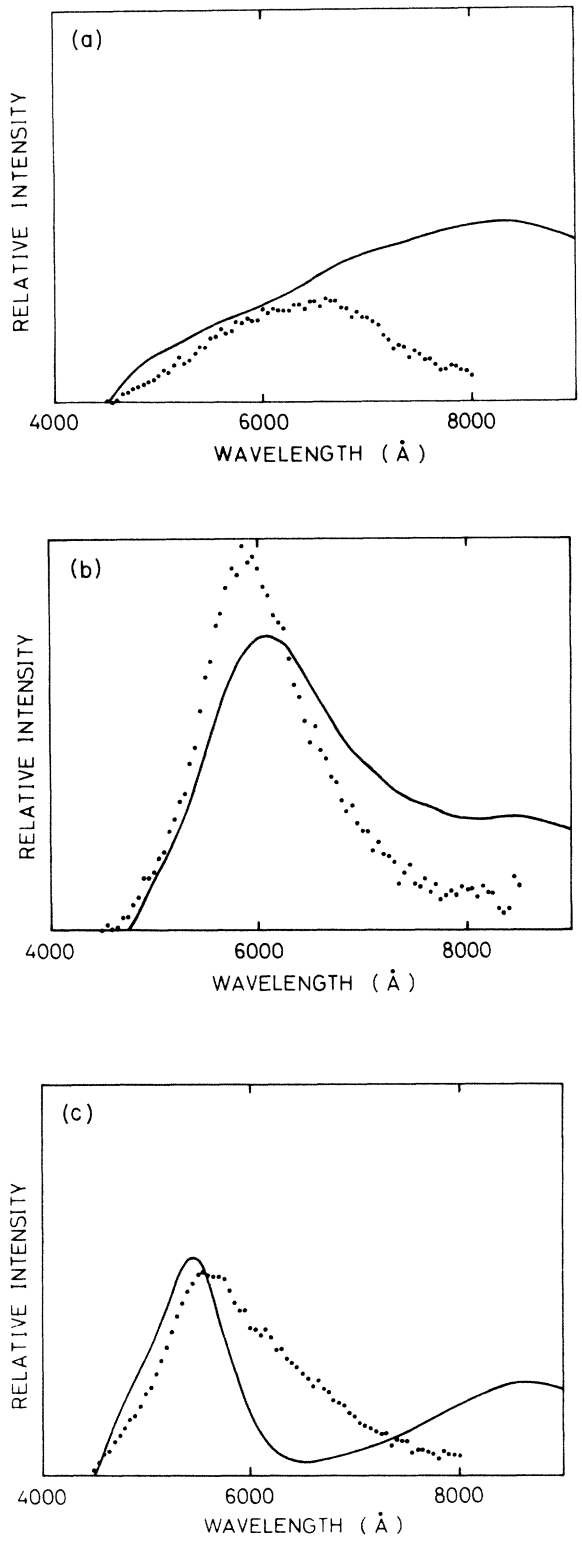


FIG. 8. (a) Measured (dots) and calculated emission spectra at 41°. Bias voltage: 2.75 V. Film thicknesses used in the calculation are Al (180 Å), Ox (20 Å), and Au (320 Å). (b) Measured (dots) and calculated emission spectra at 43.7°. Bias voltage: 2.6 V. Film thicknesses are the same as in (a). (c) Measured (dots) and calculated emission spectra at 46°. Bias voltage: 2.75 V. Film thicknesses are the same as in (a).

tensities for different emission angles show reasonable agreement between theory and experiment, as we have already seen in Figs. 6 and 7.

Since there is no adjustable parameter other than the single normalization constant, the discrepancies shown in Fig. 8 clearly demonstrate the inadequacy of the present theory. In earlier comparisons of emission data with theory[9,14] the degree of surface roughness was used as an adjustable factor to fit the data. However, in the present experiment with smooth junction surfaces, there is no ambiguity about the conditions of the surface, and as we have said earlier, this result allows rigorous evaluation of theories. The comparisons shown in Figs. 8(a), 8(b), and 8(c) clearly indicate that improvements of the theory are required to explain the observed frequency dependence of the emitted light.

In order to understand the cause for the discrepancies between experimental and calculated spectra, we have examined the contour map of $P(k_{\|},\omega)$ of Eq. (36) in the $(k_{\|},\omega)$ plane, as shown in Fig. 9. Note that the scale unit for $k_{\|}$ is 10^6 cm^{-1}, while the unit used in Fig. 1 and Fig. 3 is 10^5 cm^{-1}. At this scale all three light lines for vacuum, prism, and oxide are squeezed near the vertical axis shown by dashed lines. The height of the contours are normalized to the maximum value of $P(k_{\|},\omega)$ that occurs near $k_{\|}=1.5\times10^6$ cm^{-1} and $\hbar\omega=2$ eV. This is the region of $(k_{\|},\omega)$ space where the asymptote of the dispersion curve of the slow mode lies (beyond the right-hand edge of Fig. 1). Compared with this high contour region at large-$k_{\|}$ values, the height of contours in the radiative region to the left of the prism light line is a factor of 10^{-4} to 10^{-5} lower. The large value of $P(k_{\|},\omega)$ in this region reflects very strong coupling between tunneling currents and the slow mode. Most of the power from the tunneling current is pumped into the slow mode and the power going directly into the fast mode is negligible in comparison. We note that this high contour region of the slow mode lies in the energy range of the observed peak of the emission spectra of Fig. 8. This is the reason why we postulated[5] a conversion process of the slow mode to the fast mode via residual junction roughness as the most important mechanism that produces the observed spectra. Mode conversion occurs in both directions, but since the slow mode is so highly populated compared with the fast mode (by a factor of 10^4 to 10^5), the dominant net effect is the conversion of the slow mode to the fast mode. Once the fast mode is created from the slow mode, it radiates free photons into the prism. Since the conversion mediated by surface roughness conserves energy and only changes the value of $k_{\|}$, the population distribution of the fast mode so created reflects the density of states of the slow modes. Hence, we expect high population density of the converted fast mode in the energy range between 1.6 and 2.2 eV. This is why we observe a narrower peak in the observed emission spectra than the calculated result. The theory presented here assumes that all modes are excited directly by the tunneling current, and does not take into account the residual roughness-induced conversion of the slow mode as an excitation process for the fast mode.

As a result from this discussion, one may ask why the slow mode is not converted directly to free photons as ef-

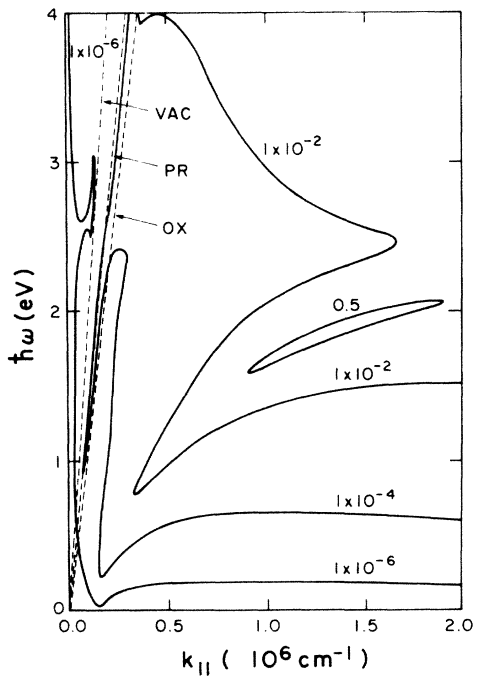


FIG. 9. Contour map of $P(k_{\|},\omega)$ in the $(k_{\|},\omega)$ plane.

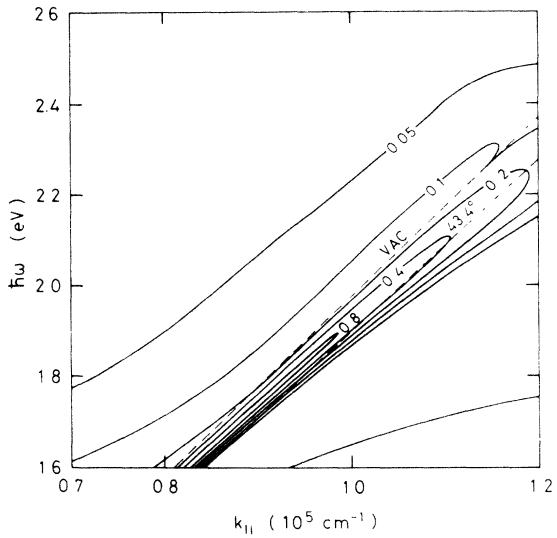


FIG. 10. Detail of the contour map of $P(k_{\|},\omega)$ near the light line, showing a narrow ridge along the fast-mode dispersion curve.

ficiently as to the fast mode? As far as the magnitude of $k_{\|}$ is concerned, both the fast mode and the free photon (vacuum light line) lie close to each other. At this point we must assume that the matrix element coupling the slow mode to free photons is much smaller than that between the slow and fast modes. Whether or not this is true must be answered by a theory that includes junction roughness.

Figure 10 shows a detailed map of contours of $P(k_{\|},\omega)$ near the vacuum light line where the narrow ridge due to the excitation of the fast mode lies. The ridge along the dispersion curve of the fast mode does not have a uniform height along its length; its height varies with frequency. This height variation arises from the structures in the dielectric constant of Au and Al. The height of the contours is normalized to unity at the maximum in this frame. The observation line for 43.4° is indicated by a dashed straight line. The height of $P(k_{\|},\omega)$ along this line is projected into the calculated spectrum at this emission angle. We see that this observation line is almost parallel to the ridge and as a result produces the calculated spectra of Fig. 8 which have a long tail on the low-energy side (long wavelength). To produce the more narrow observed spectra, we must have a high area on this ridge near 2.12 eV. However, with only the direct excitation of the fast mode included in the theory, the highest part of the fast-mode ridge lies around 1.7 eV as we see in Fig. 10. This is the reason why we were compelled to postulate conversion of the slow mode into the fast mode via surface roughness. Such a conversion process will place a high contour region on the ridge centered around 2 eV, and will produce spectra that resemble the observed ones more closely than the present theory predicts.

As mentioned earlier, the light emission process can be conveniently separated into two steps, i.e., generation of SPP modes by tunneling electrons and subsequent emission of light by SPP. In the case of prism-coupled emission, the latter step involves only the radiative fast mode and this step is completely understood. Unresolved questions remain in the first step, namely, the generation mechanism of SPP in tunnel junctions. In the present theory we assumed that fluctuations in the tunneling current act as an oscillating source of radiation. Moreover, we assumed a particular form for the Fourier component of current fluctuations at the frequency ω given by Eq. (35). Although it appears that inclusion of the effects of residual surface roughness into the present structure of the theory may lead to agreement with experimental results, we must remember that the "current fluctuation" theory is merely one possible approach to the problem and other approaches must be considered also. Kirtley *et al.*[8] have proposed that hot electrons injected into the cold film can generate SPP and that this generation mechanism can explain some of the observed facts. However, this "hot-electron" theory has not yet been formulated in a concrete fashion; thus, we cannot assess how the hot-electron effect will influence the spectrum of the emitted light. What is needed now is a carefully formulated theory of the coupling between tunneling electrons and SPP that can describe the generation process of SPP and the resultant population distribution of SPP.

IV. CONCLUSION

We have developed a theory of prism-coupled light emission from tunnel junctions. The form of the theory as well as the underlying assumptions are the same as those used for the theory by Laks and Mills, but we assumed smooth interfaces instead of roughened ones. The angle dependence of the emission intensity at a fixed frequency agrees very well with the experimental result. Thus the kinematics of the theory is correct and it is clear that the observed emission occurs from the fast mode as

expected. However, the calculated emission spectra at different angles do not match the observed spectra. The theory assumes that all SPP modes are excited directly by the fluctuations of tunneling currents, but the comparison with experiment suggests that we should include in the theory an indirect channel of fast-mode excitation by the slow mode via residual surface roughness of the tunnel junction.

We should also keep in mind that other mechanisms of SPP generation than current fluctuations, such as hot electrons, may be playing an important role in determining the population distribution of the generated SSP.

ACKNOWLEDGMENTS

We would like to thank D. L. Mills for valuable discussions and Kenji Sakamoto for assistance in computer programming. The early part of this research was supported by the U.S. National Aeronautics and Space Administration through Grant No. NAG-524, and the last stage of this work was supported by the U.S. National Science Foundation and the Japan Society for Promotion of Science through the U.S.—Japan Cooperative Science Program. Also one of us (S.U.) would like to acknowledge support by a grant from the Murata Science Foundation.

---

*Present address: Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Sendai, Miyagi 980, Japan.

[1]J. Lambe and S. L. McCarthy, Phys. Rev. Lett. **37**, 923 (1976).

[2]An up-to-date list of references is found in P. Dawson, D. G. Walmsley, H. A. Quinn, and A. J. L. Ferguson, Phys. Rev. B **30**, 3164 (1984).

[3]A. Adams, J. C. Wyss, and P. K. Hansma, Phys. Rev. Lett. **42**, 912 (1979).

[4]J. R. Kirtley, T. N. Theis, and J. C. Tsang, Appl. Phys. Lett. **37**, 435 (1980).

[5]S. Ushioda, J. E. Rutledge, and R. M. Pierce, Phys. Rev. Lett. **54**, 224 (1985).

[6]S. Ushioda and Y. Sasaki, Phys. Rev. B **27**, 1401 (1983).

[7]E. Kretschmann and H. Raether, Z. Naturforsch. **23A**, 2135 (1968).

[8]J. R. Kirtley, T. N. Theis, J. C. Tsang, and D. J. Di Maria, Phys. Rev. B **27**, 4601 (1983).

[9]B. Laks and D. L. Mills, Phys. Rev. B **20**, 4962 (1979).

[10]D. L. Mills and A. A. Maradudin, Phys. Rev. B **12**, 2943 (1975).

[11]P. B. Johnson and R. W. Christy, Phys. Rev. B **12**, 4370 (1972).

[12]H. Ehrenreich, H. R. Philipp, and B. Segall, Phys. Rev. **132**, 1918 (1963).

[13]M. A. Ordal, L. L. Long, R. J. Bell, S. E. Bell, R. R. Bell, R. W. Alexander, Jr., and C. A. Wood, Appl. Opt. **22**, 1099 (1983).

[14]K. Arya and R. Zeyher, Phys. Rev. B **28**, 4080 (1983).