PHYSICAL REVIEW B VOLUME 24, NUMBER 10 15 NOVEMBER 1981

Light emission from tunnel junctions on gratings

John Kirtley, T. N. Theis, and J. C. Tsang
IBM T. J. Watson Research Center, Yorktown Heights, New York 10598
(Received 24 April 1981)

We present measurements of light emission from Al-Al_2O_3-Ag tunnel junctions on holographically produced, sinusoidal profile gratings. Sharp angle tunable emission peaks result from coupling out of the Ag-vacuum fast surface-plasmon polariton by the grating structure. The dependences of the emission intensities on angle, frequency, bias voltage and current, and grating amplitudes agree well with a recent theory by Laks and Mills [Phys. Rev. B 22, 5723 (1980)]. The absolute intensities and dependence of intensities on Ag-film thickness do not agree as well. We propose a possible explanation for this discrepancy. PACS numbers 1977 73.40.Rw, 78.65.Ez BD2092

I. INTRODUCTION

Light-emitting tunnel junctions were first demonstrated by Lambe and McCarthy in 1976.[1,2] They biased Al-Al_2O_3-metal tunnel junctions to voltages $V_0=2-4$ V and observed broad-band optical emission. The emission was strongest for highly roughened top metal electrodes. Since roughness is required for surface plasmons to radiate, this implied that surface plasmons were being excited by the tunneling electrons. Further, there was a linear onset of the emission below a critical frequency $\omega_c=eV_0/\hbar$ which Lambe and McCarthy took to mean that the plasmons were excited through an inelastic tunneling process. They reasoned that the bulk of the emission from these devices was radiated from the "slow" or junction mode (see Sec. III) which has field strengths most highly concentrated in the junction region, and therefore is excited most strongly by the tunneling electrons. Several theoretical treatments of light emission from tunnel junctions have subsequently concentrated on radiation from the slow mode.[3]

Recently, Adams, Wyss, Hansma, and Broida[4] have shown that radiation from tunnel junctions can also couple out from localized plasmons. In these experiments the top metal electrode was partially composed of 10–30-nm-diameter, roughly spherical Au particles. At least part of the radition from these devices could be attributed to spherical shape resonances of these particles. These devices have also been theoretically modeled.[5]

Detailed comparison of theory with experiment for randomly roughened devices has been difficult because of incomplete knowledge of their surface topology. For example, theoretical calculations indicate that radiation from the slow modes must be coupled out (at optical frequencies) by roughness with Fourier components of 10 nm or less.[3] While such roughness is probably present in these devices, quantitative analysis is difficult.

We report in this paper experiments in which Al-Al_2O_3-Ag tunnel junctions have been fabricated on holographically produced gratings with periodicities from 120 to 1250 nm.[6] These junctions emit angle tunable, quasimonochromatic radiation at optical frequencies with total external quantum efficiencies comparable to those measured for randomly roughened junctions. The frequency versus emission angle dependence of these devices shows that the radiation results from the Ag-air interface surface-plasmon polaritons (SPP's), which can radiate in the visible by momentum transfer with the grating periodicity. These modes propagate at nearly the velocity of light and are therefore referred to as "fast" plasmons.

Furthermore, (1) a detailed study of the voltage and frequency dependences of the emission intensities from these devices allows us to account for the power spectral densities of the current fluctuations through the junctions to derive a universal "antenna factor" curve. The relative intensities in this curve are in turn well modeled by a theory recently proposed by Laks and Mills (LM).[7] (2) The emission linewidths and dependence of intensities on grating amplitudes also agree well with LM. (3) The experimental absolute intensities, on the other hand, appear to be at least 35 times larger than predicted by LM. (4) The dependence of the emission intensities on the Ag-film thickness L follows the relation $I\sim e^{L/t}$ where t varies from 7.5 to 22

nm depending on the details of the sample preparation procedure. The length t does not correspond to the optical screening length (~ 10 nm) in these fims, as might be expected if the SPP's were excited most strongly in a region close to the tunneling barrier.

We argue that the higher than predicted emission efficiencies, as well as the sample-dependent attenuation lengths, may imply that there is an additional, previously neglected component of the tunneling-electron—surface-plasmon-polariton interaction localized at the Ag-vacuum interface. The parameter t is then interpreted as the hot-electron mean free path in these films. Although our best demonstrated total external quantum efficiencies are $\sim 10^{-6}$, we believe that efficiencies can be improved: Techniques for doing so will be discussed.

II. EXPERIMENTAL TECHNIQUES

The light-emitting tunnel junctions (shown schematically in Fig. 1) were prepared as follows: One-inch-diameter, polished, 110 orientation, silicon wafers were oxidized to 50-nm SiO_2 and coated with 100—200-nm photoresist. The photoresist was holographically exposed[8] by beam splitting the 325-nm line from a HeCd laser, reflecting, spatially collimating, and thereby defocusing the two spherically expanding beams to interfere spatially on the wafer. The variation in grating periodicity across the wafer due to the use of spherical rather than plane-wave interfering beams was extremely small, less than one part in 10^5. The exposed films were then developed using standard techniques. The grating periodicities were varied by adjusting the

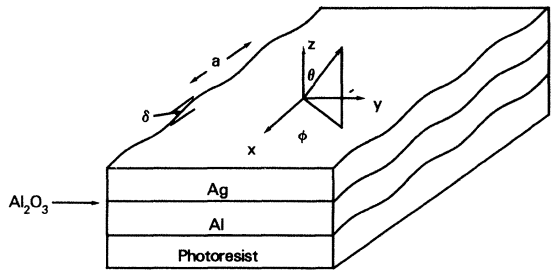

FIG. 1. Sample geometry. Junctions are evaporated on holographically produced, sinusoidal profile gratings with periodicity length a and amplitude δ. The oxide thickness is d and the Ag thickness is L. Standard spherical coordinates are used for the angular measurements.

beam geometry; the grating amplitudes were controlled by varying the exposure and development times. Grating periodicities ranging from 120—1250 nm were produced, with amplitudes ranging from 5—100 nm.

The photoresist films were dried 30 min at 70 °C, and then placed in a standard LN_2 trapped oil diffusion pumped evaporation system. The system had a base pressure of $\sim 10^{-7}$ Torr. All evaporations were done at $< 2 \times 10^{-6}$ Torr with the substrate at room temperature. Five Al strips 2-mm wide, 10-mm long, and 40-nm thick were evaporated from a resistively heated W filament onto the photoresist films. The substrate was than heated to 70 °C in room air (30—70 % relative humidity) for 30 min to oxidize the Al films, then returned to the evaporator for completion of the junctions with a 2-mm-wide, 25-mm-long, 15—60-nm-thick Ag film. Film thicknesses were measured with a calibrated 5-MHz quartz-crystal oscillator. We found that junction resistances of $> 80\ \Omega$ were required for usable light-emitting samples. This often required an additional oxidation of the complete sample in room air for a few days. The oxidation apparently occured through diffusion of oxygen through the Ag films. Junction resistances were measured with a four-terminal technique: It was clear from these measurements that the resistance between the Al and Ag films rather than the film resistances, was increasing. Junctions with thicker Ag films oxidized less rapidly; oxidation could be stopped completely by storing the samples in a dry atmosphere. Eventually all of our samples became usable. We observed no sample degration besides the gradual decrease of the sample conductance. Leads were attached to the films with In solder and the samples were mounted in a closed-cycle refrigerator and cooled to ~ 25 K. All results reported here are for junctions at low temperatures. Cooling was required to maintain sample stability at the highest bias voltages: No changes in emission spectra were observed when samples were warmed to 200 K, but samples warmed to room temperature were roughly 50% less intense than junctions at low temperatures under the same bias conditions. Near the onset frequency for a particular bias voltage, the differences between the warm and cold junction intensities were even larger. The origin of this effect is not yet clear.

The samples were biased at up to 2.7 V with up to 100 mA (dc) with the Al-film biased negative. The junctions suffered irreversible breakdown at much lower applied biases with the opposite polari-

ty. Emission spectra, however, were identical for the two polarity biases if the bias currents I_0 were normalized out. Samples could be biased up to 2.9 V with a 60-Hz square-wave bias ($+V=2.9$ V, $-V=-0.7$ V).[9] The square-wave bias also allowed us to run our sample biases up to 2.5 V at room temperature, but made quantitative normalized intensity data more difficult: All results reported here used dc bias.

The emitted light was passed through an optical window, angularly resolved with a 1.5-mm-wide, 10-mm-high slit 60 mm for the junction plane, and then focused onto the entrance slit of a single-pass Czerny-Turner spectrograph with an 1800 line/mm diffraction grating. The sample was aligned so that the grating modulation of the juncton cross sections was parallel to the resolving slit and the entrance slit of the monochromator. The observed emission angle was varied by rotating the refrigerator with respect to the optics. The polar (θ in Fig. 1) angular resolution was 1.4°, the azimuthal angle (ϕ) was held at 0° within an angular aperture of 9.5°. The solid acceptance angle was 4.2×10^{-3} sr. The monochromator was run with full-width-at-half-maximum frequency resolution of 85 cm^{-1}. Optical throughput was 10% at 632.8 nm. The dispersed light was detected with a cooled RCA 31 034 phototube and photon counting detection. The phototube quantum efficiency was estimated by the manufacturer to be $\sim10\%$. The junction bias voltage and current were continuously monitored while the spectra were being taken: Most of the data presented here was normalized with respect to the total current through the junction.

Grating periodicities were determined by measuring the first-order diffraction angle of the 632.8-nm line of a HeNe laser. Grating amplitudes were determined by measuring the first-order diffraction efficiency of the Al-Ag sandwich in p polarization and fitting it to a first-order perturbation expression by Heitman.[10] This expression is expected to be good up to amplitudes of ~30 nm. We apply it to gratings of somewhat larger amplitude (55 nm) in Sec. IV. Any errors associated with this extrapolation should not affect the qualitative conclusions of this paper.

III. EMISSION FREQUENCIES

The emission from an Al-Al_2O_3-Ag junction on 820-nm periodicity, 5-nm-amplitude photoresist grating, biased at 2.5 V, is shown in Fig. 2. The emission, primarily p polarized, is composed of a broad-band, relatively emission-angle-independent

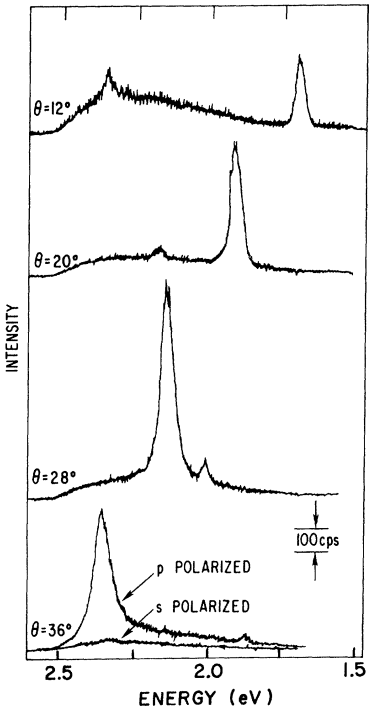

FIG. 2. Light-emission spectra from an Al-Al_2O_3-Ag tunnel junction on a 5-nm amplitude, 820-nm periodicity photoresist grating for various values of the polar angle θ. The large peak corresponds to surface-plasmon polaritons coupling to light to first order in the grating periodicity; the smaller peaks correspond to second-order scattering.

background, upon which is superposed sharp angle-tunable emission peaks. We attribute the background to coupling from residual roughness in the junction structures. The relatively small background intensity—as compared to the sharp emission peaks—as well as the good agreement between calculated and measured dispersion curves for junctions on low-amplitude gratings (as described below), imply that residual roughness in the junction structures did not dominate the radiative properties of our samples. Nevertheless, it should be pointed out that such uncontrolled roughness did exist. The sharp peaks resulted from coupling of the Ag-vacuum fast surface-plasmon polariton to light through the grating periodicity.

We can make this assignment by comparing the experimental frequency-wave-vector dispersion curve with theoretical calculations. Details of the calculations can be found elsewhere[11]: Only the relevant results will be described here. The three lowest energy normal modes of the tunneling junction structure are (see Fig. 3) (1) the slow or junction mode in which the fields are concentrated in

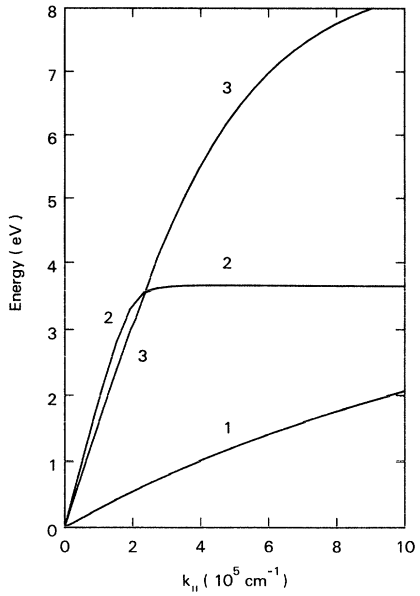

FIG. 3. Calculated surface-plasmon-polariton dispersion curves for the three lowest energy modes in our tunnel junction geometry. Mode 1 is the slow or junction mode, mode 2 is the fast Ag-air interface surface-plasmon polariton, and mode 3 is the fast Al-photoresist mode. The emission from our samples results from mode 2.

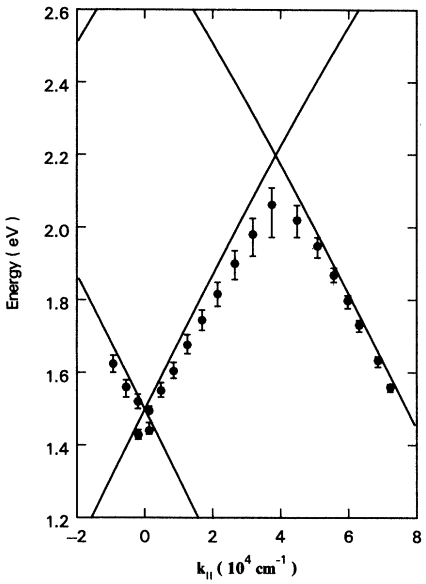

FIG. 4 Measured dispersion curves for a junction on a 81.5-nm amplitude, 815-nm periodicity grating. The error bars are the upper and lower half-maximum intensity points for each peak. The solid curves are theoretical curves as described in the text. The modes on this relatively large-amplitude grating are slowed and broadened.

the junction region, and surface charges on opposite sides of the oxides have opposite signs, (2) the fast Al-photoresist interface mode in which the propagating fields and charges are strongest at the Al-photoresist interface, and (3) the fast Ag-air interface mode. The Al-photoresist mode is significantly lower in frequency below 3 eV than the Ag-air mode because of dielectric screening by the photoresist and Si substrate. The junction mode is in the infrared for emission from grating structures with our periodicities.

The resonance condition for coupling a surface-plasmon polariton of wave vector $k_{||}^{\sigma}$ using the notation of LM (Ref. 7) to light through a grating periodicity $Q=2\pi/a$ is

$$\frac{\omega}{c}\sin\theta\cos\phi=k_{||}^{\sigma}\cos\phi^{\sigma}+\sigma Q, \tag{1}$$

$$\frac{\omega}{c}\sin\theta\sin\phi=k_{||}^{\sigma}\sin\phi^{\sigma}+\sigma Q, \tag{2}$$

where ϕ^{σ} is the azimuthal angle of the surface-plasmon-polariton wave vector with respect to the grating wave vector, θ and ϕ are the polar and azimuthal angles of the emitted photon (see Fig. 1), and $\sigma=\pm1,\pm2$, gives the scattering order. Our measurements were made for $\phi=0$ so the simpler relation

$$k_{||}^{0}\equiv\frac{\omega}{c}\sin\theta=k_{||}^{\sigma}+\sigma Q \tag{3}$$

holds.

In the energy region 1.5—3.0 eV the Ag-air fast-mode dispersion curve is well described by the simple model of a semi-infinite metal [dielectric constant $\epsilon(\omega)$] -vacuum interface. In this case the surface-plasmon-polariton dispersion curve is given by

$$k_{||}^{\sigma}=\frac{\omega}{c}\left[\frac{\epsilon(\omega)}{\epsilon(\omega)+1}\right]^{1/2}. \tag{4}$$

This expression closely follows that light line $\omega=ck$ for Ag with $\hbar\omega<3.7$ eV. The resonance condition [Eq. (3)] can be shown graphically by displacing the dispersion curve by σQ in a plot of $\hbar\omega$ vs $k_{||}^{0}$. We show in Fig. 4 such a plot for a junction on a grating of periodicity $a=815$ nm, amplitude $\delta=81.5$ nm, biased at 2.3 V. The error bars represent the upper and lower half-maximum intensity points for each energy. The diagonal solid lines in Fig. 4 result from substituting (4) into (3), using the Ag dielectric constants of Johnson and Christy.[12] Similar curves generated using the full Al-Ag structure gave indistinguishable lines in this energy range. The experimental data is some-

what lower than the theoretical curves. We believe that is due to grating induced perturbations from the smooth surface behavior as reported by other authors.[13] Note that the deviations from the theoretical curve are largest for the broadest (and presumable most strongly damped) SPP's. As the grating amplitudes were made larger the peaks became more intense, broadened, and shifted to lower energies, especially near the zone center ($k_{||}^{0}=\pi/a$). In addition, there were the expected energy gaps at the crossings of the $\sigma=-1$ branch with the $\sigma=2$ and 3 branches, and the higher-order peaks were always narrower than the lower-order peaks at a given emission peak energy.

The experimental data in Fig. 2 is then easily interpreted: The large peak which moves to higher energies at higher angles results from the $\sigma=-1$ resonance. The small peak which moves to lower energies at higher angles corresponds to $\sigma=2$. The two resonances cross at $\sim 25°$.

IV. INTENSITIES: PLASMON-PHOTON INTERACTION

Interpretation of emission intensities from light-emitting tunnel junctions requires detailed mathematical modeling. Recently, Laks and Mills[7] presented a theory for the light emission from tunnel junctions on gratings with nearly our geometry. In their model only the upper surface of the Ag was modulated by the grating; all interfaces in our junctions are modulated. This would not be expected to cause a significant discrepancy between theory and experiment since most of the surface-plasmon-polariton energy density is located near the modulated Ag-vacuum interface. They used a Green's-function formalism to solve Maxwell's equations for the contribution to the emitted radiation due to a first-order grating perturbation of the dielectric properties of the junction. The currents driving the radiation were the current fluctuations across the tunnel junction, using the simple expression first derived by Hone *et al.*[5]

We present the conclusions of the paper of Laks and Mills[7] in detail in order to show that the present picture of the emission mechanism, or even the most obvious extensions of this picture, while working very well for most of our experimental results, cannnot explain all of them.

LM (Ref. 7) assumed that the current fluctuations extended across the entire Ag film. This would imply that the radiation from the junctions would remain constant in intensity as the Ag films were made thicker. As we will show below, the intensity decreases dramatically with increasing Ag-film thicknesses. We have therefore modified the expression presented by LM to allow the tunneling current fluctuations to decay exponentially into the Ag films. In the notation of LM they assumed that the current fluctuation correlation function $\Delta(z,z')=1$. We take instead,

$$\Delta(z,z')=\begin{cases} e^{K(z+z'-2d)}, & z \text{ and } z'>d \\ e^{K(z+z')}, & z \text{ and } z'<0 \\ 1, & 0<z \text{ and } z'<d \end{cases} \tag{5}$$

with $\Delta(z,z')=0$ otherwise.

With this assumption, Laks and Mills's final expression for the number of photons emitted by the junction per unit time per unit frequency bandwidth per steradian is given by[14]

$$\frac{d^3N}{d\Omega d\omega dt}=\frac{\omega\delta^2|\epsilon_1(\omega)-1|^2}{16\pi\hbar c^3}eI_0\left[1-\frac{\hbar\omega}{eV_0}\right]\cos^2\theta \times\sum_{\sigma}\frac{J(\vec{k}_{||}^{\sigma},\omega)A(\vec{k}_{||}^{\sigma},\omega)k_{||}^{\sigma 4}}{|D(\vec{k}_{||}^{\sigma},\omega)|^2|k_0^{\sigma}|^2}, \tag{6}$$

where δ is the amplitude of the grating, $\vec{k}_{||}^{\sigma}$ is given by Eqs. (1) and (2),

$$J(\vec{k}_{||}^{\sigma},\omega)=|E_x^{>}r_x\cos(\phi-\phi^{\sigma})+E_z^{>}r_z\tan\theta|^2 + |E_x^{>}r_y\sin(\phi-\phi^{\sigma})/\cos\theta|^2, \tag{7}$$

and

$$E_x^{>}=-\frac{k_0^{\sigma}}{k_{||}^{\sigma}}\exp[ik_0^{\sigma}(L+d)], \tag{8}$$

$$E_z^{>}=\frac{1}{\epsilon_1(\omega)}\exp[ik_0^{\sigma}(L+d)], \tag{9}$$

$$k_0^{\sigma}=\left[\left[\frac{\omega+i\eta}{c}\right]^2-k_{||}^{\sigma 2}\right]^{1/2}, \quad \mathrm{Im}(k_0^{\sigma})>0,\ \eta<<\omega \tag{10}$$

$$r_x=n_x/d_x,\quad r_z=n_z/d_x,\quad r_y=n_y/d_y, \tag{11}$$

$$n_x=\frac{k_3}{k_0}\cos(k_1L)\cos(k_2d)+\frac{ik_1\epsilon_3}{k_0\epsilon_1}\cos(k_2d)\sin(k_1L)$$
$$+\frac{ik_2\epsilon_3}{k_0\epsilon_2}\sin(k_2d)\cos(k_1L)-\frac{k_1k_3\epsilon_2}{k_0k_2\epsilon_1}\sin(k_2d)\sin(k_1L), \tag{12}$$

$$n_y=\cos(k_1L)\cos(k_2d)+i\frac{k_3}{k_1}\cos(k_2d)\sin(k_1L)+i\frac{k_3}{k_2}\cos(k_1L)\sin(k_2d)-\frac{k_2}{k_1}\sin(k_1L)\sin(k_2d), \tag{13}$$

$$n_z=\epsilon_3\cos(k_2d)\cos(k_1L)+\frac{i\epsilon_1k_3}{k_1}\cos(k_2d)\sin(k_1L)+\frac{ik_3\epsilon_2}{k_2}\sin(k_2d)\cos(k_1L)-\frac{\epsilon_1k_2\epsilon_3}{k_1\epsilon_2}\sin(k_2d)\sin(k_1L), \tag{14}$$

$$k_i=\left[\epsilon_i(\omega)\left[\frac{\omega}{c}\right]^2-k_{||}^2\right]^{1/2},\quad \mathrm{Im}(k_i<0) \tag{15}$$

$$d_x=\left[\left[\epsilon_3-\frac{k_3}{k_0}\right]\cos(k_2d)\cos(k_1L)+i\left[\frac{\epsilon_1k_3}{k_1}-\frac{k_1\epsilon_3}{k_0\epsilon_1}\right]\cos(k_2d)\sin(k_1L)\right.$$
$$\left.+i\left[\frac{k_3\epsilon_2}{k_2}-\frac{k_2\epsilon_3}{k_0\epsilon_2}\right]\sin(k_2d)\cos(k_1L)+\left[\frac{k_1k_3\epsilon_2}{k_0k_2\epsilon_1}-\frac{k_2\epsilon_1\epsilon_3}{k_1\epsilon_2}\right]\sin(k_1L)\sin(k_2d)\right]e^{ik_0(L+d)}, \tag{16}$$

$$d_y=\left[\left[1-\frac{k_3}{k_0}\right]\cos(k_1L)\cos(k_2d)+i\left[\frac{k_3}{k_1}-\frac{k_1}{k_0}\right]\cos(k_2d)\sin(k_1L)\right.$$
$$\left.+i\left[\frac{k_3}{k_2}-\frac{k_2}{k_0}\right]\cos(k_1L)\sin(k_2d)+\left[\frac{k_1k_3}{k_0k_2}-\frac{k_2}{k_1}\right]\sin(k_2d)\sin(k_1L)\right]e^{ik_0(L+d)}, \tag{17}$$

$$A(k_{||}^{\sigma},\omega)=\frac{1}{ik_3+K}+\frac{1}{k_2}\left[\frac{\epsilon_3}{\epsilon_2}\sin(k_2d)+i\frac{k_3}{k_2}[1-\cos(k_2d)]\right]$$
$$+\frac{1}{2}\left\{\frac{1}{ik_1-K}(e^{ik_1Le^{-KL}}-1)\left[\left[\frac{\epsilon_3}{\epsilon_1}+\frac{k_3}{k_1}\right]\cos(k_2d)+i\left[\frac{k_3\epsilon_2}{k_2\epsilon_1}+\frac{k_2\epsilon_3}{k_1\epsilon_2}\right]\sin(k_2d)\right]\right\} \tag{18}$$
$$-\frac{1}{2}\left\{\frac{1}{ik_1+K}(e^{-ik_1Le^{-KL}}-1)\left[\left[\frac{\epsilon_3}{\epsilon_1}-\frac{k_3}{k_1}\right]\cos(k_2d)+i\left[\frac{k_3\epsilon_2}{k_2\epsilon_1}-\frac{k_2\epsilon_3}{k_1\epsilon_2}\right]\sin(k_2d)\right]\right\},$$

where the subscripts $i=0,1,2,3$ correspond, respectively, to the vacuum, Ag, Al_2O_3, and Al; n_x, n_y, n_z, d_x, and d_y are evaluated with $k_{||}=k_{||}^0=(\omega/c)\sin\theta$, $k_0=k_0^0=(\omega/c)\cos\theta$, and $D(k_{||}^{\sigma},\omega)=d_x$ with the k_i's evaluated for $k_{||}=k_{||}^{\sigma}$ as determined by Eq. (3).

In the expressions above $J(\vec{k}_{||}^{\sigma}\omega)$ is a term containing the angular dependence of the emission intensities, and results from matching the momentum of the surface-plasmon polariton on the surface with that of the emitted photon in free space through the grating periodicity. The first term in $J(\vec{k}_{||}^{\sigma},\omega)$ represents p-polarized radiation (the electric field vector of the exiting photon in the plane defined by the surface normal and the photon wave vector); the second term represents s-polarized radiation (the electric field vector of the exiting photon parallel to the surface). Since our measurements were made at $\phi=\phi^{\sigma}=0$, we only observed the p-polarized radiation. The denominator $D(\vec{k}_{||}^{\sigma},\omega)$ contains the resonance factors and reduces the radiation from the bulk of the junction by electro-

magnetic screening. The factor $A(\vec{k}_{\|}^{\sigma},\omega)$ describes the interaction of the current fluctuations in the junction with the surface-plasmon-polariton fields. It is this factor which we have modified from Laks and Mills's expression to take into account the finite extent of the tunneling current fluctuations in the metal films. We should note at this point that the theory of Laks and Mills is to first order in the grating structure perturbation of the field distributions of the surface-plasmon polaritons. It therefore does not take into account line broadening and lowering of the dispersion curve due to finite grating amplitudes. It also quite naturally does not take into account additional loss mechanisms such as scattering from impurities and residual roughness in the junction structures.

As noted above, for a given junction at bias voltage V_0 and current I_0, the emission peaks moved in energy as the observation angle was changed. As the resonance energy increased the peaks became more intense and broader until the cutoff frequency $\omega_0 = eV_0/\hbar$ was approached. As the bias voltage was increased the peak intensities as well as the cutoff frequency increased. We show in Fig. 5 the integrated intensities for an Al-Al_2O_3-Ag junction on an 815-nm-periodicity grating for five different bias voltages.

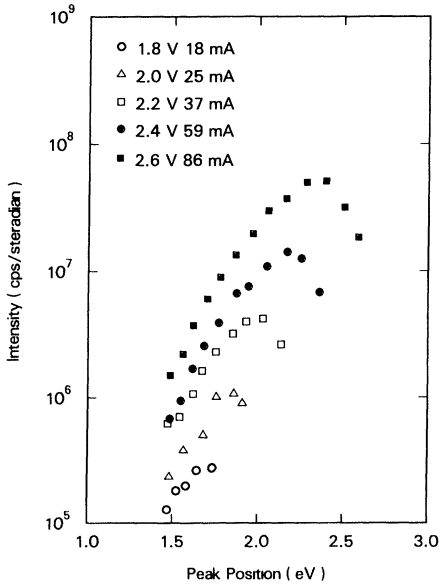

FIG. 5. Integrated intensities for a 20-nm-Al-film thickness junction on a 815-nm periodicity grating, as a function of peak emission energy (for the $\sigma=-1$ peak) for five different bias voltages. As the bias voltage increases the peaks become more intense and extend to a higher-energy linear cutoff. There is a nearly exponential falloff in integrated intensities at low energies.

For a given junction geometry at a given emission angle the bias characteristics of the junction emission are given by the power spectral density of current fluctuations across the tunneling junction.[5]

$$|I(\omega)|^2 = eI_0(1-\hbar\omega/eV_0). \tag{19}$$

Note the appearance of this expression as a factor in Eq. (6). We obtain the emission efficiency, the density of states weighted probability that an electron will be coupled into a photon, by dividing the experimentally measured photon rate by $|I(\omega)|^2/e^2$. The result is shown in Fig. 6 for the junction of Fig. 5. The emission efficiencies fall on a universal curve for all of the bias voltages and currents measured. This should not be surprising since Eq. (19) results from a density of states argument: Electrons losing small energies have more available final states than electrons losing large energies. In particular, these measurements cannot make a distinction between energy loss in the oxide, in the Ag film, or at the Ag surface.

The solid line in Fig. 6 is the theoretical prediction of Laks and Mills, taking the thickness of the oxide $d=3$ nm, the Ag-film thickness 20 nm, $Q=7.2\times10^{-3}$ nm^{-1}, measured dielectric constants for Al (Ref. 15) and Ag (Ref. 12), and setting the dielectric constant of the oxide $\epsilon_2=3$. The theoretical curve is normalized to experiment at 2.3 eV. The agreement for relative intensities be-

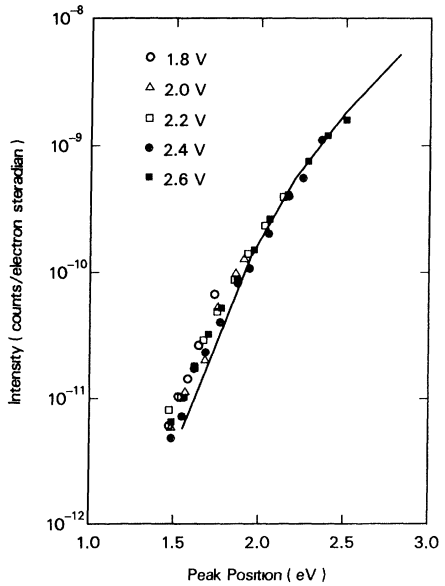

FIG. 6. Integrated intensites of the data of Fig. 5 divided by the power spectral density of current fluctuations. [Eq. (19)]. The data lie on a universal "antenna factor" curve. The solid line is the theory of Laks and Mills normalized to the data at 2.3 eV.

tween theory and experiment is quite good. However, as will be discussed below, the absolute normalization factor used in Fig. 6 is at least 35 times larger than predicted by the theory of Laks and Mills.

The most striking feature of this curve is that the efficiency of the grating as an antenna increases dramatically as the emission frequency increases. This apparently occurs because the surface-plasmon-polariton electric fields become more compacted at the Ag-vacuum interface as the resonance proceeds up the dispersion curve and away from the light line. The grating can then act as a better antenna since more energy is stored at the interface. While other antenna structures, such as those developed by Hansma *et al.*,[4] radiate more efficiently in the red, grating structures may well offer superior performance at higher energies.

One of the consequences of the higher efficiencies for higher-frequency radiation from gratings is that our light-emitting tunnel junctions become much more efficient when biased at higher voltages. We show in Fig. 7 the total predicted external quantum efficiencies (integrated over energy and emission angle), using the theory of Laks and Mills as a function of applied bias for two grating periodicities. These curves have been cut off for $eV > 3.2$ eV because of numerical difficulties. The curves can be expected to saturate for $V \sim 3.7$ V because of the large damping above 3.7 eV in Ag. These curves show graphically the unique nature of these devices: They work much better at large bias voltages and radiate much more strongly in the blue than in the red. We have attained bias voltage of 2.9 V with a square-wave bias: Voltages of 4.0 V have been reported using a pulsed source.[16]

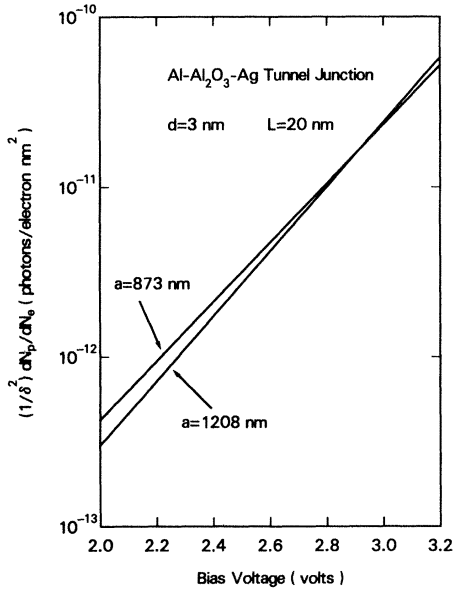

FIG. 7. Predicted total external quantum efficiencies of the theory of Laks and Mills as a function of applied bias for two grating periodicities—$Q = 2\pi/a = 7.2 \times 10^{-3}$ and 5.2×10^{-4} nm^{-1}. These devices are at least 35 times more efficient than predicted, but as predicted become much more efficient as the bias voltage is increased and radiate much more efficiently in the blue than in the red.

The best quantum efficiencies for randomly roughened light-emitting tunnel junctions with Ag-top electrodes were reported to be $\sim 10^{-5}$ at a bias voltage of 3.5 V.[1] Although we could not reliably attain such high biases, we can extrapolate our results at lower voltages by using the slope shown in Fig. 7. For instance, the total external quantum efficiency for a 55-nm grating amplitude junction is $\sim 1 \times 10^{-7}$ with a bias voltage of 2.3 V. If we extrapolate from the curve in Fig. 7, this junction would have an efficiency of $\sim 2 \times 10^{-5}$ at 3.5 V. (This particular junction could not support large bias voltages.) Extrapolation of these curves is dangerous because eventually radiation damping will limit the ultimate quantum efficiency of these devices. However, it appears that external quantum efficiencies comparable to those of randomly roughened devices can be attained with present junction preparation techniques. In addition, the emission from our devices into a particular angle is nearly monochromatic so that the efficiency per steradian per unit bandwidth can be significantly higher than that from randomly roughened devices.

The emission linewidths are also of interest. The observed linewidths varied from sample to sample, presumably varying with details of photoresist preparation, thin-film deposition, and oxidation procedures. We did not establish quantitative relationships with any of these factors although we did find that junctions with larger grating amplitudes produced larger emission linewidths. The narrowest peaks observed were those of the sample of Fig. 2.

Figure 8 shows the measured full-widths-at-half-maximum intensity for this sample as a function of emission peak energy. The solid line is the prediction of the theory of Laks and Mills using the experimentally determined dielectric responses of Ag (Ref. 12) and Al (Ref. 15). These dielectric constants were measured with evaporated films similar in thickness to ours—but on smooth substrates. The additional uncontrolled roughness of

our substrates might be expected to add to the loss in these films. Even the lowest loss sample (that of Fig. 2) had larger bandwidths than theoretically predicted. The additional widths of the emission peaks imply that the SPP resonances and the resulting field in the junction region are significantly weaker than would be expected from a simple first-order calculation. This additional loss may be due to (1) extra impurities or defect structure in the metal films composing the tunnel junctions, or (2) scattering losses from the large-scale grating structure or small-scale uncontrolled roughness. By measuring the emission linewidths we have a direct method for checking on the local-field intensities that can be expected under optical irradiation, a quantity of importance in interpreting the surface enhanced Raman effect.[17] The presence of this additional loss even at very low grating amplitudes implies that the external quantum efficiencies, as well as the efficiencies per steradian per unit bandwidth of these devices can be significantly improved by reducing unwanted roughness and scattering.

The theory of Laks and Mills[7] indicates that the intensity of light emission from our tunnel junctions should be proportional to the square of the grating amplitude. We show in Fig. 9 a comparison of the emission patterns of three junctions with grating periodicities of 1225 nm and grating amplitudes of 25, 44, and 55 nm, respectively. As the grating amplitudes increased the peaks became broader, more intense, and were shifted to slightly lower peak energies (the dispersion curve of the 55-nm junction was ~3% lower than that of the 25-nm junction).

A plot of the integrated efficiencies of a set of junctions on 1225-nm-periodicity gratings with varying amplitudes, and with the current fluctuation power spectral density [Eq. (19)] factored out, is shown in Fig. 10 for emission at 1.74 eV. The emission efficiency is proportional to the first-order diffraction efficiency which should in turn be proportional to the square of the grating amplitude. There is no appreciable saturation of the square-law dependence of integrated intensities on amplitudes up to 55 nm.

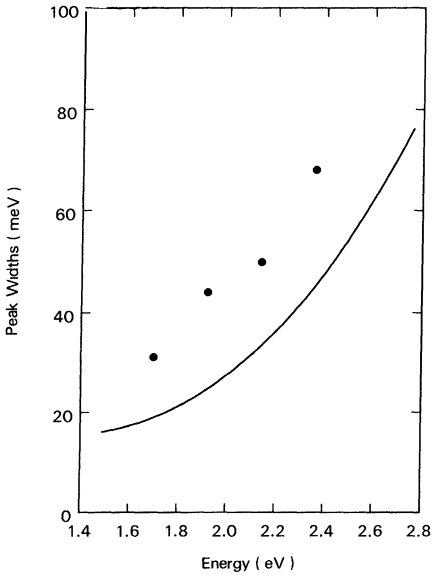

FIG. 8. Measured emission peak widths (full-width-at-half-maximum intensity) for the sample of Fig. 2 (solid circles). The lower solid line is the predicted widths of Laks and Mills assuming experimentally measured dielectric constants for smooth annealed evaporated films. Even for our lowest amplitude gratings the larger emission peak widths indicate that our films have loss mechanisms in addition to those expected for ideal thin films.

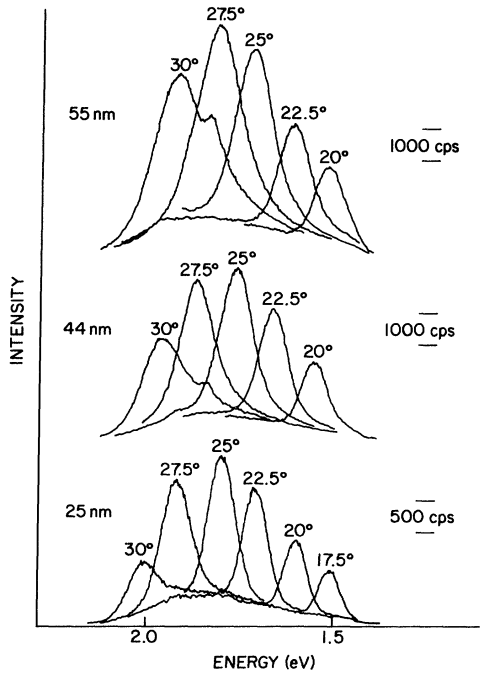

FIG. 9. Relative emission patterns for three junctions on photoresist gratings with 1225-nm periodicity and amplitudes of 25, 44, and 55 nm. As the grating amplitudes become larger than peaks become more intense, broader, and move to slightly lower energies.

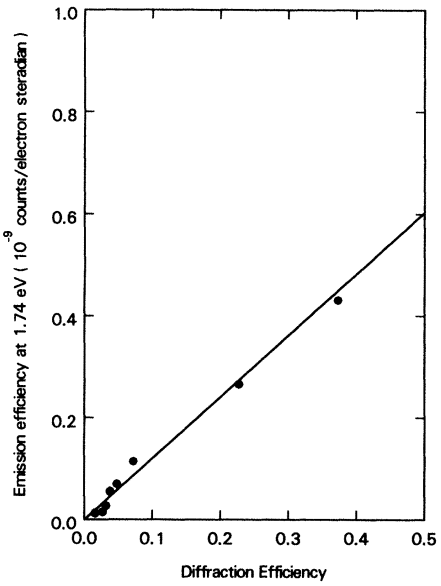

FIG. 10. Integrated emission intensities with the current fluctuation power spectral density [Eq. (19)] normalized out for a set of 1225-nm periodicity gratings with varying amplitudes, as a function of first-order diffraction efficiency. The emission efficiency is proportional to the diffraction efficiency, and therefore proportional to the square of the grating amplitude as expected.

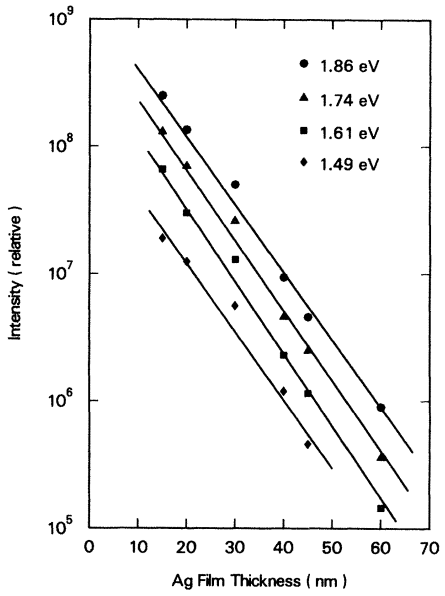

FIG. 11. Normalized emission efficiencies for a set of junctions with varying Ag-film thicknessess L on 815-nm-periodicity photoresist gratings. The emission efficiencies decrease exponentially with L with a characteristic length of ~ 8 nm for all energies.

V. INTENSITIES: ELECTRON-PLASMON INTERACTION

The frequency, angle, bias, and amplitude dependences of our experiments agree well with the theory of Laks and Mills. However, there seems to be a serious discrepancy between the predicted and measured absolute intensities. For example, the curve in Fig. 10 extrapolates to an emission efficiency of 6×10^{-10} counts per electron per steradian at a diffraction efficiency of 0.5. If we account for the total experimental optical collection efficiency of $\sim 1\%$ (see Sec. II), this implies an emitted efficiency of $\sim 6\times 10^{-8}$ photons, per electron per steradian. Using Heitman's[10] prediction for the grating amplitude at this diffraction efficiency (63 nm) and the full theory of Laks and Mills (setting the tunneling current fluctuation length $=\infty$), we find a predicted efficiency of 1.63×10^{-9} photons per electron per steradian for the appropriate experimental parameters. This means that our tunneling junctions are at least 35 times more efficient than predicted by theory. The theoretical number is an upper limit since for this calculation we have taken an infinitely long tunneling current fluctuation decay length [$K=0$ in Eq. (18)]. As we will see below, experiment shows that the current fluctuations have finite extent. For the reasonable case of current flucutation decay lengths of 20 nm, the predicted intensities become 10 times smaller because the tunneling currents do not extend to the region where the SPP fields are strongest.

Another discrepancy resulted when we studied the dependence of the emission intensity on Ag-film thickness. This was done by fabricating five junctions per Si wafer with different counter-electrode thicknesses on photoresist films with uniform grating amplitudes (within 10% as measured by the diffraction efficiency). Figure 11 shows the resultant relative intensities with all junctions biased at 2.3 V and with the current fluctuation power spherical density [Eq. (19)] normalized out. At all peak energies the emission intensity decreased exponentially ($I\sim e^{-L/t}$) as the Ag-film thickness (L) increased; these samples produced a characteristic length of $t\sim 8$ nm. To within experimental error this length was independent of emission-peak energy.

There are two important lengths involved in modeling the Ag-film thickness dependence of light-emission intensities if one assumes that the interaction occurs within the junction volume. The first is the current fluctuation attenuation length: As this length gets shorter the tunneling electrons

interact with the surface-plasmon-polariton fields in a smaller volume and in a region with weaker fields. The second length is the optical screening length: As this length gets shorter the fields which interact with the tunneling electrons become weaker. The crucial point is that the emission intensity cannot fall off more rapidly than the slowest of the two falloff rates if the interaction is in the bulk of the films. For example, if the current fluctuation attenuation length was very short and the entire interaction was within the oxide region, the emission intensity would decrease with increasing Ag thickness like the optical screening length. On the other hand, if the current fluctuation length was very much longer than the optical screening length it would dominate the emission falloff with Ag thickness. Figure 12 shows the predictions of the theory of Laks and Mills for tunneling current fluctuation decay lengths (assumed the same for both Al and Ag films) of ∞, 20,5, and 2.5 nm, at a peak emission energy of 2.5 eV. For comparison the measured ballistic mean free path of a hot electron 2 eV above the Fermi surface in a thin Ag film is ~ 18 nm.[18] Just as expected, these model

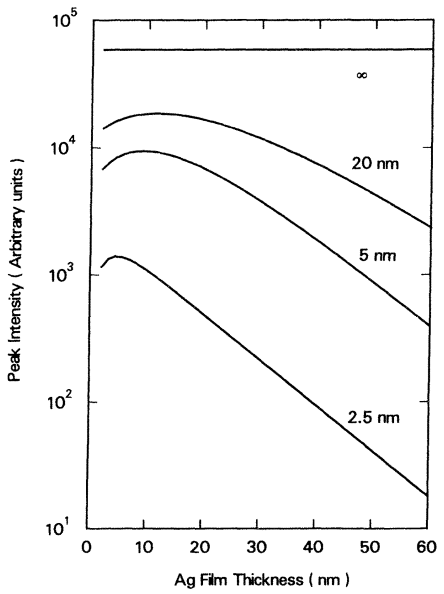

FIG. 12. Predicted dependence of emission efficiencies for the theory of Laks and Mills modified as described in the text to account for the finite extent of the tunneling current fluctuations in the junctions, for current fluctuation attenuation lengths of ∞,20,5, and 2.5 nm. In all cases the attenuation length is at least as long as the electromagnetic screening length (~ 12 nm in these films), and much longer than the lengths measured for the sample of Fig. 10.

curves never fall off more rapidly than the optical screening length of ~ 12 nm.

This analysis of Ag-film thickness dependence is made simple because of the additional scattering losses in our samples. Because Al is much more lossy than Ag at the frequencies of interest, the full theory of Laks and Mills predicts that junctions with relatively thin layers of Ag, having relatively more energy density in the Al should have broader emission lines than junctions with thick Ag. If that were the case, the analysis of emission intensities as a function of Ag-film thickness would be complex. However, for our samples the dispersion curves and emission linewidths did not vary with Ag-film thickness for $15 < L < 60$ nm. This meant that the losses did not depend strongly on the relative thicknesses of Al and Ag in the sandwich. Under these conditions the strength of the resonance does not depend on Ag-film thickness and modeling becomes relatively easy. We can then ignore the Ag thickness dependence of the resonance denominator $D(\vec{k}_{||}^{\sigma},\omega)$ (aside from an electromagnetic screening factor). The thickness dependence is then described by the tunneling-electron-surface-plasmon-polariton-field coupling factor $A(\vec{k}_{||}^{\sigma},\omega)$ which we have modified to take into account the finite attenuation length of the current fluctuations in the junction [Eq. (5)].

The experimental results of Fig. 11 show that the emission intensity does fall off more rapidly than the optical screening length. This can be understood if we hypothesize that the coupling of the tunneling electrons to surface-plasmon polaritons is predominantly localized at the Ag-vacuum interface as opposed to the junction region, or in the bulk of the Ag film. In that case the exponential falloff in emission intensity with Ag-film thickness can be dominated by the finite hot-electron mean free path rather than an algebraic sum of this length with the optical screening length. Then the relatively short mean free paths of the sample of Fig. 11 could result from small film-grain sizes.[19]

To test this hypothesis we tried to make Ag films with larger grain sizes. The junctions were fabricated on gratings etched directly into the Si substrate. The substrate was then oxidized to 80 nm of SiO_2 resulting in exceptionally smooth grating surfaces. (Small-scale roughness observable in scanning electron microscope pictures of the photoresist gratings may provide nucleation sites, producing additional grain boundaries and thereby limiting hot-electron mean free paths.) In addition the Ag films were evaporated rapidly (at a rate of

1—2 nm/sec as opposed to 0.2—0.4 nm/sec for the samples of Fig. 11). The rapid evaporation could also be expected to improve film grain sizes.[20]

The integrated emission efficiencies for such a sample on a 1.1-μm periodicity grating on Si is presented in Fig. 13. The characteristic falloff lengths are 2—2.5 times longer than for the sample for Fig. 11. There may in addition be some energy dependence to the decay lengths: Lower frequency emissions appear to fall off more rapidly in intensity with Ag-film thickness than higher-frequency emissions. The characteristic length measured for our best films, ~20 nm, agrees well with 18 nm for the ballistic mean free paths of hot electrons 2 eV above the Fermi surface in Ag, as reported by Crowell and Sze.[18]

Recently, Kroo *et al.*[21] suggested that measurements of the thickness dependence of intensities from nominally smooth Al-Al_2O_3-Ag junctions, which correlated with theoretical estimates of the mean free paths of ballistic hot electrons in these films, could be interpreted to imply that (1) the radiation was dominated by emission from the fast mode as opposed to the junction mode, and (2) that their experiments were indeed measuring the hot-electron ballistic mean free path in these films.

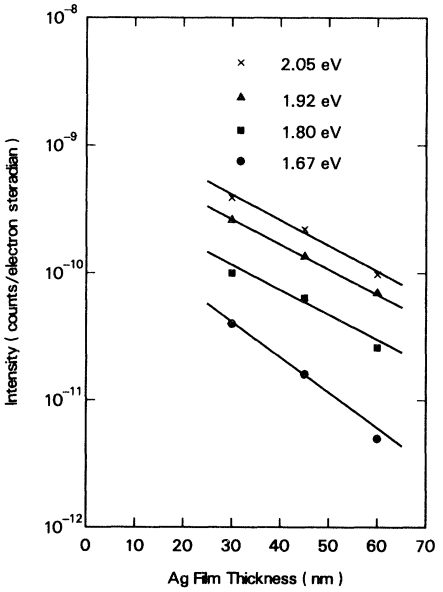

FIG. 13. Relative emission efficiencies for junctions on a 1.1-μm Si grating as a function of Ag-film thickness. The relatively longer attenuation lengths (~20 nm) than measured for the sample of Fig. 10 are attributed to improved film quality. These lengths agree with previous measurements of hot-electron mean free paths in Ag.

They obtained curves of intensity versus Ag-film thickness that gave a characteristic length t of ~16 nm. The difficulty with these experiments is that Kroo *et al.* had no control of their surface roughness and therefore had no control over which mode radiated or how strongly. Since we have controlled roughness coupling and include only the light emission due to the first Fourier component of the grating structure (the backgrounds were subtracted out in all cases), there is no such difficulty in our experiments.

There is also a major difference between our interpretation and that of Kroo *et al.* In their analysis they state that the surface-plasmon-polariton fields of the Ag-vacuum fast mode extend only a few nm into the metal. If this were true than measurements of light emission versus Ag-film thickness for samples with controlled geometry could be unambiguously associated with the hot-electron mean free paths. However, as very elementary calculations show, the field penetration depths are tens of nm's. As our modeling shows, the anomalous thickness dependence observed can only be explained in terms of the hot-electron mean free paths if the tunneling electrons interact most strongly, not with the surface-plasmon-polariton fields which penetrate throughout the junction structure, but with the field gradients which are largest at the Ag-vacuum interface.

A qualitative argument for strong tunneling-electron—surface-plasmon-polariton interaction at the Ag-air interface is as follows. The transition matrix element for a tunneling electron to emit a surface-plasmon polariton is proportional to[3]

$$\int \vec{j}_{fi}(\vec{r})\cdot\vec{A}(\vec{r})d^3r, \tag{20}$$

where

$$\vec{j}_{fi}(\vec{r})=\psi_f^*(\vec{r})\nabla\psi_i(\vec{r})-\psi_i(\vec{r})\nabla\psi_f^*(\vec{r}). \tag{21}$$

ψ_i and ψ_f are the initial and final electronic wave functions and $\vec{A}(\vec{r})$ is the vector potential associated with the surface-plasmon polariton.

Present theoretical treatments of light-emission intensities have neglected the spatial variation of the vector potential $\vec{A}(\vec{r})$ since the wavelength of light is much longer than the wavelength of the electrons. However, at the Ag-air interface the fields vary rapidly over atomic dimensions, implying that the surface term could predominate, especially when one considers that the field strength at the metal surface is $\epsilon_1 \sim -15$ times larger (and of opposite sign) than the field inside the metal. A similar large surface effect has recently been report-

ed in photoemission from aluminum.[22]

The addition of such a surface term to theories of light emission from tunnel junctions would be expected to increase the predicted absolute intensities and to change the dependence of intensities on counter-electrode thickness, but would not be expected to modify the angular, voltage, frequency, or grating amplitude dependences, all of which agree well with the theory of Laks and Mills.

VI. CONCLUSIONS

We have shown that light-emitting tunnel junctions fabricated on gratings with optical scale periodicities radiate semimonochromatic, angle tunable, optical radiation. This radiation originates from the Ag-air interface fast surface-plasmon polariton as can be demonstrated by measuring the dispersion curve of the emitted radiation. These experiments are the first to show conclusively that the radiation results from a collective electromagnetic resonance of the junction structures. The fact that the fast mode can radiate relatively efficiently shows that it should be considered in the study of emission from randomly roughened samples.

The theory of Laks and Mills for radiation from these junctions works well for the coupling of surface-plasmon polaritons to light: the dispersion curves, angular and frequency dependences of the intensities, and to some extent the linewidths all seem to be reasonably well understood. In addition, the expression for the power spectral density of current fluctuations through the junction [Eq. (19)] works well in producing a universal "antenna curve" for these devices. However, the mechanism for coupling of the tunneling electron to the surface-plasmon polaritons may not be as well understood: Both the absolute value and the Ag-film thickness dependence of emission intensities do not agree well with expeiment. We suggest that the coupling may be dominated by an additional surface term.

While these devices have extremely low quantum efficiencies at present, there are several improvements that can be made: (1) larger grating amplitudes—efficiencies so far have shown no indication of saturation as the coupling becomes stronger, and (2) improved film properties—present linewidths are much larger than the intrinsic linewidths expected from smooth annealed films.

Aside from the possibility of developing light-emitting tunneling junctions into useful display devices, they present a new, unique technique for measuring the hot-electron energy distributions in thin-film structures. For this application the fabrication of these devices on gratings is essential since only then is the out-coupling understood well enough to infer the energy distributions.

Optical emission has also been observed from Ag gratings irradiated optically[23,24] and with high-energy electron beams.[25] It is suggestive that for all three methods of excitation (tunnel injection, optical irradiation, and electorn beam irradiation), the total external quantum efficiencies are similar ($\sim 10^{-6}$). We suggest that the same process is occurring in all three cases: Hot electrons are excited, they relax in energy through the competing mechanisms of photon and plasmon emission, and the plasmons radiate light through the grating structure.

ACKNOWLEDGMENTS

We would like to thank Doug Mills for numerous enlightening conversations and for sending data prior to publication. We would also like to thank A.M. Torreson for sample preparation, G. A. Waters for ion milling the silicon wafers, and C.F. Aliota for scanning electron microscopy, our colleague J.H. Janak for several discussions, and P.K. Hansma, R.C. Dynes, H.J. Levinson, and J.J. Quinn for useful conversations.

---

[1]John Lambe and S.J. McCarthy, Phys. Rev. Lett. 37, 923 (1976); S.L. McCarthy and John Lambe, Appl. Phys. Lett. 30, 427 (1977).

[2]Light emission from metal-insulator-metal structures had previously been studied by T.W. Hickmott, J. Appl. Phys. 36, 1885 (1965); W. Pong, C. Inouye, F. Matunaga, and M. Morikawa, *ibid.* 46, 2310 (1975); Tien-Lai Hwang, S.E. Schwarz, and R.K. Jain, Phys. Rev. Lett. 36, 379 (1976).

[3]L.C. Davis, Phys. Rev. B 16, 2482 (1977); Bernardo Laks and D.L. Mills, *ibid.* 20, 2962 (1979); 21, 5175 (1980).

[4]P.K. Hansma and H. P. Broida, Appl. Phys. Lett. 32, 545 (1978); Arnold Adams, J.C. Wyss, and P.K. Hansma, *ibid.* 32, 545 (1978).

[5]Daniel Hone, B. Muhlschlegel, and D.J. Scalapino, Appl. Phys. Lett. 33, 203 (1979); R.W. Rendell, D.J. Scalapino, and B. Muhlschlegel, Phys. Rev. Lett. 41, 1746 (1978).

[6]A preliminary report has appeared in J. R. Kirtley, T. N. Theis, and J. C. Tsang, Appl. Phys. Lett. 37, 435 (1980).

[7]Bernardo Laks and D.L. Mills, Phys. Rev. B 22, 5723 (1980).
[8]I. Pockrand, Phys. Lett. 49A, 259 (1974).
[9]We would like to thank R.C. Dynes for suggesting the square-wave modulation to us.
[10]D. Heitman, Opt. Commun. 20, 292 (1977).
[11]E.N. Economou, Phys. Rev. 182, 539 (1969).
[12]P.B. Johnson and R.W. Christy, Phys. Rev. B 6, 4370 (1972).
[13]D. Heitman and H. Raether, Surf. Sci. 59, 17 (1976); R. H. Ritchie, A.T. Arakawa, J.J. Cowan, and R.N. Hamm, Phys. Rev. Lett. 21, 1530 (1968).
[14]These expressions correct several typograhical errors that appeared in Ref 7.
[15]C.J. Powell, J. Opt. Soc. Am. 60, 78 (1970).
[16]K. Parvin and William Parker (unpublished).
[17]A recent calculation of field enhancements under light irradiation of metals on gratings appears in S.S. Jha, J.R. Kirtley, and J.C. Tsang, Phys. Rev. B 22, 3973 (1980).
[18]C.R. Crowell and S.M. Sze, *Physics of Thin Films* (Academic, New York, 1967), Vol IV, p. 325.
[19]R.S. Sennett and G.D. Scott, J. Opt. Soc. Am. 40, 203 (1950).
[20]M.M. Dujardin and M.L. Theye, J. Phys. Chem. Solids 32, 2033 (1971).
[21]N. Kroo, Zs. Szentirmay, and J. Felszerfalvi, Phys. Status Solidi B 102, 227 (1980).
[22]Harry J. Levinson, E.W. Plummer, and Peter J. Feibelman, Phys. Rev. Lett. 43, 952 (1979); Harry J. Levinson and E. W. Plummer (unpublished).
[23]J.C. Tsang, J.R. Kirtley, and T.N. Theis, Solid State Commun. 35, 667 (1980).
[24]A. Girlando, W. Knoll, and M.R. Philpott (unpublished).
[25]D. Heitman, J. Phys. C 10, 397 (1977).