Long-range surface modes supported by thin films

Fuzi Yang, J. R. Sambles, and G. W. Bradberry
Thin Film and Interface Group, Department of Physics, University of Exeter, Exeter, Devon EX4 4QL, United Kingdom
(Received 22 March 1990; revised manuscript received 7 June 1991)

A detailed analysis of the surface modes of a thin slab of material of dielectric constant ϵ_2 ($=\epsilon_{r2}-i\epsilon_{i2}$) surrounded symmetrically by dielectric media is presented. Results show that in the thin-film limit, as well as the well-known long-range surface plasmon for a thin metal layer and the TM guided mode for a thin dielectric, a long-range surface mode exists for almost any value of ϵ_2. This is even true if the imaginary part of ϵ_2, ϵ_{i2}, is much larger than the real part ϵ_{r2}. We also find that a long-range surface mode may arise from the coupling between two surfaces which individually cannot support a surface mode. These are a pair of special coupled-surface modes which may exist below a certain critical film thickness and which have two separate propagation vectors each with the same field symmetry. It is also found that the inverse situation may pertain, that is for certain relative values of dielectric constants even though ordinary surface modes may exist, below a critical thickness the resulting coupled long-range mode no longer exists. The analysis has also been extended to practical situations with weakly absorbing surrounding media and to circumstances where the dielectric constants of the surrounding media are slightly different. Both of these effects modify the dispersion relations obtained for the simple case and introduce further limit thicknesses into the problem. Analytic formulas in the thin-film limit are presented for all the above situations and field distributions and energy flow (Poynting vector) profiles presented to illustrate as necessary the nature of the modes supported by these systems. Finally experimental results are presented which illustrate the rather sweeping conclusion that a long-range surface mode may exist on a thin film for almost all values of ϵ_{r2} and ϵ_{i2}. This result paves the way for a range of optics experiments on absorbing structures.

I. INTRODUCTION

A surface electromagnetic wave is defined simply as an electromagnetic wave that propagates along the interface between two media and whose amplitude decays exponentially with increasing distance from the interface into the two media. The existence or otherwise of such surface waves is established through the use of Maxwell's equations and appropriate boundary conditions. In the case of a *p*-polarized disturbance (TM wave) whose magnetic vector lies in the plane of the interface and perpendicular to the direction of propagation there will be a discontinuity in the component of the electric field normal to the interface. This results in a surface charge density and hence the surface electromagnetic mode is an electromagnetic wave coupled to these surface charges. At the interface of two nonabsorbing media the surface electromagnetic waves which may exist are generally known as Fano modes,[1] while for one medium with absorption the surface waves are called Zenneck modes.[2]

In general, in condensed-matter physics interfaces between two media can lead to a range of possible surface states involving the relevant quasiparticles (electrons, phonons, excitons, magnons, etc.). However if, as is normally the case, the decay distance of the surface vibration amplitude and the wavelength are much greater than the distance between atoms or the relevant distance scale for the particular surface mode, then such surface vibrations may be treated phenomenologically as simple alternating dipole moments. This then allows a macroscopic electrodynamic treatment of the media which are described adequately by a simple frequency-dependent dielectric constant $\epsilon(\omega)$. In the regime of frequencies close to ω_0 such that $\epsilon(\omega_0)$ displays a strong resonance, it is conventional to describe the electromagnetic surface wave as a coupled mode which is an admixture of the electromagnetic field and the elementary excitation in the medium which gives rise to the resonance in $\epsilon(\omega_0)$.[3] Such electromagnetic waves are therefore commonly called surface polaritons and we then find surface plasmon-polaritons, surface exciton-polaritons, etc. As we shall see later this description is somewhat misleading, as should be obvious when we try to examine the existence regime in ω, for a given resonance, for all resonances have a strong effect over a wider range in frequency than that often considered in the above descriptions.

Another general concept of solid-state physics and optics, often stated explicitly or often implicitly assumed in the literature, is that of a surface-active medium. It is commonly asserted that a material with a negative real part of its dielectric constant ($\epsilon_{r2}<0$) has the capacity to cause a surface polariton at the interface with a second medium of positive dielectric constant (ϵ_{r1}), provided that the inequality $|\epsilon_{r1}|>\epsilon_{r1}$ is satisfied. Actually this is only true for the Fano modes—those surface modes with no loss, in which case the dispersion equation describing the surface wave has only real components. In general, of course, all dielectrics have some imaginary component to ϵ; this is obviously true of the surface-active medium since associated with any resonance in $\epsilon_{r2}(\omega)$ there will

be a resonance in $\epsilon_{i2}(\omega)$. This then gives a Zenneck mode and the dispersion equation describing the surface modes has complex components. It has often been assumed that the imaginary part ϵ_i may be ignored, and this may well be true for the surface-inactive medium, but for most of the situations in which there is a surface-active medium giving rise to perhaps a highly negative ϵ_r it will be impossible to ignore the influence of the absorbing effect on the surface electromagnetic wave. Indeed, some situations arise, as we show below, in which a large ϵ_{i2} is vital to the existence of the surface mode.

In Sec. II of this paper it is shown how a medium having some absorption at a frequency ω, i.e., $\epsilon(\omega)$ has a nonzero imaginary part ϵ_{i2}, must be "surface active" and even though ϵ_{r2} does not satisfy the above criteria it can support a surface mode between itself and a second medium with a real ϵ. It is in fact independent of whether the real part of the dielectric constant of the "surface-active" medium is positive, negative, or even zero although when the positive real part of the dielectric constant is much larger than its imaginary part, the Zenneck mode almost becomes a purely radiative mode, that is, a Brewster wave.[4]

Recently research using the attenuated coupling method has shown that highly absorbing materials do indeed support sharp surface mode resonances.[5–7] In all these studies the surface-active medium has a large imaginary part ϵ_{i2} with its real part either positive or negative and also $|\epsilon_{r2}| > \epsilon_1$ or $|\epsilon_{r2}| < \epsilon_1, |\epsilon_{r2}| \ll \epsilon_{i2}$ or $|\epsilon_{r2}| \sim \epsilon_{i2}$. These studies therefore take the conventional wisdom of a surface-active medium having a large negative ϵ_r into the more general idea involving complex ϵ values. It would of course be argued that for highly absorbing systems the surface mode is likely to be a very broad resonance. While this is indeed generally true for the single interface mode, more complex geometries, and in particular the next simplest geometry of two such modes coupled together on a thin absorbing film, do not give broad resonances.

Consider a thin film of surface-active medium (now a material with ϵ_i) surrounded symmetrically by dielectric media with real positive dielectric constants. The two surface modes may couple across the thin film. This then results in two mixed modes which show dispersion with film thickness. One mode is symmetric, that is, the parallel component of magnetic field, which is labeled H_y, does not exhibit a zero inside the film, the other is antisymmetric which does give a zero in H_y inside the film. The absorption in the surface-active film results in a finite decay distance for both modes. However, for the truly symmetric system and indeed for small enough differences between the dielectric constants of the media bounding the film the attenuation of the symmetric mode decreases to zero as the film thickness decreases and the wave vector of this mode approaches that of a plane wave in one of the surrounding media. In this case, the symmetric mode is generally termed the long-range surface mode not only because the attenuation of this mode is less than that of a general single surface mode but also because there is a second antisymmetric mode which has a greater attenuation and is correspondingly called the short-range surface mode. As the film thickness decreases the propagation distance of the long range-branch increases while that of the short-range branch decreases. In absolute terms "long range," which implies a propagation distance much greater than the wavelength, will depend on the wavelength used and the particular geometry and it is really only in the limit of the film thickness tending to zero with correspondingly the propagation length tending to infinity that the term long range should most rigorously be used.

There have been several studies of both long-range and short-range modes for very thin *metal* films.[8–14] The short-range mode has been observed in highly antisymmetric structures with large differences between the dielectric constants of the surrounding media.[8–10] While for almost symmetric structures both modes may be observed and the decrease in the attenuation of the symmetric mode with film thickness confirmed.[11–14]

Ferguson, Wallis and Hauvet[15] have evaluated the dispersion curve as a function of metal film thickness for a film surrounded asymmetrically by glass and air. They find that, as well as a leaky Fano mode, above a certain thickness of metal there is a bound surface mode which transforms to a growing wave below some cutoff thickness. At the cutoff thickness the solution is very much like a Brewster field in that were a plane wave incident upon the layer it would be totally absorbed. As is seen later this cutoff thickness arises in more general situations not necessarily involving a metal film.

Burke, Stegman, and Tamir[16] have examined in some detail the case of a lossy metal film bounded by dissimilar dielectrics and find that the geometry can support a total of four waves. Leaky waves are predicted in regions of the dispersion curves usually associated only with bound modes.

The long-range modes which may be supported by thinner films are of substantial interest both theoretically and experimentally. This arises largely from their associated high local-field enhancement which has potential in nonlinear optics applications. There is already a substantial body of work examining nonlinear optics using long-range surface plasmons[17–22] (LRSP) including both second- and third-order nonlinear interactions. Some theoretical studies[23,24] have also been undertaken to try to find a configuration which would give a very long propagation distance with its associated strongly enhanced local field. As well as these LRSP studies there is also some work on the long-range surface phonon-polariton[25] and the long-range surface magneto-plasmons.[26] In all the above studies it is imagined that the surface-active medium must have a large negative real part to its dielectric constant and that the inequality $|\epsilon_{r2}| > \epsilon_1$ must be satisfied. Recently[27] it has been shown, as indicated earlier, that a long-range surface mode may be supported on a thin film which has a large imaginary part to its dielectric constant, the real part being very nearly zero. This coupled long-range mode has been labeled the long-range surface exciton polariton (LRSEP) since it is primarily in the vicinity of an excitonic reso-

nance that the condition $|\epsilon_{r2}| \ll \epsilon_{i2}$ may be found.

In Sec. III of this paper a systematic analysis of the long-range surface modes available with a thin film having any dielectric constant ϵ_{r2} $(=\epsilon_{r2}-i\epsilon_{i2})$ surrounded symmetrically by dielectric media having real dielectric constants $(\epsilon_1=\epsilon_3)$ is analyzed. It is found that as well as well as the normal long-range surface polaritons (plasmon, phonon, magnon, etc.) which exist for $\epsilon_{r2}<0$, $|\epsilon_{r2}| \ll \epsilon_{i2}$, $|\epsilon_{r2}|>\epsilon_1$, and the optical TM guided modes which are found for $\epsilon_{r2}>0$, $\epsilon_{r2} \gg \epsilon_{i2}$, $|\epsilon_{r2}|>\epsilon_1$ there are in fact long-range modes which exist for almost any values of ϵ_2 with the exception of a small area in the complex plane of $\epsilon_{r2}, \epsilon_{i2}$. In particular it is found that when $\epsilon_{i2}=0$, $\epsilon_{r2}<0$ but $|\epsilon_{r2}|<\epsilon_1$ *there is still a long-range coupled surface mode even though a single interface mode does not exist for this system.* This is a special surface mode created only by the coupling of the two surfaces, rather like the TM guided modes; it is not a mode created by the two surface modes coupling together. Even more surprisingly, when the film thickness is less than some critical value there are two coupled modes with the same field symmetry yet different propagation vectors for a single thickness.

The examination of the symmetrically surrounded thin film is extended in Sec. IV to the asymmetric case with $\epsilon_1 \neq \epsilon_3$, particularly for $|\Delta|=|\epsilon_1-\epsilon_3| \ll \epsilon_1, \epsilon_3$. This, of course, takes the idealized symmetric system much nearer to practical reality. The third step in this logical development then introduces in Sec. V finite but small absorption in the surrounding media while maintaining symmetry $(\epsilon_1=\epsilon_{r1}-i\epsilon_{i1}$, $\epsilon_{i1} \ll \epsilon_{r1}$, $\epsilon_1=\epsilon_3)$.

In the penultimate section of the paper, Sec. VI, we present some experimental results which verify the general conclusions that the long-range surface mode may exist for almost any value of complex dielectric constant of the surface active film. Finally, in Sec. VII some general conclusions and discussions are presented.

II. SURFACE ELECTROMAGNETIC MODES SUPPORTED BY A SINGLE INTERFACE

The geometry under discussion is that of two semi-infinite media, 1 and 2, joined at a planar interface. For simplicity we shall limit all media to those having relative magnetic permeability 1 and a dielectric function which has no spatial dispersion. The coordinate axes are chosen so that the z axis is perpendicular to the interface with $z=0$ corresponding to the interface. The x axis is in the surface-wave propagation direction while the orthogonal y axis lies in the interface plane. Medium 1 is a pure nonabsorbing dielectric, $\epsilon_1>0$, and medium 2 is a surface-active medium $\epsilon_2=\epsilon_{r2}-i\epsilon_{i2}$ $(\epsilon_{i2}>0)$.

For isotropic media the surface electromagnetic waves are transverse magnetic and are best described by their magnetic-field component which lies in the plane of the interface and in the y direction:

$$\mathbf{H}=H_0 f(z)\exp[i(\omega t-kx)]=H_y \ , \tag{1}$$

where $k=k_r-ik_i$ (note the negative sign to be consistent with ϵ and $+i\omega t$) is the complex propagation constant parallel to the interface and $f(z)$ describes the dependence of $\mathbf{H}$ on the distance away from the interface $[f(0)=1]$, and H_0 is a normalization constant. In the normal way the electric-field components may be calculated using Maxwell's equations: i.e.,

$$E_x=\frac{i}{\omega\epsilon}\frac{\delta H_y}{\delta z}, \quad E_z=\frac{-k}{\omega\epsilon}H_y \ . \tag{2}$$

We then write the dependence $f(z)$ for the two media in the form

$$f(z)=e^{-\alpha_1 z}, \quad z>0, \quad \text{medium 1} \ , \tag{3}$$

$$f(z)=e^{\alpha_2 z}, \quad z<0, \quad \text{medium 2} \ , \tag{4}$$

where the coefficients α_1 and α_2 are obtained from the wave equations as

$$\alpha_1^2=k^2-k_0^2\epsilon_1 \tag{5a}$$

and

$$\alpha_2^2=k^2-k_0^2\epsilon_2 \ , \tag{5b}$$

where $k_0=\omega/c$ is the free-space wave vector at frequency ω. Continuity of tangential $\mathbf{E}$, E_x, at $z=0$ leads to the dispersion relation

$$\epsilon_1\alpha_2+\epsilon_2\alpha_1=0 \ . \tag{6}$$

This equation has to be satisfied if a surface mode is to exist. Also for a nonradiative surface mode it is clear from Eqs. (3) and (4) that $\mathrm{Re}(\alpha_1)>0$ and $\mathrm{Re}(\alpha_2)>0$. There are then two simple situations to consider.

(a) If $\epsilon_{i2}=0$, that is, the surface-active medium has no absorption (a physically unrealizable situation but one often assumed in the past), then $k_1=0$. Then $\epsilon_2=\epsilon_{r2}$ and $k=k_r$ so Eq. (6) becomes an equation containing only real components. We must then have $\epsilon_{r2}<0$ and $|\epsilon_{r2}|>\epsilon_1$ for a solution to exist:

$$k_r=k_0\left[\frac{\epsilon_1\epsilon_{r2}}{\epsilon_{r2}+\epsilon_1}\right]^{1/2}=k_0\left[\frac{\epsilon_1\epsilon'_{r2}}{\epsilon'_{r2}-\epsilon_1}\right]^{1/2}, \tag{7a}$$

where $\epsilon_{r2}=-\epsilon'_{r2}$, $\epsilon'_{r2}>0$.

This is the very simple pure ideal Fano mode dispersion relation, which gives the decay constants α_1 and α_2 as purely real constants of the form

$$\alpha_1=k_0\epsilon_1\left[\frac{1}{\epsilon'_{r2}-\epsilon_1}\right]^{1/2}, \tag{7b}$$

$$\alpha_2=k_0\epsilon'_{r2}\left[\frac{1}{\epsilon'_{r2}-\epsilon_1}\right]^{1/2} . \tag{7c}$$

(b) If $\epsilon_{i2}\neq 0$, then ϵ_2, k, α_1, and α_2 are all complex, so Eq. (6) leads to more elaborate solutions of the form

$$k_r=k_0\left[\frac{\epsilon_1}{(\epsilon_{r2}+\epsilon_1)^2+\epsilon_{i2}^2}\right]^{1/2}\left[\frac{\epsilon_e^2+(\epsilon_e^4+\epsilon_1^2\epsilon_{i2}^2)^{1/2}}{2}\right]^{1/2}, \tag{8a}$$

$$k_i = k_0 \left[\frac{\epsilon_1}{(\epsilon_{r2}+\epsilon_1)^2+\epsilon_{i2}^2} \right]^{1/2} \frac{\epsilon_{i2}\epsilon_1}{[2(\epsilon_e^2+(\epsilon_e^4+\epsilon_1^2\epsilon_{i2}^2)^{1/2})]^{1/2}} , \tag{8b}$$

where

$$\epsilon_e^2 = \epsilon_{r2}^2+\epsilon_{i2}^2+\epsilon_1\epsilon_{r2} . \tag{8c}$$

Also α_1 and α_2 are both complex, leading to oscillating decaying fields in both half spaces.

Let us now examine Eqs. (8a)–(8c) for a range of different situations.

(1) If $\epsilon_{r2} = -\epsilon'_{r2} < 0$, $\epsilon'_{r2} \gg \epsilon_{i2}$, $\epsilon'_{r2} > \epsilon_1$, then from Eqs. (8a) and (8b) we find

$$k_r \simeq k_0 \left[\frac{\epsilon_1\epsilon_{r2}}{\epsilon_1+\epsilon_{r2}} \right]^{1/2} = k_0 \left[\frac{\epsilon_1\epsilon'_{r2}}{\epsilon'_{r2}-\epsilon_1} \right]^{1/2} > k_0\epsilon_1^{1/2} , \tag{9a}$$

$$k_i \simeq k_0 \frac{\epsilon_1\epsilon_{i2}}{2(\epsilon'_{r2}-\epsilon_1)} \left[\frac{\epsilon_1}{\epsilon'_{r2}(\epsilon'_{r2}-\epsilon_1)} \right]^{1/2} . \tag{9b}$$

This is then a much more realistic surface plasmon-polariton having $k_r > k_0\epsilon_1^{1/2}$ and with k_i proportional to ϵ_{i2}. As is well known for a surface plasmon on increasing ϵ'_{r2}, k_r tends to the limit value $k_0\epsilon_1^{1/2}$ and k_i decreases towards zero.

(2) If $\epsilon_{r2}=0$ then we find[27]

$$k_r = k_0 \left[\frac{\epsilon_1\epsilon_{i2}[\epsilon_{i2}+(\epsilon_1^2+\epsilon_{i2}^2)^{1/2}]}{2(\epsilon_1^2+\epsilon_{i2}^2)} \right]^{1/2} < k_0\epsilon_1^{1/2} \tag{10a}$$

and

$$k_i = k_0 \left[\frac{\epsilon_1^3\epsilon_{i2}}{2(\epsilon_1^2+\epsilon_{i2}^2)[\epsilon_{i2}+(\epsilon_1^2+\epsilon_{i2}^2)^{1/2}]} \right]^{1/2} \tag{10b}$$

which may be labeled a surface exciton-polariton.[5,6] For this situation if also $\epsilon_{i2} \gg \epsilon_1$ (which is the case for a frequency close to a dielectric resonance, for example, close to the transverse exciton frequency ω_T of the surface-active medium) then

$$k_r \to k_0\epsilon_1^{1/2}$$

and

$$k_i \to k_0 \frac{\epsilon_1^{3/2}}{2\epsilon_{i2}}$$

so that surprisingly k_i is inversely proportional to ϵ_{i2}. This is quite the reverse of the dependence found in Eq. (9b) for the surface-plasmon mode.

Notice from Eqs. (9a) and (10a) that while the surface plasmon has $k_r > k_0\epsilon_1^{1/2}$ the surface exciton has $k_r < k_0\epsilon_1^{1/2}$ so a relationship between ϵ_{r2}, ϵ_{i2}, and ϵ_1 must exist for which $k_r = k_0\epsilon_1^{1/2}$. Examination of Eq. (8a) gives this relationship as

$$\epsilon_{r2} = \epsilon_{i2}^{2/3} \left\{ \left[\frac{\epsilon_1}{8} + \left[\frac{\epsilon_1^2}{64} + \frac{\epsilon_{i2}^2}{27} \right]^{1/2} \right]^{1/3} + \left[\frac{\epsilon_1}{8} - \left[\frac{\epsilon_1^2}{64} + \frac{\epsilon_{i2}^2}{27} \right]^{1/2} \right]^{1/3} \right\} - \epsilon_1 . \tag{11}$$

When this equation is satisfied the real part of the wave vector of the surface mode just equals the bulk wave vector in the simple dielectric. If $\epsilon_{i2} \gg \epsilon_1$, then for $k_r = k_0\epsilon_1^{1/2}$ Eq. (11) gives $\epsilon_{r2} \simeq -0.75\epsilon_1$. This special condition may be achieved in the vicinity of an excitonic resonance close to ω_T.

For fixed ϵ_1 and ϵ_{i2} the dependence of k_r and k_i on ϵ_{r2} is shown in Fig. 1. From this figure it is clear that provided $\epsilon_{i2} \neq 0$, no matter how small, in principle a nonradiative surface mode exists, that is, k_i is finite and both α_1 and α_2 have real positive components. Thus over the

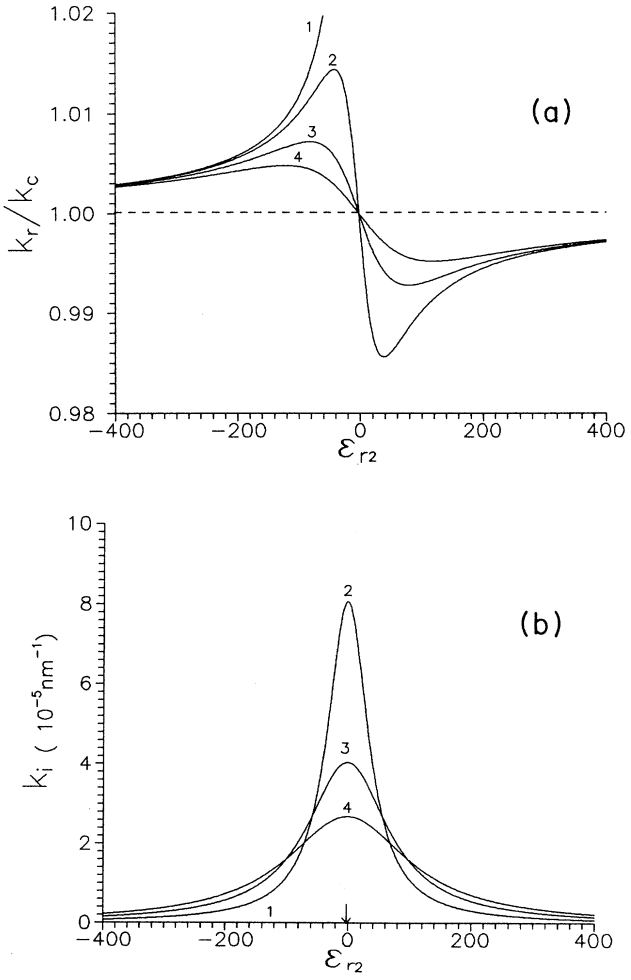

FIG. 1. Dispersion relations for a single interface surface wave. ($\lambda_0 = 3.391$ μm, $\epsilon_1 = 2.3$, $k_c = (2\pi/\lambda_0)\epsilon_1^{1/2} = 2.81$ μm^{-1}.) Curve 1, $\epsilon_{i2}=0$; curve 2, $\epsilon_{i2}=40$; curve 3, $\epsilon_{i2}=80$; curve 4, $\epsilon_{i2}=120$. (a) k_r/k_c vs ϵ_{r2}, (b) k_i vs ϵ_{r2}.

whole range of ϵ_{r2} (from $-\infty$ to $+\infty$) even though the real part of the wave vector of the surface mode, k_r, may be less than $k_0\epsilon_1^{1/2}$, a nonradiative surface mode may exist. This means that the light line, specified by $k=k_0\epsilon_1^{1/2}$ only separates radiative and nonradiative regions of the dispersion curve for $\epsilon_{i2}=0$, i.e., Fano modes. Although it is clear that when ϵ_{i2} is very small and ϵ_{r2} is positive and large enough the Zenneck modes ($k_r<k_0\epsilon_1^{1/2}$) are almost Brewster waves, because $\mathrm{Re}(\alpha_2)$ is very small and $\mathrm{Im}(\alpha_2)$ is large with α_1 tending to zero. Workers[5,6] have shown examples of the excitation of nonradiative surface modes (Zenneck modes) in the region $k_r<k_0\epsilon_1^{1/2}$.

Of course, all of the above is well established; we have placed it within this text as an introduction to the main ideas found in the following sections. It has also, we hope, clarified some of the slight misconceptions which exist regarding the rather rigid separation between surface plasmons, Fano modes, Zenneck modes, surface excitons, etc. These are often just limit concepts which are convenient ideas but which become less clearly distinguishable for real systems with complex ϵ's and complex propagation constants.

III. LONG-RANGE SURFACE MODES IN A SYMMETRIC GEOMETRY

The geometry which is analyzed in the following is that of a thin parallel-sided slab specified by ϵ_2 surrounded by semi-infinite media specified by ϵ_1 and ϵ_3, in this symmetric case $\epsilon_1=\epsilon_3$. Now, of course, the thickness d of the region 2 provides another variable in the problem being involved in the linking of the two surfaces together. Solving boundary conditions as before it is readily shown that the dispersion relation for this geometry is

$$\tanh(\alpha_2 d)=-\frac{\epsilon_2\alpha_2(\epsilon_1\alpha_3+\epsilon_3\alpha_1)}{\epsilon_1\epsilon_3\alpha_2^2+\epsilon_2^2\alpha_1\alpha_3}\ , \tag{12}$$

$$\alpha_j^2=k^2-k_0^2\epsilon_j,\quad j=1,2,3 \tag{13}$$

and

$$k=k_r-ik_i \tag{14}$$

as before.

For bound surface modes, we now require both $\mathrm{Re}(\alpha_1)>0$ and $\mathrm{Re}(\alpha_3)>0$ remembering that the depth distribution in region 3 has the form $f(z)=e^{\alpha_3 z}$ with $z<0$. When the system is symmetric, $\epsilon_1=\epsilon_3$, then Eq. (12) is split into two separate branches, one which is antisymmetric in H_y and has the form[28]

$$\tanh\left[\frac{\alpha_2 d}{2}\right]=\frac{-\epsilon_1\alpha_2}{\epsilon_2\alpha_1} \tag{15}$$

and a second, symmetric in H_y, which has the form

$$\tanh\left[\frac{\alpha_2 d}{2}\right]=\frac{-\epsilon_2\alpha_1}{\epsilon_1\alpha_2}\ . \tag{16}$$

For the long-range branch the primary interest centers on Eq. (16) which when $d\to 0$ has the solution $\alpha_1\to 0$. This is therefore the long-range mode. If d is not zero yet still small enough so that $|\frac{1}{2}\alpha_2 d|\ll 1$, then from Eq. (16) it is simple to expand to first order giving

$$\tfrac{1}{2}\alpha_2 d\simeq\frac{-\epsilon_2\alpha_1}{\epsilon_1\alpha_2} \tag{17}$$

so that the dispersion relation for the long-range coupled surface mode which exists in this thin film limit is (Appendix A)

$$k_r\simeq k_0\epsilon_1^{1/2}\left[1+\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})^2-\epsilon_1^2\epsilon_{i2}^2}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2}\right], \tag{18a}$$

$$k_i\simeq k_0\epsilon_1^{1/2}\epsilon_1^2\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2}\ , \tag{18b}$$

where we have introduced λ_0, the vacuum wavelength ($2\pi/k_0$) as the important scaling length for d. Note that the expressions both contain terms quadratic in ($\pi d/\lambda_0$), for k_r this is a correction while for k_i it is the dominant term.

Consider Eqs. (18a) and (18b) for the following circumstances.

A. Symmetrically surrounded real metal, $\epsilon_{r2}<0, \epsilon_{i2}\neq 0$

If $\epsilon_{r2}<0$ then

$$k_r\simeq k_0\epsilon_1^{1/2}\left[1+\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{(|\epsilon_{r2}|+\epsilon_1)^2}{\epsilon_{r2}^2}\right]>k_0\epsilon_1^{1/2}\ , \tag{19a}$$

$$k_i\simeq k_0\epsilon_1^{1/2}\epsilon_1^2\left[\frac{\pi d}{\lambda_0}\right]^2\frac{|\epsilon_{r2}|+\epsilon_1}{|\epsilon_{r2}|^3}\epsilon_{i2}\ . \tag{19b}$$

This is the case for a real metal and we see that k_i is proportional to ϵ_{i2}. These equations for k_r and k_i represent the realizable long-range surface plasmon for a real metal,[29] in the thin film limit given by $|\frac{1}{2}\alpha_2 d|\ll 1$.

B. Symmetrically surrounded excitonic absorber, $\epsilon_{r2}=0$

If $\epsilon_{r2}=0$ then with $\epsilon_{i2}>\epsilon_1$ we have

$$k_r\simeq k_0\epsilon_1^{1/2}\left[1+\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}^2-\epsilon_1^2}{\epsilon_{i2}^2}\right]>k_0\epsilon_1^{1/2}\ , \tag{20a}$$

$$k_i\simeq k_0\epsilon_1^{1/2}\epsilon_1^2\left[\frac{\pi d}{\lambda_0}\right]^2\epsilon_{i2}^{-1}\ . \tag{20b}$$

This is a special mode which may be found near a strong resonance feature of the active medium. For example, at the transverse exciton resonance $\epsilon_{r2}=0$ and ϵ_{i2} may be quite large. This new resonance which we label the long-range surface exciton[27] (LRSE) has the interesting feature that in this thin-film limit its half width, given by k_i is inversely proportional to ϵ_{i2}. Therefore, the stronger the absorption in the active thin layer the sharper the resonance. The H_y field and Poynting vector distribution for

such a mode show almost total exclusion of power from the film. Further the Poynting vector in the film is only along the z axis.

A situation found near a strong resonance where ϵ_{r2} is again zero but ϵ_{i2} is quite small (for example, at the longitudinal exciton frequency, ω_L) leads to a mode which, in contrast to the LRSE described above, has $k_r < k_0\epsilon_1^{1/2}$ which may also be labeled an LRSE. Once again the decay constant is inversely proportional to ϵ_{i2}, but now with $\epsilon_{i2} < \epsilon_1$ it is a much weaker resonance and is only truly a long-range mode when $d \ll \lambda_0$.

C. Symmetrically surrounded nonabsorbing dielectric, $\epsilon_{r2} > 0$

If $\epsilon_{r2} > 0$ and $\epsilon_{i2} = 0$ then

$$k_r \simeq k_0\epsilon_1^{1/2}\left[1+\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{(\epsilon_{r2}-\epsilon_1)^2}{\epsilon_{r2}^2}\right] > k_0\epsilon_1^{1/2} \tag{21a}$$

and

$$k_i \simeq 0 \ . \tag{21b}$$

In this case ϵ_{r2} has to be greater than ϵ_1 for a bound mode to exist (see Equation (A19) in Appendix A). This of course is none other than the well-known case for an ideal TM guided mode.[30] Because in this situation there is no single surface mode for such parameters [see Eq. (7)] then this long-range wave ($k_i \to 0$) arises purely from the coupling of the fields at the two surfaces of the thin film. Of course there is no corresponding antisymmetric mode.

D. Symmetrically surrounded absorbing dielectric $\epsilon_{r2} > 0, \epsilon_{i2} > 0$

If $\epsilon_{r2} > 0$ and $\epsilon_{i2} \ll \epsilon_{r2}$, then we have the practically realizable TM guided-mode situation which may be further subdivided into several different situations.

(1) If $\epsilon_{r2} > \epsilon_1$ and $\epsilon_{r2}(\epsilon_{r2}-\epsilon_1) \gg \epsilon_{i2}^2$, then k_r is still given by Eq. (21a) and

$$k_i \simeq k_0\epsilon_1^{1/2}\epsilon_1^2\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}(\epsilon_{r2}-\epsilon_1)}{\epsilon_{r2}^3} \ . \tag{22}$$

These expressions describe a normal strong waveguide with weak absorption.

(2) If $\epsilon_{r2} > \epsilon_1$ and $\epsilon_{r2}(\epsilon_{r2}-\epsilon_1) \sim \epsilon_{i2}^2$, the case for a weak waveguide with some absorption, then we need to use Eqs. (18a) and (18b).

(3) If $\epsilon_{r2} = \epsilon_1$, then

$$k_r \simeq k_0\epsilon_1^{1/2}\left[1-\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{(\epsilon_1^2-1)\epsilon_{i2}^2}{(\epsilon_1^2+\epsilon_{i2}^2)^2}\right] < k_0\epsilon_1^{1/2} \ , \tag{23a}$$

$$k_i \simeq k_0\epsilon_1^{1/2}\epsilon_1^2\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}^3}{(\epsilon_1^2+\epsilon_{i2}^2)^2} \ . \tag{23b}$$

(4) If $\epsilon_{r2} < \epsilon_1$ but $(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2}) > 0$, that is $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2} > \epsilon_1$, there is *still* a guided mode. From Eq. (A9) in Appendix A it is clear that $\mathrm{Re}(\alpha_1) > 0$. This case and that described by Eqs. (23a) and (23b) are extensions of the conventional waveguide picture. In conventional waveguide theory the requirement for a symmetric waveguide is $\epsilon_2 > \epsilon_1$, but this is strictly only true for the Fano mode. When $\epsilon_{r2} \leq \epsilon_1$, provided that $\epsilon_{i2} \neq 0$, then if $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2} > \epsilon_1$ is satisfied, there is a TM guided mode. So, provided d is small enough to satisfy Eq. (17), as $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2}-\epsilon_1$ tends to zero, because the optical fields penetrate deeper into the media on either side of the absorber, the loss progressively reduces to zero [see Eq. (18b)].

When $(\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2}) < \epsilon_1$ it is clear that $k_i < 0$ and the mode transforms to a growing mode[15,16] [$\mathrm{Re}(\alpha_1) < 0$ from Eq. (A9)] in the symmetric geometry. Because the growing surface mode is not a bound long-range surface mode and is dependent on externally incident fields supplying energy to make the total wave amplitude grow, we do not discuss this mode further. However, it is worth noting that with $\epsilon_{r2} > 0$ and $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2} < \epsilon_1$ because $\epsilon_{i2} \neq 0$ there is still a single interface mode which has now been destroyed by the coupling of the two surfaces, the converse to the situation discussed in III C above.

It is worthwhile examining the magnitudes of k_r and k_i for a possible practical case. Consider $\lambda_0 = 3.391$ μm (corresponding to a He-Ne laser line) with $d = 10$ nm, $\epsilon_1 = \epsilon_3 = 2.30$, $\epsilon_{r2} = 2.0$, and $\epsilon_{i2} = 0.8$. With these values $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2} = 2.32 > \epsilon_1 (= 2.30)$ we now expect a long-range surface mode even though $\epsilon_{r2} < \epsilon_1$. From the approximation, Eq. (18), we find the mode is very close to the critical momentum

$$\Delta k_r = k_0\epsilon_1^{1/2} - k_r = 44 \ \mathrm{m}^{-1} \ ,$$

$$k_i = 1.90\times 10^{-6} \ \mu\mathrm{m}^{-1} \ ,$$

which corresponds to a long-range mode with very low loss. Exact evaluation of Eq. (16) gives

$$\Delta k_r = 44 \ \mathrm{m}^{-1}$$

and

$$k_i = 1.90\times 10^{-6} \ \mu\mathrm{m}^{-1} \ ,$$

which shows that in this thin-film limit Eq. (18) is a very good approximation. If now we reduce ϵ_{i2} to 0.7, then because $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2} = 2.25 < \epsilon_1$, Eq. (18) no longer gives a bound solution. However, we can use Eq. (8) in Sec. II to find a single surface mode having $k_r = 1.95$ $\mu\mathrm{m}^{-1}$ and $k_i = 1.72\times 10^{-1}$ $\mu\mathrm{m}^{-1}$. This is now a strongly damped mode. Nevertheless its existence shows us that there has to be some cutoff thickness d^* such that when $d < d^*$ no symmetric bound mode exists for the above constraints on the ϵ values. The behavior for $(\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2}-\epsilon_1) \sim 0$ is complicated and higher-order approximations need to be introduced.

E. Symmetrically surrounded metal $|\epsilon_{r2}| < \epsilon_1$

For $\epsilon_{r2} < 0$ and $\epsilon_{i2} = 0$, which is the same as case (1), we now examine the influence of d for $|\epsilon_{r2}| < \epsilon_1$. From Eqs. (19a) and (19b) it is clear that we do not need $|\epsilon_{r2}| > \epsilon_1$ for this coupled mode. Yet in Sec. II, Eq. (7) the single sur-

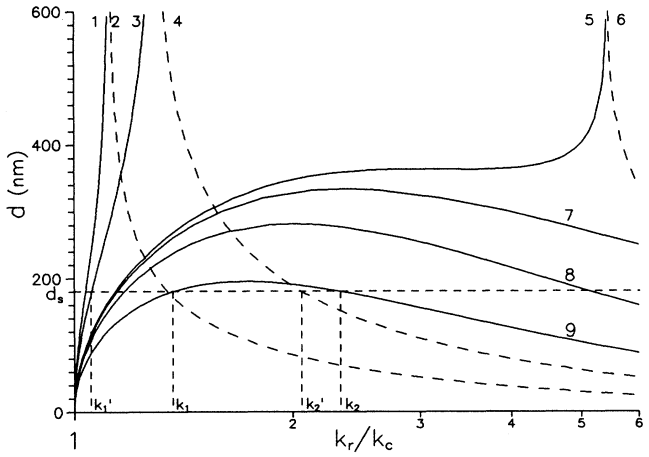

FIG. 2. Relationship between k_r/k_c and the film thickness d. ($\lambda_0=3.391$ μm, $\epsilon_1=\epsilon_3=2.3$, $\epsilon_{i2}=0$, $k_c=(2\pi/\lambda_0)\epsilon_1^{1/2}=2.81$ μm^{-1}.) Curves 1 and 2, $\epsilon_{r2}=-11.5$; curves 3 and 4, $\epsilon_{r2}=-5.75$; curves 5 and 6, $\epsilon_{r2}=-2.38$; curve 7, $\epsilon_{r2}=-2.3$; curve 8, $\epsilon_{r2}=-2.07$; curve 9, $\epsilon_{r2}=-1.50$. Solid lines and dashed lines are, respectively, for symmetric and antisymmetric modes in H_y.

face mode may only exist if $|\epsilon_{r2}|>\epsilon_1$. Thus we have a situation where a long-range mode is supported by two surfaces neither of which can independently support a mode.

We may use Eqs. (15) and (16) to examine this situation with $\epsilon_{i2}=0$ and $\epsilon_{r2}=-\epsilon_{t2}'$ ($\epsilon_{r2}'>0$). Equation (15) then becomes

$$\tanh[\tfrac{1}{2}(k_r^2+k_0^2\epsilon_{r2}')^{1/2}d]=\frac{\epsilon_1}{\epsilon_{r2}'}\left[\frac{k_r^2+k_0^2\epsilon_{r2}'}{k_r^2-k_0^2\epsilon_1}\right]^{1/2} \tag{24}$$

and since $\epsilon_{r2}'<\epsilon_1$ the right-hand side of this equation must be greater than unity so no solution may exist. This means there is no antisymmetric mode for this situation, just as for the TM guided mode case. However, Eq. (16) becomes

$$\tanh[\tfrac{1}{2}(k_r^2+k_0^2\epsilon_{r2}')^{1/2}d]=\frac{\epsilon_{r2}'}{\epsilon_1}\left[\frac{k_r^2-k_0^2\epsilon_1}{k_r^2+k_0^2\epsilon_{r2}'}\right]^{1/2}, \tag{25}$$

which has solutions with the limitation that, because the right-hand side is always less than unity, d is limited and cannot go to infinity, that is, no single surface mode exists for these conditions.

When $d\to 0$ Eq. (25) has two solutions, $k_r\to k_0\epsilon_1^{1/2}$ and

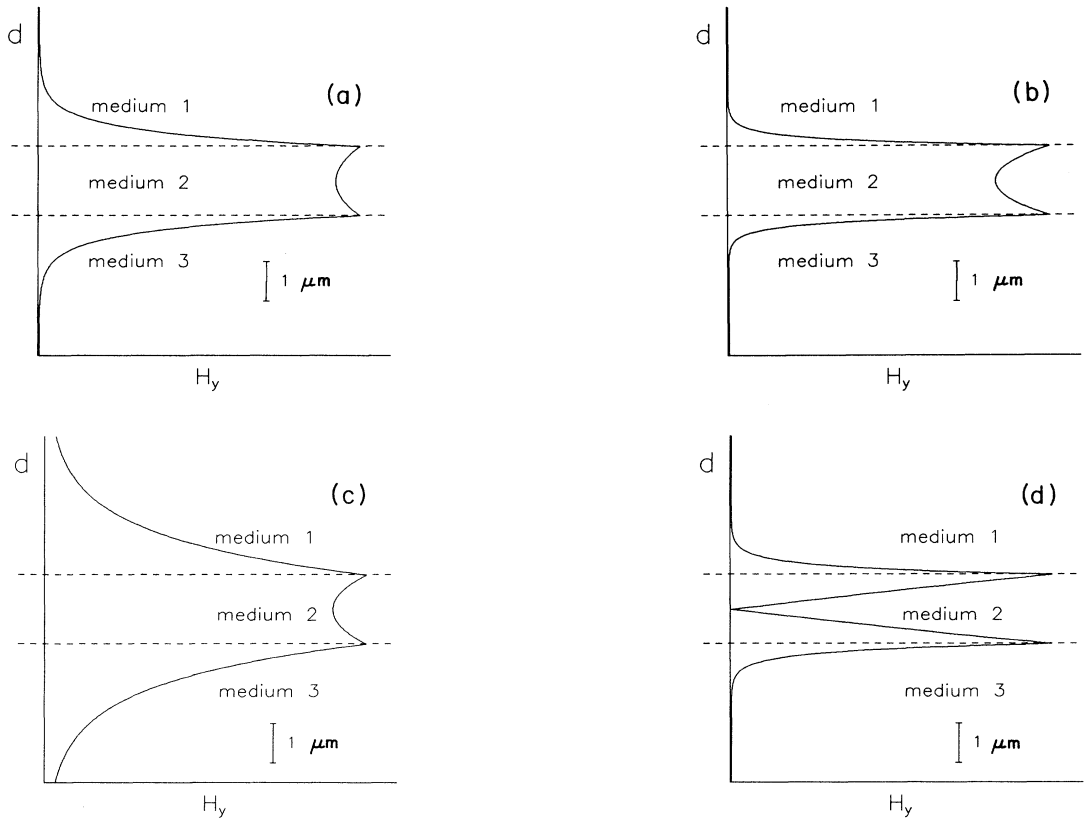

FIG. 3. The H_y field distributions for the points marked k_1, k_2, etc. for the same d, d_s in Fig. 2. ($\lambda_0=3.391$ μm, $\epsilon_1=\epsilon_3=2.3$, $\epsilon_{i2}=0$, $d_s=180$ nm.) (a) $k_1=3.84$ μm^{-1}, $\epsilon_{r2}=-1.50$, (b) $k_2=6.57$ μm^{-1}, $\epsilon_{r2}=-1.50$, (c) $k_1'=2.97$ μm^{-1}, $\epsilon_{r2}=-5.75$, (d) $k_2'=5.79$ μm^{-1}, $\epsilon_{r2}=-5.75$ (d relatively scaled by 10×).

$k_r \to \infty$ (for which $k_r d$ remains finite), this implies that there is some maximum value of d, $d_{\max}$, below which there are two modes having the same symmetry in H_y for a single d value. This simply means for any given d below $d_{\max}$ there are two values of k_r which satisfy

$$d=\frac{2}{(k_r^2+k_0^2\epsilon_{r2}')^{1/2}}\tanh^{-1}\left[\frac{\epsilon_{r2}'}{\epsilon_1}\left[\frac{k_r^2-k_0^2\epsilon_1}{k_r^2+k_0^2\epsilon_{r2}'}\right]^{1/2}\right]. \tag{26}$$

In Fig. 2 we illustrate solutions for this equation where we have also included for comparison purposes other curves calculated for $|\epsilon_{r2}|>\epsilon_1$. This shows the evolution of the curves as $|\epsilon_{r2}|$ approaches ϵ_1 and then finally $|\epsilon_{r2}|<\epsilon_1$. For the case of $|\epsilon_{r2}|<\epsilon_1$ we see that when $d=d_s(<d_{\max})$ we have solutions k_1 and k_2 for which the H_y field distributions are displayed in Figs. 3(a) and 3(b). In order to allow comparisons of these with the LRSP and short-range surface plasmons (SRSP) fields we also show in Figs. 3(c) and 3(d) the H_y distributions for the LRSP at k_1' and the SRSP at k_2'.

In general then, provided d is small enough to satisfy the approximation required to give Eq. (18), then a long-range surface mode exists for any value of ϵ_2 except for a small area described by $(\epsilon_{r2}^2+\epsilon_{i2}^2)<\epsilon_1\epsilon_{r2}$ as shown in Fig. 4. The limit to this area is a semicircle of radius $\epsilon_1/2$ centered at $\epsilon_{r2}=(\epsilon_1/2)$, $\epsilon_{i2}=0$. In fact close to this boundary higher-order terms are needed to describe the situation and the boundary is not exactly a semicircle.

In Figs. 5(a) and 5(b) we show both k_r and k_i as functions of d for several different cases, using Eqs. (15) and (16). For all these cases as d tends to zero a long-range mode is found with k_r near $k_0\epsilon_1^{1/2}$ and k_i near zero. To serve to illustrate this even more fully we have modeled an attenuated total reflection coupling situation in Fig. 6. Here the reflectivity, as a function of angle of incidence is plotted for 3.391-μm radiation coupling to a 29.2-nm film of varying ϵ_{r2} surrounded by media with $\epsilon_1=\epsilon_3=2.22$. There always exists a sharp resonance very close to the critical angle of the coupling prism ($\epsilon=2.89$). This is the long-range coupled surface mode which we note is quite insensitive to the specific value of ϵ_{r2}. However, note that this behavior is quite strongly dependent on maintaining the symmetry of the system, for as is shown in the next section, breaking the symmetry substantially changes the situation.

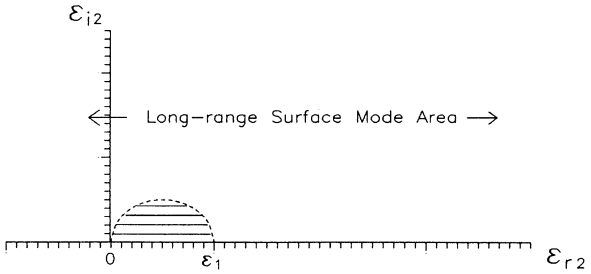

FIG. 4. A map of the region in the complex space of ϵ_2 over which in the thin-film limit the long-range mode can be created in the symmetric three-medium case.

To summarize the situations examined for the single and for the two interfaces we list the cases developed in Secs. II and III in Table I.

Finally, if we examine the power series expansion for $\tanh(x)$, viz.,

$$\tanh(x)=x-\tfrac{1}{3}x^3\cdots, \quad |x|<1, \tag{27}$$

it is clear that for the above cases we require $|\alpha_2 d/2|<1$ and we need to specify the desired precision to impose any tighter limit.

Of course, the goal of this paper is not to simply develop a series of analytic formulas for calculating the dispersion relations for the various modes (they, after all, may

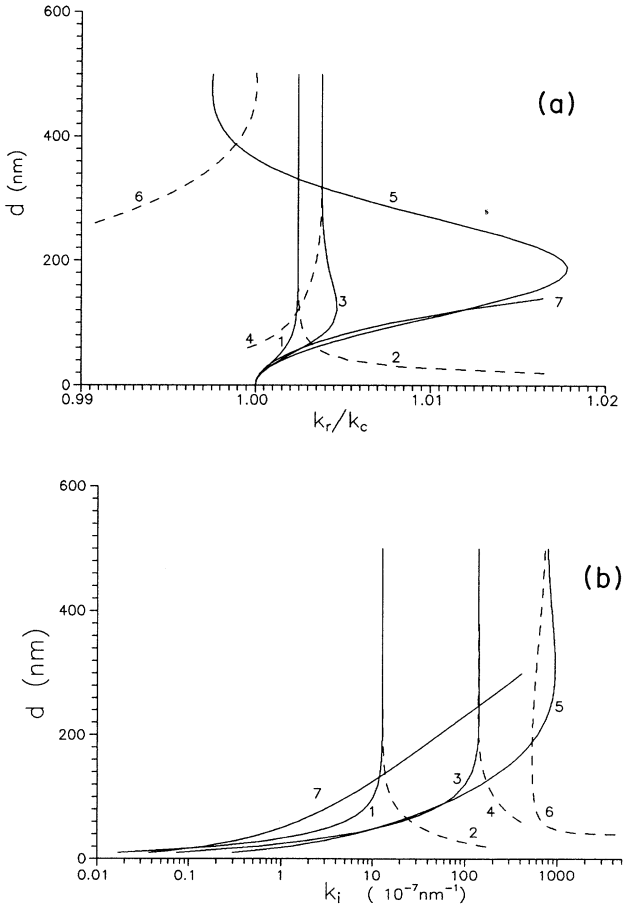

FIG. 5. (a) Relationship between k_r/k_c and the film thickness d. (b) Relationship between k_i and the film thickness d. ($\lambda_0=3.391$ μm, $\epsilon_1=\epsilon_3=2.22$, $k_c=(2\pi/\lambda_0)\epsilon_1^{1/2}=2.76$ μm^{-1}.) Curves 1 and 2, $\epsilon_2=-440-i84$ (gold); curves 3 and 4, $\epsilon_2=-105-i140$ (palladium); curves 5 and 6, $\epsilon_2=0-i40$ (vanadium); curve 7, $\epsilon_2=16.8-i1.0$ (indium antimonide). Solid and dashed lines represent, respectively, modes symmetric and antisymmetric in H_y.

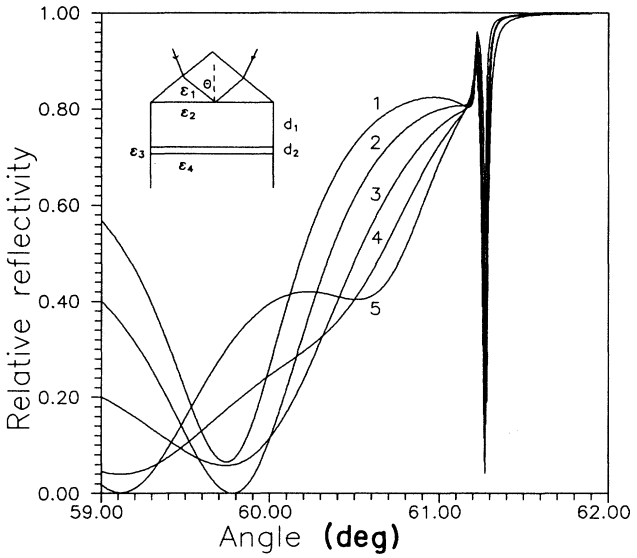

FIG. 6. Calculated angle-scanned attenuated total reflection spectra for p-polarized radiation exciting a long-range surface mode, the sharp feature, and at least one guided mode in the dielectric between the prism and the active medium. ($\lambda_0=3.391\ \mu$m, $\epsilon_1=\epsilon_3=2.22$, $\epsilon_{\text{prism}}=2.89$, $d_1=8.80\ \mu$m, $d_2=29.2$ nm.) Curve 1, $\epsilon_2=-200+i50$; curve 2, $\epsilon_2=-100+i50$; curve 3, $\epsilon_2=0+i50$; curve 4, $\epsilon_2=100+i50$; curve 5, $\epsilon_2=200+i50$.

be readily computed directly from the exact equations). Rather, the purpose of developing the analytic formulas is to allow a more detailed discussion of the long-range modes and to clarify the role of dielectric constants, thickness, and wavelength on their properties. If we wish to extend our analytic studies further in thickness then higher-order approximations may be obtained; this is not difficult using Eq. (27). For example, for $\epsilon_{r2}=0$ and $\epsilon_{i2}>\epsilon_1$ we already have the first-order formulas in Eqs. (20a) and (20b). These become, using the next-order term in the tanh expansion,

$$k_r=k_0\epsilon_1^{1/2}\left[1+\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}^2-\epsilon_1^2}{\epsilon_{i2}^2}+\frac{\epsilon_1}{3}\left[\frac{\pi d}{\lambda_0}\right]^4\frac{\epsilon_1(\epsilon_1^2-3\epsilon_{i2}^2)}{\epsilon_{i2}^2}+\frac{\epsilon_1}{18}\left[\frac{\pi d}{\lambda_0}\right]^6\frac{4\epsilon_1^2\epsilon_{i2}^2-(\epsilon_{i2}^2-\epsilon_1^2)^2}{\epsilon_{i2}^2}\right], \tag{28a}$$

$$k_i=k_0\epsilon_1^{1/2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_1^2}{\epsilon_{i2}}\left[1+\frac{1}{3}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}^2-3\epsilon_1^2}{\epsilon_1}-\frac{2}{9}\left[\frac{\pi d}{\lambda_0}\right]^4(\epsilon_{i2}^2-\epsilon_1^2)\right]. \tag{28b}$$

We compare these two pairs of formulas with the exact computed results in Figs. 7(a) and 7(b). It is clear that the first-order formula [Eq. (20a)] is adequate for k_r for thicknesses up to 50 nm, as seen by comparison of curve 2 with curve 4 on Fig. 7(a). This is not true for k_i. By a thickness of 20 nm the first-order formula, Eq. (20b), is beginning to deviate significantly from the exact calculation; however, the next approximation, Eq. (28b), is very good up to and beyond 50 nm. Thus we conclude that if one wishes to examine general relationships for thicknesses greater than 20 nm it is probably wise, particularly in the case of k_i, to use the higher-order approximation although the first-order one does show adequately the general trend in behavior of both k_r and k_i with thickness. Actually such thicknesses ($\sim$20 nm) are typical of those used to produce long-range surface modes and are also convenient for continuous metal and dielectric thin layers.

IV. LONG-RANGE SURFACE MODES IN AN ASYMMETRIC GEOMETRY

The geometry to be considered is as in Sec. III but now $\epsilon_1\neq\epsilon_3$. We shall limit our considerations to only small differences between ϵ_1 and ϵ_3, i.e., $\epsilon_1-\epsilon_3=\Delta\ll\epsilon_1,\epsilon_3$. We return to Eqs. (12)–(14). Once again in the thin-film limit with small Δ, Eq. (12) may be approximated but now by much more elaborate expressions to give [see Appendix B, Eqs. (B8a) and (B8b)]

TABLE I. Summary of conditions for the existence of surface modes at a single interface and two coupled interfaces.

parameters	single interface	two coupled interfaces
$\epsilon_{r2}<0$ $\lvert\epsilon_{r2}\rvert<\epsilon_1$ $\epsilon_{i2}=0$	no surface mode	two coupled surface modes with the same symmetry for same d value ($d<d_{\max}$)
$\epsilon_{r2}>0$ $\epsilon_{r2}>\epsilon_1$ $\epsilon_{i2}=0$	no surface mode	one long-range coupled surface mode—TM_0 guided mode
$\epsilon_{r2}>0$, $\epsilon_{r2}<\epsilon_1$ $\epsilon_{r2}+\epsilon_{i2}^2/\epsilon_{r2}<\epsilon_1$	a surface mode	no coupled surface mode ($d<d^*$)
any other cases	a surface mode	coupled long-range surface mode and short-range surface mode

$$k_r = k_0 \epsilon_1^{1/2} \left[1 + \left[\frac{\pi d}{\lambda_0} \right]^2 \frac{(\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_1 \epsilon_{r2})^2 (A - B)^2 - \epsilon_1^2 \epsilon_{i2}^2 (A + B)^2}{2\epsilon_1 \epsilon_3^4 (\epsilon_{r2}^2 + \epsilon_{i2}^2)^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]^2} \right] \tag{29a}$$

and

$$k_i = k_0 \epsilon_1^{1/2} \left[\frac{\pi d}{\lambda_0} \right]^2 \frac{\epsilon_{i2} (\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_1 \epsilon_{r2})(A^2 - B^2)}{\epsilon_3^4 (\epsilon_{r2}^2 + \epsilon_{i2}^2)^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]^2} , \tag{29b}$$

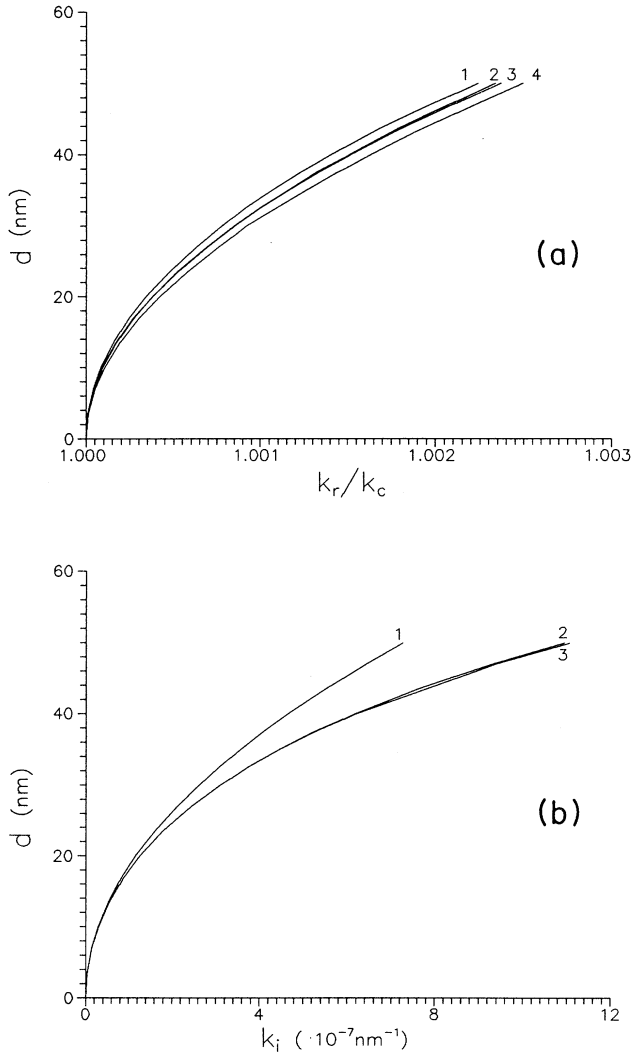

FIG. 7. (a) Relationship between k_r/k_c and film thickness, d. ($\lambda_0 = 3.391$ μm, $\epsilon_1 = \epsilon_3 = 2.22$, $k_c = (2\pi/\lambda_0)\epsilon_1^{1/2} = 2.76$ μm^{-1}.) Curve 1, TM guided mode $\epsilon_2 = 40$; curve 2, LRSEP, $\epsilon_2 = -i40$; curve 4; LRSP, $\epsilon_2 = -40$; all calculated from the exact equation (16) in Sec. III. Curve 3, LRSEP, $\epsilon_2 = -i40$ calculated from the approximate equation (21a). (b) Relationship between k_i and film thickness, d. ($\lambda_0 = 3.391$ μm, $\epsilon_1 = \epsilon_3 = 2.22$, $\epsilon_2 = -i40$.) Curve 1, calculated from the approximate equation (19b); curve 2, calculated from the exact equation (15); curve 3, calculated from the higher-order approximation, Eq. (28b).

where

$$A = \epsilon_1 \epsilon_3^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2] , \tag{29c}$$

$$B = \Delta \epsilon_1 \frac{\epsilon_{r2}^2 + \epsilon_{i2}^2}{4 \left[\frac{\pi d}{\lambda_0} \right]^2} , \tag{29d}$$

$$\Delta = \epsilon_1 - \epsilon_3 > 0 \quad (\Delta \ll \epsilon_1, \epsilon_3) . \tag{29e}$$

The appearance of the term B indicates that there must be a cutoff value for the film thickness d, that is, d can no longer go to zero and still maintain a bound mode. This is very different to the symmetric case, and there is interest in examining this cutoff for a range of situations.

A. Asymmetrically surrounded layer with $\epsilon_{i2} = 0$

If $\epsilon_{i2} = 0$, then the cutoff condition, which means $k_r = k_0 \epsilon_1^{1/2}$, gives, from Eq. (29a), $A = B$. Of course, in this case, $\epsilon_{i2} = 0$ and there is no need to use these thin-film, small-Δ, limits since it is simple to solve directly Eq. (12) with $\alpha_1 = 0$. We then have

$$d_{\text{cut}} = \frac{1}{k_0 (\epsilon_1 - \epsilon_{r2})^{1/2}} \times \tanh^{-1} \left[\frac{-\epsilon_{r2}}{\epsilon_3} \left[\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 - \epsilon_{r2}} \right]^{1/2} \right] . \tag{30}$$

Obviously there is a maximum value of $(\epsilon_1 - \epsilon_3)$, $\Delta_{\max}$, for maintaining a surface mode in a film:

$$\Delta_{\max} = \epsilon_1 - \frac{|\epsilon_{r2}|}{2(\epsilon_1 + |\epsilon_{r2}|)} \times \{ [\epsilon_{r2}^2 + 4\epsilon_1 (\epsilon_1 + |\epsilon_{r2}|)]^{1/2} + \epsilon_{r2} \} . \tag{31}$$

This result, Eq. (31), is the same as those obtained by Burke, Stegeman, and Tamar.[16]

B. Asymmetrically surrounded layer with $\epsilon_{i2} \neq 0$

If $\epsilon_{i2} \neq 0$ and $(\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_1 \epsilon_{r2}) > 0$, then $k_i = 0$ leads immediately to the conclusion that for all values of ϵ_3 (Ref. 16) there is a cutoff condition given by $A^2 = B^2$ which gives, from Eqs. (29c) and (29d), the cutoff thickness d_{cut} as

$$d_{\text{cut}} = \frac{1}{k_0 (\epsilon_1 - \Delta)} \left[\frac{\Delta (\epsilon_{r2}^2 + \epsilon_{i2}^2)}{(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2} \right]^{1/2} . \tag{32}$$

When $d < d_{\text{cut}}$ then $k_i < 0$ and the geometry is no longer able to support a bounded long-range mode; it becomes a growing mode[15,16] just as in Sec. III but now for the asymmetric situation. From Eq. (32) when $|\epsilon_{r2}| \gg \epsilon_1$, d_{cut} varies only weakly with ϵ_{i2} in agreement with the earlier numerical studies.[14] However, if $\epsilon_{r2} \sim \epsilon_1$, then ϵ_{i2} has a strong influence on d_{cut} particularly when ϵ_{i2} is not very large. Similarly when $\epsilon_{i2} \ll (\epsilon_1 - \epsilon_{r2})$ and yet $\epsilon_{r2} \simeq \epsilon_1$ then d_{cut} strongly depends on ϵ_{r2}.

From Eq. (29b), for a fixed value of d as Δ is changed so that A approaches B, then k_i approaches zero and the propagation distance of the long-range surface mode tends to infinity. Further, as d decreases to d_{cut} so k_i decreases more rapidly with changing d, this effect being further enhanced by increasing Δ. These conclusions are in accord with Wendler and Haupt's numerical results[23] for a thin metal film which shows that a slight asymmetry supports a very-long-range surface mode whose properties are dependent on Δ and d. In our case the analytic results may be used for any film so long as it is thin enough to satisfy Eq. (29).

Because the cutoff condition is $A^2 = B^2$, when $\Delta < 0$ the solution $A = -B$ prevails, then the expression for the cutoff thickness has to be changed to

$$d_{\text{cut}} = \frac{[|\Delta|(\epsilon_{r2}^2 + \epsilon_{i2}^2)]^{1/2}}{k_0(\epsilon_1 - \Delta)[(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]^{1/2}} . \tag{33}$$

If the film thickness d and ϵ_1 are fixed, then from Eqs. (28c) and (28d), the maximum value of $(\epsilon_1 - \epsilon_3)$ which will give a long-range surface mode in the film is, for $\epsilon_1 > \epsilon_3$,

$$\Delta_{\max} = \frac{4\left[\frac{\pi d}{\lambda_0}\right]^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]\epsilon_1^2}{(\epsilon_{r2}^2 + \epsilon_{i2}^2) + 8\left[\frac{\pi d}{\lambda_0}\right]^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]\epsilon_1} \tag{34a}$$

and for $\epsilon_1 < \epsilon_3$,

$$\Delta_{\max} = -\frac{4\left[\frac{\pi d}{\lambda_0}\right]^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]\epsilon_1^2}{(\epsilon_{r2}^2 + \epsilon_{i2}^2) - 8\left[\frac{\pi d}{\lambda_0}\right]^2 [(\epsilon_1 - \epsilon_{r2})^2 + \epsilon_{i2}^2]\epsilon_1} . \tag{34b}$$

From Eqs. (34a) and (34b), $\Delta_{\max}$ is approximately proportional to $1/\lambda_0^2$ so the longer the wavelength the more critical the matching condition between ϵ_1 and ϵ_3. For example, consider a silver film of 25 nm thickness with $\epsilon_1 = 2.3$ and $\epsilon_1 > \epsilon_3$. Then at $\lambda_0 = 632.8$ nm with $\epsilon_2 = -19 - 0.5i$ we find $\Delta_{\max} \simeq 0.3$ whereas at $\lambda_0 = 3.391$ μm with $\epsilon_2 = -600 - 100i$, $\Delta_{\max} \simeq 0.01$. Further, as expected on symmetry grounds, when ϵ_1 is fixed, k_r, k_i, and d_{cut} are not symmetric for $\pm\Delta$. These antisymmetric properties for $\pm\Delta$ are also given by Wendler and Haupt[23] in his numerical results for a metal film.

C. Asymmetrically surrounded layer, $|\epsilon_{r2}| < \epsilon_1$

If $\epsilon_{i2} \neq 0$ and $(\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_1\epsilon_{r2}) < 0$, we know from Sec. III there is no bound long-range surface mode for this thin-film situation with the symmetric geometry ($\Delta = 0$) because for $k_i < 0$ and $\text{Re}(\alpha_1) < 0$ the mode is a growing one. However, from Eq. (29b) in the antisymmetric case ($\Delta \neq 0$) and with $d < d_{\text{cut}}$ $[(A^2 - B^2) < 0]$ we find $k_i > 0$. Then we have to establish the signs of $\text{Re}(\alpha_1)$ and $\text{Re}(\alpha_3)$ to ensure the mode is bound. We only need consider the case of $\Delta > 0$ (for $\Delta < 0$ the conclusions are essentially the same). From Eqs. (B6) and (B12) in Appendix B it is clear that $\text{Re}(\alpha_1) > 0$ and $\text{Re}(\alpha_3) < 0$ because $A < B$ and $(\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_1\epsilon_{r2}) < 0$ so the mode is a leaky[31–33] mode. Of course if $A > B$ in the case of $(\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_1\epsilon_{r2}) < 0$, then the mode is a growing one as now $k_i < 0$ and $\text{Re}(\alpha_1) < 0$, $\text{Re}(\alpha_3) < 0$. Finally to summarize the situations examined for an asymmetrically bound film we list the cases developed in Secs. IV B and IV C in Table II.

So the three types of mode, bound, leaky, and growing, are all available from one simple system by changing the parameters of the situation such as the degree of asymmetry or the value of ϵ_{i2} or ϵ_{r2}. However we should be aware, because of the possible sensitivity of the long-range mode response to small differences between ϵ_1 and ϵ_3, that in real materials the finite ϵ_{i1} and ϵ_{i3} values may have an effect. This is considered in the next section.

V. WEAK ABSORPTION BY THE SURROUNDING MEDIA

The geometry once again is that of Sec. III but now $\epsilon_1 = \epsilon_3 = \epsilon_{r1} - i\epsilon_{i1}$ which is symmetric but now includes absorption by the surrounding media. We are only going to consider situations in which $\epsilon_{i1} \ll \epsilon_{r1}$ which is a realizable situation. The pertinent starting equation is Eq. (16) in Sec. III. Examining the thin-film and small-ϵ_{i1} limit

TABLE II. Relation between mode type and geometry parameters for both symmetrically and asymmetrically bound films.

$\epsilon_1 - \epsilon_3$	A-B	$\epsilon_{r2}^2 + \epsilon_{i2}^2 - \epsilon_{r2}\epsilon_{r1}$	k_i	$\text{Re}(\alpha_1)$	$\text{Re}(\alpha_3)$	Mode type
0	>0	>0	>0	>0		bound
0	>0	<0	<0	<0		growing
>0	>0	>0	>0	>0	>0	bound
>0	>0	<0	<0	<0	<0	growing
>0	<0	>0	<0	<0	>0	growing
>0	<0	<0	>0	>0	<0	leaky

then we find using the method outlined in Appendix A that

$$k_r = k_r^0 - k_0 \epsilon_{r1}^{1/2} \epsilon_{r1} \left[\frac{\pi d}{\lambda_0} \right]^2 \times \frac{3(\epsilon_{r2}^2+\epsilon_{i2}^2)-4\epsilon_{r1}\epsilon_{r2}}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2} \epsilon_{i2}\epsilon_{i1} , \tag{35a}$$

$$k_i = k_i^0 + k_0 \epsilon_{r1}^{1/2} \left[\left[\frac{\pi d}{\lambda_0} \right]^2 \frac{a\epsilon_{r2}^2 + b\epsilon_{i2}^2 + \epsilon_{i2}^4}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2} + \frac{1}{2\epsilon_{r1}} \right] \epsilon_{i1} , \tag{35b}$$

where

$$k_r^0 = k_0 \epsilon_{r1}^{1/2} \left[1 + \frac{\epsilon_{r1}}{2} \left[\frac{\pi d}{\lambda_0} \right]^2 \frac{(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r1}\epsilon_{r2})^2 - \epsilon_{r1}^2\epsilon_{r2}^2}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2} \right] , \tag{35c}$$

$$k_i^0 = k_0 \epsilon_{r1}^{1/2} \epsilon_{r1}^2 \left[\frac{\pi d}{\lambda_0} \right]^2 \frac{\epsilon_{i2}(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r1}\epsilon_{r2})}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2} , \tag{35d}$$

$$a = (\epsilon_{r2}-\epsilon_{r1})(\epsilon_{r2}-2\epsilon_{r1}) , \tag{35e}$$

$$b = (\epsilon_{r2}-2\epsilon_{r1})(\epsilon_{r1}+2\epsilon_{r2}) . \tag{35f}$$

The influence of ϵ_{i1} on k_r is dependent on the sign of $[3(\epsilon_{r2}^2+\epsilon_{i2}^2)-4\epsilon_{r1}\epsilon_{r2}]$ provided d is nonzero. In the limit of d tending to zero, k_r tends to an asymptotic value of $k_0\epsilon_1^{1/2}$ while k_i no longer tends to zero but to $k_0\epsilon_{i1}/2\epsilon_{r1}^{1/2}$. These limit values are, as expected, simply those belonging to a plane wave propagating in such an absorbing medium. However, if the film has a finite thickness the influence of ϵ_{i1} on k_i is complicated. For example, if $\epsilon_{i2}=0$, that is, no absorption in the thin film, then from Eq. (35b)

$$k_i = k_0(\epsilon_{r1})^{1/2} \left[\left[\frac{\pi d}{\lambda_0} \right]^2 \frac{a}{\epsilon_{r2}^2} + \frac{1}{2\epsilon_{r1}} \right] \epsilon_{i1} . \tag{36}$$

Then if $a>0$ we find $k_i > k_0\epsilon_{i1}/2\epsilon_{r1}^{1/2}$ which decreases to $k_0\epsilon_{i1}/2\epsilon_{r1}^{1/2}$ as d decreases, while if $a<0$ then $k_i < k_0\epsilon_{i1}/2\epsilon_{r1}^{1/2}$ and this increases to $k_0\epsilon_{i1}/2\epsilon_{r1}^{1/2}$ as d decreases. Clearly if $a=0$, then $k_i = k_0\epsilon_{i1}/2\epsilon_{r1}^{1/2}$ independent of d. From Eq. (35e) if $\epsilon_{r2}<0$ or $\epsilon_{r2}>2\epsilon_{r1}$ then $a>0$ while if $\epsilon_{r1}<\epsilon_{r2}<2\epsilon_{r1}$ then $a<0$ and finally if $\epsilon_{r2}=2\epsilon_{r1}$ (because $\epsilon_{i2}=0$ we must have $\epsilon_{r2}>\epsilon_{r1}$) then $a=0$. Thus we see that a weakly absorbing dielectric medium ($\epsilon_1=\epsilon_{r1}-i\epsilon_{i1}$) surrounding a thin metal film ($\epsilon_{r2}<0$) or a thin dielectric with $\epsilon_{r2}>2\epsilon_{r1}$ will support a long-range surface mode for which the loss is greater than the loss of a plane wave propagating in the absorbing medium. On the other hand, when surrounding a lossless dielectric film satisfying $\epsilon_{r1}<\epsilon_{r2}<2\epsilon_{r1}$ it will support a long-range surface mode for which the loss is less than for a plane wave in the absorbing medium.

All the above discussions centered on varying d and a, but it is also clear from Eqs. (35b) and (36) that the major contribution to k_i is linear in ϵ_{i1} and independent of d. This has a strong effect on the very-long-range surface modes which the systems previously considered may support. For example, in the numerical results of Sec. III we considered the situation of $\lambda_0=3.391$ μm, $\epsilon_1=2.3$, $\epsilon_2=2.0-i0.8$, and $d=10$ nm giving $k_r=2.81$ μm^{-1} and $k_i=1.90\times10^{-6}$ μm^{-1}. If we conclude in this calculation an imaginary part to ϵ_1 of just $\epsilon_{i1}=0.0002$ then while k_r is insignificantly different k_i has become 1.24×10^{-4} μm^{-1}, nearly 2 orders of magnitude larger. This large change is directly as a consequence of the long-range surface mode forcing the fields to lie out of the thin film and in the surrounding (now absorbing) medium. Thus we see it is vitally important that any matching fluid or substrate, used with thin films on which we wish to excite such modes, have minimal absorption.

One further consideration, as discussed in Sec. III, is the problem of the growing mode found for $\epsilon_{i1}=0$ and $(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r2}\epsilon_1)<0$. Now a small value of ϵ_{i1} can change the sign of k_i making it no longer <0 and so the mode type will depend on the sign of $\mathrm{Re}(\alpha_1)$. However, it must be stressed that in general to limit the loss of the mode, ϵ_{i1} is kept small and so it may not be that easy to find a bound mode except for $(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r2}\epsilon_{r1})$ close to zero when it is necessary to consider higher-order corrections to the formulas used thus far.

Finally consider briefly the introduction of some asymmetry between ϵ_1 and ϵ_3. To first order the effects discussed above involving Δ and ϵ_{i1} will superimpose and the mode type will be dictated by a complicated interplay between d, Δ, ϵ_{i1}, and $(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r2}\epsilon_{r1})$. The necessary calculations may be done using Eqs. (29) and (35), although the general complexity of the situation is now becoming such as to make analytic elaboration unwieldly and a numerical solution is probably more appropriate.

VI. EXPERIMENT

The purpose of this paper has been to explore the influence of the dielectric constant $\epsilon_2(=\epsilon_{r2}-i\epsilon_{i2})$ of a thin film on its ability to support a long-range surface mode. We have shown that even for a very large ϵ_{i2} such a mode may still exist. Also in general we have shown that there is nearly always a long-range coupled surface mode for a thin enough film except over a very limited range in the values of ϵ_2.

Already there exists in the literature several studies on the long-range surface plasmon[10–12,16,17] (LRSP) supported by a thin metal film. Also the TM_0-type guided mode is well documented and we need not concern ourselves with it here. Rather our aim is to give results for ϵ_2 values not normally considered; the metal film LRSP case satisfies $\epsilon_{r2}\ll0$ and $|\epsilon_{r2}|\gg\epsilon_{i2}$ while the TM_0 case satisfies $\epsilon_{r2}>0$ and $\epsilon_{r2}\gg\epsilon_{i2}$. We wish to examine situations in which ϵ_{i2} is large. We have specifically studied two cases, one in which $\epsilon_{r2}<0$ and $|\epsilon_{r2}|\sim\epsilon_{i2}$ and one in which $\epsilon_{r2}>0$ and $|\epsilon_{r2}|\ll\epsilon_{i2}$.

The experimental technique is the well-known attenuated total reflection (ATR) technique using a prism

to couple a 3.391-μm laser beam (He-Ne) on to a thin-film sample. The sample geometry is shown in the inset of Fig. 8. At the wavelength chosen the tunnel gap between prism and the active layer may be typically 5–10 μm which is readily achieved in an adjustable manner. The coupling prism which has to be suitably transparent at this wavelength is a single crystal equilateral sapphire prism cut with its symmetry axis parallel to the three equivalent faces giving $n=1.699$ for the p-polarized (TM) incident radiation. This prism is mounted in a micropositioning assembly in which a quartz optical flat with appropriate thin film may be placed adjacent to a surface of the prism. Using interference fringes from visible (632.8 nm) coherent radiation it is possible to make the gap between the sapphire prism and the quartz flat parallel to ± 0.2 μm over the incident beam diameter of 2 mm. In the present study the coupling gap which is adjustable from 4 to 30 μm is filled with a matching fluid of low absorption coefficient. A mixture of hexachloro-1,3-butadiene and tetrachloroethane is used with proportions adjusted so that the dielectric constant is very close to that of the quartz substrate. This coupling fluid thereby also acts as the upper dielectric in the nearly symmetric layer system. The active layers, in one experiment a thin film of palladium and in the other a thin film of vanadium, are deposited by vacuum evaporation on to the quartz substrate in a pressure of 10^{-4} Pa.

Reflectivity data for the 3.391-μm radiation is recorded for a range of incident angles with a resolution of 0.01°. With cooled indium antimonide bolometric detectors and phase-sensitive detection of the 1.7-kHz modulated beam a signal to noise in excess of 1000:1 is readily recorded giving very-high-quality data with which to compare theory.

At the wavelength used the two metals have different characteristics to each other and also to the noble metals conventionally used in surface plasmon studies. Palladium behaves as a metal but the real part of its permittivity while being negative is of the order of or smaller than its large imaginary part. Vanadium doesn't even behave like a metal at this wavelength, a strong interband transition gives it the permittivity more often associated with an excitonic resonance in an insulator, that is, ϵ_r is positive and much smaller than ϵ_i. In order that reflectivity data from the system described above may be compared to Fresnel theory it is first necessary to obtain as many of the independent variables in the problem as possible. In the first instance we only know λ_0. By measuring the prism the internal angles of the prism are accurately obtained (to $\pm 0.005°$) and from the reflectivity for a bare prism the critical angle is measured. By converting incident external angles to internal angles this gives an accurate ϵ value for the prism. We also know within certain tolerance limits (± 0.002) using the Brewster angle the ϵ of the quartz substrate and from the approximate proportion of fluid components the ϵ_r of the matching fluid to ± 0.005. Further, the ϵ_i of the matching fluid is known to be <0.005. This then leaves a total uncertainty for the permittivity of the metal film, ϵ_2, its thickness d, and the fluid layer thickness. By depositing at the same time the thin metal layer on the quartz substrate and a second sapphire prism it is possible to perform a secondary experiment on the film deposited on this second prism to deduce reasonable bounds on ϵ_2 and d. This second experiment involves using the combined geometry of prism, metal film, air gap, thick palladium film and examining the modes in the system using ATR.[34] From fitting Fresnel theory to this data we deduce ϵ_r, ϵ_i, and d to within a few percent and use these parameters as starting parameters with which to begin the fit to the data from the present experiment. This leaves one final unknown, the fluid layer thickness, but from experience with the micropositioner we know this lies in the region 4–7 μm. Thus the fitting of the reflectivity data such as that recorded in Figs. 8 and 9 is now relatively straightforward. In producing the full curves in these plots, which are the theoretical fits using Fresnel theory, there are seven constrained unknowns,

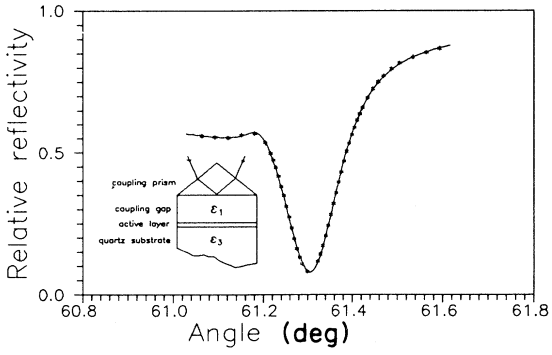

FIG. 8. Data (crosses) for the attenuated total reflection signal for TM-polarized 3.391-μm radiation fitted to Fresnel theory (full curve). Inset is a sketch of the prism-coupling system. The active film in this case is palladium which at this frequency is metallic ($\epsilon_r=-110$, $\epsilon_i=142$) giving an LRSP ($d=43.5$ nm, coupling gap = 5.74 μm).

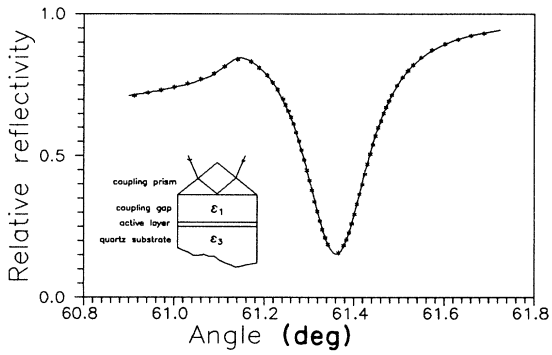

FIG. 9. Data (crosses) for the attenuated total reflection signal for TM-polarized 3.391-μm radiation fitted to Fresnel theory (full curve). The active film is vanadium which at this frequency is excitonic ($\epsilon_r=10.0$, $\epsilon_i=49.3$) giving an LRSEP ($d=50.4$ nm, coupling gap = 6.11 μm).

ϵ_{fluid} $(=\epsilon_{r1}-i\epsilon_{i1})$, ϵ_2 $(=\epsilon_{r2}-i\epsilon_{i2})$, d, ϵ_3 $(=\epsilon_{r3})$, and the fluid thickness. Of these the least constrained parameters were the fluid thickness and the metal thickness; all others were relatively tightly constrained as outlined above to avoid the fitting routine wandering aimlessly in a seven-dimensional variable space. The quality of the fitting is superb with none of the variables hitting any of the imposed bounds and all lying very neatly within sensible physical limits. In Fig. 8 the sharp dip corresponds to an LRSP on a strongly absorbing metal bounded asymmetrically with $|\Delta|=0.004+i0.001$ while the LRSEP in Fig. 9 arises from a strongly absorbing dielectric (vanadium at 3.391 μm) bounded asymmetrically with $|\Delta|=0.004+i0.001$. These sharp resonances, which cannot be simply modeled by the expressions derived earlier because of the significant perturbation introduced by the prism, have half-widths given by $\delta k_x/k_x \simeq 2\times 10^{-3}$ corresponding to propagation lengths of about 0.18 mm. Since this is of the order of the beam diameter, 0.25 mm, then the finite beam size effects may have a small influence on these results and these have not been taken into account in any way.

If we take the fitted parameters from Figs. 8 and 9 we can then model the Poynting vector and H_y distributions for the two minima recorded. It is clear from such models shown in Figs. 10 and 11 that the power is in both cases largely excluded from the active layer. Of course, it is precisely for this reason that such strongly absorbing films are able to support long-range modes. Commensurate with the long distance of propagation is the strong local-field enhancement which is, for the Poynting vector in the case of the LRSP (palladium), 38.83 and for the LRSEP (vanadium), 33.79.

These enhancement factors are by no means as large as

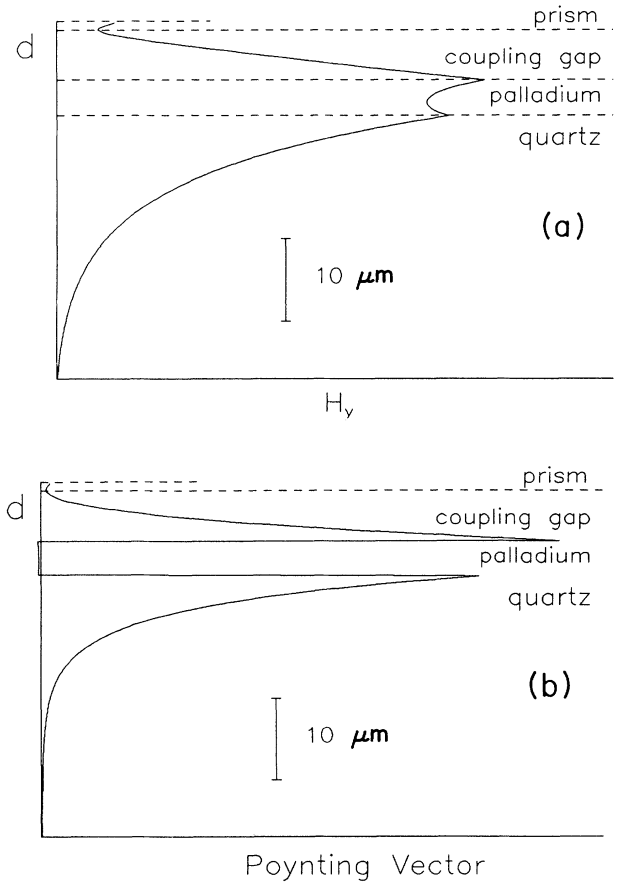

FIG. 10. The H_y field and Poynting vector amplitude distributions for the palladium film of Fig. 8 at 61.31°. The thickness of the active region is scaled by 100× relative to the remainder of the system.

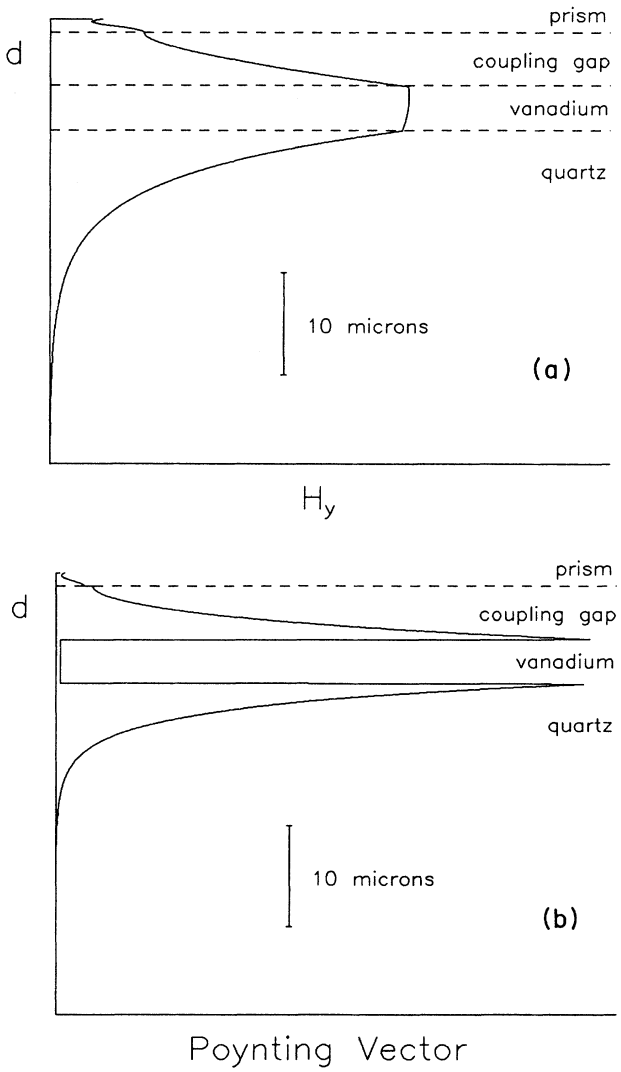

FIG. 11. The H_y field and Poynting vector amplitude distributions for the vanadium film of Fig. 9 at 61.38°. The thickness of the active region is scaled by 100× relative to the remainder of the system.

could be obtained by suitably adjusting ϵ_1 and d; however, finite beam diameter and beam divergence would then become a serious problem.

VII. DISCUSSION

A. Summary of solutions

In the preceding sections we have outlined the solutions to the thickness-dependent k vector for the branch of the dispersion curves which evolves into a long-range surface mode guided by a thin film having almost any permittivity ϵ_2 ($=\epsilon_{r2}-i\epsilon_{i2}$). For the symmetric geometry with a very thin film as well as the established long-range surface plasmon (LRSP) for $\epsilon_{r2}<0$, $|\epsilon_{r2}|\gg\epsilon_{i2}$, $|\epsilon_{r2}|>\epsilon_1$, and the TM_0 guided mode $\epsilon_{r2}>0$, $\epsilon_{r2}\gg\epsilon_{i2}$, $\epsilon_{r2}>\epsilon_1$ there are three new cases. When the imaginary part of the permittivity of the active layer, ϵ_{i2}, is quite large there is still a long-range branch supported by the film. In situations where $|\epsilon_{r2}|\ll\epsilon_{i2}$ irrespective of whether $\epsilon_{r2}\geq 0$ or $\epsilon_{r2}\leq 0$ the loss of the long-range mode is *inversely* proportional to ϵ_{i2}. More surprisingly, when $\epsilon_{i2}=0$ and $\epsilon_{r2}<0$ with $|\epsilon_{r2}|<\epsilon_1$ we also find a long-range surface mode, perhaps more sensibly labeled simply a two-surface symmetric mode because in this situation neither interface independently supports a surface mode. Further, there is a $d_{\max}$ above which no long-range branch exists and below which we can even find *two* symmetric modes having different k values for a single d value. Just like the TM_0 mode these modes do not have an antisymmetric equivalent.

For situations when $(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r2}\epsilon_1)<0$ we are unable to find the long-range branch even though a single interface may support a surface mode provided $\epsilon_{i2}\neq 0$. This again implies a cutoff thickness but now above this thickness modes exist but below it they do not. For values of $(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_{r2}\epsilon_1)$ very close to zero for thin films we find long-range surface modes with extremely low loss-independent of the magnitude of ϵ_{i2}.

In the asymmetric geometry with a small difference between the dielectric constants of the surrounding media we find a minimum cutoff thickness d_{cut} below which the long-range branch no longer exists. As d approaches d_{cut} from above, the fields penetrate progressively further into the medium with the slightly higher index, becoming progressively more like a uniform plane wave propagating parallel to the surface with progressively smaller loss. This means that for a thin film of fixed thickness a little introduced asymmetry can produce a very-long-range surface mode with low loss but increasing the asymmetry further may destroy it completely.

In reality, of course, the surrounding media will have some absorption which in the case of the very-long-range modes may have significant influence especially on k_i the imaginary part of the k vector of the surface mode. This influence is simply because for the long-range mode the fields tend to be pushed into the surrounding media. It is worth noting that if a lossless thin film is inserted into a uniform medium with weak absorption then the loss of the long-range surface mode supported by this geometry may be greater or less than the loss of a wave propagating in the lossy medium according to the exact parameters of the system.

Two new sets of experimental results are presented which support the primary conclusion of the paper; that a long-range coupled surface mode may be supported by a thin film with almost any dielectric constant.

B. The surface polariton, surface mode, and surface wave

As discussed briefly in Sec. I, the term surface polariton is in general confined to that arising from solving the real dispersion equation (6). In this case the surface polariton or electromagnetic surface mode can only be supported at an interface at which one of the media has a negative real dielectric constant whose absolute magnitude is bigger than that in the other medium. This solution is basically a steady-state solution independent of time or propagation distances. Usually this mode may be also called a normal mode of the system because one can employ normal-mode expansions which satisfy power orthogonality relations to describe the total field at any point in space. According to this rigorous definition almost all of the solutions we have considered in this paper are not to be called surface polaritons or surface modes; perhaps they should be labeled surface waves or long-range surface waves because they are solutions to a complex equation and can therefore exist and be excited in an experiment.

However, the surface polariton, or surface mode, is the result of a strong interaction (resonance) between an electromagnetic field and a surface elementary excitation. In the frequency range over which strong interaction occurs both the real and imaginary parts of the dielectric constant of the active material vary rapidly, but continuously, with frequency. Indeed while $\epsilon_i(\omega)$ is going through a sharp peak as ω is changed, $\epsilon_r(\omega)$ may go from large positive values through zero to large negative values and back again to small positive values. So there is in a real situation a combination of all forms of $\epsilon_r(\omega)$ and $\epsilon_i(\omega)$ through a strong resonance, with the exception of $\epsilon_i(\omega)=0$, the one often assumed in simple treatments of this problem. Of course, at some frequency $\epsilon_i(\omega)\ll|\epsilon_r(\omega)|$ and it may be possible to largely ignore it, but the principle is still the same, that is, that for real media the dispersion equations and all the fields will be complex. Thus one has to broaden the definition of the surface mode to include the damping effects. It has often been assumed in the past that while this is a valid point, the only pertinent regime is for $\epsilon_r(\omega)<0$. This we dispute; it is much more appropriate to include the whole of the frequency range over which the given elementary surface excitation is dominant. Thus through an excitonic resonance we may have regions where $\epsilon_r(\omega)<0$ and $\epsilon_r(\omega)>0$, regions where $\epsilon_i(\omega)<|\epsilon_r(\omega)|$ and regions where $\epsilon_i(\omega)>|\epsilon_r(\omega)|$. In all these regions a coupled long-range surface mode may exist for a thin enough film and the correct film geometry, and to subdivide the regions into sections with particular constraints on ϵ is misleading. The important criterion is simply whether a

long-range coupled surface mode exists and whether it may be externally excited and detected.

C. The imaginary part of the dielectric constant

As pointed out above any resonant mode will always have an associated imaginary $\epsilon_i(\omega)$. This ϵ_i basically describes the loss in the material through its influence on the phase of the wave. For an infinite medium one may view absorption as the progressive destructive interference between the exciting radiation and the reemitted radiation from the material. This is also described by making k_i nonzero. In general the larger $\epsilon_i(\omega)$, the larger k_i and for a plane wave the greater is the loss. However, in a nonuniform structure such as the layered structure discussed here, there are other causes of interference and these may redistribute the fields in such a way that the influence of ϵ_i is largely removed. This is particularly obvious in the situation where $|\epsilon_{r2}(\omega)| \ll \epsilon_{i2}(\omega)$ so there would be expected to be a large k_i yet in the thin-film case interference between the two surface modes redistributes the fields so they are small in the active medium. Furthermore, the larger ϵ_{i2} for this situation the smaller the absorption, ϵ_{i2}'s role being then reversed from that normally encountered simply by other interference effects. Thus we see the long-range coupled surface mode is really very different from any single surface mode and it suggests that for complex multilayered structures conventional understanding of the influence of ϵ_i may have to be substantially modified. We also find, as must now be expected, that when $\epsilon_i(\omega) > 0$ the light line defined no longer clearly separates the radiative and nonradiative regions of the dispersion curve just as Burke, Stegeman, and Tamir[16] report.

We have not discussed the short-range branch of the dispersion curve since in the thin-film limit, where the analytic forms might be used, it becomes extremely lossy and is of no potential significance although in general almost all long-range modes do have their short-range counterparts.

Interesting extensions of the present work to consider are the following.

(a) Real material $\epsilon(\omega)$ values as ω is swept through a resonance to study the influence of ω alone using a fixed thickness layer. A material such as ZnO with a strong excitonic resonance is a good example.[35]

(b) Nonlinear effects. The strong-field enhancement with the field being excluded from the strong absorber is a very interesting system for analysis and it also lends itself to experimental study.

(c) Anisotropic media where not only p modes are available but also s modes may couple via off-diagonal matrix elements to the surface charges.

These three extensions of the present study are currently being investigated.

ACKNOWLEDGMENTS

F.Y. acknowledges the financial support of the University of Exeter and the Science and Engineering Research Council (SERC). J.R.S. and G.W.B. also thank the SERC for financial support.

APPENDIX A

The fundamental equation is

$$\tanh\left[\frac{\alpha_2 d}{2}\right]=\frac{-\epsilon_2\alpha_1}{\epsilon_1\alpha_2} . \tag{A1}$$

If d is small enough so that $\alpha_2 d/2 \ll 1$, then

$$\frac{\alpha_2 d}{2}\simeq -\frac{\epsilon_2\alpha_1}{\epsilon_1\alpha_2} , \tag{A2}$$

where the α_j's are defined by Eq. (13). Because d is small, then $k_r \to k_0\epsilon_1^{1/2}$ and $k_i \ll 1$, so defining a small quantity $\delta = k_r - k_0\epsilon_1^{1/2}$ and

$$\alpha_1=[2k_0\epsilon_1^{1/2}(\delta - ik_i)]^{1/2} \tag{A3}$$

and

$$\alpha_2=[k_0^2(\epsilon_1-\epsilon_2)+\alpha_1^2]^{1/2} . \tag{A4}$$

Now from Eq. (A2) we have

$$\frac{\epsilon_1 d\alpha_1^2}{2}+\epsilon_2\alpha_1+\frac{k_0^2\epsilon_1 d(\epsilon_1-\epsilon_2)}{2}=0 ; \tag{A5}$$

thus

$$\epsilon_1 d\alpha_1=-(\epsilon_{r2}-i\epsilon_{i2})\pm[(\epsilon_{r2}-i\epsilon_{i2})^2 - k_0^2\epsilon_1^2 d^2(\epsilon_1-\epsilon_{r2}+i\epsilon_{i2})]^{1/2} . \tag{A6}$$

Since $(\epsilon_{r2}-i\epsilon_{i2})$ is not very small, then because $k_0^2 d^2$ is quite small, we choose the positive sign of the square root in (A6) so that α_1 is a small quantity also. Then

$$\alpha_1=-\frac{k_0^2\epsilon_1 d(\epsilon_1-\epsilon_{r2}+i\epsilon_{i2})}{2(\epsilon_{r2}-i\epsilon_{i2})} . \tag{A7}$$

From this equation we have

$$\mathrm{Re}(\alpha_1)=\frac{k_0^2\epsilon_1 d}{2}\left[\frac{\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2}}{\epsilon_{r2}^2+\epsilon_{i2}^2}\right] \tag{A8}$$

and

$$\alpha_1^2=\frac{k_0^4\epsilon_1^2 d^2}{4}\left[\frac{\epsilon_1-\epsilon_{r2}+i\epsilon_{i2}}{\epsilon_{r2}-i\epsilon_{i2}}\right]^2 . \tag{A9}$$

Then using (A3) and (A9) we find

$$k_r\simeq k_0\epsilon_1^{1/2}\left[1+\frac{\epsilon_1}{2}\left[\frac{\pi d}{\lambda_0}\right]^2 \times\frac{(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})^2-\epsilon_1^2\epsilon_{i2}^2}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2}\right] , \tag{A10a}$$

$$k_i\simeq k_0\epsilon_1^{1/2}\epsilon_1^2\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})}{(\epsilon_{r2}^2+\epsilon_{i2}^2)^2} . \tag{A10b}$$

APPENDIX B

The basic relationship (12) takes the form

$$\tanh(\alpha_2 d)=-\frac{\epsilon_2\alpha_2(\epsilon_1\alpha_3+\epsilon_3\alpha_1)}{\epsilon_1\epsilon_3\alpha_2^2+\epsilon_2^2\alpha_1\alpha_3} \tag{B1}$$

with, as in Appendix A,

$$\alpha_1=[2k_0\epsilon_1^{1/2}(\delta-ik_i)]^{1/2}\ , \tag{B2a}$$

$$\alpha_2=[k_0^2(\epsilon_1-\epsilon_2)+\alpha_1^2]^{1/2}\ , \tag{B2b}$$

$$\alpha_3=(k_0^2\Delta+\alpha_1^2)^{1/2}\ , \tag{B2c}$$

and

$$\Delta=(\epsilon_1-\epsilon_3)\ll\epsilon_1,\epsilon_3\ . \tag{B2d}$$

If $k_0^2\Delta$ is small and of the same order as $2k_0\epsilon_1^{1/2}\delta$, then both α_1 and α_3 are of the same order. So from (B1) in the thin-film limit, neglecting higher-order terms, we have

$$\epsilon_1\epsilon_3\alpha_2^2 d=-\epsilon_2[\epsilon_1(k_0^2\Delta+\alpha_1^2)^{1/2}+\epsilon_3\alpha_1]\ , \tag{B3}$$

i.e.,

$$[2k_0^2\epsilon_1^2 d^2\epsilon_3^2(\epsilon_1-\epsilon_2)+\epsilon_2^2(\epsilon_3^2-\epsilon_1^2)]\alpha_1^2+2\epsilon_1\epsilon_2\epsilon_3^2k_0^2(\epsilon_1-\epsilon_2)d\alpha_1+[k_0^4d^2\epsilon_1^2\epsilon_3^2(\epsilon_1-\epsilon_2)^2-\epsilon_1^2\epsilon_2^2k_0^2\Delta]=0\ . \tag{B4}$$

Then if $\epsilon_2=\epsilon_{r2}-i\epsilon_{i2}$ and $(\epsilon_1-\epsilon_2)=\epsilon_1-\epsilon_{r2}+i\epsilon_{i2}$ are not too small, from Eq. (B4) we have a solution

$$\alpha_1=-\frac{k_0^2\epsilon_1 d\epsilon_3^2(\epsilon_1-\epsilon_2)^2-\epsilon_1\epsilon_2^2\Delta d^{-1}}{2\epsilon_2\epsilon_3^2(\epsilon_1-\epsilon_2)} \tag{B5}$$

which yields

$$\mathrm{Re}(\alpha_1)=\frac{k_0^2 d(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})(A-B)}{2\epsilon_3^2(\epsilon_{r2}^2+\epsilon_{i2}^2)[(\epsilon_1-\epsilon_{r2})^2+\epsilon_{i2}^2]}\ , \tag{B6}$$

where

$$A=\epsilon_1\epsilon_3^2[(\epsilon_1-\epsilon_{r2})^2+\epsilon_{i2}^2]\ , \tag{B7a}$$

$$B=\Delta\epsilon_1(\epsilon_{r2}^2+\epsilon_{i2}^2)\Big/4\left[\frac{\pi d}{\lambda_0}\right]^2\ . \tag{B7b}$$

Using (B5) with (B2) we find

$$k_r=k_0\epsilon_1^{1/2}\left[1+\left[\frac{\pi d}{\lambda_0}\right]^2\frac{(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})^2(A-B)^2-\epsilon_1^2\epsilon_{i2}^2(A+B)^2}{2\epsilon_1\epsilon_3^4(\epsilon_{r2}^2+\epsilon_{i2}^2)^2[(\epsilon_1-\epsilon_{r2})^2+\epsilon_{i2}^2]^2}\right] \tag{B8a}$$

and

$$k_i=k_0\epsilon_1^{1/2}\left[\frac{\pi d}{\lambda_0}\right]^2\frac{\epsilon_{i2}(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})(A^2-B^2)}{\epsilon_3^4(\epsilon_{r2}^2+\epsilon_{i2}^2)^2[(\epsilon_1-\epsilon_{r2})^2+\epsilon_{i2}^2]^2}\ . \tag{B8b}$$

If now we also note that

$$\alpha_1=(\alpha_3^2-k_0^2\Delta)^{1/2} \tag{B9a}$$

and

$$\alpha_2=[k_0^2(\epsilon_1-\epsilon_2)+\alpha_3^2-k_0^2\Delta]^{1/2} \tag{B9b}$$

from (B1) we get

$$\alpha_3\simeq-\frac{\epsilon_1^2\epsilon_3 d^2k_0^2(\epsilon_1-\epsilon_2)^2+\epsilon_3\epsilon_2^2\Delta}{2\epsilon_1^2\epsilon_2(\epsilon_1-\epsilon_2)}\ . \tag{B10}$$

Then finally we find

$$\mathrm{Re}(\alpha_3)=\frac{k_0^2 d(\epsilon_{r2}^2+\epsilon_{i2}^2-\epsilon_1\epsilon_{r2})(A'+B')}{2\epsilon_1^2(\epsilon_{r2}^2+\epsilon_{i2}^2)[(\epsilon_1-\epsilon_{r2})^2+\epsilon_{i2}^2]}\ , \tag{B11}$$

where

$$A'=\epsilon_1^2\epsilon_3[(\epsilon_1-\epsilon_{r2})^2+\epsilon_{i2}^2] \tag{B12a}$$

and

$$B'=\epsilon_3\Delta(\epsilon_{r2}^2+\epsilon_{i2}^2)\Big/4\left[\frac{\pi d}{\lambda_0}\right]^2\ . \tag{B12b}$$

[1]U. Fano, J. Opt. Soc. Am. **31**, 213 (1941).
[2]J. Zenneck, Ann. Phys. **23**, 846 (1907).
[3]D. L. Mills and E. Burstein, Rep. Prog. Phys. **37**, 817 (1974).
[4]G. Borstel and H. J. Falge, Phys. Status Solidi **83B**, 11 (1977).
[5]A. Brillante, I. Pockrand, M. R. Philpott, and J. D. Swalen, Chem. Phys. Lett. **57**, 395 (1978).
[6]Fuzi Yang, G. W. Bradberry, and J. R. Sambles, J. Mod. Opt. **36**, 1405 (1989).
[7]Fuzi Yang, G. W. Bradberry, and J. R. Sambles, J. Mod. Opt. **37**, 1545 (1990).

[8]R. Bruns and H. Raether, Z. Phys. **237**, 98 (1970).
[9]A. B. Buckman and C. Kuo, J. Opt. Soc. Am. **69**, 343 (1979).
[10]S. Hayashi, T. Yamada, and H. Kanamori, Opt. Commun. **36**, 195 (1981).
[11]Y. Kuwamura, M. Fukuri, and O. Tada, J. Phys. Soc. Jpn. **52**, 2350 (1983).
[12]J. C. Quail, J. G. Rabo, and H. J. Simon, Opt. Lett. **8**, 377 (1983).
[13]A. E. Craig, G. A. Olson, and D. Sarid, Opt. Lett. **8**, 383 (1983).
[14]H. Dohi, Y. Kuwamura, M. Fukui, and O. Tada, J. Phys. Soc. Jpn. **53**, 2828 (1984).
[15]P. E. Ferguson, D. F. Wallis, and C. Hauvet, Surf. Sci. **82**, 255 (1979).
[16]J. J. Burke, G. I. Stegeman, and T. Tamir, Phys. Rev. B **8**, 5186 (1986).
[17]D. Sarid, R. T. Deck, and J. J. Fasano, J. Opt. Soc. Am. **72**, 1345 (1982).
[18]G. I. Stegeman and J. J. Burke, Appl. Phys. Lett. **41**, 906 (1982).
[19]J. C. Quail and H. J. Simon, J. Appl. Phys. **56**, 2589 (1984).
[20]H. J. Simon and Zhan Chen, Phys. Rev. B **39**, 3077 (1989).
[21]R. K. Hickernell and D. Sarid, J. Opt. Soc. Am. B **3**, 1059 (1986).
[22]G. S. Agarwal and S. D. Gupta, Phys. Rev. B **34**, 5239 (1986).
[23]L. Wendler and R. Haupt, J. Appl. Phys. **59**, 3289 (1986).
[24]F. Y. Kou and T. Tamir, Opt. Lett. **12**, 367 (1987).
[25]Z. Lenac and M. S. Tomas, J. Phys. C **16**, 4273 (1983).
[26]R. K. Hickernell and D. Sarid, Opt. Lett. **12**, 570 (1987).
[27]Fuzi Yang, J. R. Sambles, and G. W. Bradberry, Phys. Rev. Lett. **64**, 559 (1990).
[28]K. J. Kliewer and R. Fuchs, Phys. Rev. **144**, 495 (1966).
[29]D. Sarid, Phys. Rev. Lett. **47**, 1927 (1981).
[30]T. Tamir, *Integrated Optics* (Springer-Verlag, Berlin 1982).
[31]T. Tamir and A. A. Oliner, Proc. IEEE (London) **110**, 310 (1963).
[32]T. Tamir and A. A. Oliner, Proc. IEEE (London) **110**, 325 (1963).
[33]J. J. Burke and N. S. Kapany, *Optical Waveguides* (Academic, New York, 1972).
[34]Fuzi Yang, G. W. Bradberry, D. J. Jarvis, and J. R. Sambles, J. Mod. Opt. **37**, 977 (1990).
[35]F. DeMartini, M. Colocci, S. E. Kohn and Y. R. Shen, Phys. Rev. Lett. **38**, 1223 (1977).