Nonlocal corrections to Fresnel optics: Comparison of exact solutions with d-parameter approximations

Wei Chen and W. L. Schaich
Department of Physics and Materials Research Institute, Indiana University, Bloomington, Indiana 47405
(Received 6 April 1989)

Various model systems whose optics can be calculated exactly are examined from the point of view of d-parameter theory [Phys. Rev. B **39**, 10 714 (1989)]. By comparing exact and approximate estimates of reflection amplitudes, we determine the range of validity of d-parameter theory for these specific models. In general terms the d-parameter formalism at clean surfaces only has difficulty near thresholds for bulk polaritons. The size of damping terms and the strength of the spatial dispersion determine whether the breakdown of d-parameter theory is severe or not.

I. INTRODUCTION

We recently presented a formal derivation of the d-parameter theory of nonlocal corrections to Fresnel optics which allows its extension beyond jellium-model calculations.[1] At several points in this derivation one must invoke the assumption that the effective thickness of the interface region is much smaller than macroscopic wavelengths of the radiation fields. Since the general argument was made without reference to specific systems, quantitative estimates of when or to what extent the perturbative approach of d-parameter theory breaks down were not made. In this paper we provide such estimates for several simple models.

The model systems are chosen more for their simplicity of analysis than for their experimental realism. Specifically we examine here only models whose optics can be exactly calculated. We compare these exact solutions with the approximate results of d-parameter theory applied to the same models. Our aim is to illustrate explicitly the range of validity of d-parameter theory, and to develop empirical measures of when it should or should not be trusted. These empirical rules will hopefully remain meaningful for more realistic models whose exact solutions are either unknown or too complicated to conveniently evaluate.

Before turning to specific models in Sec. II, we sketch the form of the d-parameter corrections to Fresnel optics. Our focus is on the reflection amplitudes produced when a monochromatic beam of light or either s or p polarization is incident from vacuum at an angle θ with respect to the normal to a flat surface. The only material property required by the Fresnel theory is ϵ, the local dielectric function of the bulk material. One finds for the Fresnel reflection amplitudes

$$r_s^0=\frac{\beta_0-\beta_1}{\beta_0+\beta_1}\ , \tag{1}$$

$$r_p^0=\frac{\epsilon\beta_0-\beta_1}{\epsilon\beta_0+\beta_1}\ , \tag{2}$$

where the β's are normal components of wave vectors. With ω the light frequency, c the speed of light in vacuum, and $Q=(\omega/c)\sin\theta$ the common projection of all wave vectors in the plane parallel to the surface, the vacuum β is

$$\beta_0=\left[\frac{\omega^2}{c^2}-Q^2\right]^{1/2} \tag{3}$$

while the bulk β is

$$\beta_1=\left[\frac{\omega^2}{c^2}\epsilon-Q^2\right]^{1/2}\ . \tag{4}$$

The d-parameter theory characterizes the first-order corrections to (1) and (2) due to any intrinsic surface response. For systems with sufficient symmetry only two d parameters can be nonzero,[1] $d_\parallel$ and $d_\perp$, and both are complex-valued functions of ω, as is ϵ in general. The corrected reflection amplitudes are

$$r_s/r_s^0=1+2i\beta_0 d_\parallel+\cdots\ , \tag{5}$$

$$r_p/r_p^0=1+2i\beta_0\left[d_\parallel-\frac{Q^2\epsilon}{\epsilon\beta_0^2-Q^2}(d_\perp-d_\parallel)\right]+\cdots\ . \tag{6}$$

These results are derived under the presumption that the corrections are small,[1-8] and the equations for the d's are usually evaluated neglecting retardation; i.e., with c set to ∞. For the models considered in Sec. II one can find the exact r's and explicitly examine the validity of the above approximate scheme. We do this with models that roughly describe in turn plasmons, optical phonons, and Frenkel excitons. Then in Sec. III we summarize our results in general terms.

II. MODEL SOLUTIONS

We begin with two comments that apply to all the models considered here. First we stress that these models provide only crude descriptions of the particular physical systems. Better models are known but their increased physical sophistication usually complicates an exact mathematical solution. One can often fit the crude mod-

els to specific experimental systems by adjusting various free parameters, but we do not pursue this option. If we do vary these parameters it is only to exhibit the sensitivity and precision of different mathematical approximations. Our second general comment is that all of the exact solutions we study were obtained by others in earlier work. Hence we do not give details of the solution methods, but just define each model and quote its exact r's. The emphasis is on comparing these exact solutions with d-parameter estimates.

A. Plasmons

The first system we consider is a model of metal optics. To understand nonlocal effects near the plasma frequency has been a principal driving force behind the renewed interest in d-parameter calculations.[9–11] The simple hydrodynamic model we examine here contains some of the relevant physics, in particular the threshold for plasmon creation at $\omega=\omega_p$. The metal is described by the constitutive equation

$$\frac{\partial \mathbf{j}}{\partial t}=\frac{\omega_p^2}{4\pi}\mathbf{E}-D_{\mathrm{L}}\nabla\delta\rho-\gamma\mathbf{j}\ , \tag{7}$$

which relates the time derivative of the electron current density $\mathbf{j}$ to the electric field $\mathbf{E}$ and gradients of the induced charge density $\delta\rho$. At the simplest level the plasma frequency ω_p, the dispersion parameter D_{L}, and the scattering rate γ are fixed constants. For this case the transverse dielectric function is given by Drude's formula

$$\epsilon_{\mathrm{T}}=1-\omega_p^2/\tilde{\omega}^2\ , \tag{8}$$

where $\tilde{\omega}^2=\omega(\omega+i\gamma)$; while the longitudinal dielectric function alone has spatial dispersion effects,

$$\epsilon_{\mathrm{L}}=1-\frac{\omega_p^2}{\tilde{\omega}^2-D_{\mathrm{L}}|\mathbf{q}|^2}\ , \tag{9}$$

where $\mathbf{q}$ is a three-dimensional Fourier wave vector. These results imply that only p-polarized light will show deviations from Fresnel theory, and in turn that only $d_\perp$ is significant.

The exact p-wave reflection amplitude for a medium described by (7) up to a matching plane with vacuum is, if we require the normal component of $\mathbf{j}$ to vanish at the surface,[12]

$$r_p=\frac{\beta_0\epsilon_{\mathrm{T}}-\beta_1-(1-\epsilon_{\mathrm{T}})Q^2/\beta_3}{\beta_0\epsilon_{\mathrm{T}}+\beta_1+(1-\epsilon_{\mathrm{T}})Q^2/\beta_3}\ , \tag{10}$$

where the normal component β_3 of the longitudinal wave is set by

$$(\beta_3)^2=\frac{\tilde{\omega}^2-\omega_p^2}{D_{\mathrm{L}}}-Q^2\ . \tag{11}$$

Equation (11) follows from the condition $\epsilon_{\mathrm{L}}=0$. If we identify the ϵ_{T} of (8) with the ϵ of (1)–(6), we find upon expanding (10)

$$r_p/r_p^0\approx 1-2i\beta_0\frac{Q^2\epsilon}{\epsilon\beta_0^2-Q^2}\frac{i}{\beta_3}+\cdots\ . \tag{10'}$$

This is to be compared via (6) with the simplest d-parameter estimate for this model:[9]

$$d_\perp^0=i/\beta_3^0\ , \tag{12}$$

where $d_\parallel$ is zero and β_3^0, given by

$$(\beta_3^0)^2=\frac{\tilde{\omega}^2-\omega_p^2}{D_{\mathrm{L}}}\ , \tag{13}$$

is the nonretarded limit of β_3.

Thus the d-parameter estimate (6) is reliable for this model if one can replace $\beta_3\to\beta_3^0$ and expand (10) into (10'). For metals D_{L} is typically 4–5 orders of magnitude smaller than c^2, so except for ω very near ω_p the above reductions are justified. We show in Fig. 1 a comparison of an effective $d_\perp$ with the simple estimate (12). The former is found by requiring Eq. (6) with $d_\parallel=0$ to reproduce (10). We have chosen $D_{\mathrm{L}}/c^2=10^{-5}$ and $\gamma/\omega_p=10^{-3}$. The latter choice is rather small, but larger values make the agreement in Fig. 1 even better.

To provide some insight into why the d-parameter theory works so well here, consider the polarization amplitude $\mathbf{P}=\mathbf{j}/(-i\omega)$. In vacuum $\mathbf{P}$ vanishes, while inside the metal for an incident p wave

$$\mathbf{P}=P_1\left[(Q,-\beta_1,0)e^{i\beta_1 x}-\frac{Q}{\beta_3}(\beta_3,Q,0)e^{i\beta_3 x}\right]^{i(\underset{\sim}{Q}\cdot\underset{\sim}{X}-\omega t)}\ , \tag{14}$$

where x is the normal coordinate and $\underset{\sim}{X}$ describes variations within planes parallel to the surface. The metal lies in $x>0$ and all the β's here (and below) are chosen to have a positive imaginary part so the various partial

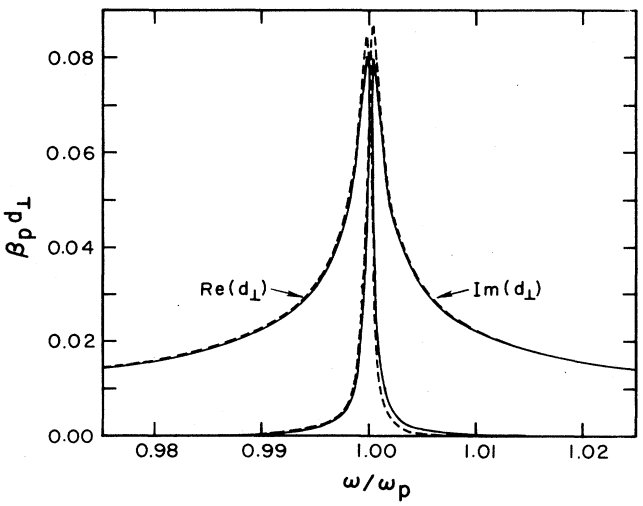

FIG. 1. Plasmon model of $d_\perp$. We plot $\beta_p d_\perp$, where $\beta_p=\omega_p/c$, vs ω/ω_p for frequencies near the plasmon threshold. Both real and imaginary parts of $d_\perp$ are shown. The solid curves are from the nonretarded result (12), while the dashed curves describe an effective $d_\perp$ that reproduces the exact p-wave reflection for an angle of incidence of 60°.

waves decay into the bulk. The vector character of $\mathbf{P}$ is described by a triplet of components along the mutually orthogonal $\hat{\mathbf{x}}$, $\hat{\mathbf{Q}}$, and $\hat{\mathbf{t}}=\hat{\mathbf{x}}\times\hat{\mathbf{Q}}$ directions. The relative strength of the transverse and longitudinal contributions to $\mathbf{P}$ has been fixed by the additional boundary condition (ABC) that $\hat{\mathbf{x}}\cdot\mathbf{P}=0$ at $x=0$. As an aside, we remark that we view the ABC required here (as well as others imposed below) as part of the specification of a model system, and refer the reader elsewhere for its rationalization.[12]

Equation (14) gives the polarization with all retardation included. The d-parameter derivations[1] assume that $\mathbf{P}$ can be separated into a Fresnel-like piece

$$\mathbf{P}^{(F)}=P_1^{(F)}(Q,-\beta_1^{(F)},0)e^{i\beta_1^{(F)}x}e^{i(\underline{Q}\cdot\underline{X}-\omega t)} \tag{15}$$

and a "surface-localized" remainder. Equation (8) implies here that $\beta_1^{(F)}=\beta_1$ and $P_1^{(F)}=P_1$. The remainder term only needs to be localized to the extent that one can extract its integral over x from a spatial range small compared to that over which $\mathbf{P}^{(F)}$ varies. For the hydrodynamic model this constraint is simply that $|\beta_3|\gg|\beta_1|$. We emphasize that it holds equally well for β_3 nearly pure real $(\omega-\omega_p\gg\gamma)$ or pure imaginary $(\omega-\omega_p\ll-\gamma)$. The minimum of $|\beta_3|c/\omega_p$ occurs near ω_p and is roughly $[(\gamma/\omega_p)(c^2/D_L)]^{1/2}$, which for our parameter choices equals 10. This is consistent with Fig. 1 where the largest error in $d_\perp^0$ is about 10% and occurs only for ω within 0.1% of ω_p. We conclude that the maximum relative error in $d_\perp^0$ is controlled by the damping. However, this inference is a special consequence of the longitudinal nature of the plasmon since the two partial waves of polarization in (14) are completely uncoupled in bulk so only the finiteness of $1/\tau$ keeps their respective β's from crossing as ω passes through ω_p.

B. Optical phonons

For the next model system we switch from the picture of itinerant charge flowing to that of localized charge polarizing. Specifically we consider an ordered array of point, dipole-polarizable molecules. Instead of (7) we write an equation for the induced dipole moment at site l

$$\mathbf{p}(l)=\alpha\mathbf{E}_{\rm loc}(l)\ , \tag{16}$$

as a product of the molecule's (presumed) isotropic polarizability α and the local electric field at the site. The latter has contributions from the externally applied field and from the induced fields produced by the other molecules

$$\mathbf{E}_{\rm loc}(l)=\mathbf{E}_A(l)+\sum_{l'\ (\neq l)}\overleftrightarrow{T}(l,l')\cdot\mathbf{p}(l')\ . \tag{17}$$

The dipole-dipole coupling matrix $\overleftrightarrow{T}$ has both short-range terms due to near (essentially nonretarded) fields and long-range terms due to far (fully retarded) fields. The latter vanish when $c\to\infty$.

This model leads to nontrivial values of both $d_\parallel$ and $d_\perp$, but if we limit ourselves to the normal incidence reflectivity we only need $d_\parallel$. We further specialize to truncate the short-range coupling beyond nearest-neighbor planes of molecules at a (100) surface of a simple-cubic array with lattice constant a. The exact s-wave reflection amplitude at $\theta=0=Q$ can then be expressed as[13]

$$r_s=-e^{i\beta_0 a}\frac{e^{-i\beta_0 a}-e^{-i\beta_1 a}}{e^{+i\beta_0 a}-e^{-i\beta_1 a}}\frac{e^{-\beta_0 a}-e^{-i\beta_2 a}}{e^{+i\beta_0 a}-e^{-i\beta_2 a}}\ , \tag{18}$$

where $\beta_0=\omega/c$ and the origin for x is at $a/2$ outside the last plane of molecules. Both β_1 and β_2 are normal components of transverse wave vectors and are to be found from

$$0=1-\alpha[V(0)+2V(1)\cos(\beta a)]+4\pi n\alpha\left[\frac{\beta_0 a}{2}\frac{\sin(\beta_0 a)}{\cos(\beta a)-\cos(\beta_0 a)}\right]\ . \tag{19}$$

Here the V describe the contribution to the local field from other molecules in the same plane—$V(0)$—or in an adjacent plane—$V(1)$. The V depend only weakly on frequency and their nonretarded values are constants: for a simple cubic lattice $V(0)/n\to+4.517=V^0(0)/n$ and $V(1)/n\to-0.164=V^0(1)/n$ as $c\to\infty$, where $n=1/a^3$ is the density of molecules.

It is not immediately obvious how (18) may be related to a corrected Fresnel theory, but if we assume that the products of βa in (19) are small, that equation can be reexpressed as

$$\beta^2\approx\beta_0^2\epsilon(\beta,\omega)\ , \tag{19'}$$

where

$$\epsilon(\beta,\omega)=1+\frac{4\pi n\alpha}{1-\alpha[V(0)+2V(1)\cos(\beta a)]}\approx\epsilon^0(\omega)=1+\frac{4\pi n a}{1-\alpha[V^0(0)+2V^0(1)]}\ . \tag{20}$$

This last result for ϵ^0 is similar in appearance to the Clausius-Mossotti formula.[14] One would only need $V^0(0)+2V^0(1)$ to equal $4\pi n/3$, which indeed is nearly true. Hence we define an effective $d_\parallel$ for this model by requiring that (5) using the ϵ^0 of (20) exactly reproduce (18). On the other hand, the $d_\parallel$ from the simplest nonretarded treatment of this model is[15]

$$d_\parallel^0=a/(e^{-i\beta_2^0 a}-1)\ , \tag{21}$$

where β_2^0 is a solution of

$$1=\alpha[V^0(0)+2V^0(1)\cos(\beta_2^0 a)]\ . \tag{22}$$

Its prediction of r_s is also to be found from (5) using the ϵ^0 of (20) to set β_1.

In Fig. 2 we plot $d_\parallel^0$ versus frequency near its resonance. We have chosen $a/\bar{\lambda}=10^{-4}$ where $\bar{\lambda}=2\pi c/\bar{\omega}$ and have set

$$4\pi n\alpha=\frac{f\bar{\omega}^2}{\bar{\omega}^2-\tilde{\omega}^2} \tag{23}$$

with $f=1$. We again write $\tilde{\omega}^2=\omega^2+i\omega\gamma$ but now

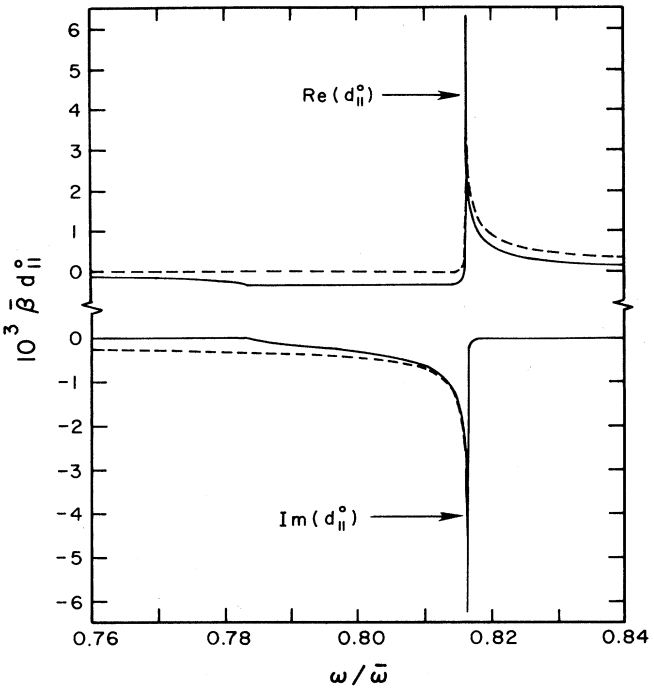

FIG. 2. Optical-phonon model of $d_{\|}$. We plot $\bar{\beta}d_{\|}^{0}$, where $\bar{\beta}=\bar{\omega}/c$, vs $\omega/\bar{\omega}$ for frequencies near the transverse resonance. Both real and imaginary parts of $d_{\|}^{0}$ are shown. The solid curves are from the point-dipole model, Eqs. (21) and (22), while the dashed curves describe the continuum model, Eqs. (27) and (28). Near ω_T the two $d_{\|}^{0}$'s differ by $a/2$.

$\gamma/\bar{\omega}=10^{-4}$. The (transverse) resonance in $\epsilon^0(\omega)$ is at $\omega_T/\bar{\omega}\approx 0.816$. The effective $d_{\|}$ defined above has not been plotted since it is nearly identical to $d_{\|}^{0}$. Only at the resonance peak does it show a small (10%) difference. Hence again d-parameter theory is quite successful.

As in Sec. II A it is useful to examine the polarization amplitude. The analogue of (14) is for an infinitesimal Q

$$\mathbf{P}(x=(l+\tfrac{1}{2})a,\ \underset{\sim}{X}=\underset{\sim}{L}a)$$
$$=P_1[(0,0,\beta_0)e^{i\beta_1(x+a/2)}-(0,0,\beta_0)e^{i\beta_2(x+a/2)}]e^{i(\underset{\sim}{Q}\cdot\underset{\sim}{X}-\omega t)}\ , \tag{24}$$

which vanishes when $l=-1$; i.e., at the location of the first "missing" layer.[13] The d-parameter theory effectively assumes that $\beta_2\approx\beta_2^0$ is much larger in magnitude than $\beta_1\approx\beta_1^F$, where β_1^F is to be calculated from (4) using the ϵ^0 of (20). This inequality is well obeyed except near ω_T, where the exact β_1 and β_2 have an avoided crossing. On the scale of Fig. 2 the consequent difficulties scarcely show, but would become apparent on an expanded scale such as used in the next section.

Before moving to that new physical system, we note that exact solutions can still be readily found for several extensions of the present model. One can allow the short-range interactions to extend beyond the nearest-neighbor plane for a finite distance for arbitrary ratios of their relative strengths;[13] or if the interactions can be described by a finite sum of terms exponential in the interplane separation, they can be kept for all separations.[16–18] One can find r_s off normal incidence[19] and can also find r_p for any θ.[20]

Exact solutions can furthermore be easily found for continuum models of the present system.[20] Such models can be imagined as the small-a limit of what we have just analyzed. One introduces say as an intermediate step in (20)

$$\epsilon_T^0(\beta,\omega)=1+\frac{4\pi n\alpha}{1-\alpha[V^0(0)+2V^0(1)]+\alpha V^0(1)a^2\beta^2}\ , \tag{25}$$

which when combined with (23) becomes

$$\epsilon_T^0(\beta,\omega)=1+\frac{S\omega_T^2}{\omega_T^2-\bar{\omega}^2+D_T\beta^2}\ , \tag{26}$$

where

$$\omega_T^2/\bar{\omega}^2=1-f[V^0(0)+2V^0(1)]/4\pi n\ ,$$

$D_T=fV^0(1)\bar{\omega}^2a^2/4\pi n$, and $S\omega_T^2=f\bar{\omega}^2$. The d's that result from such modifications are qualitatively similar to the discrete lattice answers. Figure 2 has a comparison for $d_{\|}$. The specific recipe for the continuum $d_{\|}^{0}$ plotted there is [cf. Eq. (21)]

$$d_{\|}^{0}=i/\beta_2^0\ , \tag{27}$$

where β_2^0 is defined by the pole $\epsilon_T^0(\beta,\omega)$:

$$(\beta_2^0)^2=\frac{\bar{\omega}^2-\omega_T^2}{D_T}\ , \tag{28}$$

which is an analogue of (13). The continuum $d_{\|}^{0}$ is essentially as accurate as that of the discrete lattice. The two have qualitative differences between each other away from ω_T, particularly at the other extremum of the band of real β_2^0 values for the point-dipole model. This is also evident in the nonretarded dispersion curves shown in Fig. 3. Such bands of solutions have been studied more thoroughly elsewhere.[21,22] We only stress here that the influence of these differences on reflection amplitudes can be accurately studied with d-parameter theory away from ω_T since they occur when $|\beta_2|\approx|\beta_2^0|\approx\pi/a\gg|\beta_1^F|\approx|\beta_1|$. Very near ω_T the main difference between the two model estimates of $d_{\|}^{0}$ is a constant, real-valued shift. This arises since the ABC used to produce (27) makes the polarization vanish at $x=0$, while that in (24) vanishes at $-a/2$. The extent to which this shift is noticeable is a measure of how well one should expect discrete and continuum models to agree.

C. Excitons

The last model problem we consider will in its parameter choices lead us to examine much sharper resonance structure. Although a point-dipole model can be applied to this case,[20] we use for simplicity an isotropic continuum model since it should have the same qualitative behavior near the resonance peaks, aside from the constant

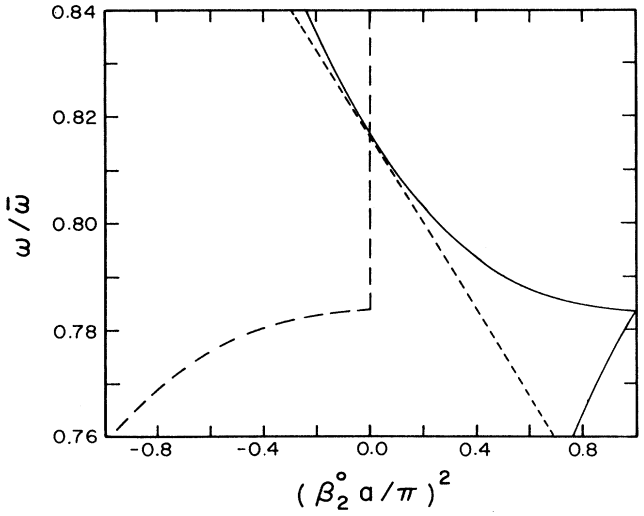

FIG. 3. Comparison of nonretarded dispersion relations near the transverse-optical phonon frequency, $\omega_T \approx 0.816\bar{\omega}$. The damping parameter γ is zero and for each value of $\omega/\bar{\omega}$ we give the value of $(\beta_2^0 a/\pi)^2$. The solid and short-dashed curves are for the point-dipole model (22) and the continuum model (28), respectively. For the former, below $\omega = 0.784\bar{\omega}$, the real part of β_2^0 remains fixed at $-\pi/a$ while the imaginary part of β_2^0 varies. The resultant imaginary part of $(\beta_2^0)^2$ is denoted by the long-dashed curve. The retarded behavior very close to ω_T is complicated; see Fig. 4.

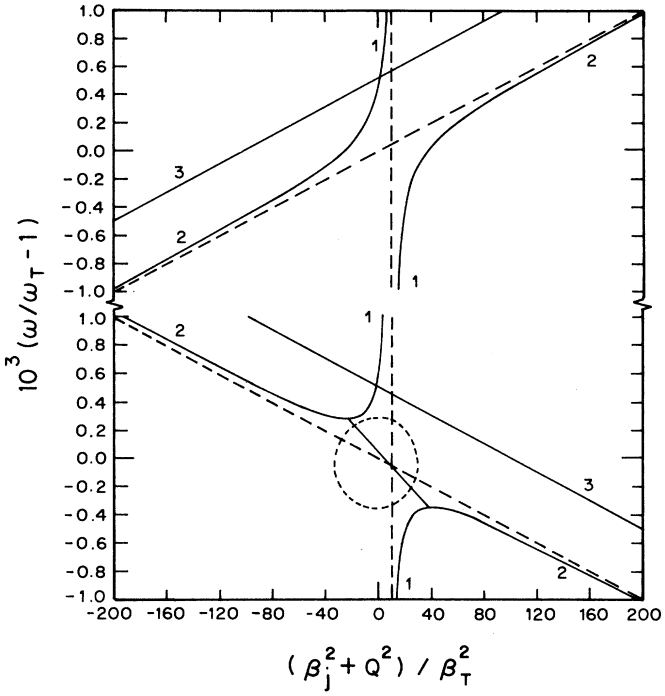

FIG. 4. Comparison of exact and asymptotic dispersion relations near the exciton resonance. The damping parameter γ is zero and we show $(\beta_j^2+Q^2)/\beta_T^2$, where $\beta_T=\omega_T/c$, vs ω/ω_T for both the longitudinal mode (30) and the two transverse modes β_1 and β_2. The various branches are labeled by the appropriate subscript of β. Note that since $Q^2/\beta_T^2 \approx \sin^2\theta < 1$, the dependence on θ is slight. The long-dashed curves give the asymptotic, uncoupled behavior of the transverse modes. The nearly vertical one is $(\beta_1^0)^2+Q^2 \approx (\omega^2/c^2)\epsilon_0$, while the oblique one is $(\beta_2)^2+Q^2 \approx (\omega^2-\omega_T^2)/D$. The two panels contrast the behavior for positive and negative sign of the spatial dispersion parameter D. When $D<0$, there is a region near ω_T where β_1 and β_2 are neither pure real nor pure imaginary, but are complex conjugates of each other. The resulting imaginary part of $(\beta_j)^2$ for $j=1,2$ is denoted by the short dashes.

shift discussed above. The model is based on the postulated bulk, nonlocal dielectric function

$$\epsilon(\mathbf{q},\omega)=\epsilon_0+\frac{S\omega_T^2}{\omega_T^2-\tilde{\omega}^2+D|\mathbf{q}|^2} . \tag{29}$$

Aside from isotropy in $\mathbf{q}$ space this differs from (26) only in allowing for a nonresonant polarization via $\epsilon_0 \neq 1$. Various partial waves of polarization may be found from (29). If we require ϵ with $\mathbf{q}=(\beta_3,Q,0)$ to vanish, we obtain one (longitudinal) solution

$$(\beta_3)^2=\frac{\tilde{\omega}^2-\omega_L^2}{D}-Q^2 , \tag{30}$$

where $\omega_L^2=\omega_T^2(1+S/\epsilon_0)$. On the other hand, the zeros of $\omega^2\epsilon/c^2-|\mathbf{q}|^2$ yield two different (transverse) solutions for the square of a normal wave-vector component, which we call either $(\beta_1)^2$ or $(\beta_2)^2$. Based on the results in Sec. II B one might expect to identify one of these with β_1^F, the Fresnel solution, and use the other (hopefully much larger) one along with β_3 to describe the surface-localized variation of the polarization. This scenario is possible for frequencies away from ω_T and ω_L, but can fail in their near vicinity.

We show in Fig. 4 how the various β^2's vary with frequency near the resonances for parameters typical of excitons. We have chosen $S=10^{-2}$ and $\epsilon_0=10$, which places ω_L close to ω_T: $\omega_L/\omega_T-1 \approx 0.5\times10^{-3}$. The dispersion parameter D/c^2 is taken as $\pm10^{-5}$, much larger in magnitude than the D_T/c^2 in (26) which is -5×10^{-9}. The physical rationale for this change is that the coupling between nearby excitonic dipoles is dominated by "mechanical," rather than electrodynamic effects. The numerical consequences of these choices are considerable. The nonretarded solutions which were sufficient in Sec. II B are now only rough guides over the frequency range of interest. Near the avoided crossing at ω_T, β_1 and β_2 are comparable and neither is close to the β_1^F found from (5) using the local limit of (30).

To find the exact r's one matches partial waves at $x=0$ using the ABC that all components of the polarization $\mathbf{P}$ vanish there.[23] This yields[20]

$$r_s=\frac{\beta_0-\beta_1+(\epsilon_1-1)\dfrac{\omega^2/c^2}{\beta_2+\beta_0}}{\beta_0+\beta_1+(\epsilon_1-1)\dfrac{\omega^2/c^2}{\beta_2-\beta_0}} , \tag{31}$$

$$r_p = r_s \frac{\epsilon_2(Q^2+\beta_1\beta_3)(\beta_1+\beta_0)-\epsilon_1(Q^2+\beta_2\beta_3)(\beta_2+\beta_0)}{\epsilon_2(Q^2+\beta_1\beta_3)(\beta_1-\beta_0)-\epsilon_1(Q^2+\beta_2\beta_3)(\beta_2-\beta_0)} . \tag{32}$$

Here β_0 is given by (3) and for $j=1,2$, $\epsilon_j=\epsilon(\mathbf{q}=(\beta_j,Q),\omega)$. We compare in Fig. 5 the reflectivities implied by (31) and (32) with the simplest Fresnel results based on (1) and (2) and using

$$\epsilon(\omega)=\epsilon_0+\frac{S\omega_T^2}{\omega_T^2-\tilde{\omega}^2} . \tag{33}$$

In both (29) and (33), $\tilde{\omega}^2=\omega^2+i\omega\gamma$ and we now set $\gamma/\omega_T=10^{-5}$. We did not show analogues of Fig. 5 in Secs. II A and II B because the differences between exact and Fresnel results are scarcely apparent there. But here the differences are large and one should not expect that d-parameter theory which is based on the lowest-order correction will be able to explain them.

The simplest estimates of d's are the nonretarded equations

$$d_{\parallel}^0=\frac{i}{\beta_2^0} , \tag{34}$$

$$d_{\perp}^0=\frac{i}{\beta_3^0} , \tag{35}$$

where

$$(\beta_2^0)^2=\frac{\tilde{\omega}^2-\omega_T^2}{D} , \tag{36}$$

$$(\beta_3^0)^2=\frac{\tilde{\omega}^2-\omega_L^2}{D} . \tag{37}$$

These are compared in Figs. 6 and 7 with two different sets of effective d's. For the first set, denoted by superscript a, we require Eqs. (5) and (6) using (33) for ϵ to exactly reproduce (31) and (32). We find from (5)

$$d_{\parallel}^{(a)}=\frac{i}{\beta_2+\beta_0} , \tag{38}$$

which is fairly simple and clearly tends to (34) as $c\to\infty$. When this result is substituted in (6), the implied $d_{\perp}^{(a)}$ does not reduce to a simple algebraic form, but is readily evaluated. Since much of the numerical difficulty near ω_T is caused by the large value of ϵ, we also consider a form of the r's found at an intermediate step in the d-parameter derivation:[1]

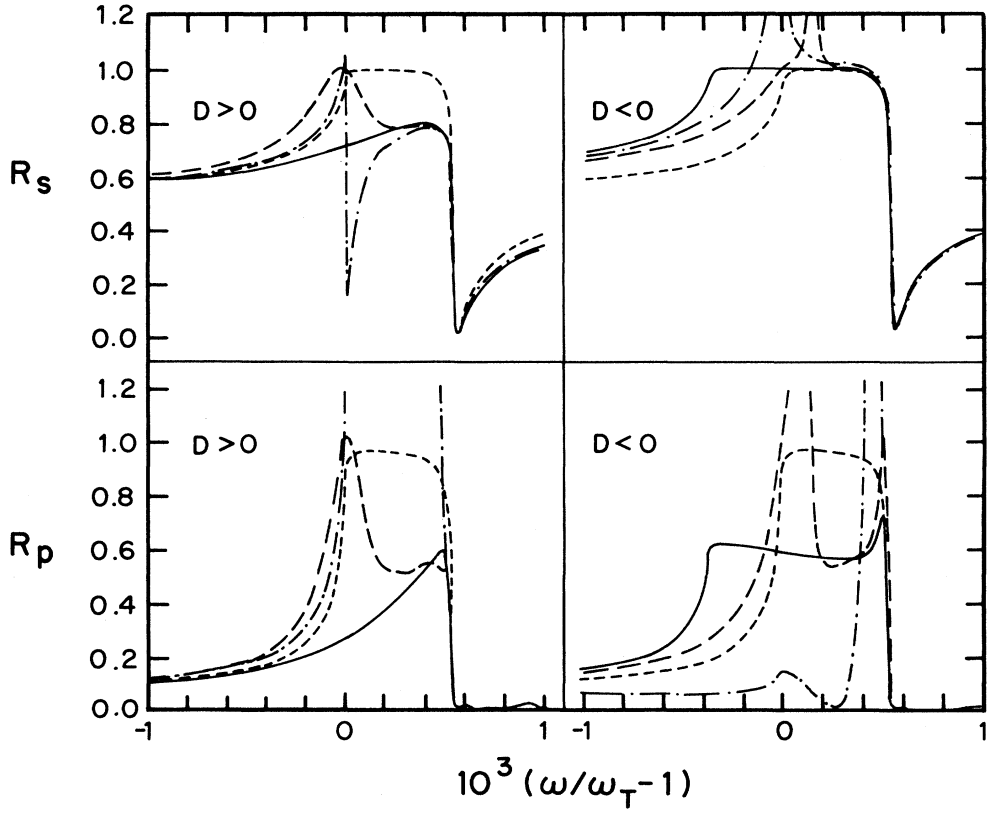

FIG. 5. Various estimates of the reflection coefficients, $R=|r|^2$, for the exciton model at an angle of incidence of 60°. The four panels are distinguished by the polarization of the light and the sign of the spatial dispersion parameter D. We plot vs frequency near the exciton resonance the exact result, solid curve; the Fresnel result, short-dashed curve; the d-parameter estimate based on (5) and (6) using (34) and (35) for the d's, dot-and-dashed curve; and the d-parameter estimate based on (39) and (40) using (34) and (35) for the d's, long-dashed curve. These last two cases do not always yield reflection coefficients less than unity.

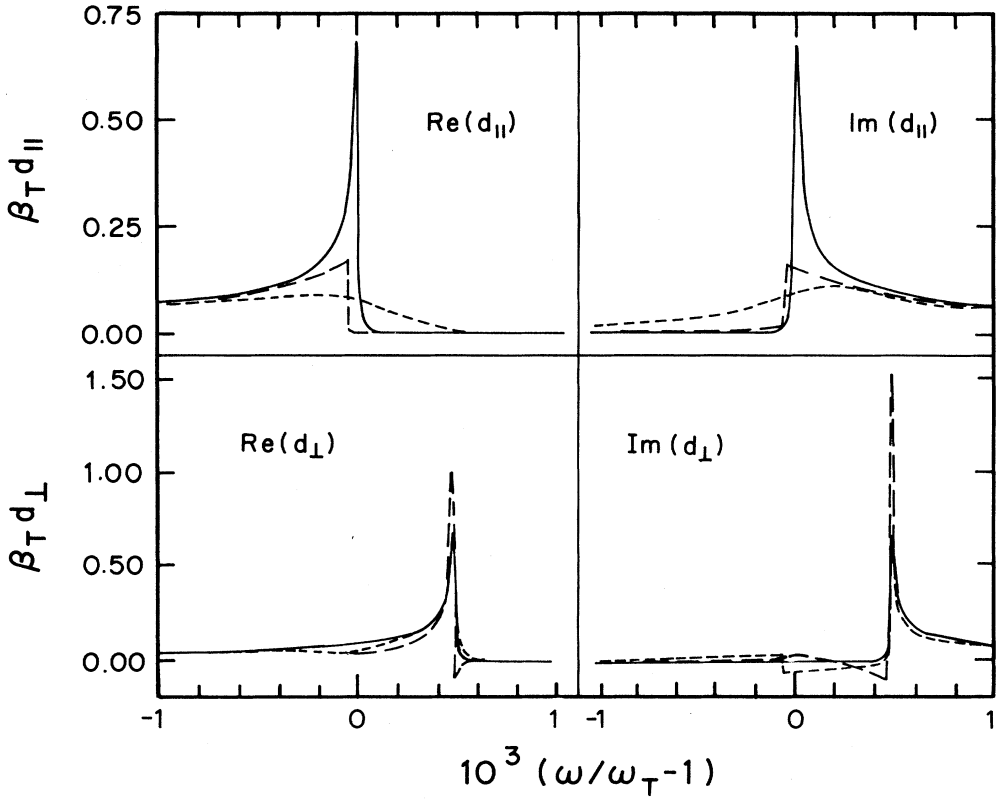

FIG. 6. Real and imaginary parts of $d_\parallel$ and $d_\perp$ in the exciton model vs frequency. The spatial dispersion parameter D is positive. We plot the nonretarded result (34) and (35), solid curve; the $d^{(a)}$'s, long-dashed curve; and the $d^{(b)}$'s, short-dashed curve. The latter two are effective d's, which when used with either (5) and (6) or (39) and (40), respectively, reproduce the exact r's.

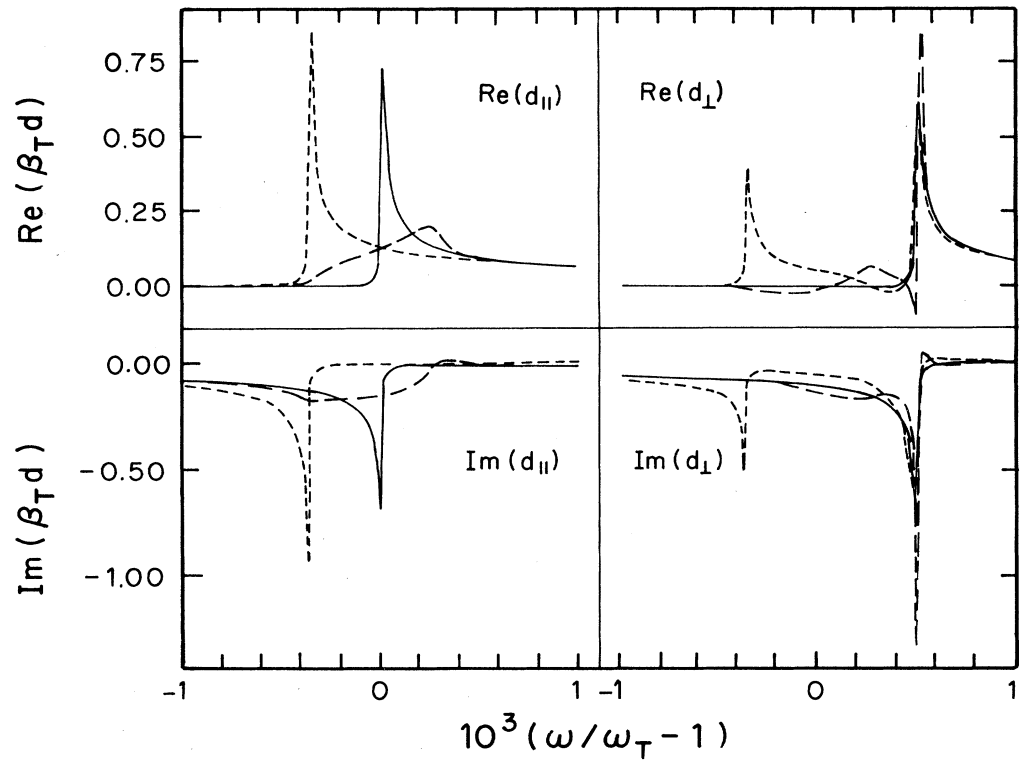

FIG. 7. Same quantities as in Fig. 6, except here D is negative.

$$r_s \approx \frac{\beta_0-\beta_1-i(\epsilon-1)\frac{\omega^2}{c^2}d_\parallel}{\beta_0+\beta_1+i(\epsilon-1)\frac{\omega^2}{c}d_\parallel}\ , \tag{39}$$

$$r_p \approx \frac{\epsilon\beta_0-\beta_1+i(\epsilon-1)(\beta_0\beta_1 d_\parallel - Q^2 d_\perp)}{\epsilon\beta_0+\beta_1+i(\epsilon-1)(\beta_0\beta_1 d_\parallel + Q^2 d_\perp)}\ . \tag{40}$$

Expanding these to first order in the d's reproduces (5) and (6), but if ϵ is divergent such an expansion seems ill-advised. Hence the second set of effective d's, which we denote by a superscript b, are defined by requiring Eqs. (39) and (40) with ϵ_1 substituted for ϵ to agree with (31) and (32). It is again possible to find a simple equation for $d_\parallel$,

$$d_\parallel^{(b)} = \frac{i}{\beta_2+\beta_1}\ , \tag{41}$$

but not for $d_\perp^{(b)}$.

From Figs. 6 and 7 it is evident that near ω_T and ω_L the d^0's, although showing some qualitative similarities to, are quantitatively quite distinct from both the $d^{(a)}$'s and the $d^{(b)}$'s, with $D<0$ being a worse case than $D>0$. The disagreement between the $d^{(a)}$'s and $d^{(b)}$'s is a measure of the error of expanding (39) and (40) into (5) and (6); but using the d^0's in either pair of equations does not account well for the differences between the Fresnel and exact reflectivities shown in Fig. 5. We further checked that all the other approximations in the d-parameter derivations are failing, too. Both β_2 and β_3 are comparable in magnitude to β_1, which is quite different from β_1^F defined via (33). Consequently the higher-order terms omitted from both the numerator and denominator of (39) and (40) are significant. Some qualitative improvement is possible by using effective Fresnel values in (5) and (6); e.g., letting $\epsilon \to \epsilon_1$, but there seems to be no quantitatively reliable way to repair d-parameter theory for the parameter choices used in Figs. 4–7.

III. SUMMARY

By comparison with several exact model solutions we have shown that d-parameter can often work quite well, but occasionally fails badly. The difficulties have only shown up near the threshold for a bulk polariton and over the troublesome frequency range one can reasonably use a continuum model to describe the behavior. Thus in effect (29) applies to all the resonances described in Sec. II, with different choices of parameters. From this point of view the question of whether d-parameter theory will be reliable hinges on the size of γ/ω_T and D/c^2. Either large damping or small spatial dispersion can preserve the accuracy of d-parameter theory through a resonance. When these conditions do not hold there does not seem to be a simple way to fix the theory, aside from going to an exact solution.

The models examined here all possess considerable simplifications, most of which are not relevant to the issue of the validity of d-parameter theory. For instance, allowing for anisotropies or near-surface modifications of coupling parameters or for different ABC's should not change our conclusions. The problems with the theory arise from bulk behavior too dissimilar from the Fresnel model, specifically when the disturbances away from the interface are not well described by a single, or narrow band of wave vectors. For the models treated here this problem occurs because β_2 and/or β_3 become comparable to β_1. More generally it would also arise when the frequency exceeds the threshold for diffraction.

It may be possible in cases where a single β_1^F is not evident to still apply the theory. We have in mind the anomalous skin-effect regime, where although Fresnel theory is wrong the disturbance that extends into bulk has only long-wavelength components. One need only reformulate the theory so that the zeroth-order result is, say, the standard anomalous skin-effect theory, with the corrections arising from variations on a spatial scale much smaller than either the skin depth or mean free path. Such a formalism is developed in Ref. 24.

Our conclusion is that although bulk spatial dispersion can cause serious difficulties for d-parameter theory, the problems are only evident over a limited frequency range. Outside this range the theory provides a tractable and accurate description, but inside it the theory is inadequate and alternate approaches are required.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF) through Grant No. DMR-85-12709. The work of one of us (W.C.) was also supported in part by the Chester Davis Foundation of Indiana University.

[1]W. L. Schaich and Wei Chen, Phys. Rev. B **39**, 10 714 (1989).
[2]D. V. Sivukhin, Zh. Eksp. Teor. Fiz. **18**, 967 (1948).
[3]D. V. Sivukhin, Zh. Eksp. Teor. Fiz. **21**, 367 (1951).
[4]D. V. Sivukhin, Zh. Eksp. Teor. Fiz. **30**, 374 (1956) [Sov. Phys.—JETP **3**, 269 (1957)].
[5]P. J. Feibelman, Phys. Rev. B **23**, 2629 (1981).
[6]P. Apell, Phys. Scr. **24**, 795 (1981).
[7]P. J. Feibelman, Prog. Surf. Sci. **12**, 287 (1982).
[8]K. Kempa and R. R. Gerhardts, Surf. Sci. **150**, 157 (1985).
[9]W. L. Schaich and K. Kempa, Phys. Scr. **35**, 204 (1987).
[10]K. Kempa, A. Liebsch, and W. L. Schaich, Phys. Rev. B **38**, 12 645 (1988).
[11]K. Kempa and W. L. Schaich, Phys. Rev. B **39**, 13 139 (1989).
[12]F. Forstmann and R. R. Gerhardts, *Metal Optics Near the Plasma Frequency* (Springer, Berlin, 1986).
[13]G. D. Mahan and G. Obermair, Phys. Rev. **183**, 834 (1969).
[14]J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[15]Wei Chen and W. L. Schaich, Surf. Sci. (to be published).
[16]J. E. Sipe and J. van Kranendonk, Can. J. Phys. **53**, 2095

(1975).
[17]C. A. Mead, Phys. Rev. B **15**, 519 (1977).
[18]C. A. Mead and M. R. Philpott, Phys. Rev. B **17**, 914 (1978).
[19]M. R. Philpott, J. Chem. Phys. **60**, 1410 (1974); **60**, 2520 (1974).
[20]M. R. Philpott, Phys. Rev. B **14**, 3471 (1976).
[21]N. Kar and A. Bagchi, Solid State Commun. **33**, 645 (1980).
[22]A. P. Lehnen and L. W. Bruch, Physica **100A**, 215 (1980).
[23]S. I. Pekar, Zh. Eksp. Teor. Fiz. **33**, 1022 (1957) [Sov. Phys.—JETP **6**, 785 (1958)].
[24]W. L. Mochan, R. Fuchs, and R. G. Barrera, Phys. Rev. B **27**, 771 (1983).