Aspects of spatial dispersion in the optical properties of a vacuum-dielectric interface*

David Linton Johnson and Peter R. Rimbey
Ames Laboratory-ERDA and Department of Physics, Iowa State University, Ames, Iowa 50010
(Received 29 March 1976)

We have examined the relationship between the polarizability for a two-phase (vacuum-dielectric) system and the use of additional boundary conditions and the like, as regards the response of systems exhibiting spatial dispersion. As a consequence we are able to derive information about induced-charge and current densities and the continuity of the field quantities across the interface. It is shown that it is not possible to resonantly excite longitudinal bulk modes with incident light in the formalism of Rimbey-Mahan. We have derived sum rules in wave-vector space on bulk polaritons in homogeneous isotropic systems. In the case of nonhomogeneous perfect crystals in which the bulk response is described by the matrix $\vec{\epsilon}(\vec{Q},\vec{Q}')$, we have solved formally for the surface impedance in terms of an assumed arbitrary $\vec{\epsilon}(\vec{Q},\vec{Q}')$, by means of an extension of the Fuchs-Kliewer formalism.

I. INTRODUCTION

Spatial dispersion in a medium (or nonlocal optical response) occurs whenever the induced polarization or current density at some point depends not only on the value of the electric field at that point but also on the value of $\vec{E}$ in a region about that point.[1]

$$\vec{P}(\vec{r}) = \int \vec{\chi}(\vec{r},\vec{r}')\cdot\vec{E}(\vec{r}')\,d^3r'. \tag{1.1}$$

If the medium is homogeneous and infinite so that $\vec{\chi}(\vec{r},\vec{r}')=\vec{\chi}(\vec{r}-\vec{r}')$, Eq. (1.1) can be Fourier transformed as

$$\vec{P}(\vec{q}) = \vec{\chi}(\vec{q})\cdot\vec{E}(\vec{q}). \tag{1.2}$$

Natural optical activity (gyrotropy),[2] the anomalous skin effect,[3] the unusual reflectivity spectra of crystals in the classically forbidden region near an isolated exciton resonance,[4] and the resonant excitation of longitudinal modes (plasmons) in thin films by p-polarized light[5-7] are all interesting consequences of spatial dispersion in the dielectric function $\vec{\epsilon}(\vec{q}) = \vec{1} + 4\pi\vec{\chi}(\vec{q})$.

In a real experiment, there is always a crystal-ambient interface so that translational symmetry is destroyed and (1.2) is invalid even if the dielectric is homogeneous. In order to handle this problem various schemes have been devised using additional boundary conditions[1,4,6] (ABC's), symmetry conditions,[8,9] or, more recently, there have been attempts to solve directly for the response by neglecting altogether the effects of the surface.[10] Roughly speaking, the problem arises because in the presence of spatial dispersion there are multiple bulk modes (polaritons) of a given frequency and polarization but having different wave vectors, and the surface couples externally incident light to each of them. In this way of thinking, Maxwell's equations do not provide sufficient boundary conditions at the surface to solve uniquely for the fields in the dielectric although it is equally clear that if the polarizability $\vec{\chi}$ in (1.1) is known, any problem of interest should be exactly solvable without the need for *ad hoc* boundary conditions. In Sec. II we point out that most of the commonly used attempts to solve problems of this nature are each equivalent to a particular, if implicit, assumption about the relationship of the polarizability $\vec{\chi}$ of the two-phase system[11] to that of the bulk; the two are not, in general, equal. The various ABC's are shown to be a consequence of the particular polarizability used. In particular, the method of Ref. 10 originates from the same general formalism as the method of ABC's; each formalism derives from different values of a single parameter indicative of a particular choice of macroscopic susceptibility functions [Eqs. (2.10), below]. However, in principle this parameter must be determined from microscopic considerations. We note that the method of symmetry-imposed conditions is particularly attractive, when applicable, because the surface impedance, which can be simply used to solve many electromagnetic problems of interest, can be evaluated by a straightforward integration over wave vector of an arbitrary bulk dielectric function, regardless of the number of different polariton modes. The nature of the two-phase polarizability in turn determines the nature of surface charge or current layers and the continuity of $\vec{E}$ and $\vec{B}$ across the interface. Some of the isolated results of Sec. II have been published[12,13] and others have been communicated to us privately. These results have not previously been derived from the gen-

eral formalism of Sec. II which has the advantage that we are able to deduce properties of the system quite generally, i.e., without having to actually solve the problem in specific instances. Throughout this article we do not wish to comment on the validity of the various schemes; instead we are focusing our attention on their consequences.

In Sec. III, the wave equation in terms of effective surface current and magnetic current densities is derived from the general polarizability function of Sec. II. We show that it is not possible to resonantly excite a longitudinal bulk mode in the Rimbey-Mahan formalism[9]; it has been demonstrated to occur for the Fuchs-Kliewer formalism,[7,8] and the effect has been observed in thin metallic films.[5] This provides a possible experimental test of the validity of the various boundary conditions used. We have explicitly derived the equivalence of, e.g., the Rimbey-Mahan formalism to the method of partial waves used by Pekar,[1] Hopfield and co-workers,[4] and subsequent investigators. We have also demonstrated that, although there may be more than one polariton of a given frequency in the dielectric (bulk), the "strengths" of these polaritons sum to unity.

In Sec. IV, the response of a two-phase system in which the dielectric is a perfect crystal but nonhomogeneous is derived from an extension of the Fuchs-Kliewer formalism, i.e., we have solved for the surface impedance (in general, a matrix) a knowledge of which allows one to solve for either the reflectivity or the dispersion of a surface excitation.[14,15]

II. BOUNDARY CONDITIONS AND THE TWO-PHASE POLARIZABILITY TENSOR

A. Relationship of the two-phase polarizability tensor to the bulk polarizability tensor

Consider a two-phase system consisting of vacuum in the half-space $z<0$ and some dielectric medium in the half-space $z>0$ which we need *not* assume to be homogeneous. We will assume, however, that $z=0$ represents a sharp boundary between the two media and that the "properties" of the crystal persist up to the surface. We will also assume that the plane $z=0$ would be a mirror plane if the crystal were continued into $z<0$. The significance of this last assumption will become apparent later. In cases where one assumes the dielectric to be homogeneous and isotropic, this represents no additional restriction because every plane is then a symmetry plane. Quite generally, one may relate the induced polarization density for the two-phase system to the electric field as

$$\vec{P}(\vec{r}) = \int_{z'>0} \vec{\chi}_{vs}(\vec{r},\vec{r}')\cdot\vec{E}(r')\,d^3r'. \tag{2.1}$$

It will be convenient to assume that all fields vary as $e^{-i\omega t}$. $\vec{P}$ is defined in terms of the induced current density as $\vec{P}=(i/\omega)\vec{J}$, and we have defined $\vec{\chi}$ in terms of the fields[16] rather than the potentials[17] so as to avoid divergences in reciprocal space. Unless $\vec{\chi}_{vs}$ is a δ function, the response of the system is nonlocal (hence the term spatial dispersion). There is, in general, more than one normal-mode solution to Maxwell's equations (bulk polariton) for a given frequency. Early attempts to solve for the reflectivity of the two-phase system assumed the field for $z>0$ could be expanded in terms of bulk normal modes, and to do this one needed an additional boundary condition (ABC), over and above those given by Maxwell's equations,[1] in order to solve uniquely for, e.g., the reflectivity $R(\omega)$. Intuitively, one ought to be able to solve for any electromagnetic quantity of interest once $\vec{\chi}_{vs}$ in Eq. (2.1) is specified, even though $\vec{\chi}_{vs}$ is not a true response function. Actually, all attempts to solve for the nonlocal response (the surface impedance, say, or the reflectivity) involve an implicit assumption about the relationship of the vacuum-solid polarizability tensor $\vec{\chi}_{vs}$ to that of the bulk polarizability $\vec{\chi}_B$; this latter quantity is much simpler to calculate directly from some realistic model (the Lindhard function,[18] one-oscillator models,[4] etc.) than is the former. The assumption is present in articles[8,9] employing a "mirroring" of the crystal into the region $z<0$, in articles[1,4,6] employing ABC's, and it is *implicitly* assumed in recent articles purporting to have avoided the use of additional boundary conditions altogether.[10] This relationship has been recognized before[19-22]; we wish to amplify on that relationship, keeping strict track of the tensorial character of $\vec{\chi}_{vs}$, so that we may draw conclusions about the continuity of the fields at $z=0$, induced-charge and current densities and the like. Neglect of the tensorial character of $\vec{\chi}_{vs}$ suffices for normal-incidence light on an isotropic medium when all vector quantities are assumed parallel to the surface; however, it fails to correctly determine the nature of any field components normal to the surface, as occur in the non-normal incidence of light (especially photoemission[23]) or in surface excitations.

Consider, for example, the method employed by Fuchs and Kliewer[8] (FK) for specular reflection of electrons at the inner surface of a metal. This is equivalent to the assumptions[20]

$$\vec{\chi}_{vs}(\vec{r},\vec{r}') = \begin{cases} \vec{\chi}_B(xyz;x'y'z') \\ +\vec{\chi}_B(xyz;x'y'-z')\cdot\vec{\alpha}, & z,z'>0, \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

where $\vec{\vec{\chi}}_B(\vec{r},\vec{r}')$ is the bulk polarizability (which we do not assume to be homogeneous) and $\vec{\vec{\alpha}}$ is the matrix for mirroring across $z=0$:

$$\vec{\vec{\alpha}}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}=\vec{\vec{\alpha}}^{-1}. \tag{2.3}$$

Substituting in (2.1) and changing the dummy variable $z'\to -z'$ in the second term, one can write

$$\vec{P}(\vec{r})=\int_{\substack{\text{all}\\\text{space}}}\vec{\vec{\chi}}_B(\vec{r},\vec{r}')\cdot\vec{E}_{\text{eff}}(\vec{r}')\,d^3r', \tag{2.4}$$

where

$$\vec{E}_{\text{eff}}(xyz)=\begin{cases}\vec{E}(xyz), & z>0\\ \vec{\alpha}\cdot\vec{E}(xy-z), & z<0\\ [E_x(xy-z),E_y(xy-z),-E_z(xy-z)], & z<0.\end{cases} \tag{2.5}$$

The problem has now been recast into a form in which, as far as the right half-space ($z>0$) is concerned, the left half-space ($z<0$) appears to be filled with the same polarizable medium and with an electric field mirrored according to (2.5). As pointed out,[8] this guarantees that if free electrons are the polarizable entities, they will be specularly reflected from the inner surface. We wish to emphasize that (2.4) and (2.5) are a mathematical trick from which one must not directly draw conclusions about real surface-charge densities or the continuity of $\vec{E}$. Similarly, Rimbey and Mahan[9] (RM) have proposed a treatment applicable to specular reflection of (Frenkel[19]) excitons from the surface, which is equivalent to

$$\vec{\vec{\chi}}_{vs}(\vec{r},\vec{r}')=\vec{\vec{\chi}}_B(xyz;x'\,y'\,z')-\vec{\vec{\chi}}_B(xyz;x'\,y'-z')\cdot\vec{\vec{\alpha}},\quad z,z'>0; \tag{2.6}$$

this is of the form (2.4) if

$$\vec{E}_{\text{eff}}(xyz)=[-E_x(xy-z),-E_y(xy-z),+E_z(xy-z)],\quad z<0. \tag{2.7}$$

Agarwal *et al.*, Maradudin and Mills, and Birman and Sein[10] have claimed to have avoided the need for any ABC, mirroring, etc. However, their treatment is simply the approximation

$$\vec{\vec{\chi}}_{vs}(\vec{r},\vec{r}')=\vec{\vec{\chi}}_B(\vec{r},\vec{r}'),\quad z,z'>0 \tag{2.8}$$

or

$$\vec{E}_{\text{eff}}(xyz)\equiv 0,\qquad z<0. \tag{2.9}$$

It appears that this assumption neglects the reflection of the excitation (exciton or electron) from the surface; it can be shown, however, that (2.8) applied to a vacuum–homogeneous-dielectric system leads to an expression for the surface impedance of s-polarized light which is identical with that obtained for diffuse scattering.[24] Equations (2.2) and (2.5)–(2.9) may be summarized as

$$z,z'>0,\quad \vec{\vec{\chi}}_{vs}(\vec{r},\vec{r}')=\vec{\vec{\chi}}_B(\vec{r},\vec{r}')+U\vec{\vec{\chi}}_B(xyz;x'\,y'-z')\cdot\vec{\vec{\alpha}} \tag{2.10a}$$

$$=\vec{\vec{\chi}}_B(\vec{r},\vec{r}')+U\vec{\vec{\alpha}}\cdot\vec{\vec{\chi}}_B(xy-z;x'\,y'\,z'). \tag{2.10b}$$

The parameter $U=+1$ for the FK formalism,[8] $U=-1$ for the RM formalism,[9] and $U=0$ for the formalism of Ref. 10. Equation (2.10b) holds only if the $z=0$ plane would be mirror plane of the bulk[16]; if $\{\vec{\vec{\alpha}}|\vec{\tau}\}$ is any member of the space group of the crystal then

$$\vec{\vec{\alpha}}^{-1}\cdot\vec{\vec{\chi}}_B(\vec{\vec{\alpha}}\cdot\vec{r}+\vec{\tau},\vec{\vec{\alpha}}\cdot\vec{r}'+\vec{\tau})\cdot\vec{\vec{\alpha}}=\vec{\vec{\chi}}_B(\vec{r},\vec{r}').$$

We note that it does not make sense to speak of specular reflection of the excitation at all unless $z=0$ is a symmetry plane. This does not represent a restriction in cases where a cubic crystal is presumed isotropic and homogeneous because every plane is a symmetry plane. Similarly

$$\vec{E}_{\text{eff}}(xyz)=U\vec{\vec{\alpha}}\cdot\vec{E}(xy-z),\quad z<0 \tag{2.11}$$

and since $\vec{B}=(1/i\omega)\nabla\times\vec{E}$

$$\vec{B}_{\text{eff}}(xyz)=-U\vec{\vec{\alpha}}\cdot\vec{B}(xy-z),\quad z<0. \tag{2.12}$$

Although there are many treatments of the nonlocal response of a two-phase system which do not fit (2.10a),[25] we will show that many of the formalisms are of the form (2.10a) or can be simply modified therefrom.

B. Limiting values $P_i(0^+)$, $dP_i(0^+)/dx_j$, and the continuity of $\vec{E},\vec{B}$ across the surface

It is of interest to see under what conditions the fields are continuous across the boundary $z=0$. Of course, the source-free Maxwell's equations guarantee that B_z and $E_{x,y}$ are continuous. Consider

$$c^2\nabla\times\vec{B}=4\pi\vec{J}_{\text{ind}}+\partial\vec{E}/\partial t \tag{2.13a}$$

$$=4\pi\frac{\partial\vec{P}}{\partial t}+\frac{\partial\vec{E}}{\partial t}. \tag{2.13b}$$

By Stokes's theorem, the tangential component of $\vec{B}$ is continuous whenever $P_{x,y}$ (and $E_{x,y}$) are finite at the surface. According to (2.13) and (2.1), $P_{x,y}(z=0)=P_{x,y}(z\to 0^+)$ whenever $\vec{\vec{\chi}}_{vs}$ is finite, or at least no worse than a δ function.

($|\vec{E}|$ may be assumed to decay into the bulk.) A *counter* example occurs for the classical local treatment of a homogeneous magnetic system,

$$\vec{M} = (1/4\pi)[c^2 - 1/\mu(\omega)]\vec{B} = \gamma(\omega)\vec{B},$$

$$\vec{P}_M(r) = \frac{i}{\omega}\vec{J}_M(r) = \frac{i}{\omega}\nabla\times\vec{M} = \frac{i\gamma(\omega)}{\omega}\nabla\times\vec{B} \quad (2.14)$$

$$= [\gamma(\omega)/\omega^2]\nabla\times\nabla\times\vec{E},$$

which has a Fourier transform

$$\vec{P}_M(q) = -[\gamma(\omega)/\omega^2]\vec{q}\times\vec{q}\times\vec{E}. \quad (2.15)$$

The (transverse) polarizability tensor has a contribution $\chi_M \sim \gamma q^2/\omega^2$ which means that

$$\chi_B(r) \sim \int \frac{\gamma q^2}{\omega^2} e^{i\vec{q}\cdot\vec{r}}\, d^3r \quad (2.16)$$

diverges worse than a δ function. In this case there is a surface current density $\vec{J}_M(\vec{r}) = \vec{M}_{x,y}\delta(z)$ and the tangential component of $\vec{B}$ is not continuous. (For this reason, it is customary to define $\vec{H} = c^2\vec{B} - 4\pi\vec{M}$ whose tangential component is continuous.) In the following, we will assume that $\vec{\chi}_B$ is no worse than a δ function; for an isotropic system this is equivalent to assuming $\lim_{q\to\infty}\chi(q)$ is finite. It is clear that the long-wavelength magnetic susceptibility in an isotropic system, $\mu(\omega)$, is derivable through (2.15) from $[d^2\epsilon_t(q,\omega)/dq^2]_{q=0}$, i.e.,

$$\frac{1}{8\pi}\left(\frac{d^2\epsilon_t}{dq^2}\right)_{q=0} = \frac{\gamma(\omega)}{\omega^2} = \frac{1}{4\pi\omega^2}\left(c^2 - \frac{1}{\mu(\omega)}\right). \quad (2.17)$$

We see that any material exhibiting spatial dispersion is "magnetic"; however, our treatment becomes invalid only if $\epsilon_t(q)$ diverges for large q.

The continuity of $E_z(z\to 0^+)$ is different for the three cases $U = \pm 1, 0$. Since there are no external charges introduced in the system

$$\nabla\cdot\vec{E} = -4\pi\nabla\cdot\vec{P}, \quad (2.18)$$

and

$$E_z(xy0^+) - E_z(xy0^-) = -4\pi[P_z(xy0^+) - P_z(xy0^-)]$$

$$= -4\pi P_z(xy0^+), \quad (2.19)$$

which means there is an induced surface-charge density

$$\sigma_s(xyz) = -4\pi P_z(xy0^+)\delta(z) \quad (2.20)$$

in addition to any induced bulk-continuum charge density. To compare the three different cases, we first assume that $\vec{\chi}_B(\vec{r},\vec{r}')$ is well behaved, i.e., that it is finite, continuous, differentiable and has no δ-function (or worse) contribution. We will come back to the case of a background local susceptibility later. The second term of (2.10a) evaluated at the surface gives

$$\vec{\chi}_B(xy0; x'y'-z')\cdot\vec{\alpha} = \vec{\alpha}\cdot\vec{\alpha}^{-1}\vec{\chi}_B[\alpha(xy0); \alpha(x'y'z')]\cdot\vec{\alpha}$$

$$= \vec{\alpha}\cdot\vec{\chi}_B(xy0; x'y'z'), \quad (2.21)$$

and it is at this point that we have used the assumption that $z = 0$ represents a mirror plane of the crystal. Equation (2.1) gives

$$\vec{P}(xy0^+) = (\vec{I} + U\vec{\alpha})\cdot\int_{z'>0}\vec{\chi}_B(xy0; x'y'z')\cdot\vec{E}(x'y'z')\,d^3r', \quad (2.22)$$

which clearly implies ($U = +1$, FK)

$$P_{x,y}(xy0^+) \neq 0, \quad (2.23a)$$

$$P_z(xy, 0^+) \equiv 0. \quad (2.23b)$$

In general, then, the tangential components of $\vec{P}$ (or $\vec{J}$) are not continuous in the FK formalism (though they may be in special cases) but the normal component always is. As a consequence of the latter, the normal as well as tangential components of $\vec{E}$ are continuous at the surface and there is never a real induced surface-charge density, in the FK formalism, if $z = 0$ represents a mirror plane of the crystal. Similarly from (2.10b) with $U = +1$ (FK)

$$\lim_{z\to 0^+}\left(\frac{\partial P_x}{\partial z} = \frac{\partial P_y}{\partial z} = \frac{\partial P_z}{\partial x} = \frac{\partial P_z}{\partial y}\right) \equiv 0, \quad (2.24)$$

and the remaining derivatives evaluated as $z \to 0^+$ are not zero except in special cases. It will be noted that $\partial P_x(0^+)/\partial z = 0$ is identical with Eq. (16) of Ref. 21 in which they have in effect derived (2.2) for a model Wannier exciton; their formalism is therefore equivalent to Fuchs-Kliewer for normal incidence. Ting *et al.*[21] have, however, approximated $\vec{\chi}_B$ as a scalar so that for non-normal incidence their results will be quite different.

Similarly, if $U = -1$ (RM) in (2.10b) it is easy to see that

$$P_{x,y}(x, y, 0^+) = 0, \quad (2.25a)$$

$$P_z(x, y, 0^+) \neq 0, \quad (2.25b)$$

and

$$\lim_{z\to 0^+}\left(\frac{\partial P_{x,y}}{\partial_{x,y}} = \frac{\partial P_z}{\partial_z}\right) = 0. \quad (2.26)$$

Equation (2.25a) was originally considered by Pekar to be the additional boundary condition pertinent to the Frenkel excitonic problem; indeed it has been derived from a model Schrödinger equation for excitons,[19] again in the approximation that $\vec{\chi}_{vs} = \chi_{vs}\vec{I}$. We have seen here that the boundary condition $P_x = 0$ is equivalent to the choice (2.6). Equation (2.25b) demonstrates

that there is an induced surface-charge density given by (2.20) except in the case of s-polarized light. The normal component of the electric field is in general, then, discontinuous. Equations (2.26) together imply that the induced charge density near (but not at) the inner surface is zero.

$$\lim_{z\to 0^+} \rho(xyz) = -4\pi\nabla\cdot\vec{P}(xy0^+) = 0. \tag{2.27}$$

A much stronger result for the case of a homogeneous cubic dielectric will be discussed in Sec. III.

Note that in the case $U \neq \pm 1$, nothing in general can be said about the limiting values for any of the aforementioned quantities; the generalized extinction theorem[10,15] does apply, but to evaluate it one needs specific and analytic information about the functional form of $\bar{\chi}_B(\vec{r}, \vec{r}')$ and this must be done on a case by case basis. The case $U = 0$ has been evaluated in terms of the two-phase eigenmodes. We can see that the induced surface-charge density is, in this case, formally half that of RM *if* the electric fields are assumed to be identical for $z > 0$.

$$\sigma(\text{surface}) = -4\pi P_z(xy0^+) = -4\pi\hat{n}\cdot\int_{z'>0}\bar{\chi}_B(xy0, x'y'z')\cdot\vec{E}(x'y'z')\,d^3r' \qquad (U=0). \tag{2.28}$$

We note that $U \neq \pm 1$ does not conform to an ABC of the form $AP_x + BdP_x/dz = 0$ except in special cases. This was explicitly demonstrated for $U = 0$, in which the ABC depends on the eigenmode.[10] We note also that Eqs. (2.10) do not correspond to the method of the "dead layer" either.[4]

The results (e.g., reflectivity or surface-excitation dispersion) for different values of U are, in general, quite different for a given $\chi_B(r, r')$. In all cases, the polarizability tensor has the property

$$\lim_{z,z'\to+\infty} \bar{\chi}_{vs}(\vec{r}, \vec{r}') = \bar{\chi}_B(\vec{r}, \vec{r}'), \tag{2.29}$$

and so it is not true that one can neglect the difference between $\bar{\chi}_{vs}$ and $\bar{\chi}_B$ just because that difference is appreciable only near the surface. In fact, in many cases of interest the bulk damping parameter is so small that the contribution from the second term on the right-hand side of (2.10a) is comparable to that from the first.

It must not be thought that because $J_z(P_z)$ does not go to zero at the surface for $U \neq +1$ that this necessarily represents a gross unphysicality in the formalism. It does not imply that the induced charge is jumping out of the surface—only that it is piling into a surface layer of zero thickness. This surface charge density is simply a result of imposing an infinitesimally sharp boundary separating vacuum and bulk dielectric. The fact that one has attempted a partial microscopic treatment of the e – m response does not necessarily negate the consequences of a sharp boundary. For example, the surface charge density always associated with a local response function

$$\bar{\chi}_{vs} = \bar{\chi}_{\text{loc}}\,\delta(\vec{r} - \vec{r}'), \quad z' > 0 \tag{2.30}$$

is eliminated with a smoothly varying, position dependent, but local, response function

$$\bar{\chi}_{vs}(\vec{r}, \vec{r}') = \bar{\chi}_{\text{loc}}(z')\delta(\vec{r} - \vec{r}') \tag{2.31}$$

in which $\bar{\chi}_{\text{loc}}(z' \to -\infty) = 0$ and $\bar{\chi}_{\text{loc}}(z' \to +\infty) = \bar{\chi}_B$ (constant); there is no discontinuity in P_z and therefore no true surface charge density. We return to this point in Sec. III B.

Is there a situation in which $\vec{P}(0^+) \equiv 0$, i.e., all components vanish at the surface, as originally envisioned by Pekar?[1] Formally, this situation can exist if

$$\bar{\chi}_{vs}(\vec{r}, \vec{r}') = \bar{\chi}_B(\vec{r}, \vec{r}') - \bar{\alpha}\cdot\bar{\chi}_B(\vec{r}, \bar{\alpha}\cdot\vec{r}')\cdot\bar{\alpha}. \tag{2.32}$$

If one further assumes $\bar{\chi}_B = \chi_B\bar{I}$, then

$$\bar{\chi}_{vs} = [\chi_B(\vec{r}, \vec{r}') - \chi_B(xyz, x'y' - z')]\bar{I}, \tag{2.33}$$

which is equivalent to the (RM) formalism (2.6) or (2.7) applied to normally incident light on a homogeneous dielectric but not otherwise. Similarly, the + sign in (2.33) would guarantee $d\vec{P}(0^+)/dz = 0$. Equation (2.32) cannot be cast in the form (2.4), (2.11); it does not represent specular reflection of the excitation for the simple reason that the polarization/current density (a vector) is not the same as, e.g., the exciton wave function (a scalar) although the two are closely related. It is not mathematically possible to choose the phase of *any* reflected p-polarized or longitudinal vector wave such that all components vanish at the surface (except at normal incidence). In articles employing the approximation $\bar{\chi}_B = \chi_B\bar{I}$ there is therefore some ambiguity as to how one should proceed for non-normally incident light.

Finally, we consider the useful situation in which part of the response is considered local and part is nonlocal, e.g.,

$$\bar{\chi}_B(\vec{r}, \vec{r}') = \bar{\chi}_L\delta(\vec{r} - \vec{r}') + \bar{\chi}_{NL}(xyz, x'y'z'), \tag{2.34}$$

which implies

$$\vec{P}(\vec{r}) = \bar{\chi}_L \cdot \vec{E}(\vec{r}) + \int [\bar{\chi}_{NL}(\vec{r},\vec{r}') + U\bar{\chi}_{NL}(\vec{r},\vec{\alpha}\circ\vec{r}')\cdot\vec{\alpha}]\cdot\vec{E}(\vec{r}')\,d^3r' \tag{2.35a}$$

$$= \vec{P}_L(\vec{r}) + \vec{P}_{NL}(\vec{r}), \tag{2.35b}$$

and therefore

$$\sigma(\text{surface}) = -4\pi\hat{n}\cdot\bar{\chi}_L\cdot\bar{\bar{E}}(xy0^+)\delta(z) \tag{2.36}$$

in addition to whatever is contributed by the remaining terms discussed previously. The quantity $\vec{P}_{NL} = \vec{P} - \vec{P}_L$ obeys the relations derived above [Eqs. (2.23)–(2.26)] as $z \to 0^+$, depending on U. For a homogeneous isotropic dielectric we note that in terms of the Fourier transform of $\bar{\chi}_B(\vec{r}-\vec{r}')$

$$\bar{\chi}_L = \lim_{q\to\infty} \bar{\chi}_B(q) \tag{2.37}$$

and this quantity, the amount of background nonlocal response, determines the entire surface charge density in FK formalism.

A primary conclusion of this section, then, is that the method of ABC's, the method of imposed symmetry conditions, and the method employed in Ref. 10 all originate from the same common formalism [Eqs. (2.10)] but with a different value of the parameter U. The appropriate value of U that makes (2.10) a valid approximation for a given system must be determined from microscopic considerations.

III. SOME RESULTS FOR A VACUUM INTERFACING A HOMOGENEOUS ISOTROPIC DIELECTRIC

A. Wave equation in the two-phase system in terms of effective surface current densities and magnetic current densities

By means of Eqs. (2.10) and (2.11) the two-phase problem may be simply related to an equivalent bulk problem in terms of effective fields (which are the true fields for $z > 0$). We make the following definitions:

$$\vec{D}(\vec{r}) = \vec{E}_{\text{eff}}(\vec{r}) + 4\pi\vec{P}(\vec{r}) = \int \bar{\epsilon}(\vec{r},\vec{r}')\cdot\vec{E}_{\text{eff}}(\vec{r}')\,d^3r', \tag{3.1a}$$

$$\bar{\epsilon}(\vec{r},\vec{r}') = \delta(\vec{r}-\vec{r}')\bar{I} + \bar{\chi}_B(\vec{r},\vec{r}'), \tag{3.1b}$$

$$\vec{\mathcal{E}}(\vec{Q}) = \int_{\text{all space}} \vec{E}_{\text{eff}}(\vec{r})e^{-i\vec{Q}\cdot\vec{r}}\,d^3r, \tag{3.1c}$$

$$\vec{\mathcal{E}}^s(\vec{Q}_t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \vec{\mathcal{E}}(\vec{Q})e^{iQ_z 0^+}\,dQ_z, \tag{3.1d}$$

where $\vec{Q}_t$ is the component of $\vec{Q}$ in the $x-y$ plane

$$\vec{Q} = (\vec{Q}_t, Q_z). \tag{3.1e}$$

It therefore follows that

$$\vec{E}_{\text{eff}}(xyz) = \frac{1}{(2\pi)^3}\int \vec{\mathcal{E}}(\vec{Q})e^{i\vec{Q}\cdot\vec{r}}\,d^3Q, \tag{3.2a}$$

$$\vec{E}_{\text{eff}}(xy0^+) = \frac{1}{(2\pi)^2}\int \vec{\mathcal{E}}^s(\vec{Q}_t)e^{i\vec{Q}_t\cdot(x,y)}\,d^2Q_t. \tag{3.2b}$$

For a homogeneous isotropic dielectric, translational symmetry implies that $\vec{Q}_t$ is an eigenvalue of the problem; whether the problem is the reflectivity of incident light $|\vec{Q}_t| = (\omega/c)\sin\theta$ or the dispersion of a surface excitation $\omega(Q_t)$, different values of $\vec{Q}_t$ do not mix. Moreover, since the plane of incidence (the plane normal to the surface and passing through $\vec{Q}_t$) is a mirror plane, the electric field may be assumed, by group theory, either to lie in this plane or be perpendicular to it. All quantities vary temporally as $e^{-i\omega t}$

Maxwell's equations reduce to

$$\frac{\omega^2}{c^2}\vec{\mathfrak{D}}(\vec{Q}) = \int [\nabla\times\nabla\times\vec{E}_{\text{eff}}(\vec{r})]e^{-i\vec{Q}\cdot\vec{r}}\,d^3r. \tag{3.3}$$

To express (3.3) in terms of $\vec{\mathcal{E}}(\vec{Q})$ one must integrate by parts[8] being careful to handle terms like, e.g. (E_x means x component of $\vec{E}_{\text{eff}}$)

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{\partial^2 E_x}{\partial z^2}e^{-iQ_z z}\,dz &= \int_{-\infty}^{0^-}\frac{\partial^2 E_x}{\partial z^2}e^{-iQ_z z}\,dQ_z + \int_{0^+}^{\infty}\frac{\partial^2 E_x}{\partial z^2}e^{-iQ_z z}\,dz \\ &= (-iQ_z)^2\int_{-\infty}^{\infty}E_x(xyz)e^{-iQ_z z}\,dQ_z + \left(\frac{\partial E_x}{\partial z}e^{-iQ_z z} - E_x(xyz)(-iQ_z)e^{-iQ_z z}\right)_{-\infty}^{0^-} \\ &\quad + \left(\frac{\partial E_x}{\partial z}e^{-iQ_z z} - E_x(xyz)(-iQ_z)e^{-iQ_z z}\right)_{0^+}^{\infty} \\ &= -Q_z^2\int_{-\infty}^{\infty}E_x(xyz)e^{-iQ_z z}\,dQ_z - (1+U)\frac{\partial E_x(xy0^+)}{\partial z} + iQ_z(U-1)E_x(xy0^+).\end{aligned} \tag{3.4}$$

Expressions like $\partial E_x(xy0^+)/\partial z - \partial E_z(xy0^+)/\partial x$ can be eliminated from the right-hand side of (3.3) in terms of the magnetic field. The result is

$$\begin{aligned}(\omega^2/c^2)\vec{\mathfrak{D}}(\vec{Q}) + \vec{Q}\times[\vec{Q}\times\vec{\mathcal{E}}(\vec{Q})] &= (-i\omega/c)(1+U)\hat{n}\times\vec{B}^s(\vec{Q}_t) \\ &\quad - i(1-U)\vec{Q}\times[\hat{n}\times\vec{\mathcal{E}}^s(\vec{Q}_t)],\end{aligned} \tag{3.5}$$

where $\hat{n}=\hat{z}$ is the unit vector normal into the bulk. To this point we have assumed $z=0$ is a mirror plane of the crystal but have not assumed the crystal to be homogeneous. Proponents of diffraction theory will recognize the first term in the right-hand side of (3.5) as due to an *effective* surface current density and the second as the curl of an *effective* magnetic current density.[26] Inasmuch as surface plasmons couple to light waves in the presence of a rough surface, an extension of Eq. (3.5) may prove useful for calculations of the properties of surface plasmons in that regard.

B. Properties of solutions for the case of homogeneous cubic dielectrics

If the bulk dielectric is assumed homogeneous, translational symmetry requires $\vec{\epsilon}(\vec{r}\ \vec{r}')=\vec{\epsilon}(\vec{r}-\vec{r}')$, and the relationship between $\vec{\mathfrak{D}}$ and $\vec{\mathcal{E}}$ is

$$\vec{\mathfrak{D}}(\vec{Q})=\vec{\epsilon}(\vec{Q})\cdot\vec{\mathcal{E}}(\vec{Q}), \tag{3.6}$$

where $\vec{\epsilon}(Q)$ is the Fourier transform of $\vec{\epsilon}(\vec{r})$. Isotropic symmetry further requires (every axis is a symmetry axis)

$$\vec{\epsilon}(\vec{Q})=\epsilon_l(|\vec{Q}|)\hat{Q}\hat{Q}+\epsilon_t(|\vec{Q}|)(\vec{I}-\hat{Q}\hat{Q}). \tag{3.7}$$

In this special case, Eq. (3.5) can be written

$$\left[\frac{\omega^2}{c^2}\epsilon_l(Q)\hat{Q}\hat{Q}+\left(\frac{\omega^2}{c^2}\epsilon_t(Q)-Q^2\right)(\vec{I}-\hat{Q}\hat{Q})\right]\cdot\vec{\mathcal{E}}(\vec{Q}) = \frac{-i\omega}{c}(1+U)\hat{n}\times\vec{B}^s(Q_t) - i(1-U)\vec{Q}\times[\hat{n}\times\vec{\mathcal{E}}^s(Q_t)]. \tag{3.8}$$

The matrix in brackets on the left-hand side can be inverted by inspection; it is essentially the Green's function for the problem.[15]

$$[\ ]^{-1}=\frac{c^2}{\omega^2\epsilon_l}\hat{Q}\hat{Q}+\frac{c^2}{\omega^2\epsilon_t-c^2Q^2}(\vec{I}-\hat{Q}\hat{Q}). \tag{3.9}$$

The solution to (3.8) is, therefore,

$$\vec{\mathcal{E}}(Q)=\frac{-ic}{\omega\epsilon_l}(1+U)\hat{Q}\cdot[\hat{n}\times\vec{B}^s(\vec{Q}_t)]\hat{Q}+\frac{i\omega c(1+U)}{\omega^2\epsilon_t-c^2Q^2}\hat{Q}\times\{\hat{Q}\times[\hat{n}\times\vec{B}^s(\vec{Q}_t)]\}-\frac{ic^2(1-U)}{\omega^2\epsilon_t-c^2Q^2}\vec{Q}\times[\hat{n}\times\vec{\mathcal{E}}^s(\vec{Q}_t)]. \tag{3.10}$$

We have the important result that all solutions to Maxwell's equations incorporating RM symmetry conditions ($U=-1$) are such that

$$\rho_{\text{ind}}(xyz)=\frac{1}{4\pi}\nabla\cdot E\equiv 0,\quad z>0^+. \tag{3.11}$$

[Equation (3.11) refers to effective fields, which are the true fields for $z>0$, but cannot be used to deduce surface charge densities.] We see from (3.11) and Sec. II that in the RM formalism all of the induced charge density (if any) resides in a zero-thickness layer at the surface; this is exactly the situation that occurs in a local theory, and as mentioned earlier, we do not consider this to be an unphysical result. However, if one recognizes that there are no sharp boundaries in nature, the RM formalism predicts that all the induced charge density is confined to the transition region (a few angstroms) between vacuum and dielectric; presumably, the RM formalism is quite sensitive to the nature of the transition region. The exact opposite case holds for the FK formalism; there is no induced surface-charge layer and there is an appreciable bulk charge density extending on the order of hundreds of angstroms into the bulk for typical electron-gas parameters.[12]

One of the more unusual aspects of the FK formalism is that it predicts the resonant excitation of longitudinal plasmons by incident (transverse) e – m waves (p polarized only).[7,8] This was predicted by Sauter, Forstmann, and by Melnyk and Harrison for thin films.[6] Such absorption has been observed in silver films and in potassium films[5] and is predicted *not* to occur by the RM formalism by (3.11). (This is anticipated by the RM expression[9] for the surface impedance Z_p which does not depend on ϵ_l.) This is not a serious objection as the RM formalism was devised for Frenkel excitons.[19] However, Eq. (3.11) predicts that it is not possible to excite longitudinal excitons either; it would be of great interest to see if obliquely incident p-polarized light would be resonantly absorbed for frequencies $\omega\simeq\omega_l(q)$ in thin slabs of materials having an appreciable exciton oscillator strength, e.g., on CuBr or ZnO in which the ratio of oscillator strength to damping is so large that surface excitons can be observed.[27]

The only exception to (3.11) occurs if, for some $\vec{Q}_t$ and ω, there is a Q_z such that $\epsilon_l(\vec{Q}_t, Q_z, \omega)=0$ and the inverse of the matrix in square brackets in (3.8) does not exist. However, it may further be shown that this situation can occur only if $\vec{Q}_t\equiv 0$. The resultant expression for the surface impedance ($U=-1$) does not depend on this "plasmon"; the plasmon does not affect the reflectivity of an external e – m wave and is itself unaffected by it. Similarly, surface excitations do not cou-

ple to this mode either.

From (3.10) and the additional relation (if $U=-1$)

$$\vec{B}(\vec{Q}) = (c/\omega)\vec{Q}\times\vec{\mathcal{E}}(\vec{Q}), \tag{3.12}$$

one can deduce the surface impedance for s- or p-polarized light, in the RM or the FK formalism,[8,9] as has already been reported for the four cases, e.g., $U=+1$ (FK)

$$Z_p = \frac{E_x(xy0^+)}{B_y(xy0^+)} = \frac{i\omega}{c\pi}\int_{-\infty}^{\infty}\left(\frac{Q_x^2}{(\omega^2/c^2)\epsilon_l(Q)} + \frac{Q_z^2}{(\omega^2/c^2)\epsilon_t(Q)-Q^2}\right)\frac{e^{iQ_z0^+}}{Q^2}\,dQ_z, \tag{3.13}$$

in which the variation of the fields in the $x-y$ plane has been assumed to be e^{iQ_xx}; $Q^2=Q_x^2+Q_z^2$, and $e^{iQ_z0^+}=1$ can be assumed because $E_x^{\text{eff}}(z)$ is continuous across $z=0$. One then matches the tangential fields in the bulk to their values in the vacuum

$$E_x(xy0^+)/B_y(xy0^+) = E_x(xy0^-)/B_y(xy0^-), \tag{3.14}$$

thereby determining either the reflectivity[8,9] for a particular frequency ω and angle of incidence $[Q_x=(\omega/c)\sin\theta]$ or the dispersion of the surface excitation $\omega(Q_x)$.[14,15] As indicated in Sec. II, such a procedure is valid only if $\epsilon_t(Q)$ does not diverge as Q^2 for large Q—otherwise B_y is not continuous across the vacuum-solid interface.

C. Sum rule on bulk polaritons

Consider the Fourier transform $F(q_z)$ of any function $f(z)$ which has a single jump discontinuity at $z=0$ but is otherwise "well-behaved." We can integrate by parts twice to obtain the following:

$$F(q_z) = \int_{-\infty}^{\infty} f(z)e^{-iq_zz}\,dz \tag{3.15a}$$

$$= \frac{-i}{q_z}\left[f(0^+)-f(0^-) - \frac{i}{q_z}\left(f'(0^+)-f'(0^-) + \int_{-\infty}^{\infty} f''(z)e^{-iq_zz}dz\right)\right]. \tag{3.15b}$$

Therefore

$$f(0^+)-f(0^-) = i\lim_{q_z\to\infty} q_zF(q_z). \tag{3.16}$$

As an example, consider

$$f(z) = \int_{-\infty}^{\infty}\frac{e^{iq_zz}q_z\,dq_z}{q^2-(\omega^2/c^2)\epsilon_t(q,\omega)}. \tag{3.17}$$

Since f is an odd function, (3.15) gives

$$\int_{-\infty}^{\infty}\frac{q_ze^{iq_z0^+}\,dq_z}{q^2-(\omega^2/c^2)\epsilon_t} = i\pi \tag{3.18}$$

if ϵ_t does not diverge as q^2 or worse, for large q. The limit 0^+ is very important in (3.18) otherwise the integral vanishes.

Equation (3.18) arises naturally in either the RM or FK formalism. Consider p-polarized light in the $x-z$ plane ($\vec{q}_t=q_x\hat{x}$). With $U=+1$ (FK) one can solve for $B_y(z)$ by means of (3.12), (3.10), and (3.2)

$$B_y(x,z) = \frac{B_y^s}{4\pi^3 i}\int_{-\infty}^{\infty}\frac{q_ze^{iq_zz}\,dq_z}{q^2-(\omega^2/c^2)\epsilon_t(q,\omega)}e^{iq_xx}. \tag{3.19}$$

Similarly for $U=-1$ (RM) one can solve for

$$E_x(x,z) = \frac{E_x^s}{4\pi^3 i}\int_{-\infty}^{\infty}\frac{q_ze^{iq_zz}\,dq_z}{q^2-(\omega^2/c^2)\epsilon_t}e^{iq_xx}. \tag{3.20}$$

In order that the theory be consistent, i.e., that

$$E_x(x,0^+) = \frac{1}{(2\pi)^2}E_x^se^{iq_xx} \tag{3.21}$$

as in (3.2b), there is a requirement that ϵ_t satisfy (3.18). This will always be the case if ϵ_t does not diverge as q^2 or worse, which is the same condition that ensures the continuity of tangential $\vec{B}$.

The "sum rule" (3.18) takes on an interesting form if $\epsilon_t(q)$ can be continued onto the entire upper-half complex q plane, as can be done with isolated oscillator models[4] or the Lindhard function.[18] $\epsilon_t(q)$ may have poles, as it does in the former, or branch cuts, as it does in the latter. Let q_i be any complex number which is a zero of the denominator of the integrand of (3.18), i.e.,

$$\lim_{q\to q_i}\left(q^2-\frac{\omega^2}{c^2}\epsilon_t(q,\omega)\right) = \frac{1}{A_i}(q^2-q_i^2). \tag{3.22}$$

[Since $\epsilon_t(q)$ is even in q, q_i and $-q_i$ are zeros of the left-hand side of (3.22) with the same value of A_i.] For each pair $(q_i,-q_i)$, there is one and only one value of $q_z=\pm(q_i^2-q_x^2)^{1/2}$ in the upper half-plane such that[28]

$$q_x^2+q_z^2-(\omega^2/c^2)\epsilon_t((q_x^2+q_z^2)^{1/2},\omega)=0.$$

We are now able to close the contour of (3.18) with an infinitely large semicircle in the upper half-plane. The resultant sum rule on the bulk

polariton modes is

$$\sum_i A_i(\omega) = 1, \tag{3.23a}$$

where A_i is defined by Eq. (3.22):

$$\frac{1}{A_i(\omega)} = 1 - \frac{\omega^2}{c^2}\frac{d\epsilon_t}{d(q^2)}\bigg|_{q_i}. \tag{3.23b}$$

[In (3.23a) integration of (3.18) around branch cuts, if any, is included in the sum.] The sum rule (3.23a) must be satisfied by any bulk dielectric function $\epsilon_t(q)$ to be used in either FK or RM formalism; the only assumptions in deriving (3.23) is that ϵ_t can be continued onto the entire complex-q plane and that ϵ_t does not diverge as q^2 or worse for large q. We interpret (3.23) as saying that the "strengths" of the bulk polariton modes must sum to unity; this must not be taken too seriously because q_i and A_i are, in general, complex, possible pure imaginary.

Are there examples of functions $\epsilon_t(q, \omega)$ which can explicitly be shown to obey (3.23)? Trivially, if ϵ_t is independent of q, then the single mode $q_1 = (\omega/c)\sqrt{\epsilon_t}$ completely exhausts the sum rule with $A_1 \equiv 1$. An important nontrivial example can be found in the commonly used one-oscillator dielectric function.[4]

$$\epsilon_t(q, \omega) = \epsilon_\infty - \frac{K}{\omega^2 - \omega_0^2 - \alpha c^2 q^2}, \tag{3.24}$$

which has two polariton modes $i = 1, 2$

$$q_i^2(\omega) = \frac{(1+\alpha\epsilon_\infty)\omega^2 - \omega_0^2 \pm \{[(1-\alpha\epsilon_\infty)\omega^2 - \omega_0^2]^2 + 4\alpha K\omega^2\}^{1/2}}{2\alpha c^2}, \tag{3.25a}$$

$$\frac{1}{A_i(\omega)} = 1 + \frac{\alpha K\omega^2}{(\omega^2 - \omega_0^2 - \alpha c^2 q_i^2)^2} = 1 + \frac{\alpha(\omega^2 - c^2 q_i^2)^2}{K\omega^2}. \tag{3.25b}$$

It is straightforward to demonstrate

$$A_1(\omega) + A_2(\omega) \equiv 1. \tag{3.26}$$

In this particular case, A_i is real and positive if there is no damping. The transverse Lindhard function, with or without generalization to finite lifetimes,[18] has branch cuts due to the logarithm; we did not directly verify (3.23) in that case.

There are, of course, many such sum rules. Consider, for example,

$$\int_{-\infty}^{\infty} \frac{[\epsilon_t(q) - 1] q_z e^{iq_z 0^+}\, dq_z}{q^2 - (\omega^2/c^2)\epsilon_t(q)} = i\pi(\epsilon_L - 1), \tag{3.27}$$

where ϵ_L is the background "local" contribution to the dielectric function,

$$\epsilon_L = \lim_{q\to\infty} \epsilon_t(q). \tag{3.28}$$

Integration of (3.27) in the upper half-plane gives

$$\sum_i [\epsilon_t(q_i) - 1] A_i = \epsilon_L - 1 \tag{3.29}$$

in the absence of branch cuts.

Sum rules for the longitudinal polaritons are less informative. Consider

$$\int_{-\infty}^{\infty} \frac{q_z e^{iq_z 0^+}\, dq_z}{q^2 \epsilon_l(q)} = \frac{i\pi}{\epsilon_l(\infty)}, \tag{3.30}$$

or

$$\sum_j \frac{B_j}{q_j^2} = \frac{1}{\epsilon_l(\infty)} - \frac{1}{\epsilon_l(0)}, \tag{3.31a}$$

where

$$\lim_{q\to q_j} \epsilon_l(q) = \frac{1}{B_j}(q^2 - q_j^2). \tag{3.31b}$$

A longitudinal function ϵ_l of the form (3.24) can easily be shown to satisfy (3.31); in this case there is only one zero, q_1.

D. Connection with the method of partial waves

In this subsection we specialize for simplicity to the case of normal incidence with RM symmetry conditions ($U = -1$). As mentioned previously, this symmetry condition is equivalent to the additional boundary condition used by Pekar[1] and later by many others (normal incidence only). The solution, the function $E(z)$ for $z > 0$, appears to be quite different for the two cases. In the former we have from (3.20)

$$E_x(z) = \frac{E_x^s}{4\pi^3 i}\int_{-\infty}^{\infty} \frac{q e^{iqz}\, dq}{q^2 - (\omega^2/c^2)\epsilon_t(q)} \quad (U = -1) \tag{3.32}$$

whereas in the latter the field was assumed to be expanded in bulk normal modes

$$E_x(z) = \sum_i E_i e^{iq_i z}, \tag{3.33}$$

where the E_i's are not known *a priori* but are determined from the boundary condition. Under the same condition as applied in Sec. III C, the contour in (3.32) can be closed in the upper half-plane to give

$$E_x(z) = \frac{E_x^s}{4\pi^2}\sum_j A_j e^{iq_j z}, \quad (U = -1) \tag{3.34}$$

where q_j, A_j are defined in (3.23). This is the same form as (3.33) except that the relative sizes of the components are already known from (3.23b). If we define E_I, E_R as the amplitude of the in-

cident and reflected waves, respectively, and $E_i = (E_x^s/4\pi^s)A_i$, then we have from the continuity of tangential $\vec{E}$:

$$E_I + E_R = \sum_i E_i \tag{3.35}$$

which appears in Pekar's and subsequent work based on ABC's. Similarly one can solve for $P_x(z)$ in the RM formalism since $\vec{P}(\vec{Q}) = [\vec{\epsilon}(\vec{Q}) - \vec{I}] \cdot \vec{E}(\vec{Q})$,

$$P_x(z) = \frac{E_x^s}{16\pi^4 i} \int_{-\infty}^{\infty} \frac{[\epsilon_t(q,\omega) - 1] q e^{iqz}}{q^2 - (\omega^2/c^2)\epsilon_t(q,\omega)}\, dq, \qquad (U = -1). \tag{3.36}$$

Integrating in the complex plane as before, and using the definition $n_j = cq_j/\omega = [\epsilon_t(q_j, \omega)]^{1/2}$, we have

$$P_x(z) = \sum_i (n_i^2 - 1) E_i e^{iq_i z}, \tag{3.37}$$

which is also of the form used in the method of partial waves. Pekar's ABC is that $P_x(0^+) = 0$, i.e.,

$$0 = \sum_i (n_i^2 - 1) E_i \tag{3.38}$$

which is simply Eq. (3.29) if we do not allow for a "background" local contribution to ϵ_t; this ABC has been derived for a model Frenkel[19] exciton, again in the neglect of the tensorial character of $\vec{\chi}_B$. If there are only two bulk normal modes (polaritons), Eqs. (3.35) and (3.38) alone are enough to solve the problem exactly. It would appear that if there are more than two bulk modes [as occurs, for example, if the dispersion in the denominator of (3.24) includes quartic terms in q] one would need another boundary condition at the surface; this is not so because the one condition $U = -1 \rightarrow P_x = 0$ is enough to solve the problem exactly with the result that the E_i's are given by (3.29) and (3.23c). This is because the polarizability (2.6) does contain more information than just $P_x(0^+) = 0$.

Similar results hold for the Fuchs-Kliewer formalism,[9] i.e., one retrieves the method of partial waves introduced by Sauter[6] by closing the contour. The situation is complicated by the continuum of electron-hole excitations which manifests itself as a branch cut in the transverse dielectric function for the free-electron gas.

To conclude this section we present the FK and the RM results for the normal-incidence surface impedance $Z_N = E_x(0^+)/B_y(0^+)$ in terms of which the normal-incidence reflectivity may easily be calculated.[8]

$$\frac{1}{Z_N} = \frac{i\omega}{\pi c} \int_{-\infty}^{\infty} \frac{\epsilon_t e^{iq0^+}\, dq}{(\omega^2/c^2)\epsilon_t(q) - q^2} \qquad (U = -1) \tag{3.39a}$$

$$= \sum_i A_i (cq_i/\omega), \tag{3.39b}$$

$$Z_N = \frac{i\omega}{\pi c} \int_{-\infty}^{\infty} \frac{e^{iq0^+}\, dq}{(\omega^2/c^2)\epsilon_t - q^2} \qquad (U = +1) \tag{3.40a}$$

$$= \sum_i A_i (\omega/cq_i). \tag{3.40b}$$

Inasmuch as ω/cq_i is the impedance of each bulk polariton, we see that the boundary condition $P_x(0^+) = 0$ (RM or $U = -1$) implies that the surface impedance consists of the individual mode impedances (weighted by their strengths) taken in parallel and the boundary condition $dP_x(0^+)/dz = 0$ (FK or $U = +1$) is equivalent to a series circuit of bulk polariton impedances.

IV. CALCULATION OF THE SURFACE IMPEDANCE MATRIX FOR A NONHOMOGENEOUS PERFECT CRYSTAL: THE LOCAL FIELD EFFECT

In this section we wish to consider the surface impedance for a two-phase system consisting of a vacuum and a crystal which is nonhomogeneous. As in previous sections we assume that $z = 0$ still represents a sharp boundary between the two and that this plane would be a mirror plane of the bulk crystal. We define the surface-impedance matrix $\vec{Z}(\vec{Q}_t)$ by

$$-\hat{n} \times [\hat{n} \times \vec{\mathcal{E}}^s(\vec{Q}_t)] = \vec{Z}(\vec{Q}_t) \cdot [\hat{n} \times \vec{B}^s(\vec{Q}_t)], \tag{4.1}$$

i.e., $\vec{Z}$ is a 2×2 matrix relating the tangential components of the surface electric field to the tangential components of the surface magnetic field. If the dielectric is assumed homogeneous and isotropic, then $\vec{Z}$ is automatically diagonal in terms of p-polarized and s-polarized components.

$$\vec{Z} = \begin{pmatrix} Z_p & 0 \\ 0 & Z_s \end{pmatrix}. \tag{4.2}$$

Even in a cubic crystal, it is possible that incident p-polarized light can induce a reflected component having s polarization and vice versa, except along symmetry directions.

We begin by considering Eq. (3.5) and we immediately specialize, for sake of definiteness, to the case of FK boundary conditions, $U = +1$

$$(\omega^2/c^2)\vec{\mathcal{D}}(\vec{Q}) + \vec{Q} \times [\vec{Q} \times \vec{\mathcal{E}}(\vec{Q})] = (-2i\omega/c)\hat{n} \times \vec{B}^s(\vec{Q}_t). \tag{4.3}$$

But here $\vec{\mathcal{D}}(\vec{Q})$ is not so simply related to $\vec{\mathcal{E}}(\vec{Q})$ as Eq. (3.6). Instead we have

$$\vec{\mathcal{D}}(\vec{Q}) = \sum_{\vec{K}} \vec{\epsilon}(\vec{Q}, \vec{Q} + \vec{K}) \cdot \vec{\mathcal{E}}(\vec{Q} + \vec{K}), \tag{4.4}$$

where $\vec{K}$ is a bulk reciprocal lattice vector. We will assume that in real space $\vec{\epsilon}(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}')\vec{I}$

$+4\pi\vec{\chi}_B(\vec{r},\vec{r}')$, where $\vec{\chi}_B$ does not have any δ-function (or worse) singularities. Although translational symmetry in any direction with a component normal to the surface is destroyed, there are still bulk translation vectors parallel to the surface giving rise to a surface reciprocal lattice (taken to lie in the $z=0$ plane). (The surface plane can be represented by integral-valued Miller indices.) Because of this symmetry, the $\vec{Q}$ vectors occurring in the Fourier analysis of the fields are always of the form

$$\vec{Q}=\vec{q}_s+\vec{G}_s+\vec{Q}_N, \tag{4.5}$$

where $\vec{Q}_N$ is normal to the surface, $\vec{G}_s$ is any surface reciprocal lattice vector, and $\vec{q}_s$ is in the first surface Brillouin zone. Without loss of generality, one may assume that $\vec{q}_s$ is an eigenvalue of the problem. For example, we may wish to calculate the dispersion of surface plasmons $\omega(\vec{q}_s)$ or the reflectivity of non-normally incident light $|\vec{q}_s|=(\omega/c)\sin\theta$; in either case only one value of $\vec{q}_s$ need be considered. It should be noted that the tangential component of every bulk reciprocal lattice vector is a surface reciprocal lattice vector, and there are no others; corresponding to every surface reciprocal $\vec{G}_s$ there is a $\vec{Q}_N' \| \hat{n}$ such that $\vec{G}_s+\vec{Q}_N'$ is a bulk reciprocal lattice vector. Therefore

$$\vec{Q}=\vec{q}_s+\vec{G}_s+\vec{Q}_N=\vec{q}_s+\vec{q}_N+\vec{K}_j \tag{4.6a}$$

$$\vec{Q}+\vec{K}_i=\vec{q}_s+\vec{G}_s'+\vec{Q}_N'=\vec{q}_s+\vec{q}_N+\vec{K}_l, \tag{4.6b}$$

and Eq. (4.3) becomes

$$\sum_{K_j}\vec{\vec{T}}(\vec{q}_s+\vec{q}_N+\vec{K}_i,\vec{q}_s+\vec{q}_N+\vec{K}_j)\cdot\vec{\mathcal{E}}(\vec{q}_s+\vec{q}_N+\vec{K}_j)=\frac{2i\omega}{c}\hat{n}\times\vec{B}_s[\vec{q}_s+(\vec{K}_i^t)], \tag{4.7}$$

where

$$\vec{\vec{T}}(\vec{q}+\vec{K}_1,\vec{q}+\vec{K}_2)=(\omega^2/c^2)\vec{\vec{\epsilon}}(\vec{q}+\vec{K}_1,\vec{q}+\vec{K}_2)-\delta_{\vec{K}_1\vec{K}_2}[\vec{\vec{I}}(\vec{q}+\vec{K}_1)^2-(\vec{q}+\vec{K}_1)(\vec{q}+\vec{K}_1)]. \tag{4.8}$$

The tangential component of $\vec{K}_i$ is a surface reciprocal lattice vector $\vec{G}_s=(\vec{K}_i^t)=(\vec{\vec{I}}-\hat{n}\hat{n})\vec{K}_i$. Note that the bulk polaritons are defined by setting the left-hand side of (4.7) equal to zero.[16]

Define $\vec{\vec{T}}^{-1}$ by

$$\sum_{K_2}\vec{\vec{T}}^{-1}(\vec{q}+\vec{K}_1,\vec{q}+\vec{K}_2)\cdot\vec{\vec{T}}(q+\vec{K}_2,\vec{q}+\vec{K}_3)=\vec{\vec{I}}\,\delta_{\vec{K}_1,\vec{K}_3}.$$

so that

$$\vec{\mathcal{E}}(\vec{q}_s+\vec{q}_N+\vec{K}_j)=\frac{2i\omega}{c}\sum_{\vec{K}_l}\vec{\vec{T}}^{-1}(\vec{q}_s+\vec{q}_N+\vec{K}_j,\vec{q}_s+\vec{q}_N+\vec{K}_l)\cdot[\hat{n}\times\vec{B}_s(\vec{q}_s+\vec{K}_l^t)], \tag{4.9}$$

and so [from Eq. (3.1d)]

$$\vec{\mathcal{E}}^s(\vec{q}_s+\vec{K}_j^t)=\sum_{\vec{K}_l^t}\vec{\vec{M}}(\vec{q}_s+\vec{K}_j^t,\vec{q}_s+\vec{K}_l^t)\cdot\hat{n}\times\vec{B}^s(\vec{q}_s+\vec{K}_l^t), \tag{4.10a}$$

where

$$\vec{\vec{M}}(\vec{q}_s+\vec{K}_j^t,\vec{q}_s+\vec{K}_l^t)=\frac{2i\omega}{c}\int_{-\infty}^{\infty}dq_N\,\vec{\vec{T}}^{-1}(\vec{q}_s+q_N\hat{n}+\vec{K}_j,\vec{q}_s+q_N\hat{n}+\vec{K}_l). \tag{4.10b}$$

If different bulk reciprocal lattice vectors differ only by a (reciprocal lattice vector) normal to the surface, this complicates the bookkeeping; the vector $\vec{K}_j^t$, the projection of $\vec{K}_j$ onto the surface, may be the same for several bulk reciprocal lattice vectors.

To recapitulate up to this point: $\vec{\vec{T}}$ is defined in terms of the dielectric response matrix for the bulk $\vec{\vec{\epsilon}}(\vec{q}+\vec{K}_i,\vec{q}+\vec{K}_j)$ by Eq. (4.8) and $\vec{\vec{M}}$ is derived from $\vec{\vec{T}}^{-1}$ by integrating over the normal component $q_n\hat{n}$ [Eq. (4.10b)]. Equation (4.10a) relates the Fourier components of the electric field at the surface $\vec{E}(xy0^+)$ [see Eqs. (3.1) and (3.2)] to those of the magnetic field $\vec{B}(xy0^+)$. This is as far as we can go mathematically because the latter are as indeterminate as the former.

We need another condition from physics and it is this: On the vacuum side of the interface there are fields which vary as $e^{i(\vec{q}_s+\vec{K}_l^t)\cdot\vec{r}}$ along the surface; with the possible exception of the component with $\vec{K}_l^t=0$ the variation of these components in a direction normal to the surface is always *outgoing* whether it be evanescent or radiative (the Borrmann effect).[16] That is, in the vacuum the fields vary as

$$\vec{E}_{\rm vac}(\vec{r})=\sum_l\vec{E}_l e^{i\vec{Q}_l\cdot\vec{r}}, \tag{4.11a}$$

where

$$\vec{Q}_l(\vec{K}_l^t\neq 0)=\vec{q}_s+\vec{K}_l^t-[\omega^2/c^2-(\vec{q}_s+\vec{K}_l^t)^2]^{1/2}\hat{n}. \tag{4.11b}$$

The normal components in (4.11b) are chosen so that (4.11a) is a solution to the source-free Maxwell's equations. Only the minus sign in front of the radical is present. In the case of the reflectivity of incident radiation, the two vectors corresponding to $\vec{K}_l^t=0$, namely, $\vec{q}_s+(\omega^2/c^2$

$-q_s^2)^{1/2}\hat{n}$ and $\vec{q}_s-(\omega^2/c^2-q_s^2)^{1/2}\hat{n}$ are both present corresponding to the incident beam and the reflected beam; in the case of a surface "plasmon" (or exciton, etc.) Eq. (4.11b) applies even for $\vec{K}_l^t=0$ and all normal components of the $\vec{Q}_l$'s are imaginary corresponding to exponentially damped (evanescent) waves.

Within the restrictions previously stated, that $\vec{\chi}_B(\vec{r},\vec{r}')$ has no strictly local electric or magnetic contributions, all components of the electric and magnetic fields are continuous across the boundary (FK formalism where the surface is a mirror plane of the crystal). We, therefore, have the important auxiliary condition

$$\vec{K}_l^t\neq 0,\quad (\omega/c)\vec{B}^s(\vec{q}_s+\vec{K}_l^t)=\vec{Q}_l\times\vec{\mathcal{E}}^s(\vec{q}_s+\vec{K}_l^t). \tag{4.12}$$

At this point it is convenient, though unnecessary, to consider the dispersion of surface excitations and the reflectivity of externally incident radiation separately. In the former case, all fields are evanescent—they damp out in a direction normal to the surface. This is true for the term $\vec{K}_l^t=0$ so that (4.12) holds for all components of the fields. Substituting in (4.10)

$$\sum_l\left(\vec{I}\,\delta_{jl}-\frac{c}{\omega}\vec{M}(\vec{q}_s+\vec{K}_j^t,\vec{q}_s+\vec{K}_l^t)\cdot\hat{n}\times\vec{Q}_l\times\right)\times\vec{\mathcal{E}}^s(\vec{q}_s+\vec{K}_l^t)=0. \tag{4.13}$$

The condition for a nontrivial solution to (4.13) is that the determinant of the matrix in square brackets must vanish. For each value of $\vec{q}_s$, this implicitly determines $\omega(\vec{q}_s)$, the dispersion relation for surface excitations. As discussed previously, in the presence of spatial dispersion, one may expect that a surface excitation is never infinitely long lived due to the possibility of mixing with bulk modes.[14,15] Equation (4.13) may, however, prove useful, or even necessary, in the discussion of the surface modes of tightly bound molecular crystals.

In order to calculate the reflectivity of externally incident light, Eq. (4.12) holds for $\vec{K}_l^t\neq 0$ but for $\vec{K}_l^t=0$ there is an incoming and an outgoing (reflected) wave. (The incident light will almost always correspond to $\vec{K}_l^t=0$.) We now have the equation

$$\sum_{\vec{K}_l^t}\left(\vec{I}\,\delta_{j,l}-\frac{c}{\omega}(1-\delta_{0l})\vec{M}(\vec{q}_s+\vec{K}_j^t,\vec{q}_s+\vec{K}_l^t)\cdot\hat{n}\times\vec{Q}_l\times\right)\times\vec{\mathcal{E}}^s(\vec{q}_s+\vec{K}_l^t)=\vec{M}(\vec{q}_s+\vec{K}_j^t,\vec{q}_s)\cdot\hat{n}\times\vec{B}^s(\vec{q}_s). \tag{4.14}$$

By inverting the matrix on the left-hand side, one can solve for any component of the electric field at the surface, $\vec{\mathcal{E}}^s(\vec{q}_s+\vec{K}_l^t)$, in terms of the *one* component of the magnetic field at the surface, $\vec{B}^s(\vec{q}_s)$. In particular, one can solve for the surface-impedance matrix $\vec{Z}(\vec{q}_s,\omega)$ defined by (4.1):

$$-\hat{n}\times[\hat{n}\times\vec{\mathcal{E}}^s(\vec{q}_s,\omega)]=\vec{Z}(\vec{q}_s,\omega)\cdot[\hat{n}\times\vec{B}^s(\vec{q}_s,\omega)]. \tag{4.1′}$$

Having determined the surface-impedance matrix, one is now able to calculate the reflectivity for an arbitrary angle of incidence and arbitrary incident polarization by matching the fields across the boundary. The direction of $\vec{q}_s$ is parallel to the surface in the incident plane and $|q_s|=(\omega/c)\sin\theta$. The incident and reflected wave vectors are

$$\vec{q}_I=\vec{q}_s+(\omega^2/c^2-|q_s|^2)^{1/2}\hat{n}, \tag{4.15a}$$

$$\vec{q}_R=\vec{q}_s-(\omega^2/c^2-|q_s|^2)^{1/2}\hat{n}, \tag{4.15b}$$

so that the total electric field in the vacuum $z<0$ is

$$z<0,\quad \vec{E}=\vec{E}_I e^{i\vec{q}_I\cdot\vec{r}}+\vec{E}_R e^{i\vec{q}_R\cdot\vec{r}} \tag{4.16}$$

and similarly for $\vec{B}$. Substituting in Eq. (4.1) we have

$$-\hat{n}\times\hat{n}\times(\vec{E}_I+\vec{E}_R)=(c/\omega)\vec{Z}\cdot\hat{n}\times(\vec{q}_I\times\vec{E}_I+\vec{q}_R\times\vec{E}_R), \tag{4.17}$$

where $\vec{Z}$ was determined from (4.14) and the definition (4.1). From (4.17) it is straightforward to solve for the tangential components of the reflected amplitude $\vec{E}_R$ in terms of those of the incident $\vec{E}_I$; the normal component is determined by the requirement that $\vec{E}_R\cdot\vec{q}_R=0$. The reason for the unfortunate matrix formulation (4.1), (4.17) is that, e.g., an incident p-polarized wave acquires a component of s polarization on reflection, and vice versa, except along a symmetry axis where group theory requires $\vec{Z}$ to be diagonal.

It is important to note that a crucial step in this derivation was the assumption that all components of the electric field, not just the tangential, are continuous across $z=0$; otherwise (4.12) is not true. The normal component of $\vec{E}$ was shown in Sec. II to be continuous across $z=0$ only if that surface represents a mirror plane of the crystal; indeed it does not make sense to speak of "specular" reflection in any other case.

ACKNOWLEDGMENTS

We are most grateful for extensive and illuminating discussions with R. Fuchs, K. L. Kliewer, and A. J. Mansure.

*Work performed for the U.S. Energy Research and Development Administration under Contract No. W-7405-eng-82.

[1]S. I. Pekar, Zh. Eksp. Teor. Fiz. 33, 1022 (1957) [Sov. Phys.-JETP 6, 785 (1958)]; *ibid.* 34, 1176 (1958) [*ibid.* 34, 813 (1958)]; V. M. Agranovich and V. L. Ginzburg, *Spatial Dispersion in Crystal Optics and the Theory of Excitons* (Wiley-Interscience, New York, 1966).

[2]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Addison-Wesley, Reading, Mass., 1960).

[3]G. E. H. Reuter and E. H. Sondheimer, Proc. R. Soc. Lond. A 195, 336 (1948).

[4]J. J. Hopfield, Phys. Rev. 112, 1555 (1958). J. J. Hopfield and D. G. Thomas, Phys. Rev. 122, 35 (1961); *ibid.* 132, 563 (1963); G. D. Mahan and J. J. Hopfield, Phys. Rev. A135, 428 (1964); J. J. Hopfield, Proceedings of the Eighth International Conference on the Physics of Semiconductors, Kyoto 1966; J. J. Hopfield, J. Phys. Soc. Jpn. Suppl. 21, 77 (1966).

[5]I. Lindau and P. O. Nilsson, Phys. Lett. A 31, 352 (1970); Phys. Scr. 3, 87 (1971); M. Anderegg, B. Feuerbacher, and B. Fitton, Phys. Rev. Lett. 27, 1565 (1971).

[6]F. Sauter, Z. Phys. 203, 488 (1967); F. Forstmann, Z. Phys. 203, 495 (1967); A. R. Melnyk and M. J. Harrison, Phys. Rev. Lett. 21, 85 (1968); Phys. Rev. B 2, 835 (1970).

[7]W. F. Jones, K. L. Kliewer, and R. Fuchs, Phys. Rev. 178, 1201 (1969).

[8]K. L. Kliewer and R. Fuchs, Phys. Rev. 172, 607 (1968). R. Fuchs and K. L. Kliewer, Phys. Rev. 185, 905 (1969).

[9]P. R. Rimbey and G. D. Mahan, Solid State Commun. 15, 35 (1974).

[10]G. S. Agarwal, D. N. Pattanayak, and E. Wolf, Phys. Rev. B 10, 1447 (1974); Phys. Rev. B 11, 1342 (1975), and references therein; J. L. Birman and J. J. Sein, Phys. Rev. B 6, 2482 (1972); A. A. Maradudin and D. L. Mills, Phys. Rev. B 7, 2787 (1973).

[11]Throughout this article we will assume for simplicity that the ambient is vacuum.

[12]A. J. Mansure, Ph.D. thesis, (Iowa State University, 1971) (unpublished).

[13]G. D. Mahan, *Antwerp Advanced Study Institute on Elementary Excitations in Solids, Molecules and Atoms, Part B,* edited by J. T. Devreese, A. B. Kunz, and T. C. Collins (Plenum, New York, 1974).

[14]R. Fuchs and K. L. Kliewer, Phys. Rev. B 3, 2270 (1971).

[15]P. R. Rimbey, Phys. Status Solidi B 68, 617 (1975); Ph.D. thesis (University of Oregon, 1974) (unpublished).

[16]D. L. Johnson, Phys. Rev. B 12, 3428 (1975).

[17]R. M. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B 1, 910 (1970).

[18]J. Lindhard, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. 28, No. 8 (1954); K. L. Kliewer and Ronald Fuchs, Phys. Rev. 181, 552 (1969).

[19]R. Zeyher, J. L. Birman, and W. Brenig, Phys. Rev. B 6, 4613 (1972).

[20]K. L. Kliewer and R. Fuchs, *Advances in Chemical Physics,* edited by I. Prigogine and S. A. Rice (Wiley, New York, 1974), Vol. 27.

[21]C. S. Ting, M. J. Frankel, and J. L. Birman, Solid State Commun. 17, 1285 (1975).

[22]V. V. Hyzhnyakov, A. A. Maradudin, and D. L. Mills, Phys. Rev. B 11, 3149 (1975).

[23]K. L. Kliewer, Phys. Rev. Lett. 33, 900 (1974); J. K. Sass, H. Laucht, and K. L. Kliewer, Phys. Rev. Lett. 35, 1461 (1975).

[24]Ronald Fuchs (unpublished).

[25]For example, the use of the dead layer in Ref. 4.

[26]J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).

[27]I. Hirabayashi, T. Koda, Y. Tokura, J. Murato, and Y. Kaneko J. Phys. Soc. Jpn. 40, 1215 (1976); J. Lagois and B. Fischer, Phys. Rev. Lett. 36, 680 (1976).

[28]Note that because $q_i^2 = (\omega^2/c^2)\,\epsilon_t(q_i, \omega)$ and Im $(\epsilon_t) > 0, q_i$ may be assumed to be in the first quadrant.