Dielectric function of a model insulator

G. A. Rezvani* and Robert J. Friauf
Department of Physics and Astronomy, University of Kansas, Lawrence, Kansas 66045
(Received 24 July 1992)

We have calculated a dielectric response function $\epsilon(q,\omega)$ using the random-phase approximation for a model insulator originally proposed by Fry [Phys. Rev. **179**, 892 (1969)]. We treat narrow and wide band-gap insulators for the purpose of using results in the simulation of secondary-electron emission from insulators. Therefore, it is important to take into account the contribution of the first and second conduction bands. For the real part of the dielectric function we perform a numerical principal value integration over the first and second Brillouin zone. For the imaginary part we perform a numerical integration involving the δ function that results from the conservation of energy. In order to check the validity of our numerical integration methods we perform a Kramers-Kronig transform of the real part and compare it with the directly calculated imaginary part and vice versa. We discuss fitting the model to the static dielectric constant and the f-sum rule. Then we display the wave number and frequency dependence for solid argon, KCl, and model Si.

I. INTRODUCTION

A linear-response function provides a powerful tool in the description of the properties of interacting many-body systems.[1] The dielectric response function $\epsilon(\mathbf{q},\omega)$, which can be used for those properties that are the result of Coulomb interaction between the particles, has many applications in solid-state physics. It has been used in the description of optical properties of solids,[2] energy loss of fast electrons in solids,[3] collective oscillations in solids,[4] and screening of the Coulomb field.[5] Our primary reason for calculating the dielectric response function for model insulators is to use the result in studying the process of secondary-electron emission (SEE) from insulators.

A. Dielectric response function

An exact expression for the dielectric response function exists, but it relies upon knowledge of exact many-electron wave functions. On the other hand, useful approximate expressions have been developed requiring knowledge of only single-electron wave functions, such as the random-phase approximation (RPA) for the longitudinal dielectric function:[6]

$$\epsilon(\mathbf{q},\omega) = 1-\frac{4\pi}{Vq^2}\sum_{\mathbf{k}ll'}\frac{|\langle \mathbf{k},l|e^{-i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}+\mathbf{q},l'\rangle|^2(n_{\mathbf{k}+\mathbf{q},l'}-n_{\mathbf{k},l})}{E_{\mathbf{k}+\mathbf{q},l'}-E_{\mathbf{k},l}-(\omega+i\eta)} . \tag{1}$$

Here l' is the band index for a Bloch state with momentum $\mathbf{k}+\mathbf{q}$, energy $E_{\mathbf{k}+\mathbf{q},l'}$, and occupation number $n_{\mathbf{k}+\mathbf{q},l'}$, and similarly for l. V is the volume of the crystal, and η is an infinitesimal positive real number that specifies how to handle singularities of the denominator. This form differs from the corresponding expression for the uniform electron gas in two respects: the presence of Bloch matrix elements in the numerator, and the occurrence of energy bands in the denominator, as indicated by the band indices l and l'. At zero temperature for insulators the occupation number $n_{\mathbf{k},l}$ is zero for states in the conduction band and one for states in the valence bands.

Next we wish to determine the characteristic energy-loss spectrum for a high-energy electron (10–2000 eV) moving through a solid. Nozieres and Pines[7] showed that the probability per unit time $W(\mathbf{q},\omega)$ that the electron transfers a momentum $\mathbf{q}$ and an energy ω to the electron system (in atomic units, with Rydberg unit of energy) is given by

$$W(\mathbf{q},\omega)=\frac{8\pi}{q^2}\,\mathrm{Im}\left[\frac{-1}{\epsilon(\mathbf{q},\omega)}\right] . \tag{2}$$

From this result we can find the total probability for an inelastic collision by summing over all possible momentum transfers $\mathbf{q}$ and integrating over all possible energy transfers ω,

$$\frac{dP}{dt}=\sum_{\mathbf{q}}\frac{8\pi}{q^2}\int_0^\infty d\omega\,\mathrm{Im}\left[\frac{-1}{\epsilon(\mathbf{q},\omega)}\right] . \tag{3}$$

Evaluation of these probabilities allows us to calculate the inelastic-scattering cross section, which is a vital ingredient for construction of a Monte Carlo simulation of SEE.

B. Secondary-electron emission

The choice of model insulator and the selection of methods of calculation of the dielectric function are influenced by our goal of simulating SEE. In this process a primary electron incident on the surface of the solid penetrates into the solid, interacting with it and consequently suffering deflections and energy losses. In classical language we can say that the electron undergoes multiple elastic and inelastic collisions as it proceeds along a jagged path inside the solid. At each inelastic-scattering

site the energy lost by the incident electron results in excitation of one of the bound electrons of the solid. This internal secondary electron also travels along a jagged path inside the solid, and can excite one or more tertiary electrons, etc. If one of the primary, secondary, tertiary, etc. electrons reaches the surface with enough energy to overcome the potential barrier at the surface, then it can leave the solid and become an external secondary electron.

The yield δ, the average number of secondary electrons emitted per incident primary electron, depends on the energy E of the primary electron and also on the properties of the bombarded material. For most materials the maximum yield δ_m occurs between 0.1 and 2.0 keV, and usually the value is less than 1.0 for metals, between 1 and 2 for semiconductors, and typically 6 to 12 or even larger for insulators. Experimental measurements show that δ_m for an insulator is determined primarily by the band gap and electron affinity of the material.

For the production of secondary electrons the topmost valence-band electrons in the insulator are the most important, because they are less tightly bound and much less localized than electrons in the lower valence bands. Consequently, we treat excitation of electrons in lower valence bands as an ionization process, as if these electrons reside in a free atom. The topmost valence-band electrons are more involved in many-body effects, and so we choose to treat their interaction with the incoming electrons by means of the dielectric theory. The details of the Monte Carlo procedure and the results of the SEE simulations will be discussed in a subsequent paper.

In order to determine the inelastic-scattering cross section we must calculate $\epsilon(\mathbf{q},\omega)$, and for this purpose we need the wave functions and energy bands of the model insulator. One can use the actual wave functions and energy bands for a specific material, and in this way study the SEE behavior of that particular material. But our objective is to study trends in SEE as we change solid-state properties such as band gap and electron affinity. For this purpose we need a model that contains the major features of an insulating solid, yet at the same time is simple enough so that we can incorporate the results in our Monte Carlo simulation program. We believe that the model insulator suggested by Fry[5] is suitable for this purpose, but find that some changes are needed in the method of calculation, as described in the next section.

II. MODEL INSULATOR

We consider typical wide-band-gap insulators such as alkali halides (8–12 eV) and solid rare-gas crystals (12–15 eV). The model suggested by Fry[5] is reasonably appropriate for these classes of insulators. For an alkali halide crystal, for instance, the lowest conduction band is composed of *ns* states of the alkali ions, whereas the topmost valence band is derived from *np* states of the halide ions. Valence-band widths are much smaller than the minimum gap, and the halide *ns* valence band usually lies several electron volts lower than the *np* band. Based on these features Fry proposed the following model: (1) A single conduction band composed of electrons represented by single orthogonalized plane waves (OPW's), occupying a broad parabolic energy band. (We shall introduce the second conduction band in Sec. V.) (2) Several valence bands composed of tightly bound electrons represented by linear combinations of np_x, np_y, np_z, and *ns* atomic orbitals, occupying narrow energy bands. (3) The *np* and *ns* valence bands are separated from each other and from the conduction band by finite energy gaps.

A. Band structure and wave functions

A sketch of the energy bands of this model is shown in Fig. 1. (In our treatment of the dielectric function we do not include the *ns* valence band because these electrons are more tightly bound. The contribution of the *ns* valence electrons to SEE will be included by means of a formula for ionization from core levels as discussed in a subsequent paper.) We choose isotropic energy bands with a parabolic form:

$$E_c(\mathbf{k}) = E_c|\mathbf{k}|^2 , \tag{4}$$

$$E_{lm}(\mathbf{k}) = -E^0_{lm} - E_{lm}|\mathbf{k}|^2 . \tag{5}$$

E_c is the reciprocal effective mass in the conduction band, and the E_{lm} are reciprocal effective masses in the valence bands. Since the topmost valence band in our model is derived from *np* states, we shall set $l=1$ in the following equations, and replace E^0_{lm} with E_g. We also assume that $E_{10}=E_{11}=E_{1-1}$, which is a reasonable assumption for our purposes.

For the valence-band electrons in the tight-binding scheme we start with a fairly well localized atomic orbital,

$$U_{lm}(\mathbf{r}) = U_l(r)Y_{lm}(\theta,\phi) . \tag{6}$$

For the radial part of the *np* orbital we use a normalized Slater orbital for $l=1$,

$$U_1(r) = \left[\frac{(2b)^5}{24}\right]^{1/2} re^{-br} . \tag{7}$$

The parameter b determines the degree of localization of the orbital (see Sec. VI for adjustment of b). Then we ob-

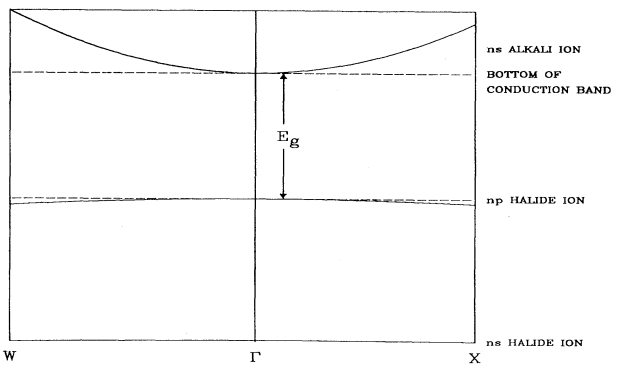

FIG. 1. The band structure of a model insulator. The energy is shown along the Γ,X and Γ,W directions. E_g is the band gap, and $E=0$ at the bottom of the (first) conduction band.

tain valence-band orbitals by forming a Bloch sum

$$\psi_{lm}(\mathbf{k})=\frac{1}{\sqrt{N}}\sum_{\nu} e^{i\mathbf{k}\cdot\mathbf{R}_\nu}U_{lm}(\mathbf{r}-\mathbf{R}_\nu)\ , \tag{8}$$

where the sum is over all N lattice vectors $\mathbf{R}_\nu$ at which atomic wave functions U_{lm} are located; m is the azimuthal quantum number. If overlap integrals between atoms are neglected, the tight-binding Bloch functions are normalized since the atomic wave functions are already normalized.

For the conduction-band wave functions we start with plane waves. Then we form a single OPW by orthogonalizing to all the lower valence states $\psi_{lm}(\mathbf{k})$,

$$\psi_c(\mathbf{k})=\frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}}-\sum_{l,m}\mu_{l,m}(\mathbf{k})\psi_{lm}(\mathbf{k})\ . \tag{9}$$

V is the volume of the crystal containing N atoms, and the $\mu_{lm}(\mathbf{k})$ are the usual orthogonalization coefficients,

$$\mu_{lm}(\mathbf{k})=(N/V)^{1/2}\int e^{i\mathbf{k}\cdot\mathbf{r}}U^*_{lm}(\mathbf{r})dV\ . \tag{10}$$

The conduction-band wave function is normalized by multiplying by a constant $A(\mathbf{k})$,

$$A(\mathbf{k})=\left[1-\sum_{m=-1}^{1}\mu^*_{lm}(\mathbf{k})\mu_{lm}(\mathbf{k})\right]^{-1/2}=\left[1-\frac{8\pi}{\Omega}\frac{(2b)^7|\mathbf{k}|^2}{(b^2+|\mathbf{k}|^2)^6}\right]^{-1/2}, \tag{11}$$

where Ω is the volume of the Wigner-Seitz cell, namely, a (primitive) unit cell of the crystal. Note that the value of the OPW normalization constant $A(\mathbf{k})$ depends on the parameter b.

For the purpose of numerical calculation it is convenient to separate the real and imaginary parts of $\epsilon(\mathbf{q},\omega)=\epsilon_1(\mathbf{q},\omega)+i\epsilon_2(\mathbf{q},\omega)$. We also convert the sum over $\mathbf{k}$ to an integral over the first Brillouin zone,

$$\epsilon_1(\mathbf{q},\omega)=1+\frac{4\pi}{q^2}\frac{2}{(2\pi)^3}\sum_{m=-1}^{1}\left[P\int_{\mathrm{BZ}}\frac{|\langle\psi_{lm}(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_c(\mathbf{k}+\mathbf{q})\rangle|^2d^3k}{E_c|\mathbf{k}+\mathbf{q}|^2+E_{lm}|\mathbf{k}|^2+E_g-\omega}+\int_{\mathrm{BZ}}\frac{|\langle\psi_c(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_{lm}(\mathbf{k}+\mathbf{q})\rangle|^2d^3k}{E_c|\mathbf{k}|^2+E_{lm}|\mathbf{k}+\mathbf{q}|^2+E_g+\omega}\right], \tag{12}$$

$$\epsilon_2(\mathbf{q},\omega)=\frac{4\pi}{q^2}\frac{2}{(2\pi)^3}\sum_{m=-1}^{1}\left[\int_{\mathrm{BZ}}|\langle\psi_{lm}(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_c(\mathbf{k}+\mathbf{q})\rangle|^2\delta[E_c|\mathbf{k}+\mathbf{q}|^2+E_{lm}|\mathbf{k}|^2+E_g-\omega]d^3k\right]. \tag{13}$$

In the expression for $\epsilon_1(\mathbf{q},\omega)$ P indicates principal value integration.

B. Method of calculation

We choose to perform the integrations over the actual Brillouin zone (BZ) for a fcc structure, which is a typical close-packed structure for insulators. The reduced zone scheme is used, in which all wave vectors are referred to equivalent vectors in the first BZ. When $\mathbf{k}+\mathbf{q}$ falls outside the first BZ, it must be reduced by a reciprocal-lattice vector, and in such cases we follow Fry[5] and speak of umklapp processes.

Since Fry wanted to use the dielectric function to study screening of extra charges in crystals, it was sufficient to include only one conduction band in his calculations.[5] He approximated the first BZ with a spherical zone, which simplifies the integrations extensively, but makes treatment of umklapp processes vague.[8] We eventually realized that higher-energy excitations are especially important for SEE, so we need to include at least the second conduction band in our calculations. The method that Fry used for treating umklapp processes is not applicable at all for higher conduction bands, but we were able to extend our direct summation method to include the second conduction band for both normal and umklapp processes (Sec. V).

Although our method of calculation of $\epsilon(\mathbf{q},\omega)$ can be used for all directions of $\mathbf{q}$ without great difficulty, the incorporation of the results in the simulation study of SEE becomes very involved. Therefore we assume isotropic behavior for $\epsilon(\mathbf{q},\omega)$ and calculate the dielectric function for only one direction of $\mathbf{q}$. Fry[5] made reference to Nara's[9] calculations for silicon to suggest that $\epsilon(\mathbf{q},\omega)$ with q along the z axis of the crystal is "probably not a bad average over directions, if an average is desired for a screening calculation."[5]

For the band structure of an actual crystal, the effective mass of an electron in the conduction band may be a function of energy. We use a parabolic band for our model insulator, however, and so we have an energy-independent effective mass in the conduction band. An internal secondary electron usually has 3–20 eV energy with respect to the bottom of the conduction band, and a primary electron usually has much higher energy. In order to unify the treatment of electrons with different energies in the conduction band, we use the free-electron mass for the effective mass of all electrons in the conduction band. Since our ultimate purpose is to study SEE for different model insulators, the choice of free-electron mass in the conduction band also unifies the treatment of insulators with different values of band gap and electron affinity.

III. EVALUATION OF THE REAL PART OF THE DIELECTRIC FUNCTION

The integral in Eq. (12) for $\epsilon_1(\mathbf{q},\omega)$ ranges over all vectors $\mathbf{k}$ in the first BZ. When the momentum transfer in an inelastic collision is nonzero, i.e., $|\mathbf{q}|>0$, there is a finite contribution to the dielectric function from umklapp processes. Thus when $\mathbf{k}+\mathbf{q}$ falls outside the first BZ, it must be replaced by $\mathbf{k}+\mathbf{q}-\mathbf{G}$, where there is one and only one reciprocal-lattice vector $\mathbf{G}$ for each point

$\mathbf{k}+\mathbf{q}$ outside the zone. Hence we divide the first BZ into two regions, normal BZ^N and umklapp BZ^U, and write each integral over the first BZ as a sum of two parts,

$$\epsilon_1(q,\omega)=1+\epsilon_1^N(q,\omega)+\epsilon_1^U(q,\omega)\ . \tag{14}$$

First, we discuss evaluation of the normal contribution, and then make a few remarks about the umklapp calculation at the end of this section.

A. Normal contribution to the real part

For each integral in Eq. (12) we set $l=1$ and write the sum over m explicitly; so we obtain one term for $lm=10$ and two equal contributions for $lm=11$ and $1-1$.

$$\epsilon_1^N(\mathbf{q},\omega)=\frac{4\pi}{q^2}\frac{2}{(2\pi)^3}\left[P\int_{BZ^N}\frac{A(\mathbf{k}+\mathbf{q})|M^N_{10,c}|^2}{E_c|\mathbf{k}+\mathbf{q}|^2+E_{10}|\mathbf{k}|^2+E_g-\omega}d\mathbf{k}+\int_{BZ^N}\frac{A(\mathbf{k})|M^N_{c,10}|^2}{E_ck^2+E_{10}|\mathbf{k}+\mathbf{q}|^2+E_g+\omega}d\mathbf{k}\right.$$
$$\left.+2P\int_{BZ^N}\frac{A(\mathbf{k}+\mathbf{q})|M^N_{11,c}|^2}{E_c|\mathbf{k}+\mathbf{q}|^2+E_{11}|\mathbf{k}|^2+E_g-\omega}d\mathbf{k}+2\int_{BZ^N}\frac{A(\mathbf{k})|M^N_{c,11}|^2}{E_ck^2+E_{11}|\mathbf{k}+\mathbf{q}|^2+E_g+\omega}d\mathbf{k}\right]\ . \tag{15}$$

The integrations are over the normal region, and the OPW normalization factor $A(\mathbf{k})$ is shown explicitly (Sec. II A). The matrix elements $M^N_{10,c}$, etc., are evaluated in Appendix A.

Let us look at the nature of the integrand in these integrals. The numerators are smooth well-behaved functions of $\mathbf{k}$ throughout the zone. The denominators of the second and fourth integrands are always positive, and therefore nonsingular. But the integrands of the first and third integrals can become singular inside the first BZ when ω is larger than a threshold value.

In order to investigate the behavior of the singularity in the first integral in Eq. (15), we set the denominator equal to zero, and after some rearrangement obtain

$$k_x^2+k_y^2+\left[k_z+\frac{E_cq}{E_c+E_{10}}\right]^2=\frac{(\omega-E_g)(E_c+E_{10})-E_cE_{10}q^2}{(E_c+E_{10})^2}\ . \tag{16}$$

When the right-hand side is positive this equation defines a sphere of singularity in reciprocal space. The center of the sphere C^N is located at $(0,0,-E_cq/(E_c+E_{10}))$ and the radius of the sphere is given by

$$R=\left[\frac{(\omega-E_g)(E_c+E_{10})-E_cE_{10}q^2}{(E_c+E_{10})^2}\right]^{1/2}\ . \tag{17}$$

The sphere of singularity appears only when ω is larger than a threshold value:

$$\omega\geq\omega_{\rm th}=E_g+\frac{E_cE_{10}}{E_c+E_{10}}q^2\ . \tag{18}$$

In Fig. 2 the first BZ is shown with a typical sphere of singularity inside it. The radius of the sphere of singularity increases as ω becomes larger, and eventually the sphere intersects the walls of the Brillouin zone. As ω becomes still larger no segment of the sphere of singularity is left inside the first zone.

When the singularity first appears at $\omega=\omega_{\rm th}$ there is only one point of singularity. In this case the denominator can be written as $|\mathbf{k}-\mathbf{k}_{C^N}|^2$. Now if we transform the integrand to a coordinate system centered at C^N, then the denominator goes to zero as k^2, and since the volume element also varies as k^2, the integral remains finite. When $\omega>\omega_{\rm th}$ we have a sphere of singularity with finite radius. Then the integrand tends to negative infinity as we approach the sphere from inside, and to positive infinity as we approach from outside. Since we expect the numerator to be a slowly varying function of $\mathbf{k}$ near the singularity, it is conceivable that the positive and negative singularities cancel each other and leave a finite value for the integral. From this visualization of the process of principal value integration, we conclude that it should be possible to devise a numerical integration scheme. But we must handle the integrand very carefully in the vicinity of the singularity so that the cancellation can take place properly.

It would be desirable to treat the integral analytically, but there are several severe difficulties. One is that the OPW normalization factor, the matrix element, and the denominator are functions of both the magnitude and polar angle of $\mathbf{k}$. If we transform the integral to a coordi-

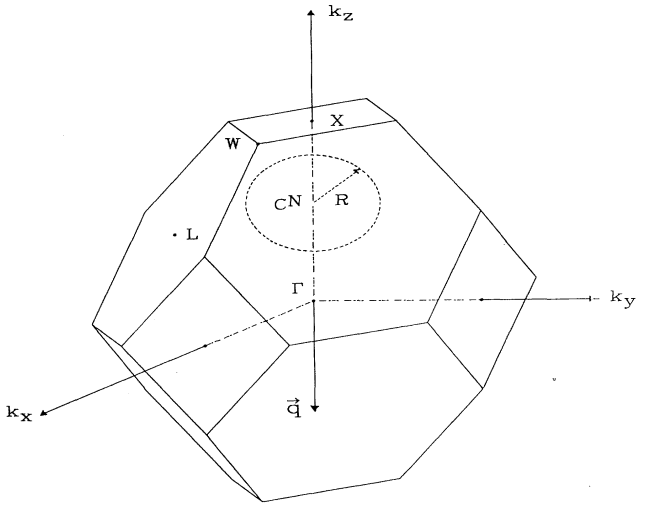

FIG. 2. The first Brillouin zone (BZ) for a face-centered-cubic lattice with the sphere of singularity shown.

nate system that is centered at the origin of the sphere of singularity, then the denominator becomes a function of only one variable k. It might then be possible to evaluate the principal-value integral analytically at each value of the polar angle. But as the sphere of singularity expands, it eventually intersects the walls of the BZ. The analytical treatment would then require expressing the boundaries of the BZ in spherical polar coordinates with an off-set origin (Fig. 2). This would make the formalism very difficult, and is another reason for not pursuing the idea of calculating the principal-value integrals analytically. Therefore we searched for another method of handling the singularity, as discussed in the following subsections.

B. Procedure for numerical integration

In order to integrate over the actual BZ in reciprocal space, we choose a Cartesian coordinate system with origin at the center of the BZ, and divide the first BZ into a number of small integration cubes with sides s. With an even number of divisions between the Γ point and X point, each cube that is truncated by a (111) face of the zone contributes $\frac{5}{6}s^3$ to the volume of the zone if the center of the cube is inside the zone and $\frac{1}{6}s^3$ if the center is outside. No small cube is truncated by a (100) face of the zone. In our calculation the number of divisions along the Γ,X direction is 10. Therefore the volume of the conventional body-centered-cubic cell in reciprocal space is $8000s^3$, and the volume of the first zone is $4000s^3$.

Using the symmetry of the BZ we can limit the region of integration to $k_x \geq 0$, $k_y \leq k_x$, and multiply the result by 8 for the entire volume of the first zone. Since q can be different from zero in these calculations, there is no symmetry about the (k_x,k_y) plane to reduce the region of integration any further. Thus by using symmetry we need to perform the summation effectively over 500 cubes. In a calculation of the real part of the dielectric function for $\omega=0$ Lipari[10] converted the integral to a sum over a discrete set of points in the first zone. He also used 4000 points in the first zone to do the summation.

When the sphere of singularity exists we divide the integration cubes in the first zone into two groups. Nonsingular cubes are those that are far from the sphere of singularity ("far" is defined below). The other group of singular cubes includes those that are either intersected by the sphere of singularity or are in the neighborhood of the sphere of singularity.

C. Nonsingular cubes

In the case of the second and fourth integrals in Eq. (14) all cubes in the zone are nonsingular. In the case of the first and third integrals when $\omega<\omega_{\text{th}}$ all cubes are again nonsingular. When $\omega>\omega_{\text{th}}$ we have a sphere of singularity with the center at C^N and radius R. But there are still some integration cubes that can be treated as nonsingular because they are far enough from the singularity so that the denominator is a slowly varying function of $\mathbf{k}$.

For the ith cube with the center at $\mathbf{k}_i$ we calculate the distance from $\mathbf{k}_i$ to the sphere of singularity along the radial line as $d=|\mathbf{k}_i-\mathbf{k}_{C^N}|-R$. We define nonsingular cubes as those for which

$$|d|>r_c=3s \ .$$

In order to choose a value for the critical distance r_c we note that an integral such as

$$I=P\int_R \frac{dk_x dk_y dk_z}{k_x^2+k_y^2+k_z^2-k_s^2}$$

can be evaluated analytically over a sphere of radius $R>k_s$. For several values of k_s we compare the analytical results to those of the numerical method described below, and find that $r_c=3s$ is a good choice both for accuracy of results and saving computer time. The integral over a nonsingular cube is approximated by the value of the integrand at the center of the cube multiplied by the volume of the cube.

D. Singular cubes

For a singular cube, and even for nearby cubes that are not cut by the sphere of singularity, careful treatment is essential for the principal-value integral. Therefore we devised a scheme that is not as difficult as the analytical approach, yet at the same time has enough features in it to allow for suitable cancellation to take place in the principal-value integration. Let us write the first integral in Eq. (15) as

$$I=\sum_{\text{BZ}^N} P\int_{i\text{th cube}} \frac{Q(k_x,k_y,k_z)}{g(k_x,k_y,k_z)}dk_x dk_y dk_z$$

$$\simeq \sum_{\text{BZ}^N} Q(\mathbf{k}_i)P\int_{i\text{th cube}} \frac{dk_x dk_y dk_z}{g(k_x,k_y,k_z)} \ . \tag{19}$$

For a sufficiently small integration cube $Q(\mathbf{k}_i)$ should be a slowly varying function, and we approximate its value throughout the cube with its value at the center of the cube at $\mathbf{k}_i$. In this way the behavior of the singular part of the integrand is separated from the smoothly varying nonsingular part.

When $g(k_x,k_y,k_x)=0$ the integrand becomes singular. In our model we use parabolic bands, and this causes the locus of singularity to become a sphere in reciprocal space. But in general the locus may have other more complicated shapes depending on the band structure used in the model. In any case if the size of the cube is small, the portion of the locus of singularity that is inside the cube can be approximated with a plane. We choose the zeroth- and first-order terms of the Taylor-series expansion of $g(k_x,k_y,k_z)$ around the center of the cube at $\mathbf{k}_i$ for this approximation. Thus the denominator has the form of a general linear function, and we can write

$$I_S=P\int_{i\text{th cube}} \frac{dk_x dk_y dk_z}{Ak_x+Bk_y+Ck_z-D} \ . \tag{20}$$

For cubes that are in the vicinity of the surface of singularity but are not intersected by the surface, we can still use the same method, namely, a Taylor-series expansion

about the center of that cube.

From the expression in the denominator of the first integral in Eq. (15) we find the expansion coefficients A,B,C, and D:

$$\begin{aligned} A &= 2(E_c+E_{10})k_{ix} \ , \\ B &= 2(E_c+E_{10})k_{iy} \ , \\ C &= 2(E_c+E_{10})k_{iz}+2E_cq \ , \\ D &= (E_c+E_{10})(k_{ix}^2+k_{iy}^2+k_{iz}^2)+2E_cqk_{iz}+E_g-\omega \ . \end{aligned} \tag{21}$$

The results of this method of evaluating the principal-value integral are given in Appendix B.

E. Umklapp contribution to the real part

The form of the integrals for $\epsilon_1^U(\mathbf{q},\omega)$ is basically the same as for $\epsilon_1^N(\mathbf{q},\omega)$, except that now $\mathbf{k}+\mathbf{q}$ is replaced by $\mathbf{k}+\mathbf{q}-\mathbf{G}$ in all places. The energy bands chosen for this model are not periodic in reciprocal space. Therefore if an energy has to be calculated at $\mathbf{k}+\mathbf{q}$ that falls outside the first BZ, we need to calculate the energy at the corresponding point inside the first BZ.

Now all integrations are over the umklapp region of the first BZ. Actually, we have a number of different subregions; each is defined by the points in the umklapp region that require the same reciprocal-lattice vector $\mathbf{G}_s$ to translate $\mathbf{k}+\mathbf{q}$ back into the first BZ. Hence for each subregion we have a different sphere of singularity with center C_s^U at $(E_c'G_{sx},E_c'G_{sy},E_c'(G_{sz}-q))$, with $E_c'=E_c/(E_c+E_{10})$, and with radius given by

$$R_s=\left[\frac{(\omega-E_g)(E_c+E_{10})-E_cE_{10}(\mathbf{q}-\mathbf{G}_s)^2}{(E_c+E_{10})^2}\right]^{1/2} . \tag{22}$$

In this case also we can define a threshold value of ω for the appearance of any singularity.

We assign each small cube to a subregion according to $\mathbf{k}_i$ at the center of the cube. Then we calculate the sphere of singularity for the corresponding $\mathbf{G}_s$, determine whether the cube is singular or nonsingular, and evaluate the contribution to the integrals by the preceding methods. We treat the entire volume of the ith cube in the subregion determined by $\mathbf{k}_i$ at the center of the cube. This procedure does not allow for the possibility that the boundary between the subregions may cut through some cubes, and may introduce some numerical errors. However, the overall consistency of results is the final test of the numerical procedures. The general behavior of $\epsilon_1(\mathbf{q},\omega)$ vs ω is smooth, and the general check with the Kramers-Kronig relations works well (Sec. VI).

IV. EVALUATION OF THE IMAGINARY PART OF THE DIELECTRIC FUNCTION

To complete the evaluation of the dielectric function we must also calculate the imaginary part in Eq. (13). As with the calculation of the real part, we distinguish between normal and umklapp processes, as in Eq. (14).

A. Normal contribution to the imaginary part

For $\epsilon_2^N(\mathbf{q},\omega)$ we again write out the sum over m in Eq. (13) and obtain

$$\epsilon_2^N(\mathbf{q},\omega)=\frac{4\pi^2}{q^2}\frac{2}{(2\pi)^3}\left[\int_{\mathrm{BZ}^N}A(\mathbf{k}+\mathbf{q})|M_{10,c}^N|^2\delta[E_c|\mathbf{k}+\mathbf{q}|^2+E_{10}|\mathbf{k}|^2+E_g-\omega]d\mathbf{k} + 2\int_{\mathrm{BZ}^N}A(\mathbf{k}+\mathbf{q})|M_{11,c}^N|^2\delta[E_c|\mathbf{k}+\mathbf{q}|^2+E_{11}|\mathbf{k}|^2+E_g-\omega]d\mathbf{k}\right] . \tag{23}$$

The arguments of the δ functions in these integrals are the same as the denominators of the first and third integrands in the expression for $\epsilon_1^N(\mathbf{q},\omega)$ in Eq. (15). At those values of ω for which the denominators vanish for the real part, the system can absorb energy, and the imaginary part of the dielectric function becomes finite. This can happen only when $\omega\geq\omega_{\mathrm{th}}$. Then the sphere of singularity is just the locus of states in the valence band that can be excited into the conduction band for a given ω and $\mathbf{q}$.

The difficulty with the integrals in Eq. (23) is that the argument of the δ function depends on both the magnitude and polar angle of $\mathbf{k}$. Therefore we transform to a new coordinate system $\mathbf{k}'=\mathbf{k}-\mathbf{k}_{C^N}$, where $\mathbf{k}_{C^N}$ is the vector to the center of the sphere of singularity discussed above. In the new coordinate system the first integral contains a δ function

$$\delta\left[(E_c+E_{10})k'^2+\frac{E_cE_{10}}{E_c+E_{10}}q^2+E_g-\omega\right]=\delta(ak'^2-k_0'^2)$$

in which the argument depends only on the magnitude of $\mathbf{k}'$. Then by moving across the sphere of singularity in the normal direction we can account properly for the "width" of the δ function. The first integral becomes

$$I=\left[\frac{8\pi^3}{\Omega}\right]^2\frac{8\pi N}{V}(2b)^7\int_{\mathrm{BZ}^N}\left[1-\frac{8\pi}{\Omega}\frac{(2b)^7|\mathbf{k}_{010}'+\mathbf{q}|^2}{[b^2+|\mathbf{k}_{010}'+\mathbf{q}|^2]^6}\right]^{-1} \times\left[\frac{\cos\theta_{\mathbf{k}_{010}'}|\mathbf{k}_{010}'|}{(b^2+|\mathbf{k}_{010}'|^2)^3}-\frac{\cos\theta_{\mathbf{k}_{010}'+\mathbf{q}}|\mathbf{k}_{010}'+\mathbf{q}|}{[b^2+|\mathbf{k}_{010}'+\mathbf{q}|^2]^3}\cdot a_2\right]\frac{1}{2(E_c+E_{10})k_{010}'}k_{010}'^2\sin\theta'd\theta'd\phi' , \tag{24}$$

where $\mathbf{k}'_{010}=\mathbf{k}'_0(E_c+E_{10})^{-1/2}$ varies only over the surface of the sphere of singularity.

B. Evaluation of the surface integral

For the purpose of evaluating the surface integral numerically we subdivide the volume of the BZ into small cubes as discussed before. Then the surface integral in Eq. (24) can be divided into integrals over the "patches" of the sphere of singularity which are inside cubes intersected by the surface. For small cubes truncated by a face of the BZ we calculate the surface integral for the entire cube, and then multiply by $\frac{5}{6}$ (if the center of cube is inside the BZ) or $\frac{1}{6}$ (if the center is outside the BZ). (See Sec. III B.)

In our calculation of ϵ_2 we approximate the value of the integral over each surface of the patch by the value of the integrand at the point of "vortex" on the patch times the area of the patch. First, we find the coordinates of the intersections of the sphere of singularity and the edges of the cube. There can be 3, 4, 5, or 6 intersection points in a typical case when the center of the sphere is well removed from the integration cube, and along with the vortex these define the corners of a three-dimensional polygon. (Special cases of 1, 2, 8, 12, or 24 points of intersection happen very seldom.) The point of "vortex" is the intersection point on the sphere of the radius that passes through the center of mass of the corners of the polygon. Then we find the area of the patch as the area of the triangles in the "tent shape" formed by the point of vortex and the corners of the polygon. To test this algorithm for approximating the area of a patch, we calculated the surface area of spheres with known radii by enclosing them in a mesh of small cubes with the same resolution as the one used in our calculations. In all cases we found errors less than 0.5%.

C. Umklapp contribution to the imaginary part

The form of the integrals involved in calculation of $\epsilon_2^U(q,\omega)$ is basically the same as in Eq. (23), except that now a reciprocal-lattice vector is subtracted from $\mathbf{k}+\mathbf{q}$ wherever it occurs in the integrals. Therefore the normalization factor, matrix elements, and energy bands are affected. Just as in the normal region we handle the δ function by moving the origin of the coordinate system to the center of the appropriate sphere of singularity. Then we can integrate the δ function analytically along the radius k', and proceed to numerical evaluation of the surface integral.

The main feature of the calculation in this case is that it is done cube by cube in the umklapp region. At each cube the vector $\mathbf{k}_i+\mathbf{q}$ is found, and for that vector the corresponding reciprocal-lattice vector $\mathbf{G}_s$ is used to translate $\mathbf{k}_i+\mathbf{q}$ back into the first BZ. This vector $\mathbf{G}_s$ defines the sphere of singularity for this particular cube. If that cube is intersected by this sphere of singularity, then the area of the patch is evaluated and multiplied by the value of the integrand at the vortex. If that cube is not intersected, then the contribution of that cube is zero, and the next cube is tried. In this way, since each cube is treated independently, we automatically take into account the different subregions of the umklapp region, as in the calculation of the real part.

V. INCLUSION OF THE SECOND CONDUCTION BAND

In order to allow for some higher-energy excitations which are vitally important for SEE, we need to include at least the second conduction band in our calculations. The appropriate expressions for the real and imaginary parts of the dielectric function are

$$\epsilon_1(\mathbf{q},\omega)=1+\frac{4\pi}{q^2}\frac{2}{(2\pi)^3}\sum_{lmi}\left[P\int_{\mathrm{BZ}}\frac{|\langle\psi_{lm}(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_{ci}(\mathbf{k}+\mathbf{q})\rangle|^2d^3k}{E_{ci}(\mathbf{k}+\mathbf{q})+E_{lm}(\mathbf{k})+E_g-\omega}+\int_{\mathrm{BZ}}\frac{|\langle\psi_{ci}(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_{lm}(\mathbf{k}+\mathbf{q})\rangle|^2d^3k}{E_{ci}(\mathbf{k})+E_{lm}(\mathbf{k}+\mathbf{q})+E_g+\omega}\right], \tag{25}$$

$$\epsilon_2(\mathbf{q},\omega)=\frac{4\pi}{q^2}\frac{2}{(2\pi)^3}\sum_{lmi}\left[\int_{\mathrm{BZ}}|\langle\psi_{lm}(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_{ci}(\mathbf{k}+\mathbf{q})\rangle|^2d^3k\,\delta[E_{ci}(\mathbf{k}+\mathbf{q})-E_{lm}(\mathbf{k})-\omega]\right]. \tag{26}$$

The summation over l and m includes the various valence bands, and the summation over i is introduced in order to include the first, second, and higher conduction bands. Since the reduced zone scheme is used in both expressions, the integrals are over the first BZ and $\mathbf{k}$ is inside the first BZ.

A. Reduced zone scheme

In order to implement the reduced zone scheme, we note the following properties of the Brillouin zones. (1) No reciprocal-lattice vector can be found to connect two different points of the same zone. (2) To every point in the first BZ there corresponds one and only one point in the second zone, and vice versa. (3) The corresponding points can be translated to each other by a reciprocal-lattice vector $\mathbf{G}_2$. (4) For a fcc crystal lattice in three dimensions there are 14 reciprocal-lattice vectors needed in order to translate all of the points in the first zone to the second zone, and vice versa. Therefore each time we refer to $\mathbf{G}_2$ we mean one of the members of this set of 14 vectors.

Because of these properties we can index each point in the second zone to a corresponding point in the first zone in a unique fashion. For each point $\mathbf{k}$ inside the first BZ we (i) find the reciprocal-lattice vector $\mathbf{G}_2$ that translates $\mathbf{k}$ to the corresponding point in the second zone; (ii) evaluate the conduction-band wave function and energy at $\mathbf{k}+\mathbf{G}_2$. In the two-dimensional scheme in Fig. 3, for example, the points in region A of the second zone are indexed to the points in segment A' of the first zone by the reciprocal-lattice vector $\mathbf{G}_2$.

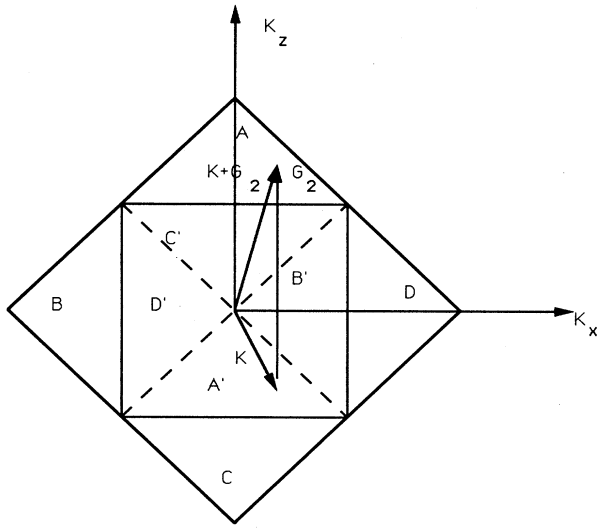

FIG. 3. The first and second BZ for a two-dimensional square lattice. We show a typical reciprocal-lattice vector $\mathbf{G}_2$, a point $\mathbf{k}$ in the first zone, and the corresponding point $\mathbf{k}+\mathbf{G}_2$ in the second zone.

Since we are going to include only the first and second conduction bands in our calculations, we set $i=2$ in the rest of this section. In the reduced zone scheme the conduction-band energy is $E_{c2}(\mathbf{k})=E_c|\mathbf{k}+\mathbf{G}_2|^2$ and the wave function is given by

$$\psi_{c2}(\mathbf{k})=\frac{e^{i(\mathbf{k}+\mathbf{G}_2)\cdot\mathbf{r}}}{\sqrt{V}}-\sum_{l,m}\mu_{l,m}(\mathbf{k}+\mathbf{G}_2)\psi_{lm}(\mathbf{k})\ . \tag{27}$$

The form of the matrix elements is similar to that for the first conduction-band calculation, except that the reciprocal-lattice vector $\mathbf{G}_2$ appears in the argument of the μ's. To clarify this point we show here the form of $M^N_{lm,c2}$ and $M^U_{lm,c2}$:

$$M^N_{lm,c2}=\left[\frac{8\pi^3}{\Omega}\right]\left[\mu_{lm}(\mathbf{k}+\mathbf{G}_2)-\sum_{l'm'}\mu_{l'm'}(\mathbf{k}+\mathbf{q}+\mathbf{G}_2)I_{lm,l'm'}(\mathbf{q})\right], \tag{28}$$

$$M^U_{lm,c2}=\left[\frac{8\pi^3}{\Omega}\right]\left[\mu_{lm}(\mathbf{k}+\mathbf{G}_2-\mathbf{G})-\sum_{l'm'}\mu_{l'm'}(\mathbf{k}+\mathbf{G}_2+\mathbf{q}-\mathbf{G})I_{lm,l'm'}(\mathbf{q})\right].$$

The reciprocal-lattice vector $\mathbf{G}$ is the vector that translates $\mathbf{k}+\mathbf{G}_2+\mathbf{q}$ back into the second BZ if it falls outside the second BZ (see Sec. V B below). The OPW normalization factor also has to be evaluated at $\mathbf{k}+\mathbf{G}_2$.

For our purposes it is important to clarify what we mean by an umklapp process in the calculation of the contribution of the second conduction band. For a given vector $\mathbf{q}$, we consider each $\mathbf{k}$ inside the first BZ. If the corresponding vector $\mathbf{k}+\mathbf{G}_2+\mathbf{q}$ falls outside the second BZ, we subtract a reciprocal-lattice vector $\mathbf{G}$ such that $\mathbf{k}+\mathbf{G}_2+\mathbf{q}-\mathbf{G}$ is inside the second BZ, and refer to it as an umklapp process. This method appears to be a reasonable and consistent method of handling umklapp processes for the second zone.

B. Evaluation of the real part

The integrals in the real part are volume integrals, which are replaced with the sum of integrals over the small cubes in the first BZ. For a cube with center at $\mathbf{k}_i$ we select $\mathbf{G}_2$ to give the corresponding point in the second BZ (Fig. 3). Then we need to check whether $\mathbf{k}_i+\mathbf{G}_2+\mathbf{q}$ is outside the second BZ. Since the second BZ is not a simply connected set of points, the checking process becomes rather complicated.

So we index back to the first BZ, and check whether $\mathbf{k}_i+\mathbf{q}$ lies outside the segment (determined by $\mathbf{G}_2$) in which $\mathbf{k}_i$ resides. When this happens two cases may occur. One is that $\mathbf{k}_i+\mathbf{G}_2+\mathbf{q}$ is outside the second BZ. In order to translate back to the second BZ, we first find $\mathbf{G}'$ such that $\mathbf{k}_i+\mathbf{G}_2+\mathbf{q}+\mathbf{G}'$ falls inside the first BZ. Then with $\mathbf{G}'_2$, which associates this last point with the second BZ, we find that $\mathbf{k}_i+\mathbf{G}_2+\mathbf{q}-\mathbf{G}$ lies inside the second BZ, where $-\mathbf{G}=\mathbf{G}'+\mathbf{G}'_2$. The other case is that $\mathbf{k}_i+\mathbf{G}_2+\mathbf{q}$ falls in a different segment of the second BZ. In this case we also follow the method just described, but with the result that $\mathbf{G}'=0$, because of property (1) of the Brillouin zones.

The procedure for integration is similar to the calculation for the first conduction band. For a normal process the equation for the sphere of singularity is

$$E_c|\mathbf{k}+\mathbf{G}_2+\mathbf{q}|^2+E_{10}|\mathbf{k}|^2+E_g-\omega=0\ . \tag{29}$$

Then we find the distance from the center of the cube to the surface of the sphere along the radial direction, decide whether this cube is singular or nonsingular, and proceed with the numerical integration as discussed previously. For an umklapp process we determine the sphere of singularity from

$$E_c|\mathbf{k}+\mathbf{G}_2+\mathbf{q}-\mathbf{G}|^2+E_{10}|\mathbf{k}|^2+E_g-\omega=0 \tag{30}$$

and proceed as before.

C. Evaluation of the imaginary part

As in the case of the calculation with the first conduction band, the problem in the evaluation of the imaginary part is in the treatment of the δ function that appears in the integrand. We integrate over the δ function by shifting to a coordinate system with origin at the center of the sphere of singularity, as described in Sec. IV. Then the remaining surface integral is evaluated numerically in the original coordinate system, proceeding cube by cube as before.

VI. ADJUSTMENT AND TESTS OF THE MODEL

A. Selection of parameters

The model insulator is described by the parameters a, E_g, E_c, E_{10}, and b. Now we are interested in applying

the dielectric function results to the Monte Carlo simulation of SEE from insulators. For this purpose we choose to model solid Ar, KCl, and Si, respectively, as representatives of wide, medium, and narrow band-gap insulators. The choice of these materials then determines the lattice parameter a, band gap E_g, and valence-band width through E_{10} and E_{11}. For uniformity we set $E_c=1$ for all materials, to correspond to the free-electron mass in the conduction band.

The remaining parameter is b, which determines the localization of the valence-band wave functions in Eq. (7). One criterion is to adjust the value of b until $\epsilon(0,0)$ is equal to the experimental value for $\epsilon_0^{(5)}$. Another criterion is to adjust b to satisfy the f-sum rule:

$$\frac{2}{\pi}\int_0^\infty \omega\epsilon_2(\omega)d\omega=\omega_p^2=\frac{4\pi ne^2}{m} . \tag{31}$$

Since the plasma frequency ω_p is directly related to the number of electrons per unit volume n in the topmost valence band,[11] this rule provides a simple physical basis for calibrating the model. We shall discuss the application of these criteria for choosing b for each material in Sec. VII. The extent to which both criteria can be satisfied by a suitable choice of b provides an overall consistency check for the simplified band structure of our model insulator.

B. Kramers-Kronig relations

In Secs. III and IV we have described our methods of calculation of the real and imaginary parts of the dielectric function. Although the geometry of the sphere of singularity is the same for both ϵ_1 and ϵ_2, the details of the numerical integration are very distinct. For ϵ_1 we need to evaluate a principal-value integral over the volume of an integration cube containing or adjacent to the surface of singularity, whereas for ϵ_2 we first perform the δ-function integration analytically and then evaluate the remaining surface integral numerically. Both of these procedures involve moderately complicated geometry and require considerable care in order to obtain reasonable, well-behaved numerical results.

Fortunately, we are able to obtain an overall independent check of the calculations by making use of the Kramers-Kronig relations, whereby the real and imaginary parts of the dielectric function are related by

$$\epsilon_1(\mathbf{q},\omega)=1+\frac{2}{\pi}P\int_0^\infty \frac{\omega'\epsilon_2(\mathbf{q},\omega')}{\omega'^2-\omega^2}d\omega' , \tag{32}$$

$$\epsilon_2(\mathbf{q},\omega)=-\frac{2\omega}{\pi}P\int_0^\infty \frac{\epsilon_1(\mathbf{q},\omega')-1}{\omega'^2-\omega^2}d\omega' . \tag{33}$$

These relations are of very general validity. From a physical point of view, they depend primarily on the assumption of causal connection and linear response between the induced polarization and electric field.[12] It is also implied that the relationship between the displacement and electric field vectors is local in space, and that $\epsilon(\mathbf{q},\omega)$ is well behaved at $\omega=0$, both of which are reasonable assumptions for insulators.[12]

In order to apply the above relations we first evaluate the real and imaginary parts of the dielectric function by the direct numerical methods discussed earlier, and then use the calculated value of ϵ_2 in Eq. (32) to find the corresponding ϵ_1, and vice versa. In practice, we must select a finite upper limit ω_T for the numerical integration of Eq. (33). For our model the maximum frequency $\omega_{\max}$ for nonzero ϵ_2 is the sum of the width of the valence band, energy gap, and the energy at the top of the second conduction band. We find $\omega_T=5\omega_{\max}$ to be a satisfactory choice, and divide the frequency range from 0 to ω_T into 300 equal intervals. We handle the principal-value integral by procedures similar in principle to those in Sec. III, and with $\omega_i'=(\omega_i+\omega_{i+1})/2$ obtain

$$\epsilon_1(\mathbf{q},\omega)=1+\frac{1}{\pi}\sum_{i=0}^{299}\epsilon_2(\omega_i')\left[\ln\frac{|\omega_{i+1}-\omega|}{|\omega_i-\omega|}-\ln\frac{\omega_{i+1}+\omega}{\omega_i+\omega}\right] \tag{34}$$

with similar results for $\epsilon_2(\mathbf{q},\omega)$.

Even though our model is incomplete, namely, it does not include all of the higher conduction bands, the Kramers-Kronig relations are still applicable to what we calculate for ϵ_1 and ϵ_2 with only two conduction bands. This conclusion follows from superposition, which is allowed by the linearity of the Kramers-Kronig relations. By the same argument, the Kramers-Kronig relations should apply separately to the normal parts ϵ_1^N and ϵ_2^N, and also to the umklapp parts ϵ_1^U and ϵ_2^U. The results of the Kramers-Kronig check will be shown in Sec. VII.

VII. RESULTS

Now we have all the ingredients to perform the numerical calculation of the dielectric function. In this section we present figures showing representative results, and also provide some comments about features of the behavior. The parameters for each model insulator are summarized in Table I.

A. Discussion of the results for argon

In this case we find that the f-sum rule and static dielectric constant can be satisfied simultaneously. We fit the static dielectric constant with $b=1.18$, and then find the sum rule satisfied, for six $3p$ electrons per atom, to

TABLE I. Lattice constant a, band gap E_g, bandwidth of the top valence band *BWVB*, bandwidth of the first conduction band *BWCB*, and the parameter of the Slater orbital b for solid Ar, KCl, and model Si.

Material	a (Å)	E_g (eV)	*BWVB* (eV)	*BWCB* (eV)	b (a_0^{-1})
Ar	5.29[a]	13.3[a]	0.6[a]	5.37[b]	1.18
KCl	6.28[a]	8.5[a]	1.2[c]	3.81[b]	0.91
Si	4.31[d]	1.12[a]	5.5[e]	8.08[b]	1.52

[a]Reference 5.
[b]Reference 17.
[c]Reference 19.
[d]Reference 18.
[e]Reference 20.

within 2%. This agreement shows that the model insulator works well for describing the dielectric behavior of argon.

In Fig. 4 we show the momentum dependence of the real part of the dielectric function at $\omega=0$ for argon, along with comparison to Fry's calculations.[5] For this purpose we carried out one calculation with only one conduction band, and with $m=2m_e$ in the conduction band. In this way the only difference between our calculation and Fry's is that we do the integration over the actual BZ, whereas Fry did his calculation over a spherical zone. The free parameter b in both calculations is adjusted by fitting the long-wavelength limit of ϵ_1 to the experimental result, which explains the agreement around $q=0$. Both results show the same general dependence on q, except that Fry's result is somewhat smaller than ours in the midvalues of q, which may be the result of his use of a spherical zone. As q becomes larger (or as wavelength becomes shorter), the response of the electron system becomes weaker, and in the limit of large q the dielectric response becomes 1.

In Fig. 5 we show the real and imaginary part of the dielectric function for solid Ar at $q=0$. For this and all subsequent calculations we include both first and second conduction bands, and also use $m=m_e$ in the conduction bands. Several important features are observable in Fig. 5. (1) The general shape of the real part of the dielectric function is appropriate. We see that ϵ_1 first increases as ω increases, then goes through a region of anomalous dispersion, and finally moves into a normal dispersion regime at higher frequencies. (2) The form of $\epsilon_1(q,\omega)$ is a smooth function of ω, and there are no noticeable irregularities. This behavior shows that the numerical procedure for handling the singularities in the integrands for ϵ_1 are reasonably adequate.

(3) The imaginary part of the dielectric function is zero for $\omega<E_g$. For values of ω large enough to excite electrons across the band gap into the conduction band, we see the expected behavior of a broad resonant absorption.

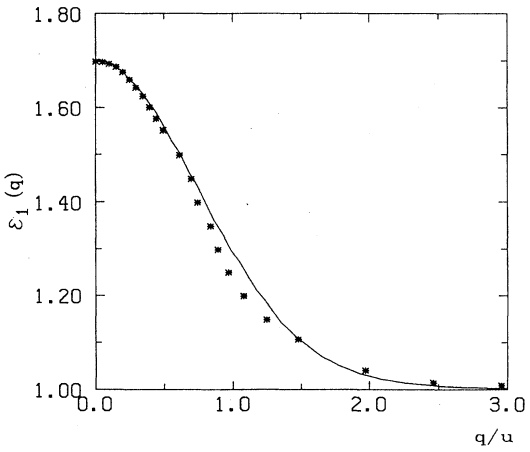

FIG. 4. Real part of the dielectric function for solid argon at $\omega=0$ vs q. (The momentum u is the distance from Γ point to X point in the first BZ.) Our results (solid line) are compared to Fry's calculations (symbols).

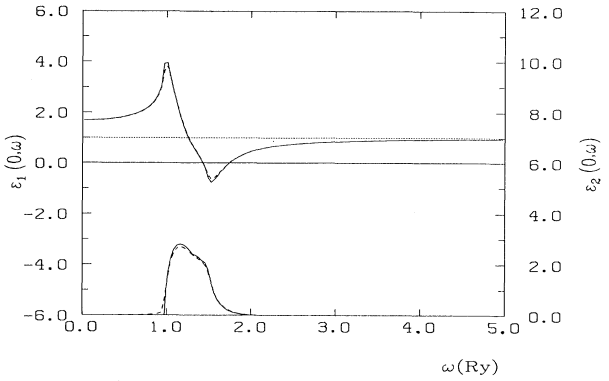

FIG. 5. The real and imaginary part of the dielectric function vs frequency in solid argon at $q=0$. In each graph the result of the Kramers-Kronig transform is shown as a dashed line.

With only two conduction bands in our model we expect ϵ_2 to become zero for ω larger than the energy of the top of the second conduction band, since no conduction-band states are available at higher energies. There is some overlap in energy between the first and second conduction bands, and the shoulder seen in ϵ_2 is apparently due to the second conduction band. Also see Fig. 7 in this regard.

(4) The other important feature in Fig. 5 is the result of the Kramers-Kronig relations. It can be seen that there is excellent agreement with the Kramers-Kronig transforms for both ϵ_1 and ϵ_2. A striking feature is that ϵ_2 from the Kramers-Kronig transform remains zero for $\omega<E_g$, and then rises sharply as ω becomes greater than E_g. The transforms do show slight rounding at the corner and very small deviations in the narrow peaks, as might be expected for numerical calculations. The overall agreement shows that the different numerical methods of integration for ϵ_1 and for ϵ_2 are both appropriate and consistent.

In Fig. 6 we show the real part of the dielectric func-

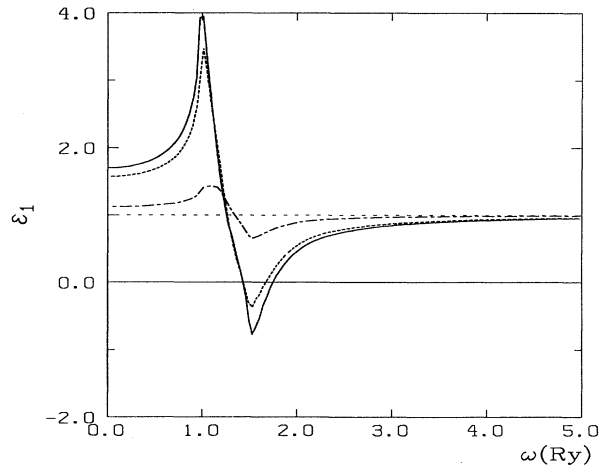

FIG. 6. Real part of the dielectric function in solid argon at different q's. The cases are $q=0$ (solid line), $q=0.5u$ (dashed line), and finally $q=1.15u$ (dot-dashed line). u is defined in Fig. 4.

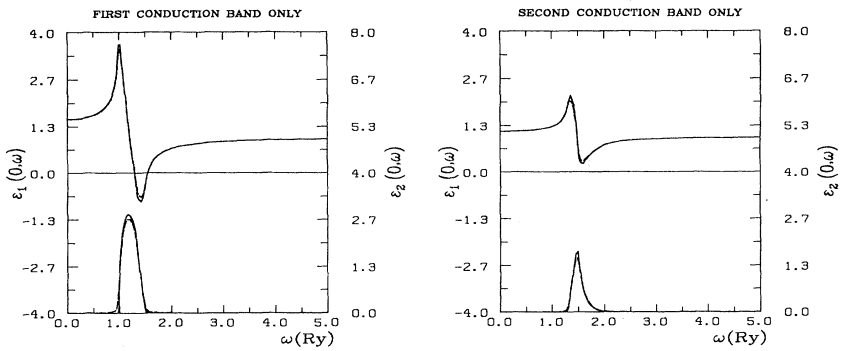

FIG. 7. The contribution of the first and second conduction bands to the dielectric function in solid argon with $q=0$.

tion for three values of q. The real part decreases as q increases, as we also saw in Fig. 4 for $\omega=0$. In Fig. 7 we show the separate contributions of the first and of the second conduction bands, along with the separate Kramers-Kronig transforms. It can be seen that the contribution of the second conduction band is smaller than the first, but is still significant. The second conduction band is especially important for SEE, since the higher-energy internal secondary electrons produced by those transitions have a better chance of escaping from the solid.

B. Discussion of the results for KCl

In the case of KCl the parameter b cannot be chosen to fit both the sum rule and the static dielectric constant simultaneously. The experimental value for the electronic contribution to the low-frequency dielectric constant is $\epsilon_0=2.13$, which is obtained with our model for $b=0.84$. But to satisfy the sum rule for six electrons per Cl^- ion, we need $b=0.91$ instead, making $\epsilon_0=1.55$ noticeably smaller.

We can understand the origin of this difference for KCl. The electronic polarizability of Cl^- is 2.947 $Å^3$, and that of K^+ is 1.133 $Å^3$.[13] Therefore the ratio of the contribution of the Cl^- ions to the total polarizability in KCl is 0.72.[14] Now consider our numbers. In order to count the number of electrons in the valence band properly, we use $b=0.91$ to fit the sum rule in the calculation of the dielectric function for KCl. But since the topmost valence band in our model includes only $3p$ electrons of Cl^- ions, we are not surprised to find our calculated value $\epsilon=1.55$ to be too small. The $3p$ electrons of K^+ ions will be included through ionization in the actual simulation for SEE.

In Figs. 8 and 9 we show the cases of $q=0$ and $q=0.8u$ for KCl. The band gap in this case is 8.5 eV, and therefore the imaginary part starts to become nonzero at $\omega=0.625$ Ry. The lattice constant in this case is smaller than argon, and therefore (with the same effective mass in the conduction band) we expect a narrower conduction bandwidth. There is some structure in the real part of the dielectric function, which also is consistent with the result of the Kramers-Kronig transform.

C. Discussion of the results for model Si

In the case of silicon we have more difficulty choosing the parameter b. For $b=1.52$ the sum rule for four electrons per silicon atom is satisfied, but the model gives $\epsilon_0=6.15$ instead of the experimental value 11.7. On the other hand, $b=1.36$ fits the experimental value of ϵ_0, but then the sum rule is off. Since we want to count the number of electrons properly, we use $b=1.52$ to satisfy the sum rule. These difficulties show that our simple model, with simply connected parabolic conduction bands, does not seem to work as well for silicon.

In Figs. 10 and 11 we show the results for $q=0$ and $q=0.5u$. In silicon the band gap is 1.12 eV (0.082 Ry), and therefore the imaginary part starts to increase from zero at much smaller values of ω (compared with argon and KCl). Perhaps the presence of a much wider valence band in silicon (width of 5.5 eV for silicon compared to 0.6 eV for argon and 1.2 eV for KCl) accounts for some of the more prominent structural features seen for both ϵ_1 and ϵ_2. The very good check of the Kramers-Kronig relations shows that the numerical calculations have been done consistently.

D. Conclusion

We have calculated the dielectric function $\epsilon(\mathbf{q},\omega)$ for a wide variety of insulating solids, based on our simple model insulator. Good values are obtained for the static dielectric constant and/or f-sum rule in most cases. The results display the dependence on wave vector and fre-

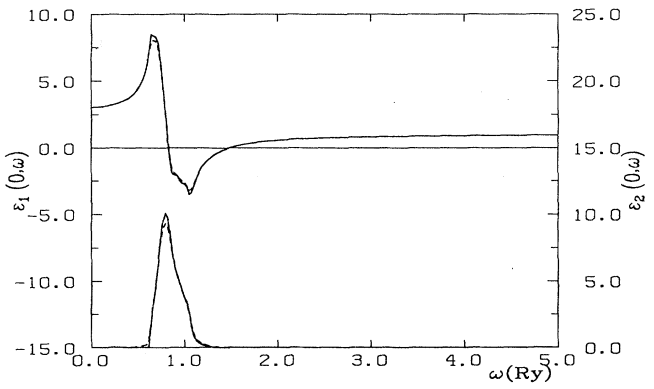

FIG. 8. The real and imaginary part of the dielectric function vs frequency in KCl at $q=0$. In each graph the result of the Kramers-Kronig transform is shown as a dashed line.

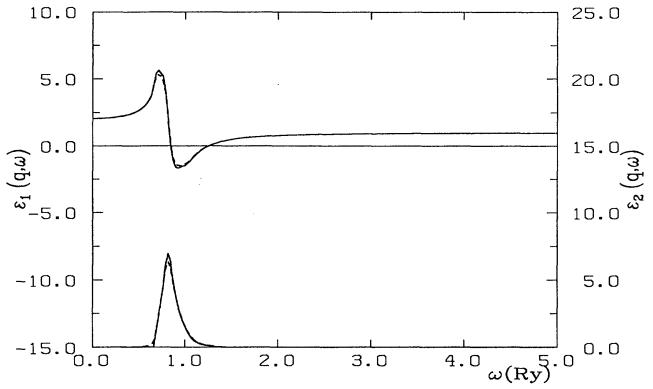

FIG. 9. The real and imaginary part of the dielectric function vs frequency in KCl at $q=0.8u$. In each graph the result of the Kramers-Kronig transform is shown as a dashed line. u is defined in Fig. 4.

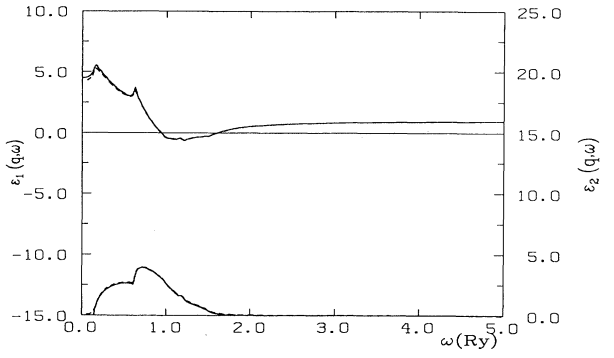

FIG. 11. The real and imaginary part of the dielectric function vs frequency in silicon at $q=0.5u$. In each graph the result of the Kramers-Kronig transform is shown as a dashed line.

quency in a reasonable fashion, in accord with the basic features of the model. An excellent self-consistency check is provided by the Kramers-Kronig relations.

The main purpose is to obtain a broad, general description of the dielectric response for typical narrow and wide band-gap insulators. Our particular interest is to use the imaginary part of $1/\epsilon(q,\omega)$ to describe inelastic collisions that produce internal secondary electrons. This information is then used, along with treatment of ionization processes and elastic scattering by core electrons, to perform a Monte Carlo simulation of SEE for a variety of insulators (to be discussed in a subsequent paper). We obtain maximum yields of 1.5 to 2.0 for Si, 7 to 9 for KCl and Ar, and 0.7 for aluminum (using the Lindhard dielectric function[15]), in good agreement with experimental observations. Thus we have been able to account for the major features of SEE for insulators, thereby demonstrating the suitability and applicability of the dielectric response function calculations.

ACKNOWLEDGMENTS

The authors express their appreciation to Ronald Bass for stimulation in formulating the problem, and to John Ralston and Douglas McKay for useful discussions about the Kramers-Kronig relations. We acknowledge partial support of the research by the NASA Lewis Research Center through Contract No. NSG 4223.

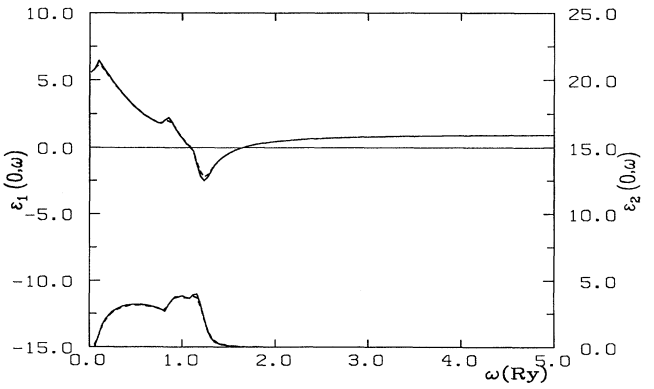

FIG. 10. The real and imaginary part of the dielectric function vs frequency in silicon at $q=0$. In each graph the result of the Kramers-Kronig transform is shown as a dashed line.

APPENDIX A: CALCULATION OF MATRIX ELEMENTS

For the calculation of the dielectric functions in Eqs. (12) and (13), matrix elements appear between states in the conduction and in the valence bands. Let us define

$$M_{lm,c}=\langle\psi_{lm}(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_c(\mathbf{k}')\rangle\ , \tag{A1}$$

$$M_{c,lm}=\langle\psi_c(\mathbf{k})|e^{-i\mathbf{q}\cdot\mathbf{r}}|\psi_{lm}(\mathbf{k}')\rangle\ , \tag{A2}$$

where $\mathbf{k}$ and $\mathbf{k}'$ are assumed to be two arbitrary wave vectors. It is also convenient to introduce

$$I_{lm,l'm'}(\mathbf{q})=\int U^*_{lm}(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}}U_{l'm'}(\mathbf{r})d\mathbf{r}\ , \tag{A3}$$

where $U_{lm}(\mathbf{r})$ is the atomic wave function from Eq. (6).

1. Evaluation of $M_{lm,c}$ and $M_{c,lm}$

With the above definitions it can be shown that[16]

$$M_{lm,c}=\frac{1}{N}\left[\mu_{lm}(\mathbf{k}'-\mathbf{q})-\sum_{l'm'}\mu_{l'm'}(\mathbf{k}')I_{lm,l'm'}(\mathbf{q})\right]\times\sum_{\nu}e^{-i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}_\nu}\ , \tag{A4}$$

where μ_{lm} are the orthogonalization coefficients from Eq. (10). With respect to the last summation two cases may happen. (i) A normal process occurs when $\mathbf{k}+\mathbf{q}$ is inside the first BZ. Then,

$$\sum_{\nu}e^{-i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}_\nu}=N\left[\frac{8\pi^3}{\Omega}\right]\delta(\mathbf{k}+\mathbf{q}-\mathbf{k}')\ . \tag{A5}$$

The effect of the δ function is that only values of $\mathbf{k}'$ contribute such that $\mathbf{k}'=\mathbf{k}+\mathbf{q}$; so we can write $M^N_{lm,c}$ for the normal process as

$$M^N_{lm,c}=\left[\frac{8\pi^3}{\Omega}\right]\left[\mu_{lm}(\mathbf{k})-\sum_{l'm'}\mu_{l'm'}(\mathbf{k}+\mathbf{q})I_{lm,l'm'}(\mathbf{q})\right]. \tag{A6}$$

(ii) An umklapp process occurs when the tip of the vector $\mathbf{k}+\mathbf{q}$ lies outside the first BZ. Then $\mathbf{k}+\mathbf{q}=\mathbf{s}+\mathbf{G}$, where $\mathbf{s}$ is a vector inside the first BZ, and $\mathbf{G}$ is a nonzero reciprocal-lattice vector. Then the sum over ν in Eq. (A5) introduces $\delta(\mathbf{s}-\mathbf{k}')$, and we obtain

$$M^U_{lm,c}=\left[\frac{8\pi^3}{\Omega}\right]\times\left[\mu_{lm}(\mathbf{k}-\mathbf{G})-\sum_{l'm'}\mu_{l'm'}(\mathbf{k}+\mathbf{q}-\mathbf{G})I_{lm,l'm'}(\mathbf{q})\right]. \tag{A7}$$

The calculation of $M_{c,lm}$ is very similar. The result for both normal and umklapp cases is

$$M^U_{c,lm}=M^N_{c,lm}=\left[\frac{8\pi^3}{\Omega}\right]\left[\mu_{lm}(\mathbf{k}+\mathbf{q})-\sum_{l'm'}\mu^*_{l'm'}(\mathbf{k})I_{l'm',lm}(\mathbf{q})\right]. \tag{A8}$$

It is interesting to note that in this case both umklapp and normal processes lead to the same form. This is because the valence-band wave functions have the periodicity of the reciprocal lattice, and therefore nothing changes upon subtraction of a reciprocal-lattice vector $\mathbf{G}$ from the argument.

2. Introduction of atomic orbitals

In order to evaluate $M_{lm,c}$, etc., we use the atomic orbital from Eq. (7). We also choose $\mathbf{q}$ along the k_z axis in reciprocal space. Only μ_{10}, μ_{11}, I_{1010}, and I_{1111} are needed for all calculations:

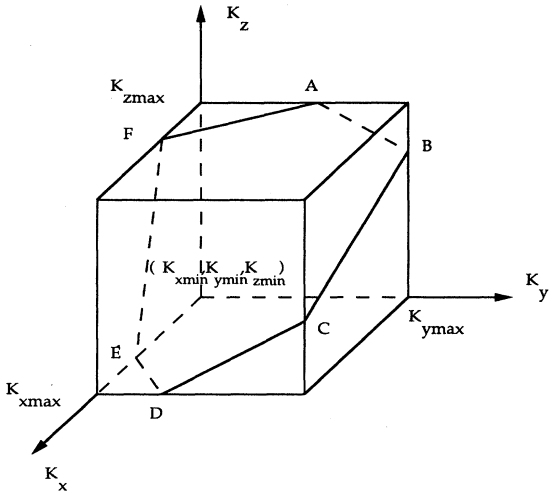

FIG. 12. A singular cube that is intersected with the plane of singularity. The intersection with the edges of the cube takes place at points A,B,C,D,E, and F.

$$\mu_{10}(\mathbf{k})=\left[\frac{12\pi N}{V}\right]^{1/2}\left[\frac{(2b)^5}{24}\right]^{1/2} i\cos\theta_{\mathbf{k}}\frac{8b|\mathbf{k}|}{(b^2+|\mathbf{k}|^2)^3}, \tag{A9}$$

$$\mu_{11}(\mathbf{k})=\left[\frac{12\pi N}{V}\right]^{1/2}\left[\frac{(2b)^5}{24}\right]^{1/2} i\sin\theta_{\mathbf{k}}e^{-i\phi_{\mathbf{k}}}\frac{8b|\mathbf{k}|}{(b^2+|\mathbf{k}|^2)^3}. \tag{A10}$$

Here $\theta_{\mathbf{k}}$ is the angle between $\mathbf{q}$ and $\mathbf{k}$. Also,

$$I_{1010}(\mathbf{q})=\frac{(2b)^6[(2b)^2-5q^2]}{[(2b)^2+q^2]^4}=a_2, \tag{A11}$$

$$I_{1111}(\mathbf{q})=\frac{(2b)^6}{[(2b)^2+q^2]^3}=a_3. \tag{A12}$$

With these forms incorporated in Eqs. (A6), (A7), and (A8) we obtain the following typical results for the matrix elements, with $C=(8\pi^3/\Omega)^2(8\pi N/V)(2b)^7$:

$$|M^N_{10,c}|^2=C\left[\frac{\cos\theta_{\mathbf{k}}|\mathbf{k}|}{(b^2+|\mathbf{k}|^2)^3}-\frac{\cos\theta_{\mathbf{k}+\mathbf{q}}|\mathbf{k}+\mathbf{q}|}{[b^2+|\mathbf{k}+\mathbf{q}|^2]^3}\cdot a_2\right]^2, \tag{A13}$$

$$|M^N_{c,10}|^2=C\left[\frac{\cos\theta_{\mathbf{k}+\mathbf{q}}|\mathbf{k}+\mathbf{q}|}{(b^2+|\mathbf{k}+\mathbf{q}|^2)^3}-\frac{\cos\theta_{\mathbf{k}}|\mathbf{k}|}{[b^2+|\mathbf{k}|^2]^3}\cdot a_2\right]^2, \tag{A14}$$

$$|M^N_{1-1,c}|^2=|M^N_{11,c}|^2=\frac{C}{2}\left[\frac{\sin\theta_{\mathbf{k}}|\mathbf{k}|}{(b^2+|\mathbf{k}|^2)^3}-\frac{\sin\theta_{\mathbf{k}+\mathbf{q}}|\mathbf{k}+\mathbf{q}|}{[b^2+|\mathbf{k}+\mathbf{q}|^2]^3}\cdot a_3\right]^2, \tag{A15}$$

$$|M^U_{1-1,c}|^2=|M^U_{11,c}|^2=\frac{C}{2}\left[\frac{\sin\theta_{\mathbf{k}-\mathbf{G}}|\mathbf{k}-\mathbf{G}|}{(b^2+|\mathbf{k}-\mathbf{G}|^2)^3}-\frac{\sin\theta_{\mathbf{k}+\mathbf{q}-\mathbf{G}}|\mathbf{k}+\mathbf{q}-\mathbf{G}|}{[b^2+|\mathbf{k}+\mathbf{q}-\mathbf{G}|^2]^3}\cdot a_3\right]^2. \tag{A16}$$

Expressions for $M^U_{c,10}$, etc., can be obtained from Eq. (A8).

APPENDIX B: EVALUATION OF THE PRINCIPAL-VALUE INTEGRAL

In evaluating Eq. (20) two cases may occur. For a singular cube that is not cut by the plane of singularity, we do not actually have a principal-value integration, but we still want to evaluate Eq. (20) analytically. Let $(k_{x\min},k_{y\min},k_{z\min})$,

$(k_{x\min}, k_{y\min}, k_{z\max})$, . . . be the coordinates of the corners of the cube centered at $\mathbf{k}_i$. Also, define the function $F(x,y,z)$ as

$$F(x,y,z)=(Ax+By+Cz-D)^2[\ln|Ax+By+Cz-D|-\tfrac{3}{2}] . \tag{B1}$$

After some algebra it can be shown that[16]

$$I_S=\frac{1}{2ABC}[F(k_{x\max},k_{y\max},k_{z\max})-F(k_{x\max},k_{y\max},k_{z\min})-F(k_{x\max},k_{y\min},k_{z\max})+F(k_{x\max},k_{y\min},k_{z\min})$$
$$-F(k_{x\min},k_{y\max},k_{z\max})+F(k_{x\min},k_{y\max},k_{z\min})+F(k_{x\min},k_{y\min},k_{z\max})-F(k_{x\min},k_{y\min},k_{z\min})] . \tag{B2}$$

In the other case, when the plane of singularity passes through the integration cube, the total number of intersections with the edges of the cube can be 3, 4, 5, or 6, depending on the orientation and location of the plane with respect to the center of the cube. Figure 12 shows a typical situation with six points of intersection.

For any point in the cube that lies on the plane of singularity, the denominator is zero and the integrand becomes singular. It can be shown, for all possible geometries, independent of the number of intersections, that the result of Eq. (B2) is obtained.[16] If the plane of singularity passes through one or more corners of the cube, it means that the argument of F for that corner would be zero. For the purpose of calculation on the computer such terms must be set equal to zero to avoid asking the computer to evaluate the logarithm of zero. In this case the squared factor before the brackets in Eq. (B1) more than cancels the mild singularity from the logarithmic term inside the brackets.

*Present address: Wafer Scale Integration, Fremont, CA 94538.

[1]Setsuo Ichimaru, *Plasma Physics* (Benjamin/Cummings, Menlo Park, CA, 1986).
[2]H. R. Philipp and H. Ehrenreich, Phys. Rev. **129**, 1550 (1963).
[3]R. A. Ferrell, Phys. Rev. **101**, 554 (1956).
[4]D. Pines, *Elementary Excitation in Solids* (Benjamin, New York, 1963).
[5]J. L. Fry, Phys. Rev. **179**, 892 (1969).
[6]H. Ehrenreich and M. Cohen, Phys. Rev. **115**, 786 (1959).
[7]P. Nozieres and D. Pines, Phys. Rev. **113**, 1254 (1959).
[8]J. L. Fry, Ph.D. thesis, University of California, Riverside, CA, 1966.
[9]H. Nara, J. Phys. Soc. Jpn. **20**, 778 (1965).
[10]N. O. Lipari, J. Chem. Phys. **53**, 1040 (1970).
[11]E. Shiles *et al.*, Phys. Rev. B **22**, 1612 (1980).
[12]J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[13]J. Tessman, A. Kahn, and W. Shockley, Phys. Rev. **92**, 890 (1953).
[14]C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976).
[15]J. P. Ganachaud and M. Cailler, Surf. Sci. **83**, 498 (1979).
[16]G. A. Rezvani, Ph.D. thesis, University of Kansas, Lawrence, KS, 1989.
[17]The width of the conduction band is calculated from the lattice constant and $m^*=m_e$ in the conduction band.
[18]The lattice constant of silicon is determined for an assumed fcc crystal structure.
[19]The width of the p_x and p_y bands in KCl is 0.7 eV each, and the width of the p_z band is 2.0 eV (Ref. 5). Since we assume that the widths of the p_x,p_y, and p_z bands are the same in our model, we just take the simple average of the three bandwidths for the width of the topmost valence band.
[20]The width of the p_x and p_y bands in Si is 3.2 eV each, and the width of the p_z band is 10.0 eV (Ref. 5). In this case also, as in the case of KCl, we just take the simple average of the three bandwidths for the width of the topmost valence band.