Shakeup in soft-x-ray emission. I. The low-energy tail

Peteris Livins and S. E. Schnatterly
Department of Physics, University of Virginia, Charlottesville, Virginia 22901
(Received 17 July 1987)

The L soft-x-ray emission spectra for Al and Si, and the K emission for graphite and diamond are analyzed, taking into account the effects of multielectron transitions which produce a low-energy tail on the main band. Through an appropriate transformation on the initial states, it is shown that the final-state rule of Von Barth and Grossman is a direct result of exchange. A "Pauli rigidity" forces the initial-state wave function to approach that of the final state, except near the Fermi energy. An assumption of random phases allows a formal solution that includes single-particle shakeup events to all orders. It is argued that a low-energy feature of silicon is structure due to the shakeup of single-particle-like excitations rather than a plasmon satellite.

I. INTRODUCTION

Interest in the spectra of soft-x-ray emission[1] has traditionally been concerned with the valence-electron density of states for a solid. However, multielectron or many-body effects are also a contributing factor. The most notable of these is the threshold singularity at the Fermi edge in metals. In addition, there is a less striking but always evident broadening of the emission bands on the low-energy side, producing a tailing extending tens of eV below the expected band minimum. Where such features may be neglected however, qualitative agreement with an independent electron picture has been good. For simple metals in particular, the bandwidths are close to what is expected in a noninteracting theory, general shapes agree with what is expected from dipole selection rules and a sharp Fermi edge is observed.

Efforts to understand soft-x-ray emission spectra which go beyond one-electron effects have enriched the entire field. An early calculation by Landsberg[2] investigated the low-energy tail for the sodium $L_{\mathrm{II,III}}$ band, taking into account both the effects of lifetime broadening and multiple transitions (shakeup or radiative Auger processes). He concludes that the lifetime broadening is the dominant effect. In a later calculation Pirenne and Longe[3] interpret the tail as due mainly to the multiple transitions. As first proposed by Ferrell,[4] and later confirmed by Rooke,[5] a plasmon satellite is often observed contributing to the low-energy tail. Since then, various many-body calculations[6–8] have been attempted and improved upon for a free-electron gas in the efforts to explain the emission for a metal. Here we will refer to shakeup processes as an accompanying electronic excitation with the radiative transition of an electron from the valence band to the empty core level. Such excitations include single particle and collective processes. As far as single-particle and collective excitations may be regarded as distinct elementary excitations, we will use the terms multielectron or radiative Auger processes to describe those shakeup events involving single-particle-like excitations.

We will be reporting on the low-energy features in the L emission of Si and Al, and the K emission of both graphite and diamond. Entangled in our analysis as two distinct effects are both the multielectron transitions and the plasmon satellite, and so two independent descriptions will be utilized in developing an understanding of the spectra. Therefore, we have divided our discussion into two companion papers, this one being referred to as I, and the paper immediately following this,[9] as II. In II, we describe a model used in the analysis of the plasmon satellite. Here, we will focus primarily on the multielectron processes, but with some comment on the consolidation of the two effects. Aberg, Utriainen[10,11] and others have looked at the radiative Auger effect in atomic transitions, but little experimental work that we are aware of has been carried out to investigate these processes in the emission from a valence band.

The neglect to consider these multielectron transitions in spectra at the band minima has hampered the determination of bandwidths and valence hole lifetime broadening in soft-x-ray emission. The present approach will try to address this difficulty. An added benefit of this study is that it results in backgrounds, not arbitrarily drawn, for the determination of the plasmon satellite strengths and shapes.

II. EXPERIMENT

The specimens were excited with a 3-kV electron beam with current densities of up to 10 mA/mm^2. The emission is then dispersed with a grazing incidence holographically etched grating. Soft x-rays are detected with a modified photodiode array which integrates the signal simultaneously over a broad wavelength region. Further details are described elsewhere.[12–14]

It is crucial in an accurate measurement of the radiative Auger tails to understand and properly subtract the continuous bremsstrahlung background. Our gratings allow diffraction in many orders. The effect produces a universal background shape due to the "pile up" of many orders of bremsstrahlung from higher energies, up to about 800 eV, where the reflectivity of our gratings drop. We have extracted this universal shape for each grating

and need then only to choose an appropriate scale factor to subtract this contribution from a given spectrum. Extending the spectra to regions far above the main band and beyond any doubly ionized cores determines the scale to an adequate degree of confidence.

Historically, the literature involving soft-x-ray emission has presented data in various forms. Different instruments involve different detector efficiencies, and therefore reduce data with different frequency dependences. Furthermore, when the valence-band density of states is the primary concern, an added frequency correction is often included. Here, in all figures that plot data, in both I and II, we plot transitions per unit time per unit energy in arbitrary units, corresponding to a transition rate, such as given by Eq. (1).

III. THEORY

In developing a theoretical understanding of our spectra, we write the observed transition rate[15] $T(\nu)$ (here ν corresponds to the energy of the emitted photon, but will often be referred to as a frequency) as

$$T(\nu) \propto \sum_f \nu \,|\, M_{fi} \,|^2 \delta(E_i - E_f - \nu) \;, \tag{1}$$

where $M_{fi} = \langle \Psi_f \,|\, R \,|\, \Phi_i \rangle$. Φ_i with E_i and Ψ_f with E_f are the initial and final many-body electronic states and energies, respectively, while ν is the photon frequency. $R = \sum_i \mathbf{A}\cdot\mathbf{p}_i$ is the radiative transition operator, $\mathbf{A}$ being the vector potential and $\mathbf{p}_i$ and ith electron's momentum. For single-particle transitions, no electron-electron interaction, and a long-lived core state (<0.1 eV), Eq. (1) reduces to an expression proportional to the valence-band density of states, a single-particle matrix element and the frequency factor. In such a single-particle approximation the remaining passive electrons are ignored. Since our concern will be with the effects of the multiple-particle excitations, we reexamine the many-particle matrix element M_{fi}.

One step beyond this simple approach, but which still retains independent particles, is to take into account the different ionic potentials felt by the valence electrons when the core state is occupied or empty. The works of Mahan,[16] Nozieres, De Dominicis,[17] and Combescot[18] have provided improved descriptions of this problem, with asymptotically exact solutions near the Fermi edge in metals. These methods involve the calculation of the appropriate correlation functions. Alternatively, one can consider the initial and final states constructed with Slater determinants. For emission, the single-particle initial states ϕ are distorted due to existence of the vacant core level, while the final states ψ are the undistorted Bloch states. In what follows we adapt this view, and furthermore assume a frozen core approximation, i.e., the core states are orthogonal to the valence states of both orthonormal sets. The two approaches are then equivalent, and although the correlation functions are more suited to large systems, the determinantal method will provide some useful results.

The matrix element M_{fi} is then a determinant[19] constructed by antisymmetrizing the initial and final single particle states. In considering this determinant, Davis and Feldkamp[20] have noted that in a simple single-particle calculation it is a good approximation to use orthogonalized final state (OFS) orbitals for the active electronic state (that state from which, in emission, or to which, in absorption, the electronic transition occurs). The OFS state to use in emission is given by

$$|\, \psi_i^{\mathrm{OFS}} \rangle = |\, \psi_i \rangle - \sum_j \langle \phi_j \,|\, \psi_i \rangle \,|\, \phi_j \rangle \;, \tag{2}$$

where j sums over the unoccupied states.

In considering absorption, Stern and Rehr[21] later suggest that a unitary transformation on the occupied valence states, that orthogonalizes the active electron's distorted state to the other occupied undistorted states, achieves a simplification of the many-particle Slater determinant. This simplification also produces a modified single-particle matrix element, but one which is modulated by the overlap of the remaining passive electrons. With the assumptions taken here, this approach is exact, and therefore useful in describing near-edge phenomena as well as multielectron transitions, but the energy dependence of the passive overlap remains a difficulty. Interestingly, as suggested by Stern and Rehr, such a transformation provides insight into the physical source behind the final-state rule.[22]

In this paper we propose a modification of the Stern-Rehr procedure. We first consider emission; absorption can then be treated as the emission of holes.[23] We consider a more symmetric transformation, which rather than considering the orthogonality of just the active electron's state, transforms all the distorted occupied states so that the new distorted states ϕ' are orthogonal to all other *occupied* undistorted states ψ, except the corresponding state ψ from which ϕ "evolved." This transformation Z, then takes the states ϕ to ϕ' through

$$\phi_i' = \sum_{j=1}^{M} Z_{ji}\phi_j \tag{3}$$

and satisfies (for $p \leq M$)

$$\sum_{j=1}^{M} Z_{ji} S_{pj} = \lambda_i \delta_{pi} \;, \tag{4}$$

where

$$\phi_i = \sum_{j=1}^{N} S_{ji}\psi_j \;. \tag{5}$$

Here M denotes the number of occupied states, N the total number of states in the orthonormal set, $S_{ji} = \langle \psi_j \,|\, \phi_i \rangle$, and $\lambda_i = \langle \psi_i \,|\, \phi_i' \rangle$ is chosen so as to retain normalization of the states ϕ'. Thus Z diagonalizes the occupied submatrix s, of the larger unitary matrix S, and therefore $Z_{ji} = s_{ji}^{-1}\lambda_i$, where s^{-1} is the inverse of s. The transformation is not unitary, except in the special cases $N = M$, or when the mixing of unoccupied states ψ in the expansion of occupied states ϕ is zero. However we do expect this mixing to be weak, except near the Fermi energy, a circumstance we shall refer to as weak mixing.

The advantage in introducing Z is the simplification it allows in the formal evaluation of the determinant of M_{fi}. The replacement terms vanish, as they would in a single orthonormal set, although now the complications are transferred to the evaluation of Z. The many-particle matrix element M_{fi} will be modified by a multiplicative factor, equal to the determinant of Z, which however, and most importantly, is independent of the final state (i.e., independent of the energy). We have

$$M_{fi}=\langle\Psi_j\mid R\mid\Phi_i\rangle=[\det(Z)]^{-1}\langle\Psi_f\mid R\mid\Phi_i'\rangle$$
$$=\frac{\det(s)}{\prod_j^M\lambda_j}\langle\Psi_f\mid R\mid\Phi_i'\rangle\ . \tag{6}$$

We first consider M_{fi} for the one-particle emission where the final state has one valence hole in the state ψ_n, and with the once vacant core filled. In this case

$$M_{fi}=\frac{\det(s)}{\prod_j^M\lambda_j}\sum_{p\le M}^{M}(-1)^p\langle c\mid t\mid\phi_p'\rangle\langle\psi_1,\ldots,(\text{no }\psi_n),\ldots,\psi_M\mid\phi_1',\ldots,(\text{no }\phi_p'),\ldots,\phi_M'\rangle\ . \tag{7}$$

Here $\mid c\rangle$ represents a core state and $t=\mathbf{A}\cdot\mathbf{p}$. The M-1 particle overlap is also a determinant, which due to the chosen transformation gives zero, except when $p=n$, in which case

$$M_{fi}=\frac{\det(s)}{\prod_j^M\lambda_j}(-1)^n\langle c\mid t\mid\phi_n'\rangle$$
$$\times\langle\psi_1\mid\phi_1'\rangle\cdots(\text{no }\langle\psi_n\mid\phi_n'\rangle)\cdots\langle\psi_M\mid\phi_M'\rangle$$
$$=(-1)^n\det(s)\frac{\langle c\mid t\mid\phi_n'\rangle}{\lambda_n}$$
$$=(-1)^n\det(s)\langle c\mid t\mid\bar{\phi}_n\rangle\ , \tag{8}$$

where we have introduced the "partially inverted initial state" (PIIS)

$$\bar{\phi}_n=\sum_j^M s_{jn}^{-1}\phi_j\ . \tag{9}$$

It is expected that a state ϕ which is far in energy from unoccupied states would include mostly occupied (near in energy) components of the ψ basis. In the limit of weak mixing we have arrived at a statement of the final-state rule. For in such a case, far enough in energy from any unoccupied states, the built-in orthogonality of the state ϕ' forces it to approach its corresponding state ψ. In addition, here we expect $|\lambda|^2\approx1$. However, the inclusion of components from the unoccupied states, although small, is not ruled out by this "Pauli rigidity;" indeed, such components are responsible for the multielectron processes. For states near the Fermi energy E_F, unoccupied components are not small, in which event the arguments forcing the final-state rule fail. Besides the x-ray edge singularity, this failure is most evident in the observation of excitonic emission.[24,25]

It becomes apparent that the final-state rule is a consequence of exchange. Put differently, far enough in energy from unoccupied states at least, using a final state potential is equivalent to properly antisymmetrizing the valence electron states. We mention that no knowledge of Z, besides its definition, has been required to note these consequences.

At this point we compare the OFS state to the present PIIS state. Equation (2) can be reexpressed,

$$\mid\psi_i^{\text{OFS}}\rangle=\mid\psi_i\rangle-\sum_{j=M+1}^{N}\langle\phi_j\mid\psi_i\rangle\mid\phi_j\rangle$$
$$=\sum_{j=1}^{M}\langle\phi_j\mid\psi_i\rangle\mid\phi_j\rangle$$
$$=\sum_{j=1}^{M}S_{ji}^{-1}\mid\phi_j\rangle\ , \tag{10}$$

where we have written $\langle\phi_j\mid\psi_i\rangle=S_{ji}^{-1}$ since S is unitary. Comparing Eqs. (10) and (9) we see that the distinction between the OFS and PIIS states lies entirely in whether one uses the inverse of the full matrix S, or the inverse of the smaller occupied submatrix s. As pointed out in Ref. 21, the OFS approximation breaks down near E_F. Nevertheless, the OFS states do include an edge enhancement, in fact, Green[26] has demonstrated that a logarithmic singularity exists with the OFS formalism. The present formalism, having considered the standard model (thus far without shakeup) exactly, should produce the correct Mahan power-law singularity.

We point out, in view of the present analysis, that model numerical computations[27,28] that determine the emission with determinantal many-particle wave functions need not evaluate any more determinants beyond one matrix inversion. This reduces the effort in calculating the single-particle-like emission, although the shakeup contribution still remains time consuming depending on the order of the shakeup desired.

As we have already mentioned, absorption can be viewed as emission of holes. One considers holes occupying the conduction band with an empty valence band. In this case the matrix s corresponds to the unoccupied N-M dimensional submatrix of S. Thus different matrices s^{-1} are involved between emission and absorption. Although in practice the difference may be insignificant, this brings up the question of a possible asymmetry in the edge singularities between emission and absorption.

For emission that includes multielectron transitions, we explicitly write Eq. (1) as

$$T(\nu)\propto\nu\left|\frac{\det(s)}{\prod_j^M\lambda_j}\right|^2\left[\sum_{\substack{y\\ y\le M}}|\langle\Psi(y)\,|\,R\,|\,\Phi_i'\rangle|^2\delta(\nu-E_y)\right.$$
$$+\sum_{n=1}^{M-1}\sum_{\substack{y,x\\ y_1,\ldots,y_{n+1}\le M\\ x_1,\ldots,x_n>M}}{}'\;[(n+1)!]^{-1}|\langle\Psi(y_1,\ldots,y_{n+1};x_1,\ldots,x_n)\,|\,R\,|\,\Phi_i'\rangle|^2$$
$$\left.\times\delta(\nu-(E_{y_1}+\cdots+E_{y_{n+1}})+E_{x_1}+\cdots+E_{x_n})\right]. \tag{11}$$

$\Psi(y_1,\ldots,y_{n+1};x_1,\ldots,x_n)$ denotes a final state with vacancies in the originally occupied states ϕ_y, with energy E_y, and with the originally unoccupied states ψ_x, with energy E_x, filled (we have assumed that the overall relaxation energy is a constant, and omit it from the delta functions). The prime on the sum indicates omission of terms where any summation index equals any other. In view of the transformation Z, the above expression becomes

$$T(\nu)\propto\nu\,|\det(s)|^2\left[\sum_{\substack{y\\ y\le M}}|\langle c\,|\,t\,|\,\bar\phi_y\rangle|^2\delta(\nu-E_y)\right.$$
$$+\sum_{n=1}^{M-1}\sum_{\substack{y,x\\ y_1,\ldots,y_{n+1}\le M\\ x_1,\ldots,x_n>M}}{}'\;[(n+1)!]^{-1}\left|\sum_{\varepsilon=1}^{n+1}(-1)^{\varepsilon}\langle c\,|\,t\,|\,\bar\phi_{y_\varepsilon}\rangle\right.$$
$$\left.\left.\times\langle\psi_{x_1},\ldots,\psi_{x_n}\,|\,\bar\phi_{y_1},\ldots,(\text{no }\bar\phi_{y_\varepsilon}),\ldots,\bar\phi_{y_{n+1}}\rangle\right|^2$$
$$\left.\times\delta(\nu-(E_{y_1}+\cdots+E_{y_{n+1}})+E_{x_1}+\cdots+E_{x_n})\right]. \tag{12}$$

To simplify this sum we make one very reasonable assumption. It will be assumed that the cross terms cancel in the sum over the x and y indices, i.e., the individual single-particle overlaps have random phases, in which case we obtain

$$T(\nu)\propto\nu\,|\det(s)|^2\left[\sum_{\substack{y\\ y\le M}}|\langle c\,|\,t\,|\,\bar\phi_y\rangle|^2\delta(\nu-E_y)\right.$$
$$+\sum_{n=1}^{M-1}\sum_{\substack{y,x\\ y_1,\ldots,y_{n+1}\le M\\ x_1,\ldots,x_n>M}}{}'\;[(n+1)!]^{-1}\sum_{\varepsilon=1}^{n+1}|\langle c\,|\,t\,|\,\bar\phi_{y_\varepsilon}\rangle|^2$$
$$\times\sum_{P_x}P_x\,|\langle\psi_{x_1}\,|\,\bar\phi_{y_1}\rangle\cdots(\text{no }\langle\psi\,|\,\bar\phi_{y_\varepsilon}\rangle)\cdots\langle\psi_{x_n}\,|\,\bar\phi_{y_{n+1}}\rangle|^2$$
$$\left.\times\delta(\nu-(E_{y_1}+\cdots+E_{y_{n+1}})+E_{x_1}+\cdots+E_{x_n})\right], \tag{13}$$

where P_x permutes the index x, and where we sum over all such permutations.

The evaluation of this sum is facilitated by defining the following function $p(x)$,

$$p(x)=\lim_{\Delta E\to 0}\frac{\sum_{m,q}|\langle\psi_q\,|\,\bar\phi_m\rangle|^2}{\Delta E}. \tag{14}$$

If $E(q)$ and $E(m)$ are the energies of the states ψ_q and ϕ_m, respectively, then the sum over q includes unoccupied states having energies such that $E(m)+x<E(q)<E(m)+x+\Delta E$. The sum over m includes occupied states such that $E(m)+x$ is an energy of an unoccupied state. $p(x)$ is a probability per energy, and in a model where $\langle\psi_q\,|\,\bar\phi_m\rangle$ is a function of just the energies $E(q)$ and $E(m)$, $p(x)$ is expressed as

$$p(x)=\int_{E_F-x}^{E_F}N(E)N(E+x)\,|\langle\psi(E+x)\,|\,\bar\phi(E)\rangle|^2dE, \tag{15}$$

where $N(E)$ is the single-particle density of states, and E_F is the Fermi energy in a metal, or the top of the band in an insulator.

With this definition we write, in going to a continuum approximation, the total transition rate

$$T(\nu)\propto \nu\,|\det(s)|^2\int_0^{\nu_F-\nu}A(x+\nu)\left[\delta(x)+p(x)+(2!)^{-1}\int_0^x p(z)p(x-z)dz\right.$$
$$\left.+(3!)^{-1}\int_0^x p(z)\int_0^{x-z}p(z')p(x-z-z')dz'dz+\cdots\right]dx\ , \tag{16}$$

where $A(\nu)=|\langle c\,|\,t\,|\,\bar{\phi}(\nu)\rangle|^2 N(\nu)$. $A(\nu)$ represents a transition density of states, and ν_F is the frequency corresponding to the Fermi energy for a metal, or the top of the band for an insulator. The matrix element $|\langle c\,|\,t\,|\,\bar{\phi}(\nu)\rangle|^2$ is an averaged matrix element corresponding to the energy ν. We will refer to $\nu A(\nu)$ as the zeroth-order emission.

This series can be summed by invoking the convolution theorem for Laplace transforms. The quantity in the larger brackets of Eq. (16), which we shall call $Q(x)$, is then given by

$$Q(x)=L^{-1}\{\exp[h(s)]\}\ , \tag{17}$$

where $h(s)=L[p(x)]$ is the Laplace transform of $p(x)$, and where L^{-1} denotes the inversion operation.

Another relevant property of the Laplace transform $f(s)=L[F(x)]$, for some function $F(x)$, relating the nth derivative $f^n(s)$ is

$$f^n(s)=L[(-x)^nF(x)]\ . \tag{18}$$

The convolution theorem with Eq. (18) provides the integral equation

$$\int_0^x(x-z)Q(z)p(x-z)dz=xQ(x) \tag{19}$$

for $Q(x)$. Since $Q(x)$ may be written as $Q(x)=\delta(x)+\bar{Q}(x)$, the above integral equation is also expressed as

$$\bar{Q}(x)=p(x)+\frac{1}{x}\int_0^x(x-z)\bar{Q}(z)p(x-z)dz\ . \tag{20}$$

The function $\bar{Q}(x)$ relates to that part of the emission due to just multielectron transitions.

IV. DISCUSSION

As it turns out, the effect of the size of the system is an interesting factor to be concerned with from this analysis. We expect that beyond a large enough volume the spectral shape should remain constant. For metals at least, the Anderson[29] theorem tells us that the zeroth-order emission (the final state with one valence hole) will be completely blocked for an infinite system. Therefore the multielectron transitions must make up for the lost strength, and do so in such a manner that the essential shape of the zeroth-order emission is retained.

With a localized core hole, the potential becomes an increasingly smaller perturbation for a wave function that has an increasingly greater extent as the size of the system grows. Thus the quantities $\langle\psi\,|\,\phi\rangle$, as functions of the energy, vary with increasing volume in such a manner which diminishes a particular shakeup event; i.e., the overlaps become narrower functions of the energy difference between the two states. At the same time the number of such states, both occupied and unoccupied, increases, favoring the multielectron transitions. The two trends must then compensate for each other "in the correct manner," at any volume. Although the relative strengths of zero shakeup, one shakeup etc. change with the size of the system, their total sum should remain unchanged, which suggests that in any model calculation with some particular volume, all significantly contributing shakeup orders need to be included. Equations (17) and (16) determine, provided our assumption of random phases is valid, the total emission to all orders for any size model system. Alternatively, Eqs. (20) and (16) determine, to all orders, just the multielectron contribution.

Naturally, without actually possessing $p(x)$ from a first-principles calculation, it is difficult to draw concrete conclusions about the behavior of Q or $\bar{Q}$. However, the above observations can lead us to the following sensible picture.

In the case of a metal and an infinite system, the delta function in Eq. (16) is completely suppressed due to the Anderson theorem. The emission shape is thus determined with Eq. (16) by convolving with the function $\bar{Q}$. The function $\bar{Q}$ is then expected to be a narrow but asymmetric shape, which in this case also gives the XPS line shape. The repeated convolutions of Eq. (16) indicate that $\bar{Q}(x)$ in Eq. (20) can be viewed as the solution to a multiple scattering process, in which $p(x)$ is the probability per energy of scattering once with energy loss x. Due to the increasing density of states as the volume increases, $p(x)$ becomes stronger, but also more localized near $x=0$ (for a metal), thus "multiple scattering" becomes more probable, but at smaller energies, such that, presumably, the shape of $\bar{Q}$ does not change appreciably in the limit of large volumes.

For band-gap materials, the multielectron processes do not dominate the zeroth-order emission, otherwise an added gap, equal to twice the ordinary band gap, would be observed between the absorption and emission for an insulator or semiconductor. Such an added gap is not observed. It would appear then that $\det(s)$ does not approach zero for a large system, and therefore the Anderson theorem does not apply in the case of a material with a band gap. This makes sense with the observation that the overlaps $\langle\psi\,|\,\phi\rangle$ become narrower functions of the energy with increasing volume. Once this width, for the states near the top of the band, is much less than the band gap, the occupied states of one orthonormal set approach a complete set for all the occupied states of the other orthonormal set, and probability becomes "trapped" within the matrix s. For a metal, there are always occupied states near unoccupied states that "leak" away this probability leading to the orthogonality catastrophe. In an insulator then, it would be expected that the relative strength of the shakeup to the zeroth-order emission approaches a constant for large systems, with

the zeroth-order emission dominating. Although a band-gap material can have a distinct multielectron contribution to the emission within the main band, such a separation in a metal is not easy to define without first specifying the size of the system.

The other characteristic emission of concern here is the plasmon satellite, which occurs when emission is accompanied by a plasma excitation, and thus the emitted photon energy is diminished by the corresponding plasma frequency. Therefore, on the lower-energy side of the emission, one expects to find a feature that resembles the main emission band; however, the plasmon dispersion and width broadens most of the structure. The satellite strengths are typically 1–3 % of the main band, and lie on a background produced by the multielectron transitions. The plasmon satellites are a direct consequence of electron-electron interaction and are not described with an independent electron approach. The separation of the shakeup spectrum into plasmon and multielectron contributions is of course an approximation, and will only work well if the plasmon is a well-defined elementary excitation.

To examine our data in light of the understanding presented here, we need the quantities $\langle \psi \mid \bar{\phi} \rangle$. Not having distorted and undistorted solid state wave functions readily available, we will need to model overlaps in some simple manner. We will not here make any attempt to find the states $\bar{\phi}$. Instead we will take $\langle \psi \mid \bar{\phi} \rangle = \langle \psi \mid \phi \rangle$, since for states $\bar{\phi}_n$ far from the Fermi energy, the mixture of states ϕ is expected to be rather localized about ϕ_n, thus we instead model the overlaps $\langle \psi \mid \phi \rangle$. Two such models will be used, each with different advantages. In each case we assume an integrated average over an energy surface.

V. DATA ANALYSIS

A. Atomic model

The first method uses self-consistent atomic wave functions, which seems a suitable approach for silicon and diamond. This method is free of any fitting parameters and thus provides an estimate of the strength of the emission due to shakeup. We write the distorted wave function ϕ,

$$\phi = \psi + \delta\psi \ , \tag{21}$$

where $\delta\psi$ represents the distortion of the Bloch state due to the core hole. The conduction state is taken to be a plane wave whose kinetic energy is measured from the bottom of the conduction band. Since the valence state ψ and the conduction states are assumed to be elements from the same orthonormal set, their scalar product is zero. Therefore the overlap we seek is the Fourier transform

$$\langle \mathbf{k} \mid \delta\psi \rangle = \frac{1}{\Omega^{1/2}} \int \delta\psi(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}) d\mathbf{r} \tag{22}$$

of the distortion. Here, Ω is the volume and $\mathbf{k}$ is the wave vector of the conduction state.

Since it is assumed that the core hole is localized at one site, one expects the distortion to be appreciable in this region only. With atomic wave functions, we presumably obtain a reasonable distortion near the vacant site, but deviate from it as the nearest neighbor is approached. Hopefully this error is tolerable, in particular since it is a difference we seek here rather than the correct wave functions themselves. In the case of silicon for instance, we represent the valence state with an sp^3 hybrid of $3s$ and $3p$ atomic functions whose s and p components vary throughout the band, and whose binding energy is also varied. The s and p components are determined from the difference in L and K emission, using the usual dipole selection rules for those emission bands. Their relative strengths are chosen to be consistent with band-structure calculations.[30] The binding energy is varied through a radial scaling linear with the valence-band energy (with respect to the vacuum level), which simulates the radial variation throughout the band that perhaps the inclusion of other atomic states in a linear combination of atomic orbitals approach would induce. It is the difference in these atomic functions, with and without a core vacancy, that are used to model the distortion. Thus we use

$$\langle \psi(E') \mid \phi(E) \rangle = \int [a(E)\delta\psi_s(\mathbf{r},E) + b(E)\delta\psi_p(\mathbf{r},E)] \exp(i\mathbf{k}\cdot\mathbf{r}) d\mathbf{r} \ , \tag{23}$$

where $\delta\psi_s$ and $\delta\psi_p$ are the differences in the atomic s and p functions, respectively, $a(E)$ with $b(E)$ satisfy the normalization condition $a^2+b^2=1$, E is the corresponding valence state energy and E' is the corresponding conduction state energy [$E'=(\hbar^2k^2/2m)$ measured from the bottom of the conduction band]. The self-consistent undistorted atomic functions for diamond and silicon are calculated[31] with the valence electrons in the configuration s^1p^3; the distorted functions are calculated with the appropriate core vacancy, and an occupied s state in the next highest shell to simulate screening. Since we use normalized atomic wave functions, the density of states is taken to be that appropriate to an atomic volume. Given these overlaps we can construct the emission due to the multielectron transitions. Equation (15) with Eq. (20) determine $\bar{Q}$, which, with Eq. (16) and some $A(\nu)$, give that part the emission shape due to multielectron transitions. For $\nu A(\nu)$ we first use the data itself. The calculated shakeup is then subtracted from the data and the result reiterated for $\nu A(\nu)$ until adequate convergence is achieved (usually two or three iterations).

The results of this approach for diamond and silicon are pictured in Figs. 1 and 2. In diamond the calculation depicts a sensible shake-up tail upon which, as we shall demonstrate in II, a plasmon satellite rests. What is obtained from this method is an estimate of the strength and shape devoid of any free parameters, which agrees reasonably with our data, reassuring that the effect we are describing is responsible for the observed strength.

The emission from silicon shows a striking amount of strength in regions below what would be reasonably called the band minimum, and in fact prompted this investigation. A feature appears near 79 eV, which has previously been identified as a plasmon satellite.[32,33] For silicon, not only does this atomic method produce a

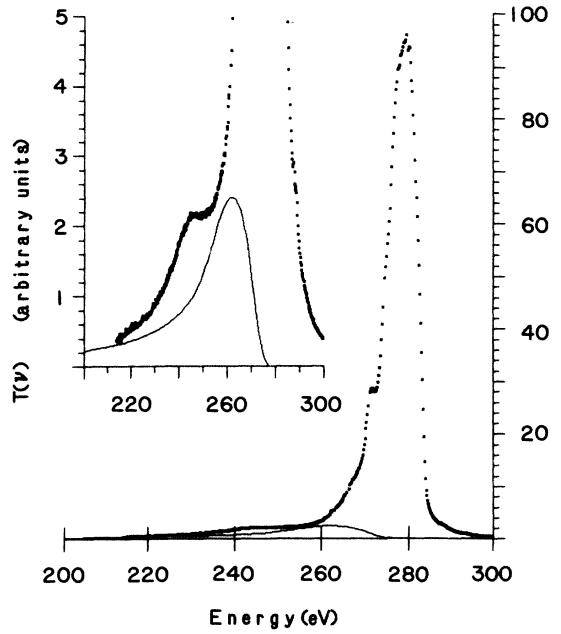

FIG. 1. Soft x-ray K emission data for diamond (dots). The solid line is the multielectron contribution to the emission calculated with atomic self-consistent wave functions. The inset is an enlargement in the region of the low-energy tail.

much stronger multielectron contribution than in diamond, consistent with the data, but it also indicates a maximum which lies several eV lower than the band minimum ($\approx$86.5 eV). In diamond the modeled maximum lies close to the band minimum, if not underneath the zeroth-order emission. This maximum in silicon, although not matching the feature at 79 eV exactly, is only 3 eV from it, suggesting that possibly this feature is not due to a plasmon satellite, but instead structure from the multielectron contribution, an interpretation we shall comment on later.

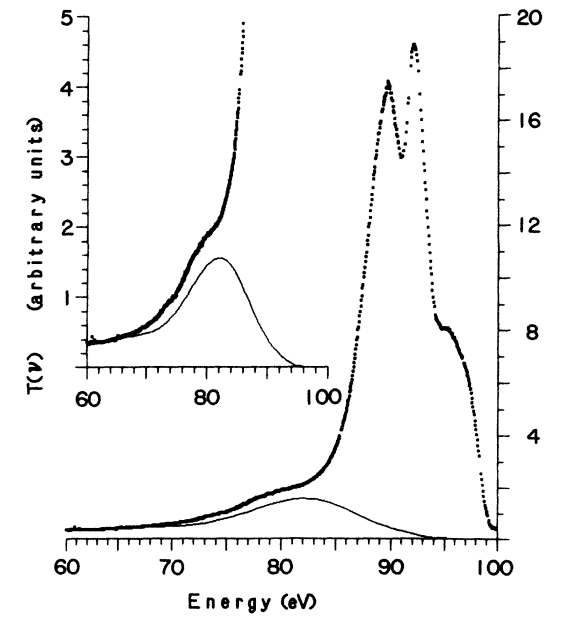

FIG. 2. Soft x-ray L emission data for silicon (dots). The solid line is the multielectron contribution to the emission calculated with atomic self-consistent wave functions. The inset is an enlargement in the region of the low-energy tail.

B. Fitting procedure

To better determine these plasmon satellite backgrounds, and as an alternative general approach to modeling the low energy tails, we employ a second method based on a fitting procedure near the band minimum. Noting the energy denominator of first-order perturbation theory, we will assume that $|\langle\psi(E')|\phi(E)\rangle|^2$ varies in a Lorentzian manner as a function of $E-E'$;

$$|\langle\psi(E')|\phi(E)\rangle|^2=\frac{K(E)}{2\pi}\frac{\Gamma}{(E-E')^2+(\Gamma/2)^2}\ , \tag{24}$$

where $K(E)$ is a normalization factor such that

$$\int|\langle\psi(E')|\phi(E)\rangle|^2N(E')dE'=1\ . \tag{25}$$

E' and E are the energies corresponding to the unoccupied conduction state and occupied valence state respectively, and Γ will be referred to as the mixing width. Given this, one can construct the total emission shape. Again Eq. (15) with Eq. (20) determine $\bar{Q}$, which, with Eq. (16) and some $\nu A(\nu)$, give that part the emission shape due to multielectron transitions. This then added to $\nu A(\nu)$ would correspond to the total emission.

Our procedure will be to fit the data in the region corresponding to the low-energy part of the zeroth-order emission (the bottom of the band). There, it is expected[1] that $\nu A(\nu)$ follows $\nu^3(\nu-\nu_0)^{n/2}$, where $n=1$ for a p core and $n=3$ for an s core. ν_0 corresponds to the energy at the band minimum. Furthermore, this power-law shape will also be lifetime broadened by an energy-dependent width.[34,35] We assume this width $\gamma(\nu)$ to have the form

$$\gamma(\nu)=\beta(\nu_F-\nu)^2\ . \tag{26}$$

Our model for the shape of $\nu A(\nu)$ near the bottom of the band requires the three parameters ν_0, β, and an overall strength.

For each material considered, the fitting regions are chosen over an extent where the power law is expected to be valid, and removed from possible interference from any plasmon satellite. The total emission at the band bottom is then taken to be a sum of the broadened power law and the multielectron contribution. This multielectron contribution is constructed with the above mentioned procedure, where we first use the data throughout the main band to represent the zeroth-order emission $\nu A(\nu)$ in Eq. (16). Here Γ will be the only parameter, thus a four-parameter fit to the bottom of the band is attempted. Once a minimum in the chi-square value is obtained, the multielectron contribution is subtracted from the first approximation for the zeroth-order emission (the data) and again the procedure is reiterated until adequate convergence is achieved.

As already mentioned, for a metal the relative strength of our power-law model to the multielectron contributions is in principle a function of the size of the system. This factor is controlled by the strength of the density of states and the mixing width Γ. Large volumes begin to

make the numerical work more difficult, since then $p(x)$ becomes a much narrower but stronger function, and Eq. (20) requires many more iterations. We again choose the density of states appropriate to an atomic volume, and let the mixing width construct various functions $p(x)$ such that the fitting procedure picks out the best one.

The results of this analysis for aluminum are pictured in Figs. 3 and 4. The band minimum ν_0 was found to be at 62.03±0.1 eV, β is 0.019±0.001 eV^{-1}, the mixing width $\Gamma=1.06$ eV, and $E_F=72.7\pm0.05$ eV, thus we report a bandwidth of 10.67±0.15 eV, which is in good agreement with the value 10.6 eV reported by Levinson *et al.*[35] using angle-resolved photoemission. The observed width is about 10% smaller than the noninteracting electron value of 11.7 eV. This trend in the simple metals has recently been discussed with regard to sodium,[36] and has been attributed to the energy dependence of the electron self-energy.[37,15,38] A useful byproduct of this fitting procedure is the background upon which the plasmon satellite rests. We emphasize that the fitting was done only in the region above the plasmon satellite threshold, in this case from 59 to 64 eV; to our satisfaction, a sensible tail below the satellite was achieved.

Results for graphite, diamond and silicon are pictured in Figs. 5, 6, and 7, respectively. The graphite emission was measured at an angle of 50° with respect to the c axis. For graphite, fitting in the region 256–265 eV, we obtain $\nu_0=263.2\pm0.5$ eV, $\beta=0.012\pm0.001$ eV^{-1}, and $\Gamma=1.21$ eV. With a Fermi energy of 284.7±0.1 eV, the bandwidth is 21.5±0.6 eV. This value falls in between 22.5 eV measured with angle-integrated photoemission[39] and 20.6 eV measured with angle-resolved photoemission,[40] and can also be compared with 20.7 eV from band-structure calculations.[41] With graphite, as well as diamond, we obtain a sensible background for the low-energy features which we shall identify in II to be plasmon satellites. Our diamond results were fit between 258 and 265 eV, giving $\nu_0=258.3\pm0.3$ eV, $\beta=0.016\pm0.001$ eV^{-1}, and $\Gamma=2.15$ eV. With the top of the band at 283.9±0.2 eV, we obtain a bandwidth of 25.6±0.5 eV. Although the lifetime

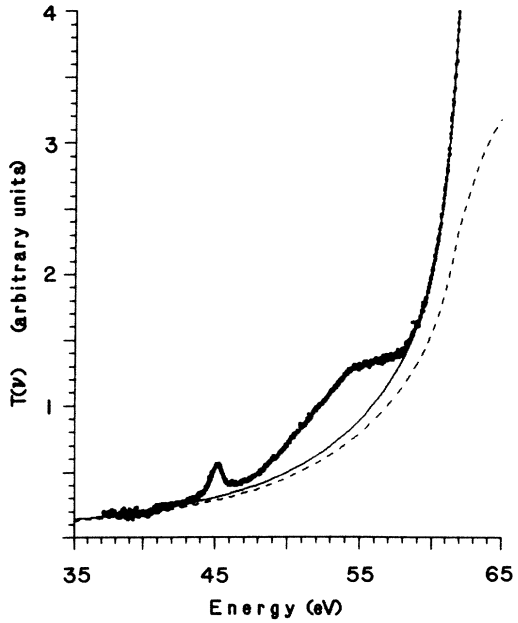

FIG. 4. Soft x-ray L emission for aluminum in the region of the plasmon satellite. The solid line is the sum of the fitting procedure. The dashed line gives the corresponding multielectron contribution.

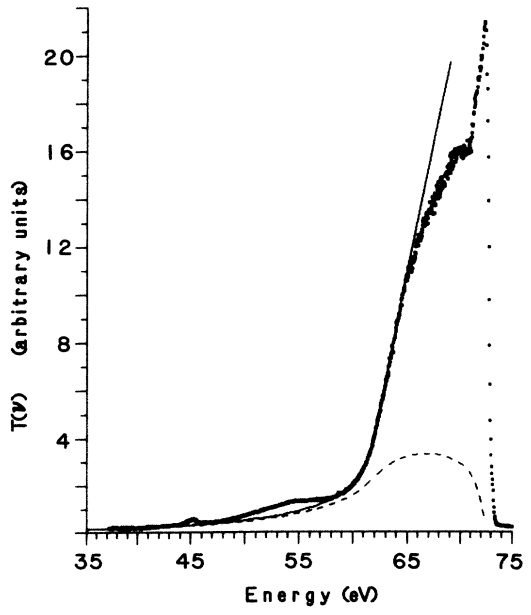

FIG. 3. Soft x-ray L emission data for aluminum (dots). The solid line is the sum of the fitting procedure. The dashed line gives the corresponding multielectron contribution.

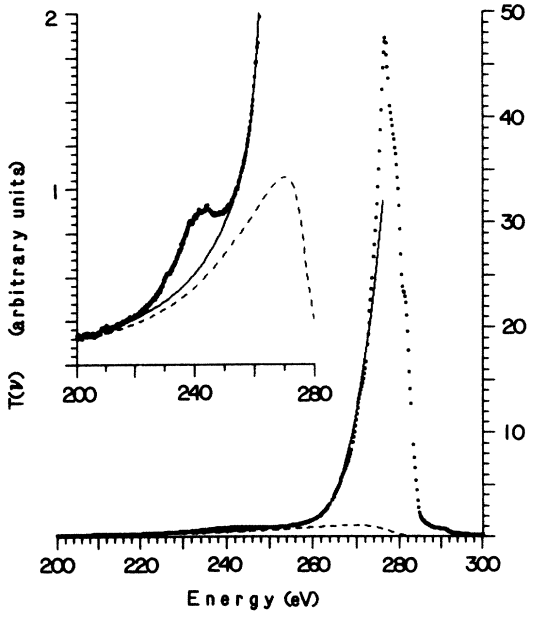

FIG. 5. Soft x-ray K emission data for graphite (dots). The solid line is the sum of the fitting procedure. The dashed line gives the corresponding multielectron contribution. The inset is an enlargement in the region of the low-energy tail.

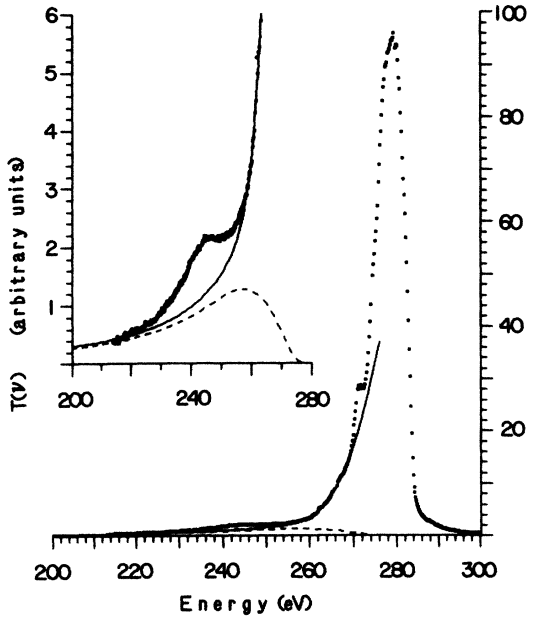

FIG. 6. Soft x-ray K emission data for diamond (dots). The solid line is the sum of the fitting procedure. The dashed line gives the corresponding multielectron contribution. The inset is an enlargement in the region of the low-energy tail.

broadening appears to be unrealistically large at the band minimum (≈ 10 eV), the bandwidth is in good agreement with 25.2 eV obtained in a recent calculation[42] that treats the dynamical aspects of correlation within a first-principles Green's function method.

In the case of silicon our fitting procedure, although depicting the general trend in the low-energy tail, did not give a satisfactory background for the purported plasmon satellite feature. We fit in the region 84–87 eV. The resulting parameters are $\nu_0=86.6\pm0.1$ eV, $\beta=0.022\pm0.001$ eV^{-1}, with a mixing width $\Gamma=2.19$ eV. With the top of the band at 98.8 ± 0.1 eV, we report a bandwidth of 12.2 ± 0.2 eV for silicon. The maximum of the shakeup is slightly higher in energy than in the atomic method, but still well below the band minimum. To further investigate whether the multielectron shakeup can be responsible for the structure at 79-eV, we consider the question of how well this feature can be attributed to a plasmon satellite.

C. Silicon low-energy feature

In II we describe a plasmon satellite model which works fairly well in the classic test case of aluminum. It satisfactorily determines the strength and relative weights to assign the plasmons of different wave vectors. The satellite shape and location is constructed with the appropriate convolutions with the main band shape. The result for silicon is shown in Fig. 8. Electron energy-loss measurements[43] put the silicon plasmon energy for the wave vector $\mathbf{k}=0$ at 17 eV; it disperses to about 20 eV at cutoff. It should be pointed out that the location of the satellite is determined mainly by the plasmon energy and its dispersion, and will not be greatly affected by the relative strengths that our model might assign each wave vector. Satellite strength corresponding to plasmon energies greater than the value for $\mathbf{k}=0$ is possible, but not corresponding to lower energies. It should then be clear from

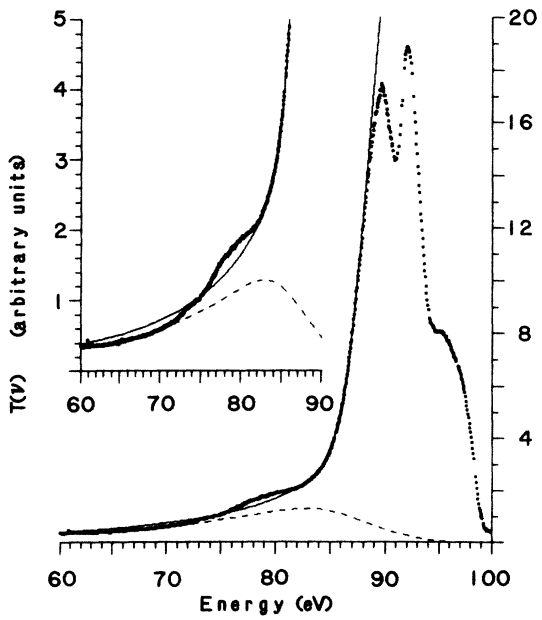

FIG. 7. Soft x-ray L emission data for silicon (dots). The solid line is the sum of the fitting procedure. The dashed line gives the corresponding multielectron contribution. The inset is an enlargement in the region of the low-energy tail.

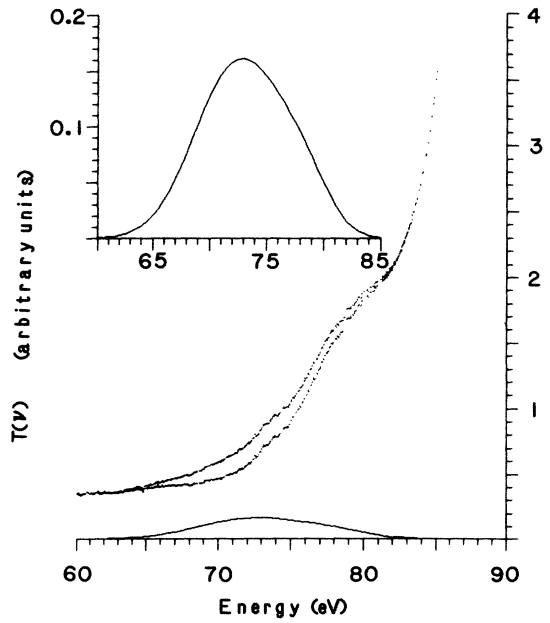

FIG. 8. The low-energy tail in silicon, where the upper set of points is the data, and the lower set of points is the data with the modeled silicon plasmon satellite subtracted. The solid line below is the modeled silicon plasmon satellite as determined by methods described in II. The inset is an enlargement of the modeled silicon plasmon satellite.

Fig. 8 that the location of the feature at 79 eV is too high by about 6 eV to be identified with a plasmon satellite. Even when dispersion is completely ignored, the modeled satellite peak falls below the observed feature by about 4 eV. Depending on how a satellite background is drawn, the feature is either too narrow or of the wrong shape to resemble the main emission band. Also, the possibility of a surface plasmon is insignificant, since emission mostly originates deep within the bulk.

Other possibilities that might explain this discrepancy would be a bound plasmon, or plasmon coupling to the upper part of the band which is much stronger than to the lower part. Sziklas[44] has suggested that bound plasmons might exist, but only on a negative impurity. In any event, the final state has no impurity whatsoever, and we find no reason for the large difference in valence band coupling needed to produce the feature observed. Also, emission corresponding to the p density of states would be quite unexpected, due to the selection rules. It would seem that a plasmon satellite in this case is a poor candidate for identification with the 79-eV feature, but an acceptable alternative is the structure in the emission due to the multielectron transitions. We view the 3-eV error of the model calculation as attributable to the simplicity of our approximations and not significant, while the 6-eV error in the plasmon satellite model is viewed as a significant inconsistent discrepancy. We should point out that a completely different explanation for the unusual strength on the low-energy side of silicon was offered by Lyapin[45] two decades ago. There, parts of the emission spectra in the low-energy tail for Si and Ge were assigned to deep "additional" levels in the valence band.

Of course, silicon does have a well-defined plasmon, as electron energy-loss measurements demonstrate. If the low-energy feature is attributed to multielectron transitions, one then might ask, where is the plasmon satellite in the emission for silicon? Figure 8 shows the silicon data in the low-energy region with and without the modeled satellite subtracted. No unusual structure appears to be introduced due to this process, which suggests that the satellite could very well in contributing to the emission, but that it sits in the "saddle" of structure due to the multielectron transitions, making it difficult to discern.

A common experimental approach to obtain the single-particle zeroth-order emission has been to subtract a linear background. In view of the present and recent[46,47] work, we have found that a simple improvement is to fit a Lorentzian to the low-energy tail. Even better, some representation of the zeroth-order emission near the band minimum should also be included in the fit. Such Lorentzian backgrounds typically have strengths that are 10–15 % of the main band, and usually have widths of the order of the main band width. Any estimate of the bremsstrahlung would be subtracted first.

VI. SUMMARY

Through a modification of a transformation suggested by Stern and Rehr, we have arrived at a single-particle formalism to describe soft-x-ray emission and absorption. This transformation provides a general explanation of the final-state rule, showing it to be a direct result of exchange. The formalism is further applied to the multielectron processes also accompanying the soft-x-ray transition, and within an assumption of random phases accounts for such processes to infinite order. Here we have modeled the required overlaps, and applied these methods in the analysis of emission data, with encouraging results.

The previous identification of a plasmon satellite in silicon is shown to be inconsistent with electron energy loss data. We instead attribute the feature to the emission due to multielectron processes which produce structure at energies below the silicon band minimum.

ACKNOWLEDGMENTS

We would like to thank P. Bruhwiler, R. Carson, D. Husk, A. Mansour, J. Nithianandam, C. Tarrio, and S. Velasquez, for helpful suggestions and discussions throughout the course of this work. We also are grateful for the earlier technical contributions of Drs. T. Aton, F. Zutavern, A. Cafolla, and C. Franck. We thank Dr. H. P. Kelly, and M. Kutzner for access to the self-consistent atomic calculations. This research was supported in part by the National Science Foundation (Grant No. DMR 85-15684).

[1]D. H. Tomboulian, *Handbuch der Physik* (Springer-Verlag, Berlin, 1958), Vol. 30.
[2]P. T. Landsberg, Proc. Phys. Soc. London, Sect. A **62**, 806 (1949).
[3]J. Pirenne and P. Longe, Physica **30**, 277 (1964).
[4]R. A. Ferrell, Rev. Mod. Phys. **28**, 308 (1956).
[5]G. A. Rooke, Phys. Lett. **3**, 234 (1963).
[6]P. Longe and A. J. Glick, Phys. Rev. **117**, 526 (1969).
[7]B. Bergersen, F. Brouers, and P. Longe, J. Phys. F **1**, 945 (1971).
[8]S. M. Bose and A. J. Glick, Phys. Rev. B **10**, 2733 (1974).
[9]Peteris Livins and S. E. Schnatterly, following paper, Phys. Rev. B **37**, 6742 (1987).
[10]T. Aberg and J. Utriainen, Phys. Rev. Lett. **22**, 1346 (1969).
[11]T. Aberg and J. Utriainen, Solid State Commun. **16**, 571 (1975).
[12]T. Aton, C. Franck, E. Kallne, S. Schnatterly, and F. Zutavern, Nucl. Instrum. Methods **172**, 173 (1980).
[13]F. Zutavern, S. Schnatterly, E. Kallne, C. Franck, and T. Aton, Nucl. Instrum. Methods **172**, 351 (1980).
[14]R. D. Carson, C. Franck, S. Schnatterly, F. Zutavern, Rev. Sci. Instrum. **55**, 1973 (1984).
[15]L. Hedin and S. Lundqvist, *Solid State Physics,* edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, London, 1969), Vol. 23.
[16]G. D. Mahan, Phys. Rev. **163**, 612 (1967).

[17]P. Nozieres and C. T. De Dominicis, Phys. Rev. **178**, 1097 (1968).
[18]M. Combescot and P. Nozieres, J. Phys. (Paris) **32**, 913 (1971).
[19]J. Friedel, Comments Solid State Phys. **2**, 21 (1969).
[20]L. C. Davis and L. A. Feldkamp, Phys. Rev. B **23**, 4269 (1981).
[21]E. A. Stern and J. J. Rehr, Phys. Rev. B **27**, 3351 (1983).
[22]U. von Barth and G. Grossman, Solid State Commun. **32**, 645 (1979).
[23]G. D. Mahan, Phys. Rev. B **21**, 1421 (1980).
[24]E. T. Arakawa and M. W. Williams, Phys. Rev. Lett. **36**, 333 (1976).
[25]R. D. Carson and S. E. Schnatterly, Phys. Rev. Lett. **59**, 319 (1987).
[26]T. A. Green, Phys. Rev. B **32**, 3442 (1985).
[27]U. von Barth and G. Grossman, Phys. Rev. B **25**, 5150 (1982).
[28]C. A. Swarts, J. D. Dow, and C. P. Flynn, Phys. Rev. Lett. **43**, 158 (1979).
[29]P. W. Anderson, Phys. Rev. Lett. **18**, 1049 (1967).
[30]D. A. Papconstantopoulos, *The Band Structure of Elemental Solids* (Plenum, New York, 1986).
[31]C. Froese-Fischer, Comp. Phys. Commun. **14**, 145 (1978).
[32]O. Aita and T. Sagawa, J. Phys. Soc. Jpn. **27**, 164 (1969).
[33]V. A. Fomichev and T. A. Zimkina, Fiz. Tverd. Tela (Leningrad) **9**, 1833 (1967) [Sov. Phys.—Solid State **9**, 1441 (1967)].
[34]G. A. Rooke, J. Phys. C **1**, 767 (1968).
[35]H. J. Levinson, F. Greuter and E. W. Plummer, Phys. Rev. B **27**, 727 (1983).
[36]E. Jensen and E. W. Plummer, Phys. Rev. Lett. **55**, 1912 (1985).
[37]B. I. Lundqvist, Phys. Kondens. Mater. **6**, 206 (1967).
[38]J. E. Northrup, M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **59**, 819 (1987).
[39]A. Bianconi, S. B. M. Hagstrom and R. Z. Bachrach, Phys. Rev. B **16**, 5543 (1977).
[40]I. T. McGovern, W. Eberhardt, E. W. Plummer, and J. E. Fisher, Physica B + C **99B**, 415 (1980).
[41]R. F. Willis, B. Fitton, and G. S. Painter, Phys. Rev. B **9**, 1926 (1974).
[42]G. Strinati, H. J. Mattaucsh, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).
[43]H. Raether, *Excitation of Plasmons and Interband Transitions by Electrons* (Springer-Verlag, Berlin, 1980).
[44]E. A. Sziklas, Phys. Rev. **138**, A1070 (1965).
[45]V. G. Lyapin, Fiz. Tverd. Tela (Leningrad) **8**, 3567 (1967) [Sov. Phys.—Solid State **8**, 2851 (1967)].
[46]A. Mansour and S. E. Schnatterly, Phys. Rev. Lett. **58**, 614 (1987).
[47]P. Bruhwiler and S. E. Schnatterly (unpublished).