PHYSICAL REVIEW B VOLUME 18, NUMBER 10 15 NOVEMBER, 1978

Coulombic resonance states in zero-gap semiconductors

B. Joós, A. K. Das, and P. R. Wallace
Physics Department, McGill University, 3600 University, Montréal, Québec H3A 2T8, Canada

(Received 16 May 1978)

We present a detailed theory of resonance states in HgTe due to charged acceptor impurities. This problem has previously been studied by Gel'mont and Dyakonov, but on the rather unrealistic assumptions of an unscreened Coulomb potential and zero electron mass. In a zero-gap (ZG) semiconductor, the screening is important and sensitive to the band parameters. Within the effective-mass formalism, we have obtained a variational solution for the lowest-order radial components in the partial-wave expansion of the wave function. Two variational methods for the phase shift are used and improved upon—the Kohn method and the Harris method, which is an interesting modification of the former and is more appropriate for the study of resonances. A simple Thomas-Fermi (TF) potential and the full random-phase approximation screened Coulomb potential have been used. We find that the resonance energy E_R is very sensitive to screening, while the relative width Γ/E_R is only weakly dependent upon it. The interband polarization, which is known to be substantial in a ZG semiconductor, dramatically reduces the resonance energy. Interestingly, within the reported range of band parameters, either the 0.7-meV or the 2.2-meV resonance can be reproduced, corresponding respectively to hole masses of $0.35m_0$ or $0.58m_0$. We also find that the resonance energy varies by at most 10% with the Fermi level E_F in the range of interest, $E_F < E_R$.

I. INTRODUCTION

Due to the absence of a gap between the conduction and the valence bands the impurity levels in zero-gap (ZG) semiconductors are degenerate with the continuum. If the interaction between a level and the continuum is strong, the impurity state will not be well defined. But if it is weak a quasilocalized or resonance state can exist, i.e., a bound state with a finite lifetime τ, or equivalently by the uncertainty principle, a finite energy width $\Gamma = \hbar/\tau$. Such a state will dramatically perturb the electrical and optical properties of the system. For example, in optical-absorption measurements an extra peak will be observed. And whenever the Fermi level will be close to the impurity state the resonant scattering of the carriers off the impurities will produce a sharp decrease in the conductivity. In HgTe, a well-known ZG semiconductor, three acceptor resonance states are experimentally known—A_0 at 0.7 meV, A_1 at 2.2 meV, and A_2 at 9 meV.[1] A_0 and A_1 have been observed in both the transport and the optical measurements but A_2 has so far been observed only in transport measurements. The first one has been attributed to charged impurities such as Cu while the other two are believed to be due to stoichiometric defects. No donor resonance states are known nor are, in fact, expected to be seen. In ZG semiconductors of the type HgTe, the mass difference between the valence and the conduction bands is large ($m_v^*/m_c^* \geq 10$). The correspondingly low density of states in the conduction band will be less effective in the broadening of the acceptor level whereas the donor levels due to the large density of states in the valence band will be highly damped.

Of the different acceptor resonances perhaps the most interesting one is the charged impurity resonance because it is this state which depends most markedly on the special features of the ZG semiconductor. It may be remarked here that the absence of a gap in semiconductors such as HgTe is not due to the accidental overlap of the conduction and the valence bands as in semimetals but is symmetry induced. The strong relativistic effects which exist in the heavy compounds such as HgTe produce an unusual ordering of the bands.[1] Contrary to ordinary semiconductors of the same crystal structure, the Γ_8 levels ($P_{3/2}$) lie above the Γ_6 level ($S_{1/2}$)—the so-called inverted band structure (see Fig. 1). Both the electron and hole states are therefore of the same symmetry, i.e., of p states with effective spin $J=\frac{3}{2}$ but different helicities. The strong coupling between the conduction and valence bands which follows, will affect all phenomena extending over several primitive cells. It will, for instance, introduce an interesting contribution to the screening of a charged impurity usually not present in ordinary semiconductors. The effects of this screening on the charged-impurity resonance states will be important and should be considered in any realistic theoretical model.

We present here the first detailed study of these resonances. Our model consists of the one-particle effective-mass Hamiltonian incorporating the many-body effects in the dielectric function. The radial equations for the ground-state resonance were solved using a method not com-

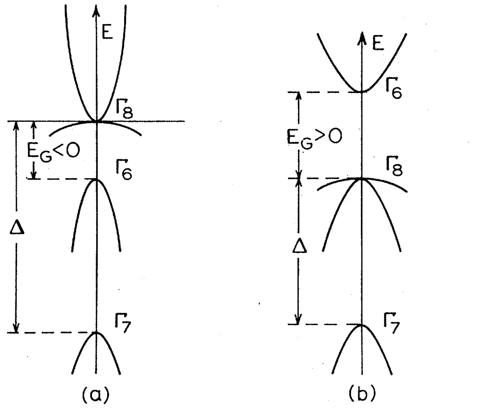

FIG. 1. Band structure at the center of the Brillouin zone of (a) zero-gap semiconductors of the type HgTe and of (b) ordinary semiconductors of the diamond and zinc-blende structure (e.g., CdTe).

monly used in this type of problems—the variational principle for the phase shift. The main result of our study is that either A_0 or A_1 may be the charged-impurity resonance depending upon whether the actual hole mass is closer to $0.35m_0$ or $0.58m_0$.

In Sec. II, we review the different theoretical works on resonances in ZG semiconductors, and briefly introduce our model. Section III describes the variational method we have used and improved upon. Section IV presents our results for a phenomenological potential and the RPA screened potential. We end with a discussion of our work and a comparison with reported experimental results.

II. THEORETICAL MODELS

Several theoretical studies have been made on the impurity states in ZG semiconductors. Liu and Brust[2] proposed a two-band model that treats the degeneracy as accidental. In this case a "smooth" potential will not couple the bands because of the distinct symmetries of their wave functions and only the central part of the potential will be responsible for the interaction of the localized levels with the continuum. States extending over several lattice sites will remain sharp. Mauger and Friedel[3] have followed up this model in the case of a potential entirely localized on the site of the impurity or defect. In our view, the model does not take into account the proper symmetry of the problem. In HgTe the degeneracy is, as discussed in the Introduction, not accidental but symmetry induced. The Liu-Brust model is a limiting case where one assumes that the interaction does not affect the structure of the wavefunction away from the central cell. Though not a realistic approximation, it enables a simple study of vacancy-type defects and yields some insight into the nature and the properties of the resonances they produce.

For a potential extending over several lattice sites a more general formalism is necessary. The most satisfactory model is the effective-mass Hamiltonian[4] where the kinematics of the electrons and holes are represented by the $\vec{k}\cdot\vec{p}$ perturbation expansion about the Γ point. In the spherical approximation, i.e., if the warping of the bands is neglected, this Hamiltonian can be written

$$H=(\gamma_1/2m_0)[(1+\tfrac{5}{4}\mu)\vec{p}^2-\mu(\vec{p}\cdot\vec{J})^2]+V(\vec{r}),$$

where m_0 is the electron mass, $\vec{p}$ the momentum operator, and $\vec{J}$ is the 4×4 angular-momentum matrix for a state $J=\frac{3}{2}$. γ_1 and μ are band parameters. In all subsequent expressions we shall use for simplicity effective Rydberg units which make $\hbar^2\gamma_1/2m_0=1$. The units of energy and length have been chosen to be, respectively,

$$R_0^{\gamma_1}=(e^4m_0/2\hbar^2)(1/\epsilon_0^2\gamma_1)\,, \tag{2.1a}$$

$$a_0^{\gamma_1}=(\hbar^2/e^2m_0)\epsilon_0\gamma_1\,, \tag{2.1b}$$

where ϵ_0 is the lattice dielectric constant. The Hamiltonian then reduces to

$$H=(1+\tfrac{5}{4}\mu)k^2-\mu(\vec{k}\cdot\vec{J})^2+V(\vec{r})\,. \tag{2.2}$$

(Note that in these units the Hamiltonian contains only one physical parameter, namely μ.) In the absence of a perturbing potential V, the eigenvalues of H are given by the relation

$$E_\pm(k)=(1\pm\mu)k^2 \quad [(+),\text{ electrons; }(-),\text{ holes}]. \tag{2.3}$$

μ should be greater than one so as to represent both the conduction and the valence bands. The electron mass m_c^* and the hole mass m_v^* are given by

$$m_c^*=1/\gamma_1(1+\mu),\quad m_v^*=-1/\gamma_1(1-\mu)\,, \tag{2.4}$$

$$\beta=m_c^*/m_v^*$$
$$=(\mu-1)/(\mu+1)\,.$$

Bastard and Nozières[5] have solved this Hamiltonian for a restricted sort of potential, namely, a separable potential in momentum space. (Though elegant, their calculation is limited by the fact that it is not clear to what type of defect their model is applicable.) Gel'mont and Dyakonov[6] have considered the same Hamiltonian for a Coulomb potential. When $V(r)$ is spherically symmetric Eq. (2.2) can be considerably simplified. The "effective total angular momentum" $\vec{F}=\vec{L}+\vec{J}$, where $\vec{L}=(1/\hbar)(\vec{r}\times\vec{p})$ now commutes with Eq. (2.2).

Any solution ψ of (2.2) can therefore be written as a linear combination of eigenstates of $\vec{F}$ where each energy level will be degenerate with respect to the projection M of $\vec{F}$ along a fixed direction

$$\psi=\sum_{F,M} c_{FM}\psi_{FM}\,,$$

$$\psi_{FM}=\sum_{l} R_{F,l}(r)\,|l,J=\tfrac{3}{2},F,M\rangle\,, \qquad (2.5)$$

$$F-\tfrac{3}{2}\leq l\leq F+\tfrac{3}{2};$$

the $|l,J=\frac{3}{2},F,M\rangle$, simultaneous eigenstates of $\vec{F}$, $\vec{L}$, and $\vec{J}$, are determined by the rules of angular-momentum addition:

$$|l,J=\tfrac{3}{2},F,M\rangle=(2F+1)^{1/2}\sum_{m,\mu}(-1)^{3/2-l+M}\times\begin{pmatrix}\frac{3}{2} & l & F\\ \mu & m & -M\end{pmatrix}\times|\tfrac{3}{2},\mu\rangle|l,m\rangle\,. \qquad (2.6)$$

$|l,m\rangle$ are the familiar spherical harmonics $Y_{lm}(\theta,\phi)$ and

$$\begin{pmatrix}\frac{3}{2} & l & F\\ \mu & m & -M\end{pmatrix}$$

are the Wigner $3j$ symbols. Inserting the above expression for ψ_{FM} into Eq. (2.2) and using the orthogonality properties of the angular part of the wave function, radial equations have been derived for these eigenstates.[6,7] The lowest quasi-localized state is expected to be the solution having a dominant zero-angular-momentum radial component. This state will be an $F=\frac{3}{2}$ state ($S_{3/2}$) of the form

$$\psi_{3/2,M}=[u_0(r)/r]\,|l=0,J=\tfrac{3}{2},F=\tfrac{3}{2},M\rangle+[u_2(r)/r]\,|l=2,J=\tfrac{3}{2},F=\tfrac{3}{2},M\rangle\,, \qquad (2.7)$$

where u_0 and u_2, coupled by symmetry, are solutions of the system of Eqs. (2.8).

$$\begin{pmatrix}\dfrac{d^2}{dr^2}+E-V(r) & \mu\left(\dfrac{d^2}{dr^2}+\dfrac{3}{r}\dfrac{d}{dr}\right)\\ \mu\left(\dfrac{d^2}{dr^2}-\dfrac{3}{r}\dfrac{d}{dr}+\dfrac{3}{r^2}\right) & \dfrac{d^2}{dr^2}-\dfrac{6}{r^2}+E-V(r)\end{pmatrix}\times\begin{pmatrix}u_0\\ u_2\end{pmatrix}=0\,. \qquad (2.8)$$

Gel'mont and Dyakonov[6] have argued that in the limit $m_c^*\to 0$ with m_v^* finite, the study of the acceptor resonances of Eq. (2.8) reduces to a bound-state problem and that for a Coulomb potential the resonance energy is proportional to the valence-band mass m_v^*. It is to be noted that the above limit cannot be taken in (2.8) since it would correspond to $\gamma_1\to\infty$ or $R_0^{\gamma_1}\to 0$. Gel'mont and Dyakonov have derived a system of equations for this case. Solving numerically their complex differential equations they find that for a Coulomb potential

$$E_R=\tfrac{4}{9}(m_v^*/m_0\epsilon_0^2)\ \text{Ry}\quad\text{or}\quad 15.11(m_v^*/m_0)\ \text{meV} \qquad (2.9)$$

(if $\epsilon_0=20$ as in HgTe). They have also obtained an expression for the relative width Γ/E_R which has the same dependence on β as that derived by Bastard and Nozières for a seemingly different potential. In both cases

$$\Gamma/E_R\propto\beta^{3/2}\,. \qquad (2.10)$$

The $\frac{3}{2}$ factor is not surprising for it indicates that in the leading order Γ/E_R is proportional to the relative density of continuum states in the conduction band. Since $\beta\ll 1$ in ZG semiconductors, Γ/E_R will most likely be very small and sharp acceptor resonances are expected. Similarly, for donor levels the relative width would be $\Gamma/E_R\propto 1/\beta^{3/2}\gg 1$ and the levels are entirely delocalized.

Our model consists of the effective-mass Hamiltonian (2.2) but a highlight of our calculation is the use of the full RPA screened-Coulomb potential. As already stated the screening effects in a ZG semiconductor are important. A strong interband polarization between the two Γ_8 bands is present due to the symmetry-induced degeneracy, and unlike in ordinary semiconductors, the absence of an ionization energy for the impurity states makes the intraband polarization due to the conduction band significant (see Sec. IV B). In a spherical approximation of the electron and hole kinematics these terms in the screening function are also spherical. The partial-wave expansion of the Hamiltonian (2.2) is therefore still possible and the ground-state resonance will be a solution of Eq. (2.8). The set of equations (2.8) with a fully screened $V(r)$ is analytically intractable. Solutions with even the simplest Coulomb potential are not known. The problem, however, lends itself well to a scattering approach. We have looked for the resonance scattering of electron states, generalizing the variational principle for the phase shift. In our approach no assumption needs to be made on the size of β and a realistic potential can be considered. The method is briefly explained in Sec. III.

III. VARIATIONAL METHOD

Since resonance is a special case of scattering phenomenon, we have applied a phase-shift analysis to solve Eqs. (2.8). As is well known in the case of a resonance, the phase shift rapidly passes and increases through $\frac{1}{2}\pi$ (modulo π), when plotted as a function of energy or momentum. With the complicated pair of Eqs. (2.8) an analytical calculation of the phase shift looks extremely difficult. One can of course integrate Eqs. (2.8) completely numerically and then obtain the phase shift. But we feel that such an involved numerical analysis is costly and will not properly display the physical subtleties that are characteristic of a resonance. Instead we have applied variational methods to calculate the phase shift. We refer the reader to some standard sources[8,9] for a detailed account of variational methods in scattering theory. In the process of our calculation we have felt that we have sufficiently explored the algorithm of the two variational methods we have used and improved upon and that our findings may be of some general interest. We have therefore decided to publish these specific results separately.[10]

In this section we shall briefly outline the essential features of these methods—the Kohn method[11] and the Harris method[12]—which is an interesting modification of the former. The variational principle in scattering can be formulated in analogy with the bound-state problem. There is however an important difference, namely, that the wave functions now do *not* decay asymptotically but behave as

$$\begin{pmatrix} u_0 \\ u_2 \end{pmatrix} \xrightarrow[r\to 0]{} 0 \tag{3.1a}$$

$$\begin{pmatrix} u_0 \\ u_2 \end{pmatrix} \xrightarrow[r\to\infty]{} (\sin kr + \lambda\cos kr)\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \tag{3.1b}$$

[(+) for electrons, (−) for holes], where the free-particle kinematics are given by (2.2) or (2.8). The wave functions u's are now characterized by the phase shift or more conveniently $\lambda = \tan\delta$. δ in our problem is the $l=0$ phase shift and part of the $l=2$ phase shift in the partial-wave expansion of ψ in Eq. (2.5). A resonance state obtains with energy $E=E_R$ when δ increases rapidly through $\frac{1}{2}\pi$ and consequently the scattering cross section

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l \tag{3.2}$$

obtains a maximum giving rise to enhanced observable quantities such as resistivity. The state has a quasilocalized character and has of course a finite lifetime and hence a width Γ. Γ is given by the slope of the δ vs k curve at the resonance point

$$\Gamma = \frac{dE}{d\delta}\bigg|_{E=E_R}. \tag{3.3}$$

(Our Γ is actually half of the width taken in usual scattering analysis.) For a well-defined resonance Γ should be $\ll E_R$. The variational method is based on the property that Eq. (2.8) with the boundary conditions (3.1) can be obtained from $\delta J = 0$, where

$$J = \langle u_0, u_2 | H-E | u_0, u_2\rangle - 2(1+\mu)k\lambda \tag{3.4}$$

with

$$\langle x\rangle = \int_0^\infty x\, dr ,$$

$H-E$ being the radial Hamiltonian in (2.8). The trial functions for the u's are a linear combination of a localized ϕ and a part that goes asymptotically to (3.1b):

$$\begin{pmatrix} u_0 \\ u_2 \end{pmatrix} = \phi + (S+\lambda C)\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

[(+) for electrons, (−) for holes], (3.5)

$$\phi = \sum_{i=1}^{N}\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}\eta_i = \sum_{i=1}^{N}\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}\left(\frac{r}{a}\right)^i e^{-r/a},$$

$$S = \sin kr ,$$

and

$$C = (1 - e^{-r/a})\cos kr$$

to make

$$C(0) = 0 \text{ and } C \to \cos kr, \quad r\to\infty.$$

The α_i's, β_i's, and the λ are the linear variational parameters. We have however treated a effectively as a variational parameter.

A. Kohn method

In this method J is made stationary with respect to all the α_i's, β_i's, and λ.

$$\frac{\partial J}{\partial \alpha_i} = \frac{\partial J}{\partial \beta_i} = 0$$

leads to a $2N \otimes 2N$ matrix equation.

$$\sum_{j=1}^{2N}(\langle\chi_i|H-E|\chi_j\rangle + \langle\chi_j|H-E|\chi_i\rangle)c_j$$
$$= -\langle\chi_i|H-E|S+\lambda C\rangle + \langle S+\lambda C|H-E|\chi_i\rangle , \tag{3.6}$$
$$i = 1, \ldots, 2N,$$

N being the size of the basis set. We have here defined, for reasons of symmetry,

$$\chi_{2i-1} = \eta_i\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \chi_{2i} = \eta_i\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$c_{2i-1} = \alpha_i \text{ and } c_{2i} = \beta_i ,$$

and $S+\lambda C$ is now the spinor

$$(S+\lambda C)\begin{pmatrix}1\\ \pm 1\end{pmatrix}.$$

Next λ_{Kohn} is obtained from the condition $\partial J/\partial\lambda = 0$. The "inner integral" $I=\langle u_0,u_2|H-E|u_0,u_2\rangle$ is not put equal to zero ($I=0$ for an exact solution). As a result λ_{Kohn} can be corrected to the second-order $[\lambda]_{\mathrm{Kohn}}$:

$$[\lambda]_{\mathrm{Kohn}}=\lambda_{\mathrm{Kohn}}-[1/2(1+\mu)k]I(\lambda_{\mathrm{Kohn}}). \tag{3.7}$$

The Kohn method suffers from the appearance of certain spurious singularities. We refer to Truhlar *et al.*[9] for an extensive discussion of these singularities and the ways to deal with them. These have been discussed also in our paper on the variational methods.[10] The essential step is to scan for a fixed k, a few values of a in order to keep away from these singularities and to have a plot of δ vs a. The plot consists of an array of plateaus, the one with $I=0$ is chosen as the optimum plateau. The process is repeated for several values of k and a curve of $a_S(k)$ is chosen for the stationary a. This process is expected to make J reasonably stationary with respect to a, i.e., $\partial J/\partial a=0$. With this $a_S(k)$ then fed back in ϕ, $\delta(k)$ is calculated, and the resonance energy and the width are obtained from the plot δ vs k.

B. Harris method

There is a subtle divergence difficulty (in addition to the spurious singularities mentioned above) in the Kohn method. This is related to the fact that E can come close to one of the eigenvalues E_μ of H in the space of η_i, i.e., the localized state. The matrix inversion in Eq. (3.6) is then not possible. The existence of these E_μ's always limit the Kohn method and conscious efforts have to be taken to avoid them. However as noted by Harris, if one works directly at $E=E_\mu$, which amounts to each side of Eq. (3.6) being separately equal to zero, the Harris phase shift $\lambda_{\mathrm{Harris}}$ can be obtained immediately

$$\lambda_{\mathrm{Harris}}=\langle\phi_\mu|H-E_\mu|S\rangle/\langle\phi_\mu|H-E_\mu|C\rangle$$

with

$$\phi_\mu=\sum_i c_i^\mu\chi_i. \tag{3.8}$$

It can be shown that the spurious singularities are avoided in this procedure and also that $\lambda_{\mathrm{Harris}}$ is the limiting value of λ_{Kohn} (without the second-order correction) at $E=E_\mu$.

In the Harris method, first the E_μ's are calculated from

$$(\underline{H}-E_\mu\underline{\chi})c^\mu=0, \tag{3.9}$$

where

$$\underline{H}_{ij}=\langle\chi_i|H|\chi_j\rangle+\langle\chi_j|H|\chi_i\rangle$$

and

$$\underline{\chi}_{ij}=\langle\chi_i|\chi_j\rangle,$$

and then for each a, $2N$ values of the phase shift, i.e., of λ are obtained—corresponding to different E_μ's. Next, values of λ are calculated for a reasonable range of a thereby getting $2N$ curves for λ vs k. The range of a in our calculation has been taken as $ak\in(0.1 \text{ to } 0.7)$ for N up to 12. As N is increased a number of curves for λ vs k converge into one and becomes stable for $N=9$. This curve then gives the resonance energy E^R_{Harris} and the width Γ_{Harris}. We have argued elsewhere[10] how in this process J is made stationary with respect to a in the Harris method. We have also shown that this method is simpler and more reliable than the Kohn method for calculating resonance energies and widths: the algorithm is simple; the difficulties with the singularities are avoided and no free-free integrals need to be calculated, i.e., those involving only the asymptotic parts of the wave functions (these integrals are often the most difficult to evaluate).

With a Thomas-Fermi potential we have calculated the resonances in both the methods. The results are nearly the same for $N=10$. With the RPA screening, however, the Kohn method becomes computationally very involved. Guided by the accuracy of the Harris method in the previous case, we have calculated the resonances now only in the Harris method. All these results are presented in Sec. IV.

IV. RESULTS

This section is divided in two parts. Section IV A deals with the resonances of a Thomas-Fermi (TF) potential, $V_{\mathrm{TF}}(r)=(2/r)e^{-r/b}$. Though this potential does not correspond to any reasonable model of screening for a ZG semiconductor since it would imply that $\epsilon(q)=\epsilon_0(1+b^{-2}/q^2)$, it is a very instructive phenomenological potential. By varying b from 0 to ∞ some feeling can be developed for the dependence of the resonance with respect to the screening length of the Coulombic potential. This enables a better understanding of the results obtained with a more realistic model for the screening, e.g., the random-phase approximation (RPA) described in Sec. IV B. Also presented in Sec. IV A are the resonance energies and their widths for a wide range of values of the band parameters and of the Fermi momentum k_F.

It should be mentioned that a reliable theoretical estimate of the resonance energy is hampered by the uncertainty which exists in the values of the band parameters. The electron effective mass has been accurately estimated at about $0.029m_0$ but the hole mass, which can be measured directly only with difficulty, is not well known. Values varying from $0.25m_0$ to $0.7m_0$ have been given. One can divide these estimates for m_v^* in two classes—those between 0.25 and $0.45m_0$ (group I) and those between 0.5 and $0.7m_0$ (group II). The former are obtained indirectly from the values of the band parameters evaluated by fitting optical-absorption measurements to a $\vec{k}\cdot\vec{p}$ band structure,[13] while the latter are obtained directly from experiments on a free gas of holes.[14] The lattice dielectric constant ϵ_0 has been taken to be 20.

A. Thomas-Fermi potential

The TF potential is found to have several resonances. Their number and energy increase with b as expected on physical grounds. In Fig. 2 only the first two resonances of symmetry $S_{3/2}$ are shown for two typical values of the band parameters ($m_v^* = 0.6m_0$ and $m_v^* = 0.4m_0$, in both cases $m_c^* = 0.029$). As b exceeds 2000 Å, E_R keeps increasing smoothly. The limit $b \to \infty$ corresponds to a pure Coulomb potential. $E_R \to 9$ meV for $m_v^* \to 0.6m_0$ and $E_R \to 6$ meV for $m_v^* = 0.4m_0$. In this limit the simple proportionality relation $E_R \propto m_v^*$ predicted by Gel'mont and Dyakonov[6] seems to hold. It is actually difficult to obtain an accurate value for E_R for a pure Coulomb potential because our variational calculation has been set up in a way that is well suited for finite-range potentials. Convergence is difficult to achieve for a potential with an infinite range. As it is clear from Fig. 2, for the values of $\beta = m_c^*/m_v^*$ we are considering, namely, 0.0725 and 0.048, the relation $E_R \propto m_v^*$ does not hold for a finite b. For this relation to hold the two curves on the graph should have coincided. The resonance energy is seen to be reduced more than expected by a decrease in the hole mass. An interesting result of our calculation is that the Coulomb potential is very sensitive to screening. It is sufficient to reduce its range below 500 Å to substantially affect its strength. For screening lengths less that 50 Å for $m_v^* = 0.6m_0$ and 80 Å for $m_v^* = 0.4m_0$ the charged impurity becomes a weak scatterer with no resonances.

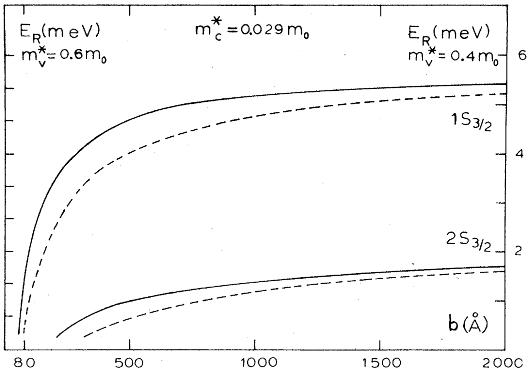

FIG. 2. Resonance energy E_R as a function of the screening length b of the TF potential.

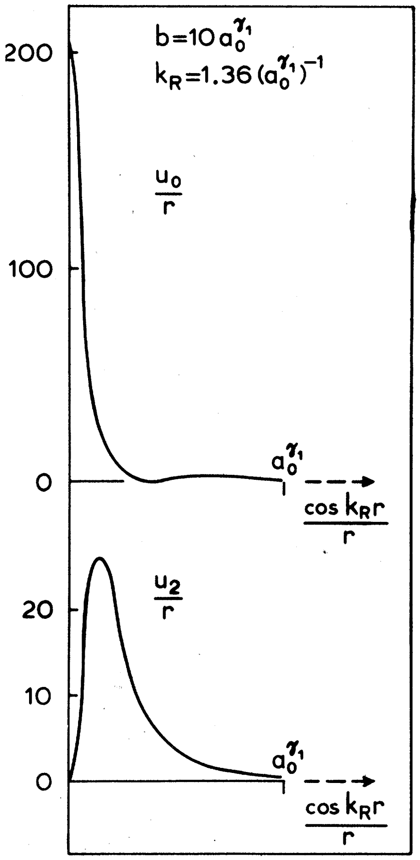

FIG. 3. Typical pair of resonance wave functions.

The width $\Gamma = (dE/d\delta)_{E=E_R}$ of these resonances has been calculated. We find that the relative width Γ/E_R is only weakly dependent on the screening length b. For the $1S_{3/2}$ resonances $\Gamma/E_R \simeq 7\%$ (7% to 7.4%) for $m_v^* = 0.4m_0$ and 4% (4.2% to 4.4%) for $m_v^* = 0.6m_0$. It increases slightly as b is reduced. These values suggest that the formula (2.10) overestimates the increase of Γ/E_R with β. The $2S_{3/2}$ resonances are significantly broader. We also show in Fig. 3 the plots of a typical pair of resonance wave functions. A striking similarity exists between these functions and the bound-state functions obtained by Baldereschi and Lipari[7] in their study of acceptor states in ordinary semiconductors.

B. Random-phase-approximation potential

As stated in Sec. III, the screening in a ZG semiconductor is more complex than in an ordinary gap semiconductor. For our calculation of charged-impurity resonances we have used the RPA screening. In this approximation, the static dielectric function is given by

$$\frac{\epsilon(q)}{\epsilon_0}=1-\frac{8\pi}{q^2}\sum_{n,n',\vec{k}}|\langle\vec{k},n|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q},n'\rangle|^2 \times\frac{f(\vec{k}+\vec{q},n')-f(\vec{k},n)}{E_{\vec{k}+\vec{q},n'}-E_{\vec{k},n}}, \qquad (4.1)$$

where $|\vec{k},n\rangle$ are Bloch states of energy $E_{\vec{k},n}$ and occupation number $f(\vec{k},n)$ with n specifying the conduction or valence band

$$\frac{\epsilon(q)}{\epsilon_0}=1-\frac{16\pi}{q^2}\sum_{\vec{k}}|\langle\vec{k},c|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q},v\rangle^2\frac{f(\vec{k}+\vec{q},v)-f(\vec{k},c)}{E_{\vec{k}+\vec{q},v}-E_{\vec{k},c}} -\frac{16\pi}{q^2}\sum_{\vec{k}}|\langle\vec{k},c|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q},c\rangle|^2\frac{f(\vec{k}+\vec{q},c)-f(\vec{k},c)}{E_{\vec{k}+\vec{q},c}-E_{\vec{k},c}}. \qquad (4.2)$$

c and v stand, respectively, for the conduction and the valence bands. The first sum is the $\Gamma_8-\Gamma_8$ interband polarization and the second is the intraband polarization of the conduction band. We are considering n-type substances only because in the actual experimental situation of interest the Fermi level lies above the point of degeneracy. The matrix elements have been evaluated using the spherical approximation of the $\vec{k}\cdot\vec{p}$ band-structure model and within the effective-mass approximation.[15]

$$|\langle\vec{k},c|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q},v\rangle|^2=X(\vec{k},\vec{q}) \qquad (4.3)$$

and

$$|\langle\vec{k},c|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q},c\rangle|^2=1-X(\vec{k},\vec{q}), \qquad (4.4)$$

where

$$X(\vec{k},\vec{q})=\tfrac{3}{4}(\vec{k}\times\vec{q})^2/|\vec{k}|^2|\vec{k}+\vec{q}|^2$$

for free electrons the second matrix element is just unity. $X(\vec{k},\vec{q})$ shows the contribution of the $p_{3/2}$ character from the band carriers. Evaluating (4.2), the dielectric function can be written[16,17]

$$\frac{\epsilon(q)}{\epsilon_0}=1+\frac{1-\beta}{\pi k_F}\left[3f_\beta\left(\frac{q}{k_F}\right)+2\frac{k_F^2}{q^2}g\left(\frac{q}{k_F}\right)\right], \qquad (4.5)$$

k_F being the Fermi momentum where $f_\beta(x)$ and $g(x)$ are the interband function and the intraband function, respectively. The former is dependent on β but the latter is not. k_F enters in both functions only as a scaling parameter. Broerman *et al.*[17] have derived an expression for $g(x)$ and Liu and Tosatti[18] have calculated $f_{\beta=0}(x)$. In the Appendix we give our own derivation of these two functions because of the relative simplicity of the expressions we have obtained and also because Liu and Tosatti have not considered the β sensitivity of $f_\beta(x)$. The plots of f and g are given in Figs. 4 and 5. When $k_F=0$, i.e., for the intrinsic substance

$$\frac{[\epsilon(q)]_{k_F=0}}{\epsilon_0}=1+\frac{3(1-\beta)}{2\beta}\left[\frac{\pi}{2}-(1-\beta)\sin^{-1}\left(\frac{1}{1+\beta}\right)^{1/2}-\beta^{1/2}\right]\frac{1}{q} =1+\lambda_\beta/q \qquad (4.6)$$

and a $1/q$ singularity is present.[15] This singularity is characteristic of the pure ZG semiconductor. The RPA therefore predicts that charged impurities are strongly screened in these substances.

As k_F is increased, the dielectric singularity is

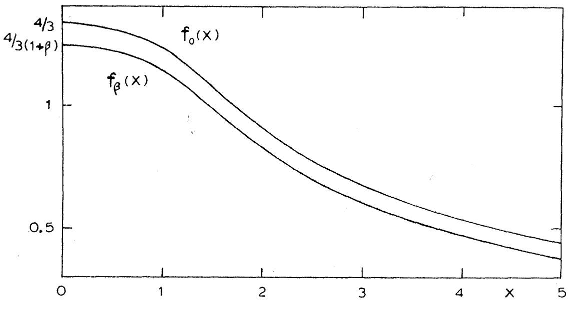

FIG. 4. Functions $f_0(x)=f_{\beta=0}(x)$ and $f_\beta(x)$ for $\beta=0.0725$ (corresponding for instance to $m_c^*=0.029m_0$ and $m_v^*=0.4m_0$).

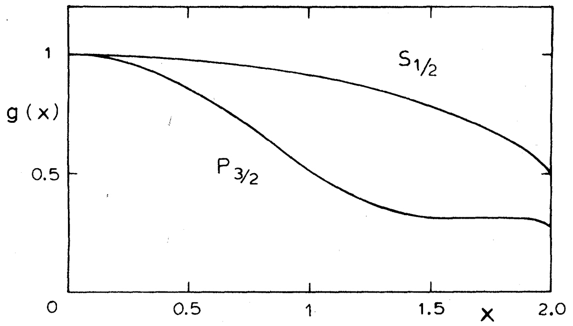

FIG. 5. Function $g(x)$ for $S_{1/2}$ electrons (the Lindhart function) and for $P_{3/2}$ electrons.

reduced but the intraband polarization becomes effective. As the plot of the g function shows, the $P_{3/2}$ electrons screen less than the $S_{1/2}$ electrons. This is expected on physical grounds. The effect on the resonance of these screening terms will now be presented.

1. $k_F = 0$

In this limit

$$[\epsilon(q)]_{k_F=0}/\epsilon_0 = 1 + \lambda_\beta/q\,.$$

The potential corresponding to a dielectric function of this form can be easily evaluated.[15]

$$V(r) = \left(\frac{2}{r}\right)\left(\frac{2}{\pi}\right)\int_0^\infty dq \frac{\sin qr}{q[\epsilon(q)/\epsilon_0]} = \frac{2}{r}F(\lambda_\beta r)\,, \quad (4.7)$$

where

$$F(x) = \cos x\left[1 - \left(\frac{2}{\pi}\right)\mathrm{Si}(x)\right] + \left(\frac{2}{\pi}\right)\sin x\,\mathrm{Ci}(x)\,. \quad (4.8)$$

In (4.8) the functions $\mathrm{Si}(x)$ and $\mathrm{Ci}(x)$ are sine and cosine integrals defined by

$$\mathrm{Si}(x) = \int_0^x \frac{\sin t}{t}\,dt$$

and

$$\mathrm{Ci}(x) = -\int_x^\infty \frac{\cos t}{t}\,dt\,.$$

A plot of $F(x)$ is given in Fig. 6. It is difficult to fit it with any simpler function. The inner part is nearly exponential. In that region the potential would correspond to a short-range TF potential of screening length of about $b = 80$ Å for both $m_v^* = 0.4m_0$ and $0.6m_0$. But the tail is of some inverse power with a slowly decaying rate ($1/r^2$). It is a much longer tail than that of a TF potential which corresponds to a dielectric function with a stronger singularity $1/q^2$.

With our choice of the trial functions, all the matrix elements entering the Harris method can be evaluated analytically if an interchange of the order of integration is made, i.e., if one performs the r integration before the q integration which transorms $V(q)$ into $V(r)$.

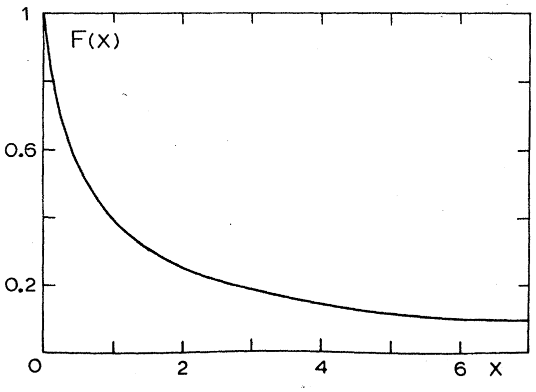

FIG. 6. Screening function $F(x)$ in the RPA screened-Coulomb potential.

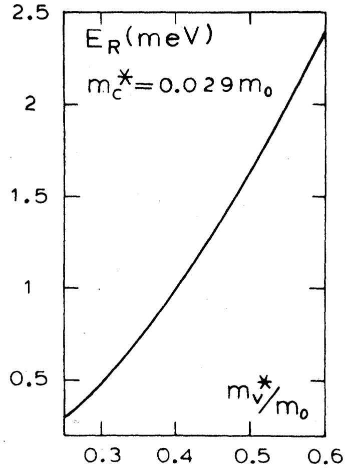

FIG. 7. Resonance energy E_R as a function of the hole effective mass m_v^* for an electron effective mass $m_c^* = 0.029m_0$ in the RPA.

In Fig. 7 the resonance energy as a function of the hole mass is plotted for the whole range of reported hole masses. The energies from group A are in the neighborhood of the 0.7-meV level (A_0) while those from group B are in the neighborhood of the 2.2-meV level (A_1).

It is instructive to note that we could have roughly reproduced Fig. 7 by simply supposing that the potential (4.7) is equivalent to a TF potential of screening length ~80 Å (see Fig. 7). The resonance energy is therefore essentially dependent upon the inner part of the potential where the $1/r$ singularity is present. This is not too surprising considering the nature of these resonances which are states quasilocalized around the impurity. The width on the other hand, not surprisingly, is slightly smaller: when $m_v^* = 0.4m_0$, $\Gamma/E_R = 6.8\%$ and when $m_v^* = 0.6m_0$, $\Gamma/E_R = 4.1\%$ instead of 7.4% and 4.4% respectively, for a TF potential of screening length $b \sim 80$ Å. The longer tail of the ZG semiconductor makes it a more efficient scatterer than the corresponding TF potential.

2. $k_F \neq 0$

When $k_F \neq 0$, the expression for $\epsilon(q)$ is complicated. As explained in the Appendix, we are able to make a fit which greatly simplifies the calculations, allowing an exact evaluation of the matrix elements in the variational calculation. In Fig. 8, we show the results for a typical value

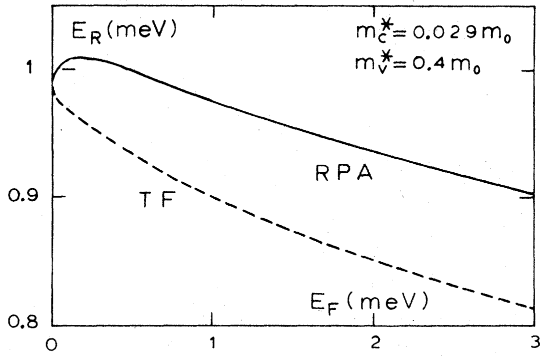

FIG. 8. Typical variation of resonance energy with increasing Fermi level. Both RPA and TF for intraband polarization. Only RPA for the interband polarization.

of $m_v^* = 0.4m_0$. The dotted line has been obtained by supposing that $g(x) = 1$ which is equivalent to taking a simple TF model for the intraband screening. Our results indicate that within our model two competing contributions cancel each other. On the one hand we have the strengthening of the intraband polarization which would increase the screening and on the other hand the weakening of the dielectric singularity which would decrease it. The cancellation effect is complete for $E_F \simeq E_R$. If we had plotted E_R vs m_v^* for $E_F = E_R$, the curve obtained would nearly coincide with the curve in Fig. 7 for $E_F = 0$. A maximum deviation of $\frac{1}{16}$ occurs at the lower mass $m_v^* = 0.3m_0$. In the interval $0 < E_F < E_R$ the variation of E_R is small (of the order of 10%).

The relative width Γ/E_R is found to be reduced from its $k_F = 0$ value as long as $k_F < k_R$. This may be attributed to the softening of the dielectric singularity due to the interband polarization which further elongates the tail of the F function in Eq. (4.7). It may be recalled that the low-q values in $\epsilon(q)$ determine the long-range part of the corresponding screened potential.

In the actual experimental situation $E_F \neq 0$ but $E_F < E_R$. Since only small changes in E_R are predicted from the intrinsic case, Fig. 7 can be used to discuss our results.

V. DISCUSSION

The physical problem we have studied is that of a singly charged acceptor impurity in an n-type zero-gap semiconductor. The charged impurity is known to be highly screened by the lattice polarization ($\epsilon_0 \simeq 20$) and by the electrons of the Γ_8 bands. The expected resonance state is a shallow one extending over several hundred angstroms.

Before comparing our theoretical results with the experimental values given in the literature let us try to assess the validity of our model. We have solved variationally for the ground-state resonance is the effective-mass approximation. The Hamiltonian is the one-particle Hamiltonian of Kohn and Luttinger neglecting the warping of the bands and the potential is the RPA screened-Coulomb potential. This effective-mass theory should be good for such extended states. We maintain that to attempt to go beyond it at this stage would be premature and that the many-body effects can be satisfactorily incorporated into the dielectric function as we have done. The neglect of the cubic terms in the Hamiltonian should not be important for HgTe. No significant anisotropy in the electronic properties of the substance has been found,[1] and as Baldereschi and Lipari[18] have shown for states of symmetry $S_{3/2}$ (as the ground state), there is no first-order correction. Therefore, any major discrepancy would be due to the inadequacy of our dielectric function. The RPA polarization terms are actually only the first terms in an expansion in powers of $(E_B/E_F)^{1/2}$ (E_B: Bohr energy of the electron) and in the specimens usually considered $E_B \sim E_F$ which seems to invalidate the approximation we have used. Abrikosov[19] however, has gone beyond the RPA and has applied renormalization-group methods to this problem (he has considered only the most sensitive case $E_F \simeq 0$). His main conclusion is that the RPA gives the correct values except in a very narrow region near the degeneracy point where the strong enhancement might be overestimated. The region is characterized by $q \ll k_B = 1/(1+\mu)$, the electron Bohr momentum. We have done a calculation where the singularity has been cut off below $q = k_B$. We find that the resonance energy increases by at most 15%. For finite k_F the RPA should be much better since the interband polarization $4\pi\alpha_{\text{inter}}(0)$ (see the Appendix) decreases as $1/k_F$ $[4\pi\alpha_{\text{inter}}(0) = 4\epsilon_0/\pi\mu k_F]$. The RPA can therefore be expected to give a reliable estimate of the resonance energy though lower than the actual value. On these grounds a fair comparison with experiment can be made in spite of the uncertainty in the hole mass m_v^*. As already discussed, the reported values of m_v^* can be separated in two groups which are consistent within themselves—group I (0.25 to 0.45) and group II (0.5 to 0.7). In our model if the true mass is within group I, A_0 should be the charged-impurity level and $m_v^* \lesssim 0.35$ but if it is within group II, A_1 should be the charged-impurity level with $m_v^* \lesssim 0.58m_0$.

At this stage it seems difficult to assert beyond doubt whether A_0 or A_1 is the charged-impurity resonance. The evidence which we have now is not sufficiently conclusive though it favors A_0.

Our model predicts that A_0 would be a charged-impurity resonance if the hole mass were about $0.35m_0$. This is the $\vec{k}\cdot\vec{p}$ theoretical estimate of the hole mass[20] and falls well within group I which comprises the values of m_v^* obtained with allowance made for the Γ_8 symmetry of the bands. An argument has been put forward to favor A_1 as a vacancy state, namely, the variation of A_1 with Cd concentration in the HgCdTe alloys.[3,21] As Cd replaces Hg, the A_1 level rises sharply from 2.2 to 15 meV at the semimetal to semiconductor (SM-SC) transition, which is what ought to happen for a vacancy state—the Γ_6 band moving up rapidly with Cd concentration pushes up the vacancy level. It is however to be noted that the charged-impurity resonance is also expected to have a similar behavior. For a charged-impurity state which is essentailly attached to the heavy hold band, the decisive factor is the change in the static screening with Cd fraction. As we have seen the screening function in pure HgTe has a strong singularity. This singularity is substantially reduced as the Γ_6 band moves up until is disappears at the SM-SC transition when HgCdTe becomes an ordinary semiconductor. The strength of the static dielectric function is hence substantially reduced and the energy of the charged-impurity resonance will increase with Cd concentration. A rough estimate of the resonance energy at the SM-SC transition can be made using Eq. (2.9). According to this formula A_0 would rise from 0.7 to about 6.5 meV and A_1 from 2.2 to 11 meV. This assumes that ϵ_0 decreases to about 18, as is experimentally known,[22] and that m_v^* remains constant. If A_0 is the charged-impurity resonance, then a resonance state should be observable around 6.5 meV. Interestingly Kozacki *et al.*[23] have reported for the alloy $Hg_{0.845}Cd_{0.155}Te$ ($E_G = -7.1$ meV $\gg -300$ meV of HgTe) an acceptor level of 5.4 meV which could well be the Coulombic level. The existence of such a level will further strengthen the case for A_0 as the charged-impurity resonance and our model will have explained why a charged-impurity level can have so low an energy as 0.7 meV when the hole mass is $0.35m_0$, i.e., within group I.

ACKNOWLEDGMENTS

This research has been financed by a National Research Council of Canada grant. One of us (B.J.) acknowledges a Centennial scholarship.

APPENDIX

In Sec. IV, the RPA dielectric function is written

$$\frac{\epsilon(q)}{\epsilon_0} = 1 + \frac{1-\beta}{\pi k_F}\left[3f_\beta\left(\frac{q}{k_F}\right) + \frac{2k_F^2}{q^2}g\left(\frac{q}{k_F}\right)\right],$$

where f depends only on β and g is independent of any band parameter. This form is obtained by supposing $q, k_F \ll K$ and for all practical purposes K infinite (all being very reasonable assumptions considering that $K > 200$). Our method of evaluating $\epsilon(q)$ is different from that of Refs. 16 and 17. In view of the relative simplicity of our expressions and the fact that they have not been published previously, we have decided to present our results. Following Eqs. (4.2)–(4.4) one can write

$$\frac{\epsilon(q)}{\epsilon_0} = 1 + \frac{4\pi\alpha_{\text{inter}}(q)}{\epsilon_0} + \frac{4\pi\alpha_{\text{intra}}(q)}{\epsilon_0}, \tag{A1}$$

where we have separated the band contribution into two parts:

$$\frac{4\pi\alpha_{\text{inter}}(q)}{\epsilon_0} = \frac{-16\pi}{q^2}\sum_k X(\vec{k},\vec{q})\frac{f(\vec{k}+\vec{q},v) - f(\vec{k},\vec{c})}{E_{\vec{k}+\vec{q},v} - E_{\vec{k},c}} \tag{A1a}$$

and

$$\frac{4\pi\alpha_{\text{intra}}(q)}{\epsilon_0} = \frac{-16\pi}{q^2}\sum_k [1 - X(\vec{k},\vec{q})]\frac{f(\vec{k}+\vec{q},c) - f(\vec{k},c)}{E_{\vec{k}+\vec{q},c} - E_{\vec{k},c}} \tag{A1b}$$

with

$$X(\vec{k},\vec{q}) = \tfrac{3}{4}|\vec{k}\times\vec{q}|^2/|\vec{k}|^2|\vec{k}+\vec{q}|^2 .$$

The domain of integration of each term is defined by the nonzero values of the occupation functions. A simpler evaluation of these polarization terms can be made if they are broken up into two and a change of variable is made in their first term so as to have only occupation functions of the form $f(\vec{k},c)$ or $f(\vec{k},v)$:

$$\frac{4\pi\alpha_{\text{inter}}(q)}{\epsilon_0} = \frac{16\pi}{q^2}\sum_k X(\vec{k},\vec{q})\left(\frac{f(\vec{k},v)}{E_{\vec{k}+\vec{q},c} - E_{\vec{k},v}} - \frac{f(\vec{k},c)}{E_{\vec{k},c} - E_{\vec{k}+\vec{q},v}}\right) \tag{A2a}$$

and

$$\frac{4\pi\alpha_{\text{intra}}(q)}{\epsilon_0} = \frac{32\pi}{q^2}\sum_k [1 - X(\vec{k},\vec{q})]\frac{f(\vec{k},c)}{E_{\vec{k}+\vec{q},c} - E_{\vec{k},c}} . \tag{A2b}$$

The evaluation of the principal parts of these integrals yields the f and g functions.

(i)

$$\frac{4\pi\alpha_{\text{inter}}(q)}{\epsilon_0} = \frac{3(1-\beta)}{\pi k_F}f_\beta\left(\frac{q}{k_F}\right), \tag{A3}$$

where $f_\beta(x) = f_0(x) - I_\beta(x)$;

$$f_0(x) = \frac{1}{2x}\int_0^x dy \frac{1}{y}\ln\left|\frac{1+y}{1-y}\right| + \frac{1}{4x}\left[\frac{1}{2}\ln\left|\frac{1+x}{1-x}\right|\left(\frac{1}{x^2} - x^2\right) + x - \frac{1}{x}\right] \tag{A3a}$$

and

$$I_\beta(x) = \frac{1}{x}\int_{-1}^{1} d\mu \frac{1-\mu^2}{(1+1/\beta-\mu^2)^{1/2}} \times\left(\frac{\pi}{2} - \tan^{-1}\frac{(1+1/\beta)1/x+\mu}{(1+1/\beta-\mu^2)^{1/2}}\right). \tag{A3b}$$

$f_0(x)$ is in fact the f function given by Liu and Tosatti.[16] The second term $I_\beta(x)$, zero when $\beta = 0$, is the correction to $f_0(x)$ due to the finite size of β. The integrand in $I_\beta(x)$ is a well-behaved function, 0 at both limits -1 and 1, and peaking near $\mu \simeq 0$. Such a function is easily integrable numerically by a simple Simpson's rule. $f_\beta(x)$ and $f_0(x)$ are plotted in Fig. 4. These functions have the asymptotic values

$$f_0(0) = \frac{4}{3}, \quad I_\beta(0) = \frac{4}{3}\frac{\beta}{1+\beta}, \quad \text{or } f_\beta(0) = \frac{4}{3}\frac{1}{1+\beta} \tag{A4}$$

and

$$f_0(x) \rightarrow \tfrac{1}{4}\pi^2(1/x) - \tfrac{4}{3}(1/x^2),$$

$$I_\beta(x) \rightarrow \frac{1}{x}\left(\frac{\pi}{2\beta}\right)\left[\beta^{1/2} - (1-\beta)\sin^{-1}\left(\frac{\beta}{1+\beta}\right)^{1/2}\right]$$

Thus,

$$f_\beta(x) \rightarrow \lambda'_\beta/x - C_\beta/x^2 \quad \text{as } x \rightarrow \infty, \tag{A5}$$

where

$$\lambda'_\beta = \lambda_\beta \frac{\pi}{3(1-\beta)} = \frac{\pi}{2\beta}\left[\frac{\pi\beta}{2} + (1-\beta)\sin^{-1}\left(\frac{\beta}{1+\beta}\right)^{1/2} - \beta^{1/2}\right]$$

$$= \frac{\pi}{2\beta}\left[\frac{\pi}{2} - (1-\beta)\sin^{-1}\left(\frac{1}{1+\beta}\right)^{1/2} - \beta^{1/2}\right]$$

and

$$C_\beta \simeq \tfrac{4}{3}$$

(ii)

$$\frac{4\pi\alpha_{\text{intra}}(q)}{\epsilon_0} = \frac{2(1-\beta)}{\pi}\frac{k_F}{q^2}g\left(\frac{q}{k_F}\right), \tag{A6}$$

where

$$g(x) = \frac{1}{2}\left[\frac{7}{4} + \ln\left|\frac{2+x}{2-x}\right|\left(1 - \frac{x^2}{4}\right)\left(\frac{1}{x} + \frac{3}{4}x\right) + \frac{3}{8}x\ln\left|\frac{1+x}{1-x}\right|\left(x^2 - \frac{1}{x^2}\right) - \frac{3}{2}x\int_{x/2}^{x} dy \frac{1}{y}\ln\left|\frac{1+y}{1-y}\right|\right]$$

The asymptotic limits are $g(0) = 1$ and

$$g(x) \underset{x\rightarrow\infty}{\longrightarrow} \frac{1.57}{x^2}.$$

Broerman *et al.*[17] do not give an expression for g. Their plot of g and our Fig. 5 compare well.

To enter the exact formulas of f_β and g into the variational calculation would be computationally prohibitive. We have for this reason approximated them with simpler functions having the proper leading asymptotic limits. This enabled us to actually evaluate all the necessary integrals exactly.

The approximate f_β and g have the form

$$f_T(x) = \begin{cases} \tfrac{4}{3}[1/(1+\beta)] - Bx^2, & x < x_0 \\ \lambda'_\beta/x - \dfrac{C}{x^2}, & x \geq x_0 \end{cases} \tag{A7}$$

and

$$g_T(x) = \begin{cases} 1 - \alpha_1 x^2 + \alpha_2 x^4, & x < x_0 \\ 1 - \alpha_1 x_0^2 + \alpha_2 x_0^4, & x \geq x_0 \end{cases}. \tag{A8}$$

The parameters x_0, B, C, α_1, and α_2 are chosen so as to optimize the fit. For instance when $m_c^* = 0.029m_0$ and $m_v^* = 0.4m_0$, i.e., $\beta = 0.0725$, the following values were chosen for the constants: $x_0 = 1.75$, $B = 0.1118$, $C = 0.6$, $\alpha_1 = 0.5236$, and $\alpha_2 = 0.0977$.

α_1 and α_2 are different from the values given by Broerman *et al.* By not attempting to fit $g(x)$ with the above polynomial up to $x = 2$, but only up to 1.75 we are able to make a better fit.

In this approximation, the dielectric function has the form

$$\frac{\epsilon(q)}{\epsilon_0} = 1 + \frac{(1-\beta)}{\pi}\frac{k_F}{q^2}G_\beta\left(\frac{q}{k_F}\right), \tag{A9}$$

where

$$G_\beta(x) = 3x^2 f_\beta(x) + 2g(x)$$

$$= \begin{cases} 2 + 2(2/(1+\beta) - \alpha_1)x^2 + (2\alpha_2 - 3B)x^4, \\ x < x_0 \\ 2(1 - \alpha_1 x_0^2 + \alpha_2 x_0^4) - 3C + 3\lambda'_\beta x, \\ x \geq x_0. \end{cases} \tag{A10}$$

This function deviates from the RPA $G(x)$ by less than 2%.

[1] See N. N. Berchenko and M. V. Poshovskii, Sov. Phys. Usp. 19, 462 (1976); B.L. Gel'mont, V. I. Ivanov-Omskii, and I. M. Tsildil'kovskii, *ibid.* 19, 879 (1976), and references therein.

[2] L. Liu and D. Brust, Phys. Rev. 157, 627 (1967).

[3] A. Mauger and J. Friedel, Phys. Rev. B 12, 2412 (1975).

[4] J. M. Luttinger and W. Kohn, Phys. Rev. 97, 869 (1955).

[5] G. Bastard and P. Nozières, Phys. Rev. B 13, 2560 (1976).

[6] B. L. Gel'mont and M. I. Dyakonov, Sov. Phys. JETP 35, 377 (1972).

[7] N. O. Lipari and A. Baldereschi, Phys. Rev. Lett. 25, 1660 (1970). A. Baldereschi and N. O. Lipari, Phys. Rev. B 8, 2697 (1973).

[8] C. J. Joachain, *Collision Theory* (North-Holland, Amsterdam, 1975).

[9] D. G. Truhlar, J. Abdallah, Jr., and R. L. Smith, Adv. Chem. Phys. 25, 211 (1974).

[10] A. K. Das, B. Joos and P. R. Wallace, J. Phys. C (to be published).

[11] W. Kohn, Phys. Rev. 74, 1763 (1948).

[12] F. E. Harris, Phys. Rev. Lett. 19, 173 (1967).

[13] J. Guldner, C. Rigaux, M. Grynberg, and A. Mycielski, Phys. Rev. B 8, 3875 (1973); S. Uchida, R. Yoshizaki, and S. Tanaka, in International Conference on the Physics of Semimetals and Narrow-Gap Semiconductors, Nice, France, 1973 (unpublished); S. H. Groves, C. R. Pidgeon, and R. N. Brown, Phys. Rev. 161, 779 (1967); C. R. Pidgeon and S. H. Groves, in *Proceedings of the International Conference on II-VI Semiconductor Compounds, Providence, R. I.* (Academic, New York, 1967), p. 1080.

[14] V. I. Ivanov-Omskii, B. T. Kolomiets, A. A. Malkova, and A. S. Mekhtiev, Phys. Status Solidi 32, K83 (1969). V. I. Ivanov-Omskii, F. P. Kesamanly, B. T. Kolomiets, A. S. Mekhtiev, and V. A. Skripkin, Phys. Status Solidi 27, K169 (1968).

[15] L. Liu and D. Brust, Phys. Rev. 173, 777 (1968); Phys. Rev. Lett. 20, 651 (1968).

[16] L. Liu and E. Tosatti, Phys. Rev. Lett. 23, 772 (1969); Phys. Rev. B 2, 1926 (1970).

[17] J. G. Broerman, L. Liu, and K. N. Pathak, Phys. Rev. B 4, 664 (1971).

[18] A. Baldereschi and N. O. Lipari, Phys. Rev. B 9, 1525 (1974).

[19] A. A. Abrikosov, Sov. Phys. JETP 39, 709 (1974).

[20] See, for instance, V. I. Ivanov-Omskii, B. T. Kolomiets, A. A. Mal'kova, Yu. F. Markov, and A. Sh Mekhtiev, Sov. Phys. Semicond. 2, 1122 (1969).

[21] L. Liu and C. Vérié, Phys. Rev. Lett. 37, 453 (1976).

[22] J. Baars and F. Sorger, Solid State Commun. 10, 875 (1972).

[23] A. Kozacki, S. Otmezguine, G. Weill, and C. Vérié, in *Proceedings of the 13th International Conference on the Physics of Semiconductors, Rome, 1976*, edited by F. G. Fumi (North-Holland, Amsterdam, 1976) p. 467.