conduction mass, are used to calculate a^* and ϵ_i. This, however, neglects any polaron effects. This is not understandable since the material is highly polar with a coupling constant $\alpha \approx 3.3$. The lack of good agreement between theory and experiment suggests that the simple semiconductor theories do not adequately describe the donor states in a polar material. The experimental Hall mobility can be fitted over a wide range of temperatures by assuming a combination of various scattering mechanisms. The polar optical mode and the acoustic mode, together, provide the fit down to 250°K. Below this temperature, ionized impurity scattering seems to be the additional mechanism.

ACKNOWLEDGMENTS

We are grateful to Dr. F. Trautweiler and his group for growing the CdF_2 crystals used in this investigation. We should like to thank Dr. D. Matz, Dr. F. Trautweiler, and F. Moser for helpful discussions during the course of this work. The valuable comments by Dr. D. C. Hoesterey on the manuscript are gratefully appreciated. We are grateful to W. Pinch and W. Selke for their technical assistance. We acknowledge the help of R. Ambrose in the atomic absorption analysis of our samples. We are thankful to Dr. D. L. Losee for evaporating titanium-gold contacts on glass plates.

PHYSICAL REVIEW VOLUME 183, NUMBER 3 15 JULY 1969

Dielectric Screening of the Electron-Hole Interaction in Small-Gap Semiconductors

JÜRGEN K. KÜBLER*
Department of Physics, Texas A&M University, College Station, Texas 77843
(Received 29 January 1969)

The dielectric function that screens the electron-hole interaction in semiconductors has been derived from first principles, and is found to differ in an essential way from the usual random-phase-approximation (RPA) dielectric function. This screening function determines the binding energy of the exciton, and hence it is important for the theory of the excitonic insulator. We have studied the binding energy of the exciton for a simple band structure, and find that the case considered does not become unstable toward exciton formation, whereas the usual dielectric function in RPA predicts the excitonic instability.

I. INTRODUCTION

THE semiconductor-semimetal phase transition with a possible intermediate excitonic insulator phase has recently received considerable attention in connection with substances where the energy gap can be changed.[1,2] The ideas developed for the excitonic insulator have subsequently been used to describe the Mott transition in heavily doped semiconductors.[3,4] The possibility of the excitonic insulator phase is based on the assumption that the binding energy of the exciton, E_B, remains nonzero for vanishing energy gaps and certain band structures. The binding energy E_B is, apart from some constants, given by ϵ_0^{-2}, where ϵ_0 is some dielectric screening constant. Thus one explains why ϵ_0 remains finite for vanishing gaps.[5] In situations where the energy gap G is smaller or equal to the binding energy E_B, the normal ground state of the semiconductor becomes unstable toward exciton formation and the new ground state has many interesting properties.[1,2] Even though it is clear that the strong electron-hole interaction ($\epsilon_0 < \infty$) is at the heart of the theory, the nature of the screening function has never been explored sufficiently, at least not to our knowledge. Haken's theory[6,7] seems inapplicable to small energy gaps, and the whole problem still seems to be as summarized by Knox[8] in 1963. From Kohn's theory[9] of the interaction of an electron with a test charge or with an impurity, it is plausible that the usual dielectric function in the random-phase approximation (RPA)[10] screens the electron-hole interaction. In this paper we investigate this assumption and find that it is not really justified. In Sec. II A we formulate the screening function using many-body techniques.[11] In Sec. II B we show how our formulation gives the usual dielectric function

* Present address: Abteilung für Physik, Ruhruniversität, Bochum, Germany.

[1] D. Jérome, T. M. Rice, and W. Kohn, Phys. Rev. **158**, 462 (1967).

[2] B. I. Halperin and T. M. Rice, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic Press Inc., New York, 1968), Vol. 21.

[3] W. Kohn, Phys. Rev. Letters **19**, 789 (1967).

[4] N. F. Mott and E. A. Davis, Phil. Mag. **17**, 1269 (1968).

[5] J. des Cloizeaux, J. Phys. Chem. Solids **26**, 259 (1965).

[6] H. Haken, Fortschr. Physik **6**, 271 (1958).

[7] H. Haken and W. Schottky, Z. Physik. Chem. (Frankfurt) **16**, 218 (1958).

[8] R. S. Knox, Solid State Phys. Suppl. **5**, 78 (1963).

[9] W. Kohn, Phys. Rev. **105**, 509 (1957); **110**, 857 (1958).

[10] D. Pines, *Elementary Excitations in Solids* (W. A. Benjamin, Inc., New York, 1963). Almost all relevant references can be found here.

[11] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Inc., Englewood Cliffs, N. J., 1963).

in the Hartree–Fock approximation if we ignore multiple-scattering processes. In Sec. II C the screening function is computed beyond the Hartree-Fock approximation, and it is shown how virtual creations of electron-hole pairs of intermediate radii give rise to a screening function which differs from the usual RPA function. In Sec. III we evaluate our screening function for a rather special case and show that for this case our screening function does not lead to the excitonic instability, whereas the usual RPA function does. However, for energy gaps of the order of those where the excitonic instability would be expected to occur on the basis of phenomenological screening, the binding energy of the exciton is found to vanish, thus indicating a possible phase change again.

II. FORMULATION OF THE SCREENING FUNCTION

A. Formal Equations

We consider a crystal with one atom per unit cell at the absolute zero of temperature. The nuclear coordinates and the core states we consider incorporated into a crystal-periodic potential and we ignore spin-dependent forces. N electrons are assumed to fill up the highest valence band v whereas the conduction bands c are assumed to be empty.

The exciton is a particle-conserving excitation of the electronic system; its binding energy is given by the pole of lowest energy of the vertex part in the gap of the one-particle spectrum.[11] It has been shown how the effective-mass equation (EME) for the Wannier exciton can be obtained from the vertex part[12,13] and the interaction term of the EME was shown to be given by a sum of polarization diagrams that can in the simplest possible approximation be written as

$$\Gamma(1,2,3,4,\omega)=\Gamma^{(0)}(1,2,3,4)+\sum_{5,6}\Gamma^{(0)}(5,2,4,6)G_0(5,6,\omega)\Gamma(1,6,3,5,\omega), \quad (1)$$

where

$$\Gamma^{(0)}(1,2,3,4)=\langle 12||\mathbf{r}-\mathbf{r}'|^{-1}|43\rangle-\langle 12||\mathbf{r}-\mathbf{r}'|^{-1}|34\rangle \quad (2)$$

denotes the antisymmetrized interaction matrix element between pairs of Hartree-Fock Bloch states. The numbers 1 to 4 stand for sets of quantum numbers needed to describe a Bloch state, i.e., they stand for a band index and a wave vector in the first Brillouin zone. $\Gamma^{(0)}$ is the simplest possible irreducible polarization diagram. $G_0(5,6,\omega)$ is the noninteracting, zero-temperature, two-particle Green's function,

$$G_0(5,6,\omega)=[\omega-(E_6-E_5)+i\eta]^{-1}$$

if 5 is occupied and 6 is unoccupied; E_5 and E_6 are Hartree-Fock energy bands, η is a positive infinitesimal, $G_0(5,6,\omega)=G_0(6,5,-\omega)$, and $G_0(5,6,\omega)=0$ if both 5 and 6 are occupied or unoccupied. To specify the interaction term of the EME further, we must give the numbers 1–4 as follows: $1=(v,\mathbf{k}-\mathbf{K})$, $2=(c,\mathbf{k}')$, $3=(v,\mathbf{k}'-\mathbf{K})$, and $4=(c,\mathbf{k})$, where $\mathbf{K}$ is the total quasimomentum of the exciton, and $\mathbf{k}$ and $\mathbf{k}'$ are wave vectors in the first Brillouin zone. It is easy to show that for any combination of band indices $\Gamma^{(0)}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4)$ satisfies the symmetry property that it vanishes unless $\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}_3+\mathbf{k}_4$. With this and the above properties of G_0, the interaction term of the EME can be written as

$$\Gamma_{vcvc}(\mathbf{k}-\mathbf{K},\mathbf{k}',\mathbf{k}'-\mathbf{K},\mathbf{k},\omega)=\Gamma_{vcvc}^{(0)}(\mathbf{k}-\mathbf{K},\mathbf{k}',\mathbf{k}'-\mathbf{K},\mathbf{k})+\sum_{c_2,\mathbf{k}_2}\Gamma_{vccc_2}^{(0)}(\mathbf{k}_2-\mathbf{q},\mathbf{k}',\mathbf{k},\mathbf{k}_2)G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)\Gamma_{vc_2vv}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)+\sum_{c_2,\mathbf{k}_2}\Gamma_{c_2ccv}^{(0)}(\mathbf{k}_2-\mathbf{q},\mathbf{k}',\mathbf{k},\mathbf{k}_2)G_0([v,\mathbf{k}_2],[c_2,\mathbf{k}_2-\mathbf{q}],-\omega)\Gamma_{vvvc_2}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega), \quad (3)$$

where we have written the band indices as subscripts, the quasimomenta we included in the arguments of the functions Γ, and where $\mathbf{q}=\mathbf{k}'-\mathbf{k}$. To complete the set of formal equations let us use (1) again to write out the equations that determine $\Gamma_{vc_2vv}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)$ and $\Gamma_{vvvc_2}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)$ occurring on the right-hand side of (3). They are

$$\Gamma_{vc_2vv}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)=\Gamma_{vc_2vv}^{(0)}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q})+\sum_{c_3,k_3}\Gamma_{vc_2vc_3}^{(0)}(\mathbf{k}_3-\mathbf{q},\mathbf{k}_2,\mathbf{k}_2-\mathbf{q},\mathbf{k}_3)G_0([v,\mathbf{k}_3-\mathbf{q}],[c_3,\mathbf{k}_3],\omega)\times\Gamma_{vc_3vv}(\mathbf{k}-\mathbf{K},\mathbf{k}_3,\mathbf{k}'-\mathbf{K},\mathbf{k}_3-\mathbf{q},\omega)+\sum_{c_3,\mathbf{k}_3}\Gamma_{c_3c_2vv}^{(0)}(\mathbf{k}_3-\mathbf{q},\mathbf{k}_2,\mathbf{k}_2-\mathbf{q},\mathbf{k}_3)G_0([v,\mathbf{k}_3],[c_3,\mathbf{k}_3-\mathbf{q}],-\omega)\times\Gamma_{vvvc_3}(\mathbf{k}-\mathbf{K},\mathbf{k}_3,\mathbf{k}'-\mathbf{K},\mathbf{k}_3-\mathbf{q},\omega); \quad (4)$$

[12] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1966).
[13] J. K. Kübler, Z. Physik **201**, 172 (1967).

$$\Gamma_{vvvc_2}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)$$
$$=\Gamma_{vvvc_2}^{(0)}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q})+\sum_{c_3,k_3}\Gamma_{vvc_2c_3}^{(0)}(\mathbf{k}_3-\mathbf{q},\mathbf{k}_2,\mathbf{k}_2-\mathbf{q},\mathbf{k}_3)G_0([v,\mathbf{k}_3-\mathbf{q}],[c_3,\mathbf{k}_3],\omega)$$
$$\times\Gamma_{vc_3vv}(\mathbf{k}-\mathbf{K},\mathbf{k}_3,\mathbf{k}'-\mathbf{K},\mathbf{k}_3-\mathbf{q},\omega)+\sum_{c_3,\mathbf{k}_3}\Gamma_{c_3vc_2v}^{(0)}(\mathbf{k}_3-\mathbf{q},\mathbf{k}_2,\mathbf{k}_2-\mathbf{q},\mathbf{k}_3)G_0([v,\mathbf{k}_3],[c_3,\mathbf{k}_3-\mathbf{q}],-\omega)$$
$$\times\Gamma_{vvvc_3}(\mathbf{k}-\mathbf{K},\mathbf{k}_3,\mathbf{k}'-\mathbf{K},\mathbf{k}_3-\mathbf{q},\omega). \quad (5)$$

Note that the set of Eqs. (4) and (5) is closed; it can in principle be solved for Γ_{vc_2vv} and Γ_{vvvc_2}, which then should be substituted into (3) to give the screened electron-hole interaction in momentum space. We shall follow this route approximately.

B. Screening Function in the Hartree-Fock Approximation

If we replace the screened scattering matrix elements Γ_{vc_2vv} and Γ_{vvvc_2} occurring on the right-hand side of (3) by the corresponding unscreened ones, which we shall do in this section, we do not take into account multiple-scattering processes, i.e., we do not sum the infinite polarization series, but rather replace it by two terms. The right-hand side of (3) can then be evaluated by approximating the matrix elements: First we Fourier transform $1/|\mathbf{r}-\mathbf{r}'|$, and then use the known symmetry properties of the Bloch functions neglecting sums over reciprocal lattice vectors. The matrix element $\Gamma_{vcvc}^{(0)}(\mathbf{k}-\mathbf{K},\mathbf{k}',\mathbf{k}'-\mathbf{K},\mathbf{k})$ then becomes

$$\Gamma_{vcvc}^{(0)}(\mathbf{k}-\mathbf{K},\mathbf{k}',\mathbf{k}'-\mathbf{K},\mathbf{k})$$
$$=-(4\pi/a^3)q^{-2}\langle v(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}'-\mathbf{K})\rangle$$
$$\times\langle c(\mathbf{k}')|e^{i\mathbf{q}\cdot\mathbf{r}}|c(\mathbf{k})\rangle$$
$$+(4\pi/a^3)K^{-2}\langle v(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{K}\cdot\mathbf{r}}|c(\mathbf{k})\rangle$$
$$\times\langle c(\mathbf{k}')|e^{i\mathbf{K}\cdot\mathbf{r}}|v(\mathbf{k}'-\mathbf{K})\rangle, \quad (6)$$

where a^3 is the volume of the unit cell. In the terminology used in the theory of excitons,[8] we call the first term the Wannier interaction and the second the Frenkel interaction, even though the second term is only approximately the usual Frenkel term. If we now remember that

$$\langle 1(\mathbf{k}-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|2(\mathbf{k})\rangle$$
$$=\delta_{1,2}-\mathbf{q}\cdot\langle\nabla_{\mathbf{k}}u_1(\mathbf{k})|u_2(\mathbf{k})\rangle+\cdots, \quad (7)$$

where $u_1(\mathbf{k})$ is the crystal-periodic function of quasi-momentum $\mathbf{k}$ and band index 1, we can see at once that for small $\mathbf{q}$ the Frenkel term is negligible compared with the Wannier term. The remaining matrix elements of (3) are similarly found to be

$$\Gamma_{1234}^{(0)}(\mathbf{k}_2-\mathbf{q},\mathbf{k}',\mathbf{k},\mathbf{k}_2)$$
$$=(4\pi/a^3)q^{-2}\langle 1(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|4(\mathbf{k}_2)\rangle\langle 2(\mathbf{k}')|e^{i\mathbf{q}\cdot\mathbf{r}}|3(\mathbf{k})\rangle$$
$$-(4\pi/a^3)|\mathbf{k}'-\mathbf{k}_2|^{-2}\langle 1(\mathbf{k}_2-\mathbf{q})|e^{-i(\mathbf{k}'-\mathbf{k}_2)\cdot\mathbf{r}}|3(\mathbf{k})\rangle$$
$$\times\langle 2(\mathbf{k}')|e^{i(\mathbf{k}'-\mathbf{k}_2)\cdot\mathbf{r}}|4(\mathbf{k}_2)\rangle \quad (8)$$

and

$$\Gamma_{1234}^{(0)}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q})$$
$$=-(4\pi/a^3)q^{-2}\langle 1(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{q}\cdot\mathbf{r}}|3(\mathbf{k}'-\mathbf{K})\rangle$$
$$\times\langle 2(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|4(\mathbf{k}_2-\mathbf{q})\rangle+(4\pi/a^3)|\mathbf{K}+\mathbf{k}_2-\mathbf{k}'|^2$$
$$\times\langle 1(\mathbf{k}-\mathbf{K})|e^{-i(\mathbf{K}+\mathbf{k}_2-\mathbf{k}')\cdot\mathbf{r}}|4(\mathbf{k}_2-\mathbf{q})\rangle$$
$$\times\langle 2(\mathbf{k}_2)|e^{i(\mathbf{K}+\mathbf{k}_2-\mathbf{k}')\cdot\mathbf{r}}|3(\mathbf{k}'-\mathbf{K})\rangle, \quad (9)$$

where in (8) $1234=vccc_2$ or c_2ccv, and in (9) $1234=vc_2vv$ or $vvvc_2$. Again with (7), the second term of (8) is negligible compared with the first and so is the second term of (9) compared with the first. Hence substituting the dominant terms of (6), (8), and (9) into (3), we get

$$\Gamma_{vcvc}(\mathbf{k}-\mathbf{K},\mathbf{k}',\mathbf{k}'-\mathbf{K},\mathbf{k},\omega)=-(4\pi/a^3)q^{-2}[\epsilon(\mathbf{q},\omega)]^{-1}$$
$$\times\langle v(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}'-\mathbf{K})\rangle\langle c(\mathbf{k}')|e^{i\mathbf{q}\cdot\mathbf{r}}|c(\mathbf{k})\rangle,$$

where the screening function is

$$[\epsilon(\mathbf{q},\omega)]^{-1}=1-4\pi\alpha(\mathbf{q},\omega), \quad (10)$$

with

$$4\pi\alpha(\mathbf{q},\omega)=-\frac{4\pi}{a^3}\frac{1}{q^2}\sum_{c_2,\mathbf{k}_2}\{|\langle c_2(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle|^2$$
$$\times G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)+|\langle v(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2-\mathbf{q})\rangle|^2$$
$$\times G_0([v,\mathbf{k}_2],[c_2,\mathbf{k}_2-\mathbf{q}],-\omega)\}. \quad (11)$$

This is seen to be the usual quantum-mechanical expression for the polarizability at $T=0$ (filled valence band and empty conduction bands), and (10) is identical with the usual Hartree-Fock dielectric function.[10]

C. Screening Function beyond Hartree-Fock

If we were justified to use $\Gamma_{vc_2vv}=\Gamma_{vc_2vv}^{(0)}/\epsilon(\mathbf{q},\omega)$ for the screened scattering matrix elements on the right-hand side of (3) and similarly for Γ_{vvvc_2} it is seen at once that we would obtain the usual RPA result,[10] i.e.,

$$\epsilon(\mathbf{q},\omega)=1+4\pi\alpha(\mathbf{q},\omega), \quad (12)$$

where the polarizability $4\pi\alpha$ is given by (11). But the particular sequence of indices of the screened matrix element Γ_{vc_2vv} indicates that this term corresponds to the scattering of a hole with the creation of an electron-hole pair, whereas Γ_{vcvc} corresponds to scattering of an electron with a hole. Since these are physically different processes, there is no reason to use the same screening function for both. This is also borne out by the formal equations, in particular (4) and (5). Let us use the same approximations as before to write out the matrix elements $\Gamma^{(0)}$ that occur in (4) and (5).

$$\Gamma_{vc_2vc_3}{}^{(0)}(\mathbf{k}_3-\mathbf{q},\mathbf{k}_2,\mathbf{k}_2-\mathbf{q},\mathbf{k}_3)$$
$$=(4\pi/a^3)q^{-2}\langle v(\mathbf{k}_3-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_3(\mathbf{k}_3)\rangle$$
$$\times\langle c_2(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle$$
$$-(4\pi/a^3)|\mathbf{k}_2-\mathbf{k}_3|^{-2}\langle v(\mathbf{k}_3-\mathbf{q})|e^{-i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle$$
$$\times\langle c_2(\mathbf{k}_2)|e^{i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|c_3(\mathbf{k}_3)\rangle \quad (13)$$

and

$$\Gamma_{c_3c_2vv}{}^{(0)}(\mathbf{k}_3-\mathbf{q},\mathbf{k}_2,\mathbf{k}_2-\mathbf{q},\mathbf{k}_3)$$
$$=(4\pi/a^3)q^{-2}\langle c_3(\mathbf{k}_3-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_3)\rangle$$
$$\times\langle c_2(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle$$
$$-(4\pi/a^3)|\mathbf{k}_2-\mathbf{k}_3|^{-2}\langle c_3(\mathbf{k}_3-\mathbf{q})|e^{-i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle$$
$$\times\langle c_2(\mathbf{k}_2)|e^{i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_3)\rangle. \quad (14)$$

$\Gamma_{c_3vc_2v}{}^{(0)}$ and $\Gamma_{vvc_2c_3}{}^{(0)}$ can be obtained from (13) and (14) by appropriately interchanging band indices. Let us limit our attention to (13). The first term we should again call the Frenkel interaction and the second the Wannier interaction. By means of (7) we see that for small $\mathbf{q}$ the Frenkel term has a finite limit and since $\mathbf{k}_3$ and $\mathbf{k}_2$ are anywhere in the Brillouin zone, the Wannier term is in general of comparable order of magnitude. Thus the scattering of the hole creates an electron-hole pair whose radius can be of any range; i.e., it is neither short nor long range. All the matrix elements $\Gamma^{(0)}$ of (4) and (5) can be analyzed and interpreted in a similar way and in each case it is seen that we need to retain both terms. Substituting these terms into (4) and (5) and replacing the screened matrix elements by the corresponding unscreened ones, we can see without further calculations that (4) and (5) do not give the Hartree-Fock screening (10) for Γ_{vc_2vv} and Γ_{vvvc_2}. Thus the usual RPA screening functions cannot be obtained either.

Let us therefore make the following ansatz:

$$\Gamma_{vc_2vv}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)=-\kappa_1^{-1}(c_2,\mathbf{k}_2,\mathbf{q},\omega)$$
$$\times(4\pi/a^3)q^{-2}\langle v(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}'-\mathbf{K})\rangle$$
$$\times\langle c_2(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle \quad (15)$$

and

$$\Gamma_{vvvc_2}(\mathbf{k}-\mathbf{K},\mathbf{k}_2,\mathbf{k}'-\mathbf{K},\mathbf{k}_2-\mathbf{q},\omega)=-\kappa_2^{-1}(c_2,\mathbf{k}_2,\mathbf{q},\omega)$$
$$\times(4\pi/a^3)q^{-2}\langle v(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}'-\mathbf{K})\rangle$$
$$\times\langle v(\mathbf{k}_2)|e^{i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2-\mathbf{q})\rangle. \quad (16)$$

Here we have defined screening functions κ_1 and κ_2 that must be determined from Eqs. (4) and (5). Let us suppose for a moment that this can be accomplished. With (15) and (16) we can then write out Γ_{vcvc} of Eq. (3) and thus determine the electron-hole screening function. In terms of κ_1 and κ_2 we find that

$$\Gamma_{vcvc}(\mathbf{k}-\mathbf{K},\mathbf{k}',\mathbf{k}'-\mathbf{K},\mathbf{k},\omega)=-(4\pi/a^3)q^{-2}$$
$$\times\langle v(\mathbf{k}-\mathbf{K})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}'-\mathbf{K})\rangle\langle c(\mathbf{k}')|e^{i\mathbf{q}\cdot\mathbf{r}}|c(\mathbf{k})\rangle$$
$$\times[1+F_1(\mathbf{q},\omega)+F_2(\mathbf{q},\omega)], \quad (17)$$

where

$$F_1(\mathbf{q},\omega)=\frac{4\pi}{a^3}\frac{1}{q^2}\sum_{c_2,\mathbf{k}_2}|\langle v(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2)\rangle|^2$$
$$\times G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)\kappa_1^{-1}(c_2,\mathbf{k}_2,\mathbf{q},\omega) \quad (18)$$

and

$$F_2(\mathbf{q},\omega)=\frac{4\pi}{a^3}\frac{1}{q^2}\sum_{c_2,\mathbf{k}_2}|\langle c_2(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_2)\rangle|^2$$
$$\times G_0([v,\mathbf{k}_2],[c_2,\mathbf{k}_2-\mathbf{q}],-\omega)\kappa_2^{-1}(c_2,\mathbf{k}_2,\mathbf{q},\omega). \quad (19)$$

Hence the electron-hole screening function is

$$1/\epsilon(\mathbf{q},\omega)=1+F_1(\mathbf{q},\omega)+F_2(\mathbf{q},\omega). \quad (20)$$

It is too hard to solve (4) and (5) directly for the functions κ_1 and κ_2. But it turns out that we can obtain approximate expressions for the functions F_1 and F_2, and this is really all we need.

To do this we substitute (15) and (16) into (4); the resulting equation we then multiply on both sides by

$$(4\pi/a^3)q^{-2}|\langle v(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2)\rangle|^2$$
$$\times G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)$$

and sum over c_2, $\mathbf{k}_2$. After some cancellations we obtain

$$F_1(\mathbf{q},\omega)=-4\pi\alpha_1(\mathbf{q},\omega)[1+F_1(\mathbf{q},\omega)+F_2(\mathbf{q},\omega)]-(4\pi/a^3)^2q^{-2}\sum_{c_2,c_3,\mathbf{k}_2,\mathbf{k}_3}|\mathbf{k}_2-\mathbf{k}_3|^{-2}\{\kappa_1^{-1}(c_3,\mathbf{k}_3,\mathbf{q},\omega)$$
$$\times\langle v(\mathbf{k}_3-\mathbf{q})|e^{-i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle\langle c_2(\mathbf{k}_2)|e^{i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|c_3(\mathbf{k}_3)\rangle\langle c_3(\mathbf{k}_3)|e^{i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_3-\mathbf{q})\rangle\langle v(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2)\rangle$$
$$\times G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)G_0([v,\mathbf{k}_3-\mathbf{q}],[c_3,\mathbf{k}_3],\omega)+\langle c_3(\mathbf{k}_3-\mathbf{q})|e^{-i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle\langle c_2(\mathbf{k}_2)|e^{i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_3)\rangle$$
$$\times\langle v(\mathbf{k}_3)|e^{i\mathbf{q}\cdot\mathbf{r}}|c_3(\mathbf{k}_3-\mathbf{q})\rangle\langle v(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2)\rangle$$
$$\times G_0([v,\mathbf{k}_3],[c_3,\mathbf{k}_3-\mathbf{q}],-\omega)G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)\kappa_2^{-1}(c_3,\mathbf{k}_3,\mathbf{q},\omega)\}, \quad (21)$$

where we have defined

$$4\pi\alpha_1(\mathbf{q},\omega)=-(4\pi/a^3)q^{-2}\sum_{c_2,\mathbf{k}_2}|\langle v(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2)\rangle|^2G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)$$

and soon shall define $4\pi\alpha_2(\mathbf{q},\omega)$ such that $4\pi\alpha=4\pi\alpha_1+4\pi\alpha_2$. Now we make the approximation that $\kappa_1(c,\mathbf{k},\mathbf{q},\omega)$ and $\kappa_2(c,\mathbf{k},\mathbf{q},\omega)$ depend only weakly on the momentum $\mathbf{k}$ and take it out of the last two sums of Eq. (21). Using the definition of F_1 and F_2, we obtain

$$F_1(\mathbf{q},\omega)(1+4\pi\alpha_1(\mathbf{q},\omega)-\gamma_1(\mathbf{q},\omega)/8\pi\alpha_1(\mathbf{q},\omega))+F_2(\mathbf{q},\omega)(4\pi\alpha_1(\mathbf{q},\omega)-\gamma_2(\mathbf{q},\omega)/8\pi\alpha_2(\mathbf{q},\omega))=-4\pi\alpha_1(\mathbf{q},\omega), \quad (22)$$

and a second equation which follows in precisely the same way from (5):

$$F_1(\mathbf{q},\omega)(4\pi\alpha_2(\mathbf{q},\omega)-\gamma_2^*(\mathbf{q},\omega)/8\pi\alpha_1(\mathbf{q},\omega))+F_2(\mathbf{q},\omega)(1+4\pi\alpha_2(\mathbf{q},\omega)-\gamma_1(-\mathbf{q},-\omega)/8\pi\alpha_2(\mathbf{q},\omega))=-4\pi\alpha_2(\mathbf{q},\omega), \quad (23)$$

where

$$\gamma_1(\mathbf{q},\omega)=2\left(\frac{4\pi}{a^3}\right)^2\frac{1}{q^2}\sum_{c_2,c_3,\mathbf{k}_2,\mathbf{k}_3}|\mathbf{k}_2-\mathbf{k}_3|^{-2}\langle v(\mathbf{k}_3-\mathbf{q})|e^{-i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|v(\mathbf{k}_2-\mathbf{q})\rangle\langle c_2(\mathbf{k}_2)|e^{i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{r}}|c_3(\mathbf{k}_3)\rangle$$
$$\times\langle c_3(\mathbf{k}_3)|e^{i\mathbf{q}\cdot\mathbf{r}}|v(\mathbf{k}_3-\mathbf{q})\rangle\langle v(\mathbf{k}_2-\mathbf{q})|e^{-i\mathbf{q}\cdot\mathbf{r}}|c_2(\mathbf{k}_2)\rangle G_0([v,\mathbf{k}_2-\mathbf{q}],[c_2,\mathbf{k}_2],\omega)G_0([v,\mathbf{k}_3-\mathbf{q}],[c_3,\mathbf{k}_3],\omega). \quad (24)$$

γ_2 is given by a similar formula obtainable from (24) by changing $v(\mathbf{k}_3-\mathbf{q})$ to $c_3(\mathbf{k}_3-\mathbf{q})$ and $c_3(\mathbf{k}_3)$ to $v(\mathbf{k}_3)$. The algebraic equations (22) and (23) can now be solved and substituted into (20) to find the electron-hole screening function.

$$\epsilon(\mathbf{q},\omega)=1+\frac{1}{2}\left[8\pi\alpha(\mathbf{q},\omega)-\frac{\alpha_1(\mathbf{q},\omega)}{\alpha_2(\mathbf{q},\omega)}\gamma_1(-\mathbf{q},-\omega)-\frac{\alpha_2(\mathbf{q},\omega)}{\alpha_1(\mathbf{q},\omega)}\gamma_1(\mathbf{q},\omega)+\gamma_2^*(\mathbf{q},\omega)+\gamma_2(\mathbf{q},\omega)\right]$$
$$\times\left[\left(1-\frac{\gamma_1(\mathbf{q},\omega)}{8\pi\alpha_1(\mathbf{q},\omega)}\right)\left(1-\frac{\gamma_1(-\mathbf{q},-\omega)}{8\pi\alpha_2(\mathbf{q},\omega)}\right)-\frac{|\gamma_2(\mathbf{q},\omega)|^2}{(8\pi)^2\alpha_1(\mathbf{q},\omega)\alpha_2(\mathbf{q},\omega)}\right]^{-1}. \quad (25)$$

Putting $\gamma_1=\gamma_2=0$ reduces $\epsilon(\mathbf{q},\omega)$ to the usual RPA function. The origin of the terms γ_1 and γ_2 are, of course, the Wannier terms in Eqs. (13), (14), etc., which we have kept. Neglecting γ_1 and γ_2 for crystals with a small band gap is not justifiable as we shall show next. Unfortunately, we can in this paper only study a rather special case of $\epsilon(\mathbf{q},\omega)$.

III. EVALUATION OF THE SCREENING FUNCTION FOR $q=\omega=0$

We have found that in $\mathbf{k}$ space the interaction of an electron with a hole in the Wannier approximation is the Coulomb attraction divided by $\epsilon(\mathbf{q},\omega)$ of Eq. (25), except that we have not yet specified the frequency ω. It can be justified that the static approximation $\omega=0$ is not a bad approximation for our problem.[12,13] Therefore, we first put $\omega=0$ in (25). Furthermore, we suppose that the $\mathbf{q}$ dependence of $\epsilon(\mathbf{q},0)$ does not alter the Coulomb interaction between the electron and hole significantly, an assumption which must of course be justified by further investigations. Hence it seems worthwhile to study first the much simpler case of the screening function for $\mathbf{q}=\omega=0$: $\epsilon(0,0)$. From (25),

$$\epsilon(0,0)=1+4\pi\alpha(0,0)$$
$$\times[1-(\gamma_1(0,0)+\gamma_2(0,0))/4\pi\alpha(0,0)]^{-1}. \quad (26)$$

The binding energy of the exciton is then of the order of $E_B=\mu/[2\epsilon(0,0)]^2$ (our units are $2m_e=\hbar=e=1$). Furthermore, let the energy gap $G>0$ be decreasing and direct, but let the first interband transition be forbidden. To have an example in mind, one may assume that the degenerate conduction and valence bands of α-Sn could be split. One could object that this is not a case of interest as far as the theory of the excitonic insulator is concerned, where one considers indirect band gaps. It still is an important case and, besides being readily solvable, has been chosen for the following reason.

The dielectric function of a zero-gap semiconductor like α-Sn has received some attention recently.[14,15] We have complemented these calculations for small energy gaps[16] and the polarization was found to be given for $G\to 0$ and $\mathbf{q}=0$ by

$$4\pi\alpha(0,0)=4\pi\alpha_0+\tfrac{3}{8}(\mu/G)^{1/2}, \quad (27)$$

where $\mu^{-1}=m_c^{-1}+m_v^{-1}$, m_c and m_v being the effective electron and hole masses, respectively. $4\pi\alpha_0$ is the background polarizability due to all bands other than the first conduction band and it can therefore be assumed to be a constant. If the electron-hole screening is given by the dielectric constant in RPA, then in view of the $G^{-1/2}$ dependence of (27) and the $\epsilon(0,0)^{-2}$ dependence of the binding energy of the exciton, there must exist a critical gap where $G=E_B$; i.e., even this case will exhibit an excitonic instability. With the data of α-Sn,[14] this critical gap is of the order of a few 10^{-3} eV only and therefore of no practical significance, even if the degenerate bands of α-Sn could be split. But it means that α-Sn should be in the excitonic insulator phase (provided the band structure of α-Sn is as simple as assumed here, which is doubtful). From all this it seems theoretically interesting to study the binding energy of the exciton using Eq. (26) for $\epsilon(0,0)$. Hence we evaluate the terms γ_1 and γ_2 using techniques developed in Ref. 16. First we set $\omega=0$ in (24) and then perform the limit $\mathbf{q}\to 0$ using Eq. (7). We then rewrite the result with the elementary relation

$$\langle u_v(\mathbf{k})|\nabla_{\mathbf{k}}u_c(\mathbf{k})\rangle=2\langle u_v(\mathbf{k})|\mathbf{p}|u_c(\mathbf{k})\rangle[E_c(\mathbf{k})-E_v(\mathbf{k})],$$

where $\mathbf{p}=-i\nabla_{\mathbf{r}}$, and $u_i(\mathbf{k})$ are the crystal-periodic functions. The matrix element $\langle u_v(\mathbf{k})|\mathbf{p}|u_c(\mathbf{k})\rangle$ is then expanded in a power series in $\mathbf{k}$, keeping only the first

[14] L. Liu and D. Brust, Phys. Rev. Letters **20**, 651 (1968); Phys. Rev. **173**, 777 (1968).
[15] D. Sherrington and W. Kohn, Phys. Rev. Letters **21**, 153 (1968).
[16] J. K. Kübler, Phys. Status Solidi **31**, 17 (1969).

nonvanishing term. We have shown earlier[16] that this gives the leading terms of integrals like those involved in γ_1 and γ_2 for $G\to 0$. The expansion coefficient $\boldsymbol{\eta}_{vc}$ can be estimated as in our earlier calculation of the polarizability by comparison with calculations of the polarizability of α-Sn,[14,15] and we set $|\boldsymbol{\eta}_{vc}|^2=\frac{3}{2}\mu^2$. The remaining integrations for γ_1 and γ_2 can now be done; the result is

$$\gamma_1(0,0)=\frac{3\mu^2}{\pi^2}\Bigg[-\frac{1}{2}\Bigg(\frac{k_B}{\mu G+k_B{}^2}-\frac{1}{\sqrt{(\mu G)}}\tan^{-1}\frac{k_B}{\sqrt{(\mu G)}}\Bigg)^2+\frac{3\pi}{8\sqrt{(\mu G)}}\Bigg(\frac{1}{\sqrt{(\mu G)}}\tan^{-1}\frac{k_B}{\sqrt{(\mu G)}}-\frac{5k_B{}^3+3k_B\mu G}{3(\mu G+k_B{}^2)^2}\Bigg)\Bigg]+\gamma_{10}$$

and

$$\gamma_2(0,0)=\frac{3\mu^2}{80\pi^2}\Bigg(\frac{1}{\sqrt{(\mu G)}}\tan^{-1}\frac{k_B}{\sqrt{(\mu G)}}-\frac{k_B}{\mu G+k_B{}^2}\Bigg)\times\Bigg(\frac{15}{\sqrt{(\mu G)}}\tan^{-1}\frac{k_B}{\sqrt{(\mu G)}}-\frac{33k_B{}^5+40\mu G k_B{}^3+15\mu^2G^2k_B}{(\mu G+k_B{}^2)^3}\Bigg)+\gamma_{20},$$

where γ_{10} and γ_{20} are due to all bands other than the first conduction band and can be treated as constants, and where k_B has been chosen to give an integration volume equal to the Brillouin-zone size. With (27) the quantities $\gamma_1(0,0)/4\pi\alpha(0,0)$ and $\gamma_2(0,0)/4\pi\alpha(0,0)$ are then found to be in leading order for small gaps

$$\frac{\gamma_1(0,0)}{4\pi\alpha(0,0)}=-\frac{1}{2}\Big(\frac{\mu}{G}\Big)^{1/2}-\frac{\mu}{\pi k_B}-\frac{\alpha_0}{2\alpha(0,0)}\Big(\frac{\mu}{G}\Big)^{1/2}$$

and

$$\frac{\gamma_2(0,0)}{4\pi\alpha(0,0)}=-\frac{3}{8}\Big(\frac{\mu}{G}\Big)^{1/2}+\frac{\mu}{5\pi k_B}-4\pi\alpha_0.$$

Hence the denominator of (26) becomes

$$1-\frac{(\gamma_1(0,0)+\gamma_2(0,0))}{4\pi\alpha(0,0)}=\frac{28\pi}{3}\alpha_0-\frac{7}{8}\Big(\frac{\mu}{G}\Big)^{1/2}.$$

Thus for a critical gap $G=E_c$ the denominator vanishes:

$$E_c=(9/16)\mu/4(4\pi\alpha_0)^2.$$

For $G=E_c$ the screening function is infinite and it is negative for $G<E_c$. We see that E_c is roughly equal to the binding energy of an improperly screened exciton, namely, an exciton screened with the background dielectric constant only. Of course, the binding energy of the real exciton, being proportional to $\epsilon(0,0)^{-2}$, goes to zero as G approaches E_c. Thus there is no gap such that $G=E_B$, and hence there is no excitonic instability in this case.

IV. DISCUSSION AND CONCLUSION

By summing an infinite but partial polarization series, treating Coulomb and exchange on an equal footing, we have obtained a screened electron-hole interaction for small-gap semiconductors from first principles. The electron-hole interaction was found to be the Coulomb attraction divided by $\epsilon(\mathbf{q},0)$ [Eq. (25)] in the static approximation ($\omega=0$). Our screening function $\epsilon(\mathbf{q},0)$ differs from the usual RPA dielectric function, since we have taken into account that the radii of electron-hole pairs created virtually in the screening process are in general of intermediate range. We have evaluated $\epsilon(\mathbf{q},0)$ for $\mathbf{q}=0$ and have studied the binding energy of the exciton, E_B, as a function of the gap G for a band structure where the energy gap is direct but the first interband transition is forbidden. The zero-gap case of this model is α-Sn, idealizing the band structure of gray tin somewhat. We found that there does not exist an excitonic instability which would be predicted on the basis of the usual dielectric constant in RPA. The binding energy was rather found to vanish for energy gaps of the order of the binding energy of an exciton screened with the background dielectric constant. This is not too surprising since our formalism treats the correlated electron-hole pairs in an approximately self-consistent fashion. For an energy gap smaller than the critical gap (where $E_B=0$) the screening constant becomes negative, which would indicate that electrons repel holes. This is, of course, very unlikely and we conclude that our calculation ceases to be valid for gaps smaller than the critical gap. We take this to indicate a phase change of some sort the discussion of which we leave open. We suspect the situation to be similar for indirect gaps; i.e., we suspect there is a phase transition at a critical gap $G=E_c$. Perhaps E_c may be obtainable with the background dielectric constant as above. But by virtue of $E_B=0$ for $G=E_c$ the new phase is not necessarily of the nature of the excitonic insulator. In connection with heavily doped semiconductors, we may speculate that this behavior of the binding energy explains the vanishing of the activation energy ϵ_2 as a function of increased doping in, e.g., n-type germanium.[17] Of course, this is a speculation because we cannot say if the random nature of the impurity centers (producing impurity bands) alters the dielectric screening in an essential way.

[17] E. A. Davis and W. D. Compton, Phys. Rev. **140**, A2183 (1965).