PHYSICAL REVIEW B VOLUME 35, NUMBER 5 15 FEBRUARY 1987-I

Effect of disorder on exciton binding in semiconductor alloys

M. A. Kanehisa
Laboratoire de Physique des Solides, Université Pierre et Marie Curie, Tour 13, 4 place Jussieu, 75252 Paris Cédex 05, France

R. J. Elliott
Department of Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, England, United Kingdom

(Received 23 April 1986)

We apply the coherent-potential approximation to the Wannier exciton and show that broadening of the conduction and valence bands, which pushes down the absorption edge relative to the virtual-crystal band position, has less influence on the exciton states, thus reducing the binding energy. We also show that the effective interaction induced by simultaneous scattering of the electron and hole by the same impurity can be either attractive or repulsive depending on the band lineup. In most semiconductor alloys, this turns out to be attractive and cancels partially the binding-energy reduction due to bowing. These results are obtained by an approximate solution to the Bethe-Salpeter equation appropriate to a disordered system. Our theory is compared with available experimental data in Ga (As,P).

I. INTRODUCTION

Optical properties of semiconductor alloys are of great technological interest. The low-energy edge of absorption spectra is determined by excitons—bound electron-hole pairs. It is therefore interesting to investigate the behavior of the exciton in disordered systems. One of the fundamental questions in this problem, whether disorder enhances the electron-hole binding, is addressed in this paper.

We shall consider an exciton associated with the direct gap of a semiconductor alloy. Then our system consists of an electron (particle 1) in the conduction band (c) with mass m_c and a hole (particle 2) in the valence band (v) with mass m_v. Our problem is to solve the Hamiltonian

$$H=\frac{\mathbf{p}_1^2}{2m_c}+\frac{\mathbf{p}_2^2}{2m_v}+u(\mathbf{r}_1-\mathbf{r}_2)+V_c(\mathbf{r}_1)+V_v(\mathbf{r}_2)\ , \tag{1.1}$$

with both the Coulomb interaction u and the potential V_μ due to disorder ($\mu=c,v$). (We have neglected disorder-induced interband mixing.) In place of the electron-hole coordinates, $(\mathbf{r}_1,\mathbf{p}_1)$ and $(\mathbf{r}_2,\mathbf{p}_2)$, one may introduce the center-of-mass and relative coordinates, $(\mathbf{R},\mathbf{P})$ and $(\mathbf{r},\mathbf{p})$, to rewrite (1.1) as

$$H=\frac{\mathbf{p}^2}{2m_r}+u(\mathbf{r})+\frac{\mathbf{P}^2}{2M}+V_c\left[\mathbf{R}+\frac{m_v}{M}\mathbf{r}\right]+V_v\left[\mathbf{R}-\frac{m_c}{M}\mathbf{r}\right]\ , \tag{1.2}$$

where m_r and M are reduced and total masses, respectively. Because of the random potential, the translational and the relative degrees of freedom cannot be decoupled. This is the essential difficulty in the two-body problem in a disordered system.

When the exciton energy is much larger than typical disorder energy,[1] one can first look for the eigenfunction ϕ for the relative part

$$H_r=\frac{\mathbf{p}^2}{2m_r}+u(\mathbf{r})\ ,$$

and then solve, for each exciton states ϕ, the translational part

$$\overline{H}_t=\frac{\mathbf{P}^2}{2M}+\overline{V}_c(\mathbf{R})+\overline{V}_v(\mathbf{R})\ . \tag{1.3}$$

Here $\overline{V}_c$ and $\overline{V}_v$ are averages of V_c and V_v with respect to the relative state ϕ; for instance,

$$\overline{V}_c(\mathbf{R})=\int d^3r\ |\phi(\mathbf{r})|^2 V_c\left[\mathbf{R}+\frac{m_v}{M}\mathbf{r}\right]\ .$$

This approach is very similar to the Born-Oppenheimer adiabatic approximation: The translational motion in the exciton problem corresponds to the atomic motion in the molecular problem. From this analogy, one can tell immediately that this method is valid if the exciton state in question is well separated from other states so that the energy spacing is much larger than the translational width and disorder. In this case, one can neglect the relative motion and apply any single-particle alloy theory solely to their translational motion (1.3).[2,3] This holds in mixed alkali halide and probably II-VI crystals.

On the contrary, when the exciton energy is comparable to the disorder energy, the "adiabatic approximation" breaks down, and it is essential to take into account the effect of disorder both on the translational and relative motions. This is the case with the Wannier exciton in III-V alloys in which we shall be interested hereafter. In this case, as far as we know, there is no theoretical approach except to start from independent electron and hole by neglecting u in (1.1) and then to take into account the

Coulomb interaction between an average electron and an average hole. A further simplified approach adopted by Mahanti[4] in solving the Bethe-Salpeter equation is to suppose free-electron-like one-particle Green's function with a built-in width to allow for the random potential due to disorder. This, of course, is not very satisfactory except for the very-small-disorder case.

We shall show in this paper an approximate method for solving (1.1) with both the Coulomb interaction and disorder. Our paper is organized as follows. In Sec. II we define our model of disordered semiconductor and introduce the one-particle Green's function for a given configuration of alloy atoms. Then we consider the configuration average of their products. In Sec. III we analyze the Bethe-Salpeter equation for an electron and a hole in a disordered system. Section IV deals with the weak disorder case, where one can obtain alloy exciton states analytically. In Sec. V we investigate the exciton state for an arbitrary strength of disorder using the coherent-potential approximation. Section VI compares our theoretical results with available experimental data. In Sec. VII we discuss the validity of our approximation and suggest some possible extensions. Finally, Sec. VIII is a resumé of the present work.

II. THE BINARY SEMICONDUCTOR ALLOY

A. Model

The disordered system we consider is a two-band generalization of the binary-alloy model frequently used by many authors.[5] We take an alloy A_xB_{1-x}, x and $1-x$ being the concentration of A and B atoms, respectively. Each band μ ($\mu=c,v$) has its center-of-mass energy ε_A^μ or ε_B^μ depending on the site n being of A or B type. The energy difference $\delta_\mu=\varepsilon_A^\mu-\varepsilon_B^\mu$ measures the disorder. The band widths w_μ are supposed to be independent of the alloy composition x. Our model is characterized by the following six quantities: concentration x, the average gap ε_0, the conduction- and valence-band widths w_c,w_v, and their site-energy differences (disorder) δ_c,δ_v.

In this model, the average (or "virtual-crystal") gap is given by

$$E_g^{\mathrm{vc}}(x)=\varepsilon_0+(\delta_c-\delta_v)(x-\tfrac{1}{2}) . \tag{2.1}$$

We shall see in the following sections that there is a strong deviation from this linear law when higher-order effects in δ_c and δ_v are taken into account.

B. One-particle Green's function

The Green's function describing the propagation of an electron or hole is defined as usual.[6] For a given alloy configuration, it obeys the equation of motion

$$G_\mu(\varepsilon,ik)=\delta_{ik}g_\mu(\varepsilon,i)+g_\mu(\varepsilon,i)\sum_j t(i-j)G_\mu(\varepsilon,jk) , \tag{2.2}$$

where

$$g_\mu(\varepsilon,i)=1/(\varepsilon-\varepsilon_i^\mu) \tag{2.3}$$

is the "locator"[6] and $t_\mu(i-j)$ is the transfer integral between site i and j, which determines the width w_μ. Usually it is very difficult to solve (2.2) to obtain the configuration-dependent Green's function G.

Fortunately, many physical quantities can be derived from the configuration average

$$\mathcal{G}=\langle G\rangle , \tag{2.4}$$

which can be obtained more easily in approximate form. Of the numerous approaches to this problem which have been proposed over the last three decades, one of the most successful is the single-site coherent-potential approximation (CPA) which is theoretically reasonable yet numerically manageable.[6] The essence of this approach is to find an average potential (or the self-energy Σ), called "coherent potential," so that the *single-site* scattering matrix due to residual potential $\varepsilon_i^\mu-\Sigma_\mu$,

$$t_i^\mu=\frac{\varepsilon_i^\mu-\Sigma_\mu}{1-(\varepsilon_i^\mu-\Sigma_\mu)F_\mu} , \tag{2.5}$$

vanishes *on average*:

$$\langle t\rangle=0 , \tag{2.6}$$

where

$$F_\mu=\langle 0|\mathcal{G}_\mu|0\rangle \tag{2.7}$$

is the site-diagonal element of the Green's function. When the condition (2.6) is used, Eq. (2.5) gives a functional relation between Σ and $\mathcal{G}$. Combining this with the Dyson equation giving $\mathcal{G}$ in terms of Σ, one can solve it for $\mathcal{G}$ and Σ.

C. The vertex parts

In the following sections, we shall encounter averages of products of several Green's functions. It is important to note that these are different from products of averages (2.4):

$$\langle GG\rangle\neq\langle G\rangle\langle G\rangle\equiv\mathcal{G}\mathcal{G} , \tag{2.8a}$$

$$\langle GGG\rangle\neq\langle G\rangle\langle G\rangle\langle G\rangle\equiv\mathcal{G}\mathcal{G}\mathcal{G} , \tag{2.8b}$$

etc. For the two-particle average (2.8a), we define the irreducible vertex part $\mathcal{L}_{\mu\nu}$ by the equation

$$\langle G_\mu G_\nu\rangle=\mathcal{G}_\mu\mathcal{G}_\nu+\mathcal{G}_\mu\mathcal{G}_\nu\mathcal{L}_{\mu\nu}\langle G_\mu G_\nu\rangle ,$$

where the second term allows for the difference between the both sides in (2.8a). We can define higher irreducible vertex parts in a similar way.

Velický[7] expresses the two-particle irreducible vertex in terms of the average of a product of single-site t matrices:

$$\mathcal{L}_{\mu\nu}(z_1,z_2)=\frac{\langle t_\mu(z_1)t_\nu(z_2)\rangle}{1+F_\mu(z_1)\langle t_\mu(z_1)t_\nu(z_2)\rangle F_\nu(z_2)} . \tag{2.9}$$

We then need to calculate explicitly the average of product of two t matrices, $\langle t_\mu t_\nu\rangle$. When the single-site CPA condition (2.6) is used, it turns out that

$$\langle t_\mu t_\nu\rangle=\frac{\delta_\nu\Sigma_\mu-\delta_\mu\Sigma_\nu}{\delta_\mu F_\mu(1+\Sigma_\nu F_\nu)-\delta_\nu F_\nu(1+\Sigma_\mu F_\mu)} . \tag{2.10}$$

Here the self-energies Σ_μ, Σ_ν are to be measured from each band center. Equations (2.9) and (2.10) give finally the required vertex part

$$\mathscr{L}_{\mu\nu}(z_1,z_2)=\frac{\delta_\nu\Sigma_\mu(z_1)-\delta_\mu\Sigma_\nu(z_2)}{\delta_\mu F_\mu(z_1)-\delta_\nu F_\nu(z_2)} . \tag{2.11}$$

Especially, in the intraband case $\mu=\nu$, we have

$$\mathscr{L}_{\mu\mu}(z_1,z_2)=\frac{\Sigma_\mu(z_1)-\Sigma_\mu(z_2)}{F_\mu(z_1)-F_\mu(z_2)} . \tag{2.12}$$

Equations (2.11) and (2.12) tell that the two-particle vertex part is completely determined by the one-particle quantities (Σ and F). This is in accordance with the general requirement of the gauge invariance which guarantees the charge conservation.[7] In fact, (2.12) seems to be a direct manifestation of the generalized Ward identity. In the limit of $z_1 \to z_2$, this reduces to the Leath formula.[8] We also note that Abe and Toyozawa[9] have obtained a formula similar to (2.11) for a model slightly different from ours.

III. THE BETHE-SALPETER EQUATION

As is well known, exciton states are obtained as poles in the two-particle Green's function K for the electron and hole.[10,11] We can write down the Bethe-Salpeter equation obeyed by K for a given alloy configuration of atoms. In the site representation ($i,j,k,l;a,b,c,d=$sites), it reads [see Fig. 1(a)]

$$K^{(\omega)}(\varepsilon\varepsilon',ij,kl)=K_0^{(\omega)}(\varepsilon,ij,kl)2\pi\delta(\varepsilon-\varepsilon')+i\int\frac{d\varepsilon''}{2\pi}\sum_{a,b,c,d}K_0^{(\omega)}(\varepsilon,ij,ab)u(ab,cd)K^{(\omega)}(\varepsilon''\varepsilon',cd,kl) . \tag{3.1}$$

Here

$$K_0^{(\omega)}(\varepsilon,ij,kl)=G_c(\varepsilon+\omega,\,ik)G_v(\varepsilon,lj) \tag{3.2}$$

is the product of one-particle Green's functions G_c and G_v for the electron in the conduction band (c) and the hole in the valence band (v), respectively. In (3.1) we have only taken the bare Coulomb interaction [see Fig. 1(b)]

$$u(ab,cd)=v(a-b)\delta_{ac}\delta_{bd}-w(a-c)\delta_{ab}\delta_{cd} , \tag{3.3}$$

including the direct (v) and exchange (w) parts.

Equations (3.1) and (3.2) tell that, in principle, if we know the one-particle Green's functions for a particular configuration, the two-particle function K for that configuration is completely determined. This program, unfortunately, cannot usually be carried out, for the configuration-dependent quantities lack translational symmetry. (However, in the Appendix, the configuration-dependent Green's function is obtained in the special case of the atomic limit.) Thus one must have recourse to configurationally averaged quantities, which are more easily obtained in approximate form.

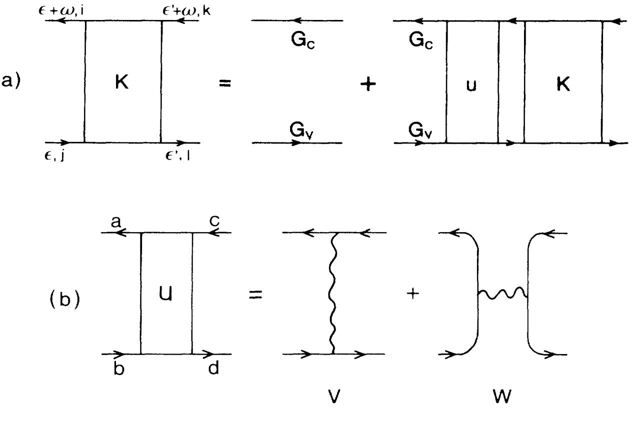

FIG. 1. (a) Diagrammatic representation of the Bethe-Salpeter equation (3.1) for an electron and hole pertinent to a particular configuration of semiconductor alloy. The solid line represents the one-particle Green's function for this configuration. The two-particle Green's function $K^{(\omega)}(\varepsilon\varepsilon',ij,kl)$ describes the process where the electron with energy $\varepsilon'+\omega$ at site k and the hole with energy ε' at site l are scattered into the electron with energy $\varepsilon+\omega$ at site i and the hole with energy ε at site j. These configuration-dependent Green's functions lack translational symmetry. (b) The kernel u of the Bethe-Salpeter equation consists of the direct (v) and the exchange (w) parts of the Coulomb interaction.

In contrast to the configuration-dependent case, the averaged two-particle Green's function

$$\mathscr{K}=\langle K\rangle$$

cannot be expressed solely in terms of the averaged one-particle Green's functions (2.4). This is because the average of Eq. (3.1) in iterated form written symbolically as

$$\mathscr{K}=\langle G_cG_v\rangle+i\langle G_cG_vuG_cG_v\rangle + i^2\langle G_cG_vuG_cG_vuG_cG_v\rangle+\cdots \tag{3.4}$$

contains averages of products of 2,4,6,... one-particle Green's functions, which are different from products of averages of Green's functions (2.4), as in (2.8). When all the vertex corrections are taken into account, (3.4) takes the form

$$\mathscr{K}=\mathscr{G}_c\mathscr{G}_v+\mathscr{G}_c\mathscr{G}_v(\mathscr{L}+iU)\mathscr{K} , \tag{3.5}$$

where $\mathscr{L}\equiv\mathscr{L}_{cv}$ is the interband irreducible vertex (2.11) and

$$U=u+(\mathscr{G}_c\mathscr{L}_{cc}\mathscr{G}_c)u+u(\mathscr{G}_v\mathscr{L}_{vv}\mathscr{G}_v) +(\mathscr{G}_cv\mathscr{L}\mathscr{G}_u)+\cdots \tag{3.6}$$

is the Coulomb interaction modified by disorder. $\mathscr{L}_{cc}$ and $\mathscr{L}_{vv}$ are *intraband* vertex parts (2.12). Since the translational symmetry is restored by averaging, it is convenient to employ the momentum representation. Using the four-vector notation, $P=(\varepsilon,\mathbf{p})$, $Q=(\omega,\mathbf{q})$, etc., we write Eq. (3.5) explicitly as

$$\mathscr{K}^{(Q)}(P,P')=\mathscr{K}_0^{(Q)}(P)\delta_{PP'}+\mathscr{K}_0^{(Q)}(P)\sum_{P''}[\mathscr{L}^{(Q)}(P,P'')+iU^{(Q)}(P,P'')]\mathscr{K}^{(Q)}(P'',P') \tag{3.7}$$

(the P'' sum includes ε'' integration). This is an integral equation involving four variables $P=(\varepsilon,\mathbf{p})$. The impurity scattering is elastic, and for the single-site CPA, which we will use subsequently, $\mathscr{L}$ does not depend on momentum:

$$\mathscr{L}^{(Q)}(P,P')=\mathscr{L}^{(\omega)}(\varepsilon)2\pi\delta(\varepsilon-\varepsilon') . \tag{3.8}$$

Further, since we are interested in concurrent effects of disorder and the Coulomb interaction in (3.7), we shall neglect in (3.6) the modification of the latter due to random potentials and take the lowest-order term, the unperturbed Coulomb potential:

$$U^{(Q)}(P,P')\approx u^{(q)}(\mathbf{p},\mathbf{p}')\equiv v_{\mathbf{p}-\mathbf{p}'}-w_{\mathbf{q}} . \tag{3.9}$$

The validity of these approximations will be discussed later. With all these simplifications, (3.7) is still an integral equation involving nontrivially four variables. Although it may not be impossible to solve it numerically, we would prefer to adopt one more assumption of short-range "Coulomb" potential

$$v_{\mathbf{p}-\mathbf{p}'}=v_0=\text{const} .$$

This assumption allows us to obtain an algebraic solution, from which the absorption coefficient is calculated. For a photon with $Q=(\omega,\mathbf{q})$, this is proportional to the imaginary part of the polarization part

$$\mathscr{P}(Q)=-i\sum_{P,P'}\mathscr{K}^{(Q)}(P,P') .$$

Without disorder, $\mathscr{G}\to G^{\text{crys}}$, $\mathscr{K}\to K^{\text{crys}}$, $\mathscr{L}\to 0$, and the solution for the polarization part $\mathscr{P}\to\Pi^{\text{crys}}$ is simply

$$\Pi^{\text{crys}}(Q)=\frac{\Pi_0^{\text{crys}}(Q)}{1+u_0^{(\mathbf{q})}\Pi_0^{\text{crys}}(Q)} , \tag{3.10}$$

where

$$u_0^{(\mathbf{q})}\equiv v_0-w_{\mathbf{q}} \tag{3.11}$$

is the coupling constant (short-range Coulomb and exchange), and

$$\Pi_0^{\text{crys}}(Q)=-i\sum_P G_c^{\text{crys}}(P+Q)G_v^{\text{crys}}(P) . \tag{3.12}$$

Equation (3.10) is quite similar to the Koster-Slater model for impurities. For each $\mathbf{q}$, bound states may occur at energy ω for which the denominator of (3.10) vanishes:

$$1+u_0^{(\mathbf{q})}\Pi_0^{\text{crys}}(\omega,\mathbf{q})=0 . \tag{3.13}$$

This gives the exciton of total momentum $\mathbf{q}$.

On the other hand, if the Coulomb coupling is absent, (3.7) yields

$$\mathscr{P}_0(Q)=-i\int\frac{d\varepsilon}{2\pi}\frac{\mathscr{R}_0^{(Q)}(\varepsilon)}{1-\mathscr{L}^{(\omega)}(\varepsilon)\mathscr{R}_0^{(Q)}(\varepsilon)} , \tag{3.14}$$

whose imaginary part is proportional to the interband optical absorption coefficient, with

$$\mathscr{R}_0^{(Q)}(\varepsilon)=\sum_{\mathbf{p}}\mathscr{G}_c(P+Q,P)\mathscr{G}_v(P) . \tag{3.15}$$

For $\mathbf{q}=\mathbf{0}$, (3.14) reduces to the result of Abe and Toyozawa.[9]

With both disorder and the Coulomb interaction, we can still solve (3.7). The result can be written as

$$\mathscr{P}(Q)=\frac{\mathscr{P}_0(Q)}{1+u_0^{(\mathbf{q})}\mathscr{P}_0(Q)} , \tag{3.16}$$

in terms of (3.14). This formula is very similar to (3.10): The effect of disorder is to replace the polarization part Π_0^{crys} by the averaged $\langle\Pi_0\rangle=\mathscr{P}_0$. This is related to our approximation (3.9) of truncating the series (3.6). We shall discuss this more in detail in Sec. VII. We also note that though the vertex function and the Coulomb interaction enter in a similar way in (3.7), they appear completely differently in the final result (3.16). This is because the vertex function is momentum-independent and diagonal in energy [there is no exchange in energy between the electron and hole by potential scattering, see (3.8)], while the Coulomb interaction does not depend on energy (static interaction).

IV. WEAK DISORDER CASE

In this section, we shall examine the consequences of Sec. III in a weak disorder case, $\delta_c,\delta_v \ll w_t,w_v$, where we can apply the perturbation theory and obtain final results analytically. But before doing this, let us first consider the case without disorder. The exciton state is determined from (3.13), and one must calculate the polarization part (3.12). We assume the Hubbard density of states[2,5] for both the conduction and valence bands with width w_c and w_v, respectively. This permits us to calculate the $\mathbf{p}$ sum of a single Green's function ($\mu=c,v$)

$$\begin{aligned}F_\mu^{\text{crys}}(z)&=\langle 0\,|\,G^{\text{crys}}(z)\,|\,0\rangle\\&=(2/w_\mu)[z-(z^2-w_\mu^2)]^{1/2} ,\end{aligned} \tag{4.1}$$

but not of the *product* of two Green's functions (3.12) without specifying the dispersion relations for the conduction and valence bands. Here we assume, following Abe and Toyozawa,[9] similar dispersion in the conduction and valence bands, which makes the $\mathbf{p}$ sum feasible for $\mathbf{q}=\mathbf{0}$. The result is

$$\Pi_0^{\text{crys}}(\omega)=(2/w)[\omega-(\omega^2-w^2)]^{1/2} , \tag{4.2}$$

with the absorption-band width given by

$$w=w_c+w_v . \tag{4.3}$$

Using the condition (3.13), one can tell that the bound state occurs when the coupling constant u_0 is sufficiently large:

$$u_0>w/2 , \tag{4.4}$$

where

$$u_0 = v_0 - w_0$$

is the coupling constant (3.11) at $\mathbf{q}=0$.[12] The binding energy (defined as the band edge minus the exciton energy) is given by

$$E_B^{\text{crys}} = u_0 + w^2/(4u_0) - w \ . \tag{4.5}$$

Let us now go back to the disordered case. To the second order in δ_μ ($\mu = c,v$), one has the Edwards formula[13] for the self-energy of the one-particle Green's functions

$$\Sigma_\mu = \bar{\varepsilon}_\mu + x(1-x)\delta_\mu^2 F_\mu \ , \tag{4.6}$$

where $\bar{\varepsilon}_\mu$ is the virtual crystal energy, and F_μ is the site-diagonal element of the Green's function (2.7). For the momentum-independent self-energy like (4.6), the Green's function $\mathscr{G}$ can be obtained by replacing z by $z-\Sigma$ in G^{crys}. Thus, using (4.1), we have

$$z - \Sigma_\mu = F_\mu^{-1} + w_\mu^2 F_\mu / 4 \ . \tag{4.7}$$

Eliminating Σ_μ from (4.6) and (4.7), one obtains a quadratic equation for F_μ, which can be easily solved. The result has the same form as (4.1) with the origin shifted by $\bar{\varepsilon}_\mu$ and the width w_μ replaced by

$$w'_\mu = w_\mu + 2x(1-x)\delta_\mu^2 / w_\mu \ .$$

The second term represents the so-called bowing. This bowing occurring for both the conduction ($\mu = c$) and valence ($\mu = v$) bands, the gap decreases by

$$E_{\text{bow}} = 2x(1-x)(\delta_c^2/w_c + \delta_v^2/w_v) \ . \tag{4.8}$$

This is the first correction to (2.1).

The polarization part $\mathscr{P}_0$ in the disordered case can be calculated easily. The result has the same form as (4.2) with the origin shifted by the second term of (2.1) and the width w replaced by

$$w' = w_c' + w_v' = w + E_{\text{bow}} \ .$$

The condition of vanishing denominator in (3.16) gives the exciton binding energy.

Besides the above effect due to the modification in single-particle motion, there arises an effective correlation between the electron and the hole because they can be scattered simultaneously by the same impurity. The magnitude of this effective interaction is the interband irreducible vertex part $\mathscr{L} \equiv \mathscr{L}_{cv}$ [Eq. (2.11)] divided by a typical width (4.3). To the lowest order in disorder, (2.11) reduces to

$$\mathscr{L} = x(1-x)\delta_c\delta_v \ , \tag{4.9}$$

which then gives the correction to the exciton binding energy in disordered system

$$E_{\text{eff}} = x(1-x)\delta_c\delta_v / w \ . \tag{4.10}$$

When this term is taken into account, one has finally the exciton binding energy in disordered solids:

$$E_B = E_B^{\text{crys}} - E_{\text{bow}} \left[1 - \frac{w}{2u_0} \right] - E_{\text{eff}} \ . \tag{4.11}$$

Here the second term represents the reduction in exciton binding due to bowing of the band edge (4.8): Owing to (4.4), the factor in the parentheses is always positive. Thus one can tell that the excitation follows only partially the bowing of the absorption edge and that this makes the exciton binding more shallow. The last term in (4.11), effective electron-hole coupling due to disorder, can be positive or negative depending on the relative sign of δ_c and δ_v [see (4.10)]. Two cases may occur in going from *A* crystal to *B* crystal: Either both the conduction and valence bands move upward (or both downward), $\delta_c\delta_v > 0$, or one band goes up while the other goes down, $\delta_c\delta_v < 0$. The latter seems to be the case in most semiconductor alloys. Then the effective interaction term $E_{\text{eff}} < 0$ *deepens* the binding.

V. EXCITON IN THE COHERENT-POTENTIAL APPROXIMATION

In order to check more quantitively the above analysis for an arbitrary strength of disorder, we have solved the two-particle problem formulated in Sec. III in the coherent-potential approximation (CPA).

We use the Hubbard density of states as in Sec. IV to calculate explicitly the one-particle Green's functions and, with the same assumption as in Sec. IV, the two-particle Green's function (3.15) for $\mathbf{q}=0$. We show in Fig. 2 the behavior of the interband vertex part (2.11) as a function of ε for a given ω at $\mathbf{q}=0$. Its real part does not vary much from the asymptotic value [which turn out to be equivalent to the second-order result (4.9)] except near resonances in the conduction and valence bands, where $\mathscr{L}$ also acquires a small imaginary part. When the denominator in (3.14) vanishes,

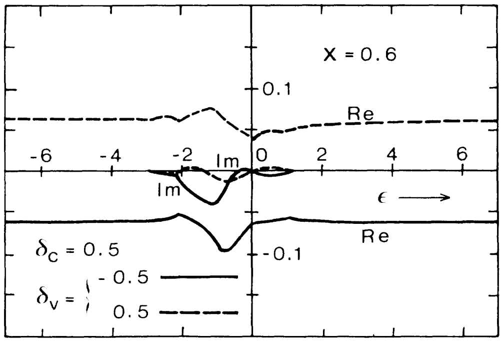

FIG. 2. Irreducible vertex part $\mathscr{L}^{(\omega)}(\varepsilon)$ for a given ω (taken as $\omega=3.03$) as a function of energy ε. Due to simultaneous scattering of the electron and the hole by the same impurity, there arises a correlation between their motions. The irreducible vertex part $\mathscr{L}^{(\omega)}(\varepsilon)$ describes this effective interaction as a function of the electron energy $\varepsilon+\omega$ and the hole energy ε. Its real part does not vary much from its asymptotic value $x(1-x)\delta_c\delta_v$ except at resonances due to conduction and valence bands, where $\mathscr{L}$ also acquires a small imaginary part. Parameters used are concentration $x=0.6$; band widths, $w_c=2$, $w_v=1$; conduction-band disorder, $\delta_c=0.5$; valence-band disorder, $\delta_v=-0.5$ (solid curve), $\delta_v=0.5$ (dashed curve). All energies are measured in the unit of valence-band width.

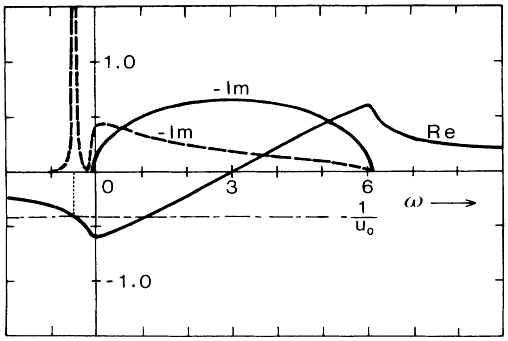

FIG. 3. Short-range exciton model in disordered semiconductors. The solid lines represent the noninteracting electron-hole polarization part $\mathscr{P}_0$, whose imaginary part is proportional to the free-carrier optical absorption intensity. When there is a short-range "Coulomb" interaction between the electron and the hole, bound state (exciton) may occur, and its position is determined by the intersection of Re$\mathscr{P}_0$ and $-1/u_0$. The optical absorption intensity $-\mathrm{Im}\mathscr{P}$ in this case is shown by the dashed curve. Parameters used are concentration $x=0.6$; band widths, $w_c=2$, $w_v=1$; Coulomb strength $u_0=2.5$; disorder, $\delta_c=0.5$, $\delta_v=-0.5$. The exciton state is infinitely sharp and a small width of 0.005 is arbitrary added to represent it on the figure. All energies are measured in the unit of valence-band width.

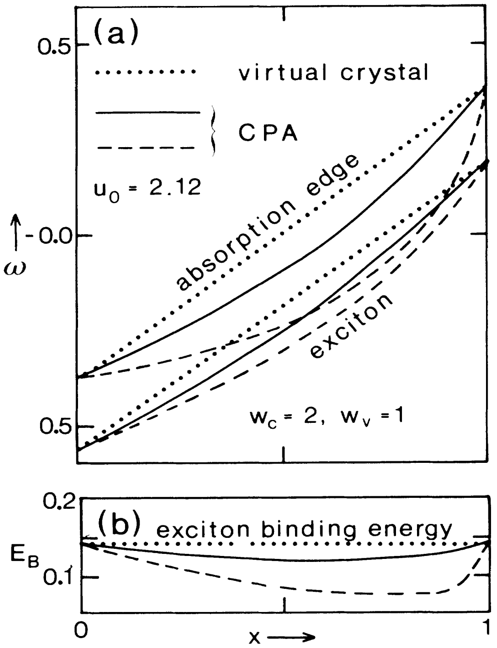

FIG. 4. Absorption edge and exciton energy (a) and the exciton binding energy (b) as a function of concentration x. Band width, $w_c=2$, $w_v=1$; Coulomb strength $u_0=2.12$. Solid curve: $\delta_c=0.5$, $\delta_c=-0.25$ (antiparallel motion). Dashed curve: $\delta_c=1$, $\delta_v=0.25$ (parallel motion). These two cases give the same virtual-crystal behavior (dotted line). All energies are measured in the unit of valence-band width.

$$1-\mathscr{L}^{(\omega)}(\varepsilon)\mathscr{R}_0^{(Q)}(\varepsilon)=0\ , \tag{5.1}$$

outside the absorption band determined by the single-particle density of states, this might introduce additional absorption band. However, in the range of disorder parameters of interest, this is not likely to happen. We checked it numerically, but could not determine analytically the threshold values for the parameters which make the condition (5.1) hold.

We have carried out numerical calculation of (3.14) for different parameter sets and then used (3.16) to determine the exciton states. All energies are measured in the unit of valence band width. A typical example is shown on Fig. 3. The solid line represents $\mathscr{P}_0$. The imaginary part vanishes abruptly at well-defined edges. This lack of "tail" states is known to be a shortcoming of the CPA. Exciton state is obtained as the intersection of Re$\mathscr{P}_0$ and $-1/u_0$. This appears outside the main band, and because the CPA does not yield tail states there, the exciton states have no width.

Figure 4 shows the concentration dependence of the absorption band edge, exciton peak, and its binding energy. Two different sets of disorder parameters are taken: antiparallel $\delta_c=0.5$, $\delta_v=-0.25$ and parallel $\delta_v=1$, $\delta_v=0.25$. In the linear (virtual-crystal) approximation (2.1), these two cases give identical behavior. We see in Fig. 4 that second- and higher-order terms give considerable difference between these two cases. We also note a striking asymmetry around $x=0.5$ unlike the prediction of the second-order theory (Sec. IV). The fact that $x>0.5$ is more "bowed" than $x<0.5$ is related to the tendency of impurity band formation near $x=1$, where atom A is host and atom B with lower site energy behaves as an impurity. (Near $x=0$, on the contrary, the impurity band would be formed at the upper band edge.)

Figure 4 also shows the exciton binding energy as a function of alloy concentration x. Excitons are more shallow in alloys, because they do not follow the bowing of the absorption edge sufficiently. This tendency is already seen in the lowest-order theory (4.11). The stronger the exciton binding (or the stronger the coupling u_0), the less the exciton sees the disorder and hence its absolute position remains unaffected. (For very strong u_0, the exciton is pinned at the virtual-crystal exciton energy.) This makes the apparent binding energy reduced, since it is defined as the difference in position of the absorption edge and the exciton energy.

VI. COMPARISON WITH EXPERIMENT

Let us compare our theory with experiment. Free-exciton peaks are observed in some III-V alloys including (In,Ga)P (Ref. 14) and (Ga,Al)As (Ref. 15). Detailed experimental results have been obtained in GaAs$_{1-x}$P$_x$ by Nelson, Holonyak, and Groves.[16] For $x=0.37$, these authors observe in the optical spectra two exciton peaks, $n=1$ and $n=2$, which allow them to determine the binding energy to be 4.8 meV. In considering the appropriate value in the virtual-crystal case, it is necessary to allow for changes in the reduced mass and dielectric constant with alloying. When this is done, the expected binding energy is 5.6 meV. The difference of 0.8 meV, 14% of the expected binding energy, is considered to be due to disorder.

In order to apply our theory to this material, we must determine the site-energy differences, δ_c and δ_v, and the band widths, w_c and w_v. The absolute position of the valence-band edge is obtained from the ionization energies. For GaAs and GaP, these are 5.4 and 5.95 eV below the vacuum level, respectively,[17] which give $\delta_v=-0.55$ eV. Next, using the direct band-gap energies, 1.52 eV for GaAs (Ref. 18) and 2.89 eV for GaP (Ref. 19), we obtain $\delta_c=0.82$ eV. The conduction- and valence-band widths are somewhat difficult to determine, as we are using the short-range disorder potential which "sees" states well above in the conduction band and deep in the valence band. This means we must use an effective width. One way to do so is to assume that the observed bowing is entirely due to disorder as was shown by Baldereschi and Maschke,[20] and then to use the second-order formula (4.8) to determine w_c and w_v by imposing $w_c:w_v=m_c^{-1}:m_v^{-1}$. Taking the average masses at $x=0.37$ as $m_c=0.11m_0$ and $m_v=0.61m_0$, and the bowing parameter [coefficient of $x(1-x)$ in (4.8)] $c=0.21$,[21] we obtain $w_c=23.1$ eV and $w_v=4.0$ eV. Finally we must choose a value for the electron-hole Coulomb potential u_0. Again, because of the short-range nature of the model, this is a delicate problem. A reasonable way would be to put $E_B^{\text{crys}}=4.8$ meV in the crystal exciton formula (4.5) to determine u_0, which gives $u_0=13.8$ eV. However, this value is too close to the threshold $w/2=13.6$ eV and, as is well known, bound-state energies in the Koster-Slater-type model are extremely sensitive to the value of the short-range interaction when they lie very close to the edge. For this reason, we have taken somewhat arbitrarily a slightly larger value of $u_0=14.5$ eV, which corresponds to $E_B^{\text{crys}}=56$ meV.

With all these parameters, at $x=0.37$ we obtain the binding energy $E_B=48$ meV, which corresponds to the reduction in binding of 14%, as is observed experimentally. Incidentally, this percentage reduction changes from 35% to 11% when u_0 is changed from 14.2 to 15.5 eV. [Though, as we have seen, the absolute value of reduction increases when u_0 is increased, the percentage decreases because the binding energy increases faster going as $(u_0-w/2)^2$.] In this way, though the absolute value of the theoretical binding energies cannot be directly comparable with observed binding energies (an order of magnitude greater than observed), we believe our main conclusion, disorder-induced shallowing, is in agreement with experiment. There is a further evidence for this: Ref. 16 reports the value of bowing parameter $c=0.186$ obtained from the exciton data is much smaller than the electroreflectance value $c=0.210$.[21] This might also indicate that the exciton is shallower in alloys.

Finally, we have also calculated the exciton binding energy without the vertex correction and obtained $E_B=42$ meV. Comparing this with the value including the vertex correction, 48 meV, we can conclude that the effective electron-hole interaction induced by simultaneous scattering from the same impurity deepens the binding. This is in agreement with our assertion deduced from the weak-disorder theory in Sec. IV, since $\delta_c\delta_v<0$ for Ga(As,P).

VII. DISCUSSION

In this section we shall examine the nature of our approximation employed in Sec. III, especially the truncation of the series (3.6) by (3.9) and consider some possible extensions. For a given configuration of impurity atoms, the exciton state is determined in principle by the pole of the formal solution to (A2):

$$\underline{\Pi}=\underline{\Pi}_0(\underline{1}+u_0\underline{\Pi}_0)^{-1} \; . \tag{7.1}$$

This is similar to (3.10), but $\underline{\Pi}$'s are considered to be matrices whose (i,j) element is $\Pi(\omega,ii,jj)$, and matrix inversion is implied on the right-hand side of (7.1). Our approximation (3.9) is thus seen to be equivalent to replacing the true configurational average of (7.1) by (3.16), or

$$\langle \underline{\Pi}_0(\underline{1}+u_0\underline{\Pi}_0)^{-1}\rangle \to \langle \underline{\Pi}_0\rangle(\underline{1}+u_0\langle \underline{\Pi}_0\rangle)^{-1} \; .$$

In other words, (3.9) results in the replacement

$$\langle \underline{\Pi}_0 u_0 \underline{\Pi}_0 u_0 \cdots \rangle \to \langle \underline{\Pi}_0\rangle u_0 \langle \underline{\Pi}_0\rangle u_0 \cdots \; . \tag{7.2}$$

Our approximation thus takes into account correctly the impurity-induced correlation of two particles but not more. In the case of the Wannier exciton, for which the Coulomb binding is loose, the electron-hole pair encounters many impurities before they get Coulomb scattered, and the replacement of the true pair propagator $\underline{\Pi}_0$ by its average $\langle \underline{\Pi}_0\rangle$ in (7.2) is justifiable.

A. Frenkel exciton limit

In the Appendix, we show explicitly that our configuration-dependent equation (3.1) can describe the Frenkel exciton in disordered solids if appropriate limits are taken. The Frenkel exciton equation (A3) turns out to be equivalent to the one-particle alloy problem (2.2). We do not know the exact solution to the Frenkel problem (A3) either, although the single-particle CPA gives a reasonable approximation in this case. However, in a limit of isolated atoms, we know the correct answer (A4) which gives the average polarization part (A5). This formula is different from our main result (3.16), where the denominator contains the averaged $\langle \Pi_0\rangle=\mathscr{P}_0$. As we have seen, this is closely related to our approximation (3.9). To describe the Frenkel exciton, it is necessary to take into account in (3.6) higher vertex corrections than the lowest (2.11) we considered. At least for short-range Coulomb interaction, there seems to be a reasonable hope that the series (3.6) or (3.4) can be summed in a way similar to the diagrammatic perturbation method[6] pioneered by Edwards and developed by many other people.

B. Impurity-bound excitons

At low concentration, when disorder is sufficiently strong, impurity states detached from the main band may be formed. In our model this occurs when $\delta_\mu/w_\mu>0.5$ for $y\equiv 1-x\to 0$. There are now two types of optical transitions, involving the host band and the impurity band. If these two types of transitions are sufficiently separated, one might expect two types of excitons. However, because in (3.16) one is using the average pair propagator $\mathscr{P}_0$, the resonance due to impurity (with intensity

$y \ll 1$) is wiped out by the host resonance with intensity $x \approx 1$, and there is no impurity-bound exciton. This has been confirmed by our numerical calculation.

C. Moving exciton

Finally, we have concentrated until now on the exciton states at zero total momentum $\mathbf{q}=0$. Our formalism permits us to investigate its properties at finite $\mathbf{q}$ if we specify explicitly the conduction- and valence-band dispersion in order to evaluate (3.15). One of the fundamental problems here is the question of whether disorder can trap the exciton (without impurity or other defect).[22] This is similar to the problem of localization of an electron in disordered systems. In our exciton case it is the translational degree of freedom represented by $\mathbf{q}$ which is to be considered. Unfortunately since the single-particle CPA cannot generate localization, there is no hope to obtain this in our CPA theory. Other theories, which describe the tail states better and emphasize localization effect,[1,23] are required.

VIII. CONCLUSIONS

In conclusion, we have developed a systematic approximate treatment of the two-body problem in disordered systems. To our knowledge, this is the first time that a Bethe-Salpeter-type problem in disordered systems has been solved starting from the solution to the one-particle Green's functions.

ACKNOWLEDGMENTS

One of the authors (M.A.K.) acknowledges financial support from the European Science Exchange Program between the Centre National de la Recherche Scientifique and the Royal Society, enabling stays at the Department of Theoretical Physics at Oxford. He would also like to thank A. Ghazali for indicating him Abe and Toyozawa's paper and for his interest in this work. The Laboratoire de Physique des Solides is laboratoire No. 154 associé au Centre National de la Recherche Scientifique.

APPENDIX: FRENKEL EXCITON IN DISORDERED SOLIDS

In this appendix, we show how the Frenkel exciton in an alloy may be derived from (3.1). Essentially, this occurs when the hopping integral $t(i-j)$ in (2.2) vanishes. Then the excitation becomes localized on atoms and its transfer takes place via the exchange interaction w.

1. Equation for the polarization part

As long as the kernel u does not depend on energy, it is possible to treat the equation for the polarization part instead of dealing with the two-particle Green's function K [Eq. (3.1)]. To do so, let us define the configuration-dependent polarization part Π in the site representation in terms of the two-particle Green's function as

$$\Pi(\omega,ij,kl)=\int\frac{d\varepsilon}{2\pi}\int\frac{d\varepsilon'}{2\pi}K^{(\omega)}(\varepsilon\varepsilon',ij,kl)\ . \tag{A1}$$

The equation satisfied by Π can be obtained by integrating (3.1) and reads

$$\begin{aligned}\Pi(\omega,ij,kl)=\Pi_0(\omega,ij,kl)\\ &-\sum_{a,b}\Pi_0(\omega,ij,ab)v(a-b)\Pi(\omega,ab,kl)\\ &+\sum_{a,b}\Pi_0(\omega,ij,aa)w(a-b)\Pi(\omega,bb,kl)\ ,\end{aligned} \tag{A2}$$

where Π_0 is the noninteracting polarization part derived from (3.2) in a similar way to (A1). In (A2), we have also used the explicit form (3.3).

2. Frenkel exciton

When the overlap of atomic wave functions is negligible, the one-particle Green's function reduces to the locator (2.3):

$$G_\mu(\varepsilon,ik)=\delta_{ik}g_\mu(\varepsilon,i)\quad(\mu=c,v)$$

and Π_0 takes the form

$$\Pi_0(\omega,ij,kl)=\delta_{ik}\delta_{jl}\pi_0(\omega,ij)\ ,$$

where

$$\pi_0(\omega,ij)=\int\frac{d\varepsilon}{2\pi}\int\frac{d\varepsilon'}{2\pi}g_c(\varepsilon+\omega,i)g_v(\varepsilon,j)\ .$$

In this case, (A2) takes a particularly simple form for the site-diagonal polarization part ($i=j$, $k=l$), namely

$$[1+\pi_0(\omega,i)v(0)]\Pi(\omega,ii,kk)=\delta_{ik}\pi_0(\omega,i)+\pi_0(\omega,i)\sum_j w(i-j)\Pi(\omega,jj,kk)\ , \tag{A3}$$

where $\pi_0(\omega,i)\equiv\pi_0(\omega,ii)$ describes the intra-atomic excitation.

Equation (A4) is very similar to (2.2)for a single electron in alloys. In fact, we only need to make a "translation" of notation in going from (2.2) to (A3):

locator	↔	$g(\varepsilon,i)$	↔	$[1+\pi_0(\omega,i)v(0)]^{-1}\pi_0(\omega,i)$
interactor	↔	$t(i-j)$ (transfer)	↔	$w(i-j)$ (exchange)
Green's function	↔	$G(\varepsilon,ik)$	↔	$\Pi(\omega,ii,kk)$

The reason for this similarity of the two-body problem to the one-body problem is clear. The electron and hole are so tightly bound that they are always at the same site, and thus their relative motion is negligible. It is then necessary to consider solely the translational degree of freedom, which corresponds to the one-particle motion described by (2.2).

3. Isolated atoms

As is the case for the one-electron problem (2.2), it is not possible to obtain an exact solution Π for (A3) pertinent to a particular configuration except for the case of $w=0$, where atoms are totally separate and there is no dipolar coupling between them due to the exchange interaction. In this case, we can trivially solve (A3) to obtain

$$\Pi(\omega,ii,kk)=\frac{\delta_{ik}\pi_0(\omega,i)}{1+\pi_0(\omega,i)v(0)} . \tag{A4}$$

At this stage, we can safely average this over all configurations. In our case of binary alloy, site i can be either A or B with probability x and $1-x$, and thus

$$\mathscr{P}(\omega,ii,kk)=\delta_{ik}\left[\frac{x\pi_0(\omega,A)}{1+\pi_0(\omega,A)v(0)}+\frac{(1-x)\pi_0(\omega,B)}{1+\pi_0(\omega,B)v(0)}\right] . \tag{A5}$$

---

[1] S. D. Baranovski and A. L. Efros, Fiz. Tekh. Poluprovodn. **12**, 2233 (1978) [Sov. Phys.—Semicond. **12**, 1328 (1978)].

[2] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. **24**, 341 (1968).

[3] P. N. Sen, Phys. Rev. Lett. **30**, 553 (1973).

[4] S. D. Mahanti, Phys. Rev. B **10**, 1384 (1974).

[5] B. Velický, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. **175**, 747 (1968).

[6] See, for a review, R. J. Elliott, J. A. Krumhansl, and P. L. Leath, Rev. Mod. Phys. **46**, 465 (1974); F. Yonezawa and K. Morigaki, Prog. Theor. Phys. Suppl. **53**, 1 (1973).

[7] B. Velický, Phys. Rev. **184**, 614 (1969).

[8] P. L. Leath, Phys. Rev. B **2**, 3078 (1970).

[9] S. Abe and Y. Toyozawa, J. Phys. Soc. Jpn. **50**, 2185 (1981).

[10] See, for instance, E. M. Lifshitz and L. P. Pitaevskii, *Relativistic Quantum Theory* (Pergamon, Oxford, 1974), Pt. 2, p. 470.

[11] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1966).

[12] Concerning the sign of u_0, $u_0>0$ corresponds (as usual) to *repulsive* electron-electron interaction. The reason why u_0 appears with a plus sign in the resonance condition (3.13) rather than a minus sign as is usual in the Koster-Slater impurity problem is that here one is dealing with the interaction of an electron with a hole.

[13] S. F. Edwards, Philos. Mag. **3**, 1020 (1958).

[14] R. J. Nelson and N. Holonyak, Jr., J. Phys. Chem. Solids **37**, 629 (1976).

[15] B. Monemar, K. K. Shih, and G. D. Pettit, J. Appl. Phys. **47**, 2604 (1976); R. Dingle, R. A. Logan, and J. R. Arthur, Jr., Inst. Phys. Conf. Ser. **33a**, 210 (1977).

[16] R. J. Nelson, N. Holonyak, Jr., and W. O. Groves, Phys. Rev. B **13**, 5415 (1976); R. J. Nelson, in *Excitons,* edited by E. I. Rashba and M. D. Sturge (North-Holland, Amsterdam, 1982), p. 319.

[17] G. M. Guichar, C. A. Sebenne, and G. A. Garry, Phys. Rev. Lett. **37**, 1158 (1976); G. M. Guichar, C. A. Sebenne, and D. C. Thuault, Surf. Sci. **86**, 789 (1979).

[18] S. B. Nam, D. C. Reynolds, C. W. Litton, R. J. Almassy, T. C. Collins, and C. M. Wolfe, Phys. Rev. B **13**, 761 (1976).

[19] C. Varea de Alvarez, J. P. Walter, M. L. Cohen, J. Stokes, and Y. R. Shen, Phys. Rev. B **6**, 1412 (1972).

[20] A. Baldereschi and K. Maschke, Solid State Commun. **16**, 99 (1975). This is, however, a point of controversy; see references cited therein.

[21] A. G. Thompson, M. Cardona, K. L. Shaklee, and J. C. Woolley, Phys. Rev. **146**, 601 (1966).

[22] S. Lai and M. V. Klein, Phys. Rev. Lett. **44**, 1087 (1980); S. T. Lai and M. V. Klein, Phys. Rev. B **29**, 3217 (1984).

[23] See, for a review, E. O. Kane, Solid-State Electron. **28**, 3 (1985).