PHYSICAL REVIEW B VOLUME 11, NUMBER 10 15 MAY 1975

Exciton dispersion in degenerate bands

Evan O. Kane
Bell Laboratories, Murray Hill, New Jersey 07974

(Received 26 September 1974)

The dispersion relations of band-gap excitons in covalent cubic crystals are investigated. For low momentum a perturbation method is used following Baldereschi and Lipari. At very low momentum the dispersion is strongly affected by a splitting of the fourfold $K = 0$ degeneracy for those materials with nonspherical conduction-band energy surfaces. In silicon and germanium this splitting leads to strong nonparabolic effects at energies in the 0.1-meV range. For large momentum a band decoupling scheme which generalizes the center-of-mass transformation method is derived and the concept of "light-mass" and "heavy-mass" excitons naturally results. Numerical values are given for germanium, silicon, and gallium arsenide.

I. INTRODUCTION AND CONCLUSIONS

The problem of obtaining the $E(\vec{K})$ (dispersion relation) for excitons is very simple for the case where the electron and hole both belong to nondegenerate bands. A center-of-mass transformation of classical form suffices to separate the translational motion from the relative motion.[1] The transformation is easily generalized to ellipsoidal energy surfaces. If one of the particles is in a degenerate band, however, the problem is not so readily solved as has been recognized since the work of Dresselhaus.[1] Unfortunately, most of the excitons in cubic covalent crystals are of the latter type since the valence-band maximum is four-fold degenerate at the Γ point. Although very few determinations of the exciton dispersion relation have been made it enters into many different problems and hence it is very desirable to have a more quantitative knowledge of its value than presently exists.

One of the most straightforward problems involving the exciton dispersion is the optical absorption, which is proportional to the exciton density of states weighted by an optical matrix element. Recent experimental and theoretical work on this problem in germanium was motivated by the present paper.[2] Here it was shown that strong nonparabolic effects associated with the "mass-reversal" effect to be described later can be very clearly identified in the energy dependence of the absorption.

Any long-range low-frequency perturbation may be discussed in terms of the dispersion relations. Hopfield and Thomas[3] have used a combined electric and magnetic field perturbation method to obtain a direct measurement of the mass in CdS. Since the material is hexagonal the present methods are not directly applicable. Another application would be to the asymptotic decay rate $e^{-\kappa r}$ of the wave function of a bound exciton, which should be related to the exciton mass M^* by $\hbar^2\kappa^2/2M^* = E$, where E is the binding energy.

In this paper we tackle this problem by two different methods whose range of validity depends on the exciton momentum. At low momentum the perturbation method is most applicable and here we generalize the method of Baldereschi and Lipari[4] to treat the dispersion. At high momentum the hole bands are nearly decoupled into light- and heavy-hole bands owing to their large kinetic-energy differences. We exploit this fact to obtain an approximate decoupling of the excitons. Our method may be regarded as a generalization of the center-of-mass transformation to degenerate bands in that it achieves an "optimal" (though not exact) separation of the problem into translational- and relative-kinetic-energy terms.

Using the perturbation method we obtain general expressions for the dispersion relations for the case where the conduction band is an ellipsoid centered on a $00k$, kkk, or 000, k point. The 000 ellipsoids are spherical and the $K=0$ excitons have the fourfold degeneracy of the hole states. The nonspherical ellipsoids lead to a splitting of the degeneracy at $K=0$ into two Kramers doublets. The splitting is a small fraction of the total exciton binding and indeed is required to be so for the validity of the perturbation method. The presence of the splitting leads to nonparabolicities in the dispersion relation for exciton kinetic energies of the order of the splitting energy. Even for the spherical conduction band, nonparabolic dispersion relations result from the coupling between the degenerate bands on an energy scale equal to the exciton binding energies. The magnitude of the effect may be estimated by the use of the high-momentum decoupling approximation. In this limit the excitons separate into two sets of exciton bands corresponding to the light and heavy holes. Each set has many excited bands corresponding to the excited states of the internal degrees of freedom (1*s*, 2*s*, 2*p*, etc., in the hydrogenic notation, although this notation is imprecise because the relative kinetic energy is ellipsoidal rather than spherical). Each excited band on the same set has the same translational mass in the decoupling approximation, as expected from a center-of-mass transformation. At some kinetic energy of the order of the exciton binding the ground state of the light-mass exciton band will intersect the 2*p* or 3*d* (or higher) excited states of the heavy-mass exciton bands. This will lead to a strong interaction between the overlapping bands and will result in nonparabolic dispersion relations. Neither of the methods presented here is valid in this case.

Detailed numerical results are given for GaAs, Si, and Ge. The perturbation method is inaccurate for germanium as evidenced by nonphysical negative masses in the [111] direction. A variational approach should be capable of improving on the calculations when perturbation theory breaks down but that method has not been tested here. A very simple variational calculation for germanium has been performed elsewhere[2] which, at least, avoids the problem of negative longitudinal masses.

In silicon and germanium there are "light" and "heavy" exciton bands roughly in accord with the concept of adding electron and hole masses to obtain the translational mass which has been justified by the high-momentum band decoupling scheme. At low $\vec{K}$ the hole contribution to the exciton mass depends on the direction of $\vec{K}$ relative to the axis of the electron ellipsoid. The lowest exciton state is heavy-hole-like (light-hole-like) parallel (perpendicular) to the ellipsoidal axis, while the situation is reversed for the upper exciton. Perpendicular to the ellipsoidal axis the exciton mass "reverses" between the lower and upper excitons as $\vec{K}$ increases, which leads to strong nonparabolic effects at energies of the order of the splitting energies, i.e., a few tenths of a millivolt in silicon and germanium. An example of this mass reversal as a function of K^2 is given in Fig. 1.

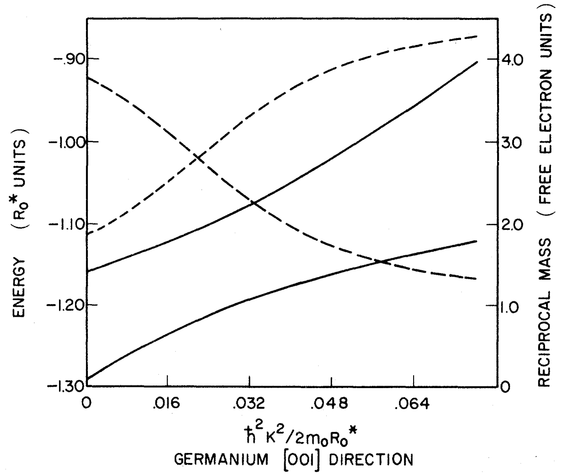

FIG. 1. Reciprocal masses and energies versus K^2 for excitons in germanium with momentum in the [001] direction. The solid curves are the energies and the dashed curves are the reciprocal masses.

We have not included the "exchange" interaction between the spin of the electron and that of the hole. It appears that this effect is quite small for indirect materials but may be more important in direct materials. Further work on this problem will be reported at a later date. We also have ignored linear *k* terms which exist in zinc-blende materials but are usually quite small.[5]

II. OUTLINE OF PAPER

In Sec. III we discuss notation and introduce a "Kramers basis" and a set of symmetry-adapted operators in which we can conveniently describe the exciton Hamiltonian. In substance these are the well-known effective-mass Hamiltonians of Kohn and Luttinger[6] with modifications similar to those introduced by Baldereschi and Lipari.[4]

In Sec. IV we introduce the perturbation due to the finite exciton momentum and treat it by the perturbative methods of Lipari and Baldereschi.[4] General expressions for $E(\vec{K})$ are obtained for conduction-band ellipsoids centered on 000, 00*k*, and *kkk* points.

In Sec. V we present the large-momentum band decoupling scheme which generalizes the center-of-mass transformation method to degenerate bands.

In Sec. VI we discuss the results obtained and give detailed numerical computations for the case of germanium, silicon, and gallium arsenide.

III. EXCITON HAMILTONIAN; NOTATION AND SYMMETRY CONSIDERATIONS

We consider the problem of excitons where the hole has the fourfold degeneracy characteristic of the uppermost valence band in tetrahedrally bonded semiconductors. We treat the problem in the effective-mass approximation and ignore the contribution from the spin-orbit split-off band. We also ignore the spin of the electron and hence do not treat the small exchange splitting of the spin degeneracy.[7] We follow the method of Lipari and Baldereschi,[4] hereafter referred to as LB, wherein the "*d*-like" part of the Hamiltonian is treated by second-order perturbation theory. We extend their results by calculating the momentum dependence of the exciton energy (dispersion relation).

We consider the effective-mass Hamiltonian for the exciton energy relative to the energy gap,

$$H_{\rm ex} = T_e - T_h + H_C \; ; \qquad (1)$$

T_e and T_h are the electron and hole kinetic energies, while H_C is their Coulomb interaction:

$$H_C = -e^2/\epsilon\,|r_e - r_h| \,. \tag{2}$$

In momentum space we have

$$T_e = \frac{\hbar^2}{2}\left(\frac{k_{e\parallel}^2}{m_\parallel} + \frac{k_{e\perp}^2}{m_\perp}\right) , \tag{3}$$

$$T_h = -\hbar^2 \gamma_1 k_h^2/2m_0 + T_{dh}(k_{hi}\, k_{hj}) , \tag{4}$$

where $\vec{k}_h$ is the hole momentum and $\vec{k}_e$ is the electron momentum relative to the conduction band minimum:

$$\vec{k}_e^{\,\text{tot}} = \vec{k}_e + \vec{K}_0 \,. \tag{5}$$

T_{dh} is a 4×4 matrix operator on the fourfold-degenerate valence band in the Kohn-Luttinger representation.[6] T_h was given first by Dresselhaus, Kip, and Kittel.[8] We prefer to utilize the symmetry of a Kramers basis defined by[9]

$$\begin{aligned}
\psi_{3/2} &\equiv -\tfrac{1}{2}\sqrt{2}\,(x+iy)\uparrow ,\\
\psi_{1/2} &\equiv \sqrt{\tfrac{2}{3}}\, z\downarrow - \tfrac{1}{6}\sqrt{6}\,(x+iy)\downarrow ,\\
K\psi_{3/2} &\equiv -\tfrac{1}{2}\sqrt{2}\,(x-iy)\downarrow ,\\
K\psi_{1/2} &\equiv \tfrac{1}{6}\sqrt{6}\,(x-iy)\uparrow + \sqrt{\tfrac{2}{3}}\, z\downarrow ,\\
K\psi_\mu &= \pm\psi_{-\mu} ,
\end{aligned} \tag{6}$$

where x, y, and z refer to the *tetrahedral* symmetry of the wave functions. All symmetry operations can be written as a tetrahedral operation $\mathcal{J}$ or as $\mathcal{J}$ times $(\tau|J)$, where $(\tau|J)(x,y,z) = (x,y,z)$ and $(\tau|J)$ interchanges the two atoms in a unit cell.

The notation x, y, z is really a shorthand for $x \to (x+a/8)f(\vec{r}+\vec{\tau}/2) - (x-a/8)f(\vec{r}-\vec{\tau}/2)$, i.e., tight-binding p functions which are bonding between the atomic sites $\pm\vec{\tau}/2$, $\vec{\tau} = (a/4)\,(1, 1, 1)$. $f(\vec{r})$ is invariant under cubic operations. Of course, we are not making the tight-binding approximation but we can use it to give a unique definition of the symmetry properties of our basis functions.

In a Kramers basis such as that of Eq. (6) a *traceless* Hamiltonian which we abbreviate as $\mathcal{H}(A,B,C)$ can be written in the form

$$\mathcal{H}(A,B,C) \equiv \begin{array}{c} \begin{array}{cccc} \psi_{3/2} & \psi_{1/2} & K\psi_{3/2} & K\psi_{1/2} \end{array} \\ \left|\begin{array}{cccc} A & B & 0 & C \\ B^* & -A & -C & 0 \\ 0 & -C^* & A & B^* \\ C^* & 0 & B & -A \end{array}\right| \end{array} . \tag{7}$$

A, B, C can be further simplified by expressing them in terms of irreducibly transforming elements $A_{J\alpha}$ according to

$$\begin{aligned}
A &= A_{2a} ,\\
B &= A_{3xz} - iA_{3yz} ,\\
C &= A_{2b} - iA_{3xy} \,.
\end{aligned} \tag{8}$$

The notation for $A_{J\alpha}$ is group theoretical, with J referring to the irreducible representation and α to the partner. Wherever possible we will abbreviate $\mathcal{H}(A,B,C)$ still further by

$$\mathcal{H}(A_{J\alpha}) \equiv \mathcal{H}(A,B,C) , \tag{9}$$

with the arguments related by Eq. (8). In this paper the "space" of $J\alpha$ will always consist of

$$\{J\alpha\} = \{2a,\ 2b,\ 3xy,\ 3yz,\ 3xz\} \tag{10}$$

and all sums over $J\alpha$ will refer to sums over this set. We also note the convenient fact that when A, B, C are numbers the eigenvalues $E_{A\pm}$ of $\mathcal{H}(A,B,C)$ are

$$E_{A\pm} = \pm\,(A^2 + |B|^2 + |C|^2)^{1/2} . \tag{11a}$$

When the $A_{J\alpha}$ are real numbers, Eq. (11a) can also be written

$$E_{A\pm} = \pm\left(\sum_{J\alpha} A_{J\alpha}^2\right)^{1/2} . \tag{11b}$$

With the use of this notation the Hamiltonian T_{dh} in Eq. (4) can be written

$$T_{dh}(k_{hi}\, k_{hj}) = \frac{\hbar^2}{2m_0}\, \mathcal{H}\big(\gamma_J\, \mathcal{O}_{J\alpha}(k_{hi}\, k_{hj})\big) \,. \tag{12}$$

The symmetry-adapted polynomials $\mathcal{O}_{J\alpha}$ are defined by

$$\begin{aligned}
\mathcal{O}_{2a}(k_i k_j) &\equiv (2k_z^2 - k_x^2 - k_y^2) ,\\
\mathcal{O}_{2b}(k_i k_j) &\equiv \sqrt{3}\,(k_x^2 - k_y^2) ,
\end{aligned} \tag{13}$$

$$\mathcal{O}_{3xy}(k_i k_j) \equiv 2\sqrt{3}\, k_x k_y ,\ \text{etc., for } xz,\ yz . \tag{14}$$

The dimensionless inverse-mass parameters γ_J refer to the Luttinger parameters[10] γ_1, γ_2, γ_3.

Similarly, the electron kinetic-energy operator T_e in Eq. (3) can be written in the manner of LB:

$$T_e = \frac{\hbar^2 \gamma_{1e}}{2m_0}\, k_e^2 + T_{de}(k_{ei}\, k_{ej}) , \tag{15}$$

$$T_{de} \equiv -\frac{\hbar^2}{2m_0}\, \gamma_{Je} \sum_\alpha c_{J\alpha}\, \mathcal{O}_{J\alpha}(k_{ei}\, k_{ej}) , \tag{16}$$

where we have

$$\begin{aligned}
&J = 2, \quad c_{2a} = 1;\ [001]\text{ellipsoid}\\
&J = 3, \quad c_{3xy} = c_{3xz} = c_{3yz} = \tfrac{1}{3}\sqrt{3}\,;\ [111]\text{ellipsoid.}
\end{aligned} \tag{17}$$

In Eq. (17) all quantities not explicitly defined are zero.

The inverse-mass parameter γ_{Je} is given by

$$\gamma_{Je} = \tfrac{1}{3}\, m_0\left(\frac{1}{m_\perp} - \frac{1}{m_\parallel}\right) , \tag{18}$$

$$\frac{\gamma_{1e}}{m_0} \equiv \frac{1}{3}\left(\frac{2}{m_\perp} + \frac{1}{m_\parallel}\right) . \tag{19}$$

IV. PERTURBATION METHOD; LOW EXCITON EXCITON MOMENTUM

In Sec. III we have simply rewritten the effective-mass Hamiltonian for excitons in terms of a notation whose symmetry properties enable the results to be stated more simply. We now divide the Hamiltonian in the manner of LB:

$$H_{\mathrm{ex}} = H_s + H_d , \tag{20}$$

$$H_s \equiv \frac{\hbar^2}{2m_0}\, \gamma_{1e}\, k_e^2 + \frac{\hbar^2}{2m_0}\, \gamma_1\, k_h^2 + H_C , \tag{21}$$

$$H_d = T_{de}(k_{ei} k_{ej}) - T_{dh}(k_{hi} k_{hj}) . \tag{22}$$

We make a "center-of-mass" transformation to express the Hamiltonian H_s in terms of hydrogenic wave functions and energies:

$$\vec{\mathrm{k}}_e = \vec{\sigma} + \beta_e \vec{\mathrm{K}} , \tag{23}$$

$$\vec{\mathrm{k}}_h = \vec{\sigma} - \beta_h \vec{\mathrm{K}} , \tag{24}$$

$$\beta_e = \gamma_1/(\gamma_1 + \gamma_{1e}) , \quad \beta_h = \gamma_{1e}/(\gamma_1 + \gamma_{1e}) , \tag{25}$$

where by (23)–(25) we have conservation of exciton center-of-mass momentum $\vec{\mathrm{K}}$,

$$\vec{\mathrm{k}}_e - \vec{\mathrm{k}}_h = \vec{\mathrm{K}} . \tag{26}$$

Actually, in accord with Eq. (5), we have

$$\vec{\mathrm{K}}^{\mathrm{tot}} = \vec{\mathrm{K}} + \vec{\mathrm{K}}_0 , \tag{27}$$

but we use $\vec{\mathrm{K}}$ throughout the remainder of the paper as the center-of-mass momentum and we wish to calculate the energy of the excitonic states as a function of $\vec{\mathrm{K}}$.

Using (23)–(25) the eigenstates of H_s have hydrogenic form with the energy spectrum

$$E_{ns} = -\frac{R_0^*}{n^2} + \frac{\hbar^2 K^2}{2M_0} , \tag{28}$$

$$R_0^* = R_0/\gamma_{1r}\, \epsilon^2 , \tag{29}$$

$$\gamma_{1r} = \gamma_1 + \gamma_{1e} , \tag{30}$$

$$M_0 = \gamma_1^{-1} + \gamma_{1e}^{-1} , \tag{31}$$

n is the usual hydrogenic quantum number, and the envelope wave functions are the usual hydrogenic functions of the relative momentum $\vec{\sigma}$. Actually n will be required to run over continuum states also as detailed in LB. R_0^* is the "effective Rydberg" of the hydrogenic problem.

The complete set of unperturbed wave functions are then given by

$$\psi_{nlm\mu}^{K} = e^{i\vec{\mathrm{K}}\cdot(\beta_e \vec{\mathrm{r}}_e + \beta_h \vec{\mathrm{r}}_h)}\, \varphi_{nlm}(\vec{\mathrm{r}}_e - \vec{\mathrm{r}}_h)\, u_c(\vec{\mathrm{r}}_e) u_{v\mu}^*(\vec{\mathrm{r}}_h) , \tag{32}$$

where φ_{nlm} are a complete set of hydrogenic functions, u_c is the periodic part of the electron wave function at the band minimum, and $u_{v\mu}$ with $\mu = \pm\frac{3}{2}, \pm\frac{1}{2}$ is one of the four hole state functions in Eq. (6).

When the center-of-mass transformation equations (23)–(25) are substituted in H_d in Eq. (22) we obtain three types of terms,

$$H_d \equiv H_{d\sigma} + H_{dm} + H_{dK} , \tag{33}$$

$$H_{d\sigma} \equiv T_{de}(\sigma_i \sigma_j) - T_{dh}(\sigma_i \sigma_j) , \tag{34}$$

$$H_{dK} \equiv \beta_e^2\, T_{de}(K_i K_j) - \beta_h^2\, T_{dh}(K_i K_j) , \tag{35}$$

$$H_{dm} \equiv 2\beta_e T_{de}[\sigma_i K_j] + 2\beta_h T_{dh}[\sigma_i K_j] , \tag{36}$$

$$[\sigma_i K_j] \equiv (\sigma_i K_j + K_i \sigma_j)/2 . \tag{37}$$

In the spirit of LB we propose to treat the "d-like" terms by perturbation theory. We define

$$\begin{aligned} a_{J\alpha} &\equiv \frac{\hbar^2}{2m_0}\, \gamma_J\, [2\beta_h\, \mathcal{O}_{J\alpha}(\sigma_i K_j) - \mathcal{O}_{J\alpha}(\sigma_i \sigma_j)] , \\ d &\equiv -\frac{\hbar^2}{2m_0}\, \gamma_{Je} \sum_{\alpha} c_{J\alpha}[2\beta_e \mathcal{O}_{J\alpha}(\sigma_i K_j) + \mathcal{O}_{J\alpha}(\sigma_i \sigma_j)] , \end{aligned} \tag{38}$$

where the $c_{J\alpha}$ are defined in Eq. (17), and then use equations analogous to (8) to define quantities a, b, c:

$$a \equiv a_{2a} , \quad b \equiv a_{3xz} - i a_{3yz} , \quad c \equiv a_{2b} - i a_{3xy} . \tag{39}$$

Then we can write

$$H_{d\sigma} + H_{dm} = d\, I + \mathcal{H}(a_{J\alpha}) = d\, I + \mathcal{H}(a, b, c) , \tag{40}$$

where I is the four-dimensional identity and $\mathcal{H}$ is defined in Eqs. (7)–(9).

We introduce the operator P which projects onto the set of states in Eq. (32) for which $\varphi_{nlm}(\vec{\mathrm{r}}_e - \vec{\mathrm{r}}_h) = \varphi_{1s}(\vec{\mathrm{r}}_e - \vec{\mathrm{r}}_h)$. We then define perturbation-theory Green's functions by

$$g_0 \equiv (1-P)(E - H_s)^{-1}(1-P) , \quad G_0 \equiv g_0 I , \tag{41}$$

where I is the four-dimensional identity. We treat $H_{d\sigma} + H_{dm}$ in Eq. (40) by degenerate second-order perturbation theory obtaining the effective 4×4 Hamiltonian

$$H_{\mathrm{pert}} \equiv P(H_{d\sigma} + H_{dm})\, G_0 (H_{d\sigma} + H_{dm})\, P . \tag{42}$$

We note the commutative relations

$$\langle 1s \,|\, a g_0 b \,|\, 1s \rangle = \langle 1s \,|\, b g_0 a \,|\, 1s \rangle , \tag{43}$$

with similar relations for all pairs of a, b, c, d. With the use of these relations and the definition of $\mathcal{H}(\ldots,\, \ldots,\, \ldots)$ in Eq. (7) we can evaluate H_{pert} in Eq. (42) as

$$\begin{aligned} H_{\mathrm{pert}} = {} & I \langle 1s \,|\, d g_0 d + a g_0 a^* + b g_0 b^* + c g_0 c^* \,|\, 1s \rangle \\ & + 2\mathcal{H}(\langle 1s \,|\, d g_0 a \,|\, 1s \rangle, \langle 1s \,|\, d g_0 b \,|\, 1s \rangle, \\ & \langle 1s \,|\, d g_0 c \,|\, 1s \rangle) . \end{aligned} \tag{44}$$

Equation (44) can then be evaluated more explicitly with the use of the relations

$$\langle 1s | a_{J\alpha} g_0(E) a_{J'\alpha'} | 1s \rangle$$
$$= - \gamma_J \gamma_{J'} [4 \beta_h^2 \hbar^2 Q_{J\alpha, J'\alpha'}(K^2) g_3(E) + g_1(E) \delta_{JJ'} \delta_{\alpha\alpha'}], \tag{45}$$

$$\langle 1s | a_{J\alpha} g_0(E) d | 1s \rangle$$
$$= \gamma_J \gamma_{Je} \sum_{J'\alpha'} c_{J'\alpha'} \times [4\beta_e \beta_h \hbar^2 Q_{J\alpha, J'\alpha'}(K^2) g_3(E) - g_1(E) \delta_{JJ'} \delta_{\alpha\alpha'}], \tag{46}$$

$$\langle 1s | d g_0(E) d | 1s \rangle$$
$$= - \gamma_{Je}^2 \sum_{J\alpha, J'\alpha'} c_{J\alpha} c_{J'\alpha'} \times [4 \beta^2{}_e \hbar^2 Q_{J\alpha, J'\alpha'}(K^2) g_3(E) + g_1(E) \delta_{JJ'} \delta_{\alpha\alpha'}]. \tag{47}$$

The polynomials $Q_{J\alpha, J'\alpha'}$ have the property $Q_{J\alpha, J'\alpha'} = Q_{J'\alpha', J\alpha}$ and are further defined by

$$Q_{2a,2a} = (4K_z^2 + K_x^2 + K_y^2)/6, \quad Q_{2b,2b} = (K_x^2 + K_y^2)/2,$$
$$Q_{2a,2b} = (-K_x^2 + K_y^2)/2\sqrt{3}; \quad Q_{2a,3xy} = -K_x K_y / \sqrt{3},$$
$$Q_{2a,3xz} = K_x K_z / 2\sqrt{3}, \; Q_{2a,3yz} = K_y K_z / 2\sqrt{3}, \tag{48}$$
$$Q_{2b,3xy} = 0, \quad Q_{2b,3xz} = K_x K_z / 2, \quad Q_{2b,3yz} = -K_y K_z / 2,$$
$$Q_{3ij,3ij} = (K_i^2 + K_j^2)/2, \quad Q_{3ij,3lj} = K_i K_l / 2.$$

The quantities $g_1(E)$ and $g_3(E)$ are defined by

$$g_1(E) = \frac{16 R_0^*}{5 \gamma_{1r}^2} S_1\left(\frac{-R_0^* - E}{R_0^*} \right), \tag{49}$$

$$g_3(E) = \frac{4}{m_0 \gamma_{1r}} S_3\left(\frac{-R_0^* - E}{R_0^*} \right). \tag{50}$$

TABLE I. Tabulation of the functions $S_1(x)$ and $S_3(x)$ used in Eqs. (49) and (50) and defined by Baldereschi and Lipari in Ref. 4. The values of S_1 for positive argument are included for convenience. They agree with values already obtained by Baldereschi and Lipari.

x	$S_1(x)$	$S_3(x)$
0	0.2246	0.1875
1	0.1587	0.0974
2	0.1267	0.0674
3	0.1069	0.0519
4	0.0932	0.0424
5	0.0829	0.0359
6	0.0750	0.0311
8	0.0632	0.0247
10	0.0550	0.0205
−0.1	0.2359	0.2086
−0.2	0.2490	0.2361
−0.3	0.2645	0.2736
−0.4	0.2831	0.3287
−0.5	0.3066	0.4200
−0.6	0.3377	0.6120
−0.7	0.3834	1.453

From LB we have the values $S_1(0) \simeq 0.2246$ and $S_3(0) = \frac{3}{16}$. LB tabulate $S_1(x)$. We have tabulated $S_3(x)$ in Table I. Finally we express the complete exciton Hamiltonian as

$$H_{\text{tot}} = I\left(-R_0^* + \frac{\hbar^2 K^2}{2M_0} \right) + H_{\text{pert}} + H_{dK}. \tag{51}$$

H_{tot} is then easily diagonalized using Eq. (11a). After some algebra the resulting energies can be expressed in the form

$$E_{1s\pm}(\vec{K}) = \mathcal{S} \pm Z^{1/2}, \tag{52}$$

where

$$\mathcal{S} = R_a + \frac{\hbar^2 K^2}{2M_a}, \tag{53}$$

$$Z = R_{Jb}^2 + 2R_{Jb}\left(\frac{\hbar^2 K^2}{2M_{Jb}} \right) + \left(\frac{\hbar^2 K^2}{2M_{Jc}} \right)^2, \tag{54}$$

$$R_a = -R_0^* - g_1(E)(2\gamma_2^2 + 3\gamma_3^2 + \gamma_{Je}^2), \tag{55}$$

$$\frac{1}{M_a} = \frac{1}{M_0} - 2g_3(E)\left(4[\beta_h^2(\tfrac{2}{3}\gamma_2^2 + \gamma_3^2) + \tfrac{1}{3}\beta_e^2 \gamma_{Je}^2] + \tfrac{2}{3}\beta_e^2 \gamma_{Je}^2 \frac{\mathcal{O}_{Ja}(K_i K_j)}{K^2} \right) - \frac{\beta_e^2 \gamma_{Je}}{m_0} \frac{\mathcal{O}_{Ja}(K_i K_j)}{K^2}, \tag{56}$$

$$R_{Jb} = -2g_1(E)\gamma_J \gamma_{Je}, \tag{57}$$

$$\frac{1}{M_{Jb}} = \frac{V_J}{3} + \left(\frac{V_J}{6} - W_J \right) \frac{\mathcal{O}_{Ja}(K_i K_j)}{K^2}, \tag{58}$$

$$V_J = 16\beta_e \beta_h \gamma_{Je} \gamma_J g_3(E), \tag{59}$$

$$W_J = \beta_h^2 \gamma_J / m_0, \tag{60}$$

$$\frac{1}{(M_{Jc})^2} = + \mathcal{C}_{J2} \frac{\mathcal{O}_{Ja}(K_i K_j)}{K^2} + \mathcal{C}_{J4}\, \mathcal{O}_{Ja}(K^4) + \mathcal{C}_J + 4W_2^2 + [\mathcal{C}_{J1} + 12(W_3^2 - W_2^2)] \frac{\mathcal{O}_1(K^4)}{K^4}. \tag{61}$$

In the above formulas $J = 2$ for a[001]ellipsoid, and $J = 3$ for a[111]ellipsoid. For a spherical conduction band the value of J is irrelevant since all quantities involving J are zero by virtue of $\gamma_{Je} = 0$.

$$\begin{aligned} &\mathcal{C}_2 \equiv \tfrac{2}{9} V_2^2, \quad \mathcal{C}_3 \equiv \tfrac{1}{6} V_3^2 \\ &\mathcal{C}_{21} \equiv -\tfrac{1}{3} V_2^2 + \tfrac{1}{6} V_3^2, \quad \mathcal{C}_{31} \equiv \tfrac{1}{9} V_2^2, \\ &\mathcal{C}_{22} \equiv \tfrac{1}{9} V_2^2 - \tfrac{4}{3} V_2 W_2, \quad \mathcal{C}_{32} \equiv \tfrac{1}{12} V_3^2 - \tfrac{1}{3} V_2 W_2 - V_3 W_3, \\ &\mathcal{C}_{24} \equiv \tfrac{1}{12} V_3^2 - 2 V_2 W_2 + 2 V_3 W_3, \\ &\mathcal{C}_{34} \equiv -\tfrac{1}{9} V_2^2 + \tfrac{1}{3} V_3^2 + 2 V_2 W_2 - 2 V_3 W_3, \end{aligned} \tag{62}$$

where

$$\begin{aligned} &\mathcal{O}_{2a}(K_i K_j) \equiv (2K_z^2 - K_x^2 - K_y^2), \\ &\mathcal{O}_{3a}(K_i K_j) \equiv 2(K_x K_y + K_x K_z + K_y K_z), \\ &\mathcal{O}_{2a}(K^4) \equiv 2K_x^2 K_y^2 - K_x^2 K_z^2 - K_y^2 K_z^2, \\ &\mathcal{O}_{3a}(K^4) \equiv K_x K_y K_z^2 + K_x K_z K_y^2 + K_y K_z K_x^2, \end{aligned} \tag{63}$$

$$\mathcal{O}_1(K^4) \equiv K_x^2K_y^2 + K_x^2K_z^2 + K_y^2K_z^2 .$$

The forms of Eqs. (55), (56), (58) and (61) are much simpler for the case of a spherical conduction band so they will be rewritten explicitly for the sake of convenience:

$$E_{1s\pm}(\vec{\mathrm{K}}) = R_a + \frac{\hbar^2K^2}{2}\left(\frac{1}{M_a} \pm \frac{1}{M_c}\right) , \tag{64}$$

$$R_a = -R_0^* - g_1(E)(2\gamma_2^2 + 3\gamma_3^2) , \tag{65}$$

$$\frac{1}{M_a} = \frac{1}{M_0} - 8g_3(E)\,\beta_h^2\,(\tfrac{2}{3}\gamma_2^2 + \gamma_3^2) , \tag{66}$$

$$\frac{1}{M_c} = \frac{2\beta_h^2}{m_0}\left(\gamma_2^2 + 3(\gamma_3^2 - \gamma_2^2)\frac{\mathcal{O}_1(K^4)}{K^4}\right)^{1/2} . \tag{67}$$

Since Eqs. (52)–(67) are our principal results we list the equations in which the symbols have been defined. γ_{Je} is the conduction-electron reciprocal "anisotropy" mass defined in the manner of LB in Eq. (18). β_e and β_h are center-of-mass parameters defined by Eqs. (25) and (19). R_0^* and M_0 are defined in Eqs. (28)–(31). $g_1(E)$ and $g_3(E)$ are defined in Eqs. (49) and (50). $\vec{\mathrm{K}}$ is the total exciton momentum apart from $\vec{\mathrm{K}}_0$ in Eq. (27) which represents the momentum separation of the band edges.

Equation (52) is expected to be valid for exciton energies small compared to the effective Rydberg R_0^* in Eq. (29).

The exciton splitting at $\vec{\mathrm{K}} = 0$ is given by $2R_{Jb}$ as defined in Eq. (57) in agreement with the result obtained by LB. Equations (52)–(54) are clearly nonparabolic owing to the splitting of the degeneracy by R_{Jb}. Simple parabolic limiting cases are easily obtained for kinetic energies large or small compared to R_{Jb}: For small $\vec{\mathrm{K}}$,

$$E_{1s\pm}(K) = R_a \pm |R_{Jb}| + \frac{\hbar^2K^2}{2}\left(\frac{1}{M_a} \pm \frac{\mathrm{sgn}(R_{Jb})}{M_{Jb}}\right) ; \tag{68}$$

for intermediate K,

$$E_{1s\pm}(K) = R_a \pm R_{Jb}\,\frac{M_{Jc}}{M_{Jb}} + \frac{\hbar^2K^2}{2}\left(\frac{1}{M_a} \pm \frac{1}{M_{Jc}}\right) . \tag{69}$$

Of course, Eq. (69) is valid only for kinetic energies small compared to the exciton Rydberg. We will derive formulas valid in the very-large-kinetic-energy regime by a kinetic-energy band decoupling method discussed in Sec. V.

When the perturbation correction is small it is sufficient to evaluate $g_1(E)$ and $g_3(E)$ in Eqs. (49) and (50) at $E = -R_0^*$. For larger perturbations a self-consistent value for E obtained by iteration is more accurate. $S_1(E)$ has been tabulated by LB. We tabulate $S_3(E)$ in Table I.

V. LARGE-EXCITON-MOMENTUM LIMIT

The relative momentum of the electron and hole about the center of mass is of the order $\hbar/a_0^*$, where $a_0^* = a_0/\gamma_{1r}\epsilon$ is the effective Bohr radius. When the exciton momentum $\hbar\vec{\mathrm{K}}$ is large compared to $\hbar/a_0^*$ it is possible to take a different approach from that of Sec. IV. We make a transformation similar to (23),

$$\vec{\mathrm{k}}_e = \vec{\sigma}_0 + \vec{\mathrm{K}}, \quad \vec{\mathrm{k}}_h = \vec{\sigma}_0, \tag{70}$$

which still satisfies Eq. (26). The idea is to treat $\vec{\sigma}_0$ as a constant which will be the average value of the hole momentum. The variable $\vec{\sigma}$ will then be expanded about $\vec{\sigma}_0$ in obtaining the effective Hamiltonian for the exciton wave function.

We determine $\vec{\sigma}_0$ by minimizing the total kinetic energy of electron and hole. For simple bands

$$E_{\mathrm{KE}} = \frac{\hbar^2}{2m_e}(\vec{\sigma}_0 + \vec{\mathrm{K}})^2 + \frac{\hbar^2\vec{\sigma}_0^2}{2|m_h|} . \tag{71}$$

The requirement $\partial E_{\mathrm{KE}}/\partial\sigma_i = 0$ yields

$$\vec{\sigma}_0 = -\frac{|m_h|\vec{\mathrm{K}}}{m_e + |m_h|} ,$$

$$\vec{\mathrm{k}}_e = \frac{m_e\vec{\mathrm{K}}}{m_e + |m_h|} , \tag{72}$$

$$E_{\mathrm{KE}} = \frac{\hbar^2K^2}{2(m_e + |m_h|} ,$$

which is the same result as would be obtained by making a center-of-mass transformation.

In the general case we substitute Eqs. (70) into Eqs. (3) and (4) to obtain

$$H_{\mathrm{KE}} = \frac{\hbar^2}{2}\left(\frac{(\sigma_{0\parallel} + K_\parallel)^2}{m_\parallel} + \frac{(\sigma_{0\perp} + K_\perp)^2}{m_\perp}\right) + \frac{\hbar^2\gamma_1\sigma_0^2}{2m_0} - T_{dh}(\sigma_{0i}\sigma_{0j}). \tag{73}$$

We diagonalize the expression for T_{dh} using Eq. (11b) and obtain

$$E_{\mathrm{KE}}^{\mp} = \frac{\hbar^2}{2}\left(\frac{(\sigma_{0\parallel} + K_\parallel)^2}{m_\parallel} + \frac{(\sigma_{0\perp} + K_\perp)^2}{m_\parallel} + \frac{\sigma_0^2\gamma_1}{m_0} \pm \frac{2}{m_0}\left[\gamma_2^2\sigma_0^4 + 3(\gamma_3^2 - \gamma_2^2)(\sigma_{0x}^2\sigma_{0y}^2 + \sigma_{0x}^2\sigma_{0z}^2 + \sigma_{0y}^2\sigma_{0z}^2)\right]^{1/2}\right). \tag{74}$$

We first treat the simple case of valence-band isotropy, $\gamma_2 = \gamma_3$. Then the conditions $\partial E_{\mathrm{KE}}/\partial\sigma_i = 0$ give

$$\sigma_{0\parallel\pm} = -\frac{K_\parallel|m_{h\pm}|}{m_\parallel + |m_{h\pm}|} , \tag{75}$$

$$\sigma_{0\perp\pm} = -\frac{K_\perp|m_{h\pm}|}{m_\perp + |m_{h\pm}|} ,$$

$$\frac{1}{|m_{h\pm}|} \equiv \frac{\gamma_1}{m_0} \pm \frac{2\gamma_3}{m_0} . \tag{76}$$

Substituting these values in Eq. (74) gives for the total kinetic energy of the average electron and hole momentum

$$E^{\pm}_{\mathrm{KE}}=\frac{\hbar^2}{2}\left(\frac{K_{\parallel}^2}{m_{\parallel}+|m_{h\pm}|}+\frac{K_{\perp}^2}{m_{\perp}+|m_{h\pm}|}\right). \tag{77}$$

To this average kinetic-energy term we must add the binding energy of the exciton calculated with the appropriate effective masses. We obtain these by writing

$$\vec{\sigma}=\vec{\sigma}_0+\delta\vec{\sigma}, \quad \vec{k}_e=\vec{\sigma}+\vec{K} \quad \vec{k}_h=\vec{\sigma}. \tag{78}$$

We must recognize that we are still dealing with a four-band Hamiltonian as given by Eq. (7) with σ_i substituted in Eqs. (13). The diagonalization of these bands is determined by $\vec{\sigma}_0$ but not by $\delta\vec{\sigma}$. We ignore interactions between bands involving $\delta\vec{\sigma}$. Because of the isotropy condition $\gamma_2=\gamma_3$ we can take any direction to be the "z direction" in Eqs. (13) and (7). The simplest is to let $\hat{z}$ be $\hat{\sigma}_0$, a unit vector in the direction of $\vec{\sigma}_0$. This is generally not the direction of $\vec{K}$ unless $m_{\parallel}=m_{\perp}$ or when $\vec{K}=K_{\parallel}$ or $K_{\perp}$. We can write an expression for the total kinetic-energy operator in the form

$$H^{\pm}_{\mathrm{KE}}=E^{\pm}_{\mathrm{KE}}+\frac{\hbar^2}{2}\left(\frac{(\delta\sigma_{\parallel})^2}{m_{\parallel}}+\frac{(\delta\sigma_{\perp})^2}{m_{\perp}}+\frac{(\delta\sigma)^2\gamma_1}{m_0}\pm\frac{\gamma_3}{m_0}\left\{3[(\delta\vec{\sigma})\cdot\hat{\sigma}_0]^2-(\delta\sigma)^2\right\}\right). \tag{79}$$

$\hat{\sigma}_0$ is a unit vector in the direction of $\vec{\sigma}_0$. Unless $K_{\parallel}=0$ or $K_{\perp}=0$ the principal axes for H_{KE} as a function of $\delta\vec{\sigma}$ in Eq. (79) are not the parallel and perpendicular directions. A transformation to principal axes must be made, after which the exciton binding energy can be estimated quite accurately by use of the Kohn-Luttinger variational form.[11]

The general principal-axis transformation in two dimensions is easily written:

$$ax^2+2bxy+cy^2=\alpha_+\xi^2+\alpha_-\eta^2, \tag{80}$$

where (x, y) has been rotated into (ξ, η) and

$$\alpha_{\pm}=\frac{a+c}{2}\pm\left[\left(\frac{a-c}{2}\right)^2+b^2\right]^{1/2}. \tag{81}$$

The substitution of (75) for $\vec{\sigma}_0$ in Eq. (79) yields a form like the left-hand side of (80) in $\delta\sigma_{\parallel}$ and $\delta\sigma_{\perp}$. The principal masses are then determined with the use of (81).

We write the results schematically:

$$H^{\pm}_{\mathrm{KE}}=E^{\pm}_{\mathrm{KE}}+\frac{\hbar^2}{2}\left[\alpha^{\pm}_{+}(\delta\sigma_+)^2+\alpha^{\pm}_{-}(\delta\sigma_-)^2\right], \tag{82}$$

where $\alpha_{\pm}$ are reciprocal masses. If $\vec{K}=K_{\parallel}$ or $\vec{K}=K_{\perp}$ the ± directions are parallel and perpendicular, and we can write

$$\alpha^{\pm}_{\parallel}=\frac{1}{m_{\parallel}}+\frac{\gamma_1}{m_0}\pm\left(\frac{\gamma_3}{m_0}\right)2, \qquad \alpha^{\pm}_{\perp}=\frac{1}{m_{\perp}}+\frac{\gamma_1}{m_0}\pm\left(\frac{\gamma_3}{m_0}\right)(-1) \tag{83}$$

for $\vec{K}=K_{\parallel}$;

$$\alpha^{\pm}_{\parallel}=\frac{1}{m_{\parallel}}+\frac{\gamma_1}{m_0}\pm\left(\frac{\gamma_3}{m_0}\right)(-1), \qquad \alpha^{\pm}_{\perp}=\frac{1}{m_{\perp}}+\frac{\gamma_1}{m_0}\pm\left(\frac{\gamma_3}{m_0}\right)2 \tag{84}$$

for $\vec{K}=K_{\perp}$.

Since the valence band is isotropic, Eqs. (83) and (84) probably represent extreme limits for α_+ and α_-.

In summary we can say that for the case of the isotropic valence band $\gamma_2=\gamma_3$ we get results, Eqs. (75)-(77), for the translational kinetic energy which are exactly analogous to the case of simple bands, Eq. (72), with the obvious generalizations. For the reduced masses the multiband nature of the valence band leads to the results given in Eqs. (83) and (84), which are more complicated than the reduced mass in the case of simple bands since they depend on the direction of $\vec{K}$.

Equation (75) shows that $\vec{\sigma}_0$ is not in general parallel to $\vec{K}$. It is the direction of $\vec{\sigma}_0$ that determines the orientation of the degenerate valence-band wave functions. The heavy-mass excitons [minus sign in Eq. (76)] will correspond to valence bands $u^*_{\nu(\pm 3/2)}(\vec{r}_h)$, where the axis of quantization is parallel to $\vec{\sigma}_0$. The choice of direction for $\vec{\sigma}_0$ would not have been obvious without use of the energy-minimization principle except when $\vec{K}$ is in a symmetry direction.

The general case $\gamma_2\neq\gamma_3$ is complicated for an arbitrary $\vec{K}$ direction even for a spherical conduction band. The symmetry directions $\vec{K}=(0,0,K)$, $(K/\sqrt{3})(1,1,1)$, $(K/\sqrt{2})(1,1,0)$ and their cubic equivalents are simple however. When these directions are parallel or perpendicular to the principal axes the results remain simple even for ellipsoidal surfaces. The results are easily shown to be

$$\sigma_0=-\frac{\vec{K}|m_{h\pm}|}{(m_{c\vec{K}}+|m_{h\pm}|)}, \tag{85}$$

$$E^{\pm}_{\mathrm{KE}}=\frac{\hbar^2}{2}\,\frac{K^2}{m_{c\vec{K}}+|m_{h\pm}|}. \tag{86}$$

The hole masses are those appropriate to the symmetry directions, namely,

$$\frac{m_0}{|m_{h\pm}|}=\gamma_1\pm 2\gamma_2, \qquad \vec{K}=(K,0,0) \tag{87}$$

$$\frac{m_0}{|m_{h\pm}|}=\gamma_1\pm 2\gamma_3, \qquad \vec{K}=\frac{1}{\sqrt{3}}(K,K,K) \tag{88}$$

TABLE II. Parameters used in calculation and $K=0$ exciton binding energies by ordinary perturbation theory. E and δE are given in units of R_0^*, while R_0^* is given in meV. Numbers in parentheses come from Brillouin–Wigner perturbation theory.

	γ_1	γ_2	γ_3	m	m	ϵ_0	$R_0^*/(\text{meV})$	$E_{\pm 1/2}$	$E_{\pm 3/2}$	δE
Si	4.280	0.375	1.450	0.9163	0.1905	11.4	12.85	−1.081	−1.103	0.0226
								(−1.079)	(−1.100)	(0.0211)
Ge	13.38	4.24	5.69	1.588	0.08152	15.36	2.65	−1.158	−1.292	0.134
								(−1.149)	(−1.265)	(0.115)
GaAs	7.65	2.41	3.28	0.067	0.067	12.5	3.86	−1.061	−1.061	0

$$\frac{m_0}{|m_{h\pm}|}=\gamma_1\pm(3\gamma_3^2+\gamma_2^2)^{1/2}\,,\ \vec{\mathrm{K}}=\frac{1}{\sqrt{2}}\,(K,K,0). \qquad (89)$$

The conduction masses are the principal masses in the direction of $\vec{\mathrm{K}}$. The formulas only apply for principal directions, i.e., silicon (0, 0, 1), (1, 0, 0), (1, 1, 0) and germanium (1, − 1, 0), (1, 1, 1). In less-symmetric directions the direction of σ_0 is not the direction of $\vec{\mathrm{K}}$ even for a spherical conduction band.

TABLE III. Reciprocal exciton masses in free-electron-mass units for principal K directions for low-, intermediate-, and high-momentum values. Values in parentheses are obtained from Brillouin-Wigner perturbation theory. At a given K the upper value refers to the exciton band of highest energy. For germanium [111] the exciton bands cross. The silicon ellipsoid is oriented in the [001] direction while the germanium ellipsoid is oriented along [111].

Substance	K	M^{*-1}	low K	M^{*-1} int K	M^{*-1} high K
Ge	[001]	1.86	(2.16)	4.54	
		3.78	(4.00)	1.10	
Ge	[111]	−0.867	(−0.426)	−0.404	0.614
		−0.404	(−0.255)	−0.867	0.479
Ge	$[1\bar{1}0]$	3.22	(3.46)	7.04	8.13
		5.87	(6.12)	2.05	2.18
Ge	[110]	0.495	(0.867)	2.45	
		1.69	(1.87)	−0.272	
Ge	$[11\bar{1}]$	2.77	(3.02)	6.39	
		5.17	(5.41)	1.55	
Si	[001]	0.681	(0.719)	0.681	0.897
		0.598	(0.629)	0.598	0.833
Si	[100]	1.87	(1.89)	2.22	2.57
		2.10	(2.12)	1.75	2.11
Si	[111]	1.47		2.22	
		1.60		0.858	
Si	[110]	1.87		2.77	2.97
		2.10		1.20	1.31
Si	[011]	1.28		1.78	
		1.35		0.841	
GaAs	[001]	5.53			6.80
		1.32			2.38
GaAs	[111]	6.29			7.30
		0.556			1.02
GaAs	[110]	6.12			7.19
		0.726			1.35

In all these cases the translational masses have the same form as in simple bands. We do not evaluate the reduced masses, which we have seen to be more complicated even in the case of valence-band isotropy. That approximation can generally be used to obtain a good estimate of the effect of Coulomb binding. For directions of lower symmetry we will not have $\sigma_0\sim\vec{\mathrm{K}}$ as in Eq. (85) and the algebraic relations are much more complicated. The symmetric results lead to extremal values which should in many cases serve as bounds for the more complicated directions.

In this kinetic-energy decoupling method we have ignored the kinetic-energy terms in $\delta\sigma$ coupling different bands. In this way each "decoupled exciton band" has its translational kinetic energy plus a complete set of excited orbital states which, however, are not hydrogenic because of the ellipsoidal mass. The two sets of bands may be referred to as "light" and "heavy" excitons according to the hole mass involved in Eq. (86). The decoupling scheme improves for the heavy 1s exciton band as its kinetic energy increases. For the light exciton bands the approximation gets worse initially because the light exciton 1s band will tend to intersect the heavy exciton excited-state bands. For very high energies compared to exciton binding energies the coupling will become unimportant. The treatment of the light excitons would be improved by retaining the coupling terms to the heavy excitons. A perturbation method might still be used but we have not attempted to carry this approach any further.

VI. DISCUSSION

We illustrate the results of Secs. IV and V by calculating numerical values for the exciton masses in the principal symmetry directions [001], [110], [111] for the case of silicon, germanium, and gallium arsenide. The mass parameters used are given in Table II and the exciton reciprocal masses are given in Table III. The three columns labeled small K, intermediate K, and large K refer to the approximations in Eq. (68), Eq. (69), and Eqs. (86)–(89), respectively.

The use of ordinary perturbation theory corresponds to setting $E=-R_0^*$ in Eqs. (49) and (50). For germanium the perturbation method is most suspect. This is indicated by the magnitude of the perturbation correction to the unperturbed value $-R_0^*$ in the starting approximation. (See Table II.) For the [111] direction the perturbation method leads to negative masses, a most unlikely result. An improvement can be obtained by using Brillouin-Wigner perturbation theory, namely, by treating E in Eqs. (49) and (50) as the exciton energy determined self-consistently from the perturbation expressions. We have done this using the value $E(\mathrm{K}=0)$ but not attempting to correct the finite K energies in this way. The [111] masses are substantially reduced in absolute value but remain negative. The lighter masses are least affected. In silicon the changes are very small and in gallium arsenide should be even smaller. Even for germanium the perturbation results are good enough to be useful. For the [111] direction the average of the two large-K reciprocal masses gives an upper bound to the average of the true reciprocal-mass values. The actual reciprocal-mass values are smaller than the values indicated owing to interband coupling terms which have been left out.

In germanium more recent work has presented a calculational alternative to the perturbation method.[2] The results are not expected to be highly accurate but are preferable to negative masses.

In Table III the upper most number of a given K refers to the mass of the exciton band whose energy is highest in that K regime. At high and intermediate K there is a heavy- and light-mass exciton band in accord with the simple idea of adding electron and light- and heavy-hole masses to get the translational mass. This result follows from the fact that the axis of quantization of the valence wave functions follows the direction of the average hole momentum at high $\vec{\mathrm{K}}$. At low $\vec{\mathrm{K}}$ the direction of quantization is parallel to the axis of the electron ellipsoid. (The direction of quantization is determined by the minimization of the total energy, which for high $\vec{\mathrm{K}}$ is dominated by the translational kinetic energy and at low $\vec{\mathrm{K}}$ by the exciton binding energy.)

At low $\vec{\mathrm{K}}$ the lower state is heavy-hole-like (light-hole-like) for a $\vec{\mathrm{K}}$ direction parallel (perpendicular) to the electron ellipsoid. For the upper exciton the situation is just the reverse. If one increases the momentum from zero in the transverse direction the direction of quantization will rotate from perpendicular to parallel to the $\vec{\mathrm{K}}$ vector. This leads to the "mass-reversal" effect which is quite noticeable in Table III and in Fig. 1. For a $\vec{\mathrm{K}}$ direction parallel to the ellipsoid axis the direction of quantization remains fixed and the bands are parabolic. (This also follows from group theory since the two excitons have different symmetry for $\vec{\mathrm{K}}$ in the parallel direction and will not mix).

In the [111] direction in germanium the exciton bands cross according to Table III. This result is not found in Ref. 2 and appears to be an artifact of the perturbation calculation. The mass-reversal effect results in strongly nonparabolic effects in the dispersion and the density of states which are described in more detail in Ref. 2.

The high-K and intermediate-K masses are not too different in Table III. This indicates that the dominant effect is the orientation of the axis of quantization. The interband coupling is neglected at high K but is included perturbatively at intermediate K. These terms do not appear to be too important, particularly for the lighter masses.

In gallium arsenide the results are simplest since there is no splitting of the exciton degeneracy at K = 0 and hence no intermediate-K regime. The large-K reciprocal masses are larger than the low-K values owing to the neglect of interband coupling terms. For the heavy-mass band this result is in accord with the true physical situation; the interband coupling is strongly reduced by the large energy denominators. For the light-mass band at very high kinetic energy compared to the exciton binding the result is also true. However, in the regime where the light-mass 1s band intersects the heavy-mass 2p, 3d, etc., bands strong interactions will occur and the masses will be very inaccurate.

The difference between the masses in the different K regimes are substantial and have lead to observable nonparabolic effects.[2] Such effects occur for energies of the order of the exciton binding in GaAs but in Ge and Si they occur for the very much smaller energies associated with conduction-band-anisotropy-induced splitting of the K = 0 exciton degeneracy. In Ge and Si these splittings are only a few tenths of a millivolt.

ACKNOWLEDGMENT

The author would like to thank A. Baldereschi for valuable discussions and A. Frova and G. A. Thomas for stimulating conversations.

---

[1] G. Dresselhaus, J. Phys. Chem. Solids 1, 14 (1956).

[2] A. Frova, G. A. Thomas, R. E. Miller, and E. O. Kane (unpublished).

[3] D. G. Thomas and J. J. Hopfield, Phys. Rev. Lett. 5, 505 (1960).

[4] A. Baldereschi and N. O. Lipari, Phys. Rev. B 3,

439 (1971); N. O. Lipari and A. Baldereschi, Phys. Rev. B 3, 2497 (1971).

[5] E. O. Kane, *Physics of III-V Compounds* (Academic, New York, 1966), Vol. 1, p. 75.

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

[7] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 22, 833 (1967).

[8] G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. 98, 368 (1955).

[9] E. O. Kane, J. Phys. Chem. Solids 1, 83 (1956).

[10] J. M. Luttinger, Phys. Rev. 102, 1030 (1956).

[11] W. Kohn and J. M. Luttinger, Phys. Rev. 98, 915 (1955); E. O. Kane, Phys. Rev. 180, 852 (1969).