Naturforsch. 21, A461 (1966).

[37] W. E. Spear and P. G. LeComber, Phys. Rev. Letters 13, 434 (1964).

[38] Y. Fukuda and M. Fukai, J. Phys. Soc. Japan 23, 902 (1967).

PHYSICAL REVIEW B VOLUME 2, NUMBER 10 15 NOVEMBER 1970

Dielectric Theory of Impurity Binding Energies. III. Group-III Acceptors in Si and Ge

J. C. Phillips
Bell Telephone Laboratories, Murray Hill, New Jersey 07974

(Received 22 May 1970)

Chemical shifts of ground-state energies and *g* factors of acceptors X in Si or Ge host crystals are analyzed within the framework of effective-mass theory. The unit cell centered on the impurity is regarded as a unit cell of a hypothetical SiX of GeX crystal. The differences in energy levels of valence and conduction bands of this hypothetical crystal at Γ, X, and L are calculated according to the semiempirical spectroscopic rules developed by Phillips and Van Vechten to describe levels of zinc-blende crystals. These energy differences are compared with those of the host crystal and are used to renormalize effective masses in the impurity unit cell. Rough estimates then show that this approach yields chemical trends in good agreement with experiment and explains several quantitative features of the data that cannot be explained by qualitative models based on ionic radii or electronegativity differences.

I. INTRODUCTION

In two preceding papers,[1,2] an analysis of the chemical shifts of ground-state energies of donor impurities in Si, Ge, and GaP has been made based on a spectroscopic theory[3,4] of the covalent bond in tetrahedrally coordinated $A^N B^{8-N}$ semiconductors. Both experiment and theory[5] agree that the effective-mass approximation (EMA) of a point-charge impurity embedded in a dielectric quasicontinuum gives an excellent account of the energies and wave functions of excited states of shallow impurities in semiconductors. However, there is now abundant experimental evidence to show that the EMA fails both quantitatively and qualitatively to account for ground-state energies. Quantitatively there is the obvious point that ground-state energies vary from one donor impurity to another (or from one acceptor impurity to another) in the same host crystal, whereas according to the EMA the binding energies of all states are determined only by properties of the host crystal.

It has been customary[5] to explain these "chemical shifts" in terms of a "central cell correction," the breakdown of the hydrogenic approximation for the effective potential in the atomic cell containing the impurity. This brings us no closer to understanding ground-state energies, but it has the convenient feature of relegating the problem to another discipline which already has its own full quota of unsolved problems. The qualitative value of this classification, however, became doubtful when Hopfield and Thomas[6] discovered that even in III-V semiconductors, isoelectronic impurities (such as N in GaP) could bind electrons and holes although the effective impurity potential was *zero* in the EMA. Thus, the EMA is qualitatively wrong in this case, and it appears that we must face up to the problem of shallow impurity states associated with donors and acceptors as well as isoelectronic impurities or isoelectronic-impurity complexes (e.g., CdO in GaP).[6]

The first point to recognize is that in the presence of a short-range potential only, one would not generally expect to find shallow impurity states. Either the potential is not strong enough to produce any bound states, or else it is likely to produce states with a binding energy which is a significant fraction of the energy gap. In practice, unless the difference in electronegativity is very large (e.g., O in GaP), one usually finds small binding energies comparable to EMA binding energies for donors on acceptors and of the same order of magnitude for isoelectronic complexes.

The explanation[7] for this behavior, which for some time made the EMA appear to be more accurate than it really is, is that because most semiconductors are highly polarizable, a strain field develops around each impurity to prevent the accumulation of electronic charge much above or more below the requirements of the valence bonds of the host lattice.

In the case[1] of shallow donor impurities in Si or Ge, the central cell corrections $\Delta E_d = E_I^d - E_0^d$ do not vary monotonically with impurity size. Here E_I^d is the donor ground-state energy and E_0^d is the EMA ground-state energy. Instead, ΔE_d is found to reach a minimum value at Sb in both Si and Ge host crystals. This suggested to us that one could account[1,2] for the chemical shifts in ΔE_d primarily as an effect arising from differences between impurity-host bonds and host-host bonds. These differences can be characterized through the spectroscopic theory[3,4] of bonding in these crystals. It is a characteristic feature of the theory, which involves no adjustable parameters, that these differences are minimal for Sb in both Si and Ge.

In the case of shallow acceptor impurities, the central cell corrections $\Delta E_a = E_I^a - E_0^a$ do vary monotonically with impurity size. This suggests that the situation for hole wave packets made up of states localized near the valence-band edge at Γ is different from that of electron wave packets formed near the $\langle 100 \rangle$ conduction-band minima in Si, and the $\langle 111 \rangle$ conduction-band minima in Ge, or the $\langle 000 \rangle$ conduction-band minima in GaAs. However, and this is basic to our dielectric approach, if these mechanisms involve mainly polarization processes, they will involve *both* valence and conduction bands in each case, although the carriers in question may be either electrons or holes. Thus, we have one set of mechanisms for both donors and acceptors.

II. MICROSCOPIC AND MACROSCOPIC ASPECTS

In order to explain chemical shifts ΔE_d or ΔE_a in ground-state binding energies, one must introduce in into the theory at some point specific properties of the host atoms and the impurity atoms. The great strength of the macroscopic or continuum approach is that by forming wave packets near band edges, one needs only the curvatures or effective masses of the host crystal to solve a hydrogenic wave equation for the envelope wave-packet amplitude. These masses have been measured in some crystals by cyclotron resonance, so that one does not have to solve the crystal Schrödinger equation to find Bloch functions and energy bands.

At the time that the EMA was being developed formally (largely 1951–1955), very little was known about the energy bands of semiconductors. Today, the energy bands of these crystals are known with great accuracy, thanks largely to theoretical calculations using pseudopotentials calibrated against detailed measurements of fundamental optical spectra.[8] Moreover, the spectroscopic theory of chemical bonding has been extended to predict interband energy differences at Γ, X and L with an accuracy that surpasses that of pseudopotential calculations yet utilizes fewer parameters.

The program of this paper is to combine the macroscopic EMA theory with microscopic band theory to identify the physical mechanisms which determine chemical shifts. We have found that acceptor binding energies and g factors in Si and Ge provide the best field for our analysis. In Sec. III, we review the elements of EMA theory which are needed in our calculations, and in Sec. IV present the microscopic information which is needed. The remaining sections demonstrate how the combined approach works and show which features of the approach are necessary to expain the experimental data.

III. EMA BACKGROUND

The energy bands of Ge as calculated by the semiempirical pseudopotential method[8] in the absence of spin-orbit effects are shown in Fig. 1, and those of Si are shown in Fig. 2. The valence-band maximum at $\vec{k}=0$ is threefold orbitally degerate and is labeled $\Gamma_{25'}$. With the addition of spin, these p-like states split into Γ_8 (corresponding to $J=\frac{3}{2}$) and Γ_7 (corresponding to $J=\frac{1}{2}$) with a spin-orbit splitting $\frac{3}{2}\lambda$, where $\lambda\vec{L}\cdot\vec{S}$ is the spin-orbit interaction. The wave function $\Psi(r)$ of the impurity state can be written[5] in the EMA as

$$\Psi(r)=\sum_{j=1}^{6} F_j(\vec{R})\Phi_j(\vec{r}) \quad , \tag{3.1}$$

where the $\Phi_j(\vec{r})$ are suitable linear combinations of the Bloch states associated with Γ_8 and Γ_7. The envelope functions $F_j(\vec{R})$ depend on the discrete lattice coordinate $\vec{R}$ and have a hydrogenic form. Explicit forms suitable for variational calculations have been given by Schechter,[9] Mendelson and James,[10] and Suzuki, Okazaki, and Hasegawa[11] (SOH). We quote here chiefly formulas taken from SOH, who used the most recent cyclotron-resonance parameters and give the most complete discussion of Zeeman effects. In Si these parameters correspond to the case[12,13] $B<0$.

The effective-mass Hamiltonian can be written in a simple form using the effective-mass momentum and orbital and spin angular momentum operators $\vec{P}$, $\vec{L}$, and $\vec{\sigma}$:

$$H_{\text{eff mass}} = A\vec{P}^2 - 3B[(L_x^2 - \tfrac{1}{3}L^2) + \text{c.p.}]$$

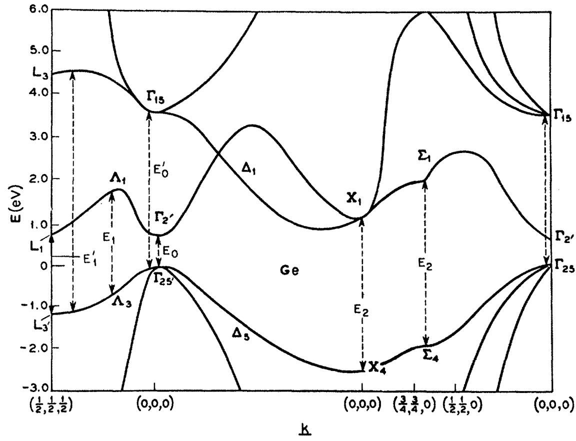

FIG. 1. The energy bands $E_n(\vec{k})$ of Ge as calculated by the semiempirical pseudopotential method [Brust, Phillips, and Bassani, Phys. Rev. Letters 9, 94 (1962)].

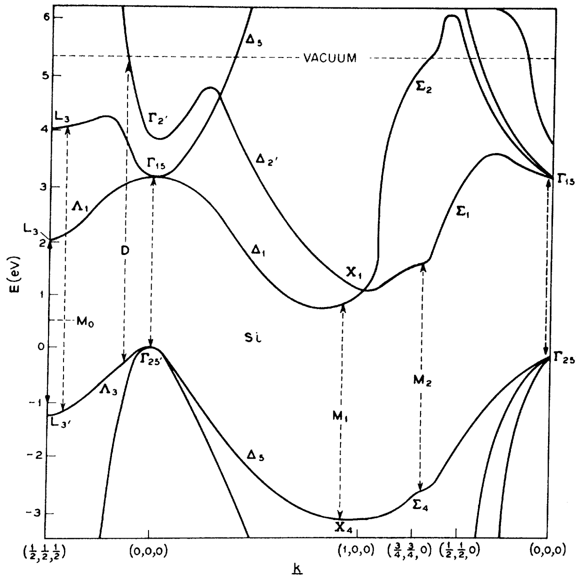

FIG. 2. The energy bands $E_n(\vec{k})$ of Si as calculated by the semiempirical method [Burst, Cohen, and Phillips, Phys. Rev. Letters 9, 389 (1962)].

$$- N[(L_xL_y + L_yL_x)P_xP_y + \text{c. p.}]$$
$$+ \tfrac{1}{3}\Lambda(\vec{L}\cdot\vec{\sigma} - 1) - e^2/\epsilon_0 r \quad , \qquad (3.2)$$

where c.p. denotes cyclic permutations. If $B = N = \Lambda = 0$, $H_{\text{eff mass}}$ reduces to the Hamiltonian of a hydrogen atom with a Coulomb interaction reduced by the reciprocal of the electronic dielectric constant ϵ_0^{-1} and effective mass $m^* = 1/2A$.

Because of the terms B and N in (3.2), even the $1s$-like ground state has some d-like components. According to SOH, numerical solution of the wave equation obtained from (3.2) gives the ground-state energies for Si and Ge shown in Table I. We can also use their wave functions to calculate the probability that an electron will be found in the unit cell of volume $\frac{1}{4}a^3$ centered on the impurity atom:

$$P_c = \tfrac{1}{4}\,|F_0(0)|^2 a^3 \quad . \qquad (3.3)$$

SOH denote the normalized amplitude of the s-wave component of $F_0(\vec{R})$ by c_0 and its hydrogenic radius by r_1. Thus, from their wave functions with $c_0 = 0.93$ in Si and 0.86 in Ge, one finds

$$P_c^a = (a/r_1)^3 c_0^2 \quad , \qquad (3.4)$$

where the superscript labels acceptors. Values of P_c calculated from $r_1 = 15$ Å in Si and 38 Å in Ge are also given in Table I.

To connect the EMA with band theory, we need to know the relations between A, B, and N and the band structure. These relations are[12]

$$A = 1 + \tfrac{1}{3}(F + 2G + 2H_1) \quad , \qquad (3.5)$$

$$3B = F + 2G - H_1 \quad , \qquad (3.6)$$

$$N = F - G + H_1 \quad . \qquad (3.7)$$

Here F and H are given by

$$F = \frac{\hbar^2}{2m}\,\frac{|\langle \Gamma_{2'}|\vec{p}|\Gamma_{25'}\rangle|^2}{E_0} \quad , \qquad (3.8a)$$

$$H_1 = \frac{\hbar^2}{2m}\,\frac{|\langle \Gamma_{15}|\vec{p}|\Gamma_{25'}\rangle|^2}{E_0'} \quad . \qquad (3.8b)$$

Here

$$E_0 = E(\Gamma_{2'}) - E(\Gamma_{25'})$$

and

$$E_0' = E(\Gamma_{15}) - E(\Gamma_{25'})$$

in the notation of Van Vechten.[14] The term G connects $\Gamma_{25'}$ with Γ_{12}, which is higher in energy and has a smaller matrix element, making G a factor of 10 smaller than F and H_1, and hence negligible for our purposes here. As discussed previously,[13] effective-mass data show that the interband momentum-matrix elements in (3.8a) and (3.8b) vary slowly from one tetrahedrally coordinated $A^N B^{8-N}$ crystal to another. Our interest therefore focuses on the interband energies E_0 and E_0'.

The g factor for the Zeeman effect is, in general, quite complicated. SOH give formulas for two limiting cases corresponding to $3B = N$ (which holds approximately for Ge) and to $\Lambda = 0$ (which is fairly good for acceptor states in Si; see Ref. 11, p. 939). When only the s-wave contribution to $F_0(\vec{R})$ is retained, i.e., $c_0 = 1$, these reduce to

$$g = -2\kappa = \tfrac{2}{3}(1 + \tfrac{1}{2}F - \tfrac{1}{2}H_1) \quad , \qquad (3.9)$$

which is also the free-hole g factor.

IV. HYPOTHETICAL BAND STRUCTURES

In the construction of a model for chemical shifts, the chief problem is to introduce the properties of the impurity into the theory in a way consitent with the properties of the electronic structure of the host lattice. There are a number of solutions to the impurity problem which, although formally correct, do not meet this requirement and must therefore be discarded. For example, one can expand the impurity wave function in Wannier or atomic orbitals,[15] but a few of those basis functions do not give an accurate representation of the wave functions of the semiconductor host crystal.[16] Thus, with energy gaps much smaller than valence bandwidths,

TABLE I. Effective-mass parameters in Si and Ge. The values of E_0^a and P_c are taken from Ref. 11.

Crystal	λ (meV)	E_0^a (meV)	P_c (10^3)
Si	30	35.7 ± 0.2	40 ± 1.5
Ge	200	10.0 ± 0.1	2.5 ± 0.1

as we have in many semiconductors, we cannot hope for success from this approach.

In discussing donors, we found it useful[1] to consider the average energy gap E_g of an impurity-host bond compared to host bonds. More generally, one can imagine constructing a hypothetical crystal AX having the sphalerite or zinc-blende structure, where X is the impurity atom and the A atoms are its four nearest neighbors. For a S donor replacing P in GaP, for example, one could consider the energy bands of a hypothetical GaS crystal. For such a crystal, one could calculate not only E_g but also the interband optical energies E_0, E_0', E_1, E_2, . . . discussed by Van Vechten.[14] He has given rules for calculating these energies in sphalerite crystals based almost entirely on the values observed for diamond-type crystals and the valences and atomic radii of the constituent atoms. We propose to use these rules to obtain the band structures of our hypothetical AX crystals.

It may be objected that Van Vechten's rules are designed from atom pairs $A^N B^{8-N}$, so that the net core charge per atom pair is also eight, and the number of s-p valence electrons per atom pair is also eight. The latter, but not the former, condition is satisfied for $A^N X^{8-N\pm1}$ pairs. However, we feel that for substitutional impurities the rules should give useful results because the impurity as well as host atoms see structurally similar tetrahedrally coordinated environments. Moreover, the algebraic simplicity of the rules guarantees the kind of uniformity which is necessary for analyzing chemical trends.

As an example, consider Sb in Si. From a table of covalent radii[17] we estimate that $d_{\text{Si-Sb}}$ is 2.578 Å compared to $d = 2.346$ Å in Si. (In this case, no correction is needed for core-effect expansion because only one of the atoms exhibits this effect. When both do, the smaller expansion parameter should be deleted.[17]) We can now calculate $E_{h,i}$, where $i = 0$, $0'$, 1, 2, from

$$E_{h,i} = (E_{h,i})_{\text{Si}}\,(2.578/2.346)^{s_i} \quad , \tag{4.1}$$

where the values of $(E_{h,i})_{\text{Si}}$ and s_i are given in Table II of Ref. 14. In this way, we obtain $E_{h,0} = 3.2$ eV, $E_{h,0}' = 2.8$ eV, $E_{h,1} = 2.9$ eV, $E_{h,2} = 3.6$ eV. Next, we compute C, the ionic energy, by arguing that this must be about half as large in a Si-Sb bond as in an Al-Sb bond. [We could also have used Eq. (1.5) of Ref. 14 with $b = 1.5$, but we prefer to rely on experiment whenever possible.] This gives $C = 1.5$ eV. Finally, we use from Ref. 14, Eq. (3.1), C, $E_{h,0}'$, and $E_{h,2}$ to compute E_0' and E_2 and Eq. (3.7), C, $E_{h,0}$, $E_{h,1}$, Δ_{av}, ΔE_0, and ΔE_1 from Table II of Ref. 14 to compute E_0 and E_1. The results are: $E_0 = 0.2$ eV, $E_0' = 3.2$ eV, $E_1 = 2.0$ eV, $E_2 = 3.9$ eV. The complete results for group-III acceptors and group-V donors are shown in Table II for Si and in Table III for Ge.

TABLE II. Interband energy differences in eV in hypothetical SiX crystals, calculated using the methods of Ref. 14.

X	E_0	E_1	E_0'	E_2
Si	4.1	3.6	3.4	4.5
P	4.6	4.1	3.8	4.9
B	8.4	6.4	4.6	8.4
Al	4.2	3.7	3.5	5.0
As	2.9	3.1	3.5	4.5
Sb	0.2	2.0	3.2	3.9
Bi	−1.5	1.2	3.0	3.8
Ga	2.9	3.3	3.6	4.6
In	0.6	2.1	3.2	4.0
Tl	−0.4	1.75	3.2	3.9

There are several remarks that should be made at this point. Although Van Vechten's rules may appear novel, it has been demonstrated[14] that they describe interband structure in $A^N B^{8-N}$ semiconductors more than twice as accurately as empirically adjusted pseudopotential band calculations do. (The respective rms errors are about 0.2 and 0.4 eV, respectively.) Thus, all the chemical shifts in the host crystals are given accurately. Moreover, in a few minutes of hand calculation, one does the equivalent of an *entire* computer band calculation, using rules of such algebraic simplicity as to assure the uniformity needed for discussion of chemical trends. Finally, if we compare the values of E_0 and E_2 in Sb : Si (0.2, 3.9) with pure Si (4.1, 4.5), we see immediately the great drop in E_0/E_2 which is characteristic of the heavy metal Sb and which occurs even more dramatically in the Pb row atoms. (The parameters for the latter are given in the Appendix to Ref. 18.)

V. ANALYSIS OF $\Delta E_a(X)$ IN Ge AND Si

It has long been customary[5] to regard the chemical shifts $\Delta E_a(X)$ as associated with a potential δV localized in the central atomic cell of the X impurity. The Fourier transform of this potential will have large components $\delta V_{\vec{k}}(X)$ for $k \lesssim \frac{1}{2}G$, where G is an average diameter of the Brillouin zone. One would therefore expect δV, regarded as a perturbation, to mix into $\Psi_a(\vec{r})$ Bloch states from the low-lying conduction bands *throughout* the Brillouin zone. The analysis of chemical shifts for acceptors given in this section shows that this is not the case. The analysis leads to quite a different physical model for the important part of $\delta V(X)$ which is discussed at the end of this section.

In Table IV, the spectroscopically measured values[19] of $\Delta E_a(X)$ in Ge are listed for X = B, Al, Ga, In, and Tl. These may be compared with the values of E_0, E_0', E_1, and E_2 listed for GeX in Table III. Note particularly, the variation

TABLE III. Interband energy differences in eV in hypothetical GeX crystals, calculated using the methods of Ref. 14.

X	E_0	E_1	E_0'	E_2
Ge	1.0	2.2	3.14	4.1
P	2.9	3.4	3.7	4.8
As	1.2	2.4	3.4	4.35
Sb	0.7	2.0	2.9	3.7
Bi	−2.1	0.9	−2.8	3.4
B	2.4	3.4	4.3	6.1
Al	2.35	3.0	3.4	4.3
Ga	1.2	2.5	3.4	4.35
In	0.5	2.0	3.0	3.7
Tl	−1.2	1.8	3.25	3.5

$$\Delta E_a(\mathrm{In}) - \Delta E_a(\mathrm{Ga}) = 0.6 \text{ meV}$$

compared to

$$\Delta E_a(\mathrm{Tl}) - \Delta E_a(\mathrm{In}) = 1.5 \text{ meV} \;.$$

Among the four interband energies listed for the hypothetical GeX band structures, *only* E_0 shows this qualitative trend. Indeed the other interband energies show a larger shift from In to Ga than from Tl to In, whereas the shifts in E_0 and ΔE_a are about 2.5 times larger in the latter case than in the former.

In Table V, the spectroscopically measured values[20] of $\Delta E_a(X)$ in Si are listed for X=B, Al, Ga, and In. Here the striking experimental fact is the large shift in $\Delta E_a(X)$ from X = In to X = Ga. Again among the four interband energies listed for the hypothetical SiX band structures, *only* E_0 shows this qualitative trend.

The interband energies E_0, E_0', E_1, and E_2 all describe direct or $\vec{k}$-conserving optical transitions of the "vertical" type shown in Figs. 1 and 2. However, using Van Vechten's rules,[14] one can also compute indirect gap energies between $\Gamma_{25'}$ and L_1 or X_1 conduction-band states. These also do not show the required chemical shifts for heavy impurities (Tl in Ge or In in Si).

The foregoing analysis demonstrates that variations in

$$E_0 = E(\Gamma_{2'}) - E(\Gamma_{25'})$$

are responsible for an important part of the X dependence of $\Delta E_a(X)$. The analysis can be made more quantitative by making the Taylor-series expansion

$$E_I^a(X) = E_0^a + \delta E_0^a + a_1 \Delta E_0(X) + a_2^2[\Delta E_0(X)]^2 \;, \tag{5.1}$$

where

$$\Delta E_0(X) = E_0(AX) - E_0(AB) \;, \tag{5.2}$$

assuming that X has substituted for B. The physical meaning of the parameters is that δE_0^a is the central-cell correction which would be present in the absence of any difference in E_0 in the AX and AB crystals; it represents an average correction coming from differences in all other energies, a difference which depends so little on X that it can be treated as a constant. The term which is linear in ΔE_0 is the one of major interest; it has the nature of a mass renormalization effect (see below).

The values of δE_0^a, a_1, and a_2 for Si and Ge are listed in Table VI. The results obtained from (5.1) are compared with the experimental values in Tables IV and V. For X in Ge, the agreement is excellent. For Si, the agreement is good, except that the actual shift for Al to Ga is ten times smaller than the predicted shift. This will be discussed in Sec. VI.

We do not attach much meaning to the value of a_2, which is included in the expansion (5.1) to give good values of δE_0^a and a_1. The ratios of the two latter parameters in Si compared to Ge are both about 25. According to Table I, the ratio $P_c(\mathrm{Si})/P_c(\mathrm{Ge})$ is about 16. Thus, both δE_0^a and a_1 scale approximately with the macroscopic parameter P_c, although there is some evidence for nonlinearity in the greater values of δE_0^a and a_1 in Si.

Now we turn to the physical meaning of the terms δE_0^a and $a_1 \Delta E_0$ in Eq. (5.1). At first sight, the appearance of these terms is paradoxical, because they seem to imply that the perturbation associated with the impurity has a very long range. Indeed, as we shall see in Sec. VI, one may formally interpret the term $a_1 \Delta E_0$ as a mass renormalization term. Such a term is, however, almost entirely absent for excited p states of donors,[21] and it is therefore difficult to see how it could be present for the ground states discussed here.

This paradox may be developed further. Suppose one rejects the analysis given here and instead retains the Kohn-Luttinger model[5] based on a strictly hydrogenic potential for $r \gtrsim a$, where a is of order the atomic radius. Then the central-cell correction should be proportional to $\Omega(X)/\Omega(\text{host})$, where $\Omega(X)$ is the atomic volume calculated from a table of covalent radii.[7] These ratios are shown in Tables IV and V, and it is clear that they do not give a satisfactory account of experimental trends in $E_I^a(X)$.

In analyzing interband spectra in more than 20 semiconductors,[8] it has been our experience that

TABLE IV. Acceptor binding energies for X in Ge. Experimental data are from Ref. 18, theoretical values from Eq. (5.1).

Ge : X	B	Al	Ga	In	Tl
$E_I^a(X)_{\mathrm{exp}}$ (meV)	10.5	10.8	11.0	11.6	13.1
$E_I^a(X)_{\mathrm{theory}}$ (meV)	10.7	10.8	11.1	11.6	13.0
$\Omega(X)/\Omega(\mathrm{Ge})$	0.34	1.00	1.00	1.50	1.69

TABLE V. Acceptor binding energies for X in Si. Experimental data are from Ref. 19, theoretical values from Eq. (5.1).

Si:X	B	Al	Ga	In	Tl
$E_I^a(X)_{\text{exp}}$ (meV)	44	69	72	155	$\cdots$
$E_I^a(X)_{\text{theory}}$ (meV)	52	69	92	153	188
$\Omega(X)/\Omega$ (host)	0.39	1.15	1.15	1.73	1.94

arguments based on chemical trends are generally much more reliable than *ab initio* theoretical models, no matter how carefully the latter are constructed. We are therefore reluctant to ignore the good agreement demonstrated here between $\Delta E_I^a(X)$ and $\Delta E_0(X)$.

To resolve this paradox, it is helpful for us to note that the heats of formation of tetrahedrally coordinated semiconductors have been shown[18] to depend very sensitively on E_0. These heats of formation involve averages over the entire valence bonds throughout the Brillouin zone, yet the neighborhood of $\vec{k}=0$ is found to play a particularly important role when E_0 is small. Because the wave packets of holes bound to acceptors are localized near $\Gamma_{25'}$, the admixture of states near $\Gamma_{2'}$ into states near $\Gamma_{25'}$ may be expected to exert a strong influence on $\Delta E_a(X)$. Within the framework of the EMA, in the impurity atomic cell one might replace the periodic part of the Bloch function $u_{nk}^{AB}(r)$ of the AB host crystal by $u_{nk}^{AX}(r)$, the periodic part of the Bloch function in the hypothetical crystal AX. This would obviously lead to matching problems at the radius of the impurity cell and would tend to generate terms of the kind contained in (5.1).

VI. OPERATOR FORMALISM

In Sec. V in order to render the numerical analysis simple, attention was focused on $\Delta E_0(X)$. However, if the position of the $\Gamma_{2'}$ level is important, then so should the position of the Γ_{15} level be important. Its effects are less striking because

$$\Delta E_0' = [E_{AX}(\Gamma_{15}) - E_{AX}(\Gamma_{25'})] - [E_{AB}(\Gamma_{15}) - E_{AB}(\Gamma_{25'})] \qquad (6.1)$$

is usually smaller than ΔE_0 and in Ge, $E_0 \ll E_0'$. In the case of Si, however, E_0' and E_0 are comparable, so that both should be important. Also compare E_0, E_0' for Al in Si (4.2, 3.5) with Ga in Si (2.9, 3.6). This reversal of levels will be used to explain the smallness of $E_I^a(\text{Ga}) - E_I^a(\text{Al})$ noted previously.

A generalization of the effective-mass equations (3.8), which is in the spirit of our analysis of chemical trends, is the following:

$$E_0 F(X) = (\hbar^2/2m)\,|\langle\Gamma_{2'}|p|\Gamma_{25'}\rangle|^2 + \alpha\Delta E_0\,, \qquad (6.2)$$

$$E_0' H_1(X) = (\hbar^2/2m)\,|\langle\Gamma_{15}|p|\Gamma_{25'}\rangle|^2 + \beta\Delta E_0'\,. \qquad (6.3)$$

Here α and β are kinematic parameters to be determined empirically, and (6.2) and (6.3) are called kinematic corrections to the EMA.

In principle, one should now repeat EMA calculations with a modified potential which for $\Delta E_0 = \Delta E_0' = 0$ will give rise to a ground-state energy $E_0^a + \delta E_0^a$ in Si and Ge. (The potential may consist of the ordinary dielectrically screened Coulomb potential plus a square well centered on the origin.) With this potential, one can then calculate $E_I^a(X)$ for a variety of values of the kinematic parameters α and β until a good fit to experiment is obtained. Such calculations lie outside the scope of this paper. One can, however, use the modified potential and Eqs. (6.2) and (6.3) to make crude estimates of the chemical trends of g factors of impurity states in order to see whether these agree qualitatively with experiment. These are known at present only for acceptors in Si, but we will make very rough predictions for Ge as well.

A. Acceptors in Si

We can compute $\partial E_I^a/\partial A$ by first-order perturbation theory using the wave functions and matrix elements of the Hamiltonian given by SOH. For Si, only c_0 and $c_2 < c_0$ are large, so for the purpose of making a rough estimate we write in Si

$$E_I^a \cong c_0^2 H_0 + 2c_0 c_2 H_{02} \qquad (6.4)$$

and note that H_0 depends only on A and that H_{02} depends only on N. The accompanying matrix elements are u_0, the s-wave kinetic energy, and u_{01}, and s-d centrifugal term.

Both the basis function amplitudes c_i and the various matrix elements such as u_0 obviously depend on the well depth of the central cell (which we associate with δE_0^a) and on changes in A, B, and N through F and H_1. In Si, in particular, $\delta E_0^a \cong E_0^a$. This will increase the s-wave amplitude c_0 at the expense of d-wave amplitudes such as c_2, so that from the EMA values[11] $c_0 = 0.93$, $c_2 = -0.35$ one would guess that better values for Si (neglecting changes in the effective values of A, B, and N) would be $c_0 = 0.98$, $c_2 \cong -0.2$. One would also expect the Bohr radius r_1 of the s-wave basis function to be reduced by a factor of about $\sqrt{2}$ from the EMA value of 15 Å to about 11 Å, or close to the Bohr radius r_2 of the d-wave function, which makes the parameter t defined by SOH have the value 1. To

TABLE VI. Parameters for the Taylor-series expansion (5.1).

Host	E_0^a (meV)	δE_0^a (meV)	$10^4 a_1$	$10^2 a_2$
Si	36	34	-150	0.5
Ge	10.0	1.2	-3	1.4

TABLE VII. g factors for acceptors in Si. Experimental values from Ref. 22, $2g_{\text{av}}=g_{\parallel}+\frac{1}{2}g_{\perp}$.

X	$g_{\parallel}$	$\frac{1}{2}g_{\parallel}$	g_{av}
B	1.21	1.22	1.21
Al	1.18	1.08	1.13
Ga	1.14	1.02	1.08
In	0.98	0.79	0.88
EMA (SOH)	0.97	1.02	1.09
Present theory	1.02	1.02	1.02

first order in c_2, Eq. (22) of SOH then reduces to

$$g_{\parallel}=\tfrac{1}{2}g_{\perp}=g=-1.96\kappa-0.1(2B+\tfrac{1}{3}N) \tag{6.5}$$

$$=0.56+0.46=1.02 \quad . \tag{6.6}$$

The first term on the right-hand side of (6.5) is the s-wave contribution, the second term the estimated d-wave one. According to Eq. (5.1), the kinematic corrections should be minimal between Al and Ga. The experimental values[22] for $g_{\parallel}$ and $\frac{1}{2}g_{\perp}$ (both of which are equal to g when Λ is small) are listed in Table VII. The average value of $g_{\parallel}$ and $\frac{1}{2}g_{\perp}$ between Al and Ga is 1.10, in good agreement with (6.6). Also shown are the calculated EMA values of SOH, which are too large and too anisotropic ($g_{\perp}\neq 2g_{\parallel}$) because the d-wave amplitudes $c_i\,(i\geq 2)$ are too large.

Returning to Eq. (6.4), we estimate H_0 and H_{02} in the presence of the modified potential without kinematic corrections as follows. According to SOH, in the EMA omission of the d-wave terms in Si [principally the second term on the right-hand side of (6.4)] reduces the binding energy by a factor of $\frac{2}{3}$. This gives approximately

$$H_{02}=-0.3E_0^a \quad , \tag{6.7}$$

and since in the modified potential $E_I^a=2E_0^a$, we obtain

$$H_0=1.5H_0^a \tag{6.8}$$

which are consistent with $c_{02}\approx -0.2$.

To estimate α and β in Si, note that E_0' changes by $+3\%$ on going from Al to Ga, while E_0 changes by about -30%. The two changes are in opposite directions and must nearly cancel in their effect on $E_I^a(X)$ in order to explain the small shift in $E_I^a(X)$ observed experimentally (Table V). Assuming for qualitative purposes that a virial relation holds for the kinetic and potential contributions to H_0, we can rewrite (6.7) and (6.8) as

$$H_{02}=-0.3E_0^a(1+\delta N/N) \quad , \tag{6.9}$$

$$H_0=1.5E_0^a(1+\delta A/A) \quad , \tag{6.10}$$

in the presence of kinematic corrections δA and δN. Substituting $c_0=0.98$, $c_2=-0.2$, (6.9), and (6.10) into (6.4) gives

$$\delta E_I^a/E_0^a\cong 1.5(1+\delta A/A)+0.6(1+\delta N/N) \quad . \tag{6.11}$$

From (3.5) and (3.7), this reduces to

$$E_I^a(X)/E_0^a\cong 2+(0.5/A+0.6/N)\delta F + (1/A+0.6/N)\delta H_1 \quad . \tag{6.12}$$

Substituting $\delta F=\alpha\,\delta E_0/E_0$ and $\delta H_1=\beta\delta E_0'/E_0'$ from (6.2) and (6.3), and setting $\delta E_I^a=0$ on going from Al to Ga, we have

$$\alpha(\text{Si})=0.17\beta(\text{Si}) \quad . \tag{6.13}$$

According to Table II, from B in In, E_0' changes from 4.6 to 3.2 eV, giving $\Delta E_0'/E_0'\approx 0.3$. Similarly E_0 changes from 44 to 155 meV, so that $\Delta E_I^a(X)/E_I^a(\text{Ga})\approx 1.5$. Thus, (6.12) gives approximately

$$-1.5=(-0.3)\alpha(-1.7)+(-0.17)(\alpha/0.17)(-0.3),$$

$$\alpha=-3, \quad \beta=-18 \quad . \tag{6.14}$$

Although these numerical estimates are quite rough, they are sufficient to enable us to estimate the chemical trends in the g factors $g(X)$ listed in Table VII. Using (3.9) and (6.6) one finds approximately

$$\delta g=\tfrac{1}{3}\delta F-\tfrac{1}{3}\delta H_1-0.2\delta B-\tfrac{1}{30}\delta N \quad . \tag{6.15}$$

Now, the first three terms on the right-hand side of (6.15) approximately cancel because of (6.13). Within the accuracy of our calculations, this almost gives $\delta g\approx 0$. We therefore estimate $\delta g(\text{B})-\delta g(\text{In})$ from (3.6), (3.7), and (6.14) hoping to get the sign right. One finds

$$\begin{aligned}\delta g(\text{B})-\delta g(\text{In}) &= [-3(1.7)/3-(-18)0.3/3] \\ &\quad -\tfrac{1}{30}[-3(1.7)+(-18)0.3] \\ &= 0.08+0.34=+0.42 \quad ,\end{aligned} \tag{6.16}$$

which is to be compared with the experimental value (average of $\delta g_{\parallel}$ and $\frac{1}{2}\delta g_{\perp}$) of $+0.33$. Considering the substantial cancellations and crudity of the approximations made, the agreement with experiment is more than satisfactory. Because δF and δH_1 enter κ and B only in the combination $\delta F-\delta H_1$ (which approximately vanishes), it is the term in δN which dominates $\delta g(X)$.

B. Acceptors in Ge

The case of Ge is simpler than Si in one respect because the corrections to the EMA are 25 times smaller. In another respect, however, it is more complicated. The d-wave components of the EMA solution found to SOH are large for c_2, c_3, c_4, and c_5. We do not expect α/β to have the same ratio in Ge as in Si. A good guess is that

$$\frac{\alpha(\text{Ge})}{\beta(\text{Ge})}=\frac{\alpha(\text{Si})}{\beta(\text{Si})}\frac{E_0(\text{Si})}{E_0(\text{Ge})}\frac{E_0'(\text{Ge})}{E_0'(\text{Si})} \quad , \tag{6.17}$$

$$\frac{\alpha(\text{Ge})}{\beta(\text{Ge})}=3.3\frac{\alpha(\text{Si})}{\beta(\text{Si})} \quad . \tag{6.18}$$

In Si, we found that the effects associated with α (changes in $\Gamma_{2'}$) were comparable to those associated with β (changes in Γ_{15}). This made $E_I^a(\mathrm{Ga}) - E_I^a(\mathrm{Al})$ ten times smaller than one would have expected from α effects alone. According to (6.18), the α effects should be about three times larger in Ge than the β ones, so that the observed shift in $E_I^a(X)$ should be about two-thirds that predicted from α alone. This is indeed the case experimentally, as can be seen from Table IV.

In order to estimate $g(X)$, we utilize Eq. (19) of SOH and retain only terms containing c_0. This gives (again $t \approx 1$)

$$g = -2\kappa c_0^2 + B(5c_1 + c_3)c_0/\sqrt{2} \approx -1.5\kappa - 0.5B \qquad (6.19)$$

using the values of c_i given by SOH. From (3.9) and (3.6), one finds

$$\begin{aligned} \delta g &\approx \tfrac{5}{12}(\delta F - \delta H_1) \\ &\approx 5\alpha\,\delta E_0/18E_0 \ . \end{aligned} \qquad (6.20)$$

From (6.13) and (6.18), one has

$$\alpha(\mathrm{Ge}) \approx 0.5\beta(\mathrm{Ge}) \qquad (6.21)$$

and on going from B to Tl, $E_I^a(X)/E_I^a(\mathrm{Ga})$ changes by about 0.25. Meanwhile E_0 changes by -3.6 eV and E_0' by 1.0 eV. Because $c_1, \ldots, c_4$ are all appreciable, $E_I^a(X)$ depends on A, B, and N. For the purpose of making a very rough estimate [justified by the absence of experimental data for $g(X)$ in Ge], assume the dependence given in (6.12). Then α can be estimated from

$$\begin{aligned} -0.25 = {} & (-0.04 - 0.02)\alpha(-3.6) \\ & + (-0.07 - 0.02)2\alpha(-1.0) \ , \end{aligned} \qquad (6.22)$$

$$\alpha(\mathrm{Ge}) = -0.6, \quad \beta(\mathrm{Ge}) = -1.2 \ . \qquad (6.23)$$

Substituting (6.23) into (6.21) and using $\delta E_0/E_0 \cong -3$ on going from B to Tl, one guesses $\delta g \simeq 0.5$ on going from B to Tl. According to the EMA values for $g_\parallel$ and $\frac{1}{2}g_\perp$ given by SOH, g_{av} should vary from about -1.5 for B in Ge to -1.0 for Tl in Ge. These estimates are very rough.

VII. DISCUSSION

As stressed in the Introduction, the aim of this paper has been to indicate how much one might hope to gain by combining the results of modern energy-band theory with classical wave-packet analysis as it appears in the EMA. To this end, we have carried out in Secs. V and VI several highly simplified calculations to show what can be expected from the combined theory. A number of encouraging results have been obtained. These are (a) chemical trends of $E_I^a(X)$ and $g(X)$ in Ge are probably dominated by shifts in $E_0(\mathrm{Ge}X)$; (b) chemical trends of $E_I^a(X)$ in Si are influenced about equally by shifts in $E_0(\mathrm{Si}X)$ and $E_0'(\mathrm{Si}X)$, or by interactions of $\Gamma_{25'}$ with $\Gamma_{2'}$ and Γ_{15}, respectively; (c) other things [E_0, E_0' in the host crystal and $\Delta E_0(X)$, $\Delta E_0'(X)$ in the hypothetical crystal] being equal, interaction with Γ_{15} is about five times more important than interaction with $\Gamma_{2'}$. In Si, one has approximate equality between E_0 and E_0', but the total change $\Delta E_0'(X)$ on going from $X = \mathrm{B}$ to $X = \mathrm{In}$ is about five times smaller than $\Delta E_0(X)$, so that in Si, chemical trends are actually affected about equally by Γ_{15} and $\Gamma_{2'}$, as mentioned in (b). (d) From the over-all trends in $E_I^a(X)$ in Si, one can predict the trends in $g(X)$. In spite of considerable cancellation between the $\Gamma_{2'}$ and Γ_{15} terms, the results are in good agreement with experiment.

These qualitative successes in separating the effects of $\Gamma_{2'}$ from those of Γ_{15} prompt one to inquire further concerning the rather arbitrary separation of kinematic and potential effects we have made. It is quite clear that the mass renormalization we have used is highly artificial and is valid only in the region of the impurity itself. However, so long as we calibrate α and β by analyzing linear shifts in $E_I^a(X)$, no great error is made in this way.

If one wishes to undertake more elaborate calculations of the type carried out by SOH in the EMA, one would notice that the effects of changing E_0 and E_0', which admix $\Gamma_{2'}$ and Γ_{15} into $\Gamma_{25'}$ are analogous to cubic phase shifts $\delta_{2'}$ and δ_{15} in the envelope functions $F_i(R)$ which are the generalization to the crystalline situation of spherical phase shifts δ_l associated with scattering from a spherically symmetric potential. In the crystalline case, one can imagine drawing a unit cell of volume $\frac{1}{4}a^3$ centered on the impurity whose edges pass through the centers of the four nearest neighbors. One-quarter of each of the four nearest-neighbor's pseudopotentials lies inside this unit cell, the remainder outside as part of the host crystal. Expand the former in tetrahedral harmonics. The first-order effect mixes Γ_{15} into $\Gamma_{25'}$, while one needs combined strain and spin-orbit effects to mix $\Gamma_{2'}$ into $\Gamma_{25'}$, since in the tetrahedral group $\Gamma_{25'}$, Γ_{15}, and $\Gamma_{2'}$ transform as xy, xy, and 1, respectively. This explains (c) above, i.e., why interaction with Γ_{15} is much more important, other things being equal, than interaction with $\Gamma_{2'}$. It also explains the signs of α and β, which depend on whether one is attracting electrons or holes.

Further discussion of this formalism would entail calculations which lie outside the scope of this paper. However, we believe that $E_0(X)$ and $E_0'(X)$ should serve as useful calibrations for chemical shifts in $\delta_{2'}$ and δ_{15}.

VIII. OTHER APPLICATIONS

Included in Tables II and III are values of E_0, E_0', E_1, and E_2 for donors in Si and Ge. In an earlier paper,[1] we compared trends in $E_I^d(X)$ in Si and Ge with trends in $E_g(X)$, which is almost the same as $E_2(X)$, the energy gap between conduction-band states near X_1 and valence-band states near X_4. According to the point of view developed here, this is a good approach for Si, where the transverse effective mass m_t is determined[23] primarily by E_2, but for Ge, where m_t is associated with the conduction-band edges at L, a better correlation should be found with E_1. Effects should also be found associated with longitudinal effective masses m_l. Unfortunately, there are two contributions to the latter in each crystal, associated with two remote levels, and these two contributions tend to cancel, leaving m/m_l close to 1 in both cases. Thus, analysis of chemical trends of donor binding energies $E_I^d(X)$ is intrinsically more complicated than is analysis of trends in $E_I^a(X)$. Moreover, because of certain selection rules associated with the neighborhood of X, the g shifts of donors in Si are very small and require elaborate analysis,[24] while g shifts of donors in Ge have not been reported.

On the basis of the information available for donors alone, it appears that no conclusive analysis can be made. However, if one accepts the qualitative conclusions of our analysis of $E_I^a(X)$ and $g(X)$ for acceptors, then we should look for chemical trends in E_2 (to describe m_t effects in Si) and E_1 (to describe m_t effects in Ge). As for m_l, the nearest levels of symmetry Δ_1(Si) or Λ_1(Ge) which contribute to this factor connect with Γ_1 (s symmetry). Only Γ_1 in the conduction band should shift appreciably with X, and its trend *may* parallel that of $\Gamma_{2'}$, although little is known about this point at present. Should this be the case, it is easy to see that as one proceeds from lighter to heavier impurities, the trends in m_t and m_l will have opposite sign, because the former is influenced most strongly by a nearby valence band of lower energy and the latter most strongly by a nearby conduction band of higher energy. This would also account for the observed minimum in $E_I^d(X) - E_0^d$ at $X = \mathrm{Sb}$.

Chemical trends in acceptor binding energies in GaP similar to those of acceptor binding energies in Si and Ge have been noted.[25] Divalent acceptors (Mg, Zn, Cd) may be introduced on the Ga sites, and tetravalent ones (C, Si, Ge) on the P sites. A striking feature of the data is that $E_I^a(\mathrm{Ge}) - E_I^a(\mathrm{Si})$ is about 100 meV, while $E_I(\mathrm{Zn}) - E_I(\mathrm{Mg})$ is only 10 meV. From an EMA viewpoint, the holes spend about 70% of the time on the P sublattice and about 30% on the Ga sublattice. Thus, a ratio of chemical shifts of 2 : 1 might be expected, compared to the ratio of 10 : 1 which is actually observed. This is the kind of problem for which a complete phase-shift analysis (of the kind sketched in Sec. VII) would be valuable.

---

[1] J. C. Phillips, Phys. Rev. B 1, 1540 (1970).

[2] J. C. Phillips, Phys. Rev. B 1, 1545 (1970).

[3] J. C. Phillips, Phys. Rev. Letters 20, 550 (1968); J. A. Van Vechten, *ibid.* 182, 981 (1969).

[4] J. C. Phillips, Rev. Mod. Phys. 42, 317 (1970).

[5] W. Kohn and J. M. Luttinger, Phys. Rev. 98, 915 (1955); W. Kohn, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1955), Vol. V, p. 257.

[6] D. G. Thomas and J. J. Hopfield, Phys. Rev. 150, 680 (1966); D. G. Thomas, Phys. Today 21, 42 (1968).

[7] J. C. Phillips, Phys. Rev. Letters 22, 285 (1968).

[8] J. C. Phillips, Solid State Phys. 18, 55 (1966).

[9] D. Schechter, J. Phys. Chem. Solids 23, 237 (1961); see also Y. Abe, J. Phys. Soc. Japan 19, 825 (1964).

[10] K. S. Mendelson and H. M. James, J. Phys. Chem. Solids 25, 729 (1964).

[11] K. Suzuki, M. Okazaki, and H. Hasegawa, J. Phys. Soc. Japan 19, 930 (1964); see also P. J. Lin-Chung and R. F. Wallis, J. Phys. Chem. Solids 30, 1453 (1969).

[12] J. C. Hensel and G. Feher, Phys. Rev. 129, 1041 (1963).

[13] J. C. Phillips, Phys. Rev. 125, 1931 (1962).

[14] J. C. Phillips and J. A. Van Vechten, Phys. Rev. Letters 22, 705 (1969); J. A. Van Vechten, Phys. Rev. 187, 1007 (1969).

[15] J. C. Slater, Phys. Rev. 76, 1592 (1949); 87, 807 (1952); G. F. Koster, *ibid.* 89, 67 (1953).

[16] J. C. Phillips, *Covalent Bonding in Crystals, Molecules and Polymers* (University of Chicago Press, Chicago, 1970).

[17] J. A. Van Vechten and J. C. Phillips, Phys. Rev. (to be published).

[18] J. C. Phillips and J. A. Van Vechten, Phys. Rev. (to be published).

[19] R. L. Jones and P. Fisher, J. Phys. Chem. Solids 26, 1125 (1965).

[20] G. Picus, E. Burstein, and B. Henvis, J. Phys. Chem. Solids 1, 75 (1956).

[21] R. A. Faulkner, Phys. Rev. 184, 713 (1969).

[22] G. Feher, J. C. Hensel, and E. A. Gere, Phys. Rev. Letters 5, 309 (1960).

[23] L. Kleinman and J. C. Phillips, Phys. Rev. 118, 1153 (1960).

[24] L. Liu, Phys. Rev. 126, 1317 (1962).

[25] P. J. Dean, E. G. Schönherr, and R. B. Zetterstrom, J. Appl. Phys. 41, 3474 (1970).