Effect of Spin-Orbit Coupling and Other Relativistic Corrections on Donor States in Ge and Si

Joachim Appel

John Jay Hopkins Laboratory for Pure and Applied Science, General Atomic Division of General Dynamics Corporation, San Diego, California

Four corrections to the effective-mass approximation are considered with first-order perturbation theory; namely, the deviation of the total perturbing potential U from the potential $U_0 = -e^2/\kappa r$ (κ = static dielectric constant) and the three relativistic corrections: (a) spin-orbit coupling, (b) s-shift correction, and (c) mass-velocity correction. The number of independent matrix elements is determined for each perturbation by means of the selection-rule theorem. There is no effect of spin-orbit coupling on the effective-mass ground state; the corresponding effect on the exact eigenstates of the nonrelativistic Hamiltonian appears to be small for donors in Si, and does not cause a splitting of the sextet state (with spin) in Ge or Si. Both of the other two relativistic corrections give rise to a shift and a splitting of the degenerate effective-mass ground state, as does the perturbation $U - U_0$. The magnitude of the relativistic corrections and their relative contribution to the observed splitting of the effective-mass ground state are discussed briefly.

I. INTRODUCTION

IN this paper, some of the corrections to the effective-mass theory[1] of shallow donor states in Ge and Si will be studied. The simple effective-mass equation for the motion of a donor electron in the perturbed periodic-potential field of a germanium or silicon crystal accounts for a qualitatively correct picture of the stationary eigenstates of donor electrons. The energy spectra of different donor impurities (P, As, Sb) in silicon have been determined experimentally from infrared-absorption measurements, and quantitative agreement has been found between theory and experiment, with the exception of the ground state.[2] There are, however, some discrepancies between theory and experiment with respect to the ground-state energy and the amplitude of its effective-mass wave function at a donor site. Kohn and Luttinger[3] have shown that the effective-mass wave function can be corrected for by a simple semiempirical procedure, using the observed energy of the lowest donor state. It has been recognized by the same authors that the discrepancies are caused by the breakdown of the effective-mass formalism in the immediate neighborhood of the impurity ion where the perturbing potential is large and varies strongly with distance from the impurity nucleus. Therefore, several corrections to the effective-mass theory are necessary. These corrections are discussed in Sec. II where the problem is also formulated so it can be treated by first-order perturbation theory. In Sec. III, a group-theoretical analysis is presented to determine the number of independent constants for the matrixes of the perturbation components with the help of the selection-rule theorem. Section IV is concerned with a brief qualitative discussion of the relative magnitudes of the different perturbations considered in this paper.

[1] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955); W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5.

[2] G. S. Picus, E. Burnstein, and B. Henvis, J. Phys. Chem. Solids **1**, 75 (1956); see also W. Kohn, Phys. Rev. **98**, 1856 (1955).

[3] W. Kohn and J. M. Luttinger, Phys. Rev. **97**, 883 (1955).

II. CORRECTIONS—FORMULATION OF THE PROBLEM

If one neglects the effects of lattice vibrations, there are three corrections which affect the effective-mass ground state:

(1) the deviation of the total perturbing potential $U(\mathbf{r})$ from $U_0(r) = -e^2/\kappa r$ due to the self-consistent potential of the impurity ion in the central cell region—being different from that of a Ge or Si atom—the effect of local strain introduced by the impurity, and polarization effects (see Ref. 7 below);

(2) the admixture of Bloch functions from higher bands into the impurity wave function caused by the strong potential U near the impurity ion; and

(3) relativistic corrections: (a) spin-orbit coupling; (b) s-shift correction, and (c) mass-velocity correction.

It is a characteristic property of these corrections that the major contribution of each of them comes from the central cell; that is, the atomic cell containing the impurity ion. Thus, in calculating their effect on the effective-mass ground state, there are two technical difficulties. First, one must know the correct form of the perturbing potential $U(\mathbf{r})$ within the central cell. This potential is known for small $r \ll a_0 (= 5.29 \times 10^{-9}$ cm) and for large $r \gg r_s$ ($\frac{4}{3}\pi r_s^3 = \frac{1}{2}\Omega$, where Ω is the unit cell volume).[3] The symmetry of $U(\mathbf{r})$ is that of a lattice site and is given by the tetrahedral point group. Second, the Bloch functions admixed into the impurity wave function must be known. However, at present one knows only the approximate Bloch functions of the lowest conduction band. In the case of Ge and Si, these functions are suitable admixtures of $4s$ and $4p$ atomic orbitals, and of $3s$ and $3p$ atomic orbitals, respectively. Thus, we shall neglect the admixtures of Bloch functions from higher bands into an impurity wave function and shall consider only the effects of corrections (1) and (3).

A primary motivation for this consideration is a group-theoretical analysis of the four perturbations associated

with corrections (1) and (3). It is of particular interest to know how the eightfold and twelvefold degenerate effective-mass ground states, including spin-degeneracy, of donor electrons in Ge and Si, respectively, split under the influence of these corrections. This question is of interest because a number of experimental investigations have been carried out in an original manner to determine the structure of the energy levels emerging from the degenerate effective-mass ground state (see Fig. 1).[4] The experimental information on the energy-level scheme of those impurity states which emerge from the degenerate effective-mass ground state gives rise to two questions: (1) to what extent do the three relativistic corrections account for the observed energy levels; and (2) how large is the additional splitting to be expected from spin-orbit coupling? This problem is formulated with the two-component Pauli equation[5]

$$\left[\frac{\mathbf{p}^2}{2m}+\mathcal{V}+\frac{\hbar}{4m^2c^2}(\nabla\mathcal{V}\times\mathbf{p})\cdot\boldsymbol{\sigma}+\frac{\hbar}{4im^2c^2}(\nabla\mathcal{V}\cdot\mathbf{p})\right.$$

$$\left.-\frac{1}{2mc^2}(E_0+\mathcal{V})^2\right]\Psi=E\Psi,\quad(1)$$

where E_0 is the lowest energy eigenvalue of the non-relativistic Hamiltonian H_0 defined below. The second term is the total potential energy of an impurity electron,

$$\mathcal{V}(\mathbf{r})=V(\mathbf{r})+U(\mathbf{r}),\quad(2)$$

where $V(\mathbf{r})$ is an effective periodic potential. The three last terms of Eq. (1) are peculiar to the relativistic theory and correspond to the corrections (3a), (3b), and (3c). We mention that the s-shift correction shifts the s terms of a free atom (pure Coulomb potential), but contributes nothing to p, d, and higher terms.[6] The stationary states of Eq. (1), its eigenvalues, and the associated two-component wave functions, cannot be

[4] The results summarized in Fig. 1 have been obtained by experimental investigations of a different nature: (1) The change of the electron spin resonance spectrum of donor electrons in Si has been investigated by D. K. Wilson and F. Feher, Phys. Rev. 124, 1068 (1961). (2) The low-temperature electrical resistivity of antimony-doped Ge under uniaxial strain and compression has been measured by F. Fritzsche, Phys. Rev. 125, 1560 (1961); Twose shows in an appendix to Fritzsche's paper, that, even in the effective-mass approximation, there is a small splitting of the degenerate ground state; this splitting is neglected here. (3) The strain-induced shifts of the absorption lines of the Lyman series for donor electrons in Ge have been observed by G. Weinreich, W. S. Boyle, H. A. White, and K. F. Rodgers, Phys. Rev. Letters 3, 244 (1959).

[5] This equation is derived from the four-component Dirac equation under the assumption that at all points in the configuration space, two components are small compared to the other two. This assumption is justified (in zero magnetic field) if the first two terms of Eq. (1) are small compared with the rest-energy of an electron.

[6] E. U. Condon and A. H. Shorley, *Theory of Atomic Spectra* (Cambridge University Press, Cambridge, England, 1958), p. 125. In the present case the potential $V(\mathbf{r})$—and in particular $U(\mathbf{r})$—is not a pure Coulomb potential and, therefore, the so-called s-shift correction does contribute very slightly to Bloch states and to impurity states which are not pure s states.
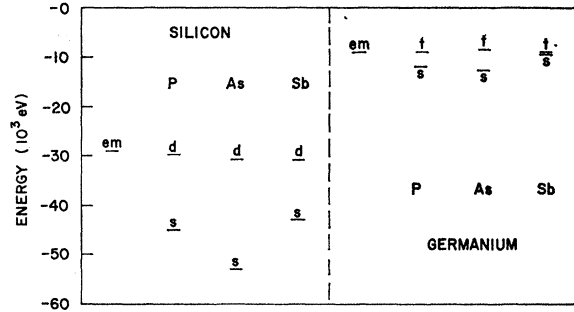
FIG. 1. Energy-level scheme, neglecting the spin-degeneracy, for singlet (s) and doublet (d) states of donors in silicon and for singlet and triplet (t) states for donors in germanium after Wilson and Feher, and Fritzsche. (See Ref. 4.) The splitting between doublet and triplet states for donors in Si is assumed to be zero (em = energy of effective-mass ground state).

calculated without an approximation. In the next section the problem is treated by perturbation theory.

III. PERTURBATION THEORY

The nonrelativistic problem with the Hamiltonian

$$H_0=(\mathbf{p}^2/2m)+V(\mathbf{r})+U_0(\mathbf{r})\quad(3)$$

has been treated by Kohn and Luttinger.[1] The impurity potential $U_0=-e^2/\kappa r$, where κ is the static dielectric constant.[7] An eigenfunction of H_0 can be developed, in terms of Bloch functions, for electrons moving in the periodic potential $V(\mathbf{r})$:

$$\psi=\sum_{n,\mathbf{k}}A_n(\mathbf{k})\psi_n(\mathbf{k},\mathbf{r}),\quad(4)$$

where n labels the conduction bands. In the effective-mass approximation, a donor wave function is written as an appropriate linear combination of wave functions

$$\psi_i=\sum_{\mathbf{k}}A_0(\mathbf{k}-\mathbf{k}_i)\psi_0(\mathbf{k}-\mathbf{k}_i,\mathbf{r}),\quad(5)$$

where $i=1, 2, \cdots, m$ correspond to the minima of the lowest conduction band, $n=0$. Assuming that A_0 has appreciable magnitude only in the vicinity of $\mathbf{k}_i$, one can rewrite Eq. (5) to a lowest approximation, in the form

$$\psi_i=F_i(\mathbf{r})\psi_0(\mathbf{k}_i,\mathbf{r}),\quad(6)$$

where

$$F_i(\mathbf{r})=\sum_{\mathbf{k}}A_0(\mathbf{k}-\mathbf{k}_i)\exp(i\mathbf{k}_i\cdot\mathbf{r})$$

is the envelope function, i.e, an eigenfunction of the effective mass equation. The wave function ψ_i is normalized as usual (see Ref. 3).

[7] K. Müller [dissertation, Braunschweig, 1960 (unpublished)] has investigated in some detail the polarization field near a point charge located at a lattice site of a cubic semiconductor. He employs a technique similar to Ewald's method of long waves to obtain the corrections to the continuum polarization. A deformation of the potential U_0 occurs for distances $r<3a$ (a=lattice constant) because of a dispersion of the polarizability for short wavelengths and because of a local dielectric anisotropy near the point change.

TABLE I. Character table for the tetrahedal double group.[a]

Character \ Class	1E	1Ē	6C₄²	8C₃	8C̄₃	6S₄	6S̄₄	12σd
$^1\Gamma_1$	1	1	1	1	1	1	1	1
$^1\Gamma_2$	1	1	1	1	1	−1	−1	−1
$^2\Gamma_3$	2	2	2	−1	−1	0	0	0
$^3\Gamma_4$	3	3	−1	0	0	−1	−1	1
$^3\Gamma_5$	3	3	−1	0	0	1	1	1
$^2\Gamma_6$	2	−2	0	1	−1	2	−2	0
$^2\Gamma_7$	2	−2	0	1	−1	−2	−2	0
$^4\Gamma_8$	4	−4	0	−1	1	0	0	0
Ge, "1s" state with spin	8	−8	0	1	−1	0	0	0
Si, "1s" state with spin	12	−12	0	0	0	0	0	0

[a] For the symmetry operations of the tetrahedral point group and the associated double group, see G. F. Koster, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5. The direct products of the double group representations are given by D. Schechter Technical Report No. 4 Carnegie Institute of Technology, March, 1958 (unpublished). The decompositions of symmetric direct products can be determined by a method found in F. D. Murnaghan, *The Theory of Group Representations* (The Johns Hopkins Press, Baltimore, 1938), p. 72.

As the perturbation to the Hamiltonian $H_0 \cdot \mathbf{1}$ (1 is the unit matrix), we take

$$H' = (U-U_0)\cdot\mathbf{1} + \frac{\hbar}{4m^2c^2}(\nabla\mathcal{V}\times\mathbf{p})\cdot\boldsymbol{\sigma}$$

$$+\frac{\hbar}{4im^2c^2}(\nabla\mathcal{V}\cdot\mathbf{p})\cdot\mathbf{1} - \frac{1}{2mc^2}(E_0+\mathcal{V})^2\cdot\mathbf{1}. \quad (7)$$

The first term of Eq. (7) is the deviation of the actual impurity potential from the potential U_0 of the effective-mass equation; the other terms represent the *relativistic corrections* which depend on the *total* potential $\mathcal{V}$. Taking the eigenfunction of H_0 given by the effective-mass approximation and multiplied by an arbitrary spin function we readily find the correct zeroth-order wave function corresponding to H' from symmetry considerations.

Germanium. The lowest effective-mass eigenstate is eightfold degenerate when the spin degeneracy is included. This state belongs to a reducible representation of the tetrahedral double group $\bar{T}_d$; a perturbation can split this state into three states associated with the representations $^2\Gamma_6$, $^2\Gamma_7$, and $^4\Gamma_8$ as seen from Table I. The correct zeroth-order wave function, which forms a basis for an irreducible representation $^t\Gamma_s$ of $\bar{T}_d$, is written in the form

$$\Psi_{rs} = \sum_i (\alpha_i{}^{rs}\phi_i + \beta_i{}^{rs}\chi_i), \quad (r=1, 2, \cdots t), \quad (8)$$

where the two-component wave functions ϕ_i and χ_i are related by time reversal symmetry; i.e., Wigner's operation K for spin one-half particles,[8]

$$\phi_i = [(u_+ + u_-)/\sqrt{2}]\psi_i, \quad \chi_i = [(u_- - u_+)/\sqrt{2}]\psi_i{}^*. \quad (9)$$

[8] The time reversal operation $K = i\sigma_y K_0$, where K_0 is the complex conjugation operator.

The wave functions ψ_i are given by Eq. (6), if not explicitly noted otherwise. The two normalized spin functions u_+ and u_-, as well as the linear combinations $u_+ + u_-$ and $u_- - u_+$, form a basis for the representation $^2\Gamma_6$.[9] The formal advantage for choosing the $2m$ spin functions ϕ_i and χ_i as components of the zeroth-order wave functions Ψ_{rs}, instead of spin-up and spin-down functions, is suggested by the time reversal symmetry between ϕ_i and χ_i. Hence, one immediately finds the time-reversed wave function of Ψ_{rs}; it is

$$K\Psi_{rs} = \sum_i [(\alpha_i{}^{rs})^*\chi_i + (\beta_i{}^{rs})^*\phi_i], \quad (10)$$

which, according to Kramer's theorem, also belongs to the basis given by Eq. (8). There is no additional degeneracy of an eigenstate associated with $^2\Gamma_6$, $^2\Gamma_7$, or $^4\Gamma_8$ because of time reversal symmetry; the above representations fall under Wigner's case (c),[10] as does the small representation of the Bloch state for the conduction band minimum. The coefficients $\alpha_i{}^{rs}$ and $\beta_i{}^{rs}$ are calculated in Appendix A; the results are: for ($^2\Gamma_6$),

$$\alpha_i{}^{16} = \tfrac{1}{2}(1,1,1,1), \quad \alpha_i{}^{26} = \beta_i{}^{16},$$
$$\beta_i{}^{16} = (0,0,0,0), \quad \beta_i{}^{26} = \alpha_i{}^{16}; \quad (11a)$$

for ($^2\Gamma_7$),

$$\alpha_i{}^{17} = (1/2\sqrt{3})(1, 1, -1, -1),$$
$$\beta_i{}^{17} = (1/2\sqrt{3})(1-i, i-1, 1-i, i-1),$$
$$\alpha_i{}^{27} = (\beta_i{}^{17})^*,$$
$$\beta_i{}^{27} = \alpha_i{}^{17}; \quad (11b)$$

and for ($^4\Gamma_8$),

$$\alpha_i{}^{18} = (1/2\sqrt{2})(1+i, -1-i, 1-i, i-1),$$
$$\beta_i{}^{18} = (0,0,0,0),$$
$$\alpha_i{}^{28} = [1/2(6)^{1/2}](-2, -2, +2, +2),$$
$$\beta_i{}^{28} = [1/2(6)^{1/2}](1+i, -1-i, 1-i, i-1),$$
$$\alpha_i{}^{38} = (\beta_i{}^{28})^*,$$
$$\beta_i{}^{38} = \alpha_i{}^{28},$$
$$\alpha_i{}^{48} = \beta_i{}^{18},$$
$$\beta_i{}^{48} = (\alpha_i{}^{18})^*. \quad (11c)$$

Silicon. The twelvefold-degenerate effective-mass ground state, including the electron spin, can split under a perturbation into four different states; the appropriate decomposition, as found from Table I, is $^2\Gamma_6 + ^2\Gamma_7 + 2\,^4\Gamma_8$. The correct zeroth-order wave functions given by Eq. (8) can be determined with the help of the table of coefficients given in Ref. 1 [Kohn, Eq. (5.46)] by proceeding as in Appendix A for the case of Ge.

[9] The complex conjugate spin functions u_+^* and u_-^*, as well as the two functions $u_+^* + u_-^*$ and $u_-^* - u_+^*$, transform according to $^2\Gamma_6^*$ under the symmetry operations of $\bar{T}_d$.

[10] E. Wigner, Nachr. Acad. Wiss. Goettingen, Math.-Physik. Kl. 1932, 546.

The Ψ_{rs} are the correct zero-order wave functions for the perturbation H'. Thus, the influence of the perturbation on the effective-mass energy levels of the unperturbed Hamiltonian is obtained by evaluating the expectation value of H' in the wave functions Ψ_{rs}. If one rewrites Eq. (7) in the form

$$H' = \sum_{\kappa} H_{\kappa}, \tag{12}$$

where $H_1 =$ potential correction, $H_2 =$ spin-orbit coupling correction, $H_3 = s$-shift correction, $H_4 =$ mass-velocity correction, the matrix elements of the four perturbations are given by

$$\int_{\text{a.s.}} \Psi_{rs}^* H_{\kappa} \Psi_{r's'} d\tau, \tag{13}$$

where a.s. indicates integrations over all space. In order to find the number of independent constants for each perturbation, assuming that the Ψ_{rs} are given in terms of the effective-mass wave functions ψ_i, Eq. (6), we proceed in the following way. The matrices in the Ψ_{rs} representation can be written as a linear combination of matrices in the ϕ_i, χ_i representation; the latter are given by

$$\int_{\text{a.s.}} \phi_i^* H_{\kappa} \phi_j d\tau \quad \text{and} \quad \int_{\text{a.s.}} \phi_i^* H_{\kappa} \chi_j d\tau. \tag{14}$$

Taking the case of germanium as an example, we know that the wave functions ϕ_i, χ_i transform according to the representation $(^2\Gamma_6 + {}^2\Gamma_7 + {}^4\Gamma_8)$. The four perturbations H_{κ} are scalar quantities and transform according to $^1\Gamma_1$. Thus, in the ϕ_i, χ_i representation, the selection rule theorem[11] gives at most three independent constants for each perturbation H_{κ}. For the spin-independent perturbations ($\kappa = 1, 3, 4$) two of the three constants must be equal to one another; that is, the energy eigenvalues associated with the representations $^2\Gamma_6$ and $^4\Gamma_8$ are not separated by these perturbations since the matrices between ϕ_i and χ_j vanish (the spin functions are orthogonal to one another). The spin-independent perturbation ($\kappa = 2$), however, could lift the degeneracy. To find out whether this is actually the case, the spin functions are eliminated from the matrices (14). Then we are left with the expressions

$$M_{ij,\kappa} = \int_{\text{a.s.}} \psi_i^* H_{\kappa} \psi_j d\tau, \quad (\kappa = 1, 3, 4), \tag{15}$$

and

$$M_{ij,2} = \int_{\text{a.s.}} \psi_i^* h_x \psi_j d\tau, \quad \mathfrak{M}_{ij,2} = \int_{\text{a.s.}} \psi_i^* h_x \psi_j^* d\tau, \tag{16}$$

where h_x is the x component of a pseudo-vector operator

[11] V. Heine, *Group Theory in Quantum Mechanics* (Pergamon Press, Ltd., London, 1960).

TABLE II. Number of independent constants obtained with the effective-mass approximation [ψ_i is given by Eq. (6)] for the matrices given by Eqs. (15) and (16).

Semi-conductor \ Perturbation	Potential correction	Spin-orbit coupling	s-shift correction	Mass-velocity correction
Germanium	2	0 (1)	2	2
Silicon	3	.0 (3)	3	3

h defined by the equation[12]

$$H_2 = h \cdot \sigma. \tag{17}$$

The number of independent constants contained in Eqs. (15) and (16) is determined with the selection rule theorem. The results are found in Table II; they apply also—as can be shown with a lengthy calculation —if the effective-mass wave functions are given by $\psi_i = \sum_k A_0(k) \exp(i k \cdot r)[u_{k_i} + (k - k_i) \cdot \nabla_{k - k_i} u_k]$, where $|k - k_i| \ll |k_i|$; the derivatives of u_k are taken at $k = k_i$. If one goes beyond this effective-mass approximation by assuming that the impurity wave functions contain admixtures of Bloch functions from higher conduction bands, the spin-orbit coupling matrices (16) give one and three independent constants for Ge and Si, respectively, as indicated in Table II.[13] It is emphasized that, in any case, the spin-orbit matrices $\mathfrak{M}_{ij,2}$ resulting from those matrices (14) which connect time reversal wave functions, such as ϕ_i and χ_j, vanish.[14]

In order to discuss the influence of the perturbations H_{κ} on the unperturbed energy levels, let us write down the solutions of that secular problem obtained by taking the scalar product of the two-component wave equation

$$(H_0 \cdot 1 + H') \Psi_{rs} = E_s \Psi_{rs} \tag{18}$$

with χ_i and ϕ_i. The eigenvalues of the secular matrix are found with the help of the eigenvectors α_i^{rs}, β_i^{rs} ($i = 1, 2, \cdots m$). For germanium we find

$$E_6 = E_0 + \sum_{\kappa} (M_{11,\kappa} + M_{21,\kappa} + M_{31,\kappa} + M_{41,\kappa}),$$

$$E_7 = E_0 + \sum_{\kappa} (M_{11,\kappa} + M_{21,\kappa} - M_{31,\kappa} - M_{41,\kappa}), \tag{19}$$

$$E_8 = E_7.$$

[12] Thus $h_x \psi_i = -(\hbar^2/4m^2c^2)(\nabla \upsilon \times i \nabla \psi_i)_x$, etc.

[13] For Ge the basis ψ_i transforms according to the representation $\Gamma_1 + \Gamma_4$ of the single group T_d. The symmetric direct product $[\Gamma + \Gamma_4]_{\text{sym}}^2$ contains the trivial representation Γ_1, according to which H_1, H_3, and H_4 transform, twice. The three components of the pseudovector operator $h(h_x, h_y, h_z)$ transform according to Γ_5. Since h is a vector operator, but not a vector, we must take the direct product $\Gamma_5 \times (\Gamma_1 + \Gamma_4)$ and see how often it contains Γ_1 and Γ_4 in order to determine the number of independent constants for the matrices $M_{ij,2}$ and $M_{ij,2}$. (If, however, h were a vector, and not a vector operator, we should take the symmetric direct product $[\Gamma_1 + \Gamma_4]_{\text{sym}}^2$ and see how often it contains Γ_5.)

[14] Because h is imaginary and Hermitian, one has $\mathfrak{M}_{ij,2} = -\mathfrak{M}_{ji,2}$. On the other hand, the symmetry requires $\mathfrak{M}_{ij,2} = \mathfrak{M}_{ji,2}$; that is, the solution of the secular problem given below shows that the matrices $\mathfrak{M}_{ij,2}$ must be real.

IV. DISCUSSION OF THE CORRECTIONS

A. Potential Correction

The total perturbing potential U can be written in terms of those harmonic functions which are invariant under the symmetry operations of the tetrahedral point group. The first few terms of an expansion of U are given by

$$U(\mathbf{r})=f(r)+g(r)\frac{xyz}{r^3}+h(r)\frac{x^4+y^4+z^4-3r^4/5}{r^4}+\cdots, \quad (20)$$

where[3]

$$\lim_{r\to 0}f(r)=-\frac{e^2(Z_d-Z_0)}{r}, \quad \lim_{r\to\infty}f(r)=U_0(r),$$

Z_d and Z_0 being the atomic numbers of a donor atom and of a Ge or Si atom, respectively. The harmonic functions of the tetrahedral group are linear combinations of surface spherical harmonics Y_l^m, with the same subindex l. Since spherical harmonics are solutions of the Laplace equation, one finds

$$\lim_{r\to 0}g(r)=Ar^3, \quad \lim_{r\to\infty}g(r)=Br^{-3}, \quad (21)$$

where A and B are constants.

The potential correction $U-U_0$ causes a shift and a splitting of the degenerate ground state. The shift increases the binding energy and is proportional to the diagonal elements

$$M_{ii,1}=\int_{a.s.}|\psi_i|^2(U-U_0)d\tau$$

$$\simeq\frac{1}{2}|F(0)|^2\Omega\int_{c.c.}|\psi_0(\mathbf{k}_i,\mathbf{r})|^2f(r)\frac{2d\tau}{\Omega}, \quad (22)$$

where $F(0)=F(\mathbf{r}=0)$ (the index i is omitted) and c.c.=central cell. The integral on the right side of Eq. (22) is the average value of the potential energy of a conduction electron; it is of the order of 10 eV. The factor before the integral is given by

$$\frac{1}{2}|F(0)|^2\Omega=\frac{4}{3}(r_s/a_0^*)^3=0.78\times10^{-4} \text{ for } Ge(a_0^*=45 \text{ Å}),$$
$$=0.77\times10^{-3} \text{ for } Si(a_0^*=20 \text{ Å}).$$

Thus, the shift of the effective-mass ground state amounts to 10^{-3} eV and 10^{-2} eV for donors in Ge and Si, respectively. The shifts are larger by one order of magnitude if one takes the amplitudes of the corrected envelope functions (see Ref. 1, Kohn). The splitting of the degenerate effective-mass ground state depends on the off-diagonal elements $M_{ij,1}$ which contain the anisotropic contribution $U(\mathbf{r})-f(r)$ to the perturbing potential. We are not able to estimate the off-diagonal elements, since the anisotropic contributions to U given by Eq. (20), which vanish with r^n ($n\geq 3$), are not known.

B. Spin-Orbit Coupling

The spin-orbit matrix elements (14) between wave functions ϕ_i and χ_j vanish. Therefore, the degeneracy of the sextet consisting of the two states $^2\Gamma_7$ and $^4\Gamma_8$ remains unlifted [see Eq. (19) for the case of Ge]. This result is not restricted to the effective-mass approximation and holds also if the wave functions ϕ_i and χ_j are given in terms of the exact eigenfunctions of H_0 which can be written in the form $\psi_i=\sum_{n,\mathbf{k}}A_n(\mathbf{k}-\mathbf{k}_i)\times\psi_n(\mathbf{k}-\mathbf{k}_i,\mathbf{r})$. However, spin-orbit matrix elements (14) between wave functions ϕ_i and ϕ_j can be different from zero, according to the selection-rule theorem. Whether this is actually the case, if the ψ_i are given by the effective-mass approximation Eq. (6) with $\psi_0=\exp(i\mathbf{k}\cdot\mathbf{r})u_\mathbf{k}(\mathbf{r})$, can be seen from the matrices[15]

$$M_{ij,2}=-\frac{\hbar^2}{4m^2c^2}\Bigg[\int_{u.c.}\exp(i(\mathbf{k}_j-\mathbf{k}_i)\cdot\mathbf{r})$$
$$\times u_{\mathbf{k}_i}^*(\nabla V\times i\nabla u_{\mathbf{k}_j})_x\frac{d\tau}{\Omega}+\int_{a.s.}\exp(i(\mathbf{k}_j-\mathbf{k}_i)\cdot\mathbf{r})$$
$$\times F_i^*F_j u_{\mathbf{k}_i}^*(\nabla U\times i\nabla u_{\mathbf{k}_j})_x d\tau\Bigg]. \quad (23)$$

It is assumed that the envelope functions F_i and their first derivatives are smooth functions of $\mathbf{r}$. The first integral (over a unit cell) accounts for the effect of spin-orbit coupling on a Bloch state $\mathbf{k}_i$; it is zero for all i, j as can be shown with the help of the selection-rule theorem.[16] Furthermore, the corresponding diagonal matrix element for an arbitrary state of the lowest ("s-like") conduction band vanishes, as can be seen by writing $u_\mathbf{k}(\mathbf{r})$ and $V(\mathbf{r})$ as a Fourier series and then applying the relation $u_\mathbf{k}^*=au_{-\mathbf{k}}$ ($aa^*=1$) which follows from the inversion symmetry of $V(\mathbf{r})$.[17] The second integral of Eq. (23) describes the effect of spin-orbit coupling on the effective-mass ground state; it can be shown that this integral also vanishes for all i, j.[18]

---

[15] The integral can diverge in its present form. This divergence is a consequence of the approximation applied in reducing the four-component Dirac equation to the two-component Pauli equation. The difficulty is removed, as in the theory of atomic spectra, by replacing ∇v with $(1-v/2mc^2)^{-2}\nabla v$.

[16] The Bloch state $\mathbf{k}(\pi|a,\pi|a,\pi|a)$ (of the conduction band of Ge) transforms according to the small representation L_1; the x, y, and z components of a pseudo-vector transform according to L_3 and L_2, respectively. The direct product $(L_3+L_1)\times L_1$ does not contain L_1. However, in the second-order approximation of perturbation theory, the energy levels are determined by matrix elements between an L_1 state and corresponding states of higher conduction bands which may not vanish.

[17] R. J. Elliott, Phys. Rev. 96, 266 (1954).

[18] Since the first integral of Eq. (23) vanishes, we have $M_{i,j2}=\int\psi_i^*h_x(U)\psi_j d\tau$, where $h_x(U)=-(\hbar^2/4m^2c^2)(\nabla U\times i\nabla)_x$. This expression for $M_{ij,2}$ leads to the second integral of Eq. (23), with our assumptions on F_i and on its first derivatives. For the case of germanium it is readily seen that $M_{ij,2}$ vanishes. Using Elliott's relation and denoting with ψ_{-i} the effective-mass wave function for $-\mathbf{k}_i$ which contains $F_{-i}(\mathbf{r})=F_i(\mathbf{r})$, we find $M_{ij,2}=-M_{-j-i,2}$. The wave vectors $\mathbf{k}_i$ and $-\mathbf{k}_i$ differ by 2π times a vector of the reciprocal lattice, thus $M_{-j-i,2}=M_{ji,2}$. Furthermore, since the

Hence, for Ge, there is no effect of spin-orbit coupling in first-order perturbation theory if the zero-order wave functions are taken from the effective-mass approximation Eq. (6). The same result holds for the effective-mass ground state of donors in Si. An upper limit for the second-order change of the energy of the effective-mass ground state is estimated roughly in Appendix B; it is found to be 5×10^{-5} eV for donor electrons in Ge.

The first-order effect on the exact eigenstates of H_0 cannot be calculated, since it is not known to what extent Bloch functions of higher bands are admixed into the impurity wave functions because of the strong impurity potential in the central cell. An order of magnitude estimate for the ratio of the coefficients $A_n (n \neq 0)$ and A_0, defined by Eq. (4), is given by Kohn.[1] From this estimate it is seen that the admixture for donors in Ge is much larger than for donors in Si. The different extent of admixture appears to be a characteristic distinction between impurity wave functions in Ge and Si. This distinction can be important for the g factors of donor electrons.[19]

C. s-Shift and Mass-Velocity Correction

These corrections[20] do not contain the electron spin and, therefore, they cannot cause a splitting of the

---

$M_{ij,2}$ are elements of a secular matrix and since these elements are real according to Eq. (19), we have $M_{ij,2} = M_{ji,2}$ and, therefore, the matrix elements vanish.

[19] A difference between the g factor of donor electrons and that approximative g factor for donor electrons given by L. M. Roth and B. Lax [Phys. Rev. Letters 3, 217 (1959), Eq. (5)] in terms of the tensor components $g_{\parallel}$ and $g_{\perp}$ calculated for a single ellipsoidal energy surface of conduction electrons, may be caused to some extent by the admixture of Bloch functions from higher bands into the impurity wave function. For antimony, phosphorus, and arsenic impurities in Si, the g factors of donor electrons have been measured by D. K. Wilson and F. Feher [Phys. Rev. 124, 1068 (1961)]. For conduction electrons in Si, experimental values for the anisotropic g factor do not seem to be available. However, L. M. Roth [Phys. Rev. 118, 1534 (1960)], H. Hagesawa [Phys. Rev. 118, 1523 (1960)], and L. Liu [Phys. Rev. Letters 6, 683 (1961)] have calculated the g factor of conduction electrons in Si. The experimental result for the isotropic g factor of donor electrons and the theoretical result given in terms of $g_{\parallel}$ and $g_{\perp}$ are in quantitative agreement. Thus, it appears that the impurity wave functions in Si are well described with the effective-mass approximation which neglects the admixture of Bloch functions from higher conduction bands. On the other hand, the experimental values of the g factors for donor electrons of phosphorus or arsenic impurities in Ge measured by G. Feher, D. K. Wilson, and E. A. Gere [Phys. Rev. Letters 3, 25 (1959)] are not in quantitative agreement with the theoretical value derived from $g_{\parallel}$ and $g_{\perp}$ for an ellipsoidal energy surface of conduction electrons. The difference may be caused, as in the more subtle case of antimony-doped Ge (see Feher et al., Roth and Lax), to some extent by the admixture of Bloch functions from higher conduction bands into the impurity wave functions. An excellent review article on g factors (and spin-lattice relaxation) has been presented by Y. Yafet, in Solid-State Physics, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1963), Vol. 14, p. 1.

[20] F. Herman (private communication) has solved the Hartree-Fock-Slater wave equations for neutral germanium and silicon atoms and then applied first-order perturbation theory to account for the relativistic corrections. It is a pleasure to thank Dr. Herman for a number of interesting comments concerning the relativistic corrections on Bloch states and for making available to us his numerical results for the three relativistic corrections on atomic orbitals of germanium and silicon atoms.

---

degenerate state $^2\Gamma_7$ and $^4\Gamma_8$. However, since both the s-shift correction H_3 and the mass-velocity correction H_4 have the full tetrahedral symmetry, they can partly lift the $2m$-fold degeneracy of the effective-mass ground state, as does the potential correction H_1. Let us compare the shift and the splitting caused by H_3 and H_4 with the corresponding effect of H_1. If one substitutes the effective-mass wave function into Eq. (15), the matrix elements of the s-shift correction are given by

$$M_{ij,3} = -\frac{\hbar^2}{4m^2c^2}\left[\frac{1}{\Omega}\int_{u.c.} \psi_0^*(\mathbf{k}_i\mathbf{r})\boldsymbol{\nabla} V \cdot \boldsymbol{\nabla}\psi_0(\mathbf{k}_j,\mathbf{r})d\tau\right.$$

$$\left. + |F(0)|^2 \int_{c.c.} \psi_0^*(\mathbf{k}_i,\mathbf{r})\boldsymbol{\nabla} U \cdot \boldsymbol{\nabla}\psi_0(\mathbf{k}_j,\mathbf{r})d\tau\right]. \quad (24)$$

Similarly, the matrix elements of the mass-velocity correction are found as

$$M_{ij,4} = -\frac{1}{2mc^2}\left[\frac{1}{\Omega}\int_{u.c.} \psi_0^*(\mathbf{k}_i,\mathbf{r})(E_0+V)^2\psi_0(\mathbf{k}_j,\mathbf{r})d\tau\right.$$

$$\left. + |F(0)|^2 \int_{c.c.} \psi_0^*(\mathbf{k}_i,\mathbf{r})U(\mathcal{U}+V+2E_0)\psi_0(\mathbf{k}_i,\mathbf{r})d\tau\right]. \quad (25)$$

The second contribution to each matrix element has been restricted to the central cell where the potential U has a singularity. The first term of each of the two diagonal elements [Eqs. (24) and (25)] corresponds to the shift of the Bloch state of the conduction band minimum. For these terms there are no quantitative values available at present. However, Herman[20] has calculated the effect of all three relativistic corrections on the $3s$ and $3p$ atomic orbital of a free Si atom and on the $4s$ and $4p$ atomic orbital of a free Ge atom. His result for the s-shift correction to the $3s$ and $4s$ atomic states of Si and Ge is $+0.16$ and $+0.93$ eV, respectively, and the corresponding numbers for the mass-velocity correction are -0.62 and -1.62 eV. Hence, there is a partial cancellation of the two corrections. Assuming that the net shift of the conduction band minimum of Ge is $\frac{1}{2}$ eV, we estimate roughly that the second terms of both diagonal elements, which describe the relative shift of the ground state with respect to the conduction band minimum, give a net shift of the order of magnitude $\frac{1}{2}|F(0)|^2\Omega$ eV, for As and Sb donors. Thus, with the corrected envelope function, the shift amounts to $\sim 5 \times 10^{-4}$ and $\sim 5 \times 10^{-3}$ eV for As and Sb donor states in germanium and silicon, respectively.

The nondiagonal elements of Eqs. (24) and (25) contribute to the observed splitting of the degenerate effective-mass ground state (see Fig. 1). In the case of donor electrons in Ge, the relative contribution of the two relativistic corrections to the splitting between the $^2\Gamma_6$ state and the $^2\Gamma_7 + ^4\Gamma_8$ state is determined by the

ratio

$$-\frac{1}{2mc^2}\int_{c.c.}\psi_0^*(\mathbf{k}_i)U(\mho+V+2E_0)\psi_0(\mathbf{k}_j)d\tau \; - \; \frac{\hbar^2}{4m^2c^2}\int_{c.c.}\psi_0^*(\mathbf{k}_i)\nabla U\cdot\nabla\psi_0(\mathbf{k}_j)d\tau$$
$$\overline{\int_{c.c.}\psi_0^*(\mathbf{k}_i)(U-U_0)\psi_0(\mathbf{k}_j)d\tau} \tag{26}$$

Since it is the anisotropic part of the potential $U(\mathbf{r})$ which determines the integrals and since this part has no singularity at $\mathbf{r}=0$, the contribution of the relativistic corrections to the observed splitting is small compared to the valley-orbit contribution, even in the case of Sb donors in germanium where the splitting is an order of magnitude smaller than for P and As donors.

V. CONCLUSION

The main purpose of this paper has been to discuss the possible importance of spin-orbit coupling and other relativistic corrections on the degenerate effective-mass ground state of shallow impurities in Ge and Si.

With the help of perturbation theory and the relevant selection-rule theorem, one finds that in first order, spin-orbit coupling does not affect the impurity ground states and, in particular, one finds that there is no splitting of the sextet state in Ge or Si. In second order, the effect of spin-orbit coupling is estimated to be of the order 5×10^{-5} eV or smaller. It is emphasized that these results depend on an assumption of the effective-mass theory; namely, the admixture of Bloch functions from higher conduction bands into the impurity wave function can be neglected.[21] If one could go beyond this approximation, one should expect a first-order effect of spin-orbit coupling on the impurity ground states in Ge or Si, since the Bloch states of higher "non-s-like" conduction bands are also affected by spin-orbit coupling in first order.

At present it is not known to what quantitative extent Bloch functions from higher bands are admixed into the impurity wave function.[22] Kohn's order of magnitude estimate indicates, however, that the admixture is considerably larger for impurity states in Ge than it is for those in Si. The admixture will determine, to some extent, the difference between the g factor of a donor electron and the appropriate g factor of the conduction electrons. For Si there is no such difference (see Ref. 19). For Ge there is a small difference (to our knowledge there are no available experimental g values for the conduction electrons).

As for the two other relativistic corrections, namely the s-shift correction and the mass-velocity correction, they do not depend on the spin and they have the symmetry of $U(\mathbf{r})$. Their effect on the degenerate effective-mass ground state consists of a binding energy increase of the order of 10%. The splitting caused by these corrections is determined by the anisotropic part of U which—due to the tetrahedral symmetry of U— vanishes with $r^n(n\geq3)$ for $r\to0$. Therefore, the splitting is small compared with the valley-orbit splitting.

Thus, it can be seen that, although spin-orbit coupling and the other relativistic corrections are of some importance for the binding energy of the impurity states, and for the g factors of donor electrons, the relativistic corrections cannot account for the observed energy level structure of the impurity ground states. Spin-orbit coupling will not result in an additional splitting of degenerate energy levels large enough to be observed with present experimental techniques. As for the relative importance of corrections (1), (2), and (3), it appears that the potential correction $U-U_0$, i.e., correction (1), which lowers the effective-mass binding energy and causes the valley-orbit splitting, is the most important one, at least for donor states in Si. Thus, an accurate computation of the U's for different impurities in Si would be of considerable interest. For Ge the admixture of Bloch functions from higher bands may also be of some importance (see Ref. 19 for g factors of donor electrons). In such a case the small observed valley-orbit splitting of antimony donors, which is by an order of magnitude smaller than for phosphorus and arsenic donors, can be caused by a partial compensation of the non-relativistic corrections (1) and (2).

APPENDIX A: THE COEFFICIENTS α_i^{rs} AND β_i^{rs} (Ge)

The correct zero-order wave functions given by Eq. (8) can be determined by symmetry considerations. To this end, let us write the wave functions of the Kohn-Luttinger Hamiltonian H_0 which are obtained in the effective-mass approximation and which trans-

---

[21] Since the Bloch states of a higher "non-s-like" conduction band will be affected to a different degree, depending on the particular $\mathbf{k}$, by spin-orbit coupling (or by the other relativistic corrections), the effect of the admixture will certainly depend on the composition of the impurity wave function in terms of higher band Bloch states.

[22] The admixture of a higher band Bloch state $\mathbf{k}$ into the impurity wave function will depend on the strength of the impurity potential in the central cell and on the energy gap between the energy of the state $\mathbf{k}$ and the conduction band minimum.

form according to the representations T_1 and T_3 of the tetrahedral point group T_d, in the form

$$\Psi^{(1)} = \sum_i a_i{}^{(1)}\psi_i, \quad \Psi_j{}^{(3)} = \sum_i a_{ij}{}^{(3)}\psi_i, \quad (j=1,2,3), \quad \text{(A1)}$$

where ψ_i is given by Eq. (6). If we number the minima of the conduction band in the $(1,-1,-1)$, $(-1,1,-1)$, $(1,1,1)$, $(-1,-1,1)$ directions by 1, 2, 3, 4, the coefficients $a_2{}^{(i)}$ and $a_{3j}{}^{(i)}$ $(i=1,2,3,4)$ are given by:

for (T_1)

$$a_i{}^{(1)} = \tfrac{1}{2}(1,1,1,1),$$

and for (T_3),

$$a_{i1},{}^{(3)} = \tfrac{1}{2}(1,-1,-1,1), \quad \text{(A2)}$$
$$a_{i2},{}^{(3)} = \tfrac{1}{2}(1,-1,1,-1),$$
$$a_{i3},{}^{(3)} = \tfrac{1}{2}(1,1,-1,-1).$$

With the help of Eqs. (A1) and (A2) the two-component wave functions transforming according to the representations $^2\Gamma_6$, $^2\Gamma_7$, and $^4\Gamma_8$ can be determined immediately. The first case ($^2\Gamma_6$) is trivial. In order to determine the wave functions associated with $^2\Gamma_7$ and $^4\Gamma_8$, let us write

$$X = \sum_i a_{i1}{}^{(3)}\phi_i, \quad Y = \sum a_{i2}{}^{(3)}\phi_i, \quad Z = \sum a_{i3}{}^{(3)}\phi_i, \quad \text{(A3)}$$

where ϕ_i is given by Eq. (9). In terms of these functions and their time reversal conjugates (indicated by a bar), the correct zero-order wave functions are given by: for $(^2\Gamma_7)$

$$\Psi_{17} = (1/\sqrt{3})(\bar{X} - i\bar{Y} + Z), \quad \Psi_{27} = K\Psi_{17}$$

and for $(^2\Gamma_8)$

$$\Psi_{18} = (1/\sqrt{2})(X+iY), \qquad \Psi_{38} = K\Psi_{28}, \quad \text{(A4)}$$
$$\Psi_{28} = (1/6^{1/2})(\bar{X}+i\bar{Y}-2Z), \quad \Psi_{48} = K\Psi_{18}.$$

APPENDIX B: SPIN-ORBIT COUPLING IN SECOND ORDER (Ge)

An upper limit for the second order change of the energy of the effective-mass ground state is given by

$$E_2 = \int_{a.s.} (\psi_i u_\alpha)^* H_{so}{}^2 \psi_i u_\alpha d\tau \Big/ \Delta E, \quad \text{(B1)}$$

where $\Delta E \sim 1$ eV. Let us assume that in the vicinity of an impurity nucleus the effective-mass wave function can be written in the form

$$\psi_i(\text{Ge}) = F(\mathbf{r})(\Omega/2)^{1/2}[a(\mathbf{k}_i)\psi_{4s} + b(\mathbf{k}_i)\psi_{4p}], \quad \text{(B2)}$$

where ψ_{4s} and ψ_{4p} are normalized atomic orbitals of a free Ge atom. The dominant contribution to the matrix element of Eq. (B1) comes from inside the central cell. In the limit where $r \to 0$, the potential has the form

$$U(\mathbf{r}) = -[e^2(Z_d - Z_0)/r] + \text{const}\,xyz. \quad \text{(B3)}$$

Clearly, it is the Coulombic term of the potential which, together with the $4p$ orbital in ψ_i, determines the value of E_2; one finds

$$E_2 = (|F(0)|^2 \Omega b^2/2)[(Z_d - Z_0)/Z_0]^2 \mathscr{E}_2(4p), \quad \text{(B4)}$$

where $\mathscr{E}_2(4p)$ is an upper limit for the second order change of the atomic $4p$ orbital given by

$$\mathscr{E}_2(4p) = \left(\frac{\hbar^2}{4m^2c^2}\right)^2 \frac{1}{\Delta E a_0{}^4}\left(\frac{Z_0 e^2}{a_0}\right)^2$$

$$\times \frac{4\pi}{3}\int_{a.s.} \frac{(R_{4p})^2 dr}{r^4(1+Z_d r_0/2r)^4}, \quad \text{(B5)}$$

where a_0 is the Bohr radius. This quantity can be calculated without difficulty if one substitutes for the radial eigenfunction R_{4p}, the approximative formula which applies for $r/a_0 \ll n^2/Z$.[23] The resulting upper limit $\mathscr{E}_2(4p)$ gives an unrealistic large value for the actual second-order change of the atomic $4p$ orbital. Therefore, E_2 is estimated roughly by substituting for $\mathscr{E}_2(4p)$ the value found by Herman[20] for the first-order correction of the $4p$ orbital, 0.071 eV. Taking $b^2 = 1/10$ and assuming that the corrected value for $|F(0)|^2$ is larger by a factor of ten than the corresponding value for the effective-mass envelope function, one finds $E_2 < 5 \times 10^{-5}$ eV.

[23] H. A. Bethe and E. E. Salpeter, in *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1957), Vol. 35, p. 1.