PHYSICAL REVIEW B VOLUME 11, NUMBER 2 15 JANUARY 1975

Ground-state wave function for shallow-donor electrons in silicon. II. Anisotropic electron-nuclear-double-resonance hyperfine interactions

Jerry L. Ivey*
Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio 45433

Robert L. Mieher
Department of Physics, Purdue University, West Lafayette, Indiana 47907
(Received 12 August 1974)

We report in this work the calculation of the anisotropic hyperfine constants for the Si^{29} lattice nuclei surrounding the shallow donor impurities P, As, Sb in silicon. We have used a Slater-Koster linear combination of atomic orbitals representation of the Bloch functions of the silicon lattice. These Bloch functions are used in the Brillouin-zone expansion of the impurity wave function. With these calculations we are able to further substantiate our earlier matchings of the Si^{29} lattice sites and the electron-nuclear-double-resonance shells measured by Hale and Mieher. The formalism developed is sufficiently general so that it might be used in future calculations of other electronic defects.

I. INTRODUCTION

The impurity center formed when a silicon atom is replaced by an element from the fifth column of the periodic table (e.g., P, As, Sb, Bi) is called a donor. Four of the five outer electrons of the impurity form covalent bonds and do not contribute to any paramagnetic resonance. The fifth electron, however, when bound to the impurity center at low temperatures, exhibits a paramagnetic resonance. This resonance was first studied by Fletcher *et al.*,[1,2] and later by Feher,[3] by Hale and Mieher,[4] and by Hale and Castner.[5] It has been well established by these workers that the spin resonance of donor electrons can be represented by the electron-spin-resonance (ESR) spin Hamiltonian

$$H_{\mathrm{ESR}} = g\mu_B \vec{H}_0 \cdot \vec{S} + a_D \vec{I}_D \cdot \vec{S} \tag{1}$$

and by the electron-nuclear-double-resonance (ENDOR) spin Hamiltonian

$$H_{\mathrm{ENDOR}} = \sum_l \vec{I}_l \cdot (\vec{\vec{A}}_l \cdot \vec{S} - g_n \mu_N \vec{\vec{H}}_0). \tag{2}$$

In these equations $g\mu_B \vec{H}_0 \cdot \vec{S}$ and $g_n \mu_N \vec{I} \cdot \vec{H}_0$ are the electronic and the nuclear Zeeman terms, l is a parameter denoting Si^{29} lattice nuclei with nuclear spin $\vec{I}_l$, and $\vec{\vec{A}}_l$ is the hyperfine interaction tensor. The term $a_D \vec{I}_D \cdot \vec{S}$ represents the interaction of the electron with the donor nucleus.

The tensor $\vec{\vec{A}}_l$ is separated into two components

$$\vec{\vec{A}}_l = a_l \vec{\vec{1}} + \vec{\vec{B}}_l , \tag{3}$$

where a_l is the isotropic Fermi-contact interaction constant and $\vec{\vec{B}}_l$ is the anisotropic dipole-dipole interaction tensor. These components can be expressed as

$$a_l = (\tfrac{8}{3}\pi) g g_n \mu_B \mu_N |\Psi(\vec{R}_l)|^2 , \tag{4}$$

$$(\vec{\vec{B}}_l)_{ij} = g g_n \mu_B \mu_N \langle \Psi(\vec{r}) | D_{ij}(\vec{r} - \vec{R}_l) | \Psi(\vec{r}) \rangle , \tag{5}$$

$$D_{ij}(\vec{r}) = (3x_i x_j - r^2 \delta_{ij})/r^5 , \tag{6}$$

where $\Psi(\vec{r})$ is the wave function of the bound electron. The ENDOR experiment furnishes the numbers a, B_{zz}, B_{xy}, and B_{xz} for 22 values of l, or for 22 shells of equivalent Si^{29} lattice nuclei. These values are given in the work of Hale and Mieher.[4]

Even with the additional information furnished by the dependence of the a_l with uniaxial stress,[5] it was found to be impossible[3,6-10] to use the effective-mass theory (EMT) to correlate all the experimental values with the shells of lattice nuclei which give rise to the individual interactions. There have also been four previous attempts to calculate the anisotropic hyperfine tensor for the shallow donors which utilized different approximations to the Bloch functions $\psi_n(\vec{k}, \vec{r})$: (i) Mozer,[11] who used a tight-binding approximation using s and p functions; (ii) Hale and Mieher,[7] who used an equivalent orbital representation; (iii) Ivey and Mieher,[12] who used pseudopotential Bloch functions which were orthogonalized to the core states; and, (iv) Castner,[8] who used plane waves and a tight-binding approximation for the cellular part of the Bloch function. All of these approaches suffered from two major limitations—they all assumed the validity of the EMT and they all failed to account completely for the anisotropic hyperfine constants of all the experimentally observed shells.

The resolution of the difficulty of the shell assignment came in the work of Ivey[13] and Ivey and Mieher[14,15] who were able to assign all the experimental ENDOR shells to their respective shells of

lattice nuclei on the basis of the isotropic hyperfine constants and the strain dependence of these constants. In that work it was found necessary to take proper consideration of the matrix elements of the impurity potential between Bloch functions and of the screening of the impurity potential by the dielectric response function of the silicon lattice. This represented a significant advance over previous works which used effective-mass and single-band concepts in that a full Brillouin zone numerical integration was performed. The wave function of the electron bound to the impurity can be represented in terms of the Bloch functions of the unperturbed lattice as

$$\Psi(\vec{r}) = \sum_{n\vec{k}} A_n(\vec{k})\psi_n(\vec{k},\vec{r}) . \tag{7}$$

With the use of pseudopotential Bloch functions the electronic wave-function density at the lattice sites and hence the isotropic hyperfine constants could be calculated. The summation over n and $\vec{k}$ were performed over eight bands (four valence and four conduction) and 200 values of $\vec{k}$ in $\frac{1}{48}$ of the Brillouin zone. In this work we will use the same $A_n(\vec{k})$ as in the previous work.[15] To calculate the anisotropic hyperfine constants, however, a better representation of the Bloch functions of the silicon lattice is needed.

There are two major criteria which any calculation of these hyperfine constants must meet. The first is to explain the rather large experimental values for the B_{xy} of the shells which lie along the [100] crystal axis (e.g., the A and K shells). It was noted by Feher[3] in his original work and the work that he reported to have been done by Mozer that s and p functions alone in a tight-binding or linear-combination-of-atomic-orbitals (LCAO) approximation could not explain these observations. (It has been found by the present authors that this remains true even if a detailed Brillouin-zone integration is performed.) The second criterion for a good theory is that it must explain the rather large differences between the anisotropic hyperfine constants of inversion-related shells such as the L-$(2,2,4)$ and M-$(2,2,\bar{4})$, the I-$(2,2,\bar{8})$ and P-$(2,2,8)$, and the H-$(4,4,\bar{4})$ and O-$(4,4,4)$ which have been previously reported.[13-15] The isotropic Fermi-contact constants for these pairs of shells never differ by as much as 10%, while their anisotropic hyperfine constants can differ by almost a factor of 10.

The question posed then is "does there exist an adequate representation for the Bloch functions such that the anisotropic hyperfine constants can be explained, consistent with the prior agreement of the isotropic hyperfine constants and their stress dependence?" In Sec. II of this paper we will demonstrate how to develop such a representation. In Sec. III we show how to express the hyperfine constants in terms of an auxiliary function $f(\vec{R}_l)$, and in Sec. IV go into the details of the computation of the $f(\vec{R}_l)$. In Sec. V we present the quantitative results and compare them with experiment. In Secs. VI and VII we present a discussion of our results and the conclusions of our work.

II. SLATER-KOSTER LCAO BLOCH FUNCTIONS

The Slater-Koster method[16] assumes Bloch functions of the form

$$\psi_n(\vec{k},\vec{r}) = \sum_t C_{n,t}(\vec{k})\Phi_t(\vec{k},\vec{r}) , \tag{8}$$

$$\Phi_t(\vec{k},\vec{r}) = N^{-1/2}\sum_{\vec{R}_l} e^{i\vec{k}\cdot\vec{R}_l}\phi_t(\vec{r}-\vec{R}_l) . \tag{9}$$

In this paper we use s, p, and d functions such that

$$t = s, p_x, p_y, p_z, d_{xy}, d_{yz}, d_{zx}, d_{x^2-y^2}, d_{3z^2-r^2} . \tag{10}$$

There are 18 ϕ_t's necessary because there are two basis sites in the unit cell of the diamond lattice (O_h^7). The ϕ_t's are taken to be products of radial functions and spherical harmonics on each basis site:

$$\begin{aligned}
&\phi_s(\vec{r}) = U_s(r)Y_0^0 ,\\
&\phi_{p_y}(\vec{r}) = i\,U_p(r)(Y_1^{-1}+Y_1^1)/\sqrt{2} ,\\
&\phi_{d_{xy}}(\vec{r}) = i\,U_d(r)(Y_2^{-2}-Y_2^2)/\sqrt{2} ,\\
&\phi_{d_{zx}}(\vec{r}) = U_d(r)(Y_2^{-1}-Y_2^1)/\sqrt{2} ,\\
&\phi_{p_x}(\vec{r}) = U_p(r)(Y_1^{-1}-Y_1^1)/\sqrt{2} ,\\
&\phi_{p_z}(\vec{r}) = U_p(r)Y_1^0 ,\\
&\phi_{d_{yz}}(\vec{r}) = i\,U_d(r)(Y_2^{-1}+Y_2^1)/\sqrt{2} ,\\
&\phi_{d_{x^2-y^2}}(\vec{r}) = U_d(r)(Y_2^{-2}+Y_2^2)/\sqrt{2} ,\\
&\phi_{d_{3z^2-r^2}}(\vec{r}) = U_d(r)Y_2^0 .
\end{aligned} \tag{11}$$

All the matrix elements of the Hamiltonian of the lattice can be evaluated in terms of the two-center approximation using the results given in the paper of Slater and Koster.[16] These are taken to be the two-center integrals $(ss\sigma)$, $(sp\sigma)$, $(pp\sigma)$, $(pp\pi)$, etc. With the approximation of the neglect of higher than second-nearest neighbor interactions and the use of s, p, and d basis functions, there are 23 of these numbers. In this work we have determined these numbers by fitting to the pseudopotential energy bands of silicon used in the calculation of the $A_n(\vec{k})$ in the previous work. The same values of $A_n(\vec{k})$ as in the previous work[15] are used in this work. The values of the two-center

integrals are given in Table I of this work. We performed our fitting procedure by the use of a modified "method-of-steepest-descent" computer subroutine called MINUM[17] which minimizes a function of several variables. In this work the least-squares error of the calculated Slater-Koster LCAO energy bands from the values of the energies being fit was minimized. Figure 1 shows the energy bands for silicon calculated from the pseudopotential Bloch functions and the values calculated with the two-center integrals given in Table I.

TABLE I. Two-center integrals from Slater-Koster method.[a]

$(ss\sigma)_0 = -2.386$	$(ss\sigma)_1 = -2.177$	$(ss\sigma)_2 = -0.120$
$(pp\sigma)_0 = 4.227$	$(sp\sigma)_1 = 2.601$	$(sp\sigma)_2 = 0.182$
$(dd\sigma)_0 = 17.365$	$(pp\sigma)_1 = 3.244$	$(pp\sigma)_2 = 0.525$
	$(pp\pi)_1 = -1.710$	$(pp\pi)_2 = 0.073$
	$(pd\sigma)_1 = -3.194$	$(pd\sigma)_2 = -.733$
	$(pd\pi)_1 = 2.231$	$(pd\pi)_2 = 0.161$
	$(dd\sigma)_1 = -6.465$	$(dd\sigma)_2 = -1.300$
	$(dd\pi)_1 = 3.935$	$(dd\pi)_2 = 0.272$
	$(dd\delta)_1 = -0.407$	$(dd\delta)_2 = -0.111$
	$(sd\sigma)_1 = -3.012$	$(sd\sigma)_2 = -0.349$

[a]All values of integrals are in eV. The subscripts refer to the origin, the first-nearest neighbor, and the second-nearest neighbor.

III. CALCULATION OF HYPERFINE CONSTANTS

The advantage of the LCAO formalism for the Bloch functions is the concise form possible for both the isotropic and anisotropic hyperfine constants. By the definition

$$f_t(\vec{R}_l) = N^{-1/2}\sum_{n\vec{k}} A_n(\vec{k}) C_{n,t}(\vec{k}) e^{i\vec{k}\cdot\vec{R}_l} , \tag{12}$$

one can write Eqs. (4) and (5) as

$$a_l = (\tfrac{8}{3}\pi) g g_n \mu_B \mu_N \sum_{\substack{l' l'' \\ t t'}} f_t(\vec{R}_{l'}) f_{t'}(\vec{R}_{l''}) \times \phi_t(\vec{R}_l - \vec{R}_{l'}) \phi_{t'}(\vec{R}_l - \vec{R}_{l''}) , \tag{13}$$

$$(\vec{B}_l)_{ij} = g g_n \mu_B \mu_N \sum_{\substack{l' l'' \\ t t'}} f_t(\vec{R}_{l'}) f_{t'}(\vec{R}_{l''}) \times \langle \phi_t(\vec{r}_l - \vec{R}_{l'}) | D_{ij}(\vec{r}_l) | \phi_{t'}(\vec{r}_l - \vec{R}_{l''}) \rangle . \tag{14}$$

Note that the f_t's do not contain a dependence upon the radial variable $\vec{r}$. They are to be calculated from Eq. (12) once the $C_{n,t}(\vec{k})$ and $A_n(\vec{k})$ are determined. They are always real valued because $C_{n,t}(-\vec{k}) = C^*_{n,t}(\vec{k})$ and $A_n(-\vec{k}) = A^*_n(\vec{k})$. If one used an explicit form of the $U_t(r)$, the various products and matrix elements in Eqs. (13) and (14) could be determined. At this time, however, we choose to consider only the most important contributions to these values. For a_l these contributions are those for $\vec{R}_{l'} = \vec{R}_l$ and $\vec{R}_{l''} = \vec{R}_l + \vec{R}_i$, where $\vec{R}_i = 0$ and also where $\vec{R}_i$ are the vectors to the four nearest neighbors of $\vec{R}_l$. For B_l these contributions are those only for $\vec{R}_l = \vec{R}_{l'} = \vec{R}_{l''}$. We may then write

$$a_l \simeq (\tfrac{8}{3}\pi) g g_n \mu_B \mu_N \left[f_s^2(\vec{R}_l)\phi_s^2(0) + f_s(\vec{R}_l)\phi_s(0) \sum_{t,\vec{R}_i} f_t(\vec{R}_l + \vec{R}_i)\phi_t(-\vec{R}_i) \right] , \tag{15}$$

$$(\vec{B}_l)_{ij} \simeq g g_n \mu_B \mu_N \sum_{t,t'} f_t(\vec{R}_l) f_{t'}(\vec{R}_l) \langle \phi_t(\vec{r}) | D_{ij}(\vec{r}) | \phi_{t'}(\vec{r}) \rangle . \tag{16}$$

A detailed consideration of the various terms in Eq. (15) together with Eqs. (11) results in the following expression for a_l:

$$\begin{aligned} a_l = \tfrac{2}{3} g g_n \mu_B \mu_N \{ & f_s^2(\vec{R}_l) U_s^2(0) + f_s(\vec{R}_l) U_s(0) U_s(R_B) [f_s(\vec{R}_l + \vec{R}_1) + f_s(\vec{R}_l + \vec{R}_2) + f_s(\vec{R}_l + \vec{R}_3) + f_s(\vec{R}_l + \vec{R}_4)] \\ & + (-1)^M f_s(\vec{R}_l) U_s(0) U_p(R_B) [f_{p_x}(\vec{R}_l + \vec{R}_1) + f_{p_x}(\vec{R}_l + \vec{R}_2) - f_{p_x}(\vec{R}_l + \vec{R}_3) - f_{p_x}(\vec{R}_l + \vec{R}_4) \\ & + f_{p_y}(\vec{R}_l + \vec{R}_1) - f_{p_y}(\vec{R}_l + \vec{R}_2) + f_{p_y}(\vec{R}_l + \vec{R}_3) - f_{p_y}(\vec{R}_l + \vec{R}_4) \\ & + f_{p_z}(\vec{R}_l + \vec{R}_1) - f_{p_z}(\vec{R}_l + \vec{R}_2) - f_{p_z}(\vec{R}_l + \vec{R}_3) + f_{p_z}(\vec{R}_l + \vec{R}_4)] \\ & + f_s(\vec{R}_l) U_s(0) U_d(R_B) \sqrt{\tfrac{5}{3}} [f_{d_{xy}}(\vec{R}_l + \vec{R}_1) - f_{d_{xy}}(\vec{R}_l + \vec{R}_2) - f_{d_{xy}}(\vec{R}_l + \vec{R}_3) + f_{d_{xy}}(\vec{R}_l + \vec{R}_4) \\ & + f_{d_{yz}}(\vec{R}_l + \vec{R}_2) + f_{d_{yz}}(\vec{R}_l + \vec{R}_2) - f_{d_{yz}}(\vec{R}_l + \vec{R}_3) - f_{d_{yz}}(\vec{R}_l + \vec{R}_4) \\ & + f_{d_{zx}}(\vec{R}_l + \vec{R}_1) - f_{d_{zx}}(\vec{R}_l + \vec{R}_2) + f_{d_{zx}}(\vec{R}_l + \vec{R}_3) - f_{d_{zx}}(\vec{R}_l + \vec{R}_4)] \} , \end{aligned} \tag{17}$$

where $\vec{R}_1=(\frac{1}{4}a)(1,1,1)$, $\vec{R}_2=(\frac{1}{4}a)(1,\bar{1},\bar{1})$, $\vec{R}_3=(\frac{1}{4}a)(\bar{1},1,\bar{1})$, $\vec{R}_4=(\frac{1}{4}a)(\bar{1},\bar{1},1)$ for $M=1$ (denoting an even-numbered lattice site), and $\vec{R}_i=-\vec{R}_i$ for $M=2$ (denoting an odd-numbered lattice site), $R_B=\sqrt{3}\,(\frac{1}{4}a)$, and $a=5.43$ Å for Si.

An evaluation of the matrix elements in Eq. (16) (which are independent of $\vec{R}_l$) results in only three types of terms for a given component of the hyperfine tensor in terms of $B_0=gg_n\mu_B\mu_N$:

$$B^{pp}_{zz}(\vec{R}_l)=\tfrac{2}{5}B_0B_1[2f^2_{p_z}(\vec{R}_l)-f^2_{p_x}(\vec{R}_l)-f^2_{p_y}(\vec{R}_l)], \tag{18a}$$

$$B^{sd}_{zz}(\vec{R}_l)=\sqrt{\tfrac{16}{5}}\,B_0B_2[f_s(\vec{R}_l)f_{d_{3z^2-r^2}}(\vec{R}_l)], \tag{18b}$$

$$B^{dd}_{zz}(\vec{R}_l)=\tfrac{2}{7}B_0B_3\{2[f^2_{d_{3z^2-r^2}}(\vec{R}_l)-f^2_{d_{xy}}(\vec{R}_l)-f^2_{d_{x^2-y^2}}(\vec{R}_l)]+f^2_{d_{zx}}(\vec{R}_l)+f^2_{d_{yz}}(\vec{R}_l)\}, \tag{18c}$$

$$B^{pp}_{xy}(\vec{R}_l)=\tfrac{6}{5}B_0B_1[f_{p_x}(\vec{R}_l)f_{p_y}(\vec{R}_l)], \tag{19a}$$

$$B^{sd}_{xy}(\vec{R}_l)=\sqrt{\tfrac{12}{5}}\,B_0B_2[f_s(\vec{R}_l)f_{d_{xy}}(\vec{R}_l)], \tag{19b}$$

$$B^{dd}_{xy}(\vec{R}_l)=(\tfrac{2}{7}\sqrt{3}\,)B_0B_3[-2f_{d_{xy}}(\vec{R}_l)f_{d_{3z^2-r^2}}(\vec{R}_l)+\sqrt{3}\,f_{d_{zx}}(\vec{R}_l)f_{d_{yz}}(\vec{R}_l)], \tag{19c}$$

$$B^{pp}_{xz}(\vec{R}_l)=\tfrac{6}{5}B_0B_1[f_{p_x}(\vec{R}_l)f_{p_z}(\vec{R}_l)], \tag{20a}$$

$$B^{sd}_{xz}(\vec{R}_l)=\sqrt{\tfrac{12}{5}}\,B_0B_2[f_s(\vec{R}_l)f_{d_{zx}}(\vec{R}_l)], \tag{20b}$$

$$B^{dd}_{xz}(\vec{R}_l)=(\tfrac{2}{7}\sqrt{3}\,)B_0B_3[f_{d_{zx}}(\vec{R}_l)f_{d_{3z^2-r^2}}(\vec{R}_l)+\sqrt{3}\,f_{d_{xy}}(\vec{R}_l)f_{d_{yz}}(\vec{R}_l)+\sqrt{3}\,f_{d_{zx}}(\vec{R}_l)f_{d_{x^2-y^2}}(\vec{R}_l)], \tag{20c}$$

$$B^{pp}_{yz}(\vec{R}_l)=\tfrac{6}{5}B_0B_1[f_{p_y}(\vec{R}_l)f_{p_z}(\vec{R}_l)], \tag{21a}$$

$$B^{sd}_{yz}(\vec{R}_l)=\sqrt{\tfrac{12}{5}}\,B_0B_2[f_s(\vec{R}_l)f_{d_{yz}}(\vec{R}_l)], \tag{21b}$$

$$B^{dd}_{yz}(\vec{R}_l)=(\tfrac{2}{7}\sqrt{3}\,)B_0B_3[f_{d_{yz}}(\vec{R}_l)f_{d_{3z^2-r^2}}(\vec{R}_l)+\sqrt{3}\,f_{d_{xy}}(\vec{R}_l)f_{d_{zx}}(\vec{R}_l)-\sqrt{3}\,f_{d_{yz}}(\vec{R}_l)f_{d_{x^2-y^2}}(\vec{R}_l)], \tag{21c}$$

$$B^{pp}_{xx}(\vec{R}_l)=\tfrac{2}{5}B_0B_1[2f^2_{p_x}(\vec{R}_l)-f^2_{p_y}(\vec{R}_l)-f^2_{p_z}(\vec{R}_l)], \tag{22a}$$

$$B^{sd}_{xx}(\vec{R}_l)=\sqrt{\tfrac{4}{5}}\,B_0B_2f_s(\vec{R}_l)[\sqrt{3}\,f_{d_{x^2-y^2}}(\vec{R}_l)-f_{d_{3z^2-r^2}}(\vec{R}_l)], \tag{22b}$$

$$B^{dd}_{xx}(\vec{R}_l)=\tfrac{2}{7}B_0B_3[f^2_{d_{xy}}(\vec{R}_l)+f^2_{d_{zx}}(\vec{R}_l)+f^2_{d_{x^2-y^2}}(\vec{R}_l)-f^2_{d_{3z^2-r^2}}(\vec{R}_l)-2f^2_{d_{yz}}(\vec{R}_l)-2\sqrt{3}\,f_{d_{x^2-y^2}}(\vec{R}_l)f_{d_{3z^2-r^2}}(\vec{R}_l)], \tag{22c}$$

$$B^{pp}_{yy}(\vec{R}_l)=\tfrac{2}{5}B_0B_1[2f^2_{p_y}(\vec{R}_l)-f^2_{p_z}(\vec{R}_l)-f^2_{p_x}(\vec{R}_l)], \tag{23a}$$

$$B^{sd}_{yy}(\vec{R}_l)=-\sqrt{\tfrac{4}{5}}\,B_0B_2f_s(\vec{R}_l)[\sqrt{3}\,f_{d_{x^2-y^2}}(\vec{R}_l)+f_{d_{3z^2-r^2}}(\vec{R}_l)], \tag{23b}$$

$$B^{dd}_{yy}(\vec{R}_l)=\tfrac{2}{7}B_0B_3[f^2_{d_{xy}}(\vec{R}_l)+f^2_{d_{yz}}(\vec{R}_l)+f^2_{d_{x^2-y^2}}(\vec{R}_l)-f^2_{d_{3z^2-r^2}}(\vec{R}_l)-2f^2_{d_{zx}}(\vec{R}_l)+2\sqrt{3}\,f_{d_{x^2-y^2}}(\vec{R}_l)f_{d_{3z^2-r^2}}(\vec{R}_l)], \tag{23c}$$

The numbers $B_1=\langle U_p|1/r^3|U_p\rangle$, $B_2=\langle U_s|1/r^3|U_d\rangle$, $B_3=\langle U_d|1/r^3|U_d\rangle$, $A_1=U_s^2(0)$, $A_2=U_s(0)U_s(R_B)$; $A_3=U_s(0)U_p(R_B)$, and $A_4=U_s(0)U_d(R_B)$ depend only upon the lattice and not upon any details of the impurity problem. The details of the impurity problem are contained only in the $A_n(\vec{k})$ and thus in the $f_t(\vec{R}_l)$.

IV. EVALUATION OF $f_t(R_l)$

The evaluation of $f_t(\vec{R}_l)$ as given in Eq. (12) is performed by a numerical integration over 200 $\vec{k}$ values in a fundamental wedge which is $\frac{1}{48}$ of the Brillouin zone and for eight bands. The particular wedge used is defined by $k_x\geq k_y\geq k_z\geq 0$. From this wedge can be constructed a larger wedge D_x which consists of all the $\vec{k}$ values which can be obtained from those in the fundamental wedge by the rotations about the k_x axis and reflections through $k_y=k_z$ and $k_y=-k_z$ planes. Each large wedge D_x is a factor of 8 greater in volume than the fundamental wedge since there are six possible wedges D_x, D_{-x}, D_y, D_{-y}, D_z, and D_{-z} in the Brillouin zone. We can then define six functions $F_{t,j}(\vec{R}_l)$ as those parts of Eq. (12) which arise from the integration over the wedge D_j. While it is true that the $f_t(\vec{R}_l)$ in Eq. (12) must be real valued, this is not true of the individual $F_{t,j}(\vec{R}_l)$. As shown in the previous work[15] for the calculation of $A_n(\vec{k})$, we are assured that $A_n(-\vec{k})=A_n^*(\vec{k})$ because of our choice of phase of the Bloch functions. It must also be true that $C_{n,t}(-\vec{k})=C^*_{n,t}(\vec{k})$, so that $F_{t,-j}(\vec{R}_l)=F^*_{t,j}(\vec{R}_l)$. Also, as in the previous work we define "interpolation envelope functions" by

$$F_{t,j}(\vec{R}_l)\,e^{i\vec{k}_{0j}\cdot\vec{R}_l}=N^{-1/2}\int_{D_j}\sum_{n=5,6}A_n(\vec{k})\,C_{n,t}(\vec{k})\,e^{i\vec{k}\cdot\vec{R}_l}\,d\vec{k}, \tag{24}$$

$$G_{t,j}(\vec{R}_l)=N^{-1/2}\int_{D_j}{\sum}'A_n(\vec{k})\,C_{n,t}(\vec{k})\,e^{i\vec{k}\cdot\vec{R}_l}\,d\vec{k}, \tag{25}$$

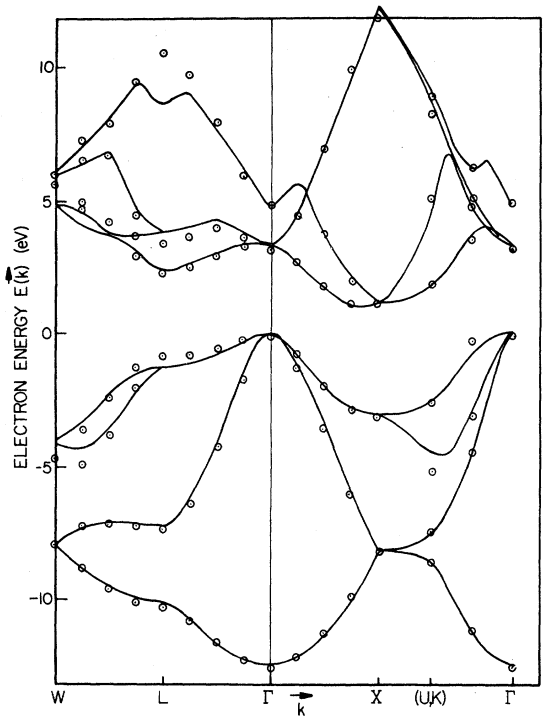

FIG. 1. Solid lines give the energy bands of silicon as calculated by the pseudopotential method as discussed in Ref. 15. The dotted circles give energy values as calculated by the Slater-Koster method with the two-center integral values of Table I. The zero of energy is placed at the top of the valence bands as calculated by the pseudopotential method.

where the prime on the sum indicates that the two lowest conduction bands, $n=5$ and 6, are not included.

These two equations are simply the generalization of Eqs. (37) and (38) of the previous work.[15] Now, however, we have 9 $f_t(\vec{R}_l)$ corresponding to the nine basis functions on each lattice site. We use the symbol $\vec{R}_l$ to refer to the position of an arbitrary lattice site and not to a given unit cell. We thus have two sets of $C_{n,t}(\vec{k})$ corresponding to the two sublattices of the diamond O_h^7 space group for the host lattice. These are given as the solutions of the Slater-Koster LCAO problem as discussed in Sec. II. We may combine Eqs. (24) and (25) as

$$\begin{aligned} f_t(\vec{R}_l) = 2[&F^R_{t,x}(\vec{R}_l)\cos k_0 x_l - F^I_{t,x}(\vec{R}_l)\sin k_0 x_l \\ &+ F^R_{t,y}(\vec{R}_l)\cos k_0 y_l - F^I_{t,y}(\vec{R}_l)\sin k_0 y_l \\ &+ F^R_{t,z}(\vec{R}_l)\cos k_0 z_l - F^I_{t,z}(\vec{R}_l)\sin k_0 z_l \\ &+ G^R_{t,x}(\vec{R}_l) + G^R_{t,y}(\vec{R}_l) + G^R_{t,z}(\vec{R}_l)], \end{aligned} \tag{26}$$

where the R and I superscripts refer to the real and imaginary parts of the $F_{t,j}(\vec{R}_l)$ and $G_{t,j}(\vec{R}_l)$. Equation (26) is now in a form similar to the EMT envelope function[3-7,10] for which the only nonzero terms are the $F^R_{s,j}(\vec{R}_l)$. For the shallow donors (up to about 50-meV binding energy) the $F^I_{s,j}(\vec{R}_l)$ are generally much more important than $G^R_{s,j}(\vec{R}_l)$ in contributing to $f_s(\vec{R}_l)$. A general rule of thumb is that $G^R_{s,j}(\vec{R}_l)$ seldom exceeds 10% of $F^R_{s,j}(\vec{R}_l) \times \cos k_0 x_j$, while $F^I_{s,j}(\vec{R}_l)\sin k_0 x_j$ may become as large as or even exceed $F^R_{s,j}(\vec{R}_l)\cos k_0 x_j$.

V. DETERMINATION OF PARAMETERS $A_1, A_2, A_3, A_4, B_1, B_2, B_3$

To make a determination of these parameters we use a least-squares fitting routine[18] to obtain the particular set which together with Eqs. (17)–(20) will explain the experimental hyperfine data of the shallow donors. After a short investigation it became apparent that contributions to the isotropic hyperfine constant due to the terms containing A_2, A_3, and A_4 were of the same order of magnitude as the over-all level of accuracy of the calculation. In the present work, then, we will set these coefficients to zero. In future work where analytic functions are used for the radial basis functions, these terms will have to be reconsidered. A second point which was discovered just from the calculation of the various combinations of the $f_t(\vec{R}_l)$ in Eqs. (18)–(23) is that the terms involving products of d functions are generally much smaller in magnitude than either the $p-p$ or $s-d$ terms. The coefficient $B_3=\langle U_d|1/r^3|U_d\rangle$ is also expected to be small because radial d functions tend to zero at small r at a faster rate than "p" functions.

From a prior knowledge of the matching of the ENDOR shells to shells of lattice nuclei in the previous work,[15] we have a handle on any fitting procedure used to determine our coefficients A_1, B_1, B_2, and B_3. To determine A_1 we use the rms of the fractional deviation of the calculated from the experimental isotropic hyperfine constants

$$\Delta = [N_i^{-1} \sum_i |(a_{\text{calc}} - a_{\text{expt}})/a_{\text{expt}}|^2]^{1/2}. \tag{27}$$

The value of A_1 is chosen by the minimization of Δ for our value of $k_0/k_{\max}=0.854$, which was used in the calculation of our pseudopotential Bloch functions in the previous work[15] and thus in the determination of $A_n(\vec{k})$. Given in Columns 1 and 2 of Table II are the experimental and calculated values of the isotropic hyperfine constants for the value $A_1=|U_s(0)|^2=137.9$. This is to be compared

TABLE II. Comparison of calculated[a] and experimental[b] values of isotropic and anisotropic hyperfine constants.[c]

Shell		Expt $\frac{1}{2}a$	Calc $\frac{1}{2}a$	Expt B_{zz}	Calc B_{zz}	Expt B_{xy}	Calc B_{xy}	Expt[d] B_{xz}	Calc B_{xz}
	As	3860	3722	57.0	38.3	55.8	24.8	0.0	0.0
A-(0, 0, 4)	P	2981	2902	41.4	39.0	41.4	19.1	0.0	0.0
	Sb	3101	2553	46.0	29.3	34.2	16.6	0.0	0.0
	As	3000	2955	−41.6	−12.6	148.2	94.0	−72.2	−56.6
B-(4, 4, 0)	P	2254	2080	−34.0	−16.5	106.2	67.1	−39.8	−28.3
	Sb	1833	1732	−28.6	−16.8	79.2	56.4	−20.2	−18.6
	As	2037	2223	0.0	0.0	5.8	5.0	−5.8	−5.0
C-(3, 3, $\bar{3}$)	P	1649	1570	0.0	0.0	5.0	8.7	−5.0	−8.7
	Sb	1397	1312	0.0	0.0	6.0	9.7	−6.0	−9.7
	As	1292	1272	4.2	−10.4	20.4	15.0	25.4	21.3
D-(3, 3, $\bar{7}$)	P	1117	1021	3.6	−9.1	16.6	14.9	22.2	16.9
	Sb	1003	909	0.0	−8.4	17.4	14.6	17.4	15.0
	As	642	230	0.0	0.0	1258	932	1258	932
E-(1, 1, 1)	P	270	20	0.0	0.0	700	492	700	492
	Sb	293	0	0.0	0.0	522	346	522	346
	As	1121	1010	151.0	149.8	−41.2	−5.0	−4.6	−51.7
F-(3, 3, 1)	P	840	639	116.4	82.7	−28.2	7.2	−11.8	−47.7
	Sb	504	503	37.0	58.8	0.2	11.4	−30.2	−44.1
	As	806	908	7.0	13.5	−1.0	−6.2	5.6	8.2
G-(7, 7, $\bar{3}$)	P	764	803	5.0	11.2	−1.2	−5.9	5.0	6.4
	Sb	761	747	4.6	10.1	−1.2	−5.7	5.2	5.5
	As	801	1160	0.0	0.0	62.4	71.7	−62.4	−71.7
H-(4, 4, $\bar{4}$)	P	689	931	0.0	0.0	50.6	52.7	−50.6	−52.7
	Sb	703	829	0.0	0.0	44.6	45.0	−44.6	−45.0
	As	718	613	−20.4	−18.7	31.8	33.9	12.0	20.9
I-(2, 2, $\bar{8}$)	P	685	564	−17.4	−12.6	27.8	26.0	13.8	18.8
	Sb	643	534	−14.0	−10.1	22.6	22.6	14.4	17.5
	As	694	583	0.0	0.0	5.0	9.4	5.0	9.4
J-(5, 5, 5)	P	739	622	0.0	0.0	5.4	8.1	5.4	8.1
	Sb	761	627	0.0	0.0	···	7.5	···	7.5
	As	758	679	16.0	18.1	7.6	6.9	0.0	0.0
K-(0, 0, 8)	P	663	589	14.0	15.0	6.6	8.7	0.0	0.0
	Sb	629	543	···	13.6	···	5.7	0.0	0.0
	As	741	646	−14.6	−26.7	9.8	4.1	−8.2	−13.7
L-(2, 2, 4)	P	582	496	−11.2	−17.7	2.6	3.9	−7.6	−13.0
	Sb	525	432	−5.8	−14.1	3.8	3.8	−4.4	−12.5
	As	777	741	−60.8	−66.2	65.2	72.2	41.4	46.3
M-(2, 2, $\bar{4}$)	P	612	566	−40.6	−42.8	49.6	49.2	35.6	35.9
	Sb	559	493	−37.2	−34.0	44.0	40.5	30.6	31.3
	As	607	656	0.0	0.0	4.2	−0.4	−4.2	0.4
N-(7, 7, $\bar{7}$)	P	612	640	0.0	0.0	3.4	−0.4	−3.4	0.4
	Sb	629	621	0.0	0.0	···	−0.4	···	0.4
	As	739	945	0.0	0.0	35.8	29.7	35.8	29.7
O-(4, 4, 4)	P	598	764	0.0	0.0	32.6	27.9	32.6	29.7
	Sb	670	681	0.0	0.0	31.4	26.7	31.4	26.7
	As	696	603	−4.8	−2.9	0.4	−1.3	−4.2	−5.9
P-(2, 2, 8)	P	662	555	−6.0	−1.8	···	−1.2	···	−6.0
	Sb	629	526	···	−1.3	···	−1.1	···	−6.0

Table II *(continued)*

Shell		Expt $\frac{1}{2}a$	Calc $\frac{1}{2}a$	Expt B_{zz}	Calc B_{zz}	Expt B_{xy}	Calc B_{xy}	Expt[d] B_{xz}	Calc B_{xz}
	As	566	503	−48.0	−69.6	101.6	132.6	27.6	41.2
Q-(1, 1, 5)	P	524	406	−40.2	−50.8	84.0	100.7	22.8	35.8
	Sb	387	363	−36.0	−43.1	78.6	87.3	31.6	33.0
	As	428	548	24.8	32.1	0.0	−6.0	25.0	15.8
R-(7, 7, 1)	P	379	441	21.0	23.2	0.2	−5.4	19.4	10.6
	Sb	332	392	15.0	19.5	−2.8	−5.1	12.0	8.5
	As	377	420	−5.4	−4.5	9.2	9.9	0.0	3.0
S-(2, 2, $\bar{12}$)	P	410	429	−5.4	−4.4	7.4	8.9	0.0	3.2
	Sb	425	425	...	−4.3	...	8.4	...	3.2
	As	364	411	0.0	−1.2	4.4	1.9	4.4	1.4
T-(2, 2, 12)	P	398	421	3.8	−1.5	3.8	1.5	4.2	1.1
	Sb	425	417	...	−1.6	...	1.3	...	1.0
	As	338	376	...	3.4	...	0.6	...	−1.6
U-(5, 5, 9)	P	383	408	...	3.1	...	0.4	...	−1.4
	Sb	425	413	...	2.9	...	0.3	...	−1.3
	As	242	203	40.0	51.5	30.0	27.2	...	47.8
X-(5, 5, 1)	P	317	290	29.4	34.5	23.6	20.9	32.2	33.9
	Sb	437	325	20.4	28.0	19.0	18.2	20.8	28.4
	As	...	277	...	7.1	...	−0.9	...	−0.7
-(10, 10, 0)	P	...	269	...	4.5	...	0.2	...	−0.8
	Sb	...	260	...	3.4	...	0.8	...	−0.9
	As	...	347	...	−18.1	...	−29.7	...	10.0
-(1, 3, 7)	P	...	308	...	−15.2	...	−26.8	...	8.9
	Sb	...	287	...	−13.8	...	−25.1	...	8.3
	As	...	356	...	1.8	...	−9.6	...	−10.1
-(0, 4, 8)	P	...	315	...	3.0	...	−6.0	...	−6.8
	Sb	...	294	...	3.3	...	−4.6	...	−5.4
	As	...	504	...	−61.7	...	28.6	...	31.7
-(2, 2, 0)	P	...	367	...	−50.2	...	22.2	...	19.1
	Sb	...	311	...	−44.3	...	19.7	...	14.0
	As	...	275	...	−1.7	...	5.0	...	−3.4
-(9, 9, 5)	P	...	296	...	−1.9	...	5.4	...	−3.5
	Sb	...	300	...	−2.0	...	5.4	...	−3.5

[a]All calculated values are for $k_0/k_{\max} = 0.854$, $A_1 = 137.88$, $B_1 = 2.351$, $B_2 = 0.594$, and $B_3 = 0.0$.

[b]Experimental values are from Hale and Mieher, Phys. Rev. 184, 739 (1969).

[c]All values of $\frac{1}{2}a$ and B_{ij} are in kHz.

[d]Only the magnitude of B_{xz} is determined by the experimental work in Ref. 4. The signs of the values listed in this table are considered the most probable in view of the calculated values.

to the value 136.2 which is equivalent to the value $\eta = 178$ which was used in the previous work.[15] The point to be emphasized here is that our $C_{n,s}(\vec{k})$ as determined by our Slater-Koster method are consistent with the previous work in which the $A_n(\vec{k})$ were determined. Given in Columns 3–8 of Table II are the experimental and calculated values of the anisotropic hyperfine constants. The calculated values are obtained from fitting Eqs. (18)–(20) so as to minimize the deviation of the calculated from the experimental values

$$\Delta = \left(N_l^{-1} \sum_l |B_{\text{calc}} - B_{\text{expt}}|^2 \right)^{1/2} . \tag{28}$$

The fractional deviation was not used in this case because there are both positive and negative values of B_{ij}. The use of the fractional deviation would tend to weigh more heavily those numbers which are small in magnitude, possibly due to cancellation effects among the three separate terms given in Eqs. (18)–(20) for each of the B_{ij}. To further avoid the possibility of favoring either very small B's, which might have interference

effects, or very large B values [e.g., the E-(1, 1, 1) shell], we determined the fitted values by considering only those experimental values greater than 5 kHz and less than 400 kHz. The best fit for $k_0/k_{\max} = 0.854$ was found to be

$$A_1 = 137.9, \quad B_1 = 2.351, \quad B_2 = 0.594, \quad B_3 = 0.0 . \tag{29}$$

VI. DISCUSSION OF RESULTS

With the use of the interpolation envelope functions $F_{t,j}(\vec{R}_l)$ and $G_{t,j}(\vec{R}_l)$ of Sec. IV we investigated any possible dependence of the anisotropic hyperfine constants upon the location of the conduction-band minimum k_0. No significant decrease in the fractional deviation as calculated from Eq. (28) was possible. In view of the earlier excellent agreement[15] of the calculated isotropic hyperfine constants and the strain ID parameters with the experimental values, we have again concluded that the value of $k_0/k_{\max} = 0.854$ used in both works is undoubtedly the best choice and no further attempt to vary k_0 was made. Nevertheless, the formulation of the calculation of the $f_t(\vec{R}_l)$ into the $F_{t,j}(\vec{R}_l)$ and $G_{t,j}(\vec{R}_l)$ makes possible interesting and informative comparisons of this work which utilizes full-zone integrations and earlier effective-mass calculations.

A careful study of the anisotropic hyperfine values in Columns 3–8 of Table II shows that our calculated values are in good agreement with experimental values. In a few cases, such as B_{zz} for the D-(3, 3, $\bar{7}$) shell and B_{xy} and B_{xz} for the F-(3, 3, 1) shell, we are off by a fairly large amount. This error is due to the contributions of the B^{pp} and B^{sd} terms being of the opposite sign. In both cases we have reasonable agreement for other components [i.e., B_{xy} and B_{xz} for D-(3, 3, $\bar{7}$) and B_{zz} for F-(3, 3, 1)] for which the contributions of the B^{pp} and B^{sd} are either of the same sign or one of them is much greater in magnitude than the other.

Certainly one of the most satisfying aspects of our results is the explanation of the lack of inversion symmetry of the impurity wave function as measured by the anisotropic hyperfine interactions for the various lattice shells L-(2, 2, 4) and M-(2, 2, $\bar{4}$), I-(2, 2, $\bar{8}$) and P-(2, 2, 8), H-(4, 4, $\bar{4}$) and O-(4, 4, 4), and S-(2, 2, $\bar{12}$) and T-(2, 2, 12). Although previous work[7,8] based upon the EMT did predict that the B_{xy} and B_{xz} are not equal for inversion-related shells, this is the first successful explanation of the different B_{zz} values for such shells. The magnitude of the difference between B_{zz} values for pairs of shells such as I-(2, 2, $\bar{8}$) and P-(2, 2, 8) dramatically illustrate the inability of the EMT to make a clear distinction between inversion-related shells which have rather different pattern styles (as discussed in Ref. 4). This limitation made it impossible for earlier workers to make any shell matchings on the basis of B_{zz} considerations.

The Hale-Mieher work in Ref. 7 used the B_{xy}-(0, 0, 4) value to determine fitting parameters for their calculation of the anisotropic hyperfine constants. However, there was an erroneous choice made of the sign of the experimental value by those workers on the basis of unpublished work by the present workers (this was reported as Ref. 19 in Ref. 7 of this paper). This choice affected all their calculated values of B_{xy} and B_{xz} so that it is impossible to make a meaningful comparison of their published results with the present results.

VII. CONCLUSIONS

In view of the over-all accuracy of the calculation of the isotropic hyperfine constants from the $A_n(\vec{k})$ which were obtained in the previous work,[15] we feel that the agreement of the calculated and experimental values of the anisotropic hyperfine constants is satisfactory for both the confirmation of our matchings of the ENDOR shells to particular shells of lattice nuclei and the understanding of the nature of shallow-donor wave functions in semiconductors. We feel that our work points to the need for a more rigorous calculation of defect wave functions from the point of view of complete Brillouin-zone expansions and not effective-mass calculations. We have found that in general, corrections to a single-band treatment for the shallowest of donors are of the approximate order of 10%. To consider effective-mass treatments[9,19,20] for defects such as the Si(S^+) or other deeper states is out of the question. Computational techniques for these systems will have to be developed along the lines of the present work where improved treatment of the basis functions and the calculation of the $A_n(\vec{k})$ are considered.

We emphasize strongly the fact that the complex nature of the $A_n(\vec{k})$ and of the cellular part of the Bloch functions $\psi_n(\vec{k}, \vec{r})$ [see Eq. (7)] is important for the calculation of all the hyperfine constants for the shallow donors. Also, the use of d basis functions in the representation of the cellular part of the Bloch function is essential for the correct explanation of the anisotropic hyperfine constants. This illustrates the limitations inherent in the use of too small a basis set in calculating electronic wave functions and energy bands.[16,21] The values of the two-center in-

tegrals in Table I illustrate the importance of the nearest-neighbor terms involving d functions relative to any of the possible third-nearest-neighbor interactions which are expected to be still smaller than the second-nearest-neighbor integrals which are shown.

---

*National Academy of Sciences-National Research Council Post-Doctoral Research Associate.

[1]R. C. Fletcher, W. A. Yager, G. L. Pearson, A. N. Holden, W. T. Read, and F. R. Merritt, Phys. Rev. 94, 1392 (1954).

[2]R. C. Fletcher, W. A. Yager, G. L. Pearson, and F. R. Merritt, Phys. Rev. 95, 844 (1954).

[3]G. Feher, Phys. Rev. 114, 1219 (1959).

[4]E. B. Hale and R. L. Mieher, Phys. Rev. 184, 739 (1969).

[5]E. B. Hale and T. G. Castner, Jr., Phys. Rev. B 1, 4736 (1970).

[6]E. B. Hale and R. L. Mieher, Phys. Rev. 184, 751 (1969).

[7]E. B. Hale and R. L. Mieher, Phys. Rev. B 3, 1955 (1971).

[8]T. G. Castner, Jr., Phys. Rev. B 2, 4911 (1970).

[9]T. H. Ning and C. T. Sah, Phys. Rev. B 4, 3468 (1971).

[10]W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5.

[11]B. Mozer (unpublished) as reported in Ref. 3 above.

[12]J. L. Ivey and R. L. Mieher, Bull. Am. Phys. Soc. 15, 1342 (1970).

[13]J. L. Ivey, Ph.D. thesis (Purdue University, 1971) (unpublished).

[14]J. L. Ivey and R. L. Mieher, Phys. Rev. Lett. 29, 176 (1972).

[15]J. L. Ivey and R. L. Mieher, preceding paper, Phys. Rev. B 11, xxx (1975).

[16]J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).

[17]C. E. Pearson and W. S. Williams, Sperry Rand Research Center report, 1965 (unpublished). This computer subroutine computes the minimum value of a function of an arbitrary number of arguments.

[18]The computer subroutine LLSQ is available from the IBM Scientific Subroutine Library.

[19]T. H. Ning and C. T. Sah, Phys. Rev. B 4, 3482 (1971).

[20]S. Pantelides, Ph.D. thesis (University of Illinois, 1972) (unpublished).

[21]G. Dresselhaus and M. S. Dresselhaus, Phys. Rev. 160, 649 (1967).