PHYSICAL REVIEW B VOLUME 23, NUMBER 8 15 APRIL 1981

Bond-bending force constant of silicon by an $\vec{r}$-space method

Evan O. Kane
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 25 August 1980)

We generalize our method of calculating Wannier functions in $\vec{r}$ space to treat the bond-bending force constant, $C_{11} - C_{12}$, in silicon. We find a value 20% too low compared to experiment which we consider satisfactory in view of the simplicity of the method. We find that the Pauling hybrids rotate in the *opposite* direction to the bonds when the crystal is distorted, and we explain this result quantitatively as resulting from the wave-function overlap with nearest neighbors. This counter rotation has a marked symmetry-lowering effect on the charge distribution of a single bond, a smaller symmetry-lowering effect on the bond quadrupole moment, and a very small effect on the crystal quadrupole moment per unit cell.

I. INTRODUCTION AND CONCLUSIONS

This paper is a continuation of a series of papers in which the orthogonality chemical pseudopotential (OCP) has been developed and tested.[1,2,3] We define a "chemical pseudopotential" as any member of a broad class of potentials which leads to a complete set of localized states in a situation where the true potential would give extended states.

The theoretical basis for the OCP was first given by Koster.[4] He showed that the orthogonality condition between different localized states can be used to determine a pseudopotential which has localized solutions. The method is simply to supplement the energy functional which leads to the Schrödinger equation by the orthogonality overlaps multiplied by Lagrange parameters. In a series of papers we have implemented Koster's prescription and showed that it leads to reasonably accurate results with a minimum of effort and in a representation (Pauling bonds) which allows much better contact between theoretical methods and intuitive chemical concepts than do $\vec{k}$-space methods.

We hope to apply this method ultimately to problems such as vacancy and surface reconstruction. In such problems the semiempirical concepts of bond bending and bond stretching have been very useful in giving simple descriptions of the displacements from the perfect crystal coordinates. Consequently, we are very interested to know how well our method predicts the numbers from which the bond-stretching and -bending parameters are derived. In the latest paper of this series[3] (hereafter referred to as I) we showed that the bond-stretching (compressibility) parameter could be derived with reasonable accuracy. In the current paper we study bond bending to complete the picture.

The bond-bending force constant (C_{11}-C_{12}) has recently been calculated by Wendel and Martin[5] using a self-consistent potential similar to ours but employing a $\vec{k}$-space method. Although their self-consistent potential was not nearly as well converged as ours, they obtain better agreement with experiment. We also believe that $\vec{k}$-space methods are better suited for this particular calculation. Our motivation is not to obtain the best value of C_{11}-C_{12} but to test our method which is more easily applied to problems of lower symmetry.[6(a)]

Values of C_{11}-C_{12} have also been predicted by the bond-orbital method of Harrison and coworkers.[6(b)] This method is very much simpler than the present approach and the accuracy is comparable. However, the information content of the two calculations is very different. We find detailed charge distributions for the bond charge and how it changes under strain. We evaluate and correct for the electron-electron interaction which is double counted in the energy eigenvalues. The ion-ion interaction is calculated and makes an important contribution to C_{11}-C_{12}. These quantities are assumed to be unimportant in the bond-orbital method, but no estimate of their size is given. If calculations of the bond-orbital type are really as accurate as comparison with experiment suggests, there is a real need to understand the reasons for their success at a more fundamental level. The present calculation is a step in this direction, although quite far from providing a satisfactory explanation.

In Sec. II we review our method for calculating Wannier functions (Pauling bonds) and indicate the extensions required to apply it to the case of a tetragonally distorted crystal. In Sec. III we discuss the results in detail. We find a bond-bending force constant (C_{11}-C_{12}), which is 20% lower than experiment. In paper I we found a bond-stretching force constant (compressibility) which was 30% too low for the same level of approximation. We showed there that orthogonality corrections and the addition of antibonding functions

reduced the discrepancy to 15%. These corrections have not been made in the current work.

We investigated the response of the hybrid wave functions to the crystal distortion. Simple Pauling theory (intra-atomic constraints only) implies that the hybrids would retain their tetrahedral angles.[7] Naive intuition suggests that the hybrids would try to compromise by rotating partway from the tetrahedral angles toward the bond directions. We find just the opposite. The hybrids rotate in the *opposite* direction by an amount roughly 0.4 of the bond rotation angle. We derive a quantitative explanation for this effect based on wave-function overlaps between first and second neighbors and the requirement that neighboring bonds remain orthogonal. We find that counter-rotation is not peculiar to silicon but is always to be expected. The magnitude of the counter-rotation goes to zero with the overlap to yield the Pauling result.

The hybrid rotation lowers the symmetry of the individual bonds from axial with inversion to C_{2h}. We find that the symmetry lowering has a marked effect on the charge distribution. We also calculate the quadrupole moment of the charge distribution. We find that the principal axis of the quadrupole tensor rotates with the bonds by 0.84 of the bond rotation angle. The other two axes become nondegenerate. When the quadrupole moment of the crystal unit cell is calculated, the symmetry-breaking effects nearly cancel. A model rotating the axial quadrupole moment of the cubic case with the bonds would give a good approximation to the moment of the unit cell.

II. $\vec{r}$-SPACE METHOD FOR THE TOTAL ENERGY

A. Bond Wannier functions

We follow the procedure described in paper I very closely and will omit many details that are presented there. We define bonding Wannier functions

$$b(\vec{r}) = [\phi(\vec{r}_1) + \phi(\vec{r}_2)]/\sqrt{2(1+S)} \tag{1}$$

$$\phi(\vec{r}) = \sum_i c_i \phi_i(\vec{r}),$$

where ϕ_i is an s-, p-, or d-type Gaussian given explicitly by

$$\phi_s = e^{-\alpha r^2},\ \phi_{px} = xe^{-\alpha r^2},\ \phi_{dx} = yze^{-\alpha r^2},$$

$$S = \int \phi^*(\vec{r}_1)\,\phi(\vec{r}_2)\,d\vec{r}, \tag{2}$$

$$\vec{r}_1 = \vec{r} - \vec{R}_1,\ \vec{r}_2 = \vec{r} - \vec{R}_2,$$

where $\vec{R}_1$ and $\vec{R}_2$ are the coordinates of nearest-neighbor atoms.

We impose a tetragonal distortion on the diamond lattice so that the cube axes are stretched or compressed according to

$$a_x = a_y = as_x,\ a_z = as_z$$
$$2s_x^2 + s_z^2 = 3. \tag{3}$$

We have used $s_x = 0.95$ in the computations. This corresponds to a rotation of the bonds in the $(1\bar{1}0)$ or (110) planes by 0.0659 radians toward the z axis. The constraint on s_x, s_z ensures that the nearest-neighbor distance is unchanged. This constraint is important since the bond-stretching potential minimum is not exactly at the equilibrium lattice constant for the model we are using.

For a bond between atoms at $\vec{R}_1 = 0$, $\vec{R}_2 = \frac{1}{4}(a_x, a_y, a_z)$ we have used as independent functions

$$\phi_s,\ \phi_{p(x+y)} = \frac{(x+y)}{\sqrt{2}}\,e^{-\alpha r^2},\ \phi_{pz} = ze^{-\alpha r^2}$$
$$\phi_{d(x+y)} = \frac{(yz+xz)}{\sqrt{2}}\,e^{-\alpha r^2},\ \phi_{dz} = xye^{-\alpha r^2}. \tag{4}$$

We now have two independent p and d functions per α where in the undistorted lattice we had only one, e.g., $\phi_p = (x+y+z)e^{-\alpha r^2}/\sqrt{3}$. The p functions are multiplied by the phase $+1(-1)$ on atom 1(2). All bonds are equivalent but pairs of bonds fall into two inequivalent sets of which b_1b_2 and b_1b_3 are representatives:

$$\hat{\boldsymbol{b}}_1 = (a_x, a_y, a_z)/a;\ \hat{\boldsymbol{b}}_2 = (-a_x, -a_y, a_z)/a;$$
$$\hat{\boldsymbol{b}}_3 = (a_x, -a_y, -a_z)/a. \tag{5}$$

We have used 3 different values of α for each s, p, and d function, namely $\alpha = 0.2$, 0.4, and 0.6 in atomic units, for a total of 15 independent functions. We have determined the coefficients c_i in Eq. (1) by minimizing the Hamiltonian energy $\langle b_1|H|b_1\rangle$ plus the orthogonality constraints $\lambda_2\langle b_1|b_2\rangle + \lambda_3\langle b_1|b_3\rangle$, the latter terms constituting the chemical pseudopotential. We use a conventional pseudopotential in describing H (see I). In the computation of λ_2, λ_3 we have used a starting guess based on the value of λ of the undistorted crystal and find rapid convergence using two-dimensional linear extrapolation of λ_2, λ_3 versus the overlaps $\langle b_1|b_2\rangle$, $\langle b_1|b_3\rangle$ which become zero for the correct values of λ_2, λ_3.

B. Charge density and potential

Having determined the bond wave functions as just described using the orthogonality chemical pseudopotential, we compute the charge density by summing over all lattice sites in the fcc lattice out to third neighbors [i.e., including $(2,1,1)a/2$] with four bonds per site. The charge density

is determined in this manner for a grid of ~130 inequivalent points in the tetragonally distorted unit cell of the diamond lattice. The charge density on these points is then least-squares fit to four independent sets of spherical Gaussians centered on lattice sites, bond sites, lattice sites $+\{(a_x/8, 0, 0)\}$ and lattice sites $+\{(0, 0, a_z/8)\}$. The symmetry equivalent sets $\{\}$ are $\{(a_x/8, 0, 0)\} = (\pm a_x/8, 0, 0), (0, \pm a_y/8, 0)$, and $\{0, 0, a_z/8\} = (0, 0, \pm a_z/8)$. Three Gaussians per site whose exponential coefficients were $\alpha = 0.4, 0.8, 1.2$ in atomic units were used for a total of twelve fitting parameters. The rms error was 0.5% of the bond-site charge density. Use of a Lagrange parameter ensured that the total charge was accurately four electrons per atom. The Gaussian charge then gave the potential in terms of error functions. The potential was summed over neutral units associated with each lattice site and lattice sites were summed out to 8th neighbors. Each neutral unit possesses a quadrupole moment in the distorted lattice but in spite of this, convergence was rapid. We estimate the truncation error to be of the order of 0.001 eV.

The above procedure yields the potential on ~130 points in the unit cell which was then least-squares fit to a sum of 6 Gaussians on lattice and bond sites plus a constant for a total of 13 fitting parameters. The rms error was 0.07 eV. The Gaussian exponential coefficients used were 0.2, 0.4, 0.6, 0.8, 1.2, and 1.6 in atomic units.

The bare pseudopotential was fit to the potential of a Gaussian smeared charge density with $Z = 4$ and Gaussian coefficient $\alpha = 1$ a.u. plus five potential Gaussians with exponential coefficients 0.4, 0.6, 0.8, 1.2, and 1.6. This was the bare Berkeley pseudopotential, exactly the same as in I. The exchange and correlation were treated as in I using the Slater $X\alpha$ local density potential with $X\alpha = \frac{2}{3}$. The potential was then iterated to self-consistency which required six iterations starting from the self-consistent unstrained potential and using $s_x = 0.95$ in

$$a_x = a_y = a s_x, \quad a_z = a s_z$$
$$2 s_x^2 + s_z^2 = 3 . \tag{5a}$$

The bond-energy change decreased by about a factor of three with each iteration, oscillating above and below the asymptotic value. The estimated error in the bond energy was 0.001 eV with equal or better accuracy in the correlation energy.

C. Ion-ion and electron-electron interactions

In calculating the total crystalline energy we need the ion-ion and electron-electron interaction energies in addition to the one-electron-energy eigenvalues (given by the bond energy). We treat these in exactly the same manner as in I. The total energy is written

$$E_{\text{tot}} = E_{II} + E_{eeC} + E_{eeX} + T_e + E_{eISC} + E_{eIG} , \tag{6}$$

where E_{II} is the ion-ion energy, E_{eeC} the electron-electron Coulomb energy, E_{eeX} is the electron exchange and correlation energy, and T_e is the electronic kinetic energy. The electronic interaction with the core pseudopotential is divided into two pieces, E_{eISC} the interaction with the smeared Coulomb ionic charge, and E_{eIG}, the interaction with the Gaussian part.

The one-electron eigenvalue energy summed over all occupied states ($= 4E_{\text{bond}}$ per atom) is given by

$$4E_{\text{bond}} = T_e + E_{eISC} + E_{eIG} + 2E_{eeC} + \tfrac{3}{4} E_{eeX} \tag{7}$$

when exchange and correlation are treated by Slater's $X\alpha$ local density method. In an infinite crystal E_{II}, E_{eeC}, and E_{eISC} are infinite. We eliminate the infinity in favor of an indeterminacy by defining

$$E'_{II} = E_{II} + \tfrac{1}{2} E_{ISC}$$
$$E'_{eeC} = E_{eeC} + \tfrac{1}{2} E_{ISC} . \tag{8}$$

The total energy may then be written

$$E_{\text{tot}} = 4E_{\text{bond}} - E'_{eeC} + E'_{II} - \tfrac{1}{3} E_{eeX} . \tag{9}$$

The quantities in Eqs. (8) may be expressed in terms of ionic and electronic charges times a potential. The potential is highly convergent when summed over symmetric neutral units but is indeterminate to within a constant which depends on how the neutral units are defined. The constant cancels in the expression for the total energy.

In this paper we use a different potential zero from that adopted in I. We define a "Coulomb" electronic potential energy by

$$V_{NC}(\vec{r}) = V_{eC}(\vec{r}) + V_{ISC}(\vec{r}) + \text{const} \tag{10}$$

with the constant defined such that $V_{NC}(\vec{r}_B) = 0$ where $\vec{r}_B$ is the bond site. V_{eC} is the potential energy of one electron interacting with the other electrons and $V_{ISC}(\vec{r})$ is the energy of interaction of an electron with the Gaussian smeared ionic charge. In paper I we chose the constant such that the average value of $V_{NC}(\vec{r})$ was zero.

III. RESULTS

A. Bond-bending energy

Following the method of Sec. II we have introduced a tetragonal distortion into the cubic silicon lattice by stretching the axes $a_x = a_y = a s_x$, a_z

$= as_z$, where a is the perfect crystal-lattice constant. We have used the constraint $2s_x^2 + s_z^2 = 3$ because this keeps the nearest-neighbor distance constant. In our model (see I), the bond-stretching potential is not accurately a minimum for the equilibrium lattice constant, hence bond stretching would introduce a term proportional to $(s_x - 1)^2$ in U. This term is eliminated by the above constraint which only introduces corrections of order $(s_x - 1)^3$. We have estimated these terms along with those from the third-order elastic constants and found them to be negligible.

The energy increase per atom, U, due to the distortion is given to second order in $(s_x - 1)$ by

$$U = \tfrac{3}{8}(C_{11} - C_{12})a^3(s_x - 1)^2 = 0.0954 \text{ eV/atom} \qquad (11)$$

for $s_x = 0.95$. C_{11} and C_{12} are elastic constants as tabulated by R. M. Martin.[8] Comparing this experimental value to our calculated value, 0.0764 eV, we see that our value is 20% too low. The prinicipal errors in our work are probably due to the neglect of overlap of bonds beyond first neighbors and neglect of antibonding corrections. Some error may also be due to the representation of the charge density by spherical Gaussians and the Gaussian representation of the potential. These errors could easily be reduced by using more centers and hence more fitting parameters.

It is interesting to note the breakdown of the distortion energy, E_{tot}, into the individual components as given in Table I. The total energy increase is almost exactly the increase in the kinetic energy. The Slater exchange contribution is small compared to the electron-electron and ion-ion contributions, but still a significant part of the total energy. The contribution from the Gaussian core potential E_{eIG} is very small because the valence charge at the nucleus is small and relatively constant. The numbers assigned to E_{bond}, E'_{eeC}, and E'_{II} are strongly dependent on the zero of potential although E_{tot} is independent of the zero. We have fixed the potential by requiring the Coulomb potential [V_{NC} in Eq. (10)] to be zero at the bond site. This choice has the advantage of making most of the difference energies comparable.

TABLE I. Decomposition of the total energy for cubic and tetragonally distorted silicon. The symbols are those of Eqs. (6) and (8). All energies in eV/atom. The bonds have been rotated by 0.0659 radians toward the z axis in the $(1\bar{1}0)$ plane.

		$0.95a$, $0.95a$, $1.09316a$	
	Cubic	Tetrag. dist.	Tetrag.-cubic
T_e	40.6872	40.7595	0.0723
$4E_{\text{bond}}$	4.2905	4.5940	0.3035
E'_{eeC}	−2.9165	−2.7891	0.1274
E_{eeX}	−27.4736	−27.4931	−0.0194
E'_{II}	−118.2601	−118.3660	−0.1059
E_{eIG}	6.0678	6.0701	0.0023
E_{tot}	−101.8943	−101.8179	0.0764

B. Tight-binding interpretation

A great deal of effort has been based on tight-binding methods using overlap integrals of atomic-like wave functions, either taken as empirical parameters or computed from atomic wave functions. Our method is closely related in spirit to these methods, hence it is interesting to calculate the overlap integrals using our self-consistent potential and wave functions.

The results for overlap integrals of the wave function and of the Hamiltonian are given in Table II. The definitions of the matrix elements are given in Fig. 1. As expected, it is seen that the wave-function overlaps are quite large, especially $\langle p_\sigma^0 | p_\sigma^1 \rangle = 0.45$. In treatments where the overlaps are treated empirically, the wave-function overlaps are usually ignored because of the large increase in fitting parameters required to treat them. We see that this would be a very bad approximation and it is therefore hard to compare our energy overlaps with the empirically determined values. A related indeterminacy is that the H overlaps depend on the zero of potential when the S overlaps are nonzero. ($\Delta H_{ij} = \Delta V S_{ij}$ when the potential is shifted by a constant ΔV.)

TABLE II. Electron energy and wave function overlap integrals between zeroth and first neighbors for the self-consistent potential. Second-neighbor overlaps from the potential of Ref. 1. Energies in electron volts.

	S_{cubic}	H_{cubic}	S_{tetrag}	H_{tetrag}
$\langle s^0 \| s^0 \rangle$	1	−0.8337	1	−0.7858
$\langle p^0 \| p^0 \rangle$	1	+4.4423	1	4.5123
$\langle s^0 \| s^1 \rangle$	0.1386	−1.7673	0.1387	−1.7611
$\langle s^0 \| p_\sigma^1 \rangle$	0.2956	−1.8229	0.2948	−1.8081
$\langle p_\sigma^0 \| p_\sigma^1 \rangle$	0.4493	0.6427	0.4476	0.6457
$\langle p_{\pi 2}^0 \| p_{\pi 2}^1 \rangle$	0.1525	−0.2765	0.1533	−0.2458
$\langle p_{\pi 1}^0 \| p_{\pi 1}^1 \rangle$	0.1525	−0.2765	0.1518	−0.3111
$\langle s^0 \| s^2 \rangle$	0.0007			
$\langle s^0 \| p_\sigma^2 \rangle$	0.0139			
$\langle p_\sigma^0 \| p_\sigma^2 \rangle$	0.0701			
$\langle p_\pi^0 \| p_\pi^2 \rangle$	0.0079			

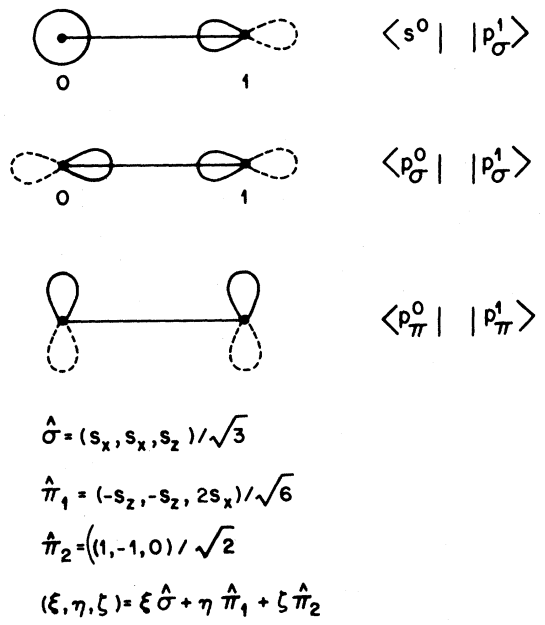

FIG. 1. Definitions of overlap integrals. For nearest neighbors the perpendicular directions are specified. For second neighbors we average over perpendicular directions. s_x, s_z describe the tetragonal distortion as defined in Eq. (5a). The directions $\hat{\sigma}$, $\hat{\pi}_1$, $\hat{\pi}_2$ are also used to define the bond coordinates ξ, η, and ζ.

It is seen that the overlaps change by only a small amount as the crystal is distorted. This is particularly true of the S overlaps. An interesting question is how the hybrid wave functions respond when the crystal is distorted. If first- and second-neighbor overlaps are neglected the bonding hybrids are fully occupied. The only degree of freedom would be the rotation of the tetrahedron formed by the four hybrids. Such a rotation is incompatible with the (001) tetragonal distortion we have assumed which keeps all hybrids equivalent. The zeroth-order expectation would be that the hybrids would remain rigidly tetrahedral as the atoms shifted.[6] Intuitively, we would expect that the hybrids would try to follow the directions connecting nearest-neighbor atoms. Surprisingly, we find that just the opposite occurs.

When we consider first- and second-neighbor overlaps, the problem posed above can be answered accurately if we use the orthogonality requirement between nearest-neighbor bonds as described in Sec. II. We have defined the overlap integrals in the following manner. For convenience, we redefine the atomic wave function $\phi(\vec{r})$ in Eq. (1) as follows

$$\phi(\vec{r}) = \sum_{i,j} c_{ij}\,\phi_{ij}(\vec{r}), \tag{11}$$

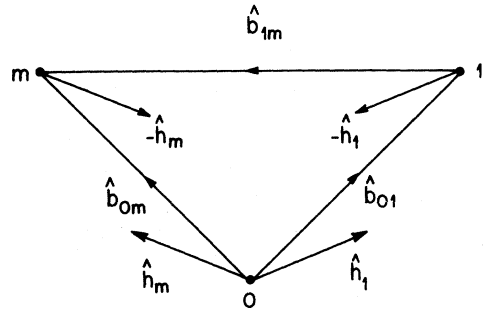

FIG. 2. Hybrids on adjacent bonds with a (001) tetragonal distortion. Atoms at $(a/4)$ (s_x, s_x, s_z), $(a/4)$ $(-s_x, -s_x, s_z)$, $(a/4)(s_x, -s_x, -s_z)$, $(a/4)(-s_x, s_x, -s_z)$ are labeled 1 to 4. The pairs (1,2) and (1,3) are inequivalent. Inversion symmetry yields the $\hat{h}_m$, $-\hat{h}_m$ relation.

where i refers to S, X, Y, and Z, and j refers to the three values of the exponent α for s. For X, j includes the three α's for x and the three α's for the cotransforming d function yz as well. Thus X, Y, Z are "effective" p functions.

We then rewrite Eq. (11) as

$$\phi(\vec{r}) = \sum_i C_i F_i(\vec{r})$$

$$\langle F_i | F_j \rangle = \delta_{ij}$$

$$C_i^2 = \sum_{j,k} c_{ij} c_{ik} \langle \phi_{ij} | \phi_{ik} \rangle \tag{12}$$

$$C_p^2 = C_x^2 + C_y^2 + C_z^2 .$$

The overlap integrals are then defined in terms of these "effective s and p" functions in the usual way as detailed in Fig. 1. Our definitions are chosen to make all overlap integrals positive. The results are given in Table II using the self-consistent potential except for second-neighbor integrals where the empirical potential in Ref. 1 was used. These latter values may be in error by ~20%. They make a very small contribution to the hybrid angle.

We can easily calculate the overlap between nearest-neighbor bonds in terms of the overlap integrals in Table II. The hybrid configuration on neighboring bonds is shown in Fig. 2. 180° rotation symmetry about x, y, and z leads to the invariance of any scalar expression when the pair indices l, m are interchanged. Using this fact, the overlap between bonds b_l and b_m may be written

$$\langle b_l | b_m \rangle = C_S^2(\langle s^0|s^0\rangle + 2\langle s^0|s^1\rangle + \langle s^0|s^2\rangle) + C_s C_p[2\langle s^0|p_\sigma^1\rangle \hat{h}_l \cdot (\hat{b}_{0l} + \hat{b}_{0m}) + \langle s^0|p_\sigma^2\rangle(\hat{h}_m - \hat{h}_l)\cdot \hat{b}_{lm}]$$
$$+ C_p^2\{\langle p^0|p^0\rangle \hat{h}_l \cdot \hat{h}_m + 2\langle p_\sigma^0|p_\sigma^1\rangle(\hat{h}_l \cdot \hat{b}_{0l})(\hat{h}_m \cdot \hat{b}_{0l}) - \langle p_\sigma^0|p_\sigma^2\rangle(\hat{h}_l \cdot \hat{b}_{lm})(\hat{h}_m \cdot \hat{b}_{lm})$$
$$+ 2\langle p_\pi^0|p_\pi^1\rangle[-\hat{h}_l \cdot \hat{h}_m + (\hat{h}_l \cdot \hat{b}_{0l})(\hat{h}_m \cdot \hat{b}_{0l})] + \langle p_\pi^0|p_\pi^2\rangle[\hat{h}_l \cdot \hat{h}_m - (\hat{h}_l \cdot \hat{b}_{lm})(\hat{h}_m \cdot \hat{b}_{lm})]\}\,. \tag{13}$$

We have averaged over perpendicular components in deriving Eq. (13). The symbol $\hat{v}$ denotes a unit vector. The orthogonality constraints are then $\langle b_1|b_2\rangle = 0$, $\langle b_1|b_3\rangle = 0$, all other pairs are equivalent to one of these. The constraints are easily solved if we linearize the equations about the tetrahedral direction $\hat{h}_l^0 = \hat{b}_{0l}^0 = (1,1,1)/\sqrt{3}$. We write

$$\hat{h}_l = (1-\delta, 1-\delta, 1+2\delta)/\sqrt{3}$$
$$\hat{b}_{0l} = (1-\epsilon, 1-\epsilon, 1+2\epsilon)/\sqrt{3}\,. \tag{14}$$

$\hat{h}_m$, $\hat{b}_{0m}$ is found by applying the symmetry operator T^m:

$$T_{ij}^m = T_i^m \delta_{ij};\ m = 1, 4$$
$$\{T^m\} = \{(111),\ (1\bar{1}\bar{1}),\ (\bar{1}1\bar{1}),\ (\bar{1}\bar{1}1)\}\,, \tag{14a}$$

e.g., $v_4 = T^4 v_1 = (-v_{x1}, -v_{y1}, v_{z1})$. We define

$$A = C_S^2(1 + 2\langle s^0|s^1\rangle + \langle s^0|s^2\rangle) + C_s C_p\left(\tfrac{4}{3}|\langle s^0|p_\sigma^1\rangle + 2\left(\tfrac{2}{3}\right)^{1/2}\langle s^0|p_\sigma^2\rangle\right)$$
$$+ \frac{C_p^2}{3}\left(-1 + \langle p_\pi^0|p_\pi^2\rangle - 2\langle p_\sigma^0|p_\sigma^1\rangle + 2\langle p_\sigma^0|p_\sigma^2\rangle\right) \tag{15}$$
$$B = \tfrac{8}{3}(\epsilon N + \delta D)$$
$$N = C_s C_p\langle s_0|p_\sigma^1\rangle + C_p^2(\langle p_\sigma^0|p_\sigma^1\rangle + \langle p_\pi^0|p_\pi^1\rangle)$$
$$D = C_s C_p\left(\langle s^0|p_\sigma^1\rangle - \frac{\sqrt{3}}{2\sqrt{2}}\langle s^0|p_\sigma^2\rangle\right) + C_p^2(1 + \langle p_\sigma^0|p_\sigma^1\rangle - \langle p_\pi^0|p_\pi^1\rangle - \tfrac{1}{2}\langle p_\sigma^0|p_\sigma^2\rangle + \tfrac{1}{2}\langle p_\pi^0|p_\pi^2\rangle)\,. \tag{16}$$

We then find

$$\langle b_1|b_2\rangle = A + B;\ \langle b_1|b_3\rangle = A - B/2,$$
$$\text{therefore, } A = 0,\ B = 0;\ \delta = -\epsilon N/D\,. \tag{17}$$

The constraint $A = 0$ yields a quadratic equation for the ratio C_p/C_s. We give the values for δ and C_p/C_s as found from Eq. (17) in Table III. Of course, the same quantities can be computed directly from the values of C_s, C_x, C_y, C_z, C_p, defined in Eq. (12), e.g., $\hat{h}_{lx} = C_x/C_p$. These values are also given in Table III. The variation in δ comes partly from linearization but is mainly due to our definition of "effective p functions" in Eq. (12).

TABLE III. Hybrid parameters as given by Eqs. (12) and (17). In Pauling's theory the ratio of p to s would be $C_p/C_s = \sqrt{3}$. The nearest-neighbor directions are rotated by $\sqrt{2}\epsilon$ toward the z axis by the tetragonal distortion. The hybrid responds by rotating $\sqrt{2}\delta$ which is seen to be in the opposite direction.

δ	C_p/C_s	
-0.536ϵ	1.881	Eq. (17)
-0.400ϵ	1.872	Eq. (12)
	1.860	cubic Eq. (12)

We note the curious result that $\delta \sim -\epsilon/2$. The hybrids rotate in the direction opposite to the bonds. Equations (16) and (17) show that this is not the result of an exceptional situation but will always be the case. All quantities in N and D are positive except some of the second-neighbor interactions and $\langle p_\pi^0|p_\pi^1\rangle$ which is also small. In the limit that overlaps are small we recover the Pauling answers $\delta = 0$, $C_p/C_s = \sqrt{3}$.

We have been arguing as if the overlaps in Table II were fixed quantities. Actually they are determined in the process of minimizing the energy of the Hamiltonian subject to the orthogonality constraints. It might be thought that different overlaps would result from the desire of the hybrid to "follow" the bond. This is not the case. The overlaps change very little from cubic to tetragonal. Equation (17) shows that this is because very drastic changes in overlaps would be required to allow a more energetically favorable hybrid orientation.

C. Charge distribution and quadrupole moment

In order to study the effects produced by the counter-rotated hybrids it is easiest to study

the charge distribution of two electrons in a single bond. We take a coordinate system with axes parallel and perpendicular to the bond as shown in Fig. 1. The charge density along various lines parallel and perpendicular to the bond is given in Table IV. Column $\rho_1(x)$ is the charge density in the unstrained crystal along the bond axis (line connecting nearest neighbors). Column $\rho_2(x)$ gives the change in the charge density along the bond axis when the bonds are rotated by 0.066 radians toward the cubic z axis [$s_x = 0.95$ in Eq. (3)]. As we have seen, the hybrid at the origin rotates by 0.026 radians in the opposite direction. The charge density along the bond axis is changed relatively little, the charge at the bond site being reduced about 0.4% by the reduction of the hybrid-hybrid overlap. (The effect is partly compensated by the requirement of normalization.)

The rotation of the hybrid also lowers the symmetry of the charge distribution which is axial when the hybrids point along the bond direction. The charge distribution retains inversion symmetry about the bond site and has reflection symmetry in the $\hat{\xi}$ plane. This leads to the symmetry relations $\rho(\xi,\eta,\zeta) = \rho(-\xi,-\eta,-\zeta)$; $\rho(\xi,\eta,\zeta) = \rho(\xi,\eta,-\zeta)$. The asymmetry between the lines $(\xi, c, 0)$ and $(\xi, -c, 0)$ is a maximum for $c = 0.3(\sqrt{3}a/8)$. The difference in the charge along these two directions is given in the column headed ρ_6 while the cubic charge along $\rho(\xi, c, 0) = \rho(\xi, -c, 0)$ is given in column ρ_3. The charge differences are quite large relative to the cubic charge density, becoming as great as 5% of the bond-site charge density. The sign reverses for $x > 1.2(\sqrt{3}a/8)$ which is reasonable since the displacement associated with the hybrid rotation changes.

The inequivalence of the $\hat{\eta}$ and $\hat{\zeta}$ planes is indicated by columns ρ_8 and ρ_9 where the charge is tabulated along lines passing through the bond site perpendicular to the bond. Columns ρ_4 and ρ_5 also show the $\hat{\eta}, \hat{\zeta}$ plane symmetry breaking. The charge differences in ρ_4, ρ_5, ρ_8, and ρ_9 are small compared to ρ_6.

a. Fitting the charge distribution. We have fit the single-bond-charge distribution to a sum of spherical Gaussian charges about 11 sites, 6 of which are independent. The site coordinates and the charge per site are given in Table V. Sites 5 and 6 were taken to lie roughly along the hybrid direction, distant by $\pm 0.1\,a$ from either atom.

TABLE IV. Charge density of a single bond along various lines for unstrained (cubic) and strained (tetragonal) cases. The coordinates in $\rho(\xi,\eta,\zeta)$ are the bond coordinates defined in Fig. 1. The bond site is at the origin and the atoms are at $(\pm 1, 0, 0)$. The charge density is in atomic units, electrons per Bohr radius cubed. Two electrons are contained in the bond.

x	$\rho_1(x)$	$\rho_2(x)$	$\rho_3(x)$	$\rho_4(x)$	$\rho_5(x)$	$\rho_6(x)$
0	0.095 34	0.000 35	0.045 39	−0.000 38	+0.000 72	0
0.2	0.084 95	0.000 30	0.042 12	−0.000 38	+0.000 67	0.002 90
0.4	0.058 43	0.000 01	0.033 07	−0.000 36	+0.000 57	0.004 56
0.6	0.029 00	0.000 08	0.021 04	−0.000 34	+0.000 43	0.004 62
0.8	0.009 51	0.000 02	0.010 26	−0.000 28	+0.000 27	0.003 58
1.0	0.001 66	0.000 00	0.003 37	−0.000 17	+0.000 11	0.002 12
1.2	0.000 01	0.000 00	0.000 39	−0.000 06	−0.000 03	0.000 68
1.4	0.000 52	0.000 00	0.000 14	+0.000 03	−0.000 10	−0.000 44
1.6	0.001 86	0.000 01	0.001 23	+0.000 08	−0.000 11	−0.001 09
1.8	0.003 40	0.000 03	0.002 39	+0.000 10	−0.000 10	−0.001 28
2.0	0.004 26	0.000 04	0.002 78	+0.000 10	−0.000 07	−0.001 13
	$\rho_7(x)$	$\rho_8(x)$	$\rho_9(x)$			
0	0.095 34	0.000 35	0.000 35			
0.2	0.087 98	0.000 22	0.000 43			
0.4	0.068 85	−0.000 10	0.000 60			
0.6	0.045 39	−0.000 38	0.000 72			
0.8	0.025 20	−0.000 50	0.000 70			
1.0	0.011 89	−0.000 45	0.000 53			

$\rho_1(x) = \rho_{\text{cubic}}(x, 0, 0)$; $\rho_2(x) = \rho_{\text{cubic}}(x, 0, 0) - \rho_{\text{tetrag}}(x, 0, 0)$
$\rho_3(x) = \rho_{\text{cubic}}(x, 0.3, 0)$; $\rho_4(x) = \rho_{\text{cubic}}(x, 0.3, 0) - \rho_{\text{tetrag}}(x, 0, 0.3)$
$\rho_5(x) = \rho_{\text{cubic}}(x, 0.3, 0) - [\rho_{\text{tetrag}}(x, 0.3, 0) + \rho_{\text{tetrag}}(x, -0.3, 0)]/2$
$\rho_6(x) = \rho_{\text{tetrag}}(x, 0.3, 0) - (x, -0.3, 0)$; $\rho_7(x) = \rho_{\text{cubic}}(0, x, 0)$
$\rho_8(x) = \rho_{\text{cubic}}(0, x, 0) - \rho_{\text{tetrag}}(0, 0, x)$;
$\rho_9(x) = \rho_{\text{cubic}}(0, x, 0) - \rho_{\text{tetrag}}(0, x, 0)$

TABLE V. Fitting charges resulting from fitting the single-bond charge distribution with spherical Gaussian charges, Gaussian exponents 0.4, 0.8, 1.2. $q_2 + 2(q_1 + q_3 + q_4 + q_5 + q_6) = 2$ electrons. The coordinates in the table are those of Fig. 1.

	q_1	q_2	q_3	q_4	q_5	q_6
Cubic	−0.7852	3.4353	−0.3772	−0.3772	0.4226	0.3994
Tetragonal	−1.1124	−0.4025	0.0983	0.2000	1.5506	0.4647

Sites 1: $a(\pm\sqrt{3}/8, 0, 0)$ atom sites
Sites 2: $(0, 0, 0)$ bond site
Sites 3: $a(0, \pm 0.1, 0)$
Sites 4: $a(0, 0, \pm 0.1)$
Sites 5: $a(\pm 0.116\,51, \pm 0.01, 0)$ inner hybrid sites
Sites 6: $a(\pm 0.316\,51, \mp 0.01, 0)$ outer hybrid sites

Three Gaussians per site were used with exponential coefficients $\alpha = 0.4$, 0.8, 1.2 atomic units.

The fitting charges vary greatly between the cubic and tetragonal cases, far more than suggested by the changes in the charge distribution given in Table IV. This clearly indicates a large amount of linear dependence between the different sites. (The charge distributions have large overlap.) Nevertheless, the fit is quite accurate. The rms error is 0.000 10 (0.000 43) for the cubic (tetragonal) cases.

b. Quadrupole moment. Although the fitting charges are relatively meaningless because they vary so much between the cubic and tetragonal cases, they are useful for calculating the quadrupole moments. The traceless quadrupole moment tensor is defined by

$$Q_{ij} = \sum_{n=1}^{6} \left(\xi_i^n \xi_j^n - \frac{1}{3} \sum_{k=1}^{3} \xi_k^n \xi_k^n \right) (-q_n + \delta_{n1}) w_n \,, \tag{18}$$

where w_n are the number of equivalent atoms per site (which contribute equally to Q_{ij}). The term δ_{n1} represents the unit positive charge on the atoms. (Each bond is assigned one unit of positive charge on each atom to give charge neutrality.) We use the bond coordinates of Fig. 1.

The axial symmetry of the cubic case guarantees that the quadrupole tensor is diagonal in the bond coordinates with $Q_{22} = Q_{33}$. In the tetragonal case Q_{12} is finite, hence to diagonalize the tensor we must rotate the tensor axis by an angle θ from 1 to 2 where

$$\tan 2\theta = \frac{2Q_{12}}{Q_{11} - Q_{22}}$$

$$Q_{1'1'} = Q_{11}\cos^2\theta + 2Q_{12}\sin\theta\cos\theta + Q_{22}\sin^2\theta \tag{19}$$

$$Q_{2'2'} = Q_{11}\sin^2\theta - 2Q_{12}\sin\theta\cos\theta + Q_{22}\cos^2\theta \,.$$

$Q_{1'1'}$, $Q_{2'2'}$, Q_{33} are the diagonal components of Q in the rotated frame. The new axis $\hat{1}'$ is given in cubic coordinates by

$$\hat{1}' = (x_q, x_q, z_q) \tag{20}$$

and the axes of the four equivalent tensors are given by the symmetry transformations T^m in Eq. (14a). Summing over the four equivalent quadrupole tensors to give Q^c, the total quadrupole moment of the crystal per unit cell, it is easy to show that

TABLE VI. Quadrupole moment tensor for single bonds and for the unit cell (4 bonds). Numerical subscripts refer to the bond coordinate axes in Fig. 1. [The tetragonal principal axes are rotated about the (1, −1, 0) direction by θ. Positive θ is a rotation toward the cubic z axis.] Units are $|e|a^2$ where a is the lattice constant.

Q_{11}	Q_{22}	Q_{33}	Q_{12}	
0.045 55	−0.022 78	−0.022 78	0	Cubic
0.044 01	−0.021 19	−0.022 82	−0.000 67	Tetragonal
0.044 02	−0.021 20	−0.022 82	0	Tetragonal principal axes

$\theta = -0.0103$; $-2(s_z - 1)Q_{22} = 4.2 \times 10^{-3}$; $-2\sqrt{2}\theta Q_{22} = -6.2 \times 10^{-4}$

$(Q_{22} - Q_{33})/3 = 5.4 \times 10^{-4}$

$Q^c_{zz} = 1.65 \times 10^{-2}$

$$Q^c_{zz} = 4[(1-3z_q^2)Q_{2'2'} + z_q^2(Q_{2'2'} - Q_{33})]$$
$$Q^c_{xx} = Q^c_{yy} = -\tfrac{1}{2}Q^c_{zz};\ Q^c_{ij} = Q^c_{ii}\delta_{ij}\,. \tag{21}$$

In the undistorted crystal $z_q^2 = \frac{1}{3}$ and $Q_{2'2'} = Q_{33}$, hence $Q^c_{ij} = 0$ as expected. z_q is given by

$$z_q = \frac{s_z}{\sqrt{3}} + (\tfrac{2}{3})^{1/2} s_x \sin\theta\,, \tag{22}$$

where s_x, s_z give the tetragonal distortion defined in Eq. (3). Treating $s_z - 1$, Q_{12}, $Q_{22} - Q_{33}$ as small, it is easy to show that, to lowest order in these quantities, Eq. (21) becomes

$$Q^c_{zz} = 4\{[-2(s_z-1) - 2\sqrt{2}\,\theta]Q_{22} + \tfrac{1}{3}(Q_{22} - Q_{33})\}$$
$$\theta = Q_{12}/(Q_{11} - Q_{22})\,. \tag{23}$$

These quantities are given in Table VI in units of $|e|a^2$ where e is the electronic charge and a is the lattice constant.

On going from cubic to tetragonal the quadrupole tensor changes by only a small amount and the symmetry breaking is even smaller. Equation (23) identifies three contributions to the quadrupole moment of the unit cell. The terms in Q_{22} come from the tilt of the bond quadrupole moment relative to the cube axes consisting of the tetragonal distortion term $(s_z - 1)$ and θ, the tilt of the principal axes due to Q_{12}. The terms involving θ and $(Q_{22} - Q_{33})$ are directly related to the symmetry breaking due to the rotation of the hybrid away from the bond direction. Table VI shows that the two symmetry-breaking terms are small and, furthermore, nearly cancel.

The unit-cell quadrupole moment would be given to a good approximation simply by rotating the quadrupole moment of the bond in the cubic case by the amount of the tetragonal distortion.

---

[1] E. O. Kane and A. B. Kane, Phys. Rev. B 17, 2691 (1978).

[2] E. O. Kane and M. Schlüter, Phys. Rev. B 19, 5232 (1979).

[3] E. O. Kane, Phys. Rev. B 21, 4600 (1980). This paper is referred to as I.

[4] G. F. Koster, Phys. Rev. 89, 67 (1953).

[5] H. Wendel and R. M. Martin, Phys. Rev. B 19, 5251 (1979).

[6] (a) M. T. Yin and M. L. Cohen have very recently obtained excellent values of cohesive energy and elastic constants using a $\vec{k}$-space method. (b) W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), p. 180 ff.

[7] J. W. Linnett and P. J. Wheatley, Trans. Faraday Soc. 45, 33 (1949); J. R. Easterfield and J. W. Linnett, J. Chem. Soc. Faraday Trans. 2 70, 317 (1974). Linnet *et al.* have considered cases where the hybrid directions could partially follow the bonds. They did not consider nearest-neighbor overlap and hence concluded for the case we have considered that the hybrid tetrahedron would not distort.

[8] R. M. Martin, Phys. Rev. B 1, 4005 (1970).