The temperature dependence of the observed recovery at ~ 5 °K gave an apparent migration energy of 0.004 eV.

A discussion of the general features of defect production and recovery of Ge and of possible energy level models was given. The relationship of the present results to the work of Zizine,[17] and the extension of this model to explain Calcott and MacKay's recovery-versus-energy data[3] were also discussed.

[†] Research supported in part by the U. S. Atomic Energy Commission under Contract No. AT(11-1)-1198.

[*] Present address: Varian, Radiation Division, Euclid, Ohio.

[1] J. W. MacKay and E. E. Klontz, J. Appl. Phys. 30, 1269 (1959).

[2] E. E. Klontz and J. W. MacKay, J. Phys. Soc. Japan Suppl. 18, 1216 (1963).

[3] T. A. Callcott and J. W. MacKay, Phys. Rev. 161, 698 (1967).

[4] J. Zizine, in Proceedings of the Symposium on Radiation Effects in Semiconductor Components (unpublished).

[5] J. W. MacKay and E. E. Klontz, in *Proceedings of the Conference on Radiation Effects in Semiconductors*, edited by L. Vook (Plenum, New York, 1968), p. 175.

[6] I. Arimura and J. W. MacKay, in Ref. 5, p. 204.

[7] W. A. Schottky, Z. Physik 118, 539 (1942).

[8] C. O. Thomas, D. Kahng, and R. C. Manz, J. Electrochem. Soc. 109, 1055 (1962).

[9] N. I. Meyer and T. Guldbrandsen, Proc. IEEE 51, 1631 (1963).

[10] A. M. Goodman, J. Appl. Phys. 34, 329 (1963).

[11] D. P. Kennedy, P. C. Murley, and W. Kleinfelder, IBM J. Res. Develop. 12, 399 (1968).

[12] D. P. Kennedy and R. R. O'Brien, IBM J. Res. Develop. 13, 212 (1969).

[13] W. Shockley, Bell System Tech. J. 28, 435 (1949).

[14] J. Bardeen, Bell System Tech. J. 28, 428 (1949).

[15] P. Baruch, J. Appl. Phys. 32, 653 (1961).

[16] J. E. Whitehouse, T. A. Callcott, J. A. Naber, and J. S. Raby, Rev. Sci. Instr. 36, 768 (1965).

[17] J. Zizine, in Ref. 5, p. 186.

[18] A. C. Damask and G. J. Dienes, *Point Defects in Metals* (Gordon and Breach, New York, 1963).

[19] R. C. Fletcher and W. L. Brown, Phys. Rev. 92, 585 (1953).

[20] T. R. Waite, Phys. Rev. 107, 471 (1957).

[21] H. M. Simpson and R. L. Chaplin, Phys. Rev. 178, 1166 (1969).

[22] J. H. Cahn, J. Appl. Phys. 30, 1310 (1959).

[23] G. D. Watkins, in *Radiation Damage in Semiconductors* (Dunod, Paris, 1964), p. 97.

[24] T. M. Flanagan and E. E. Klontz, Phys. Rev. 167, 789 (1968).

PHYSICAL REVIEW B VOLUME 4, NUMBER 6 15 SEPTEMBER 1971

Need for a Nonlocal Correlation Potential in Silicon

E. O. Kane
Bell Telephone Laboratories, Murray Hill, New Jersey 07974

(Received 11 January 1971)

An attempt is made to fit cyclotron masses and principal energy gaps for silicon using a Heine-Abarenkov–type determination of the core-valence interaction fitted to the atomic spectra of Si^{3+}. The valence-valence exchange and correlation potential is approximated by a local potential. The masses and gaps are found to obey a "$\vec{k}\cdot\vec{p}$-type" product relation under variations of the local potential. The theoretical product is 10–25% smaller in absolute value than the experimental product. We conclude that a local approximation to exchange and correlation is inadequate for silicon. If the masses are fitted the gaps are in error by 0.5–0.7 eV. We suggest that screened Hartree-Fock exchange may provide the nonlocality required to overcome these fitting difficulties.

I. INTRODUCTION AND CONCLUSIONS

A basic problem in all band calculations is the choice of potential. Most so-called *a priori* calculations use the Slater "$\rho^{1/3}$" approximation[1] to exchange and correlation. Variations on the $\rho^{1/3}$ method have been suggested such as the use of $\nabla\rho$ corrections and the empirical adjustment of the prefactor multiplying $\rho^{1/3}$.[2] All these methods are characterized by being "local" potentials, i.e., the potential operator may be taken as $V(r)$ rather than the more general $V(r, r')$. The Hartree-Fock approximation involves a nonlocal potential. However, the nonlocality is strongly reduced by screening[3] so that quantitative estimates of the importance of nonlocality have not been made.

In the empirical pseudopotential method[4] the effective potential is assumed to be local and so rapidly convergent in momentum space that it can be adequately represented in silicon by the three lowest Fourier coefficients $\tilde{V}_{111}$, $\tilde{V}_{220}$, and $\tilde{V}_{311}$, all others being taken to be zero. Empirical values for these coefficients in germanium and silicon were first obtained by Brust.[5]

Authors doing *a priori* calculations have also used empirical adjustments to obtain optimum agreement with experiment.[6]

We have taken an approach which is intermediate between the *a priori* and the totally empirical methods. We divide the problem into two parts in a manner very analogous to the Heine-Abarenkov method.[7] A valence electron interacts with the ion cores and also with the other valence electrons. We assume that we can write the Hamiltonian of a valence electron as a term involving the interaction of a single electron with the Si^{4+} cores plus a term representing the interaction with the other valence electrons.

The interaction with the $silicon^{4+}$ cores is determined empirically by using a core potential which reproduces the two lowest atomic energy levels of Si^{3+} for $l = 0$, 1, and 2. The valence-valence interaction is represented by a local potential whose Fourier coefficients $\tilde{V}_{111}$, $\tilde{V}_{220}$, $\tilde{V}_{311}$, $\tilde{V}_{222}$, and $\tilde{V}_{400}$ are treated as adjustable parameters. Although we have used only two more "pseudopotential" parameters than the conventional empirical pseudopotential treatment for silicon,[5] we expect considerably better convergence with our method because we parametrize only the valence-valence interaction and not the total potential. Because the electronic kinetic energy acts to suppress short-range fluctuations the potential due to the valence electrons converges much more rapidly in momentum space than does the potential of the ion cores. To take possible cumulative effects of high Fourier coefficients into account we treat the ionic charge Z as a parameter also. The zero-order value is $Z = 4$.

We use Brinkman and Goodman's[8] calculation for silicon as a starting point and vary each potential parameter by a small amount computing the change in the cyclotron mass parameters F', G', H', $m_0/m_{c\perp}$ and the energy gaps E_G, E_s, E_p, and E_L. F', G', and H' are valence-band mass parameters,[9] and $m_{c\perp}$ refers to the conduction band. E_G is the indirect gap, E_s is the $\Gamma_{25'} - \Gamma_{2'}$ gap,[10] E_p is the $\Gamma_{25'} - \Gamma_{15}$ gap, and E_L is the direct gap at L.[11] These parameters are discussed in more detail in Sec. IV.

We find if we form the "$\vec{k}\cdot\vec{p}$" products $-F'E_s$, $-H'E_p$, and $(m_0/m_{c\perp})(E_G + 3.5\text{ eV})$ that these quantities are relatively insensitive to the potential and sufficiently far from the experimental values (10–25%) that we are unable to fit them with reasonable changes in the potential. Large changes in the potential would be outside the linear regime. While the possibility of such large changes cannot be rigorously excluded we feel that the most likely explanation is that a local valence-valence exchange and correlation potential is not an adequate model. If the masses F', H', and $m_0/m_{c\perp}$ are fit we find $E_s = 3.55$ eV, $E_p = 2.67$ eV, and $E_G = 0.59$ eV, whereas the experimental values are $E_s = 4.00$ eV, $E_p = 3.40$ eV, and $E_G = 1.15$ eV. The errors are quite large by present standards of accuracy. They are "consistent" in the sense that all the theoretical gaps are too small.

Since completing the study reported here we have calculated the screened Hartree-Fock exchange interaction for silicon using pseudopotential wave functions and the Penn dielectric function.[12] We find an increase in the direct energy gap at Γ, X, and L which varies from 1.4–1.9 eV. The origin of this gap enhancement is the greater exchange overlap of the valence states compared to the conduction states. Since this is an occupancy effect it is not easily included in a local potential. We believe this is the origin of our fitting difficulties.

Outline of Paper

In Sec. II we discuss the method of calculation which uses an augmented plane wave (APW) method in which the "nonflat" potential is treated exactly while the nonspherical corrections are treated perturbatively.

In Sec. III we discuss the division of the crystal potential into valence-core and valence-valence interactions. The empirical determination of the valence-core interaction from the atomic energy levels of Si^{3+} is described.

In Sec. IV we discuss the experimental band-structure parameters which we use as input to determine the effective valence-valence potential. The linear interpolation matrix which is used to fit the potential to experiment is presented and discussed.

Section V contains a brief discussion of our conclusions.

II. METHOD OF CALCULATION

We use an APW method which does not rely on the "muffin-tin" approximation. This approach has been described in detail in a recent publication.[12a] The "nonflat" part of the potential is treated exactly while the "nonspherical" terms are treated by first-order perturbation theory. Varying the atomic sphere radius from $2.15a_0$ to $1.35a_0$ led to variations of less than 0.05 eV in the energy levels for sufficiently large numbers of plane waves. In order to reduce computing time, the sphere radius $2.15a_0$ was generally used and a Löwdin[13] partition into states A and B was made. States in A had kinetic energies $\leqslant (2\pi/a)^2$ 9 Ry. Their mutual interactions were treated exactly. States in B had kinetic energies $\leqslant (2\pi/a)^2$ 19.5 Ry. Interactions between A and B were treated perturbatively. The convergence errors for this partition were < 0.05 eV.

We computed the secular determinant of $H - E$ for three energies spaced 0.002 a.u. apart and used quadratic interpolation. The error here was negligible. In computing masses where energies linear in $\vec{k}$ were absent by symmetry we used finite differences, calculating $E(\vec{k}_0)$, $E(\vec{k}_0+\delta\vec{k})$, and $E(\vec{k}_0+2\delta\vec{k})$. Terms in k^4 were eliminated and higher powers neglected. We used $\delta k=(2\pi/a)\,0.026$. Errors of 0.1 to 0.2% in the larger curvatures were estimated by computing four curvatures at the Γ point which are expressible in terms of three mass parameters.

III. FORM OF THE CRYSTAL POTENTIAL

A. Core-Valence and Valence-Valence Partition

We are primarily interested in one-electron energy states of the crystal which lie within 10 eV of the Fermi level. We refer to the electrons in such states as "valence" electrons to distinguish them from the "core" electrons which are much more tightly bound in states of an atomic character.

We assume that we can describe the effective potential for the valence electrons as a sum of two potentials. The first is the "ionic" potential $V^{\text{ionic}}(\vec{r})$, which includes the interaction with the nuclear charge and the core electrons. For $\vec{r}$ outside the atomic spheres $V^{\text{ionic}}(\vec{r})$ is assumed to be accurately represented as a sum of Coulombic potentials. The second is the valence-valence interaction $V^{\text{val}}(\vec{r})$, assumed to be given by a Fourier series which converges rapidly in momentum space. Although Brinkman and Goodman's calculation[8] indicates a very small role for high Fourier coefficients of the valence charge there is the possibility that a small component of the charge associated with the core region converges very slowly in $\vec{k}$ space and has a significant cumulative effect. To allow for this possibility we permit the ionic charge Z to deviate from its zero-order value, $Z=4$. This gives us a sixth degree of freedom for the local potential.

B. Valence-Valence Potential

The valence-valence potential consists of a Hartree contribution plus exchange and correlation terms. The latter are most often treated by Slater's "$\rho_v^{1/3}$" approximation

$$V^{\text{val}}_{\text{corr}}(\vec{r})=-\tfrac{3}{2}\alpha\,[(3/\pi)\rho_v(\vec{r})]^{1/3}\,, \tag{3.1}$$

where $\rho_v(\vec{r})$ is the total valence-electron charge density. $\alpha=1$ is Slater's original potential[1] while $\alpha=\frac{2}{3}$ was suggested by Kohn and Sham.[14] Proposals to treat α as a variable to be fit by experiment have also been made.[2]

More elaborate potentials involving the gradient of ρ have also been proposed. A common characteristic of all such potentials is that they are "local," i.e., they may be represented by a Fourier expansion

$$V(\vec{r})=\sum_{\vec{K}_n}\tilde{V}(\vec{K}_n)\,e^{i\vec{K}_n\cdot\vec{r}}\,, \tag{3.2}$$

where the Fourier transform $\tilde{V}(\vec{K}_n)$ is evaluated on the discrete set of points $\vec{K}_n$ of the reciprocal lattice. Since such an expansion is expected to converge rapidly in momentum space and since its effect on the valence-band energies converges even more rapidly we expect to be able to represent the correlation potential in terms of a small number of parameters. For diamond-type crystals we believe that six should be sufficient.

The method we use bears a close resemblance to the empirical pseudopotential method.[4] The difference is that the potential which we determine empirically is that of the valence-valence interaction rather than the total potential. The ionic potential is much more slowly convergent in momentum space than is the valence-valence interaction. In the effective Hamiltonian of the APW method the ionic potential is greatly reduced by subtracting off the part within the spheres. Even so, the convergence is much slower than for the valence-valence potential so our approach should give better convergence than the empirical pseudopotential method.

We have used the ionic charge Z as a parameter to allow for possible cumulative effects of high Fourier coefficients of the valence charge.

C. Core-Valence Potential

The ionic part of the potential can also be determined empirically as in the method of Heine and Abarenkov[7] by fitting to the atomic spectra of the free ion with a single electron outside the core. If the core states are deep they will not change much between the ion and the neutral atom and even less between the atom and the solid.

We have determined an effective l-dependent (hence nonlocal) ionic potential by fitting to the atomic spectra of Si^{3+}.[15] To simplify the discussion, let us suppose that we can define a core radius R_{core} much less than the atom sphere radius R of the APW method. The potential is assumed to be Coulombic outside R_{core}. Using the known atomic levels, the quantity $u_l'(r,E)/u_l(r,E)$ at $r=R_{\text{core}}$ can be determined by integration of the Schrödinger equation from ∞ in the Coulomb field as in the Heine-Abarenkov method.[7] The values of u'/u can be determined at the energies of interest by extrapolation. If R_{core} is very small, u'/u at R_{core} will not be appreciably different in the solid than in the Si^{3+} ion. These values can then be used as boundary conditions for the integration of the Schrödinger equation for u out to the atom sphere radius R. The boundary condition completely characterizes the interaction of the valence electron with the core.

In order to avoid the assumption of R_{core} being very small and also to reduce the uncertainties associated with the extrapolation of u'/u we adopted a method which is more time consuming than that described above but which we felt should give better results. We constructed an empirical potential with two adjustable parameters per l value which were then fit to give agreement with the two lowest atomic levels for $l = 0$, 1, and 2. The empirical potential consisted of a Hartree part constructed from Clementi's Hartree-Fock core wave functions[16] calculated for Si^{3+} plus a term of the Slater "$\rho_{\text{core}}^{1/3}$" type with coefficient α_l plus a term proportional to ρ_{core} with β_l. ρ_{core} is the core charge density determined from Clementi's wave functions:

$$V^{\text{ionic}}(\vec{\mathrm{r}}) = V^{\text{Hartree}}(\vec{\mathrm{r}}) - \tfrac{3}{2}\left[(3/\pi)\rho_{\text{core}}(\vec{\mathrm{r}})\right]^{1/3}$$

$$\times \textstyle\sum_l |\, l\rangle \alpha_l \langle l\,| + \rho_{\text{core}}(\vec{\mathrm{r}}) \textstyle\sum_l |\, l\rangle \beta_l \langle l\,| \;. \qquad (3.3)$$

The values of α_l and β_l for $l = 0$, 1, and 2 are given in Table I. For $l \geqslant 2$, $\alpha_l = \alpha_2$, $\beta_l = \beta_2$.

We have also set $\beta_l = 0$ and determined α_l by a one-state fit. These values are also noted in the table. The most striking feature is that the values of α_l are much smaller than the customary $\frac{2}{3}$–1 range. They also depend strongly on whether $\beta_l = 0$ or not. Note, however, that we are not following the customary procedure since we did not include the "self-charge" of the outer electron in $\rho^{1/3}$. We are not prepared to interpret the meaning of the α and β parameters further at this point. We regard them merely as a device for obtaining an empirical determination of the effective core-valence interaction. The difference between the experimental energies and the Hartree-Fock energies[16] in Si^{3+} is much less than the Hartree to Hartree-Fock energy difference. We feel, therefore, that any attempt to understand or improve on the present model should be based on a Hartree-Fock rather than a Hartree treatment.

It is interesting to ask how much the core-valence interaction might change in going from Si^{3+} to Si^0. One would expect the difference between Si^0 and Si^{3+} to be greater than the change from Si^0 to the crystal. We use the values of α_l and β_l in Eq. (3.3), which were determined by fitting to Si^{3+}, and we use the core charge density determined from Clementi's core functions for Si^{3+} and for Si^0.

TABLE I. Empirical potential parameters α_l and β_l used to fit atomic spectra of Si^{3+}. See Eq. (3.3).

Two-state fit	
$\alpha_0 = 0.19019$	$\beta_0 = -0.03508$
$\alpha_1 = 0.24017$	$\beta_1 = -0.04075$
$\alpha_2 = 0.13544$	$\beta_2 = -0.26659$
One-state fit	
$\alpha_0 = 0.36696$	$\beta_0 = 0$
$\alpha_1 = 0.35076$	$\beta_1 = 0$
$\alpha_2 = 0.24158$	$\beta_2 = 0$

TABLE II. Dependence of energy levels at Γ on the ionic potential. The first column shows the dependence on the core wave functions going from Si^0 to Si^{3+}. The second column gives the dependence on the functional form of the fitting potential. Subscript a refers to the functional form in Eq. (3.3) while b refers to the potential $\alpha_l r^2 \times \exp(-4r) + \beta_l r^4 \exp(-8r)$. Energies in a.u. (double Ry).

	$\text{Si}^0_{\text{core}}-\text{Si}^{3+}_{\text{core}}$	$\text{Si}^{3+}_a-\text{Si}^{3+}_b$
$\delta\Gamma_{5+}$	-0.0020	-0.0022
$\delta\Gamma_{4-}$	-0.0016	-0.00086
$\delta\Gamma_{2-}$	-0.0030	-0.00047

The change in the three most important levels at Γ as ρ_{core} is varied from Si^{3+} to Si^0 is given in Table II in a.u. The maximum change is 0.08 eV. To the extent that α_l and β_l are physically meaningful one might hope that a large part of the observed change was correctly accounted for. However, we find that changes of the same order of magnitude result when the form of the fitting potential is changed. We have used the two-parameter potential $\alpha_l r^2 e^{-4r} + \beta_l r^4 e^{-8r}$, also fitted to the two lowest levels of Si^{3+}. The difference between the Γ levels for this potential and the Si^{3+} potential based on the charge density shows variations as great as 0.06 eV in Table II. This appears to be the level of accuracy attainable by the present method. We believe further improvement will require a Hartree-Fock method. The question appears to be unimportant at the present time in view of the much larger errors which result from the use of a local approximation to the valence-valence interactions.

IV. EMPIRICAL CORRELATION POTENTIAL FOR Si

In Sec. III we proposed to determine the effective valence-valence potential empirically. Our approach is similar in spirit to the empirical pseudopotential method but should be more rapidly convergent since we parametrize only the valence-valence interaction which converges more rapidly than the ionic part of the potential.

A. Experimental Input Data

The most firmly established experimental data which relate directly to band structure are the indirect gap E_G[17] and the cyclotron-resonance parameters.[18,19] In addition, the location of the conduction-band mimimum $\vec{\mathrm{k}}_{\text{min}}$ is fairly well known.[20] Reasonable estimates of the direct gap at L[11]

$$E_L = E(L_1) - E(L_{3'}) \;, \qquad (4.1)$$

and the "s" and "p" gaps at Γ[10,11] have also been given,

$$E_s = E(\Gamma_{2'}) - E(\Gamma_{25'}) ,$$
$$E_p = E(\Gamma_{15}) - E(\Gamma_{25'}) . \qquad (4.2)$$

These numbers are less accurate and the interpretation on which they are based is less generally accepted. However, they are very crucial to our discussion so we will make use of the values which seem most reasonable to us.

The question of "phonon-renormalization" effects on the masses naturally arises. We have made crude estimates which give reciprocal-mass corrections of the order of 0.01 m_0^{-1}. We will therefore neglect this effect.

The conduction-band minimum is described by two masses $m_{c\perp}^{-1}$ and $m_{c\parallel}^{-1}$, whose experimental values are given in Table III. $m_{c\parallel}^{-1}$ is very much smaller than $m_{c\perp}^{-1}$. We will only use the latter in attempting to determine the empirical correlation potential.

The valence-band maximum can be described in terms of three mass parameters. We use the parameters F', H', and G' which are related to the parameters F, G, H_1, and H_2 of Dresselhaus *et al.*,[9] by the definitions

$$F' = (F - \tfrac{4}{3} H_2)\, 2m_0/\hbar^2, \quad H' = (H_1 + H_2)\, 2m_0/\hbar^2,$$
$$G' = (G + \tfrac{2}{3} H_2)\, 2m_0/\hbar^2. \qquad (4.3)$$

H_2 is small and should not play any important role.

TABLE III. Cyclotron-mass parameters and principal energy gaps. Experimental values and values computed with adjusted potential V_2 of Table IV. See text for definition of symbols.

	Expt	V_2
F'	-5.04^a	-4.96
H'	-4.53^a	-4.09
G'	-0.87^a	-0.90
$m_0/m_{c\perp}$	5.25^b	4.62
E_G (eV)	1.15^c	1.15
E_s(eV)	4.00^d	3.61
E_p(eV)	3.40^e	2.98
E_L (eV)	3.45^e	3.01
$m_0/m_{c\parallel}$	1.09^b	
$ak_{\min}/2\pi$	0.86^f	

[a] Reference 19.
[b] Reference 18.
[c] Reference 17.
[d] Reference 10.
[e] Inferred from Ref. 11 on the assumption that Λ and Δ transitions are close in energy to L and Γ separations.
[f] Reference 20.

The quantities F, H_1, and G refer to $\vec{k}\cdot\vec{p}$ interactions between the valence band $\Gamma_{25'}$ and states of $\Gamma_{2'}$, Γ_{15}, and $\Gamma_{12'}$ symmetry, respectively. They are given by the sums[9]

$$F = \frac{\hbar^2}{m^2} \sum_n \frac{|\langle \Gamma_{25'} | p_x | \Gamma_{2'}^n \rangle|^2}{(E_{25'} - E_{2'}^n)} ,$$
$$H_1 = \frac{\hbar^2}{m^2} \sum_n \frac{|\langle \Gamma_{25'} | p_x | \Gamma_{15}^n \rangle|^2}{(E_{25'} - E_{15}^n)} , \qquad (4.4)$$
$$G = \frac{\hbar^2}{m^2} \sum_n \frac{|\langle \Gamma_{25'} | p_x | \Gamma_{12'}^n \rangle|^2}{(E_{25'} - E_{12'}^n)} .$$

The quantities F and H_1 are dominated by the term in the sum with the smallest energy denominator, namely, $E_{25'} - E_{2'}^1 = -E_s$ and $E_{25'} - E_{15}^1 = -E_p$ [see Eq. (4.2)]. Furthermore, it is well known that the momentum matrix element is much less sensitive to changes of the potential than is the corresponding energy gap.[21] Using these facts we derive an approximate relation between masses and gaps of the form

$$-F'E_s = C_s , \quad -H'E_p = C_p , \qquad (4.5)$$

where C_s and C_p are constants proportional to the square of momentum matrix elements. We refer to Eq. (4.5) as the SBCM $\vec{k}\cdot\vec{p}$ relation where the initials refer to the single-band constant matrix element approximation. The experimental values of F', H', and G' are given in Table III.

We should emphasize that the values of F', H', and G' computed here were determined as described in Sec. II and not by use of the $\vec{k}\cdot\vec{p}$ relations. These relations are introduced in order to explain the data, not to compute it.

The location of the conduction-band minimum is known experimentally to be $\vec{k}_{\min} = (2\pi/a)(0.85, 0, 0)$.[20] No attempt was made to fit this parameter. Some calculations of E_G and $m_0/m_{c\perp}$ were made for $(2\pi/a)(0.825, 0, 0)$ and others for $\vec{k}_{\min}$. The variation in E_G was < 0.02 eV and the variation in $(m_0/m_{c\perp})$ was <0.01.

The direct gap E_s has been determined to be 4.0 eV using electroreflectance and extrapolation for Ge-Si alloys.[10] Zucca and Shen[11] have reported wavelength-modulated reflectance structure on pure silicon at 3.45 and 3.40 eV. On the basis of Ge-Si alloy extrapolations[10] the 3.45-eV structure refers to E_1 transitions while 3.40 corresponds to E_0'. The location of E_1 at Λ or L is still being debated while E_0' may arise from Δ or $\Gamma_{25'}$-Γ_{15} transitions. It appears to us that independent of the resolution of these questions the E_L and E_p energies are not far from 3.45 and 3.40 eV, respectively. The difficulties we shall describe in reconciling masses and gaps appear to be much larger than the uncertainty of the above assignments.

B. Local-Correlation-Potential Difficulties

We have used the orthogonalized-plane-wave

TABLE IV. Fourier coefficients of the valence-valence interaction potential BG potential is that of Brinkman and Goodman in Ref. 8. V_2 is the potential used in our calculations. Energy in double rydbergs.

$(a/2\pi)\vec{K}$	BG potential	V_2
(1, 1, 1)	−0.036 398	−0.037 398
(2, 2, 0)	−0.006 084	−0.006 834
(3, 1, 1)	−0.002 575	−0.001 790
(2, 2, 2)	0.002 132	0.003 532
(4, 0, 0)	−0.002 850	−0.002 440
(3, 3, 1)	−0.002 177	−0.002 177
(4, 2, 2)	0.001 029	0.001 029
(3, 3, 3)	0.000 627	0.000 627
(5, 1, 1)	0.000 107	0.000 107
(4, 4, 0)	−0.000 046	−0.000 046

(OPW) calculation of Brinkman and Goodman[8] as a starting point for parameter fitting.

They have conveniently listed the Fourier components of the valence charge density as well as the Slater $\rho_v^{1/3}$ approximation to exchange and correlation where ρ_v is the valence charge density. We use the value $\alpha = 1$ in Eq. (3.1) which they found gave good results. We have listed the Fourier coefficients of their total valence-valence potential in Table IV along with the potential designated V_2. The latter is an adjusted potential which we have used in most of the calculations reported here.

Using the potential V_2 we have computed the cyclotron mass parameters F', H', G', $m_{c\perp}^{-1}$, the indirect gap E_G, and the principal gaps E_s, E_p, and E_L. These values are presented in Table III together with the experimental values.

As can be seen from Table III, the V_2 model gives a reasonably good approximation to the band structure. In order to determine the best possible fit to experiment, we compute the change in the band parameters due to a 0.005 double rydberg increment in each Fourier coefficient $\tilde{V}(\vec{K}_n)$ for $(a/2\pi)\vec{K}_n = (1,1,1)$, $(2,2,0)$, $(3,1,1)$, $(2,2,2)$, and $(4,0,0)$.

We give the resulting "band-parameter matrix" in Table V. We note that the column referring to the (4, 0, 0) Fourier component is much smaller than the others. It is reasonable to suppose that high Fourier components have still smaller effects. We have attempted to take some account of possible cumulative effects of high Fourier components by allowing the ionic charge Z to deviate from the value 4 assumed in the "V_2" potential. The resulting parameter changes appear in the column headed $\delta Z = -0.2$. We feel that with the aid of these six parameters we have been able to represent adequately the maximum freedom possible for a "local" valence-valence potential. Of course, we are only able to treat linear variations. However, the deviations in Table III between experiment and the potential V_2 are individually quite small and any one of them could easily be corrected by a small variation of the potential well within the linear range. We will shortly discuss incompatibilities between mass parameters and energy gaps which formally lead to very large corrections of the potential. Such potential changes would be both well outside the accuracy of the numbers in Table V and the validity of the linear regime. We prefer to believe that V_2 is reasonably close to the "optimum" potential and that the difficulties we encounter indicate the need for nonlocal and perhaps energy-dependent corrections.

We have tested the SBCM $\vec{k}\cdot\vec{p}$ relation of Eq. (4.5) by calculating the products $-H'E_p$ and $-F'E_s$

TABLE V. Matrix for band parameter changes resulting from changes in the local potential. Fourier components of the potential and the effective ionic charge Z are varied. $\delta B = B(V_2 + \delta V) - B(V_2)$ where B is a band parameter. $\delta\tilde{V}(\vec{K}) = 0.005$ a.u. Also plotted are the products $-H'E_p$, $-F'E_s$, and $(m_0/m_{c\perp})(E_G + 3.5\text{ eV})$ and the experimental values of these products. All band parameter energies in eV.

Band parameter $(a/2\pi)\vec{K}$	(1, 1, 1)	(2, 2, 0)	(3, 1, 1)	(2, 2, 2)	(4, 0, 0)	($\delta Z = -0.2$)	V_2	Expt
$\delta G'$	−0.0014	0.0002	−0.0079	−0.024	−0.0054	0.035		
$\delta H'$	0.262	0.121	−0.444	−0.436	−0.102	−0.827		
$\delta F'$	−0.012	−0.360	−0.746	−0.476	−0.028	0.752		
$\delta m_0/m_{c\perp}$	−0.296	0.000	0.204	0.147	0.052	0.941		
δE_G	0.319	0.003	−0.300	−0.253	−0.051	−0.831		
δE_s	−0.016	−0.272	−0.484	−0.314	−0.022	0.776		
δE_p	0.179	0.057	−0.304	−0.324	−0.053	−0.462		
δE_L	0.149	−0.113	−0.278	−0.175	−0.004	−0.12		
$-H'E_p$	12.11	12.07	12.15	12.04	12.29	12.39	12.20	15.40
$-F'E_s$	17.90	17.79	17.86	17.94	17.92	18.48	17.93	20.16
$(m_0/m_{c\perp})(E_G+3.5)$	21.49	21.50	20.99	20.96	21.49	21.24	21.49	24.4

as shown in Table V. Whereas H', E_p, F', and E_s vary individually by 10–15% the products vary by only 1 or 2% or less. On the other hand, the experimental values of $H'E_p$ and $F'E_s$ are seen to lie significantly outside the range of theoretical values. If we assume the variation of $H'E_p$ and $F'E_s$ is linear in the potential changes it can be seen that a variation of 16 times the maximum tabulated increment would be required to fit the experimental value of $H'E_p$. For $F'E_s$ the $\delta Z = -0.2$ value is quite large, but even here the required change is four times the maximum tabulated increment. For comparison, the differences between experiment and the V_2 potential in Table III are less than the maximum tabulated increment in Table V. The simultaneous fitting of gap and related mass requires at least two parameters per pair and results in larger variations of the potential than the value required to fit only the product. For instance, we can fit both F' and E_s if we take $\delta V(3,1,1) = 7.0 \times 0.005$ a.u. and $\delta Z = 4.9 \times (-0.2)$. These numbers are both unrealistically large and involve large cancellations in the net value of E_s. It seems most unlikely that the proper fitting of F' and E_s can be explained in this way.

In addition, we note that E_G and $m_0/m_{c\perp}$ are strongly correlated, a result which was not anticipated. Assuming that the SBCM $\vec{k}\cdot\vec{p}$ relation is valid for $m_{c\perp}$ also, we infer that the energy $E(\Gamma_{25'}) - E(\Delta_5(\vec{k}_{\min}))$ is relatively insensitive to the local valence-valence potential. We have found empirically that the $\vec{k}\cdot\vec{p}$ -like relation $(m_0/m_{c\perp}) \times (E_G + 3.5) = \text{const}$, with E_G in eV, is nearly satisfied as indicated in Table V. Again, the experimental value of this quantity lies far outside the theoretical spread, being 5.5 times greater than the maximum tabulated variation.

We take the three cases noted above involving F', H', and $m_0/m_{c\perp}$ as indicating a fundamental incompatibility between masses and $\vec{k}\cdot\vec{p}$-related energy gaps within the range of variation of a local valence-valence potential.

If we take the products $-H'E_p$, $-F'E_s$, and $(m_0/m_{c\perp})(E_G + 3.5 \text{ eV})$ computed for the V_2 potential in Table V and use the experimental values for H', F', and $m_{c\perp}$ we obtain $E_s = 3.55$ eV, $E_p = 2.67$ eV, and $E_G = 0.59$ eV. The error in the gaps amounts to 0.5–0.7 eV, which is quite large by the standards of accuracy which are frequently claimed in band calculations. If the gaps are fitted, the masses are in error by 10–20%.

The possibility that these discrepancies could be reduced by making large changes in the potential cannot be absolutely excluded. Nevertheless, it seems to us that the most likely explanation involves the failure of approximating the valence-valence interaction by a local potential.

We should also emphasize that in computing the product $(m_0/m_{c\perp})(E_G + 3.5 \text{ eV})$ we have not actually tested the $\vec{k}\cdot\vec{p}$ relation for the direct gap at Δ. The variables in this equation are E_G and $m_{c\perp}$. The constant 3.5 eV has not been checked as actually yielding the direct gap.

V. SUMMARY

We believe we have presented a strong case for the inadequacy of a local approximation to the valence-valence interaction. In view of the widespread use of Slater's $\rho^{1/3}$ approximation [see Eq. (3.1)] in *a priori* band calculations we believe this is a significant result. The magnitude of the inadequacy of the local approximation appears to be 0.5–0.7 eV in energy gaps or 10–20% in masses. In view of the possible uncertainty in some of the experimental gaps such as the $\Gamma_{25'}$-Γ_{15} separation and the large amount of computation involved in a study of this type we cannot claim these results to be definitive. We hope that they will serve to focus more attention on the use of the screened Hartree-Fock exchange interaction rather than on the local Slater $\rho^{1/3}$ approximation. We should note the recent work of Lipari and Kunz[22] which has also reached this conclusion for the alkali halides.

We have recently completed work which will be reported in more detail later[12] in which we have calculated screened exchange for silicon in the random-phase approximation (RPA) approximation. We used pseudopotential wave functions and the Penn dielectric function.[23] The direct gap was found to be increased by 1.4–1.9 eV at Γ, X, and L. The origin of the effect is that valence-band states have greater exchange overlap with the filled states than do the conduction-band states. Since this is an occupancy effect it is not easily included in a local potential. The effect is in the right direction and more than adequately large to explain the difficulties we encountered with the local approximation. A detailed assessment of this calculation will be given in a subsequent publication.[12] We refer to it here because we believe it to be the most likely explanation of the difficulties encountered in the present attempt to obtain an empirically adjusted band structure with a local potential.

We should emphasize that we have used a nonlocal valence-core potential fitted to atomic spectra as described in Sec. II C. We might have been able to improve on our empirical fit to the experimental band structure by varying the parameters in the core-valence interaction. We have shown by use of Clementi's core wave functions for Si^{3+} and Si^0 that such corrections should be small and hence that the core-valence interaction (including core-valence exchange) can be quite well determined empirically from the free-ion (Si^{3+}) spectra as in the Heine-Abarenkov method.[7]

[1] J. C. Slater, Phys. Rev. 81, 385 (1951).

[2] J. C. Slater, T. M. Wilson, and J. H. Wood, Phys. Rev. 179, 28 (1969).

[3] J. C. Phillips and L. Kleinman, Phys. Rev. 128, 2098 (1962); L. Hedin, *ibid.* 139, A796 (1965).

[4] M. L. Cohen and V. Heine, *Solid State Physics*, Vol. 24, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970).

[5] D. Brust, Phys. Rev. 134, A1337 (1964).

[6] F. Herman, R. L. Kortum, C. D. Kuglin, J. P. Van Dyke, and S. Skillman, in *Methods in Computational Physics*, Vol. 8, edited by B. Alder *et al.* (Academic, New York, 1969), p. 193.

[7] V. Heine and I. Abarenkov, Phil. Mag. 9, 451 (1964).

[8] W. Brinkman and B. Goodman, Phys. Rev. 149, 597 (1966).

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

[10] J. S. Kline, F. H. Pollak, and M. Cardona, Helv. Phys. Acta 41, 968 (1968).

[11] R. R. L. Zucca and Y. R. Shen, Phys. Rev. B 1, 2668 (1970).

[12] E. O. Kane (unpublished).

[12a] E. O. Kane, following paper, Phys. Rev. B 4, 1917 (1971).

[13] P. Löwdin, J. Chem. Phys. 19, 1936 (1951).

[14] W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).

[15] Charlotte E. Moore, *Atomic Energy Levels*, Vol. I (U.S. GPO, Washington, D.C., 1949), Natl. Bur. Std. U.S. Circ. 467.

[16] E. Clementi, IBM J. Res. Develop. Suppl. 9, 2 (1965).

[17] A. Frova and P. Handler, Phys. Rev. Letters 14, 178 (1965).

[18] J. C. Hensel, H. Hasegawa, and M. Nakayama, Phys. Rev. 138, A225 (1965).

[19] J. C. Hensel (private communication). The values quoted are to be regarded as tentative. Note that they differ significantly from earlier published values such as J. C. Hensel and G. Feher, Phys. Rev. 129, 1041 (1963).

[20] E. B. Hale and T. G. Castner, Phys. Rev. B 1, 4763 (1970).

[21] H. Ehrenreich, Phys. Rev. 120, 1951 (1960).

[22] N. O. Lipari and A. B. Kunz, Phys. Rev. B 3, 491 (1971).

[23] D. R. Penn, Phys. Rev. 128, 2093 (1962).

PHYSICAL REVIEW B VOLUME 4, NUMBER 6 15 SEPTEMBER 1971

Non-Muffin-Tin Augmented-Plane-Wave Method for Diamond and Zinc-Blende Lattices

E. O. Kane
Bell Telephone Laboratories, Murray Hill, New Jersey 07974

(Received 11 January 1971)

The augmented-plane-wave method is adapted to the diamond and zinc-blende lattices by including both "nonflat" and "nonspherical" corrections to the usual "muffin-tin" potential. The former are treated exactly while the latter are treated perturbatively. The method is tested for silicon where "nonflat" corrections are as large as 3.4 eV with "nonspherical" corrections less than 0.4 eV. Plane wave convergence is superior to the orthogonalized-plane-wave method. The atomic-sphere radius was varied from $2.15a_0$ to $1.35a_0$ with energy changes of less than 0.03 eV. Nonflat matrix elements are easily computed by use of a spherical-harmonic expansion of the potential due to point charges. Multipole-lattice-sum coefficients are given for the fcc, diamond, and zinc-blende lattices.

I. INTRODUCTION

Until quite recently all calculations using the augmented-plane-wave (APW) method[1] assumed a potential of the "muffin-tin" form, namely constant outside touching atomic spheres and spherical inside. For materials with high coordination numbers this approximation should be quite good and the APW method has been very widely applied to such cases.

For materials with tetrahedral coordination the muffin-tin approximation is very poor and band-structure calculations have usually been done with the orthogonalized-plane-wave (OPW) method.[2]

It has always been recognized that the muffin-tin approximation was merely a computational convenience which was in no way essential to the APW method. Yet, it has been only in the past year that practical "non-muffin-tin" applications have been made. There are two separate contributions to the non-muffin-tin corrections: the "nonflat" part of the potential outside the atomic spheres and the "nonspherical" part of the potential within the spheres. The nonflat part of the problem involves computing plane-wave matrix elements of the potential over the region outside the atomic spheres. This is easily done for plane waves. However, a Fourier representation of the potential is too slowly convergent to be useful because of the Coulombic nature of the ionic potential. Hence the "obvious" approach to the nonflat problem cannot be made. The nonspherical part of the problem can be straightforwardly approached via perturbation theory but the number of terms which couple goes up very rapidly with the angular momentum.

The non-muffin-tin calculations which have ap-