The nature of interplanar interactions in SiC polytypes

J J A Shaw and Volker Heine

TCM, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

Abstract. The stability of SiC polytypes has been discussed elsewhere in terms of successive, pairwise, interplanar interactions J_n between Ising-like pseudo-spins. Here the model is first critically re-examined to resolve some doubts about the generality of the energy expression and the meaning of the J_n. The original expression is validated as the correct most general form, with the four-spin coupling term K re-interpreted as a modification of the third-neighbour interaction depending on the local geometry. Secondly we note the moderately long range of the interaction to third-neighbour atomic double layers found in density functional calculations on Si and SiC. We argue that this can be interpreted as the remnant in the semiconductor of what would be Friedel oscillations in a metal.

1. The questions and principal conclusions

This paper is part of a series [1–5] to further our understanding of why some materials form polytypes, notably SiC. Other work has considered the energies of polytypes at temperature $T = 0$ K [1–3], the contribution of the phonon free energy [4, 5], and the relaxation of bond angles and interatomic distances from the ideal tetrahedral values of the cubic zincblende structure [3, 4]. The present purpose is to tidy up two issues running through all of the above. One is to analyse more thoroughly than before the general energy expression for an arbitrary polytype in terms of interplanar interactions to still certain doubts that have arisen. This model is basic to the analysis of polytype energies at $T = 0$ K [1, 2] and their phonon free energies at higher T [5], so its validity and generality need to be established critically. The other issue is to discuss the physical origin of the range of the interplanar interaction which is found to extend out to about six bond lengths.

 The SiC polytype structures consist of atomic double layers (hereafter 'layers' for short), each of which can be stacked in two different orientations onto the one below, as has been described many times, e.g. in [2] and references given there. It is usual [1–5] to denote the two orientations of layer i by a pseudo-spin $s_i \equiv \{\uparrow \; = 1, \; \downarrow \; = -1\}$ (hereafter a 'spin' for short). The energy including interactions up to the third-nearest-neighbour layer for a system of N layers is then written (per layer) as

$$E_0 - \frac{1}{N}\sum_i (J_1 s_i s_{i+1} + J_2 s_i s_{i+2} + J_3 s_i s_{i+3}). \tag{1a}$$

An additional four-spin term with coefficient K

$$- K s_i s_{i+1} s_{i+2} s_{i+3} \tag{1b}$$

was considered separately by Cheng *et al* [2], although it was found to be smaller than
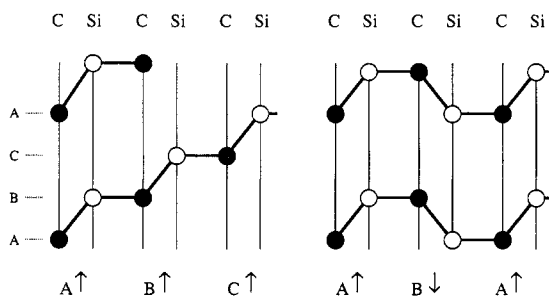
Figure 1. Different relative lateral positions of the final double layer—it is an 'up' layer in both cases—in the two sequences (left) A ↑ B ↑ C ↑ and (right) A ↑ B ↓ A ↑ .

J_3. We are confining ourselves to 'ideal' polytypes in which all bond lengths and angles equal their values in the cubic structure. A layer is then uniquely characterised by its s_i: relaxation effects [3] are not encompassed by (1). The expression can be applied to the polytype energy at $T = 0$ K or to the phonon free energy at arbitrary temperature by making the coefficients functions of T. We emphasise that here a spin refers to a whole (perfect) layer, not to a pair of Si and C atoms as it might in some Ising model.

The energy expression (1) seems obvious enough: interactions between layers up to three layers apart. Moreover it has the required symmetry of being invariant under simultaneous inversion of all the $s_i \rightarrow -s_i$, which corresponds to rotating the whole polytype by 180° about the stacking direction.

One might question, however, whether this was indeed the most general expression up to and including next-next-nearest interactions for four reasons. First, the { ↑ , ↓ } is only a shorthand for a relative change in layer type as a function of position. That is, there are really three different absolute positions at which a Si or C atom could lie, corresponding to the A, B and C sites of the usual hexagonal lattice. Thus the lateral positions of the outer two layers in the set ↑ ↑ ↑ (i.e. A ↑ B ↑ C ↑) and the set ↑ ↓ ↑ (A ↑ B ↓ A ↑) are different (see figure 1). Yet in the interlayer model these outer spins would interact with the same coupling J_2: should not there be two coupling constants J_2' and J_2'' depending on whether we are dealing with the pair of layers AC or AA in this example, i.e. depending on the orientation of the layer in between? For instance, the difference between them could be represented by a three-spin term proportional to $s_i s_{i+1} s_{i+2}$, although such a possibility with an odd number of spins breaks the global spin-inversion symmetry already mentioned.

Second, a problem arises from the geometry of stacking layers one on top of another: if one layer in the structure is reversed in orientation, say the third in figure 2, then all the subsequent layers to the right also undergo a translation from one hexagonal site to another, e.g., layer 4 is translated from A ↑ to B ↑ . The consequent change in interaction energy between layers to the left of the flipped spin and layers to the right must therefore be correctly taken into account in the coefficients of (1).

The third reason concerns the lower symmetry of SiC compared with the corresponding (although hypothetical) polytypes of pure Si or C. To see this, consider a three-layer block as shown in figure 3. The two arrangements ↑ ↑ ↓ and ↓ ↑ ↑ are related by a 180° rotation of the block and hence possess, if all the atoms are identical, the same energy. This is true in pure Si, but in SiC the rotation is not a symmetry operation because it interchanges Si and C atoms and so the two blocks have different
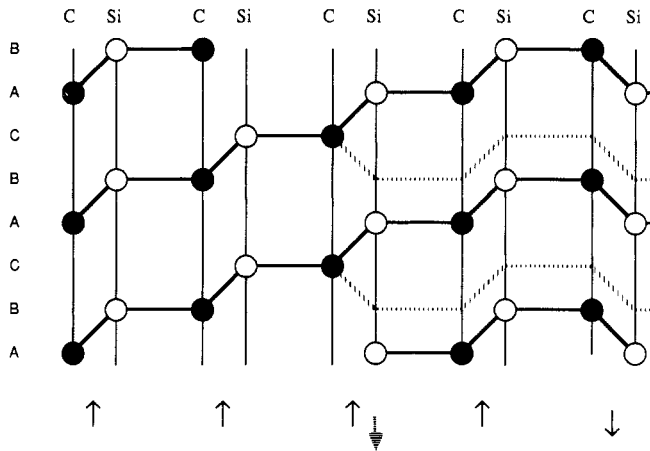
Figure 2. Stacking layers to form SiC polytypes: the five layers are ↑ ↑ ↑ ↑ ↓ (full line) and ↑ ↑ ↓ ↑ ↓ (dashed line), the lines tracing the bond directions in what would be the (110) plane of the cubic polytype. Note that reversing the direction of a single layer—here it is the third—causes a lateral shift of all subsequent layers.
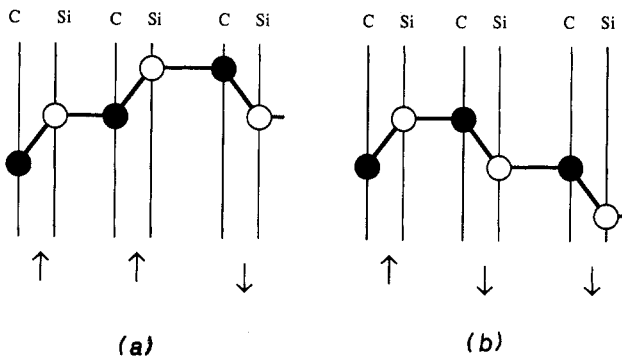

Figure 3. Two configurations (a) ↑ ↑ ↓ and (b) ↑ ↓ ↓ of a three-layer block. Note that if the C atoms are replaced throughout by Si atoms then (a) and (b) are related by a 180° rotation about a vertical axis in the plane of the paper.

energies. Hence, we might expect an energy expression for SiC to require a greater number of parameters than for Si. The paradox is that the energy expression (1), containing five parameters, appears to apply to both systems since the polytypes of both can be uniquely described in terms of the pseudo-spins s_i. The resolution of this paradox is revealing: the energy of an isolated three-layer block of either Si or SiC can be expressed as

$$E^{(3)} = E_0 - J_1(s_1 s_2 + s_2 s_3) - J_2 s_1 s_3 - \alpha(s_1 s_2 - s_2 s_3). \tag{2a}$$

This involves four constants, corresponding to the four different energies present when spin-inversion symmetry is taken into account. For a (hypothetical) Si polytype the last term of (2a) is identically zero by the symmetry shown in figure 3, but not so for SiC.

Now consider an arbitrary set of three layers in a polytype, the energy of which can be written, in analogy with (2a), as

$$E_n^{(3)} = E_0 - J_1(s_n s_{n+1} + s_{n+1} s_{n+2}) - J_2 s_n s_{n+2} - \alpha(s_n s_{n+1} - s_{n+1} s_{n+2}). \tag{2b}$$

To obtain the total energy all such three-blocks are summed over, in which case the α-term sums to zero. The Hamiltonian (1) therefore need contain only the first three terms of (2b) when interactions are truncated beyond next-nearest neighbour. The J_1 in (2b) and (1) differ by a factor of two because of overcounting when the sum over all blocks is taken.

The fourth point concerns the significance of the four-spin term (1b). In [2] it was assumed to be a higher-order interaction, of negligible magnitude relative to J_3 but perhaps comparable to J_4. However, from the preceding discussion, particularly our first point, this is clearly incorrect: interactions between next-next-nearest layers depend on what orientations are possessed by intermediate layers and this is precisely the information carried by a four-spin term. We can combine the terms J_3 and K to write

$$- (J_3 + K s_{i+1} s_{i+2}) s_i s_{i+3} \tag{3}$$

which we interpret as an interaction between layers i and $i + 3$ of magnitude $J_3 + K$ or $J_3 - K$ depending on whether the intervening layers have the same spin or opposite spin. The four-spin term may in addition of course contain a genuine fourth-order multilayer interaction. Similarly if interactions up to fourth-neighbour layers were to be included the J_4-term and two different four-spin terms would be required with coefficients K_4 and K_4', say.

In section 2 we will address these issues by starting in full generality with a finite block of four layers only, and building up the energy expression for the full polytype. We can already understand from the preceding discussion the conclusion that will emerge. Equation (1) is indeed the correct general expression up to third-neighbour interactions, as one might expect from the general validity of symmetry arguments. There are additional interaction constants as seen in (2), but when appropriately formulated they sum to zero for a polytype of infinite extent as shown from (2b). The J_n in (1) are therefore not simple interactions between two layers, but have many effects swept up into them including the change of the interaction with the 'downstream' layers discussed in connection with figure 2. The interpretation of the four-spin term (1b) has already been given as part of the third-neighbour interaction (3) depending on the geometry of the layers in between.

We turn now to the other topic in the present work concerning the range of the interactions that are inherent in the electronic structure energy at $T = 0$ K. When the research project on SiC was started, opinion was very sceptical about significant interactions beyond first- or second-neighbour bonds. Such was the accepted wisdom from organic chemistry for perfect saturated tetrahedral bonds. In reality interactions to third-neighbour layers were found, i.e. to sixth-neighbour bonds [1, 2]. A similar manifestation from the electronic structure can be seen in the Hellmann–Feynman forces on nearby atomic layers around an antiphase boundary (calculated for the ideal structure in the sense already defined) [3]. These forces cause the atomic relaxations observed in the polytypes [3]. From the calculations on the polytypes $\langle 2 \rangle$, $\langle 3 \rangle$ and $\langle 23 \rangle$ one can deduce that these forces extend to about the fourth atomic double layer. What is the origin of this unexpectedly long range? In section 3 we shall discuss the point that SiC is a good example of a nearly free-electron gas, as recognised for diamond by Mott and Jones [6]. We shall argue that the range of the interaction found in SiC is the remnant

in the semiconductor of the long range Friedel oscillations in a metal [7, 8]. In a sense this follows almost by definition from the current understanding of electronic structure. An antiphase boundary or a single reversed layer is a perturbation of the system. This results in a local screening by the electron gas, plus a longer range disturbance of the electron density which we call the Friedel oscillations although they may differ considerably from their simple form in a free-electron gas.

2. Deriving the general energy expression

In this section we derive the most general form of the energy expression for an arbitrary polytype up to and including interactions between third neighbours. To proceed, consider subdividing the bulk configuration of spins $\{s_i\}$ into blocks of four-spins, the size of block being chosen to include the maximum range of the microscopic interactions. Since the blocks overlap there will be some multiple counting when we sum over all blocks to obtain the total energy but this is easily corrected. In making the division into a set of four-blocks, the translational symmetry is lost, and one obtains 2^4 different four-blocks with a total of eight different energies once spin-inversion symmetry has been accounted for. Clearly the five parameters of the energy model (1), E_0, J_1, J_2, J_3, and K can no longer distinguish these eight energies uniquely so new terms with three new parameters are required. Odd-order combinations of spins are not included because they are not invariant under spin inversion $s_i \rightarrow -s_i$ for all i.

The three new parameters α, β, γ are not uniquely defined but a natural choice is the following: three different types of J_1 can be identified in a single four-block, each arising from the difference in energy between different pairs of spins:

$$J_1^a = E(\uparrow \uparrow \cdot \cdot) - E(\uparrow \downarrow \cdot \cdot)$$
$$J_1^b = E(\cdot \uparrow \uparrow \cdot) - E(\cdot \uparrow \downarrow \cdot) \qquad (4)$$
$$J_1^c = E(\cdot \cdot \uparrow \uparrow) - E(\cdot \cdot \uparrow \downarrow).$$

Here the dots denote averaging over up and down spins on those particular sites in the four-block. In SiC all three are distinct, while in pure Si a rotation by 180° about an axis perpendicular to the plane of the paper in figures 1 to 3 is a symmetry operation giving $J_1^a \equiv J_1^c$ in this case. We define the parameter α as the coefficient of the spin term that represents this difference between J_1^a and J_1^c (see (2) and (3)), and β corresponds to the difference $(J_1^a - J_1^b) - (J_1^b - J_1^c)$ while J_1 is the average of J_1^a, J_1^b and J_1^c. Similarly a third new parameter γ is introduced to represent the difference between the two next-nearest neighbour J_2s that can be defined, just as α represents the difference for nearest-neighbour interactions. Combining these additional terms to form a Hamiltonian describing the energy of an arbitrary four-block one obtains:

$$4E_0 - J_1(s_i s_{i+1} + s_{i+1} s_{i+2} + s_{i+2} s_{i+3}) - \alpha(s_i s_{i+1} - s_{i+2} s_{i+3}) - \beta[(s_i s_{i+1} - s_{i+1} s_{i+2})$$
$$- (s_{i+1} s_{i+2} - s_{i+2} s_{i+3})] - J_2(s_i s_{i+2} + s_{i+1} s_{i+3}) - \gamma(s_i s_{i+2} - s_{i+1} s_{i+3})$$
$$- J_3 s_i s_{i+3} - K s_i s_{i+1} s_{i+2} s_{i+3}. \qquad (5)$$

We can now obtain the total energy (1) by summing (5) over successive blocks of four layers along the polytype, dividing the E_0-, J_1- and J_2-, terms by 4, 3 and 2 respectively to correct for multiple counting. We obtain precisely the form (1), the terms with the other parameters α, β and γ having cancelled out (ignoring end effects).

The analysis can be carried one step further. All of the eight parameters in (5) can be defined in terms of the energies of an isolated four-block. For instance from (5), the four-block of spins $\{\uparrow\ \uparrow\ \downarrow\ \uparrow\}$ possesses an energy

$$E(\uparrow\ \uparrow\ \downarrow\ \uparrow) = 4E_0 + J_1 - 2\alpha - 2\beta + 2\gamma - J_3 + K. \tag{6}$$

Equation (6), and the seven similar energy equations, can then be inverted to enable the parameters to be expressed in terms of the energies themselves. The five bulk parameters E_0, J_1, J_2, J_3 and K are then (with $E_1 = E(\uparrow\ \uparrow\ \uparrow\ \uparrow)$, $E_2 = E(\uparrow\ \uparrow\ \uparrow\ \downarrow)$, $E_3 = E(\uparrow\ \uparrow\ \downarrow\ \uparrow)$, $E_4 = E(\uparrow\ \uparrow\ \downarrow\ \downarrow)$, $E_5 = E(\uparrow\ \downarrow\ \uparrow\ \uparrow)$, $E_6 = E(\uparrow\ \downarrow\ \uparrow\ \downarrow)$, $E_7 = E(\uparrow\ \downarrow\ \downarrow\ \uparrow)$, $E_8 = E(\uparrow\ \downarrow\ \downarrow\ \downarrow)$)

$$\begin{aligned}
J_1 &= \tfrac{1}{8}(E_6 - E_1) + \tfrac{1}{24}(E_3 + E_5 + E_7 - E_2 - E_4 - E_8) \\
J_2 &= \tfrac{1}{8}(E_4 + E_7 - E_1 - E_6) \\
J_3 &= \tfrac{1}{8}(E_2 + E_4 + E_6 + E_8 - E_1 - E_3 - E_5 - E_7) \\
K &= \tfrac{1}{8}(E_2 + E_3 + E_5 + E_8 - E_1 - E_4 - E_6 - E_7).
\end{aligned} \tag{7}$$

Shaw [7] has also given the expressions for α, β, γ.

The point of all this exercise is that we believe it validates and illuminates the generality of the expression (1). In section 1 we mentioned various complications which one can certainly see in discussing the finite four-block, particularly the 'downstream' effects shown in figure 2 and the difference in symmetry between pure Si and SiC polytypes. By our procedure we have shown that they are incorporated for a long polytype into the five parameters of (1). For instance the greater symmetry of a Si polytype would imply $\alpha = \gamma = 0$ in that case, but that makes no difference to the final form (1). Thus the J_1, J_2, J_3 and K are seen to be *effective* interactions between pairs of layers, including much else besides. One knows that the energy of the diamond structure cannot be represented physically as the sum of pairwise central forces. Note that in (7) the various E_r are arbitrary unrelated energies: there is no assumption in the analysis about the E_r coming from pairwise forces only, so the E_r can include arbitrarily complicated many-atom covalent effects. Yet the final result (1) has the formal appearance of pairwise interlayer interactions, explicitly for the terms in J_1, J_2, J_3 and implicitly in the term (1b) if one accepts our interpretation of K in section 1 as part of the third-neighbour interaction in the sense of (3).

One can add some further flesh on the dry bones of this discussion by calculating the E_r in (7) *as an approximation* by perturbation theory as a sum of pairwise interactions. We consider two atomic double layers with a separation z placed in a free-electron gas. Their interaction $I(\mu, \nu, n)$ varies with three factors: $\mu = $ p or a, depending on whether the spins of the layers are parallel or antiparallel; $\nu = $ A, B or C if the lateral position of the second layer is A, B or C relative to the first as A, where we take the designation A, B or C from that of the left-hand atom of the double layer; $n = z/d_0 = 1$, 2 or 3 is the separation in units of the interlayer spacing d_0. The total energy E_r of any block of four layers in (7) can now be expressed as a sum of these pairwise interactions $I(\mu, \nu, n)$ where we have six separate interactions for each n. Shaw [7] has done this and hence obtained from (7) lengthy expressions for the J_n and K in terms of the $I(\mu, \nu, n)$.

These expressions confirm two points already advanced in the general preceding discussion. Firstly J_1 does not only involve $I(\mu, \nu, n)$ with $n = 1$ but also terms with $n = 2$ and 3. The latter are the 'downstream' effects mentioned in connection with figure 2. Secondly Shaw [7] obtained a non-zero expression for K, although the approximation
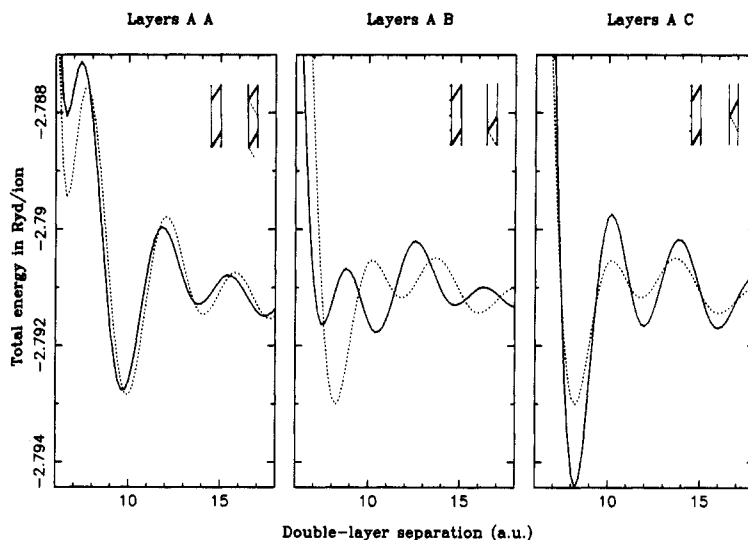
Figure 4. Interactions between planes AA, AB and AC of Si at varying separation, computed using free-electron perturbation theory. The first A layer is always up, while the second is either up (full curve) or down (dotted curve). The ideal interplanar separation in Si is $d_0 = 5.92$ au.

involves pairwise interactions only, with no genuine four-layer interaction. This confirms our interpretation of K as part of the third-neighbour interaction in the sense of (3).

Figure 4 shows the six interactions or rather the total energy of the two interacting layers as a function of separation z in the six geometries. These were calculated by second-order perturbation theory on a free-electron gas as for a metal [10]. They were calculated for pure Si because it was felt that perturbation theory was even less valid for SiC than for Si. There is no suggestion that pure Si should form polytypes: the parameters J_1 and J_2 have quite the wrong ratio [1, 2]. Moreover, we shall argue in section 3 that in the semiconductor SiC the interaction should be damped exponentially as a function of distance and an order of magnitude smaller than in figure 4. Nevertheless we draw two qualitative conclusions. Firstly, all the six values of $I(\mu, \nu, n)$ for given n are comparable, from which it follows that K is in general comparable to J_3. Secondly, dI/dz gives a force between the layers. This is the origin of the Hellmann–Feynman forces found in calculations around antiphase boundaries [3]. The forces, the J_n, the phonon force constants and the interaction between two antiphase boundaries due to the $I(\mu, \nu, n)$ all have the same origin and hence the same range. The latter has been shown to extend to third-neighbour layers in the proper calculation of the J_n [1, 2], and about one layer further in the Hellmann–Feynman forces [3] and the phonon force constants [5].

3. The range of the J_n

In the present section we discuss the origin of the range of the J_n (and of the Hellman–Feynman forces: see the end of section 2) which somewhat surprisingly extends to about third or fourth atomic double layers, i.e. over a line of about six or eight bonds.

Table 1. Interlayer interactions J_1 and J_2, in eV, calculated for Si by free-electron perturbation theory (FEPT) and by total energy calculations with density functional theory (DFT).

	FEPT	DFT†
J_1	0.056	0.0078
J_2	0.012	−0.0014

† After [1], [2] and [9].

Figure 4 shows the bare interaction that determines the J_n, calculated by second-order perturbation theory for Si double layers as if in a metal of the same mean valence electron density, as discussed in section 2. The curves show the Friedel oscillations in the interaction. The main point is that their amplitude decreases very slowly with distance for a layer–layer interaction. The distance in the figure extends to third layers but clearly the amplitude continues with little decrease far beyond that.

Our next point is that the amplitude of the interaction in a metal is an order of magnitude larger than computed from total energy calculations for a semiconductor. We choose to work with pure Si instead of SiC for the reason given in section 2. Although Si always has the cubic structure because of the magnitude and sign of J_1 (relative to J_2), this does not invalidate our discussion of the longer range interactions. The results are compared in table 1, from a second-order perturbation theory calculation as for a metal and from full total energy calculations [9]. The perturbation calculations [7, 10] were carried out to be as analogous to the total energy calculations as possible, namely the energies of several polytype structures were calculated and then analysed by fitting the form (1). We are not interested here in how well the two calculations compare in detail, only in the qualitative point that the proper results for J_1, J_2 are an order of magnitude smaller than what they would be in a free-electron-like metal.

At the end of section 1 we gave our reason for interpreting the long range part of the interaction as Friedel oscillations in a generalised sense. We can consider it to be the remnant in the semiconductor of metallic Friedel oscillations. The range, although unexpectedly long for a semiconductor, is much less than in the metal, and the amplitude is an order of magnitude less.

We now consider in more detail how the results for the semiconductor should compare with those for a metal. Turning a layer over or creating an antiphase boundary is a perturbation in the semiconductor and should be calculable, at least approximately, by semiconductor perturbation theory. This differs from metallic perturbation theory in three ways: (i) in the existence of the band gaps in the energy spectrum; (ii) in the absence of a Fermi radius; and (iii) in the shape of the occupied region of k-space. Let us amplify these points in turn. Second-order perturbation theory for the energy involves expressions of the form (the energy/wave-number characteristic in the sense of [10])

$$\chi(q) = \sum_k \frac{f(k) - f(k + q)}{E(k) - E(k + q)} \, |M(k, q)|^2 \tag{8}$$

where E is the band structure energy, M some matrix element, and f the Fermi occupation factor that ensures (8) is non-zero only if the state k is occupied and $k + q$ unoccupied (or vice versa). In free-electron perturbation theory the $\chi(q)$ has a moderately sharp S-shape in the range $q = 1.6k_F$ to $2.4k_F$ where k_F is the Fermi radius because the energy denominator of (8) goes to zero for states k, $k + q$ spanning a Fermi diameter. The
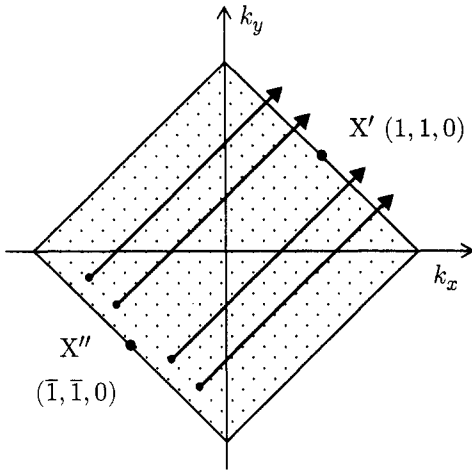
Figure 5. The $k_z = 0$ section of the Jones zone for tetrahedral semiconductors. The points X′, X″ map onto the point X (001) in the Brillouin zone, all in units of $2\pi/a$. The arrows join occupied states k inside the zone to unoccupied states at $k + g$ (220).

Fourier transform of this S-shape results in the Friedel oscillations of the interaction I in real space evident in figure 4. Regarding point (i) above, in a semiconductor the χ(equation (8)) and resulting I should be reduced because the energy denominator cannot go to zero, its minimum value being the band gap. Figure 5 shows the Jones zone in the sense of [6, 11, 12] for a diamond-type semiconductor. The whole Jones zone is roughly spherical, contains exactly four electrons per atom, and replaces the Fermi sphere of a metal. Regarding point (ii), the wave vector spanning the Jones zone is now the reciprocal lattice vector g (220), shown in figure 5 and its analogues across other pairs of zone faces. The resultant Friedel-like oscillations are now commensurate with the lattice. Those from different pairs of zone faces will add up on some atoms and cancel on others to build up a coherent table 1. Figure 5 shows by arrows various points k and $k + g$ (220), and we note (regarding point (iii)) that, unlike in a free-electron metal with a Fermi sphere, we now have a large region of k-space over which the energy denominator attains its minimum value, the band gap E_G across the zone face. The situation is analogous to that of 'nesting' of a (non-spherical) Fermi surface [13], which enhances the amplitude of the Friedel-like oscillations. There is some empirical evidence for the nesting effect in that the phonon force constants are particularly strong and longer ranged in the (110) direction in Ge [14], though less so in SiC [5]. Incidentally the Jones zone, drawn originally for the cubic structure, is not destroyed in polytypes. It is a curious fact that the x-ray structure factor in reciprocal space for any regular or irregular stacking sequence has delta functions at what are some of the reciprocal lattice vectors of the diamond structure, whereas others are split for a regular polytype or spread into lines in reciprocal space for irregular stackings [15]. We therefore expect something like the Jones zone of figure 5 with its (220) zone faces to apply to any polytype.

Unfortunately there appears to be no analytical model that can help us very much to estimate the $\chi(q)$ and I for a semiconductor or even the asymptotic form. The isotropic Penn model [16] replaces the Jones zone (figure 5) by a sphere, retaining the band gap, but the integrals are intractable. Sokel and Harrison [17] evaluate (8) with M = constant for a semiconductor with a direct band gap, i.e. where k in (8) is near a valence band maximum and $k + q$ around a conduction band minimum. They obtain an exponentially decaying interaction $I(z)$, which would become modulated by an oscillation if the band

extrema were separated in k-space. However, in our case the valence band around X'' (figure 5) is a saddle point, corresponding to an approximately constant band gap across the zone face [12]. Resta [18] and Liu [19] also obtain exponentially decaying interactions from other models, but in all cases the exponent is more or less the ratio of band gap to total valence band width E_{F0}.

In view of the shortcomings of the simplified models we have to resort to computation. Table 1 compares the values of J_1 and J_2 for Si from free-electron perturbation theory (FEPT) [8] and from total energy calculations using density functional theory (DFT) with the local density approximation for exchange and correlation [9, 10]. In both cases the energies of some simple structures were calculated, and J_1, J_2 obtained by fitting the form (1), the FEPT ones being our own and the DFT ones from [1], [2] and [9]. The FEPT gives quite a reasonable account of short range bonding, even in Si [20], because it is quite a good example of a nearly free-electron gas [6, 11, 12] with a vertical band gap small compared with the total valence band width, and with a band structure that conforms to the expected form [12]. However, even J_1 depends on the interaction of atoms which are a minimum of three bond lengths apart. It is therefore not surprising to see in table 1 that the DFT values are an order of magnitude smaller than those from FEPT. Incidentally the values for FEPT are only a rough indication: the way they are calculated includes the effect of some more distant J_n which is not negligible for FEPT because of the long range of the Friedel oscillations in that case.

We conclude that the magnitude and range of the interactions found by DFT calculations in SiC [1–5], Si [1, 2, 9] and Ge [14] can plausibly be interpreted as the remnant in the semiconductors of what would be Friedel oscillations in a free-electron metal. In broad terms the semiconductors are like a nearly free-electron gas [6, 11, 12], so one would expect a remnant. The magnitude would be reduced by the band gap, counterbalanced to some extent by the nesting effect, and an order-of-magnitude reduction of J_1, J_2 is found compared with metallic values (table 1). The interaction is expected to decay exponentially with distance, unlike the metallic case of figure 4, with a decay length proportional to (E_{F0}/E_G) or $(E_{F0}/E_G)^{1/2}$ depending on the model [17–19] with various numerical factors. For SiC we estimate the vertical gap $E_G \simeq 5$ eV [21] giving $E_{F0}/E_G \simeq 4$. The observed interactions have a decay of order e^{-1} per double layer spacing d_0 [1–5], i.e. vanishing to a few per cent in three or four double layers, where $k_F d_0 \simeq 4.5$. These figures are entirely consistent with the expected picture of exponentially decaying, Friedel-like oscillating interactions in the semiconductor.

Acknowledgments

References

[1] Cheng C, Needs R J, Heine V and Churcher N 1987 *Europhys. Lett.* **3** 475
[2] Cheng C, Needs R J and Heine V 1988 *J. Phys. C: Solid State Phys.* **21** 1049
[3] Cheng C, Heine V and Needs R J 1990 *J. Phys.: Condens. Matter* **2** at press

[4] Cheng C, Heine V and Needs R J 1990 *Europhys. Lett.* at press

[5] Cheng C, Heine V and Jones I L 1990 *J. Phys.: Condens. Matter* at press

[6] Mott N F and Jones H 1935 *The Theory of Metals and Alloys* (Oxford: Clarendon)

[7] Shaw J J A 1988 The form and origin of incommensurate structures and polytypes *Thesis* University of Cambridge

[8] Friedel J 1977 *Electron–Phonon Interaction and Phase Transition* ed T Riste (New York: Plenum)

[9] Chou M Y, Cohen M L and Louie S G 1985 *Phys. Rev.* B **32** 7979

[10] Heine V and Weaire D L *Solid State Physics* vol 24 (New York: Academic) p 249

[11] Heine V and Weaire D L *Solid State Physics* vol 24 (New York: Academic) p 354

[12] Heine V and Jones R O 1969 *J. Phys. C: Solid State Phys.* **2** 719

[13] Heine V and Weaire D L *Solid State Physics* vol 24 (New York: Academic) p 288

[14] Fleszar A and Resta R 1986 *Phys. Rev.* B **34** 7140

[15] See, for example,
Betteridge G P 1973 *J. Phys. C: Solid State Phys.* **6** L427

[16] Penn D R 1962 *Phys. Rev.* **128** 2093

[17] Sokel R and Harrison W A 1976 *Phys. Rev. Lett.* **36** 61

[18] Resta R 1977 *Phys. Rev.* B **16** 2717

[19] Liu L 1982 *Phys. Rev.* B **26** 975

[20] Heine V and Hafner J 1989 *Many-Atom Interactions in Solids* ed M Puska and R Niemenen (Berlin: Springer)

[21] Madelung O, Schultz M and Weiss (ed) 1982 *Landolt–Börnstein: New Series* Group III, vol 17a (Berlin: Springer)