First-principles calculations of self-interstitial defect structures and diffusion paths in silicon

1999 J. Phys.: Condens. Matter 11 10437

(http://iopscience.iop.org/0953-8984/11/50/332)

J. Phys.: Condens. Matter **11** (1999) 10437–10450. Printed in the UK PII: S0953-8984(99)06618-7

First-principles calculations of self-interstitial defect structures and diffusion paths in silicon

R J Needs

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

Received 3 August 1999, in final form 7 October 1999

Abstract. A first-principles pseudopotential study of neutral self-interstitial defects in silicon is reported, together with calculations for Pandey's concerted exchange mechanism for self-diffusion. The energies and structures of the fully relaxed hexagonal, tetrahedral, split-⟨110⟩, 'caged' (Clark S J and Ackland G J 1997 *Phys. Rev.* B **56** 47), split-⟨100⟩, and bond-centred interstitials are calculated using supercells with up to 128 + 1 atoms. We present results obtained using the local density approximation (LDA) and the PW91 generalized gradient approximation (GGA) for the exchange–correlation energy. Both the LDA and PW91–GGA functionals give the hexagonal and split-⟨110⟩ defects as the lowest-energy self-interstitials. The hexagonal and split-⟨110⟩ defects are essentially degenerate in energy with formation energies of about 3.3 eV (LDA) and 3.80 eV (PW91–GGA). Energy barriers are studied by calculating saddle-point structures using a simple 'ridge-walking' method. The energy barrier for a diffusive jump between the hexagonal and split-⟨110⟩ interstitial sites is calculated to be 0.15 eV (LDA) and 0.20 eV (PW91–GGA) and the barrier between neighbouring hexagonal sites is 0.03 eV (LDA) and 0.18 eV (PW91–GGA), but we have not found a low-energy path between split-⟨110⟩ interstitial sites. The results suggest that self-interstitial diffusion in silicon occurs via diffusive jumps between the hexagonal sites and between hexagonal and split-⟨110⟩ defects.

1. Introduction

The most important intrinsic point defects in silicon are vacancies and self-interstitials [1,2]. In thermal equilibrium the concentration of vacancy and self-interstitial defects in silicon is small because their formation energies are several eV. Interstitials are often created during processing, such as in ion implantation [3], plasma etching, or chemical treatments, and they may also be created by irradiation. Both vacancies and interstitials are mobile, so they can annihilate one another, or form other defects such as pairs or larger complexes, or they may combine with dopant or impurity atoms. One of the important problems encountered in improving electronic devices is that the diffusion of dopant atoms during thermal processing of the device may limit how small the device can be made [2]. As this diffusion is intimately linked with the presence of intrinsic point defects it is important to have a good understanding of vacancy and self-interstitial defects in silicon.

In this paper we study the neutral silicon self-interstitial. Silicon self-interstitials have not been detected directly, but their presence has been inferred from various measurements [1]. The self-diffusion constant or self-diffusivity of silicon can be measured at high temperatures using radioactive isotopes of silicon as tracers, but of course this measurement does not tell us which diffusive mechanism or mechanisms are operating. The experimental data for self-diffusion have established an Arrhenius behaviour with an activation energy in the range 4.1–5.1 eV [4]. The self-diffusivity, D_{SD}, is usually written as the sum of contributions from independent

0953-8984/99/5010437+14$30.00

diffusive mechanisms. The contribution of a particular microscopic mechanism can be written as the product of the diffusivity, D_i, and the concentration, C_i, of the relevant defect, i.e.,

$$D_{\mathrm{SD}} = \sum_i f_i D_i C_i \tag{1}$$

where the f_i are correlation factors of order unity. In a recent review, Gösele *et al* [5] gave experimental estimates of the contributions to the self-diffusivity from self-interstitials and vacancies as

$$\begin{aligned} D_{\mathrm{I}}C_{\mathrm{I}} &= 914\exp(-4.84/k_BT)\ \mathrm{cm}^2\ \mathrm{s}^{-1} \\ D_{\mathrm{V}}C_{\mathrm{V}} &= 0.6\exp(-4.03/k_BT)\ \mathrm{cm}^2\ \mathrm{s}^{-1} \end{aligned} \tag{2}$$

where k_BT is in units of eV. These results predict that self-interstitial diffusion is important at higher temperatures and is dominant above ≈1200 K. Similar results have been obtained by Bracht *et al* [6] who give

$$\begin{aligned} D_{\mathrm{I}}C_{\mathrm{I}} &= 2980\exp(-4.95/k_BT)\ \mathrm{cm}^2\ \mathrm{s}^{-1} \\ D_{\mathrm{V}}C_{\mathrm{V}} &= 0.92\exp(-4.14/k_BT)\ \mathrm{cm}^2\ \mathrm{s}^{-1} \end{aligned} \tag{3}$$

which predicts that self-interstitial diffusion dominates above about ≈1100 K.

The experimental situation regarding self-diffusion in silicon is, however, still highly controversial, especially when it comes to the individual values of D_i and C_i. Indeed, experimental data have been used to support values of the diffusivity of the silicon self-interstitial, D_{I}, which differ by ten orders of magnitude at the temperatures of around 800 °C at which silicon is processed [2].

Clearly self-diffusion in silicon is both technologically important and imperfectly understood, and our aim is to perform an accurate first-principles study which can help clarify the situation. Because of the paucity of direct experimental information about silicon self-interstitials the reliance on calculated properties is substantial. We have performed a series of first-principles density functional theory (DFT) calculations using the local density approximation (LDA) [7] and the PW91 generalized gradient approximation (GGA) [8] for the exchange–correlation energy. We have considered both the relaxed structures and the diffusive paths for self-interstitial atoms. Our aim is to perform a reasonably comprehensive study at the LDA and PW91–GGA level, being careful to check the convergence of the calculations. This allows us to make more detailed comparisons between the various defects and possible diffusion processes than has previously been possible.

There have been a large number of previous studies of silicon self-interstitials using methods similar to ours [9–16], and it is legitimate to ask why another one is needed. A few of the more important reasons are:

(1) In two recent LDA studies of silicon self-interstitials the lowest-energy defects were found to have energies around 2.2 eV in one study [15] and 4.2 eV in the other [16]. This discrepancy is large enough to affect the comparison with experimental diffusion data, and it must be resolved.
(2) In a recent LDA study [15] a new low-energy interstitial configuration was found and named the 'caged' interstitial. This needs to be tested by further calculations.
(3) The GGA approach to DFT has gained in popularity in recent years and, as far as we are aware, this is the first GGA study of self-interstitial defects in silicon.
(4) The nature of the diffusive paths between the low-energy structures is very important and our understanding of them is far from complete. These paths are necessary inputs to thermodynamic integration calculations of the self-diffusion constant, and they can also be used to understand the diffusive jumps that occur in molecular dynamics simulations.

2. Calculational methods and convergence tests

Our calculations are based on the Kohn–Sham density functional theory (KS–DFT) using LDA [7] and PW91–GGA [8] density functionals. Within the LDA the contribution to the exchange–correlation energy at a point $\boldsymbol{r}$ in space is taken to be that in a uniform electron gas of charge density $n(\boldsymbol{r})$. In GGAs the contribution to the exchange–correlation energy at $\boldsymbol{r}$ is approximated by a non-linear function of $n(\boldsymbol{r})$ and $|\nabla n(\boldsymbol{r})|$. The LDA works well for nearly uniform systems but less well for systems with highly inhomogeneous charge densities. GGAs tend to lower the energies of systems with more inhomogeneous charge densities.

We have used a norm-conserving pseudopotential for the Si^{4+} ion generated from all-electron atomic LDA calculations via the method of Troullier and Martins [17]. A supercell approach is used in which a unit cell containing a single defect is repeated throughout space to retrieve three-dimensional periodicity. We used the experimental value of the lattice constant of 5.429 Å for all calculations. The one-electron Schrödinger equations are solved using iterative methods. The positions of the atoms are relaxed using the forces on the atoms. A review of these methods can be found in reference [18].

We will not discuss our calculational methods in detail because they are standard; however, we will devote some space to the convergence of our results with respect to the size of the basis set, the quality of the Brillouin-zone (BZ) integrations, and the size of the supercell. Our calculations were performed using a plane-wave basis set. The BZ integrations were performed by sampling a regular mesh of points in reciprocal space. We used fcc supercells whose translation vectors were integer multiples of the primitive translation vectors of the diamond structure. These integer multiples are denoted by n, so an $n \times n \times n$ supercell of the perfect diamond structure contains $2n^3$ atoms.

First we checked the convergence with respect to the size of the basis set. LDA formation energies for the tetrahedral, hexagonal, and split-$\langle 110 \rangle$ interstitials for an $n = 2$ supercell with various different basis-set energy cut-offs up to a maximum of 36 Ryd are shown in table 1. These results indicate that the energy differences between the defect structures are converged to 0.03 eV with the 12 Ryd cut-off, which was used for most of the results reported in this paper. However, each of the defect formation energies calculated with a 12 Ryd cut-off is subject to a basis-set correction of about -0.3 eV.

Table 1. Self-interstitial defect formation energies in eV calculated within the LDA as a function of the plane-wave cut-off energy for the $n = 2$ simulation cell and using the $m = 4$ k-point mesh.

Defect	12 Ryd	24 Ryd	36 Ryd
Split-$\langle 110 \rangle$	3.49	3.28	3.23
Hexagonal	3.73	3.50	3.44
Tetrahedral	3.64	3.37	3.31

Second we checked the convergence with respect to the quality of the BZ integration. We have tested the quality of the BZ integration on the perfect diamond structure, and the tetrahedral, hexagonal, and split-$\langle 110 \rangle$ interstitials. We define the density of the BZ sampling by an integer, m, and we sample $m \times m \times m$ meshes of k-points in the BZ. These grids are of the equally spaced Monkhorst–Pack type [19], but are centred on an L point of the BZ. This centring gives the same k-point meshes as the Monkhorst–Pack definition for even values of m, but gives superior BZ integrations for odd values. Our meshes also have the advantage that the quality of the BZ integration varies smoothly with m. Some LDA results from these tests are shown in table 2. After testing various values of n and m on the above structures we have selected $m = 4$ for $n = 2$, $m = 3$ for $n = 3$, and $m = 2$ for $n = 4$ for the calculations reported

Table 2. Self-interstitial defect formation energies in eV calculated within the LDA as a function of the k-space integration scheme for the $n = 2$ simulation cell with a 12 Ryd plane-wave cut-off energy. Γ indicates that only the Γ point of the simulation cell was included, while L indicates that only the L point was included, and the other data are for $m \times m \times m$ meshes of the type described in the text.

Defect	Γ	L	$m = 2$	$m = 4$	$m = 8$
Split-$\langle 110 \rangle$	2.77	3.02	3.61	3.49	3.51
Hexagonal	1.96	3.34	3.75	3.73	3.71
Tetrahedral	0.80	2.34	3.57	3.64	3.63

here, which our tests show give errors in the energy differences between structures of about 0.03 eV.

We also tested the convergence of the energies and relaxed structures with respect to the size of the supercell, i.e., with respect to n. Some results from these tests are given in table 3, which show that the LDA total energies converge quite rapidly with the size of the supercell. The defect formation energies for the $n = 2$ and $n = 4$ supercells differ only by about 0.1 eV. The relaxed structures change by a small amount on going from the $n = 2$ to $n = 3$ supercells, but the further change on increasing the supercell to $n = 4$ is extremely small. Finally we tested the effect on the structures of replacing the LDA by the PW91–GGA. This effect was very small and we decided to use the relaxed LDA structures for all the calculations presented here.

Table 3. Self-interstitial defect formation energies in eV calculated within the LDA as a function of the size of the simulation cell, denoted by n (see the text). A plane-wave cut-off energy of 12 Ryd was used and the BZ integrations are performed using $m \times m \times m$ k-point meshes with $m = 4$ for $n = 2$, $m = 3$ for $n = 3$, and $m = 2$ for $n = 4$, which give very good BZ integrations in each case.

Defect	$n = 2$	$n = 3$	$n = 4$
Split-$\langle 110 \rangle$	3.49	3.63	3.57
Hexagonal	3.73	3.67	3.60
Tetrahedral	3.64	3.69	3.76

From the above results we estimate that the errors due to the BZ integrations and from the finite supercell are negligible for the $n = 4$, $m = 2$ results given in table 3. These results still show a significant error in the formation energies of about -0.3 eV due to the use of a 12 Ryd basis-set energy cut-off, but because it appears to be an almost constant error it is easily corrected for.

3. Energies and structures of the self-interstitial

A large number of possible structures for the silicon self-interstitial have been proposed. The most widely studied are the hexagonal, tetrahedral, split-$\langle 110 \rangle$, split-$\langle 100 \rangle$, and bond-centred structures. Recently Clark and Ackland [15] performed a first-principles LDA study of the silicon self-interstitial and found a new low-energy interstitial site which they named the 'caged' interstitial. There are disagreements between the various calculations concerning the energies of the various defects, but the general consensus among first-principles LDA calculations is that the hexagonal, tetrahedral, and split-$\langle 110 \rangle$ interstitials are low in energy, while the split-$\langle 100 \rangle$ and bond-centred structures are significantly higher in energy. Clark and Ackland [15] found their 'caged' interstitial to belong to the low-energy group.

Unless stated otherwise, the results quoted in this section are for the $n = 4$ cells with $m = 2$ BZ sampling and a 12 Ryd cut-off, with a basis-set correction applied. These corrections are obtained as the difference between results for $n = 2$, $m = 2$ calculations with 12 Ryd and 36 Ryd cut-off energies. Results for such calculations are reported in table 1 for the lowest-energy defects. In each case we used structures which were fully relaxed within the LDA. The final energies of the interstitial defects are reported in table 4 and the LDA energies are compared with those obtained by other authors in table 5.

Table 4. Best estimates of the self-interstitial defect formation energies, E_F, calculated within the LDA and PW91–GGA. These results were obtained from $n = 4$, $m = 2$, 12 Ryd calculations (see table 3 for the LDA data), with an additional basis-set correction as described in the text.

Defect	E_F^{LDA} (eV)	E_F^{PW91} (eV)
Split-⟨110⟩	3.31	3.84
Hexagonal	3.31	3.80
Tetrahedral	3.43	4.01

Table 5. Comparison of defect formation energies from various LDA calculations.

Defect	E_F^{LDA} (eV)
Split-⟨110⟩	3.31[a], 3.3[b], 2.16[c], 4.19[d]
Hexagonal	3.31[a], 2.45[c], 4.23[d]
Tetrahedral	3.43[a], 4.35[d]
Concerted exchange	4.58[a], 4.3[e]

[a] This work.
[b] Blöchl *et al* [14].
[c] Clark and Ackland [15].
[d] Lee, Lee, and Chang [16].
[e] Pandey [23].

The first proposal that the split-⟨110⟩ interstitial might be the most stable self-interstitial in silicon was by Bar-Yam and Joannopoulos [11], who called it the X configuration. Our calculations give the formation energy of the split-⟨110⟩ interstitial as 3.31 eV (LDA) and 3.84 eV (PW91–GGA). Our LDA value is in excellent agreement with the first-principles calculations of Blöchl *et al* [14], who obtained an LDA formation energy of 3.3 eV for this defect using a 64-atom supercell, but is not close to the LDA values of 2.16 eV obtained by Clark and Ackland [15] and 4.19 eV by Lee *et al* [16]. The structure of the split-⟨110⟩ interstitial is illustrated in figure 1. Below we list the atomic positions in a Cartesian basis and units of Å for our relaxed split-⟨110⟩ defect. Atoms 1 and 2 share the lattice site at the origin and atoms 3 to 6 are the near neighbours.

1:	−0.862	−0.862	0.787
2:	0.862	0.862	0.787
3:	−1.413	−1.413	−1.406
5:	−1.449	1.449	1.433
4:	1.413	1.413	−1.406
6:	1.449	−1.449	1.433

The bond between the two atoms forming the split interstitial is 2.44 Å long, which is slightly longer than the bond length in bulk silicon of 2.35 Å. The atoms forming the split interstitial

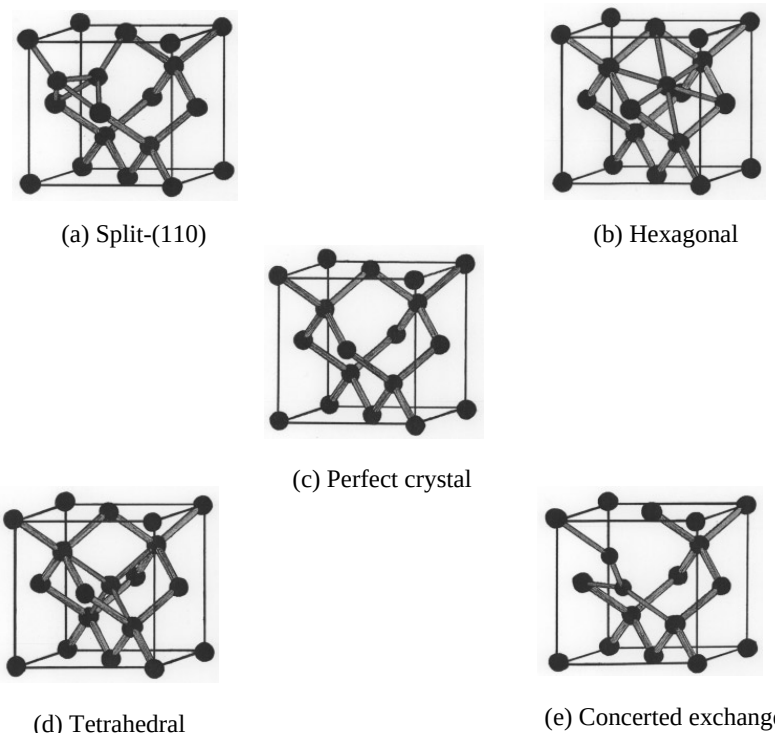

Figure 1. Illustrations of the split-⟨110⟩, hexagonal, and tetrahedral self-interstitial defects, together with the perfect crystal and the saddle point of Pandey's concerted exchange.

are fourfold coordinated, while two of the neighbouring atoms are fivefold coordinated. The LDA band gap for the relaxed structure calculated from the band energies at the k-points used for the self-consistent calculation was 0.51 eV, which is not much smaller than the band gap of 0.81 eV for the perfect diamond structure at these k-points. (The calculated minimum gap for the perfect diamond structure considering all k-points was 0.55 eV within the LDA.) To determine whether any of the states with energies around the gap region were localized on the defect we integrated the modulus squared of each wave function over spheres of various radii centred on the defect. We were unable to find any indication of localized defect states. We have also found that the band gap for the defect structure is sensitive to the size of the simulation cell, which also indicates that there are no strongly localized states in the gap. We conclude that the split-⟨110⟩ interstitial has either very shallow states in the gap or no states in the gap. We note that, on the basis of their first-principles calculations on the split-⟨110⟩ interstitial, Car *et al* [13] have stated that 'no states in the gap are associated with this structure'.

We have studied the new 'caged' interstitial proposed by Clark and Ackland [15]. Clark and Ackland performed a first-principles LDA study of silicon self-interstitials, obtaining energies of formation of 2.16 eV for the split-⟨110⟩ interstitial and 2.45 eV for the hexagonal interstitial. Note that these energies are about 1 eV lower than our LDA results and are also significantly different from the energies found in previous studies. Clark and Ackland stated that the reason for the differences from previously published first-principles results was that the previous studies used an insufficient number of k-points in the Brillouin-zone integrations. Our calculations do not support this conclusion. We have found that the various self-interstitial energies converge fairly rapidly with the number of k-points. The results of our convergence

tests are in agreement with the statements of Blöchl *et al* [14]. For their best calculations of the static defect structures Blöchl *et al* [14] used 27 k-points for a cell containing $64 + 1$ atoms, which our tests indicate gives an excellent BZ integration.

The starting coordinates for the caged-interstitial calculation were obtained from the relaxed coordinates of Clark and Ackland†. We took the positions of the interstitial and neighbouring atoms from their set and added further shells of atoms at the perfect atomic positions. All the atomic positions were then relaxed. Clark and Ackland's caged interstitial can be viewed as a distorted split-$\langle 110 \rangle$ interstitial. In our calculations the caged structure relaxed to the split-$\langle 110 \rangle$ interstitial, and we could not find a separate energy minimum for the caged structure.

Our calculations show that the hexagonal self-interstitial has a formation energy of 3.31 eV (LDA) and 3.80 eV (PW91–GGA), so it is essentially degenerate with the split-$\langle 110 \rangle$ interstitial. The structure of the hexagonal interstitial is illustrated in figure 1. Below we list the relaxed atomic positions in Å for the hexagonal self-interstitial and its six nearest neighbours. The origin of coordinates is on the hexagonal interstitial, which is the centre of inversion of the structure.

1:	0.000	0.000	0.000
2:	0.684	0.684	−2.153
3:	−0.684	2.153	−0.684
4:	0.684	−2.153	0.684
5:	−0.684	−0.684	2.153
6:	2.153	−0.684	−0.684
7:	−2.153	0.684	0.684

The relaxations of the six nearest neighbours are quite significant in this structure as, for instance, the unrelaxed position of atom 2 would be (0.679, 0.679, −2.036). The hexagonal interstitial atom is sixfold coordinated, while its six neighbours are fivefold coordinated. The calculated LDA band gap for the relaxed hexagonal structure is 0.57 eV, which is slightly larger than for the split-$\langle 110 \rangle$ defect. We carried out the same tests for localized defect states mentioned above, but found no sign of them, indicating that the hexagonal interstitial has either no states in the gap or only shallow states in the gap. This appears to be a new result, which we have not seen elsewhere in the literature. The absence of states in the gap of the uncharged defect does not necessarily mean that charged states of the defect do not exist, as they could be induced by further lattice distortions. We have performed calculations for charged defects by adding and subtracting one or two electrons from the supercell and including a uniform neutralizing background charge density, but we have not been able to locate localized electronic states for either the hexagonal or split-$\langle 110 \rangle$ defects. This result is in conflict with a number of earlier studies [9–12, 16, 20, 21] which report LDA calculations on various charged states of the hexagonal self-interstitial. As remarked earlier, the band gaps of the defect structures are sensitive to the size of the simulation cell, and we believe that our use of a larger cell is the reason for this difference from earlier results.

Calculations for the tetrahedral interstitial were performed with the constraint that the structure retained the tetrahedral symmetry. We found the tetrahedral self-interstitial to have a formation energy of 3.43 eV (LDA) and 4.07 eV (PW91–GGA), which are slightly higher than the energies of the hexagonal and split-$\langle 110 \rangle$ interstitials. The structure of the tetrahedral interstitial is illustrated in figure 1. Below we list the relaxed atomic positions in Å for the

† We are grateful to the authors of reference [15] for sending their atomic coordinates for the caged interstitial to us.

tetragonal self-interstitial and its four nearest neighbours, with the origin of coordinates on the tetrahedral interstitial.

1:	0.000	0.000	0.000
2:	−1.410	1.410	−1.410
3:	−1.410	−1.410	1.410
4:	1.410	1.410	1.410
5:	1.410	−1.410	−1.410

The relaxations of the nearest neighbours for the tetrahedral self-interstitial are a little smaller than for the hexagonal structure as, for instance, the unrelaxed position of atom 2 of the tetrahedral structure would be (−1.357, 1.357, −1.357). The tetrahedral interstitial atom is fourfold coordinated while its four neighbours are fivefold coordinated. In common with earlier studies we find that the tetrahedral position is not a minimum of the energy surface because it is unstable, for example, with respect to a small displacement towards the hexagonal positions, e.g. along the $[\bar{1}\bar{1}\bar{1}]$ direction for the above atomic coordinates. The fully relaxed tetrahedral interstitial has states within the fundamental band gap.

The split-$\langle 100 \rangle$ interstitial is calculated to have a formation energy of 4.56 eV (LDA) and 4.98 eV (PW91–GGA). The bond length is 2.19 Å, indicating that it is significantly compressed compared with the bulk bonds. The split-$\langle 100 \rangle$ interstitial has electronic states within the fundamental band gap. We have found that there is no stable split-$\langle 111 \rangle$ interstitial.

The bond-centred interstitial is calculated to have a large formation energy of 4.78 eV (LDA) and 5.39 eV (PW91–GGA). In the bond-centred configuration an atom is inserted in the middle of a bond, and the distance to the nearest neighbours after relaxation is calculated to be 2.25 Å. The energy of the bond-centred defect is lowered if the central atom is moved off the bond centre in directions perpendicular to the bond. Therefore it is not a minimum in the energy or a first-order saddle point, but is some higher-order saddle point. This means that it is unlikely to be involved in the self-diffusion process. The bond-centred interstitial has electronic states within the fundamental band gap.

As the split-$\langle 100 \rangle$ and bond-centred interstitial defects have such large formation energies we will not consider them further here.

4. Migration energies

The first stage of the investigation has been to determine the equilibrium structures and energies of some possible interstitial defect structures. We now investigate the possible diffusion paths between the low-energy structures. These calculations are performed at zero temperature and are complementary to finite-temperature calculations. Molecular dynamics (MD) simulations using first-principles calculations of the atomic forces have been performed, including two studies of self-diffusion in silicon [14, 15]. However, first-principles MD studies of self-diffusion in silicon are only practical at temperatures close to the melting point, since at lower temperatures the diffusion events are too rare to be observed on the timescale of the MD simulation. An alternative method was employed by Milman *et al* [22] to study Li, Na, and K ions diffusing in silicon, which involves thermodynamic integration along the diffusion path. However, to use this method one has to know the diffusion path beforehand.

We have developed a technique for locating saddle points which is simple and robust and which is described in the appendix. This method finds a saddle point at which the forces on the atoms are zero, although without extensive further computation one cannot determine whether the order of the saddle point is one or greater than one. Also, one cannot rule out

the possibility that some lower-energy path exists, and therefore our results are upper bounds to the true barrier heights. We have used our method to locate saddle points on the path between the hexagonal and split-⟨110⟩ configurations, and between neighbouring hexagonal configurations. Our coordinates for the saddle point between the hexagonal and split-⟨110⟩ interstitial are given below with the origin on an atomic site of the crystal. Atom 1 will become the hexagonal interstitial, while atoms 1 and 2 will form the split-⟨110⟩ interstitial and atoms 3 to 6 are the near neighbours.

1:	−1.094	−1.094	0.950
2:	0.498	0.498	0.298
3:	−1.387	−1.387	−1.351
4:	−1.343	1.466	1.449
5:	1.449	1.449	−1.527
6:	1.466	−1.387	1.450

The bond between atoms 1 and 2 is 2.25 Å long, i.e., slightly compressed from the bulk bond length. The energy of the saddle-point configuration is 0.15 eV (LDA) and 0.20 eV (PW91–GGA) above those of the essentially degenerate hexagonal and split-⟨110⟩ interstitials†. An earlier LDA estimate of this barrier obtained using first-principles calculations on a 16 + 1 atom cell but without full relaxation to the saddle point gave 0.4 ± 0.4 eV [11], whereas a more modern LDA calculation gave 0.15 eV [16], in agreement with our value.

We have given the coordinates of the saddle-point and split-⟨110⟩ structures with respect to one origin of coordinates, whereas the coordinates of the hexagonal interstitial were given with respect to another origin of coordinates on the interstitial atom. The nature of the diffusive path in configuration space is more clearly seen if we use the same origin of coordinates for each of the structures. This can be achieved by adding the vector (−0.679, −0.679, +2.036) to the hexagonal coordinates. From this we can see that the relaxation of the near neighbours is considerably smaller than the movements of the two atoms which form the split-⟨110⟩ defect. The path can therefore be illustrated by the coordinates of these two atoms. For the split-⟨110⟩ we have:

1:	−0.862	−0.862	0.787
2:	0.862	0.862	0.787

while we have at the saddle point:

1:	−1.094	−1.094	0.950
2:	0.498	0.498	0.298

and at the hexagonal configuration:

1:	−0.679	−0.679	2.036
2:	0.005	0.005	−0.117

These sets of coordinates clearly show that the minimum-energy diffusive path is strongly curved in the configuration space.

We have not been able to locate a low-energy path for a diffusive jump between split-⟨110⟩ defects centred at different sites. We have, however, found a very low-energy path for diffusion

† Because the energy difference between the hexagonal and split-⟨110⟩ interstitial defects is of order the estimated convergence error of our calculations we quote the barrier height with respect to the average of the two defect energies.

between neighbouring hexagonal sites. We consider a diffusive jump between a hexagonal site at the origin of coordinates and one of its neighbouring hexagonal sites at $(-0.25, -0.25, 0)a$, where a is the cubic lattice constant of 5.429 Å. This mechanism is essentially a single-atom jump, i.e., only the interstitial atom itself moves by a large amount during the diffusive process. Using units of ångströms, the hexagonal interstitial is initially at

$$1: \quad 0.000 \quad 0.000 \quad 0.000$$

while at the saddle point:

$$1: \quad -0.725 \quad -0.725 \quad -0.451$$

and at the neighbouring hexagonal site:

$$1: \quad -1.357 \quad -1.357 \quad 0.000$$

The diffusive path is strongly curved. At the saddle point the diffusing atom is quite close to the tetrahedral site, which is at $(-0.679, -0.679, -0.679)$ Å. The energy of the saddle-point structure is 0.03 eV (LDA) and 0.18 eV (PW91–GGA) higher than that of the hexagonal interstitial.

5. The concerted exchange mechanism

There are many possible mechanisms for self-diffusion which do not involve self-interstitial atoms. It has long been recognized that self-diffusion can occur without the involvement of intrinsic defects if nearby atoms of the perfect lattice exchange with one another. Pandey [23] proposed a specific path for such an exchange in diamond- and zinc-blende-structure semiconductors and he performed first-principles calculations of a similar type to those reported here for the case of silicon. In Pandey's concerted exchange mechanism a pair of nearest-neighbour atoms interchange via a path involving a complicated three-dimensional motion which allows the atoms to avoid large barriers. Pandey described the motion of the atoms during the exchange process in terms of two angles, θ and ϕ [23]. Consider the concerted exchange of two atoms joined by a bond in the [111] direction of the diamond lattice. The angle θ represents a rotation of the pair of atoms about their bond centre in the $(1\bar{1}0)$ plane, while the angle ϕ is an independent rotation about the original direction of the bond [23]. Along the concerted exchange path the bond direction is

$$\begin{bmatrix} \cos\theta - \sqrt{\frac{1}{2}}\sin\theta\cos\phi + \sqrt{\frac{3}{2}}\sin\theta\sin\phi \\ \cos\theta + \sqrt{2}\sin\theta\cos\phi \\ \cos\theta - \sqrt{\frac{1}{2}}\sin\theta\cos\phi - \sqrt{\frac{3}{2}}\sin\theta\sin\phi \end{bmatrix}.$$

The end points correspond to $(\theta = 0^\circ, \phi = 0^\circ)$ and $(\theta = 180^\circ, \phi = 60^\circ)$, in which cases the bond vectors are along the [111] and $[\bar{1}\bar{1}\bar{1}]$ directions, respectively. The saddle point of the exchange is at $(\theta = 90^\circ, \phi = 30^\circ)$, in which case the bond points in the $[01\bar{1}]$ direction.

The initial unrelaxed atomic coordinates for the saddle-point configuration were obtained from those of the perfect diamond structure by rotating a pair of nearest-neighbour atoms joined by a bond in the [111] direction about the centre of the bond so that the new bond points in the $[01\bar{1}]$ direction. The positions of all the atoms were then relaxed with the constraint that the bond continued to point in the $[01\bar{1}]$ direction. The structure of the saddle-point configuration is illustrated in figure 1.

Our calculated coordinates for the saddle point of the concerted exchange are given below in units of Å with the origin on the middle of the bond between the two exchanging atoms, which is the centre of inversion of the structure. The two exchanging atoms are numbered 1 and 2, while atoms 3 to 8 are the near neighbours.

1:	0.000	−0.759	0.759
2:	0.000	0.759	−0.759
3:	−0.592	1.907	1.907
4:	0.592	1.907	1.907
5:	−1.980	2.069	0.831
6:	1.980	0.831	2.069
7:	1.980	−0.831	2.069
8:	−1.980	−2.069	0.831

The relaxed bond length between the exchanging atoms at the saddle point is 2.15 Å, which is somewhat shorter than the bulk bond length and is in reasonable agreement with the value of 2.19 Å given by Kaxiras and Pandey [24]. At the saddle point two bonds are broken, so the exchanging atoms and two other atoms are threefold coordinated.

The original LDA calculations of Pandey [23] gave a saddle-point energy for a 54-atom supercell of 4.5 eV, which he estimated would be reduced by 0.2 eV by relaxations in a very large supercell. Our LDA value for the 54-atom cell of 4.71 eV is in good agreement with Pandey's value. For our 128-atom cell we obtained saddle-point energies of 4.58 eV (LDA) and 4.93 eV (PW91–GGA).

6. Conclusions

Both the LDA and PW91–GGA density functionals give the lowest-energy self-interstitial defects as the split-⟨110⟩ and hexagonal configurations, with the tetrahedral interstitial being slightly higher in energy. We predict the split-⟨110⟩ and hexagonal interstitials to be essentially degenerate in energy to within the convergence of the calculations using either of the density functionals. Our LDA energy for the split-⟨110⟩ interstitial is in good agreement with that of reference [13], but is significantly different from those of references [15] and [16]. Our results do not support the existence of an energy minimum for a 'caged'-interstitial structure [15].

We have not found any well-localized electronic states within the fundamental band gap of the low-energy hexagonal and split-⟨110⟩ configurations, although we cannot rule out the possibility that shallow states exist within the gap. It is an open question whether this is a consequence of the famous 'band-gap problem' of DFT [25] or whether it is a correct physical result.

We calculate the activation energy for Pandey's concerted exchange to be 4.58 eV (LDA) and 4.93 eV (PW91–GGA). Although these values are in reasonable agreement with experimental measurements of the activation energy for self-diffusion we have found the sum of formation and migration energies for interstitial self-diffusion to be 1 eV smaller, and therefore we expect self-interstitial diffusion to dominate.

We have found a low-energy path for diffusive jumps from the split-⟨110⟩ to hexagonal self-interstitial sites with a migration energy of 0.15 eV (LDA) and 0.20 eV (PW91–GGA). A path also exists for diffusive jumps between neighbouring hexagonal sites with a migration energy of 0.03 eV (LDA) and 0.18 eV (PW91–GGA). The existence of two almost degenerate low-energy defect structures joined by two paths with low migration energies may be responsible for the large prefactor of the self-interstitial contribution to the self-diffusivity in silicon. The

low-migration-energy paths include both an 'interstitial' mechanism in which an atom diffuses via the hexagonal sites without exchanging with the host lattice and 'interstitialcy' mechanisms which involve an exchange of interstitial atoms with the host lattice via the hexagonal–split-$\langle 110 \rangle$ path.

Combining our results for the defect formation and migration energies yields activation energies for interstitial-mediated self-diffusion of about 3.5 eV (LDA) and 3.8 eV (PW91–GGA). These values are both significantly lower than the best current experimental estimate of 4.84 eV [5]. There are a number of possible reasons for this discrepancy:

(1) It would be more appropriate to compare our calculated results for the activation energy with measurements of the equilibrium self-diffusion constant at low temperatures, but unfortunately such experimental data do not exist.
(2) The experimental data for the self-diffusivity at high temperatures are well established, but the division into the contributions from self-interstitial and vacancy diffusion presented in references [5] and [6] may be less well founded.
(3) The self-diffusion may be slowed by trapping of self-interstitials, for instance by the carbon atoms which are normally present in silicon.
(4) The LDA and PW91–GGA functionals may be insufficient to give accurate values of the formation and/or migration energies. As the LDA and PW91–GGA formation energies differ by about 0.5 eV it is clear that the role of electron correlation is important and a recent fixed-node diffusion quantum Monte Carlo study has yielded even larger formation energies [27].

A complete understanding of self-diffusion in silicon is clearly some distance away. However, we hope that by clarifying the predictions of DFT calculations we have made some progress towards achieving this goal.

Acknowledgments

We thank Graeme Ackland, Patrick Briddon, Malcolm Heggie, Chris Van de Walle and Paul Kent for helpful discussions. Financial support was provided by the Engineering and Physical Sciences Research Council (UK).

Appendix. Location of saddle points

Consider a system of N atoms whose atomic positions $\{\boldsymbol{r}_i\}$ are described by the $3N$-dimensional vector $\boldsymbol{r}$. At finite temperatures the set of atomic positions will remain in the vicinity of $\boldsymbol{r}$ for some time until a diffusive jump occurs to a nearby point $\boldsymbol{r}'$. In general there will be many possible paths in configuration space between the start and end points of a diffusive jump; however, those for which the maximum energy along the path is smallest contribute the most to the diffusion. The barrier height is the smallest maximum energy along any path joining two locally stable configurations. The energy of the system at this point in the configuration space is a saddle point where the forces on the atoms are zero. Sometimes the most likely position of the saddle point can be deduced from symmetry arguments. In other cases one can guess the diffusion path and calculate the energy along the path to determine the saddle point. Neither of these methods works for all cases, and a more general way to identify saddle points is desirable.

Mathematically, a minimum of a function occurs at a point $\boldsymbol{r}^*$ if the gradient of the function is zero at $\boldsymbol{r}^*$ and all of the eigenvalues of the matrix of second derivatives (the Hessian) are

greater than zero. A first-order saddle point occurs when the gradient of the function is zero at $\boldsymbol{r}^*$ and the Hessian has one negative eigenvalue, with the rest being greater than zero. Both minima and saddle points can be located using methods based on calculating the Hessian but, being a $3N \times 3N$ matrix, it is expensive to evaluate and we would like to avoid such a calculation. Efficient methods have been developed for locating minima requiring only the value and first derivative of the function, such as the conjugate gradients algorithm. The problem of finding saddle points is more difficult. We have devised a simple method of the 'ridge-walking' [26] type for locating saddle points, which uses only the energies and forces. The idea is as follows: we find two points in the configuration space, $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, on either side of the ridge on which the saddle point is located, and we then move these configurations in the direction of the projection of the forces on the atoms in the plane perpendicular to $\boldsymbol{r}_2 - \boldsymbol{r}_1$.

To locate the saddle point we proceed as follows. First we guess a path between the initial and final configurations, $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$. We have chosen the path $\boldsymbol{P} = (1-\alpha)\boldsymbol{R}_1 + \alpha\boldsymbol{R}_2$, where $\alpha = 0, 1$ denote the end points. Next we locate two points, $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, which bracket the point of maximum energy along the path. The initial guess for the position of the saddle point is $\boldsymbol{r}^* = (\boldsymbol{r}_1 + \boldsymbol{r}_2)/2$. We now start an iterative search for the saddle point:

(i) Calculate the energy and forces on the atoms for the configurations $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, and $\boldsymbol{r}^*$.

(ii) Update the positions of the atoms according to:

$$\begin{aligned} \boldsymbol{r}_1' &= \boldsymbol{r}_1 + \lambda\left[\boldsymbol{f}_1 - \frac{(\boldsymbol{f}_1 \cdot \boldsymbol{S})\boldsymbol{S}}{\boldsymbol{S}\cdot\boldsymbol{S}}\right] \\ \boldsymbol{r}_2' &= \boldsymbol{r}_2 + \lambda\left[\boldsymbol{f}_2 - \frac{(\boldsymbol{f}_2 \cdot \boldsymbol{S})\boldsymbol{S}}{\boldsymbol{S}\cdot\boldsymbol{S}}\right] \\ \boldsymbol{r}^{*'} &= \frac{(\boldsymbol{r}_1' + \boldsymbol{r}_2')}{2} \end{aligned} \tag{A.1}$$

where $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are the $3N$-dimensional vectors of the forces on the atoms in configurations $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, respectively, $\boldsymbol{S} = \boldsymbol{r}_2 - \boldsymbol{r}_1$, and λ is a parameter chosen by trial and error. By construction $\boldsymbol{S}\cdot(\boldsymbol{r}_2' - \boldsymbol{r}_2)$ and $\boldsymbol{S}\cdot(\boldsymbol{r}_1' - \boldsymbol{r}_1)$ are zero, i.e., the displacements of the atoms are perpendicular to $\boldsymbol{S}$.

(iii) If the forces on the atoms are small enough we stop the procedure as the saddle point has been located. Otherwise we set $\boldsymbol{r}_1 = \boldsymbol{r}_1'$, $\boldsymbol{r}_2 = \boldsymbol{r}_2'$, and $\boldsymbol{r}^* = \boldsymbol{r}^{*'}$, and go back to (i).

If at any time during the procedure the energy of $\boldsymbol{r}^*$ is not greater than the energies of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, then the maximum along the path has not been bracketed and new positions must be used. In practice this problem has not arisen. Occasionally we reduce the length of the vector $\boldsymbol{S}$ by making shifts such as $\boldsymbol{r}_2' = \boldsymbol{r}_2 - \beta\boldsymbol{S}$ (and similarly for $\boldsymbol{r}_1$), where β is a parameter, being careful to ensure that $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ still bracket the maximum along the path. This helps to converge on the true saddle point. This method tends to move away from maxima, and minima are easily distinguished because at a minimum the energy of $\boldsymbol{r}^*$ is less than the energies of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. The order of the saddle point obtained is not necessarily equal to unity, and this would have to be determined by separate calculations.

References

[1] Fahey P M, Griffin P B and Plummer J D 1989 *Rev. Mod. Phys.* **61** 289

[2] Eaglesham D 1995 *Phys. World* **8** (11) 41

[3] Stolk P A, Gossmann H-J, Eaglesham D, Jacobson D C, Rafferty C S, Gilmer G H, Jaraiz M, Poate J M, Luftman H S and Haynes T E 1997 *J. Appl. Phys.* **81** 393

[4] Frank W, Gösele U, Mehrer H and Seeger A 1985 *Diffusion in Crystalline Solids* ed G E Murch and A S Nowick (Orlando, FL: Academic) p 64
[5] Gösele U, Plössl A and Tan T Y 1996 *Process Physics and Modeling in Semiconductor Technology* ed G R Srinivasan, C S Murthy and S T Dunham (Pennington, NJ: Electrochemical Society) p 309
[6] Bracht H, Haller E E and Clark-Phelps R 1998 *Phys. Rev. Lett.* **81** 393
[7] We used the LDA of
Perdew J P and Zunger A 1981 *Phys. Rev.* B **23** 5048
who parametrized the diffusion quantum Monte Carlo data of
Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566
[8] Perdew J P 1991 *Electronic Structure of Solids '91* ed P Ziesche and H Eschrig (Berlin: Akademie) p 11
[9] Bar-Yam Y and Joannopoulos J D 1984 *Phys. Rev. Lett.* **52** 1129
[10] Bar-Yam Y and Joannopoulos J D 1984 *Phys. Rev.* B **30** 1844
[11] Bar-Yam Y and Joannopoulos J D 1984 *Phys. Rev.* B **30** 2216
[12] Car R, Kelly P J, Oshiyama A and Pantelides S T 1984 *Phys. Rev. Lett.* **52** 1814
[13] Car R, Blöchl P E and Smargiassi E 1992 *Mater. Sci. Forum* **83–87** 443
[14] Blöchl P E, Smargiassi E, Car R, Laks D B, Andreoni W and Pantelides S T 1993 *Phys. Rev. Lett.* **70** 2435
[15] Clark S J and Ackland G J 1997 *Phys. Rev.* B **56** 47
[16] Lee W-C, Lee S-G and Chang K J 1998 *J. Phys.: Condens. Matter* **10** 995
[17] Troullier N and Martins J L 1991 *Phys. Rev.* B **43** 1993
[18] Pickett W E 1989 *Comput. Phys. Rep.* **9** 115
[19] Monkhorst H J and Pack J D 1976 *Phys. Rev.* B **13** 5189
[20] Baraff G A, Schlüter M and Allan G 1983 *Phys. Rev. Lett.* **50** 739
[21] Kelly P J and Car R 1992 *Phys. Rev.* B **45** 6543
[22] Milman V, Payne M C, Heine V, Needs R J, Lin J S and Lee M H 1993 *Phys. Rev. Lett.* **70** 2928
[23] Pandey K C 1986 *Phys. Rev. Lett.* **57** 2287
[24] Kaxiras E and Pandey K C 1993 *Phys. Rev.* B **47** 1659
[25] Perdew J P and Levy M 1983 *Phys. Rev. Lett.* **51** 1884
Sham L J and Schlüter M 1983 *Phys. Rev. Lett.* **51** 1888
[26] Ionova I V and Carter E A 1993 *J. Chem. Phys.* **98** 6377
[27] Leung W-K, Needs R J, Rajagopal G, Itoh S and Ihara S *Phys. Rev. Lett.* **83** 2351