C. 2

ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Lattice Vacancies and Diffusion Processes in Crystalline Silicon: a First-Principles Molecular Dynamics Study

Thesis submitted for the degree
of Doctor Philosophiæ

CANDIDATE:

Enrico Smargiassi

SUPERVISOR:

Prof. Roberto Car

Academic Year 1989–90

TRIESTE

SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
I STUDI AVANZATI
TRIESTE
Strada Costiera 11

Ad Anna

Lattice Vacancies and Diffusion Processes in Crystalline Silicon: a First-Principles Molecular Dynamics Study

Thesis submitted for the degree
of Doctor Philosophiæ

CANDIDATE:

Enrico Smargiassi

SUPERVISOR:

Prof. Roberto Car

Academic Year 1989–90

Contents

Chapter 1

Introduction

The concept of perfectly periodic crystals, although useful, is an idealization. In nature, in any crystal several defects are always found. Besides the defects generated during the growing process, and the extrinsic defects (i. e. impurities) both introduced as dopant in man-made crystals and unwillingly penetrated as dirt, the existence of intrinsic defects (i. e. involving only host crystal atoms) is a simple consequence of the laws of thermodynamics: at nonzero temperature some disorder must be allowed.

The presence of defects affects a number of different physical properties, as examples we can cite:

- *Charge transport and diffusion:* the presence of impurities allows to modulate the resistivity of semiconductors. In metals, they govern the low temperature resistivity.
- *Levels in the energy gap:* the introduction of additional levels in the gap of insulators lowers the threshold for thermal and optical excitations: this allows, for instance, many insulators to be coloured.
- *Paramagnetism:* the defect levels in the gap can be singly occupied, allowing the material to be paramagnetic.
- *Diffusion, precipitation, etc.:* the defects provide a way by which atoms in solids can move. This is the issue of concern in this work.

Diffusion in semiconductors is an important field of research in modern solid state physics. It involves the knowledge of many properties of solids,

both electronic and vibrational. It is also of major importance in the industrial processing of electronic devices: annealing processes, dopant impurity diffusion, etc..

In this thesis we will focus on self-diffusion in crystalline silicon. Self-diffusion, i. e. the process by which the host crystal atoms move, is one of the most basic diffusion processes. It is also a necessary step to understand substitutional impurity diffusion, which is a very relevant process technologically.

A number of possible mechanisms may be considered for self-diffusion [1]: *i)* a vacancy mechanism, in which atoms diffuse by means of jumps into an empty site [2], *ii)* a divacancy mechanism [3,4], in which two bonded vacancies are considered (more complex aggregates of defects have been also proposed [5]), *iii)* an interstitial mechanism [5], in which an off-site atom wanders in the crystal (if it exchanges with an in-site atom it takes the name of interstitialcy mechanism), *iv)* intrinsic mechanisms (i. e., without defects), such as the "concerted exchange" model [6] in which two atoms directly interchange their positions. Of course more than one mechanism may be effective simultaneously.

From an experimental point of view, the situation for silicon is far from being clear. The self-diffusion coefficient data show a considerable spread.

To estimate theoretically the self-diffusion coefficient D at a given temperature due to a specific mechanism, e. g. a simple vacancy jump, it is necessary to evaluate:

- the vacancy concentration c, proportional to $e^{-\frac{G_f}{k_B T}}$, G_f being the Gibbs free energy of formation of the vacancy;
- the vacancy jump frequency, proportional to $e^{-\frac{G_m}{k_B T}}$, G_m being the Gibbs free energy of migration.

The coefficient D then will be proportional to the product of these two factors.

The evaluation for a realistic model semiconductor of quantities like G_f and G_m is very difficult and, as a matter of fact, no such calculation exists to date. Only the defect energetics at $T = 0$ has been worked out.

A possible way to solve the problem consists in performing Molecular Dynamics (MD) simulations, in which finite temperature effects are fully

taken into account. However, MD relies on the assumption that a satisfactory interatomic potential can be found.

A standard approach consists in modelling the interatomic potential in terms of empirical few-body potentials, such as the two-body Lennard-Jones potential, which works reasonably well for rare gases. This approach is affected by severe difficulties in the case of covalently bonded materials, where subtle chemical effects depending on the electronic ground state are very important. It seems very difficult to model in terms of fixed empirical potentials processes like bond breaking and bond formation which may occur as a consequence of the atomic motion.

Actually, to simulate silicon a wealth of potentials have been introduced in the last few years but none of them seems capable of an accurate description of all the different physical situations [7].

The First Principles Molecular Dynamics scheme [8] allows to solve the problems illustrated above. In this scheme, the forces on the ions are directly obtained by the electronic configuration in the instantaneous ground state.

In this thesis, we apply this method to study the silicon vacancy and its relevance to self diffusion. Both the problem of the vacancy formation and of its migration can be treated by first-principles MD: with specific methods it is possible to extract from the MD runs the value of G_f, so computing c, while the vacancy induced atomic diffusivity can be obtained by direct simulation. In both cases we take fully into account entropic, anharmonic and dynamical effects.

The results for the vacancy migration indicate that dynamical effects influence the high temperature behaviour of the vacancy at all temperatures studied. In particular we noted that, at $T = 1200 - 1400\ ^\circ K$, the vacancy tends to move in a correlated way: in fact it jumps several times along the same direction, instead of performing an uncorrelated random walk. When the temperature is $1600\ ^\circ K$, close to the experimental melting point of $1685\ ^\circ K$, the vacancy motion assumes a different character. Different migration mechanisms appear, such as second-neighbours jumps or concerted jumps of several atoms. The self-diffusion coefficient increases due to the onset of these mechanisms, similar to what is experimentally observed.

The formation free energy has been computed using two different methods: the local harmonic approximation [9,10] and a fully anharmonic cal-

culation using MD. Both methods yield a large formation entropy. In particular, the first method gives $S_f = 5\ k_B$. The MD calculations are still in progress; but preliminary results obtained at 1000 $^\circ K$ indicate $S_f \sim 6 - 10\ k_B$. These values are within the experimental range for S_f [11].

The self-diffusion coefficient computed from our data is still an order-of-magnitude estimate. It compares well with the results of the experiments, which are affected by a similar uncertainty.

Such a large formation entropy is a typical feature of silicon, since in metals the reported values lie in the range $1 - 4\ k_B$. This indicates that the presence of a vacancy in the silicon lattice entails a substantial local softening of the lattice.

In conclusion, our calculated vacancy concentration agrees with the results of positron annihilation experiments [11]. The self-diffusion coefficient also seems to be in good agreement with the experimental data. This indicates that the vacancy mechanism is a likely candidate for explaining self-diffusion in silicon.

Of course this does not rule out the possibility that other mechanisms may be active at the same time; in particular, calculations based on the local harmonic approximation indicate that the interstitial mechanism is another likely candidate. Further work is needed in order to elucidate this question, in particular accurate fully anharmonic calculations.

This thesis is organized as follows:

- In chapter 2, our present understanding of the silicon vacancy and of the self-diffusion process is shortly reviewed.
- In chapter 3 the First Principles Molecular Dynamics method is illustrated and discussed.
- In chapter 4 the method used to compute the vacancy concentration is exposed.
- In chapter 5 the results of calculations of the vacancy concentration and of the vacancy induced self-diffusion are exposed and compared with the available experimental data.
- In the last chapter we present our conclusions.

Chapter 2

Self-Diffusion and Lattice Vacancies

2.1 Self-diffusion in silicon

2.1.1 Thermodynamical aspects

From a macroscopic point of view, the diffusion phenomena are described by Fick's laws:

$$\begin{aligned} \mathbf{j} &= -\mathbf{D}\cdot\nabla c \\ \frac{\partial c}{\partial t} &= -\nabla\cdot\mathbf{j} = \nabla\cdot(\mathbf{D}\cdot\nabla c) = D\nabla^2 c \end{aligned} \tag{2.1}$$

where c is the concentration of the diffusing particles and $\mathbf{j}$ their flux. The coefficient D is called "diffusion coefficient" and, in principle, is a position dependent second-rank tensor. However, for systems with a sufficient degree of symmetry, such as cubic materials, D is a multiple of the identity and therefore we shall consider it as a scalar. The last equality in the equations 2.1 holds only for homogeneous systems of this kind, which are of concern here.

An early analysis of Einstein, based on a random walk model [12], links

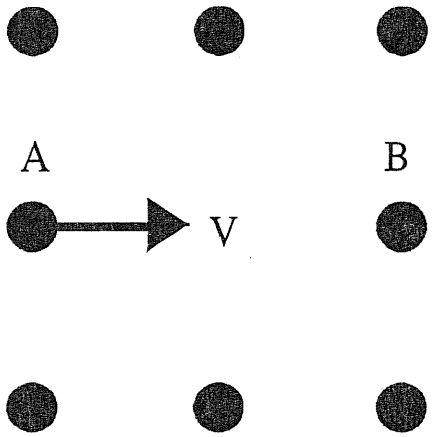

Figure 2.1: Simple vacancy jump in a simple cubic lattice.

D to the mean square displacement of the particles[1]:

$$D = \frac{\langle x^2(t) \rangle}{6t} = \frac{f\Gamma d^2}{6} \tag{2.2}$$

where the angular brackets indicate average over all the atoms and the second equality refers to a lattice: Γ is the jump frequency from one site to another and d is the jump amplitude. If more than one kind of process are present, and they are not correlated, D is the sum of all the individual coefficients.

The factor f, called the *correlation factor*, is a manifestation of the fact that, even if the defect follows a random walk, the atoms displaced in this way do not need to move randomly [14]. In fact, let us suppose that, as in figure 2.1 a vacancy is moving in a simple cubic lattice via simple jumps (first neighbour atoms jumping into the vacant site at separate times). Let us suppose that the atom A jumps into the vacancy V, by moving in the positive x direction. This atom has probability $1/6$ to jump back into the vacancy, since the defect is supposed to be randomly walking. But for this atom to jump again in the positive x direction, it is necessary to wait

[1]For systems having less than cubic symmetry the factor 6 in the right-hand side should be changed [13,14].

until the vacancy will be at the site B: this of course happens with a much smaller probability, and so the assumption of equivalence of each jumping direction, necessary to have a random walk, is not verified for the atoms.

It is clear that in the cases in which the defect coincides with the diffusing atom, for example in the interstitial diffusion, or for diffusion mechanisms not involving defects, $f = 1$. In all the other cases, f may be computed with the formula [14]

$$f = 1 + 2 \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \langle r_i r_{i+j} \rangle / \sum_{i=1}^{n} \langle r_i^2 \rangle \tag{2.3}$$

where $\mathbf{r}_i$ is the displacement of an atom due to its i-th jump, and n is the number of jumps considered.

For self diffusion in elemental materials f is (neglecting isotope effects) a mechanism dependent geometrical factor [14]. Assuming a defect mechanism in which defects move by a random sequence of uncorrelated simple jumps, f has been exactly calculated for several lattices, where it turns out to be less than 1 [15,16] ($f = 1/2$ for the vacancy in the diamond lattice).

The assumption of random defect jumps is not always justified: since the jumping particle has a peculiar dynamical situation, it probably has a different jump probability, generally higher, than the mean. In the case of the vacancy, for example, the possibility of an immediate return jump may be large, so reducing the diffusion coefficient, but it is even possible an increase of the probability for a forward jump. Such dynamical effects should be included in the factor f. This can be done only if the true dynamics of the system is known.

Let us analyse the equation 2.2 in the case of self-diffusion. We can write $\Gamma = \varphi c$, where φ is the frequency of the process inducing self-diffusion (e. g. the jump of an atom into a vacancy) and c is the concentration of sites available for it. For intrinsic self-diffusion, $c = 1$; for defect-mediated processes, we know from Statistical Mechanics that the defect concentration is [1]

$$c = \mathrm{e}^{-G_f/k_B T} = \mathrm{e}^{S_f/k_B} \mathrm{e}^{-H_f/k_B T} \tag{2.4}$$

where G_f is the Gibbs' free energy of formation, i. e. the free energy of the system with defect minus the free energy of the system without defect at constant number of particles (with similar definitions for H_f and S_f).

The frequency φ can be written as $\varphi = x\nu$. Here x is the probability for an atom to reach the saddle-point for migration [4], i. e. $x = e^{-G_m/k_BT} = e^{S_m/k_B}e^{-H_m/k_BT}$, where H_m and S_m are the enthalpy and the entropy of the saddle point configuration with respect of the minimum free energy configuration, and ν is the attempt frequency (an unknown quantity of the order of the Debye frequency of the system). Thus

$$\Gamma = \varphi c = \nu e^{(S_f+S_m)/k_B}e^{-(H_f+H_m)/k_BT} = \nu e^{-G_d/k_BT} \tag{2.5}$$

Inserting this equation into eq. 2.2 we see that diffusion is a thermally activated process with

$$D = D_0 e^{-Q/k_BT} \tag{2.6}$$

where $Q = H_f + H_m$ is the activation enthalpy and the pre-exponential factor $D_0 = \frac{1}{6}fd^2\nu e^{(S_f+S_m)/k_B}$ contains the entropic contribution. What then remains is to determine D_0 and Q from a microscopic analysis.

Exponentially-behaved phenomena like this are usually described in terms of an Arrhenius plot. This is a graph in which one plots $\ln D$ versus $1/T$. The curve defined by equation 2.6, if the values of Q and D_0 do not depend on T, is a straight line with a slope proportional to $-Q$ and which crosses the ordinate axis at $\ln D_0$.

Deviations from the linear behaviour are due either to a temperature dependence of Q and S, or to the co-existence of two or more processes with different Q and D_0. In the latter case two or more temperature ranges usually exist in which each single process is dominant: a bending or knee in the curve indicates the temperature at which the cross-over between two different regimes occurs. Even the presence of non-thermal processes, such as quantum tunneling phenomena or ionization-enhanced migration, may cause non linear behaviour.

Deviations from perfect Arrhenius behaviour are commonly found since the early studies of diffusion in metals [17], where the measurements are affected by smaller error bars than in semiconductors.

2.1.2 Experimental data

The experimental study of self-diffusion in semiconductors is not an easy task [18]. The main way to perform experiments is to diffuse radioactive isotopes in the sample at a fixed temperature and to detect their position

after some time. The results are then fitted to an appropriate solution of eq. 2.1 to extract D at this temperature.

The first studies [19,20,21] were done by directly diffusing the only isotope available: ^{31}Si, with an half-life of 2.6 hours. It is clear that during such a short time a very small distance can be covered by the tracers, so high temperatures and refined microsectioning techniques are needed. A possible alternative is the use of the stable isotope ^{30}Si, transmuting it *in situ* in the ^{31}Si isotope [22], or exploiting its resonance in the proton scattering cross section [23,24,25], provided that the natural background of ^{30}Si (a concentration of about 3% in the natural silicon) is properly subtracted. In any case, the error on the single values of D is of the order of 50%.

Another possibility is to diffuse germanium which, being chemically very similar to silicon, should yield values at least in qualitative agreement with those of self-diffusion, with the advantage of having a longer lived isotope (^{71}Ge, half-life 11.2 days).

The experimental results for intrinsic silicon are shown in Table 2.1 and in figure 2.2. We shall divide, here and later, the temperature range into three parts: "high" temperatures are intended to be over 1300 $^{\circ}K$, "intermediate" temperatures between 1300 $^{\circ}K$ and 1000 $^{\circ}K$, and "low" below 1000 $^{\circ}K$,.

The fit to eq. 2.6 of the various sets of data shows that high temperature measurements yield values of Q and D_0 *systematically* higher than the values obtained from intermediate temperature data. Measurements of diffusion of Ge in Si [33] show a similar behaviour.

For comparison, self-diffusion in Ge has $Q = 3\ eV$, which is definitely smaller than the values of Q in Si at any temperature. In metals the values for Q range from 1 to 4 eV, and only in exceptional cases the values of D_0 are greater than 10 cm^2/s, indicating a low diffusion entropy ($2 - 6\ k_B$) [34], while in Si it ranges from 8 to 16 k_B [1].

Finally, Demond *et al.* [25] showed that the values of Q and D_0 decrease when the data are fitted without including the highest T point in some intermediate T sets of data. This suggests that the difference between low and high temperature values for Q and D_0 is probably higher than what can be deduced from Table 2.1.

This latter observation stresses that an uncertainty is inserted anyway by the fitting procedure. For example, choosing for the data in ref. [19] a fitting value of 4.1 eV instead of 4.77 one get $D_0 \simeq 100\, cm^2 s^{-1}$ instead of

Reference	D_0 $(cm^2\ s^{-1})$	Q (eV)	Temperature range $(°K)$
Si self-diffusion			
Peart [19]	1800	4.77	1473–1673
Ghoshtagore [22]	1200	4.72	1451–1573
Mayer *et al.* [26]	1460	5.02	1320–1660
Fairfield and Masters [20,21]	9000	5.13	1373–1573
Sanders and Dobson [27]	5.8	4.1	1243–1343
Kalinowski and Seguin [24,28]	154	4.65	1128–1448
Demond *et al.* [25]	20	4.4	1103–1473
Hirvonen and Anttila [23]	8	4.1	1173–1373
Ge diffusion in Si			
McVay and Du Charme [29,30,31]	1500	4.7	1473–1653
Petrov *et al.* [32]	630000	5.28	1423–1623
Hettich *et al.* [33]	2505	4.97	1300–1573
Hettich *et al.* [33]	0.35	3.93	1130–1273

Table 2.1: Experimental enthalpy and pre-exponential factor for self-diffusion in silicon and for germanium diffusion in silicon.

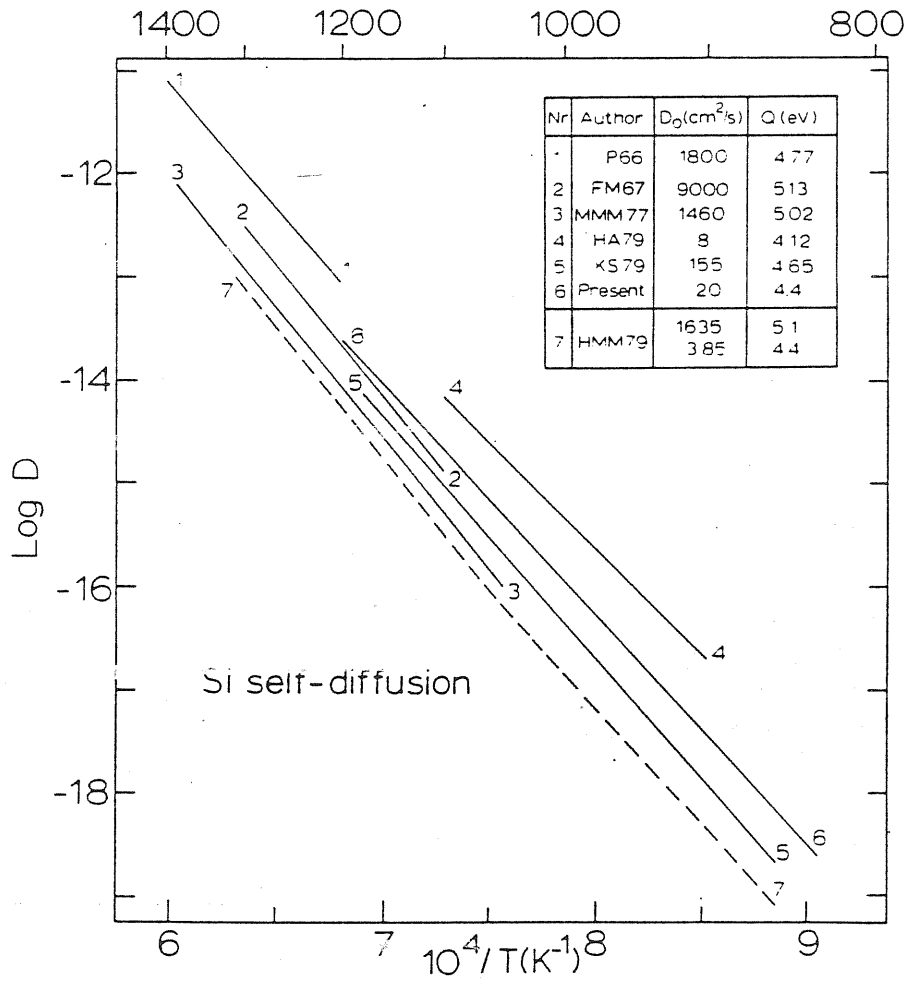

Figure 2.2: Experimental data on self-diffusion coefficient in silicon. Curve n.1: ref. [19]; n.2: ref. [20,21]; n.3: ref. [26]; n.4.: ref. [23]; n.5: ref. [24]; n.6: ref. [25]; n. 7: ref. [33] (^{71}Ge diffusion).

1800. An error of this type could explain the exceedingly high values for D_0 found in refs [20,21,32]. The fitting procedure in fact, due to the relatively small temperature range in which measurements are done, is very delicate and so even an error of 100% in evaluating D_0 is not surprising. This partly explains the large scatter in the values of D_0 at about the same T.

2.1.3 Interpretation in terms of microscopic mechanisms

A number of microscopic mechanisms has been proposed for self-diffusion in silicon. Here we limit ourselves to a short survey, referring the reader to the many existent reviews [1,4,35,36,37,38,39] for a more detailed discussion.

Each mechanism has to explain some major features of the self-diffusion data:

i) The high temperature values of D_0 are very large, implying a very large entropy of diffusion $S_d = S_f + S_m$.

ii) It seems likely that there is a curvature in the Arrhenius plot. At the lowest experimentally accessed temperatures (1100 – 1300 $°K$) the pre-exponential factor D_0 is about two orders of magnitude smaller than in the higher temperature range. The value of Q, for which a smaller error is expected, ranges from about 4 eV to about 5 eV.

A first possibility is that of an *intrinsic*, i. e. defect independent, mechanism, which could be a direct exchange or a ring exchange mechanism [1]. For a long time this has been regarded as unlikely, because of the high energy required to break six covalent bonds. However, Pandey [6] has worked out a mechanism that requires the breaking of only two bonds at a time, limiting the energy cost to less than 5 eV, a value within the experimental range.

Other conceivable mechanisms are the *interstitial* or the *interstitialcy* [37] mechanisms. Both involve off-site atoms, the difference being in the way in which the migration happens. In the interstitial mechanism, the same ion wanders through the lattice hopping from one interstitial site to another, while in the case of the interstitial mechanism it pushes away a "regular" atom which so becomes the interstitial one. Two points have to be pointed out: a number of migration paths are probably open, so increasing

somewhat the migration entropy; and both stable and saddle-point configurations may be not at all obvious, requiring very detailed investigations.

The *vacancy* mechanism is one of the strongest candidates to the role of dominant mechanism in the self-diffusion in silicon. Efforts have been made [3,40,41] in order to make plausible that the atomic relaxation and the phonon softening around the vacancy can bring the formation entropy of the vacancy to values compatible with the values for S_d reported in the previous section. Such high values of S_f are supported also by the positron annihilation experiments of Dannefaer *et al.* [11], which are independent of the self-diffusion measurements.

A possible way to obtain a high value for S_d is to choose a mechanism which strongly perturb the surrounding lattice. A *divacancy* mechanism is a possible one [3,4,19,22], but it is not clear if the divacancy formation enthalpy is low enough to allow a significant concentration to be present [11,35]. Further, it has yet to be shown that the migration of a divacancy may happen without quick dissociation into a pair of single vacancies [4,5].

The concept of strong relaxation has been extended [5] both for vacancy and for the interstitial up to the point of supposing that some sort of *local melting* may take place around these defects. The presence of such a disordered zone would imply an high entropy and an high enthalpy of migration, as experimentally observed. This phenomenon would occur in the high temperature regime; at lower temperatures "normal" vacancies and/or interstitials would be found. Due to the complexity of this mechanism, however, no detailed microscopic models have been proposed to date.

2.1.4 Theoretical tools

A standard way which is often used to compute diffusion coefficients is the so called "rate theory", also called "transition state theory" [1,13,42]. The task of the rate theory is to calculate φ in eq. 2.5, once a migration mechanism has been assumed. If c is known, this is sufficient to compute Γ and hence D according to eq. 2.2. If we assume that the system remains in thermodynamical equilibrium, before and throughout the jump, we get, by neglecting dynamical correlations and quantum effects [1,43]

$$\varphi = K\mathrm{e}^{-H_m/k_BT} \tag{2.7}$$

where K is given by

$$K = \frac{1}{2\pi} \frac{\prod_i^{3N-3} \omega_i}{\prod_i^{3N-4} \omega_i'} \tag{2.8}$$

where $\{\omega_i\}$ are the eigenfrequencies of the system around the minimum of the potential energy (here and later we shall subtract the center-of-mass coordinates from the degrees of freedom), and $\{\omega_i'\}$ are the eigenfrequencies of the system around the saddle point. The sum in the denominator excludes the mode perpendicular to the saddle plane (which has imaginary frequency).

To get a migration free energy we have to extract an entropy factor from K: if we put [43]

$$\overline{\omega}^{3N-3} = \prod_i^{3N-3} \omega_i \tag{2.9}$$

we can write

$$S_m = k_B \sum_i^{3N-4} \ln(\overline{\omega}/\omega_i') \tag{2.10}$$

obtaining $K = \frac{\overline{\omega}}{2\pi} e^{S_m/k_B}$. This choice however is by no means unique and so the value of S_m obtained in this way is to a certain extent arbitrary.

We then see that if the defect concentration, the eigenfrequencies around the minima and the saddle point, and H_m are known, this theory provides the value of the diffusion coefficient for the considered mechanism.

The rate theory has important limitations. First, it neglects dynamical effects, for instance, the possibility of immediate return jumps, or the possibility of multiple jumps. In other words, it does not provide any way to evaluate the factor f defined in section 2.1.1.

In addition the true relaxation of the system in computing H_m during the jump may be not well taken into account. In fact, the migration enthalpy H_m is usually computed allowing full relaxation of the lattice. But, while for slow events a full relaxation is expected, for high velocities of the jumping particles the surrounding atoms may have not enough time to fully relax. This may deeply change the picture obtained by $T = 0$ calculations [44].

In summary, the rate theory simply disregards the dynamical nature of the diffusion processes, i. e. one of their important characteristics. Its results are exact (if we neglect quantum effects) only in the limit of low temperatures. This is disappointing since as we have seen in the previous

sections, there are experimental indications that temperature effects play a role on the silicon self-diffusion.

A well known way to take fully into account all the dynamical effects is to simulate the true time evolution of the system by means of computer experiments: this is the Molecular Dynamics (MD) method [15,43].

In principle, the method seems to be plainly applicable to our problem. It seems sufficient to start with a perfect crystal system, heat it up and wait for the formation of a sufficient number of the the wanted defects; then their migration could be monitored, and the formula 2.2 will directly give us the sought physical quantities.

But this simple strategy cannot be practically used. Left apart considerations on the interatomic potential, on which we shall return soon, the creation rate of defects is far lower than detectable from today's simulations. In addition, in a periodically repeated MD cell point defects can be created only in the form of "Frenkel pairs", further rising the formation energy. These problems are less serious in the case of defect jump in which the probabilities, at least near the melting point, are usually much higher: using modern high-speed computers a direct simulation of defect migration is sometimes possible.

One is then forced to an indirect evaluation of the concentration of defects by computing thermodynamical quantities like G_f and using eq. 2.4. The microscopic calculation of G_f is a difficult problem of statistical mechanics. Even in this case it is possible to use the MD method; however, the direct computation of the relevant thermodynamical quantities like the free energy or the entropy is not practically possible because they are not cast in the form of equilibrium averages of functions of positions and velocities. They are written instead as configurational integrals with rapidly varying integrands which are very hard to compute [45]. A number of strategies have been proposed to overcome these problems: we shall briefly expose them in the chapter 4.

The MD approach has been successfully applied to several simple cases, namely when a satisfactory potential function can be found: mainly noble gases (Lennard-Jones systems), ionic systems and some metals. For covalent semiconductors the problem is harder: the difficulty to have a good interatomic potential has not allowed accurate calculations so far. Simplified approaches, such as valence-force schemes [7], are only useful near equilibrium position—not just our case!, and so it is not clear what should

be the choice of parameters in the case of dangling or severely distorted bonds.

An example of the difficulty in defining an adequate potential for silicon is provided by the Tersoff potential [46]. Such potential gives the right formation energy but a *negative* migration energy for the vacancy. For a review of applications of empirical potentials to defects in silicon, see e. g. the reference [7].

2.2 Lattice vacancies in silicon

2.2.1 Experimental

The first experimental studies of vacancy migration in Si were made by Watkins [47,48]. In this work the defects were generated by means of irradiation at low temperature (less than 40 $°K$) and detected via electron paramagnetic resonance (EPR). The annealing behaviour has been studied by measuring the time rate of disappearance of the EPR signal, and the use of the technique of isochronous annealing [49] allowed to extract the migration enthalpy H_m.

The observed values of H_m depend on doping. The assignment of these values to different charge states of the vacancy depends on the ordering of the gap levels introduced by the vacancy. After the discovery of the negative-U behaviour of the silicon vacancy [50,51,52] the accepted value for the migration enthalpy of the neutral vacancy is 0.45 eV.

The direct detection of the formation free energy and so of the equilibrium density of vacancies (and of other thermally activated defects) is a more difficult task. Due to the low concentration of intrinsic defects even close to the melting temperature, the physical quantities which can be affected by the presence of defects are much more altered by thermal effects. For example, since defects behave as sources or sinks of conduction electrons the electrical transport properties are sensitive to their presence; however, the defect concentration in silicon near the melting point is several order of magnitude smaller than the concentration of thermally excited electrons.

For a long time the only way to proceed consisted in heating the crystal to the desired temperature T, followed by a rapid cooling ("quench") to a temperature low enough to reduce the number of thermally excited electrons to a very small value. The reduction of the free-carrier concentration

in the sample with respect to the concentration measured in the same sample before the heat treatment may then be attributed to the trapping by defects, allowing to measure the defect concentration. It must be assumed that the number of defects does not change during the cooling or, at most, that the change is proportional to c.

Recently the refinement of the positron-annihilation technique [53] has allowed to avoid the uncertainties implied in the previous method. In the positron-annihilation experiments one exploits the fact that a positron has a lower probability of annihilation when sitting in a defect like a vacancy, which is a region of low electronic density. The consequent increase in the positron lifetime can be measured and related to the density of defects. Using a suitable model for the temperature dependence of the lifetime of positrons in the vacancy it is also possible to estimate S_f.

A few quenching measurements have been performed, we refer here to the work of Elstner and Kamprath [54], who obtained $H_f = 2.4 - 2.5\ eV$. No estimate of S_f was possible. The positron-annihilation technique has been used in a set of measures [11] in a wide temperature range (300–1523 $°K$). The reported value for H_f is $3.6 \pm 0.2\ eV$. The authors also evaluate the formation entropy of the vacancy to be as high as $6 - 10\ k_B$. This value is subject in particular to the uncertainty on the value of the dielectric constant in the neighbourhood of the vacancy: it has been estimated to be $\epsilon \sim 6$, but actually it is not known; choosing $\epsilon = 11.9$ (bulk value) the entropy would further rise by $\sim 1.4\ k_B$.

The values obtained by these two experiments are quite different. However the results obtained by quenching experiments are inconsistent also with other experimental results. In fact, together with the value of $0.45\ eV$ for H_m obtained by Watkins, and generally considered reliable, we obtain for the activation energy for self-diffusion $Q \simeq 2.9\ eV$, far lower than the values reported in the previous section (Q about $4\ eV$).

A big source of error in quenching experiments may be identified in the fact that the silicon vacancy is found to migrate even at very low temperature [47,48]. Even cooling rates as high as $10^3\ °K/s$ do not exclude the possibility of substantial annealing and clustering of defects. In particular, Bourgoin [55] argued that the defects observed in quenching experiments come from voids formed by coalesced vacancies during the cooling period. The observed H_f, according to this analysis, is the formation energy of vacancies generated by these voids. At high temperature, instead, the defects

would generate from the surface, with a different, higher, H_f.

The presently accepted values for H_f and S_f are therefore the values obtained by the positron annihilation data. From these values the vacancy concentration at melting temperature may be estimated to lie in the range $10^{-6} - 10^{-9}$.

Varotsos *et al.* [56] examined the experimental results for the vacancy in silicon under the point of view of the $cB\Omega$ theory on diffusion in solids [34]. This theory, whose results have been verified for a number of materials, assumes that the Gibbs free energy associated to the phenomenon i, where i may indicate diffusion processes, or defect formation or migration, is $G_i = c_i B\Omega$, c_i being a constant, B being the bulk modulus, and Ω the atomic volume of the crystal. Starting from this assumption, several relationships may be written for the thermodynamical quantities associated to the process i. The authors of ref. [34] found that, although this value for the formation entropy seems to be quite high, it is consistent with the $cB\Omega$ theory.

2.2.2 Theoretical studies

Modern computations on defect energetics in solids are based on refined self-consistent methods and need large availability of computer resources. In these schemes the system is thought to be not only at thermodynamical equilibrium, but also at $T = 0$. Then one carefully computes the thermodynamical functions at the critical configurations, namely the local minima and the migration saddle points, which for most mechanisms are not obvious and have to be selected by performing computations for a few most likely candidates. Finally one tries to make connections with the dynamical properties of interest.

The main results regarding the silicon vacancy are listed in table 2.2. The results of Bar-Yam and Joannopoulos [57] and those of Nichols *et al.* [58] were obtained by means of a supercell (up to 32 atoms) method, the others [59,60,61] using a self consistent Green's function method. All these authors included atomic relaxation; the relaxation of the nearest neighbours of the vacancy was estimated by self-consistent total energy minimization, while the relaxation of more distant neighbours was usually estimated by using a Keating model [62]. The Fermi level μ_f is supposed to lie at the middle of the gap for intrinsic silicon.

Reference	H_f	H_m	Q
Bar-Yam *et al.* [57]	3.6	0.5°	4.1
Nichols *et al.* [58]	3.5	0.3	3.8
Car *et al.* [59,60]			4.2
Kelly *et al.* [61]	3.92	0.27	4.19

° experimental value

Table 2.2: Theoretical enthalpies for the neutral silicon vacancy.

The formation and migration enthalpy data are directly comparable with experiments. Moreover, from section 2.1.1 we know that the experimental activation energy Q is the sum of the microscopic H_f and H_m. So we have also listed the sum of these enthalpies which have to be compared to the values of $\sim 4 - 5\ eV$ quoted in the section 2.2.1. The agreement is in general quite good.

What cannot be done in this manner is the evaluation of the contribution of a single mechanism to self-diffusion. In fact, these results do not allow to estimate the formation and migration entropies entering the pre-exponential factor D_0.

Chapter 3

First Principles Molecular Dynamics

3.1 Classical ionic dynamics via the electronic ground-state

The problem of finding the time evolution of a many-body quantum system is very difficult. In the case of a system including electrons and ions a first simplification can be achieved by the so-called Born-Oppenheimer (BO) approximation. This approximation allows to reduce the complex quantum problem of electron and ions to the much simpler case of the evolution of the ions in the potential energy surface resulting from the instantaneous electronic ground state. In many cases it is then possible to regard the ions as classical particles.

The justification for this approximation lies in the large mass difference between electrons and ions. As a consequence, the time scales of the ionic and electronic motion are well separated and the fast electrons have time to follow the slow ionic motion. If the electronic wavefunctions are in the ground state at a certain time, departures from the instantaneous ground state during the ionic evolution usually remain very small.

According to the BO approximation, it is possible to write an effective classical Hamiltonian for the ions as

$$H_I = K_I(\{\dot{\mathbf{R}}_I\}) + V_I(\{\mathbf{R}_I\};\{Z_I\}) + U(\{\mathbf{R}_I\}) \tag{3.1}$$

where $\{\mathbf{R}_I\}$, $\{\dot{\mathbf{R}}_I\}$ are the ionic coordinates and velocities, K_I is the ionic kinetic energy, V_I consists of the electrostatic interionic interaction (depending on the ionic charges $\{Z_I\}$) plus eventually the interaction with external fields acting on the ions, and $U(\{\mathbf{R}_I\})$ is the ground-state energy of a system of electrons in the presence of fixed ions at positions $\{\mathbf{R}_I\}$.

The ionic evolution is given by the Newton's equations:

$$M_I\ddot{\mathbf{R}}_I = \mathbf{F}_I = -\frac{\partial U}{\partial \mathbf{R}_I} - \frac{\partial V_I}{\partial \mathbf{R}_I}. \tag{3.2}$$

The first term in the right-hand side is the so-called Hellman-Feynman (HF) force [63]. In the following we shall refer to the dynamics generated by the forces in the equation 3.2 as the BO dynamics.

The main problem is to compute U. This is a very difficult issue. The Density Functional Theory [64,65] (DFT) provides a way of doing that, formally transforming the many body electronic problem into a self consistent single particle problem of the Hartree type.

According to the Hohenberg-Kohn theorem [64], the total energy of an inhomogeneous system of electrons having density $\rho(\mathbf{r})$ in an external potential $v_{ext}(r)$ can be written as a functional of of the density:

$$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + \int \mathrm{d}\mathbf{r}\, v_{ext}(\mathbf{r})\rho(\mathbf{r}) + \int\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}[\rho(\mathbf{r})] \tag{3.3}$$

where $T[\rho(\mathbf{r})]$ is the kinetic energy of a system of non interacting electrons with the same density $\rho(\mathbf{r})$ and $E_{xc}[\rho(\mathbf{r})]$ is the so called exchange and correlation energy. The universal functional $E_{xc}[\rho(\mathbf{r})]$ contains all the many-body difficult part of the problem. It can be written as

$$E_{xc}[\rho(\mathbf{r})] = \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\varepsilon_{xc}\left[\rho(\mathbf{r})\right] \tag{3.4}$$

where $\varepsilon_{xc}[\rho(\mathbf{r})]$ is the exchange and correlation energy functional per electron.

At the actual value of the density $\rho(\mathbf{r})$ the functional E is minimum and is equal to the exact ground-state electronic energy.

The exact form of $E_{xc}[\rho(\mathbf{r})]$ is of course unknown. The simplest and most used approximation is the so called "Local Density Approximation" (LDA) in which $E_{xc}[\rho(\mathbf{r})]$ is written as

$$E_{xc}[\rho(\mathbf{r})] = \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\varepsilon^h_{xc}\left(\rho(\mathbf{r})\right) \tag{3.5}$$

where $\varepsilon_{xc}^{h}(\rho(\mathbf{r}))$ is the exchange and correlation energy per electron of an electron gas of uniform density $\rho(\mathbf{r})$. The function $\varepsilon_{xc}^{h}(\rho(\mathbf{r}))$ has been extracted and fitted to an analytic expression [66] from accurate ground state Monte Carlo computations [67] for the homogeneous electron gas.

In the LDA each electron state is doubly (spin) degenerate. In the following, when we shall refer to electrons, we will sometimes actually mean electron states.

We are left with the problem of finding the density $\rho(\mathbf{r})$ corresponding to the minimum energy, since

$$U[\{\mathbf{R}_I\}] = \min_{\rho(\mathbf{r})} \{E[\rho(\mathbf{r}), \mathbf{R}_I]\}\,. \tag{3.6}$$

This equation defines the BO surface.

By applying the stationary condition ($\delta E/\delta\rho(\mathbf{r}) = 0$) to the functional of eq. 3.3, Kohn and Sham [68] were able to work out the so called KS equations for a set of single-particle electron orbitals ψ_i in an effective potential v_{eff}

$$\hat{H}_{KS}|\psi_i\rangle = \left\{-\frac{1}{2}\nabla^2 + v_{eff}\right\}|\psi_i\rangle = \varepsilon_i|\psi_i\rangle \tag{3.7}$$

where

$$v_{eff}(\mathbf{r}) = v_{ext}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\mathrm{d}}{\mathrm{d}\rho}\left\{\rho(\mathbf{r})\varepsilon_{xc}\left(\rho(\mathbf{r})\right)\right\}. \tag{3.8}$$

The wavefunctions are subjected to the constraint of orthonormality

$$\langle\psi_i|\psi_j\rangle = \delta_{ij}. \tag{3.9}$$

In this case, the density is given by

$$\rho(\mathbf{r}) = \sum_i f_i\langle\psi_i|\psi_i\rangle. \tag{3.10}$$

The ψ_i and ε_i are the so called KS orbitals and KS levels respectively, while the f_i are occupation numbers. The energy is

$$\begin{aligned} E_{KS}[\{\psi_i(\mathbf{r})\}] &= -\frac{1}{2}\sum_i \int \mathrm{d}\mathbf{r}\, f_i\psi_i\nabla^2\psi_i + \int \mathrm{d}\mathbf{r}\, v_{ext}(\mathbf{r})\rho(\mathbf{r}) \\ &+ \frac{1}{2}\iint \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\,\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}[\rho(\mathbf{r})]. \end{aligned} \tag{3.11}$$

The explicit expression of the formulae used in our calculations is given in the appendix B.

3.2 First-Principles Molecular Dynamics

The FPMD method aims at approximating the parametric evolution of eq. 3.2 through the classical dynamical evolution of the electronic and ionic variables generated by the lagrangian [8]

$$\begin{aligned} \mathcal{L} &= \sum_i \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle + K_I(\{\dot{\mathbf{R}}_I\}) - V_I(\{\mathbf{R}_I\}) - E_{KS}\left[\{\psi_i\}, \mathbf{R}_I\right] \\ &+ \sum_{i,j} \Lambda_{ij} \left(\langle \psi_i | \psi_j \rangle - \delta_{ij} \right) \end{aligned} \tag{3.12}$$

where the ψ_i are regarded as classical fields, M_I are the ionic masses, μ_i are mass-like parameters of appropriate units, and the last term accounts for the orthonormality constraint. In most cases the electronic masses are chosen to be equal: $\mu_i = \mu$.

The system described by the FPMD lagrangian is a *classical* conservative system. When the wavefunctions are expanded in a finite basis set it has a finite (although usually large) number of degrees of freedom. Therefore all the machinery of the classical analytical dynamics can be exploited to get additional information on its dynamical behaviour.

The constraints $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ on the KS orbitals are holonomic and scleronomous and are completely equivalent to the well-known rigid body constraints of classical mechanics. Thus, in every dynamical formalism, even more general than the present lagrangian one, these constraints cannot do work on the system and no dissipation can occur due to their presence.

In systems with few degrees of freedom it is customary to eliminate this kind of constraint by a transformation to a suitable system of independent coordinates (lagrangian coordinates) which automatically satisfy the constraints. In cases where the number of degrees of freedom is large such a procedure is usually not practical and one is forced to introduce an additional term in the lagrangian, containing the constraint equations multiplied by suitable Lagrange multipliers Λ_{ij}. The matrix Λ is hermitian if the lagrangian has real values.

The equations of motion resulting from the lagrangian 3.12 are:

$$\mu_i |\ddot{\psi}_i(t)\rangle = -f_i \hat{H}_{KS} |\psi_i\rangle + \sum_j \Lambda_{ij} |\psi_j\rangle \tag{3.13}$$

$$M_I \ddot{\mathbf{R}}_I = -\left.\frac{\partial E_{KS}}{\partial \mathbf{R}_I}\right|_\rho - \frac{\partial V_I}{\partial \mathbf{R}_I} \tag{3.14}$$

$$N_{ij}(t) = \langle\psi_i(t)|\psi_j(t)\rangle = \delta_{ij} \tag{3.15}$$

from which the unknowns $\psi_i(\mathbf{r},t), \mathbf{R}_I(t)$ and $\Lambda_{ij}(t)$ can be obtained for any choice of the initial conditions $\mathbf{R}_I(0), \dot{\mathbf{R}}_I(0), \psi_i(\mathbf{r},0), \dot{\psi}_i(\mathbf{r},0)$ such that $N_{ij}(0) = \delta_{ij}$ and $\dot{N}_{ij}(0) = 0$.

If one is not interested in doing MD, but only in finding a local minimum of the potential energy surface, the equations 3.13, 3.14 can be effectively substituted by other equations. Perhaps the simplest dynamical approach that one can use to find a minimum of a function with many variables is provided by the steepest descent method, which in the present case leads to the following equations of motion:

$$\begin{aligned}
\overline{\mu}_i|\dot{\psi}_i\rangle &= -f_i\hat{H}_{KS}|\psi_i\rangle + \sum_j \Lambda_{ij}|\psi_j\rangle \\
\overline{M}_I\ddot{\mathbf{R}}_I &= = -\left.\frac{\partial E_{KS}}{\partial \mathbf{R}_I}\right|_\rho - \frac{\partial V_I}{\partial \mathbf{R}_I} \\
N_{ij}(t) &= \langle\psi_i(t)|\psi_j(t)\rangle = \delta_{ij}
\end{aligned} \tag{3.16}$$

where the "masses" $\overline{\mu}_i$, $\overline{M}_I$ have been introduced in order to account for the difference in time scales associated with the different parameters.

Of course it is possible to keep the ions fixed and use the equations 3.16 to find the minimum of the electronic energy. We stress here that even under this point of view the FPMD scheme presents advantages with respect to standard methods. Suppose we want to solve eq. 3.7 for a system of M electrons; if we use N_{pw} plane waves to represent a single wave function, the solution of eq. 3.7, using standard methods, requires the diagonalization of a rank N_{pw} matrix that gives rise to $O(N_{pw}^3)$ operations. Equation 3.16 instead requires $O(M \times N_{pw} \times \log N_{pw})$ and $O(M^2 \times N_{pw})$ operation (the last due to the orthonormalization process) that in case of $N_{pw} \gg M$ is a considerable computational advantage making possible to treat large systems.

Some comments are in order.

- The dynamics generated by equations 3.13 and 3.15 for the electronic degrees of freedom is in general different from the adiabatic evolution on the BO surface. Thus the forces on the ions (equation 3.14) are generally different from the BO forces (equation 3.2). Only if the evolution of the electronic degrees of freedom is rapidly oscillating

around the instantaneous ground state the solution of equation 3.14 will be a good approximation of the BO ionic evolution. This implies that the electronic configuration at the initial time of the simulation must be very close to the ground state.

- An explicit, formal expression for the $\Lambda_{ij}(t)$ can be obtained as (we suppose here that the electronic masses are all equal):

$$\Lambda_{ij} = \frac{1}{2}(f_i + f_j)\langle\psi_j|\hat{H}_{KS}|\psi_i\rangle - \mu\langle\dot{\psi}_j|\dot{\psi}_i\rangle. \qquad (3.17)$$

  This expression is not suitable for numerical integration of the equations of motion, as discussed in ref. [69], but we stress that this is the *only* form of the matrix of the Lagrange multipliers compatible with the equations of motion and a conservative hamiltonian dynamics. Notice that a different dynamical structure of the equations (e.g. in connection with a different lagrangian) would give, in general, a different expression for the matrix Λ [69]. For details about the numerical integration scheme, see refs. [69,70,71]. Note however that the orthogonalization procedure exposed in these works assumes that the electronic masses are equal. If the masses depend on i, the procedure has to be changed. The details are reported in the appendix D.

- Since the FPMD lagrangian is time independent, an obvious constant of motion is the total energy of the whole system defined as

$$E_{const} = K_f + K_I + V_I + E_{KS} \simeq K_f + K_I + V_I + U = K_f + H_I \quad (3.18)$$

  where $K_f = \sum_i \mu_i\langle\dot{\psi}_i|\dot{\psi}_i\rangle$ is the "fake" kinetic energy of the electronic degrees of freedom and H_I is the (physically relevant) hamiltonian of the electron-ion system (the other constants of motion can be found in appendix A). The approximate equality holds only under conditions of adiabaticity for the electronic subsystem. The constant E_{const} has no direct physical interpretation. However, as long as the fake kinetic energy of the KS orbitals remains negligible with respect to the other terms — $K_f \ll H_I$ — the total energy H_I is almost constant. From the statistical mechanical point of view, this means that the ionic dynamics can be used to sample the ionic microcanonical ensemble ($H_I = constant$), provided that the ergodic hypothesis is satisfied in the subspace of the ionic coordinates.

- From the previous points it results that the FPMD is equivalent to a conventional MD for the ionic system if the system is in a metastable state corresponding to a temperature of the electronic degrees of freedom much smaller than the ionic temperature (the lifetime of such metastable state must be much larger than the simulation time), with electrons rapidly oscillating around the instantaneous ground state.

  One has to work far from the equipartition of energy since, if one allows the system to reach the state of thermodynamical equilibrium (equipartition), the ionic dynamics will be incorrect due to the large departures of the electrons from the BO ground state, which would give the wrong forces.

- It may happen that during the dynamical evolution of the FPMD system a systematic increase of K_f is observed. This means that, first, the electrons are gradually drifting away from the BO surface, and that, second, since the total system (electron plus ions) is conservative, the electrons are subtracting energy from the ions.

  In this case it is possible to stop the simulation after each n_d dynamical steps in order to put back again the electrons in the ground state, using n_m electronic energy minimization steps. Using this method the total system gradually loses energy, since the energy transferred from the ions to the electrons is periodically thrown away. Therefore the ionic subsystem gradually cools down, and some corrective action has to be taken.

- It is possible to use one of the standard thermostatting methods on the ionic degrees of freedom to obtain a canonical ensemble. This is useful even to avoid the cooling problem illustrated in the previous point. A continuous thermostat is preferable, since rescaling of velocities or similar methods [72] introduce discontinuities in the ionic trajectories which could perturb the evolution of the electrons. The Nosé thermostat [73] is a good choice under many respects. The method amounts to add to the system a degree of freedom s in such a way that the ionic equations of motion read

$$M_I \ddot{\mathbf{R}}_I = -\left.\frac{\partial (U + V_I)}{\partial \mathbf{R}_I}\right|_\rho - M_I \dot{\mathbf{R}}_I \frac{\dot{s}}{s} \tag{3.19}$$

The specific equation of motion for s is

$$\mathcal{Q}\ddot{s} = s\sum_I M_I\dot{\mathbf{R}}_I^2 - gk_BT_w s + \mathcal{Q}\frac{\dot{s}^2}{s} \tag{3.20}$$

where the free parameter $\mathcal{Q}$ plays the role of "mass" for s, g is the number of ionic degrees of freedom, and T_w is the desired temperature. The previous equations are written in terms of "real variables" so the dot indicates the derivative with respect to the "real time" [73].

It can be proven that the introduction of the Nosé thermostat in a microcanonical ensemble allows to compute the correct canonical averages within the intrinsic uncertainty due to fluctuations [73]. Moreover, even the correlation functions are correct to order $1/N$ [74]. One of these quantities is the self-diffusion coefficient; actually, even for small systems it can be shown [75] that the values obtained for the self-diffusion coefficients are quite satisfactory.

When using the Nosé thermostat the exact constant of motion E_{const} becomes $E'_{const} = E_{const} + H_s$. Even if K_f is negligible, H_I is no longer a constant of motion. The quantity conserved at its place is $H'_I = H_I + H_s$ instead. Here H_s is the energy of the Nosé degree of freedom s:

$$H_s = \frac{1}{2}\mathcal{Q}\frac{\dot{s}^2}{s^2} + gk_BT_w \ln s. \tag{3.21}$$

- It must be remembered that, as in any MD calculation, the use of a finite MD cell (on which periodic boundary conditions are imposed) introduces an additional periodicity on the system. If the system is not homogeneous, as in the case of the defect calculations, or disordered, that is always the case if $T \neq 0$, this periodicity is spurious and can be made harmless only by enlarging to a sufficient extent the cell itself. In particular, when studying a defective crystal the MD cell introduces an unphysical, large-scale ordered structure of defects. The interaction between the elements of such an array shows up, for example, in the formation of a band of localized defect states, which, for a silicon vacancy, can have a width of about $1\,eV$, even for relatively big supercells [76].

3.3 Computational details

The use of a MD cell with periodic boundary conditions allows one to exploit the standard techniques developed to compute the electronic structure of periodic solids. We briefly recall here the main technical points, with particular reference to the implementation in our program.

The wavefunctions have the Bloch form:

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} u_{i,\mathbf{k}}(\mathbf{r}) \tag{3.22}$$

where the functions $u_{i,\mathbf{k}}(\mathbf{r})$ have the periodicity of the MD cell and $\mathbf{k}$ belongs to the Brillouin Zone of the MD cell. The electron density may be computed as

$$\rho(\mathbf{r}) = \sum_{i,\mathbf{k}\in BZ} f_i \psi^*_{i\mathbf{k}}(\mathbf{r})\psi_{i\mathbf{k}}(\mathbf{r}). \tag{3.23}$$

Using special points in the Brillouin Zone one can reduce the infinite sum appearing in the 3.23 to a sum of only few terms [77,78] still having a good accuracy in representing the density. Making the MD cell larger the Brillouin Zone becomes smaller and fewer k-points in eq. 3.23 are enough; for infinite cells only the Γ point ($\mathbf{k} = 0$) appears in the sum 3.23. Then, approximating the sum 3.23 with just the Γ point gives satisfactory results if a sufficiently large cell is chosen.

A great advantage of the use of the Γ point is that the periodic functions $u_{i,\mathbf{k}}(\mathbf{r})$ may be made real. This halves the number of fields to be included in the lagrangian 3.12.

The electronic wavefunctions must be expressed in terms of a finite number of degrees of freedom. This amounts to choose a set of basis functions and to expand the $\{\psi_i\}$ on it: the expansion coefficients are the desired degrees of freedom. We choose a plane waves basis set (the formal details may be found in refs. [79,80]) and we indicate the expansion coefficients by $c^i(\mathbf{G})$. The vectors $\mathbf{G}$ belong to the reciprocal lattice of the Bravais lattice of the MD cell.

The number of $\mathbf{G}$ vectors is usually expressed by means of the cutoff energy E_{cut}, which represents the (quantum) kinetic energy of a plane wave with the highest G. This number is $N_{pw} \simeq \frac{1}{6\pi^2}\Omega {E_{cut}}^{3/2} = \frac{1}{6\pi^2}\Omega {G_w}^3$, Ω being the cell volume.

Every quantity which, in real space, can be written as a product of two functions, such as the electron density, in G space become a convolution:

$$\rho(\mathbf{G}) = \sum_{i,\mathbf{k}\in BZ} f_i \sum_{\mathbf{G}'} \psi^*_{i\mathbf{k}}(\mathbf{G})\psi_{i\mathbf{k}}(\mathbf{G}-\mathbf{G}'). \tag{3.24}$$

The G vectors involved in the latter sum have a maximum absolute value of $G_c = 2G_w$. In many cases it has been found that it is possible to substitute G_c with some G'_c such that $G'^2_c = \gamma G^2_w$, with $1 \leq \gamma \leq 4$. Actually, also the use of a value of γ smaller than 4 gives the correct results in the limit of high E_{cut}.

We choose to treat the electron-ion interaction within the first-principles norm-conserving pseudopotential approach [81]. This approach amounts to consider the core electron wavefunctions to be "frozen", i. e. completely insensitive to the environment, and therefore they may be computed once and for all. The effective bare interaction potential between the "pseudoion" (nucleus plus core electrons) and a valence electron is given by a suitable pseudopotential. Note that within this approach only the valence electrons have to be included into the electronic degrees of freedom.

The potentials obtained in this way are non-local, in the sense that components of the $\{\psi_i\}$ having different angular momentum character feel a different potential. The pseudopotential is usually written as $\hat{V}_{ps,nl} = \sum_l v_l \hat{P}_l$, $\hat{P}_l$ being the projection operator on the state having angular momentum l. By approximating $v_l = v_{l_c}$ for $l > l_c$, and using the completeness relation $\sum_l \hat{P}_l = \hat{1}$ we obtain $\hat{V} \simeq \sum_{l<l_c}(v_l - v_{l_c})\hat{P}_l + v_{l_c}$. For sufficiently large l_c this introduces a negligible error because high-l components of the wavefunctions do not enter appreciably in the core. In the case of silicon, where the bonds have s-p character, the choice $l_c = 2$ gives a very accurate pseudopotential, but even the choice $l_c = 1$ is sufficiently accurate for many purposes.

We use the form given by Bachelet *et al.* [82] in the so-called "fully-nonlocal" form of Kleinman and Bylander [83]. See the appendix B for details.

3.3.1 Choice of the parameters

From the previous discussion we see that in our implementation of the FPMD method there are some parameters which must be suitably chosen

in order to have an accurate description of the physics of the system. They are the set of **k** points, the size of the MD cell, the cutoff energy E_{cut}, the parameter γ, and the value of l_c.

The choice of the values of the parameters must result from a compromise between the requirements of accuracy and the available computer resources.

Our test calculations (see appendix E) indicate that a well converged formation energy (within $\sim 0.1\ eV$) is obtained with $E_{cut} = 12\ Ryd$ and $l_c = 2$. For a 64 atom cell with Γ point sampling of the Brillouin Zone such values of E_{cut} and l_c give a formation energy of 3.3 eV for the fully relaxed neutral vacancy. However finite size effects related to Γ point sampling are still not completely negligible for a 64 atom cell: on the basis of calculations with a more extended set of **k** points we conclude that this results in an underestimate of E_f by $\sim 0.3\ eV$. A larger cell with 128 atoms reduces this underestimate to $\sim 0.1\ eV$.

Putting together our corrections for finite energy cutoff and for finite size effects, we obtain $E_f = 3.5 \pm 0.2\ eV$ for our best estimate of the vacancy formation energy, where the error bars reflect the uncertainty in our extrapolation. This value is in good agreement with experimental estimates [11] and with previous theoretical results [57,58]. In all our subsequent calculations we will then assume $E_f = 3.5\ eV$. This coincides with the formation free energy F_f at $T = 0$.

The finite temperature corrections to F_f depend on the modification of the vibrational spectrum of the crystal due to the introduction of the vacancy and represent a contribution of the order of less than 25% of E_f. Since very expensive computations are required to get F_f at finite temperature we used in this case $E_{cut} = 8\ Ryd$, $l_c = 2$, and $\gamma = 1.5$. We explicitly verified that the use of $\gamma = 1.5$ introduces in this case negligible errors. The cutoff $E_{cut} = 8\ Ryd$ is sufficient to obtain the vibrational frequencies of the perfect crystal with an accuracy of better than 10%. We expect similar errors on the formation entropy. As we shall see, larger errors ($\sim$ 25–30%) result from the stochastic method used to compute F_f at finite temperature.

For the vacancy jump simulation, we are more interested in the accuracy of the migration energy. We find that $l_c = 1$, $E_{cut} = 12\ Ryd$ and $\gamma = 1.5$ give an excellent convergence for this energy; the value computed is 0.3 eV, not far from the experimental value of 0.45 eV.

A remark is in order. In all our calculations, both in static and in dynamic runs, we use constant volume calculations, while experiments are always done under constant pressure conditions. Therefore we compute energies while the experiments measure enthalpies. Under standard pressure conditions the contribution of the additional term $p\Delta V$, where p is the pressure and ΔV is the volume change due to a temperature change, or to defect creation or migration, is much smaller than the error due to other sources.

We checked the effect of using a temperature independent cell by computing the migration energy with a cell scaled to the value expected for a temperature of 1400 $^{\circ}K$. We found differences smaller than 0.01 eV.

3.4 An analysis of the FPMD method

3.4.1 Classical adiabatic behaviour and the averages method

The metastable two-temperature regime set up in the FPMD dynamics is extremely efficient to approximate the constraint of maintaining the electronic energy functional at the minimum without the need of explicit minimization at each step.

The origin of this behaviour can be traced back to the classical adiabatic nature of the dynamics of the electronic degrees of freedom. At the beginning of a numerical simulation, the electronic subsystem is at the minimum of the energy surface. When the ions start moving, their motion causes a change of the instantaneous position of the minimum in the parameter space of the electronic degrees of freedom. The electrons experience forces and start moving. Since they start from a stable equilibrium position, there will be a range of initial condition such that a regime of small oscillations is originated. In many cases, with a suitable value of the fictitious electronic masses μ_i the frequencies of the normal modes of the electronic component can be made higher than those of the ionic one.

This high frequency motion has two main consequences:

1) the electronic degrees of freedom execute fast small oscillations superimposed to an average drift which effectively follows the slow evolution of the ionic variables;

2) the long-time exchanges of energy between fast and slow variables become very small.

These arguments can be put on a sound basis by making connection between the FPMD system and the general theory of adiabatic hamiltonian systems [84,85,86].

Let's recall in an informal way the two main results of classical mechanics that are relevant to understand the reasons why the FPMD method is so effective in maintaining the electronic degrees of freedom on the BO surface.

The first is the so called "Averaging Method" [85,87]. This gives a justification of the fact that the hamiltonian of a system containing fast and slow variables can be approximated very closely by an "averaged" hamiltonian obtained by performing a phase averaging of the fast coordinates in the original hamiltonian.

The second related result, from which the Averaging Method can be derived, is the existence of high order adiabatic invariants in almost integrable systems [85,88]. This implies the existence of quantities which remain almost constant over times very rapidly increasing with the rate of change of the hamiltonian.

The formal statement of the previous results as well as the conditions for their validity and their range of applicability are the subject of contemporary research in the field of dynamical systems. It is difficult to claim with certainty that they are applicable without qualification to the behaviour of the FPMD system. Rather, taking a more empirical attitude it can be shown that the applicability of the "Averaging Principle" and of the classical adiabatic approximation is a natural way to explain the two main features of the FPMD dynamics observed in actual simulations.

The Averaging Method can explain point 1). In an informal way, this principle states that a system of fast oscillating variables perturbed by a slow variation of the parameters (at a rate ϵ) can be closely approximated by the system obtained by performing an average over the fast coordinates. The trajectory of the "average system" remains close to that of the real system within order ϵ for times proportional to some inverse power of ϵ.

If, during the evolution of the FPMD system there is no catastrophic change in the structure of the potential seen by the electrons (like the softening of a mode) the average system will correspond to the desired (adiabatic) system evolving with electronic degrees of freedom strictly bound

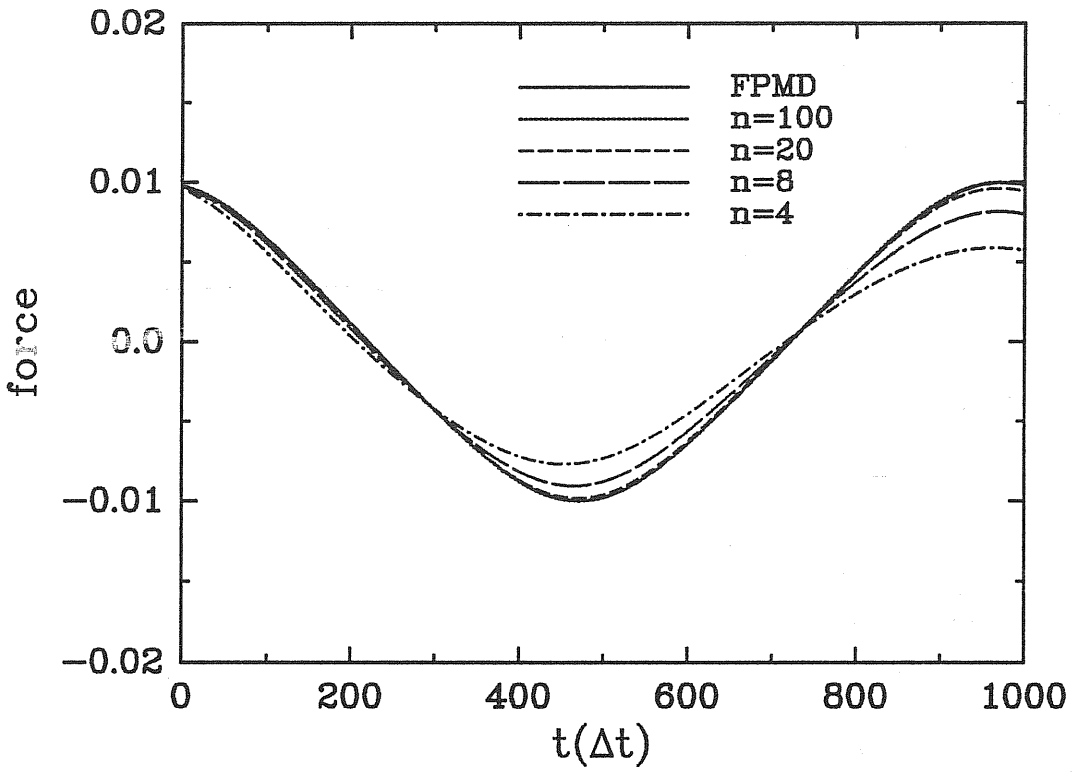

Figure 3.1: x-component of the force on the ions in a two-atom silicon cell (see text).

to the BO surface. The simulation system instead, will rapidly oscillate around this average system for very long times.

As a result of the continuous dependence of the forces on the electronic degrees of freedom in formula 3.14 the fast evolution of the electronic degrees of freedom close to the BO surface will also imply that the forces on the ions in the FPMD system will quickly oscillate around the exact HF forces.

As an illustration of the way the averaging method works for the FPMD lagrangian we computed the x-component of the force (figure 3.1) between two silicon atoms in a two-atom cell. The wavefunctions were expanded in a plane-waves basis set with a cutoff of 6 *Ryd* and the Brillouin Zone of the cell was sampled using the Γ point only. This choice gives a poor description of the system, for example the stable crystal structure turns out to be

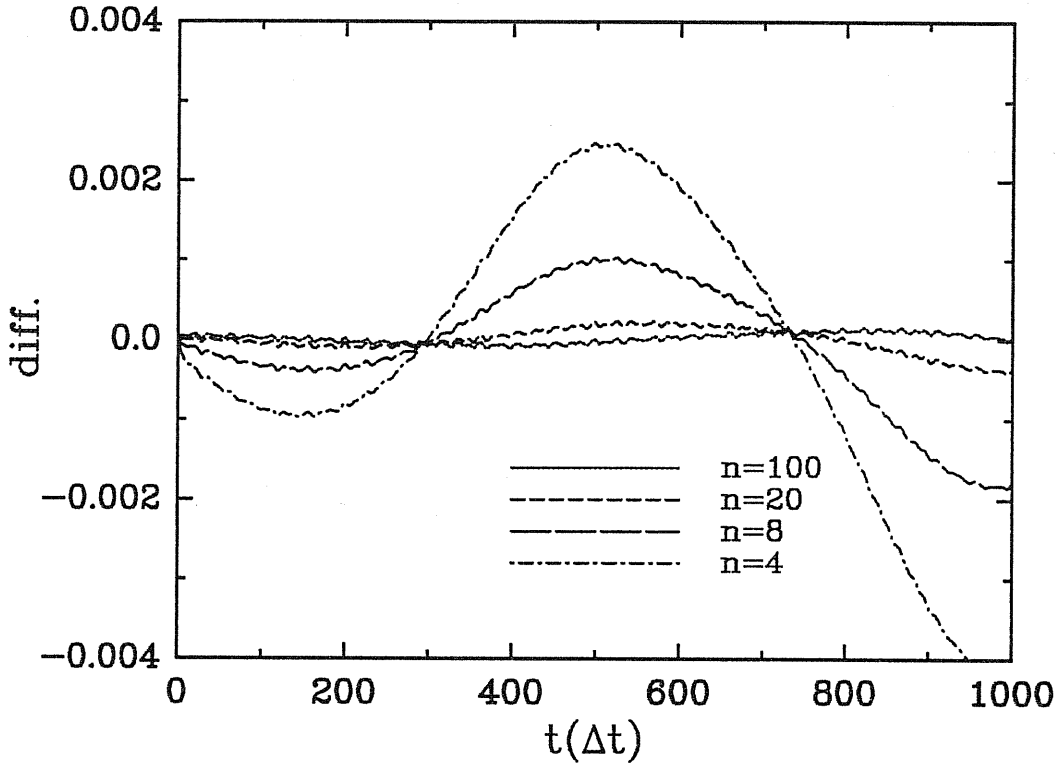

Figure 3.2: Difference between the forces of figure 3.1 and the FPMD force (see text).

simple cubic and not diamond. However, for the purpose of illustrating the behaviour of the FPMD system with realistic potentials, the accuracy of the representation of the experimental data is irrelevant. The timestep Δt was set to 5 $a.u.$ and the fictitious electron mass $\mu = 300$ $a.u.$.

The full curve in figure 3.1 refers to the force evaluated by the FPMD dynamics (equations 3.13 and 3.14), the other curves refer to a dynamics obtained by minimizing E_{KS} at each time step by using n steepest-descent steps (equation 3.16) and computing the forces on ions via the formula 3.14. We checked the convergence in n by performing a computation with $n = 350$; the results are visually indistinguishable from the continuous line.

We see that with a smaller number n the errors on the trajectory add up, leading to a systematic damping of the ionic motion. A plot of the differences of the forces plotted in fig. 3.2 with respect to the FPMD force stresses better this point.

The existence of adiabatic invariants, consequence of the fast dynamics of the electronic degrees of freedom, explains the quasi-microcanonical behaviour of the FPMD system: in a regime of small oscillations, for each normal mode there will be an adiabatic invariant (E_i/ω_i) where E_i is the energy of the mode and ω_i its frequency.

Adiabatic invariance of such quantities means that

$$\left| \frac{E_i(T)}{\omega_i(T)} - \frac{E_i(0)}{\omega_i(0)} \right| < \epsilon \tag{3.25}$$

for times of the order of ϵ^{-r}. This equation can be interpreted as the condition for approximate reversibility of the unavoidable exchanges of energy with the rest of the system.

If the electronic frequencies are too small, and in particular if they become comparable with the ionic frequencies, we can expect a "disordered" transfer of energy from the ionic to the electronic degrees of freedom followed by a rapid equilibration of the whole system. The effectiveness of this energy transfer depends on the strength of the electron-ion interaction: the weaker is the latter, the smaller can be the former.

3.4.2 A simplified model

The discussion in the previous subsection shows how the fast dynamics of the electronic degrees of freedom allows the system to follow on average the

adiabatic evolution of the electronic ground-state.

However, in some cases, such as metallic and level crossing situations, the lowest frequencies of the spectrum of the electronic degrees of freedom may get arbitrarily close to zero. In such cases it is not possible to find masses μ_i such that the adiabatic and the averaging principles apply.

Thus, it seems necessary to investigate more in details the dynamics of the oscillations of the electronic degrees of freedom to understand the possible mechanisms that break classical adiabaticity. A direct analysis of the dynamics of the electronic degrees of freedom of the full FPMD lagrangian would be necessary, but, due to the complex non-linearity of the KS hamiltonian and to the large number of electronic degrees of freedom which requires the use of Lagrange multipliers, it is impossible to obtain exact results. To simplify the problem we shall study the problem of the FPMD normal modes corresponding to a hamiltonian matrix H_{ij} defined in an n-dimensional Hilbert space. Full details can be found in reference [89].

The elements of H slowly depend on time. In addition, at variance with the KS case, H is a linear (not self-consistent) operator. In our view this simplified model contains most of the relevant physics of the real case.

The functions $\Phi_i\{i = 1, \ldots, n\}$ are an orthonormal set of basis functions which, for simplicity, may be taken as eigenfunctions of H at the time $t = 0$:

$$H(0)\Phi_i = \epsilon_i \Phi_i \tag{3.26}$$

Let the first m ($m < n$) levels be occupied with occupation numbers $0 \leq f_i \leq 1$. The ground state at $t = 0$ will be described in terms of the first m eigenstates of $H(0)$. We investigate the behaviour of the classical system obtained by writing the FPMD lagrangian at fixed ions (without loss of generality we can assume that the states are represented by real functions):

$$\mathcal{L}'(\psi, \dot{\psi}) = \frac{1}{2}\sum_i \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle - \frac{1}{2}\sum_i f_i \langle \psi_i | H | \psi_i \rangle + \sum_{i,j} \Lambda_{ij}(\langle \psi_i | \psi_j \rangle - \delta_{ij}) \tag{3.27}$$

Notice the factor $1/2$ for real degrees of freedom, different from that originally introduced by Car and Parrinello [8]. If the occupation numbers vary from a state to another, the fictitious masses μ_i are suitably defined as $\mu_i = \mu f_i$. This choice makes the eigenfrequencies of the electronic degrees of freedom independent of the occupation, with an exception we shall see later.

We can exploit the possibility offered by the small number of variables to reduce explicitly the electronic degrees of freedom to the independent ones by a change of variables. The algebra is straightforward but becomes quickly cumbersome increasing m and n. We have worked out in detail the cases $n = 2$, $m = 1$ and $n = 3$ with $m = 1$ or 2. However, we believe that they are actually typical enough to represent the full class of models.

The case $n = 3$, $m = 1, 2$

Let's start with the case $m = 1$. We express the occupied state ψ_1 as

$$\psi_1 = (\cos\theta\cos\varphi)\Phi_1 + (\sin\theta\cos\varphi)\Phi_2 + (-\sin\varphi\cos\theta)\Phi_3 \tag{3.28}$$

After evaluation of the lagrangian 3.27 and linearization of the equations of motion, we have the following two equations for the small oscillations of the variables φ and θ:

$$\begin{aligned} \mu\ddot{\theta} &+ (\epsilon_2 - \epsilon_1)\theta = 0 \\ \mu\ddot{\varphi} &+ (\epsilon_3 - \epsilon_1)\varphi = 0 \end{aligned} \tag{3.29}$$

The interpretation is straightforward: θ and φ control the overlap of ψ_1 with the excited states Φ_2 and Φ_3 respectively. If we start with the system in the minimum energy state then $\varphi = \theta = 0$. A perturbation will make ψ_1 and ψ_2 acquire components on the excited state Φ_3. The restoring forces induce oscillations with frequencies proportional to $\sqrt{\epsilon_3 - \epsilon_1/\mu}$ and $\sqrt{\epsilon_3 - \epsilon_2/\mu}$.

With this linearized hamiltonian, these two oscillations are independent and their frequencies are proportional to the square root of the difference of the eigenvalues of the empty and filled states.

The case $m = 2$ allows to study the effect of more than one occupied level. We have the following expression for the two occupied states ψ_1 and ψ_2 as a function of the lagrangian coordinates θ, φ and ψ:

$$\begin{aligned} \psi_1 \;=\; & (\cos\theta\cos\varphi)\Phi_1 + (\cos\theta\sin\varphi\sin\psi + \sin\theta\cos\psi)\Phi_2 \\ & + (-\sin\varphi\cos\psi\cos\theta + \sin\psi\sin\theta)\Phi_3 \end{aligned} \tag{3.30}$$

$$\begin{aligned} \psi_2 \;=\; & (-\sin\theta\cos\varphi)\Phi_1 + (-\sin\theta\sin\varphi\sin\psi + \cos\theta\cos\psi)\Phi_2 \\ & + (\sin\varphi\cos\psi\sin\theta + \sin\psi\cos\theta)\Phi_3 \end{aligned} \tag{3.31}$$

We have chosen the phases in such a way that for $\varphi = \theta = \psi = 0$ ψ_1 corresponds to the lowest eigenstate Φ_1 and ψ_2 to the next Φ_2. We arrive to the following simple expressions for the linearized equations of motion:

$$\begin{aligned} \mu(f_1 + f_2)\ddot{\theta} + (f_1 - f_2)(\epsilon_2 - \epsilon_1)\theta = 0 \\ \mu\ddot{\varphi} + (\epsilon_3 - \epsilon_1)\varphi = 0 \\ \mu\ddot{\psi} + (\epsilon_3 - \epsilon_2)\psi = 0 \end{aligned} \tag{3.32}$$

The physics behind equations 3.32 is analogous to that of the previous case. For equally occupied states ($f_1 = f_2$) one of the non-linear equations of motion can be integrated immediately to give the constant of motion:

$$\dot{\theta} + \dot{\psi}\sin\varphi = const \tag{3.33}$$

corresponding to a free "rotation" in the two-dimensional space of the occupied states (see equation A.6). In the linearized equations this reduces to $\dot{\theta} = const$, which is obvious since the θ coordinate is cyclic.

If $f_1 > f_2$ the quantity written in 3.33 is no more a constant of motion. A restoring force exists which tends to keep the wavefunction ψ_2 the highest in energy: it is easy to show that this is the minimum energy state for the system. If $f_1 \simeq f_2$ this force constant is quite small and can give rise to a low frequency mode.

In the case in which the amplitude of the oscillations increases, the linearized equations of motion will not be able to describe the real evolution of the system. The non-linear equations in the case $n = 3$ are very complicate, so the simpler case $n = 2$ appears more suitable for analytic investigation. We will show in the next subsection that this case, although simplified, is relevant to discuss the actual calculations.

The case $n = 2,\ m = 1$

In this case, we have only one lagrangian coordinate — $\theta(t)$ — in terms of which the exact equation of motion is

$$\ddot{\theta}(t) = \frac{\Delta E}{\mu}\sin(\theta(t) - \theta_0) \tag{3.34}$$

i. e. is the non-linear oscillator (pendulum) equation. The gap $\Delta E = \epsilon_2 - \epsilon_1$ and the equilibrium angle θ_0 appearing in the eq. 3.34 can be expressed in

terms of the matrix elements of H. In particular, they can be varied in an almost independent way.

For slow variations of the matrix elements, the electronic degrees of freedom will be able to readjust themselves to follow the instantaneous values. However, when ΔE will decrease we can expect two regimes where, at different levels of severity, the classical adiabatic behaviour of the system will be destroyed.

In the first place, the electronic dynamics will slow down and at some point the flux of energy between the ionic and the electronic component may find efficient channels.

In the second place, as soon as ΔE becomes negative (at the crossing of the two levels) a real dynamical instability appears in the sub-system of the electronic degrees of freedom. In analogy with the physical pendulum, this corresponds to a sudden change in the equilibrium angle from θ_0 to $\theta_0 + \pi$ or to a gravitational force pointing upwards. The effect is the appearance of an instability in the small oscillatory motion and a rapid increase in the kinetic energy of the electronic degrees of freedom.

Finally, we observe that the simple equation 3.34 can be exactly integrated in term of elliptic functions if the parameters ΔE and θ_0 are constant. For a specific choice of the time dependence of the gap, this equation would reduce to the Mathieu's equation. However, as far as we know, analytical studies relevant to the case of instability (ΔE which becomes negative) are lacking.

3.5 Connection with realistic systems

The figure 3.3 shows the power spectrum of the dynamics of the electronic degrees of freedom computed for the same system of figure 3.1. This spectral function is the Fourier transform of the time autocorrelation function of the electronic velocities:

$$\gamma(\omega) = \int_0^\infty \mathrm{d}t \sum_{i,\mathbf{G}} \langle \dot{c}^{i*}(\mathbf{G};\, t)\dot{c}^{i}(\mathbf{G};\, 0)\rangle \cos(\omega t) \qquad (3.35)$$

where the time average is performed over 3000 time steps ($\Delta t = 5\ a.u.$). The ions were taken at fixed positions and the electronic degrees of freedom started the dynamics with zero velocity and with a maximum random

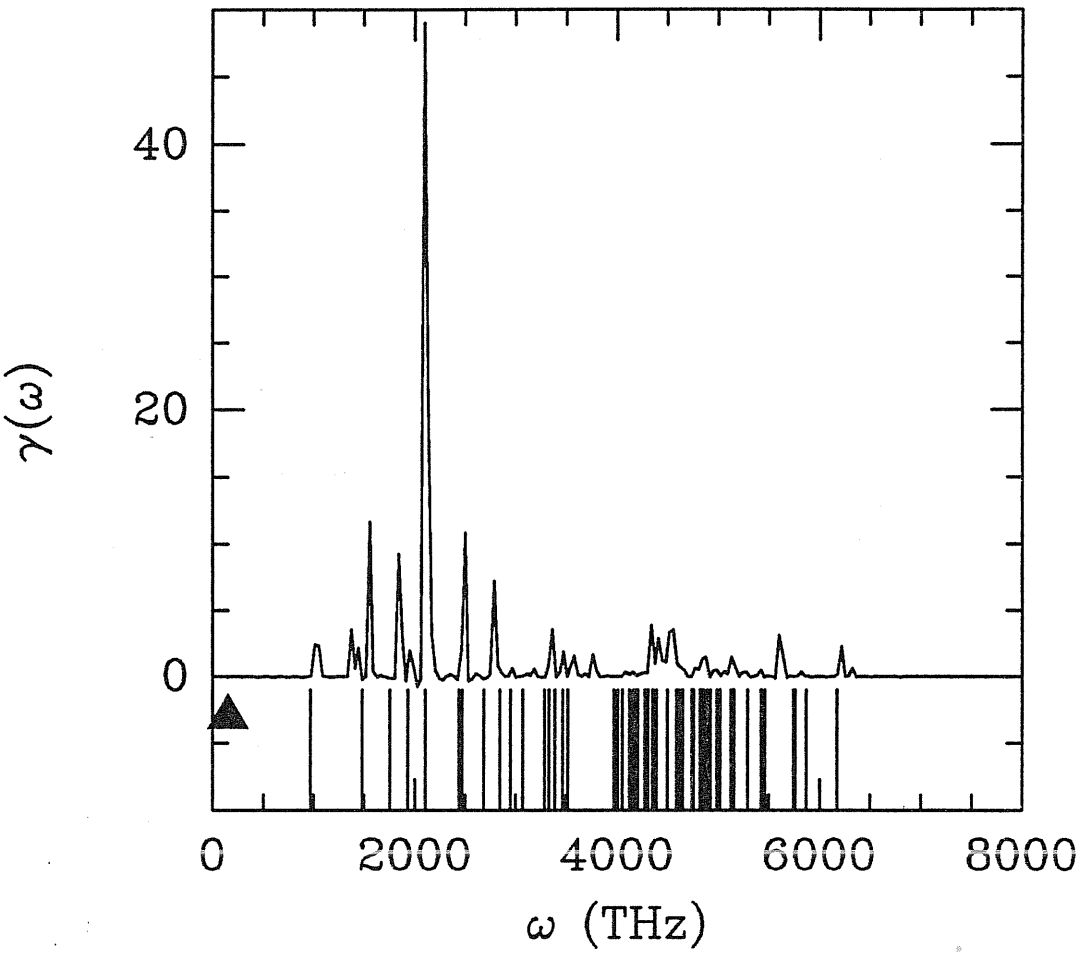

Figure 3.3: Power spectrum of the electronic dynamics of the system of figure 3.1. The triangle indicates the maximum ionic frequency.

displacement of a 4% from the ground-state configuration, a displacement which gives oscillations in the linear regime. For visual help a mark shows the position of the highest ionic frequency (22 THz).

In the present case, the frequency spectrum starts from about 1000 THz with a maximum frequency of about 6700 THz. As a matter of comparison, we have reported as vertical segments under $\gamma(\omega)$ the positions of the frequencies resulting by the analysis of the previous chapter, i. e. assuming a non-selfconsistent hamiltonian:

$$\omega_{ij} = \left((\epsilon_i^{unoccupied} - \epsilon_j^{occupied})/\mu \right)^{\frac{1}{2}} . \tag{3.36}$$

It can be seen that the position and the bandwidth of the calculated power spectrum are in reasonable agreement with the simple linear approximation.

We can conclude that, even in realistic systems, the lowest electronic frequency depends on the minimum difference between the occupied and unoccupied eigenvalues of the Hamiltonian. In the case of a bulk semiconductor, the lowest electronic frequency is thus given by $\sqrt{E_{gap}/\mu}$. If the system has gap levels, E_{gap} is substituted by the energy difference between the highest occupied gap level and the lowest unoccupied gap level, which implies a substantial reduction of the minimum electronic frequency. Therefore the adiabatic behaviour of the electrons worsens.

A possibility for restoring the regime of high frequency electronic oscillations is to reduce the electronic mass μ. However, the whole spectrum scales with $\mu^{\frac{1}{2}}$. Therefore the highest electronic frequencies would be raised too, and a smaller time step would be required for integrating correctly the electronic equations of motion.

Chapter 4

Calculation of Free Energies by Molecular Dynamics

4.1 A short review of methods for computing free energies

A direct evaluation of F, the Helmholtz free energy is not practically possible because it is written as a configurational integral with a rapidly varying integrand, and it is therefore very hard to compute [45]. For a classical system:

$$\mathrm{e}^{-F/k_BT} = \mathcal{Z} = \mathcal{Z}_p\mathcal{Z}_q \tag{4.1}$$

$$\mathcal{Z}_p = \int \mathrm{d}^{3N}p\,\mathrm{e}^{-K_I/k_BT} \tag{4.2}$$

$$\mathcal{Z}_q = \int \mathrm{d}^{3N}q\,\mathrm{e}^{-U_I/k_BT} \tag{4.3}$$

Here K_I is the kinetic energy of the ionic system and U_I is its potential energy (in the BO approximation $U_I = V_I + U$). The kinetic part $\mathcal{Z}_p$ of the partition function can be computed analytically. The difficulties in calculating $\mathcal{Z}$ lie in the so called "configurational part" $\mathcal{Z}_q$.

The free energy F can be calculated in practice by using the harmonic approximation, which gives [43]:

$$F = \phi_0 + \frac{1}{2}\sum_{I=1}^{3N-3}\hbar\omega_I + k_BT\sum_{I=1}^{3N-3}\ln\left(1-\mathrm{e}^{-\frac{\hbar\omega_I}{k_BT}}\right) \tag{4.4}$$

$$\simeq \phi_0 + k_B T \sum_{I=1}^{3N-3} \ln\left(\frac{\hbar\omega_I}{k_B T}\right) \tag{4.5}$$

where ϕ_0 is the potential energy at the minimum energy configuration, ω_I are the system eigenfrequencies, and the approximate equality refers to the high-temperature case.

Although possible (see the work of Bachelet *et al.* [90]) this approach is computationally very expensive when used in combination with first-principles calculations of the electronic ground state. In fact, in order to obtain an accurate evaluation of the thermodynamic quantities rather big cells have to be used. The computation of the full dynamical matrix $D_{IJ} = \partial^2 E/\partial\mathbf{R}_I\partial\mathbf{R}_J$, which is needed to obtain $\{\omega_I\}$, requires the knowledge of the energy variations as a function of the displacements of each *pair* of atoms. Let n be the number of values of the energy needed for an accurate evaluation of a single matrix element of D_{IJ}: to obtain the full dynamical matrix one should, in principle, perform $O(nN^2)$ energy minimizations, with a huge consumption of computer time.

Bachelet and De Lorenzi [91] suggested a possibly more efficient way of performing such calculations, but their method has not yet been sufficiently assessed.

A simpler procedure has been proposed by LeSar *et al.* [9]. In their "local harmonic approximation" (LHA) the coupling between vibrations of different atoms is neglected. In this case, the three vibrational frequencies $\omega_{1J}, \omega_{2J}, \omega_{3J}$ for each particle J in the cell can be determined by diagonalizing the local dynamical matrix D_J. In this case, the problem of calculating F is reduced to the diagonalization of N matrices of rank 3, which implies $O(nN)$ energy minimizations. This is a major computational advantage with respect to the computation of the full dynamical matrix.

In the LHA the equation 4.5 becomes

$$F \simeq \phi_0 + 3k_B T \sum_{J=1}^{N} \ln\left(\frac{\hbar(\omega_{1J}\omega_{2J}\omega_{3J})^{1/3}}{k_B T}\right) = \phi_0 + 3k_B T \sum_{J=1}^{N} \ln\left(\frac{\hbar \det D_J^{1/3}}{k_B T}\right) \tag{4.6}$$

The LHA lies on a pair of approximations. The first approximation is, of course, the harmonic approximation. In the case of the silicon crystal with a vacancy a substantial softening of the vibrational frequencies around the defect is found. This fact indicate that anharmonic contributions may be important.

The second approximation consists in supposing that the frequencies of the individual ions are decoupled. In ref. [9] this approximation was found to give excellent results for a Lennard-Jones system. It is not clear how good this approximation for semiconductors, where more complicated interatomic interactions exist.

The harmonic approximation is a good starting point but not always satisfactory. The anharmonic effects are responsible for several important physical effects, such as the thermal expansion. In the quasi-harmonic approximation the anharmonicity is introduced by taking into account the volume dependence of ϕ_0 and ω_I.

The computation of free energy differences is possible, taking fully into account the anharmonic effects, by means of statistical simulation methods such as Montecarlo or MD. However, since the free energy is not cast in the form of an equilibrium average of functions of positions and velocities, its calculation is not straightforward as the computation, e. g., of the internal energy. Specific methods have been developed to accomplish this task.

A first method is the "acceptance ratio" method [92], useful for simulations in the canonical ensemble. In this method one may compute the difference in the free energy between two systems, for example a perfect crystal and a defective crystal having not too different hamiltonians. This requirement is hardly achieved in systems, like ours, in which the formation free energy F_f is $\sim 25\, k_B T_{melting}$.

Another way of directly computing free energy differences is to use one of the so-called "Hamiltonian manipulation" methods. In this method one writes a fictitious hamiltonian $\mathcal{H}(\lambda) = K_I + \mathcal{U}(\lambda)$, where λ is a parameter such that $\lambda \in [0,1]$. One then may exploit two facts:

- The derivative of the free energy with respect to λ has the form of a canonical average, since

$$\begin{aligned} \frac{\partial F(\lambda)}{\partial \lambda} &= \frac{\partial \mathcal{Z}_q(\lambda)}{\partial \lambda} = -k_B T \frac{\partial}{\partial \lambda} \ln \int \mathrm{d}^{3N} q \, \mathrm{e}^{-\mathcal{U}(\lambda)/k_B T} \\ &= \frac{1}{\mathcal{Z}_q(\lambda)} \int \mathrm{d}^{3N} q \, \Phi(\lambda) \, \mathrm{e}^{-\mathcal{U}(\lambda)/k_B T} = \langle \Phi(\lambda) \rangle \end{aligned} \tag{4.7}$$

  where

$$\Phi = \frac{\partial \mathcal{U}}{\partial \lambda} \tag{4.8}$$

and the angular brackets indicate canonical ensemble average. If the averages are taken over MD runs, in order to get canonical sampling a thermostatting method must be used.

- The free energy is a function of state, so irrespective of the functional form of $\mathcal{U}(\lambda)$ at intermediate values of λ (provided it is a continuous function) we have

$$F(1) - F(0) = \int_0^1 \mathrm{d}\lambda \, \langle \Phi \rangle \, . \tag{4.9}$$

In other words, one may use the parameter λ to drive the system via a reversible thermodynamic transformation from the state corresponding to the hamiltonian $\mathcal{H}(0)$ to another corresponding to $\mathcal{H}(1)$ (or vice versa). It is not necessary that this transformation be experimentally realizable. This is a great advantage of computer simulations!

Let us apply this method to the computation of the free energy of formation of the vacancy. We start from the perfect crystal (N atoms) and decouple an atom (say, the atom # 0) from it; i. e., we pick up the atom 0 and put it in a state described by an hamiltonian H_0^1, reversibly. The hamiltonian H_0^1 depends only on the coordinates of the atom 0 and therefore its free energy can be computed separately:

$$F_0^1 = -k_B T \ln \int \frac{\mathrm{d}^3 p_0 \, \mathrm{d}^3 q_0}{\Lambda^3} \, \mathrm{e}^{-H_0^1/k_B T} \tag{4.10}$$

where Λ = De Broglie thermal length $= (h^2/2\pi m k_B T)^{1/2}$. We choose λ as the coupling parameter of this atom to the rest of the crystal, so

- $\mathcal{H}(1) = H_B^N$ = hamiltonian of the perfect crystal,
- $\mathcal{H}(0) = H_V^{N-1} + H_0^1$ = hamiltonian of the crystal with a vacancy (N-1 atoms) plus the hamiltonian of the atom 0.

Therefore the free energies of the bulk system and of the system with a vacancy are, respectively

$$\begin{aligned} F_B^N &= F(\lambda = 1) \\ F_V^{N-1} &= F^N(\lambda = 0) - F_0^1 \end{aligned} \tag{4.11}$$

Then, for the formation free energy of the vacancy we get

$$\begin{aligned} F_f &= F_V^{N-1} - \frac{N-1}{N} F_B^N = F^N(0) - F^N(1) + \frac{1}{N} F_B^N - F_0^1 \\ &= -\int_0^1 \mathrm{d}\lambda \, \langle \Phi \rangle + \Delta F. \end{aligned} \tag{4.12}$$

It is now apparent that the state H_0^1 is arbitrary: its free energy is implicitly added to F_f when computing $F^N(0)$ and explicitly subtracted in the term ΔF. It is therefore advisable to choose H_0^1 in such a way that the corresponding free energy F_0^1 is analytically computable.

The value of F_B^N refers to a bulk system, and therefore can be obtained from experimental data or, if the harmonic approximation holds, from accurate calculations of phonon spectra using eq. 4.4.

The integral in the eq. 4.12 must be computed numerically. In order to reduce the number of λ points needed for it, one could use higher order derivatives of $\Phi(\lambda)$. This would lead to different methods ("interpolation methods" [92]) which, however, in practice pose problems due to the poor statistical quality of the data for higher order derivatives.

The simplest choice for the reference state of the particle 0 is that of a free particle. This was the choice of Squire and Hoover [93] in their pioneering work on the vacancy in Lennard-Jones crystals. However, this choice introduces difficulties in the computation of Φ. To see this, let us suppose to have a two body potential $\phi(\mathbf{r}_{ij})$; in this case the potential energy is $U_I = \frac{1}{2} \sum_{ij, i \neq j} \phi(\mathbf{r}_{ij}) + \lambda \sum_{i>0} \phi(\mathbf{r}_{i0})$. It is apparent that, as $\lambda \to 0$, the particle 0 can approach very closely any other particle, since the potential energy is very small. But, since $\Phi = \sum_{i>0} \phi(\mathbf{r}_{i0})$ does not vanish with λ, during these "close approaches" it can become very large. This leads to a cusp in the function $\langle \Phi(\lambda) \rangle$, as reported in [93]. This fact forced the authors of reference [93] to use many λ-points to achieve sufficiently high accuracy.

A better choice is to put the 0 atom in an harmonic potential well, i. e.

$$H_0^1 = \frac{1}{2} M \dot{\mathbf{r}}_0^2 + (1-\lambda) \left[E_0 + \frac{1}{2} M \Omega^2 (\mathbf{r}_0 - \mathbf{a})^2 \right] \tag{4.13}$$

M being the ionic mass, and $\mathbf{a}$, E_0, and Ω arbitrary quantities. In this case F_0^1 may be evaluated analytically, giving, in the classical case:

$$F_0^1 = E_0 - 3 k_B T \ln \left(2\pi k_B T / \Lambda (1-\lambda) \Omega \right). \tag{4.14}$$

The advantage of this choice is that it confines the atom in a limited region of space. This is effective if a is taken to be the position vector of the vacant site and Ω a typical crystal frequency: in this case the atom 0 will behave roughly as a bulk atom. Note that this choice for the λ dependence of H_0^1 does not affects the energy at $\lambda = 1$, i. e. in the perfect crystal case.

4.2 Calculation of free energies by FPMD

The previous discussion applies to any MD simulation. In the case of FPMD some more work is needed for implementing the method. In fact the ionic potential is not a preset function of the ionic coordinates, but it is explicitly computed from the instantaneous electronic ground state. This implies that to compute the free energy of the vacancy not only an ion has to be adiabatically (here the word "adiabatically" is used in the meaning it assumes in thermodynamics) decoupled from the others and from the electrons, but even a corresponding number of electrons. In the case of silicon, which we treat in the pseudopotential scheme, we have therefore to decouple a pseudoion plus four electrons (two states). This amounts to the following:

i) Multiply by λ the ionic charge and the bare pseudopotential of the ion to be decoupled.

ii) Modify the occupation numbers of the two uppermost electronic states so that they become $f_i = 2\lambda$.

This leads to write:

$$\rho_\lambda(\mathbf{r}) = 2\left(\int \mathrm{dr} \sum_{i=1}^{M-2} |\psi_i|^2 + \int \mathrm{dr} \sum_{i=M-1}^{M} \lambda |\psi_i|^2\right)$$

$$H_I = K_I[\{\dot{\mathbf{R}}_I\}] + V_I[\{\mathbf{R}_I\}; \lambda Z^0, Z^1] + E_{KS}[\{\psi_i\}, \{\mathbf{R}_I\}; \lambda]$$
$$+\frac{1}{2}M\dot{\mathbf{r}}_0^2 + (1-\lambda)\left[E_0 + \frac{1}{2}M\Omega^2(\mathbf{r}_0 - \mathbf{a})^2\right]$$

$$E_{KS}[\{\psi_i\}, \{\mathbf{R}_I\}; \lambda] = -2\left(\int \mathrm{dr} \sum_{i=1}^{M-2} \psi_i \nabla^2 \psi_i + \int \mathrm{dr} \sum_{i=M-1}^{M} \lambda \psi_i \nabla^2 \psi_i\right)$$

$$+2\left(\sum_{i=1}^{M_2}\langle\psi_i|\hat{V}^{ps}(\lambda)|\psi_i\rangle+\sum_{M_1}^{M}\langle\psi_i|\hat{V}^{ps}(\lambda)|\psi_i\rangle\right)+\frac{1}{2}\int \mathrm{d}\mathbf{r}d\mathbf{r}'\frac{\rho_\lambda(\mathbf{r})\rho_\lambda(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$
$$+\int \mathrm{d}\mathbf{r}\rho_\lambda(\mathbf{r})\varepsilon_{xc}[\rho_\lambda(\mathbf{r})].$$

Here E_0 is put equal to the energy per particle for the corresponding bulk system, and Ω roughly equal to the empirical Debye frequency ($2 10^{-3}\, a.u.$). The resulting expression for $\frac{\partial E_{KS}}{\partial\lambda}$ is reported in the appendix C.

Since the occupation numbers in this case are not all equal to 2, we scale the electronic masses by the occupation numbers as we did at page 38. The resulting lagrangian then reads

$$\begin{aligned}\mathcal{L} &= 2\mu\sum_{i=1}^{M-2}\langle\dot{\psi}_i|\dot{\psi}_i\rangle+2\mu\sum_{i=M-1}^{M}\lambda\langle\dot{\psi}_i|\dot{\psi}_i\rangle+K_I[\{\dot{\mathbf{R}}_I\}]-V_I[\{\mathbf{R}_I\};\lambda Z^0,Z^1]\\ &- E_{KS}\left[\{\psi_i\},\{\mathbf{R}_I\};\lambda\right]+\sum_{i,j}\Lambda_{ij}(\langle\psi_i|\psi_j\rangle-\delta_{ij}).\end{aligned}\tag{4.15}$$

4.3 Tests

We have performed a number of tests to assess the feasibility of the method illustrated above.

Firstly, we did some static ($T=0$) tests. In such a zero-temperature calculation the formation free energy reduces to the formation energy. Of course, no statistical uncertainty is introduced, since the thermal average reduce to the value of Φ (equation 4.8) corresponding to the minimum energy configuration.

We took a small cell ($N=32$), used a plane-waves cutoff of 6 *Ryd*, and the Γ point of the cell. Only the s non locality ($l_c=1$) was used. We computed the formation energy of the vacancy with the usual method (like in the appendix E) and compared that value with the one obtained by using equation 4.12.

We verified that to obtain the correct value for E_f ($3.62\ eV$) with an accuracy of about $0.05\ eV$ it is sufficient to use a 3-point Gauss-Legendre formula (as found in reference [94]). The points in this case are

$$\begin{aligned}\lambda_1 &= 0.1127017, \text{ weight } 5/18\\ \lambda_2 &= 0.5\qquad\quad, \text{ weight } 8/18\\ \lambda_3 &= 0.8872983, \text{ weight } 5/18.\end{aligned}$$

We checked also that the orbital energies assigned to the states M, $M-1$ were the highest among the occupied states. From equation 3.32 we may argue that, in order to minimize the energy, these states must be the highest energy states. Actually, this can be shown in the general case. Of course, this does not necessarily happen during the dynamical runs since the simulated system does not stay exactly on the BO surface; however we verified that even in these runs the average values of the orbital energies had the right ordering.

In figure 4.1 the results of a calculation similar to that illustrated in figure 3.1 are shown. The full line shows the value of $\Phi(\lambda_1)$ computed by evolving the system according to the FPMD dynamics (equations 3.13, 3.14). The broken curves show the values obtained by evolving the system by minimizing the electron energy E_{KS} at each time step by using n steepest descent steps (equations 3.16) and computing the forces on ions via the formula 3.14.

The slow oscillation, which is present in all the curves, comes from the ionic dynamics. The fast oscillation present only in the continuous curve (FPMD dynamics) is due to the fast electronic oscillation around the instantaneous electronic ground state. Apart from this fast oscillation, which should average to zero as in the case of the forces on the ions, as n increases the broken curves converge to the full curve. The curves obtained with $n = 20$ or more lie within the oscillation band due to the fast part of the electron dynamics. This shows that, if the system is evolved via the FPMD dynamics and care is taken that the electrons do not drift away from the BO surface, the values obtained in this way for $\Phi(\lambda)$ are the correct values.

Although we found the system to maintain a good adiabaticity, i. e. to keep a small value for K_f, this it is not sufficient by itself to obtain accurate enough values for Φ. In fact, this quantity is dominated by the few electrons ($i = M, M-1$) with occupation numbers multiplied by λ. They are in addition in the highest energy states, that in any system are the most delicate from the point of view of adiabaticity since they oscillate with the lowest frequency (see the equation 3.32). The consequence is that any small deviation from the desired behaviour of the system is reflected in the value of Φ. For instance, we have seen in the figure 4.1 that the small deviation of the electronic system from the BO surface due to the fast part of the electron dynamics, whose effect usually can be hardly seen

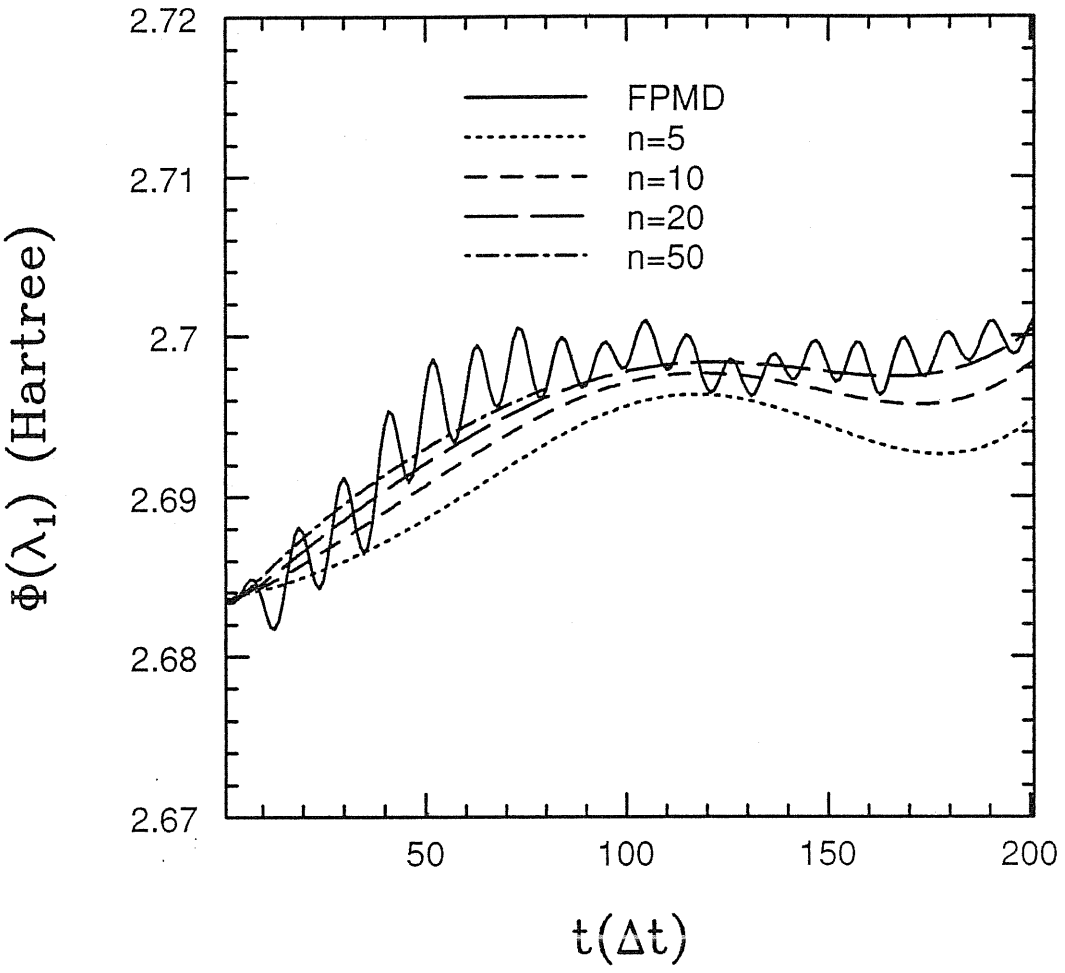

Figure 4.1: $\Phi(\lambda_1)$ computed via FPMD dynamics (continuous line) and computed via n electronic energy minimization steps at each ionic time step, as a function of time. $\Delta t = 10$ *a.u.*, $\mu = \overline{\mu} = 875$ *a. u.*.

in quantities such as E_{KS}, affects in a perceivable way the value of Φ. Actually, differential quantities, like this, are often prone to big numerical errors.

The implication is that a departure from the BO surface which could be easily tolerated in most simulations cannot be allowed here. Therefore a special care must be used in choosing the parameters for the electron dynamics, i. e. Δt, μ, and the schedule of the periodic minimizations.

We choose $\Delta t = 10$ $a.u.$, $\mu = 875$ $a.u.$. By a careful analysis of several 1000 $^{\circ}K$ runs, we found that using these parameters and putting the electrons back again on the BO surface each 100 steps, the error on F_f due to the departure from the BO surface is less than 0.1 eV, which implies an error of less than 1 k_B in the formation entropy. This error is definitely smaller than the statistical error.

Finally, we used the Nosè thermostat to obtain canonical averages. The "mass" $\mathcal{Q}$ was chosen according to the Nosè recipe [73], which in the present case gives $\mathcal{Q} = 250000$ $a.u.$.

Chapter 5

Results

5.1 Vacancy concentration

Recently, results for the vacancy formation free energy based on the LHA [9] described in chapter 4 have been obtained [10]. The vibrational frequencies have been obtained by using the FPMD method, a 64 atom cell, Γ point sampling of the Brillouin Zone and $E_{cut} = 8\ Ryd.$ The calculated formation entropy was $S_f \simeq 5\ k_B$, a quite high value.

This calculation does not suffer of any statistical uncertainty, but its validity depend on the validity of the approximations of harmonicity and of independence of the atomic vibrations discussed in chapter 4, which for silicon have not yet been tested.

We performed a calculation of F_f by using the Hamiltonian manipulation method described in chapter 4. This methods takes fully into account the anharmonic effects. Such calculations need very long simulation runs in order to reduce the statistical error to acceptable values. The simulations are still in progress and we can show only the results obtained by a 1000 $^\circ K$ run, an intermediate temperature according to the definition we gave in chapter 2. The values of $\Phi(\lambda)$ as a function of the simulation time for the three values of λ considered (see section 4.3) are shown in figure 5.1. The initial equilibration period of about 1 ps is not shown here. As it can be observed in figure 5.1, the data collected in the run with $\lambda = \lambda_1$ show larger fluctuations than the others. In order to better visualize these fluctuations, we plot in figure 5.2 the data of a longer part of the simulation with λ_1, together with the cumulative average.

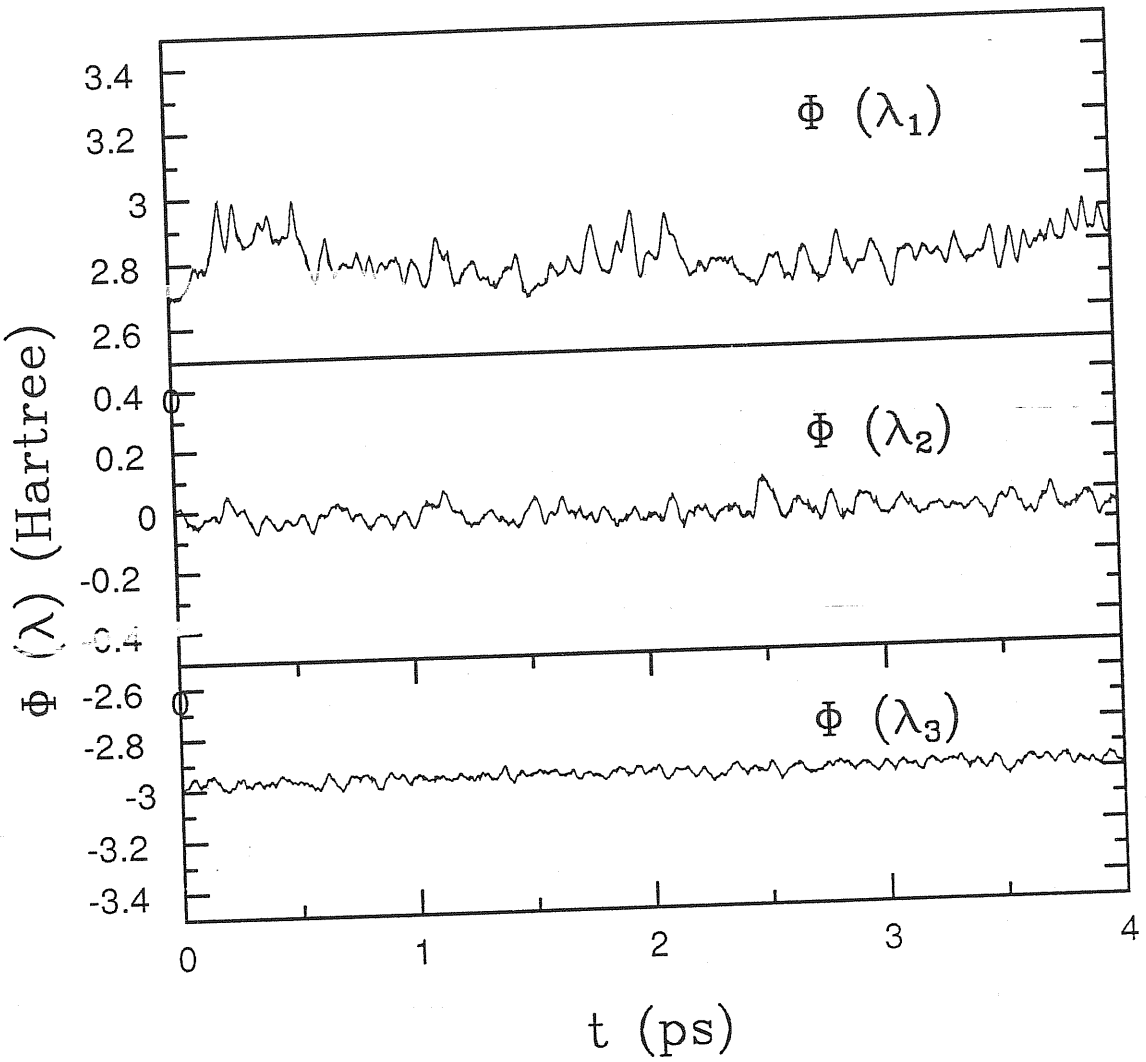

Figure 5.1: Values of $\Phi(\lambda_1)$ (top), $\Phi(\lambda_2)$ (center), $\Phi(\lambda_3)$ (bottom), as a function of time. The total simulation time shown here is 4 ps.

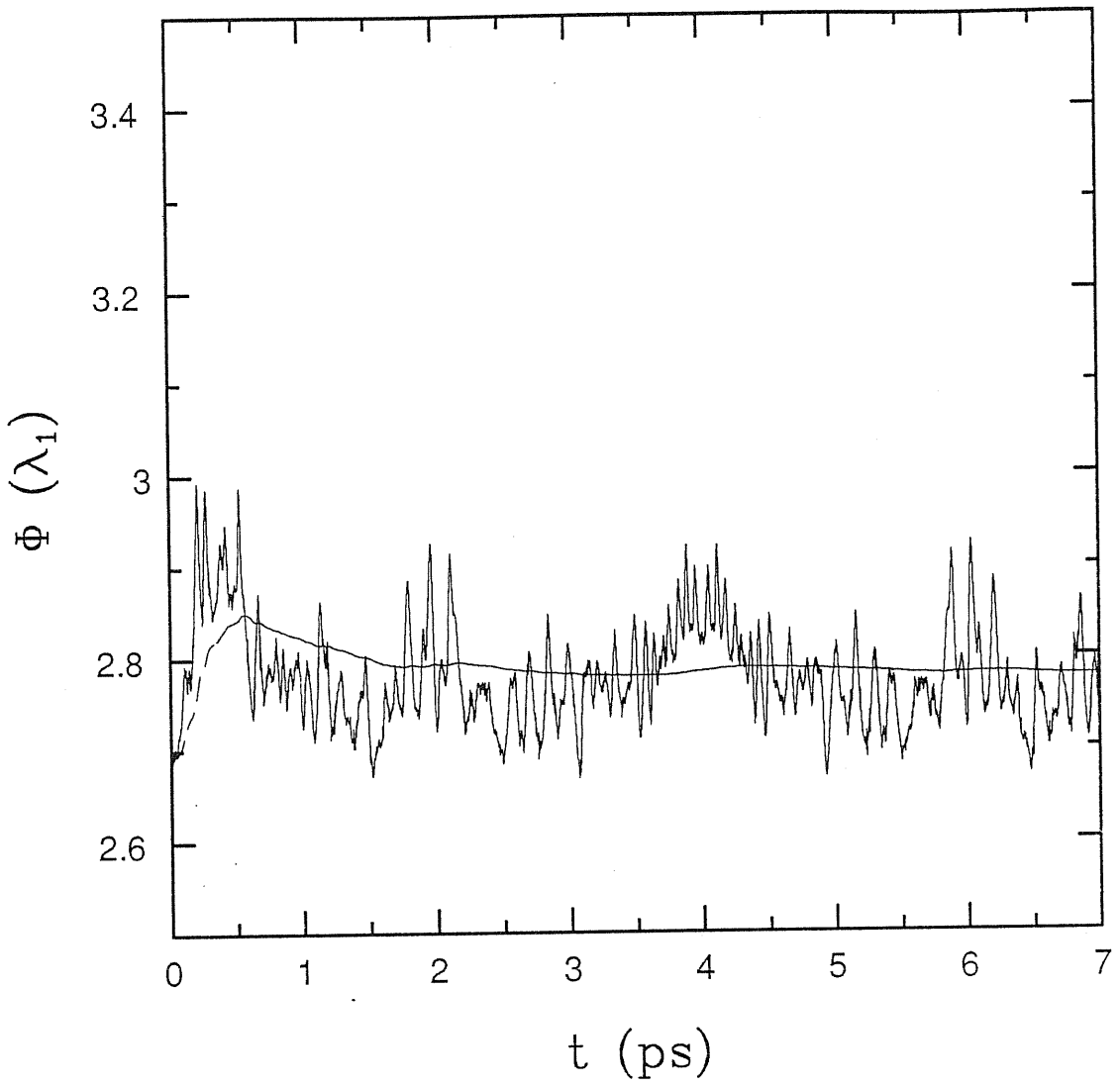

Figure 5.2: Values of $\Phi(\lambda_1)$ as a function of time. The total simulation time shown here is 6 ps.

The total error on F_f can be estimated as

$$\sigma_{tot} = \int_0^1 \mathrm{d}\lambda\, \sigma_m(\lambda) \tag{5.1}$$

where $\sigma_m(\lambda)$ is the standard deviation of the mean of the statistical sample $\Phi(\lambda)$. A rough estimate of σ_m was obtained by directly observing the fluctuations of the mean value of Φ computed over blocks of 1000 steps each (0.24 ps). This provides us with the estimate $\sigma_{tot} \simeq 0.2\ eV$, which is equivalent to an error of $\sim 2\ k_B$ on S_f.

The factor $\Delta F = \frac{1}{N} F_B^N - F_0^1$ was estimated in two ways. First we used the harmonic approximation together with phonon frequencies obtained by accurate first-principles calculations [95] to compute F_B^N. Alternatively, the value of F_B^N can be directly extracted from experimental data [96]. Both methods give $\Delta F \simeq 1\ k_B$.

The value we obtain for the formation entropy is $\sim 8\ k_B$, a value that lies within the experimental range. In figure 5.3 we plot the calculated vacancy concentration vs. $1/T$. We used the value of 3.5 eV for the zero-temperature E_f and both the LHA and our fully anharmonic results for F_f. We also report an extrapolation of our "exact" results at 1000 $^\circ K$ based on the assumption that both E_f and S_f are independent of the temperature. The error bars in the LHA results are due to our estimated $\sim 0.2\ eV$ uncertainty in E_f. The error bars in the fully anharmonic result at 1000 $^\circ K$ are due to both the uncertainty to E_f and the statistical error. The experimental range [11] is also shown by the shaded area; we see that both the theoretical results agree well with the experiment.

5.2 Vacancy diffusivity

5.2.1 Rate theory

A largely used method to compute the migration rates is the rate theory described in section 2.1.4. This theory gives a value for the migration rate which becomes more and more correct as the temperature is lowered. The inputs of the theory are, once a migration mechanism has been chosen, the migration enthalpy (usually approximated with the migration energy) and the vibrational frequencies around the minimum energy point and the saddle point.

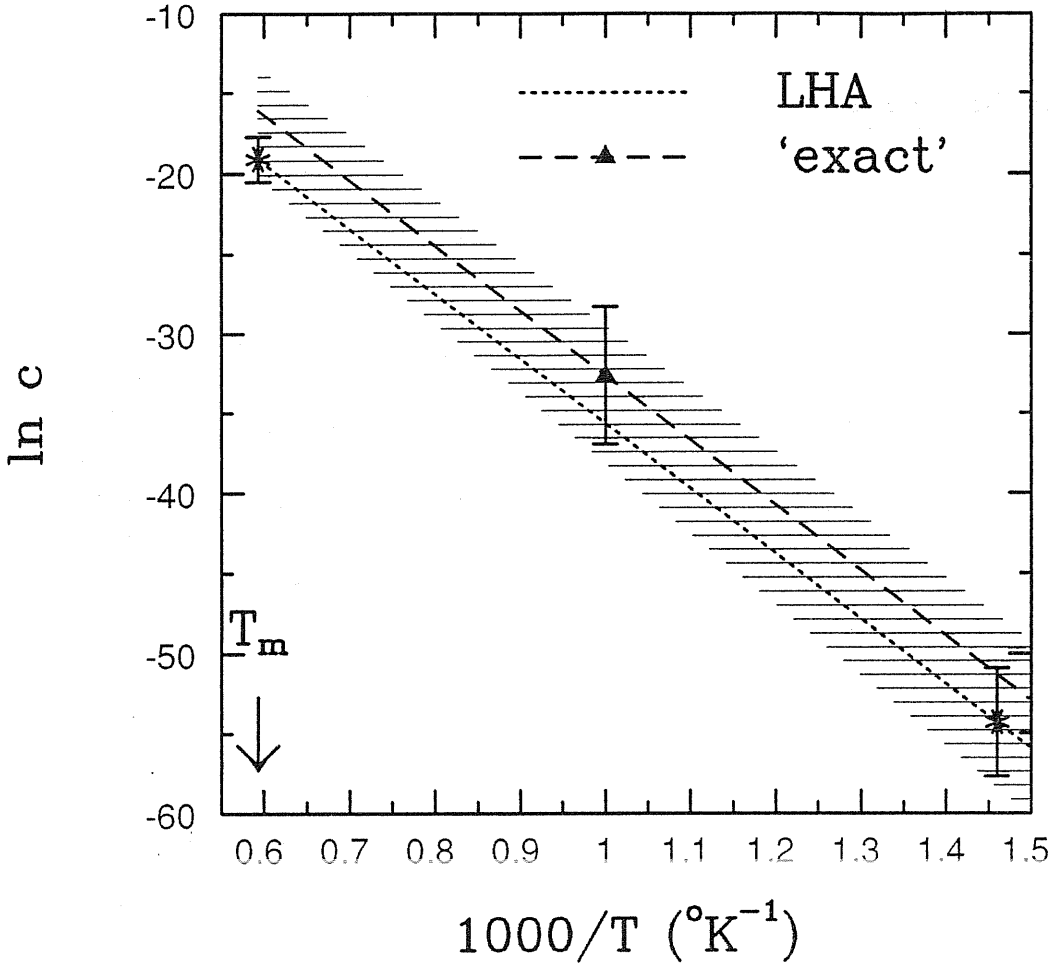

Figure 5.3: Vacancy concentration as a function of the temperature. Square: present fully anharmonic calculation. Dashed line: extrapolation (see text). Dotted line: LHA results. Shaded region: experimental range.

For our rate theory computation we choose as the migration mechanism the simple vacancy jump, i. e. the jump of an atom, first neighbour of the vacancy, in the vacant site (see figure 5.5 B). This is the most likely migration path for the vacancy at low temperature. For this mechanism, the lowest-energy migration path is a straight line connecting the vacant site to the original site of the jumping atom. We indicate by $\hat{\jmath}$ the unit vector along this line and we label with the number 1 the jumping atom.

We define a "reaction coordinate" [15] for the simple vacancy jump the quantity

$$\xi = \frac{1}{6}\sum_{I=2}^{7}(\mathbf{r}_I - \mathbf{r}_1)\cdot\hat{\jmath} \tag{5.2}$$

where the sum involves the 6 atoms that, before the jump, are either first neighbours of the vacancy or of the atom 1. The value of ξ ranges from a negative value $-d/2$ to a positive one $d/2$ (d being the bond length), which both correspond to a vacancy configuration. The value $\xi = 0$ indicates a plane including the so-called "split vacancy" configuration, i. e. the configuration in which the atom 1 sits halfway between the two end points of the migration path. This configuration corresponds to the saddle point configuration.

A computation of the vacancy jump frequency according to the full rate theory is quite expensive if one uses a first-principles scheme, since it requires knowledge of the full vibrational spectrum of the system (see the discussion in chapter 4). We therefore limit ourselves to a very simplified form of the rate theory. We suppose that the vibrational mode along $\hat{\jmath}$ is an eigenmode and that the system eigenfrequencies corresponding to all other vibrational modes are independent of ξ. Therefore equation 2.7 may be rewritten as

$$\Gamma_{RT} \simeq 4\frac{\nu}{2\pi}\,\mathrm{e}^{-E_m/k_BT}. \tag{5.3}$$

where the factor 4 has been introduced in order to take into account that there are 4 saddle points for the simple vacancy jump: this is in fact the number of first neighbours of the vacancy. Here E_m is the migration energy and ν the frequency along $\hat{\jmath}$, i. e. the attempt frequency. Both quantities can be computed with static total energy minimizations. Figure 5.4 shows the total energy as a function of ξ. In the computations we allowed full relaxation of all the atomic coordinates at fixed ξ. From the curvature of $E(\xi)$ we extracted a frequency $\nu = 5\,10^{-4}\ a.u. = 20\ THz$, which is about

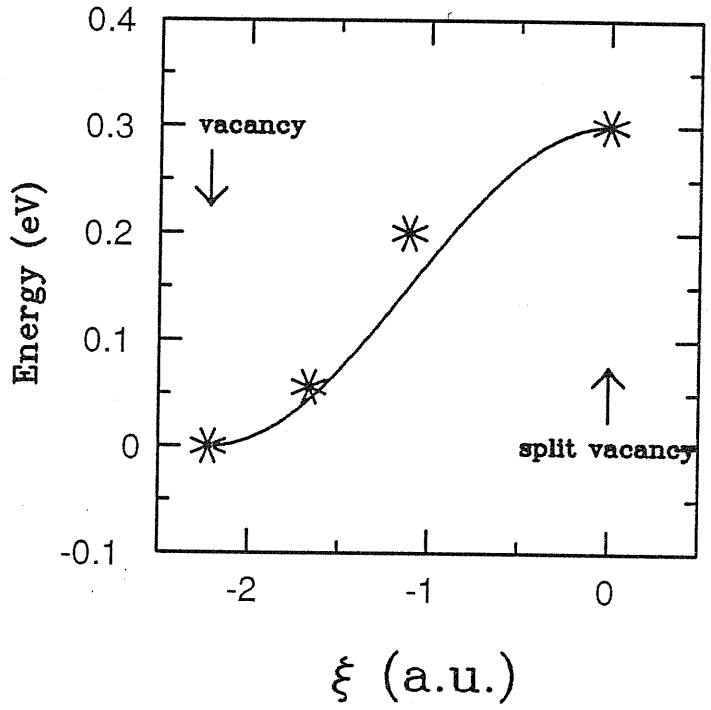

Figure 5.4: Total energy of the system, in eV, as a function of the reaction coordinate 5.2. The continuous line is a guideline for the eye and represents a cosine-like potential with the same activation energy.

1/4 of the Debye frequency of bulk silicon and lies within the acoustic peak of the phonon density of states.

Having ν and E_m we can compute the self-diffusion coefficient using formula 2.2:

$$D_{RT} = c\frac{f}{6}4d_1^2\frac{\nu}{2\pi}\,\mathrm{e}^{-E_m/k_BT} \tag{5.4}$$

where c is the vacancy concentration, $f = 1/2$ (random walk of the vacancy), and $d_1 = 4.44\ a.u.$ is the first-neighbour distance. The resulting self-diffusion coefficient for unit vacancy concentration is $D'_{RT} = D_{RT}/c = 6.1\,10^{-4}\ e^{-3480/T}\ cm^2/s$ (temperature in degrees Kelvin).

5.2.2 Dynamical runs: generalities

To study the dynamical behaviour of the vacancy we performed four MD runs on the defective system at different temperatures.

In these computations the requirement of adiabatic behaviour, although important, is less stringent than in the computation of the vacancy formation free energy, which requires the evaluation of energy derivatives. We have used a timestep of $\Delta t = 20\ a.u. = 0.48410^{-3}\ ps$ and an electronic mass $\mu_i = \mu = 3500\ a.u.$. These parameters, taking for example the 1400 $°K$ run, give after 500 time steps an increase of K_f of about 0.04 $Hartree$, which is slightly less than 1/10 of the ionic temperature. From our experience and from the experience of FPMD calculations in other systems, this is enough for an accurate description of the atomic trajectories.

The temperature chosen for the runs were 1200, 1400, and 1600 $°K$. In this way we span the experimentally accessed temperature range of self-diffusion. The temperature of the system was controlled by a Nosè thermostat with $\mathcal{Q} = 250000\ a.u.$.

In order to improve on the statistics, we performed longer runs at the lower temperatures, where the jump probability is smaller. The run lengths were: 1200 $°K$ = 13.55 ps, 1400 $°K$ = 13.06 ps, 1600 $°K$ = 7.74 ps, and 1750 $°K$ = 1.74 ps.

We checked approximately the thermal equilibration of the system by observing the power spectrum of the system dynamics $G(\omega)$:

$$G(\omega) \propto \int \mathrm{d}t\,\mathrm{e}^{i\omega t}\langle \mathbf{v}(0)\cdot\mathbf{v}(t)\rangle \tag{5.5}$$

where $\mathbf{v}$ are the ionic velocities and the angular brackets indicate canonical averages. We found that a run length of about 2 ps was sufficient to obtain a stable $G(\omega)$. Therefore in all our runs we excluded an initial time of 2 ps from our measurements.

A first way to analyse the data consists in displaying the absolute values of the atomic displacements from the perfect crystal positions as a function of time. We call a plot of this kind an $r-t$ plot. In figure 5.5, we illustrate schematically some of the important processes that may affect atoms close to a vacancy. We also show how these processes appear on the $r-t$ plot.

The processes illustrated in figure 5.5 are:

A a thermal vibration of an atom;

B a jump of a first-neighbour into a vacancy;

C a jump of a second neighbour into a vacancy;

D an almost simultaneous jump of a first and a second neighbour of a vacancy (the time delay between the jumps is less than or of the order of the optical phonon period: $2\pi/\omega_{optical} \sim 0.07\ ps$).

Notice that both process C and process D can be viewed as vacancy double jumps. However their effect on atomic diffusion is quite different. In case C the net contribution to the total atomic square displacement, hence to self-diffusion, is $d_2^2 = 52.7\ a.u.$, while in the case D the contribution is $2d_1^2 = 39.4\ a.u.$. Here $d_1 = 4.44\ a.u.$ is the first-neighbour distance, and $d_2 = 7.26\ a.u.$ is the second-neighbour distance. We see that the net contribution of process C is about 30% larger than that of process D.

5.2.3 1200 $°K$ and 1400 $°K$ runs

Let us examine the $r-t$ plots concerning the two lowest temperature runs (figures 5.6 and 5.7). We observe a certain tendency of the jumps to lump together on a relatively short time span followed by relatively long periods where no jumps were detected. The clustering of the jumps in time may indicate the presence of dynamical or memory effects. It may happen that two consecutive jumps are separated by about 2 ps, while in certain time spans it happens several times that two jumps are almost simultaneous (these are vacancy double jumps: see figure 5.5 D).

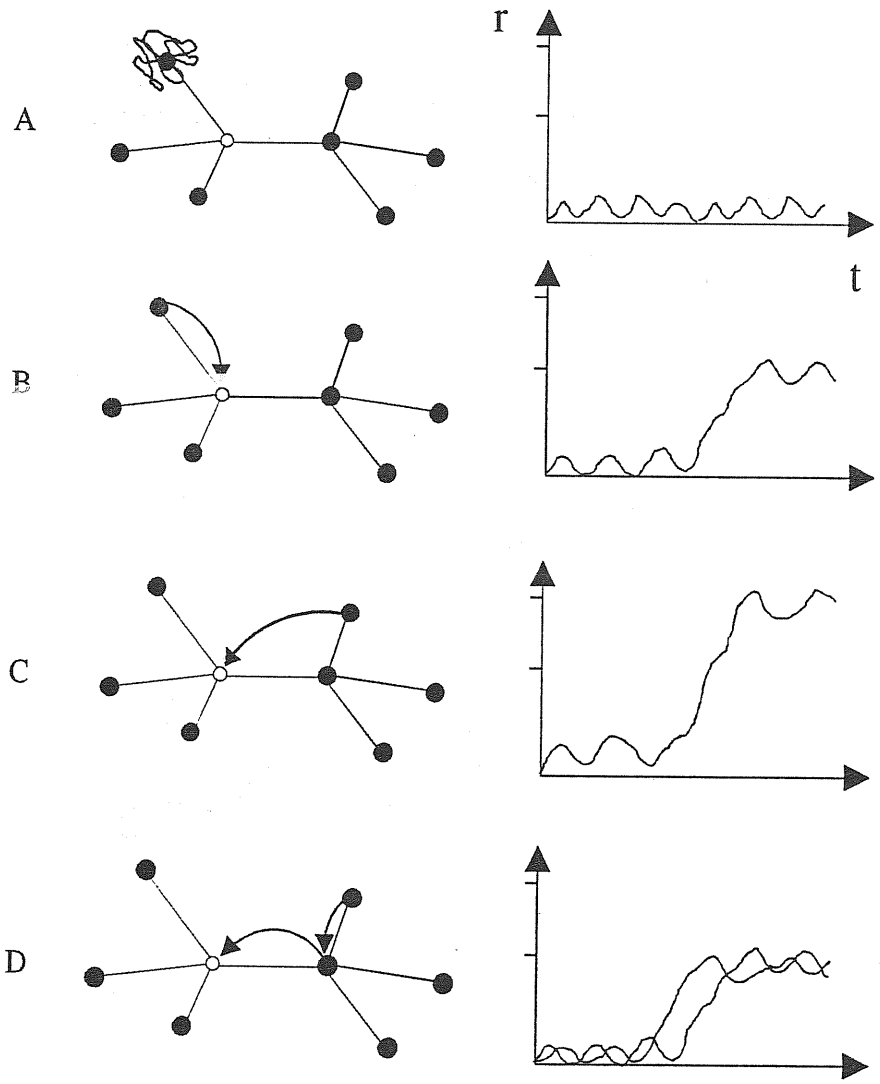

Figure 5.5: Dynamical behaviour of atoms in a diamond lattice in the presence of a vacancy. On the right the corresponding representation on a $r - t$ plot is shown; r is the absolute value of the atomic displacement as a function of time. The ticks on the r axis represent the first and second neighbours distance in the lattice.

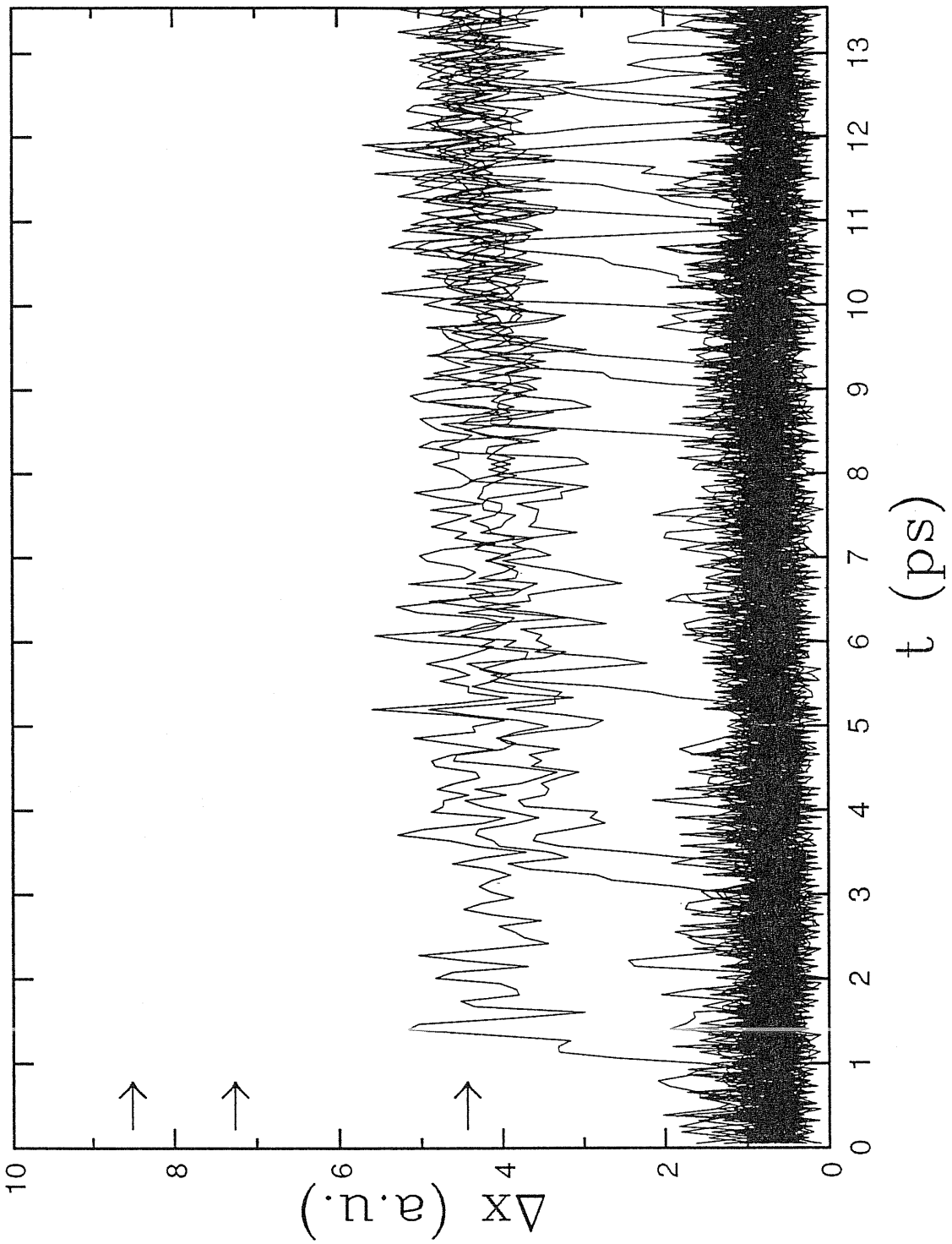

Figure 5.6: $r - t$ plot for the 1200 $^{\circ}K$ run. The arrows indicate the first, second and third neighbour distances.

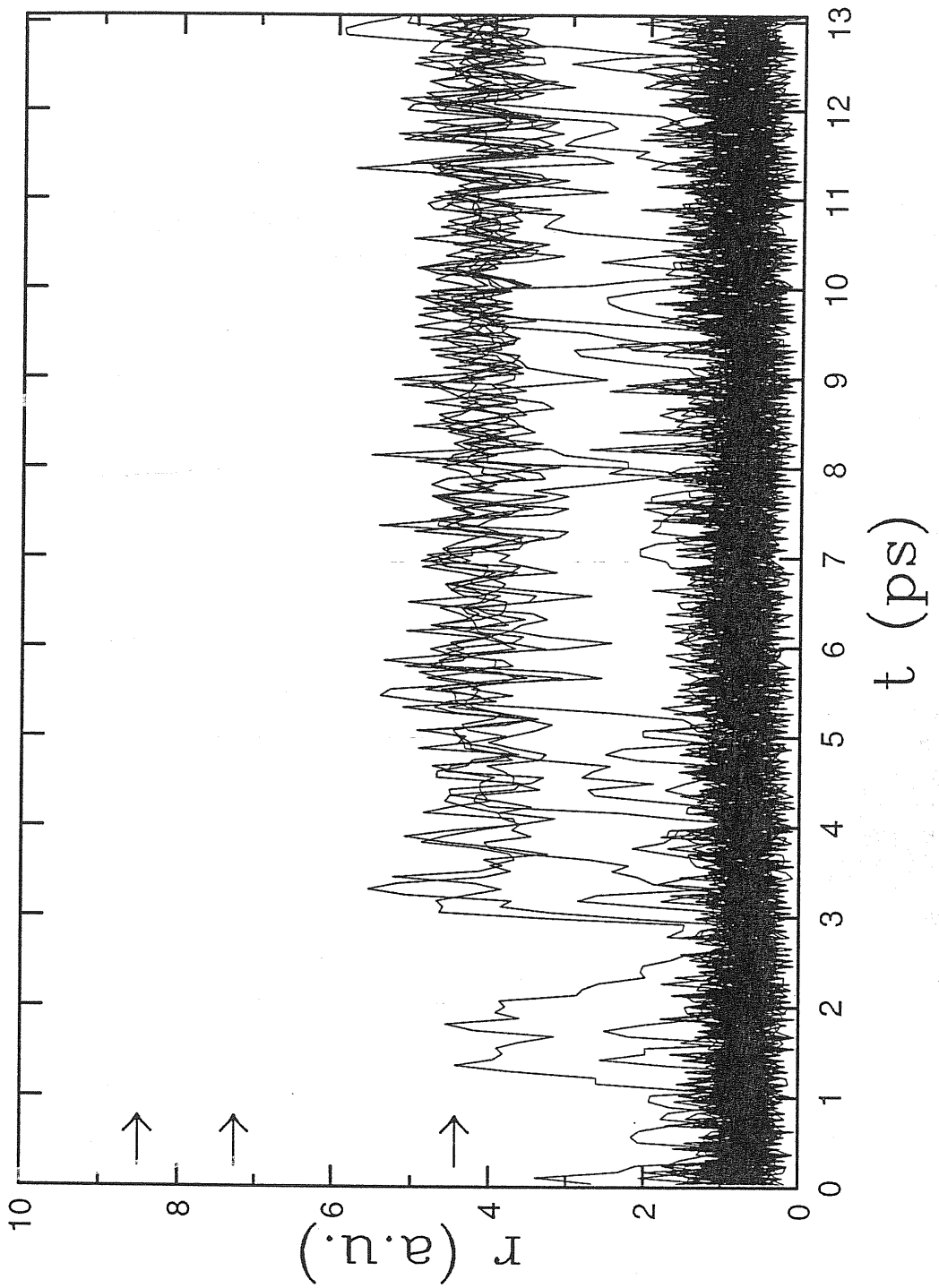

Figure 5.7: $r - t$ plot for the 1400 $^{\circ}K$ run.

Although our jump statistics is not sufficient to make an accurate analysis, we may argue that the probability distribution of the time delays among the jumps does not follow a Poisson distribution, which is expected in the case of uncorrelated jumps. This leads to suspect that dynamical effects are at work.

Actually this kind of memory effect has been observed several years ago in MD simulations of vacancy induced self-diffusion in sodium and aluminum [97]. In these simulations an anomalously high frequency of double atomic jumps was observed, i. e. a deviation from a Poisson distribution at very short times, which the authors attributed to a kind of ballistic effect.

Another, more interesting, indication of the presence of dynamical effects shows up from the observation of the spatial distribution of the jumps. Figure 5.8 represents the trajectories of the atoms lying on a crystallographic plane. We see that 7 vacancy jumps take place *along the same chain of bonds*. This means that the vacancy is not moving via a random sequence of uncorrelated simple jumps. The tendency of the vacancy to perform several jumps along the same chain of bonds does not seem to be an occasional effect, but rather seems to be the most common situation in our runs at 1200 $^{\circ}K$ and 1400 $^{\circ}K$. The correlation effect may be quantified by computing the correlation factor f as discussed in section 2.1.1. We get $f_{1400} = 0.75$ and $f_{1200} = 0.63$, while in absence of dynamical effects, f should be equal to 0.5. In the present case, the increase of f compared to a pure random walk is due mainly to a reduction of the ratio of return jumps to forward jumps.

We notice that the dynamical effects observed here seem different from those reported in reference [97], which depended on the correlation between a jump and the previous one. In our case there is a correlation at least with the *two* previous jumps, since two jumping directions are necessary to identify a specific chain of bonds lying on a crystallographic plane. We argue that this peculiar dynamical behaviour in silicon is linked to the specific structure of the diamond lattice and of the covalently bonded network.

5.2.4 1600 $^{\circ}K$ run

Apart form the effect of correlations, the diffusion mechanisms at the two temperatures considered above does not appear to be radically different from what one expects at lower T, i. e. a sequence of single vacancy jumps.

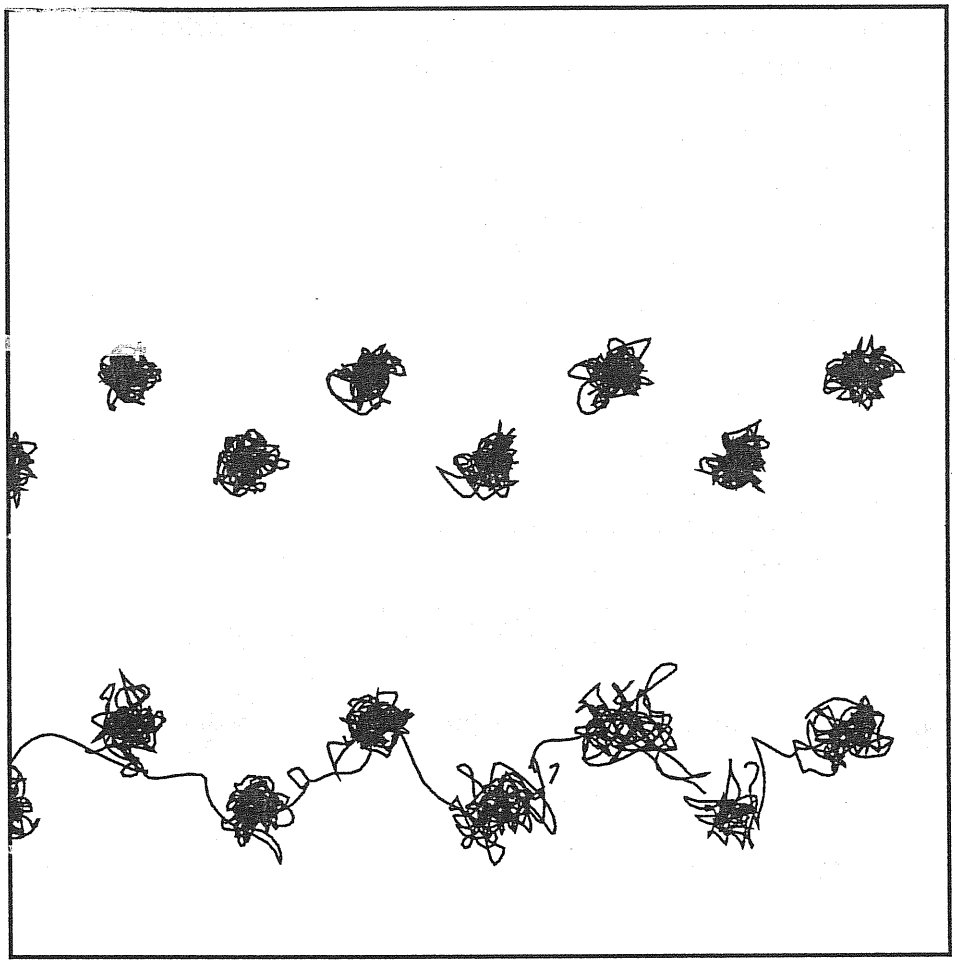

Figure 5.8: Atomic trajectories in the 1400 $^{\circ}K$ run for the atoms in the $(1\bar{1}0)$ plane. The trajectories have been projected on the plane $(1\bar{1}0)$ itself. The time interval during which the trajectories are recorded is $2.7 < t < 8.5$ ps.

At the temperature $T = 1600\ ^{\circ}K$, which is very close to the experimental melting point, the picture changes dramatically.

This can be seen in figure 5.9. During the limited time of this run we have observed, in addition to some standard jumps, two novel and distinct processes:

i) An atom jumps directly into a second-neighbour vacant site.

ii) Three atoms move simultaneously away from their equilibrium sites to end up in new crystalline sites after $\sim 2\ ps$. The total jump length for the three atoms corresponds respectively to a first-neighbour, a second-neighbour, and a third-neighbour distance.

We now examine these two peculiar processes in greater detail.

i) Second neighbour jump

In this case a second-neighbour of a vacancy jumps directly into it. This is pictorially illustrated in figure 5.10. In other words, at variance with what was sometimes observed in the lower temperature simulations, where a "vacancy double jump" consisted of an almost simultaneous jump of *two* atoms, a first neighbour and a second neighbour of a vacancy, here a *single* atom moves to a second neighbour site: this is a different migration channel. It seems likely that this process is favoured entropically (the neighbourhood of the vacancy during this process shows wide atomic rearrangement) but is characterized by a larger activation barrier than the nearest neighbour jump. However, rebonding effects may substantially reduce the height of the barrier.

We computed the pair correlation function $g_i(r)$, defined as the probability of finding an atom at a distance r from the atom i, for several situations. The atom i was chosen to be an atom A) far from the vacancy (see figure 5.11), B) one of the first neighbours of the vacancy in a time span when no jumps occurred (see figure 5.12), C) the jumping atom in a simple vacancy jump (see figure 5.13), and D) the jumping atom for the process *(ii)* (see figure 5.14). The coordination number of the atom i is defined as the integral of $g_i(r)$ from $r = 0$ to $r = r_c$, where r_c is the position of the first minimum of $g_i(r)$.

An atom in the bulk has coordination 4, and $r_c \simeq 5\ a.u.$. One could expect that the first neighbours of the vacancy have coordination 3, due

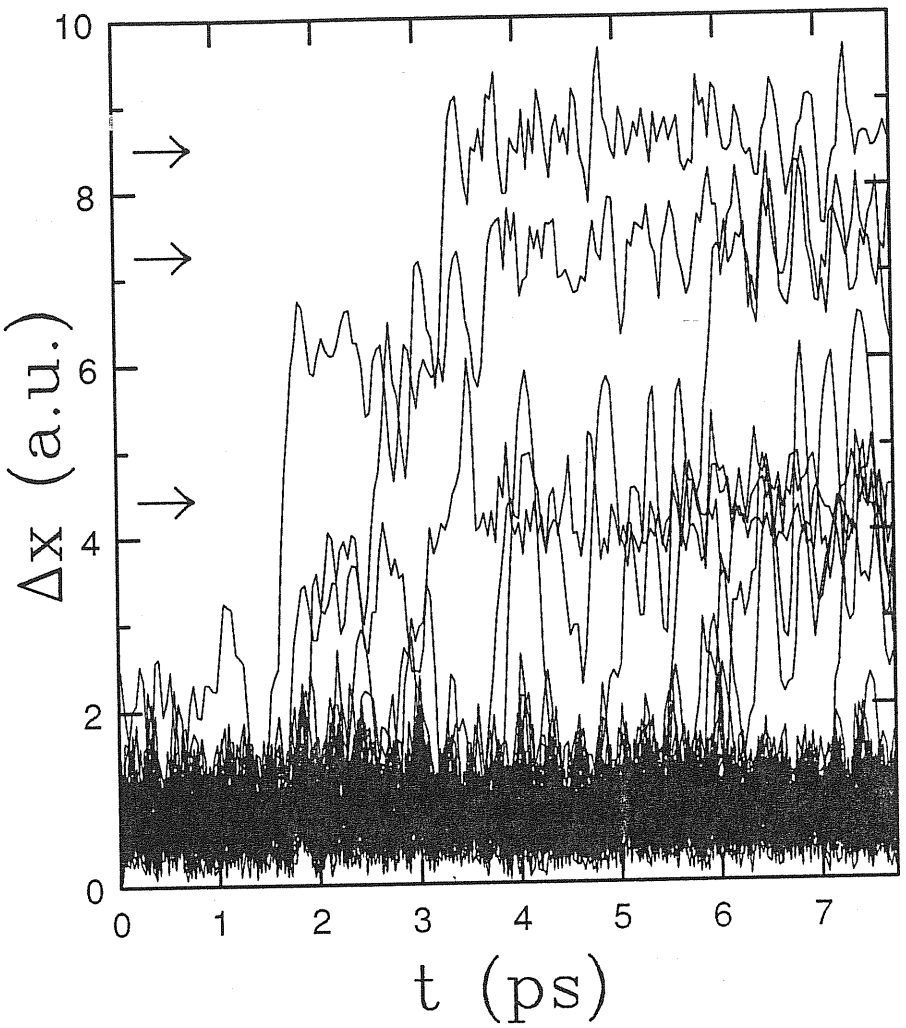

Figure 5.9: $r - t$ plot for the 1600 $^{\circ}K$ run.

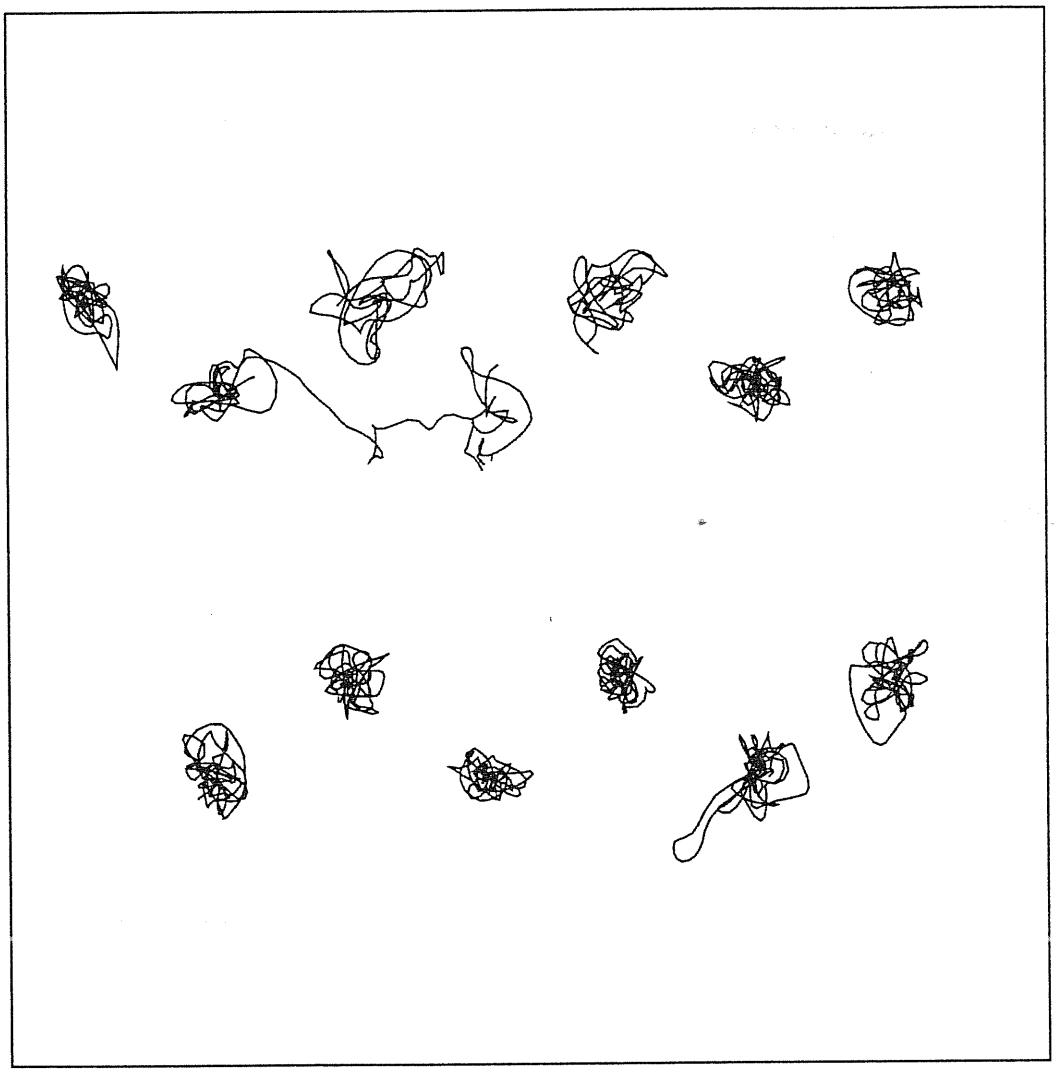

Figure 5.10: Atomic trajectories in the 1600 $^{\circ}K$ run of the atoms in the $(1\bar{1}0)$ plane. The trajectories have been projected on the plane $(1\bar{1}0)$ itself. The time interval during which the trajectories are recorded is $3.9 < t < 6.1$ ps.

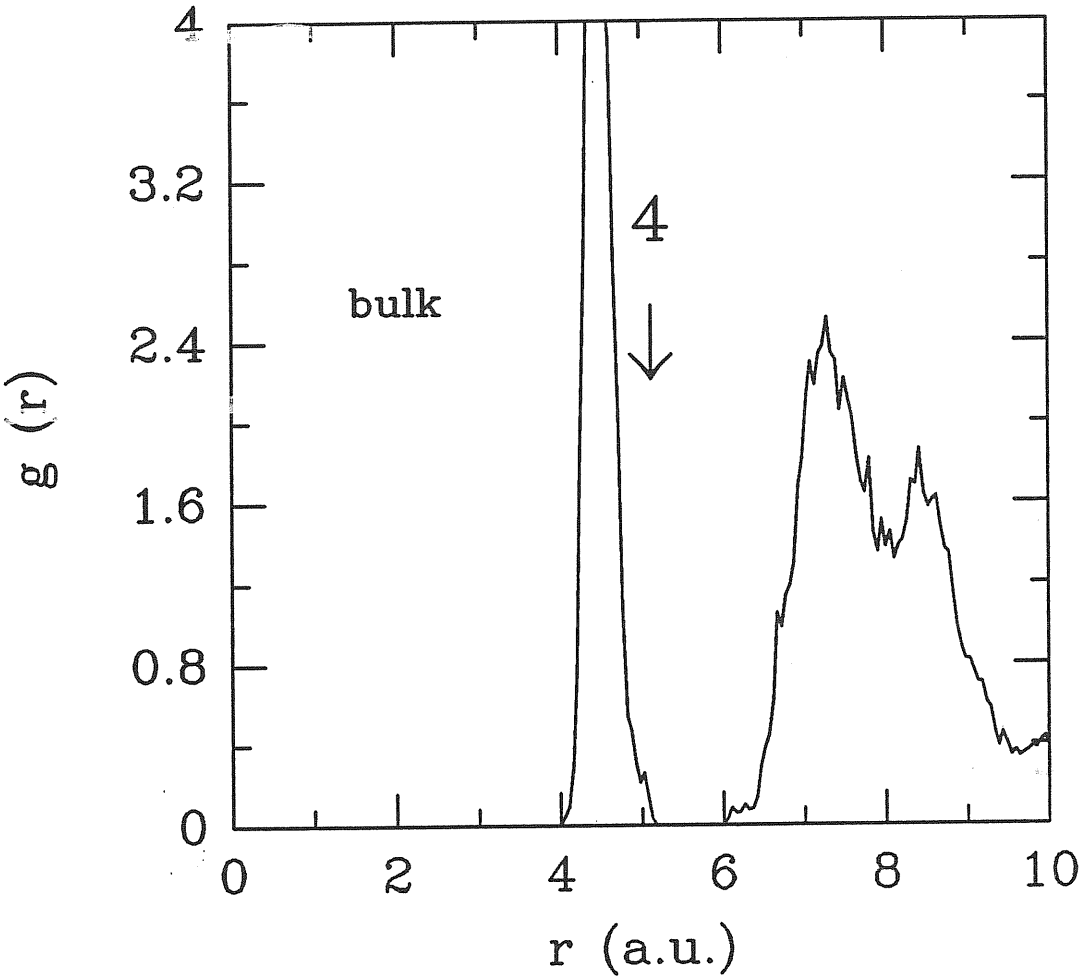

Figure 5.11: Pair correlation functions $g_i(r)$ for a bulk atom.

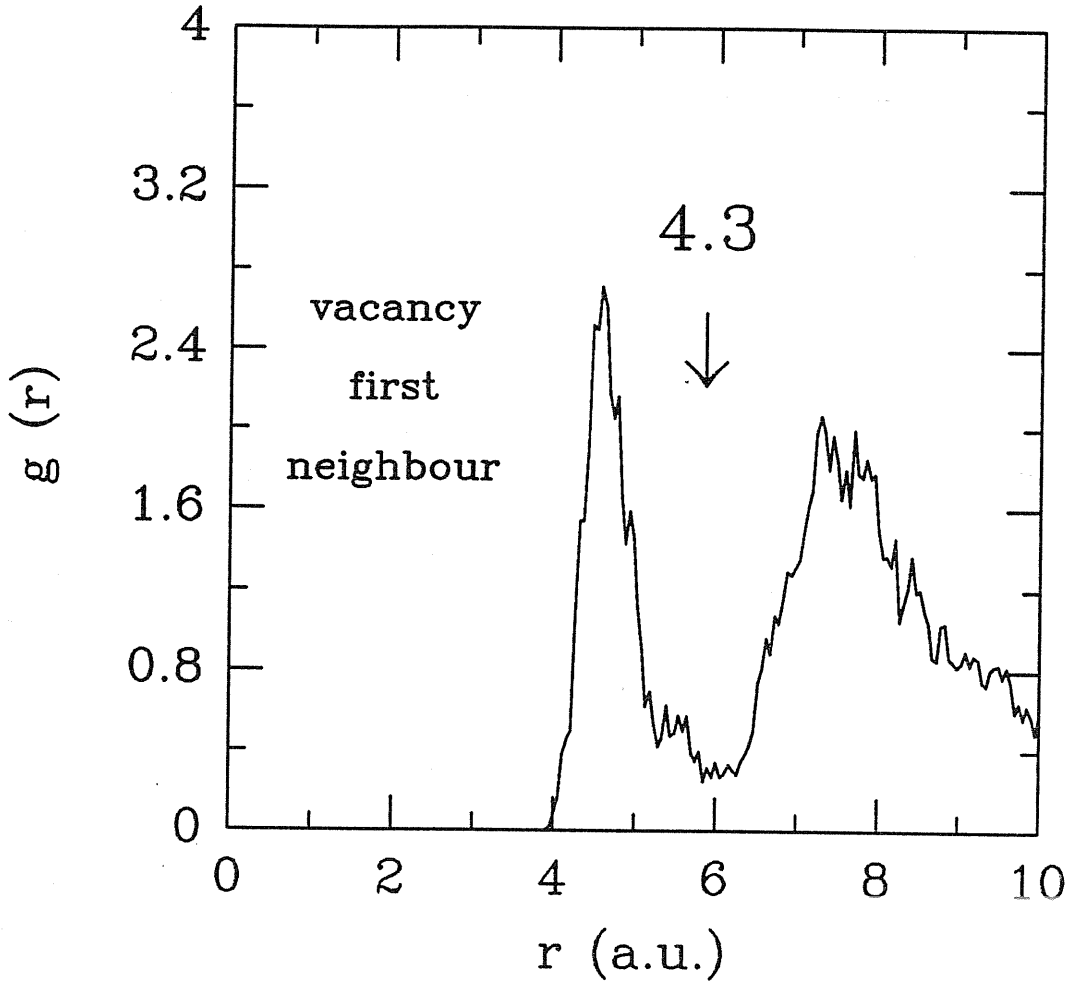

Figure 5.12: Pair correlation functions $g_i(r)$ for a first neighbour of the vacancy.

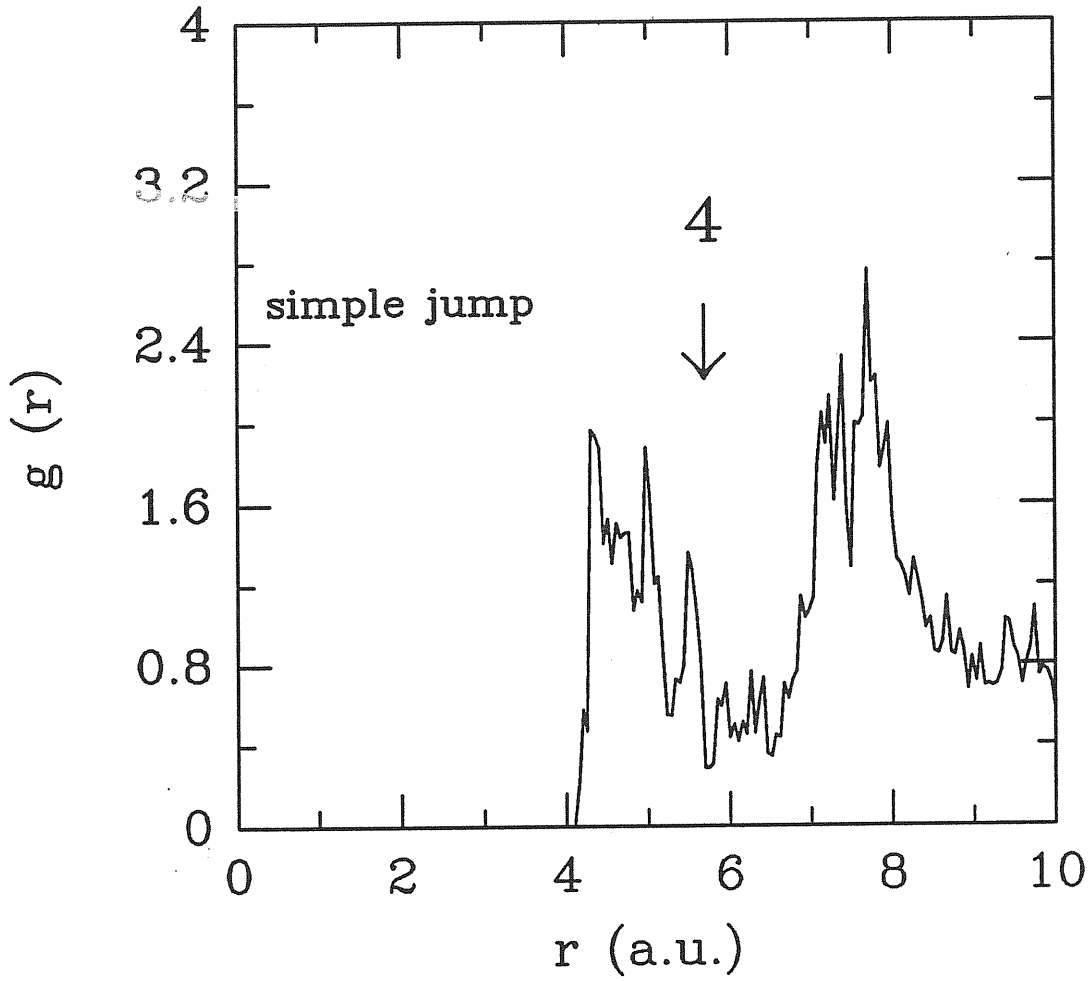

Figure 5.13: Pair correlation functions $g_i(r)$ for the jumping atom during a simple vacancy jump.

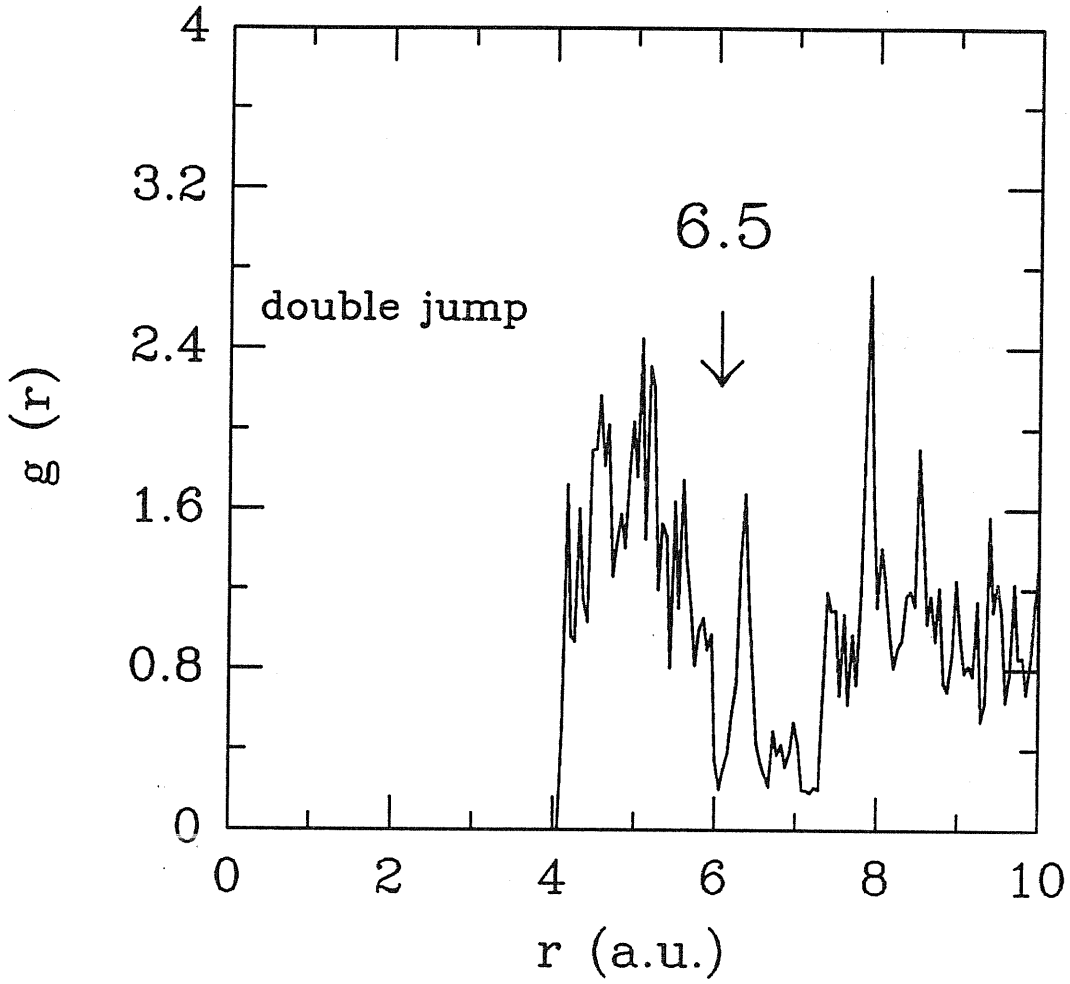

Figure 5.14: Pair correlation functions $g_i(r)$ for the jumping atom during the process *(ii)*.

to the lack of a neighbour. Instead, their coordination number is greater than 4, corresponding to $r_c \simeq 5.5 - 6$ $a.u.$. This indicates that part of the energy used to create the vacancy is recovered via the formation of weak bonds. It is interesting to note that $g(r)$ for an atom performing a simple vacancy jump show similar values for r_c and for the coordination number; this explains the low migration barrier for this mechanism.

The function $g(r)$ for the second neighbour jump was computed during the time interval in which the atom stays in the interstitial position. The plot shows that the first peak of the $g(r)$ is broadened with respect to that of a bulk atom and the coordination is raised to about 6.5, with $r_c \simeq 6$ $a.u.$. We can argue therefore that even in this case a certain amount of energy has been recovered by the formation of weak bonds, bounding the migration energy for this jump to a value not too much larger than the values for a simple vacancy jump.

ii) Concerted jump

The situation for the mechanism labeled *(ii)* is more complex. A plot of the relevant atomic trajectories (figure 5.15) suggests that the crystal is locally disordered.

A detailed analysis of the atomic trajectories allows to understand the kinematics of this concerted process. A schematic plot of the process is shown in figure 5.16. In summary, we can say that three atoms of a six-atom "ring" (shown in figure 5.16 as it were planar for clarity), which includes the vacancy, move away from their original crystal position breaking many of the original bonds, while others are created.

Within a short time, these atoms form a relatively stable structure lasting for about 2 ps, a substantial amount of the total simulation time (figure 5.16 B). The average bond angle $\langle\theta\rangle$ among the three atoms in this configuration is close to the bulk value of 109.47°, and the average coordination number as deduced by the figure 5.17, which shows the sum of the pair distribution functions around the three atoms, stays quite close to 4.

At the very end, the crystal structure is recovered (see fig. 5.16 C), and although each atom results displaced from the initial position, giving the most important contribution to the self-diffusion coefficient at this temperature ($\sum_I \Delta \mathbf{r}_I^2 = 145$ $a.u.$ over a total of 296 $a.u.$), the vacancy at the end

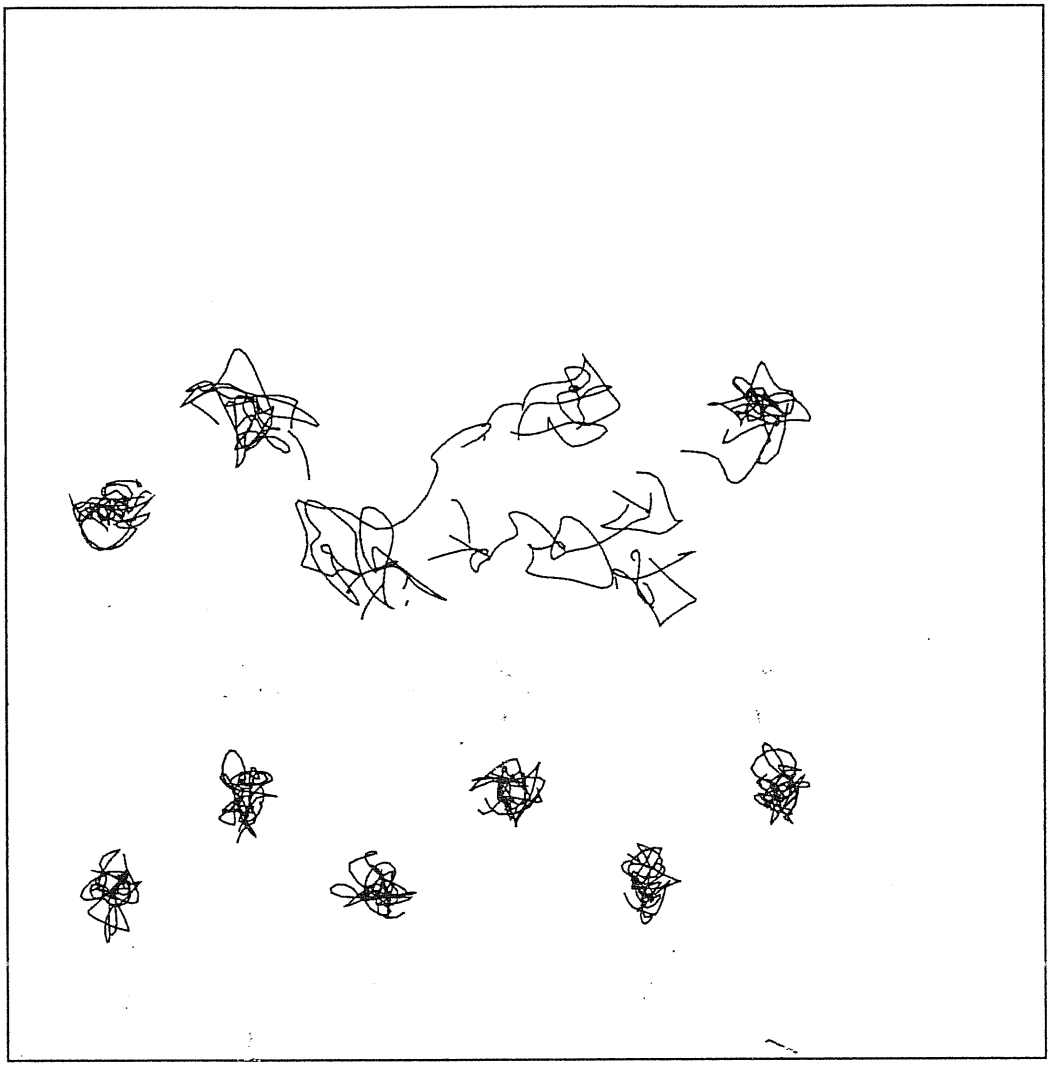

Figure 5.15: Atomic trajectories in the 1600 $^{\circ}K$ run of the atoms involved in the multiple jump and of some of their neighbours. The trajectories have been projected on the plane $(01\bar{1})$. The time interval during which the trajectories are recorded is $1.3 < t < 3.4$ ps.

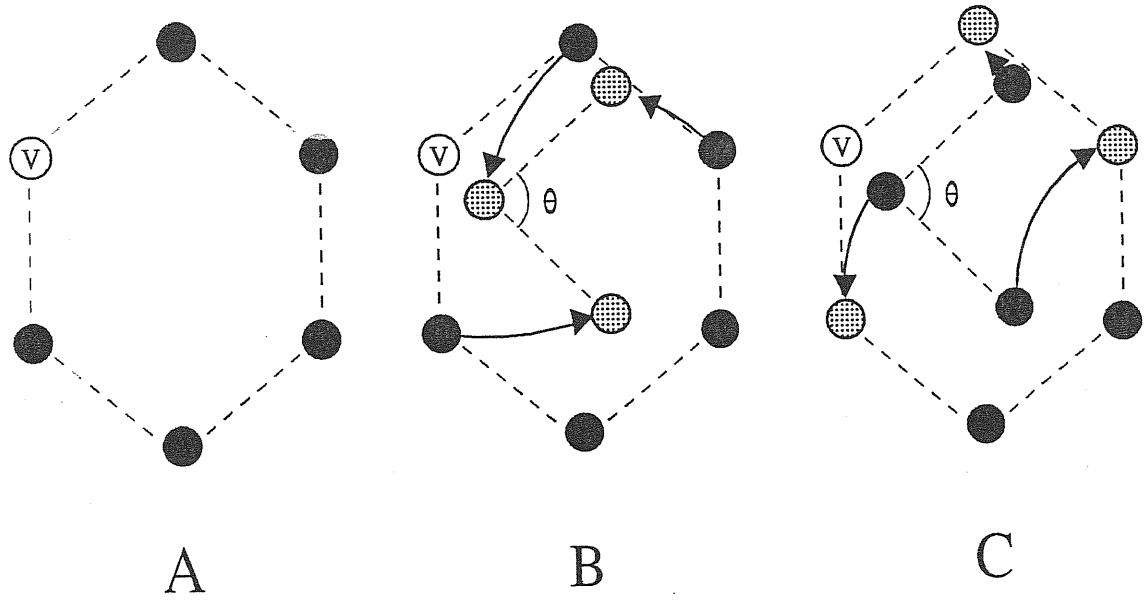

Figure 5.16: Schematic plot of the process *(ii)*. A: the "hexagon" involved in the process. B: the atoms leave their crystalline position ($t = 1.3 - 2\ ps$). C: the crystalline structure is recovered ($t = 3 - 3.4\ ps$).

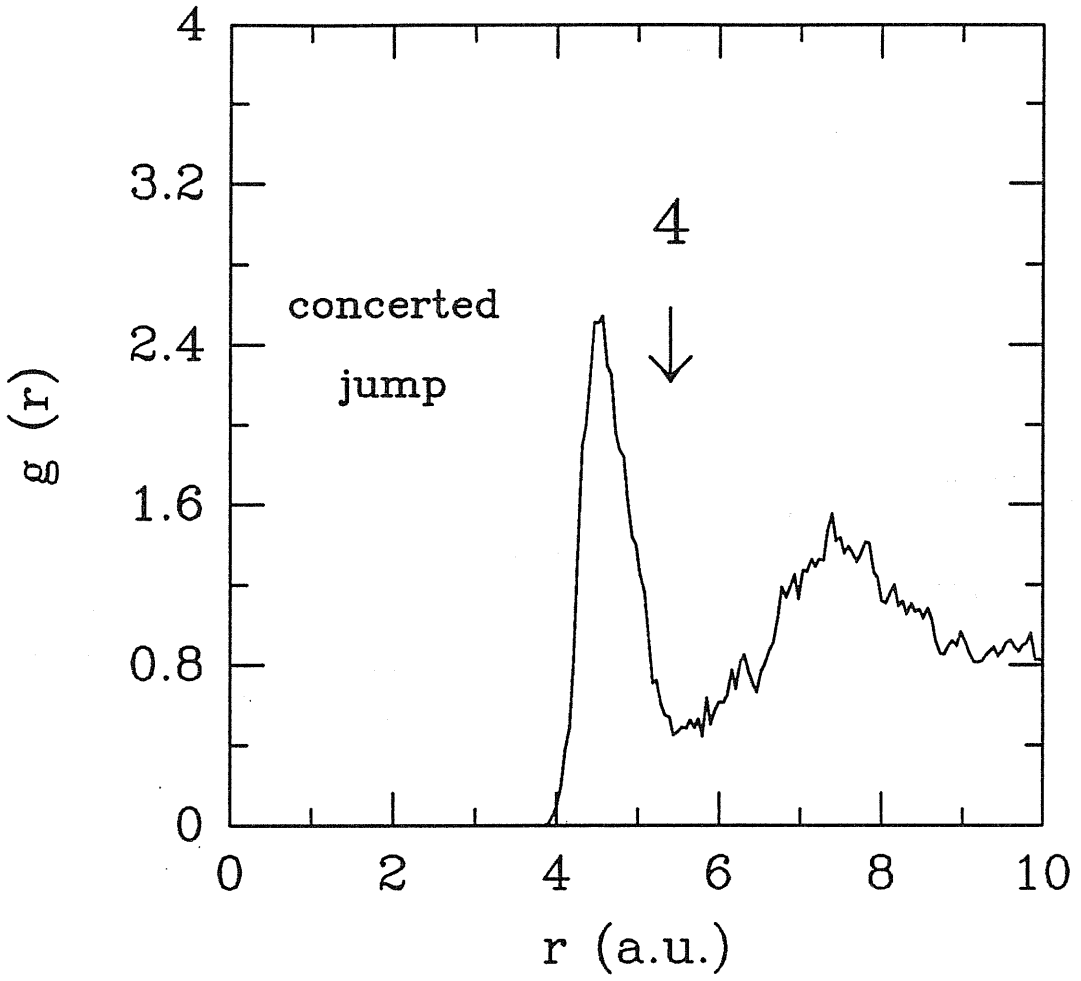

Figure 5.17: Sum of the pair correlation functions of the three atoms involved in process *(ii)*.

of this process did not move from the original position (during the process the vacancy is not even well defined). This process therefore has little to do with the vacancy mechanism as usually thought. Instead, the presence of the vacancy serves to provide a global softening of the lattice and, in addition, to make some additional room to allow atomic movements.

An analysis similar to that of the process *(i)* leads us to argue that this process, which involves the simultaneous motion of a large number of atoms, is favoured entropically but is characterized by a quite large activation barrier.

The mechanism *(ii)* is reminiscent of the so called "extended vacancy" mechanism, proposed many years ago by Seeger and Swanson [5]. According to these authors, the self-diffusion in silicon could happen via a locally melted zone around vacancies or interstitials. In these cases the defect would be "spread" over a number of lattice sites. This was claimed to allow a big diffusion entropy, in agreement with the high temperature experiments on the self-diffusion. The main problem with this proposal was the absence of a detailed mechanism for the motion of the spread vacancy.

The process detected in our simulation provides such a mechanism. The vacancy has a finite lifetime in the spread form. During this time it causes an extensive atomic diffusion around it. When the defect returns in the simple vacancy configuration it migrates via simple or double jumps. Then it may return in the spread vacancy state and cause large atomic motion at another place. This mechanism produces a large increase in the self-diffusion coefficient.

Recently, the presence of local molten zones around impurities was also suggested in a different context [98,99], to explain the anomalous high temperature behaviour observed around impurities in lead and tin.

Implications for self-diffusion

These high temperature mechanisms seem to dominate the high temperature self-diffusion behaviour. In our run at 1600 $^{\circ}K$ about 50% of the total atomic square displacement is due to them. The dynamical effects become extremely important. So we can conclude that near the melting point the picture of vacancy migration and of the related diffusivity need to be deeply changed. This is reflected in the value of $f = 1.88$ that we obtain in this case. This is very far from the value expected for a simple random walk of

the vacancy. As a consequence we observe a substantial increase of D with respect to the rate theory value.

5.2.5 1750 $^{\circ}K$ run

We tried also to simulate the diffusive behaviour of a slightly overheated silicon crystal at 1750 $^{\circ}K$. Actually we do not know at present what is the melting temperature T_m of our LDA silicon crystal; its calculation would require complicated free energy computations, both for the crystalline and the liquid phase, which at present have not been done. Indirect indications exist [100] that T_m in our system should not be very far from the experimental temperature $T_m = 1685\ ^{\circ}K$.

The presence of a vacancy in our cell favours the nucleation of melting. While in the case of the perfect crystal we can heat the system at 1750 $^{\circ}K$ this was not possible in the present case. After heating the system to 1750 $^{\circ}K$ we observed a rapid onset of diffusive behaviour (see figure 5.18) signaling transition to molten state. This temperature may be considered an upper bound for T_m of our system. The diffusion coefficient is of the same order of that computed in reference [100].

5.3 Self-diffusion coefficient

The values of $\sum_I \Delta R_I^2(t)$, where $\Delta R_I(t)$ is the distance at time t of the atom I from its original crystal position, as a function of time for the various temperatures are shown in figure 5.19. Fitting these data with a straight line it is possible, by using equation 2.2, to extract the value of D' at the various temperatures.

The estimated error of this fit is not a good estimate of the real error bar on D'. In fact, our data show long time correlations; in other words they are not independent from one another. In order to have an estimate of the true error on the result, one needs to know the value of the correlation time [101]. To this end, different methods have been developed. However they require much longer runs than ours to estimate the correlation time. To have a rough estimate of the error on D' we resorted to the following method. We cut each run in smaller parts and computed D' for each part; the difference between these values for D' give an estimate of the error bar.

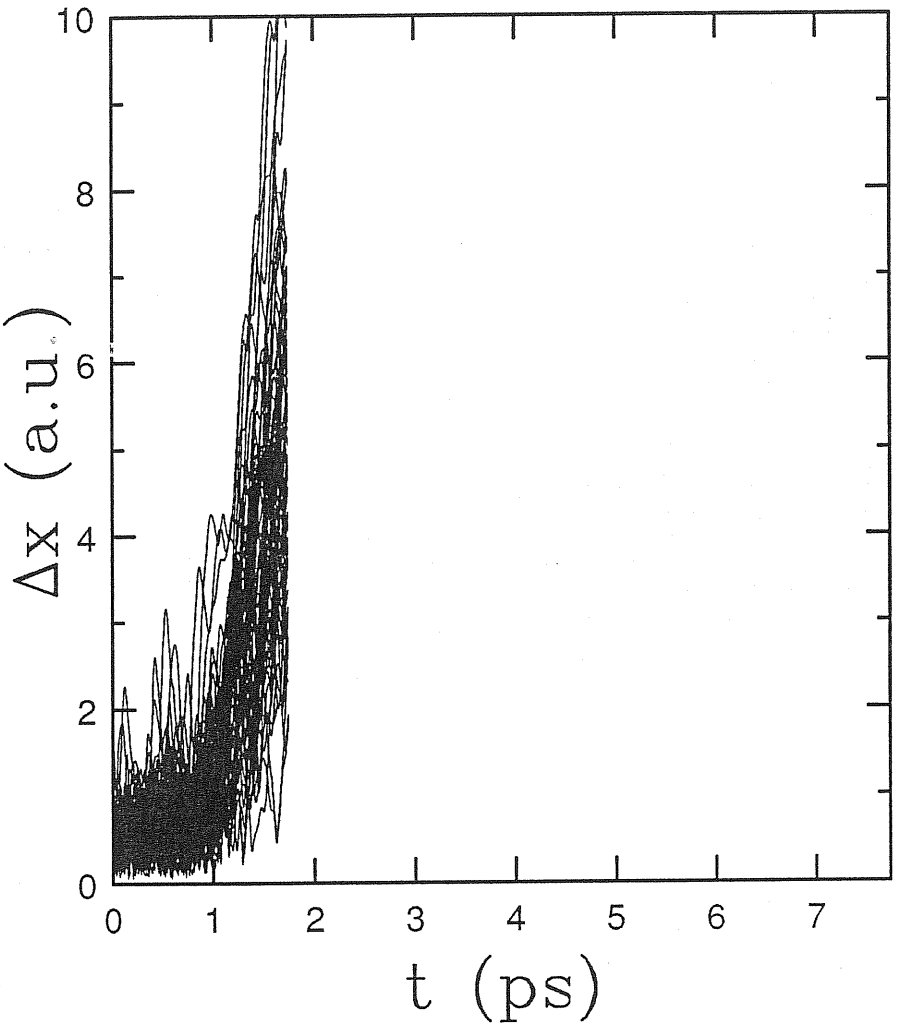

Figure 5.18: $r - t$ plot for the 1750 $^{\circ}K$ run.

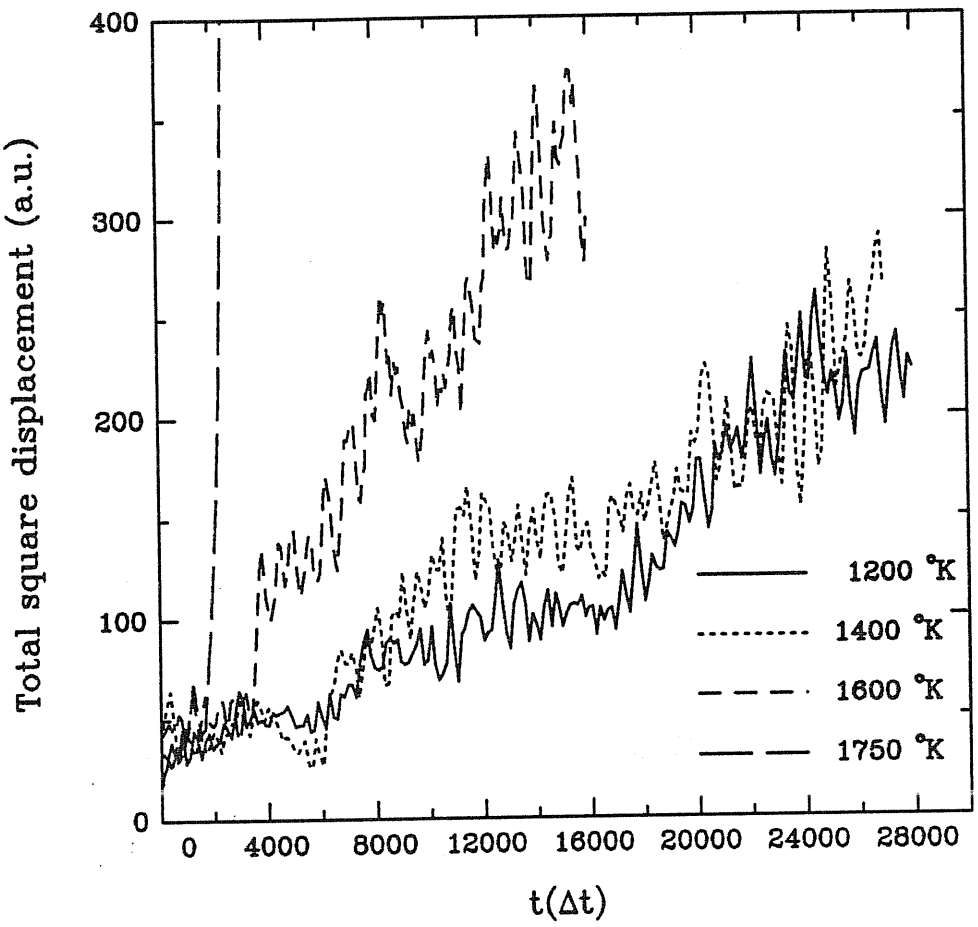

Figure 5.19: Total squared atomic displacement as a function of time.

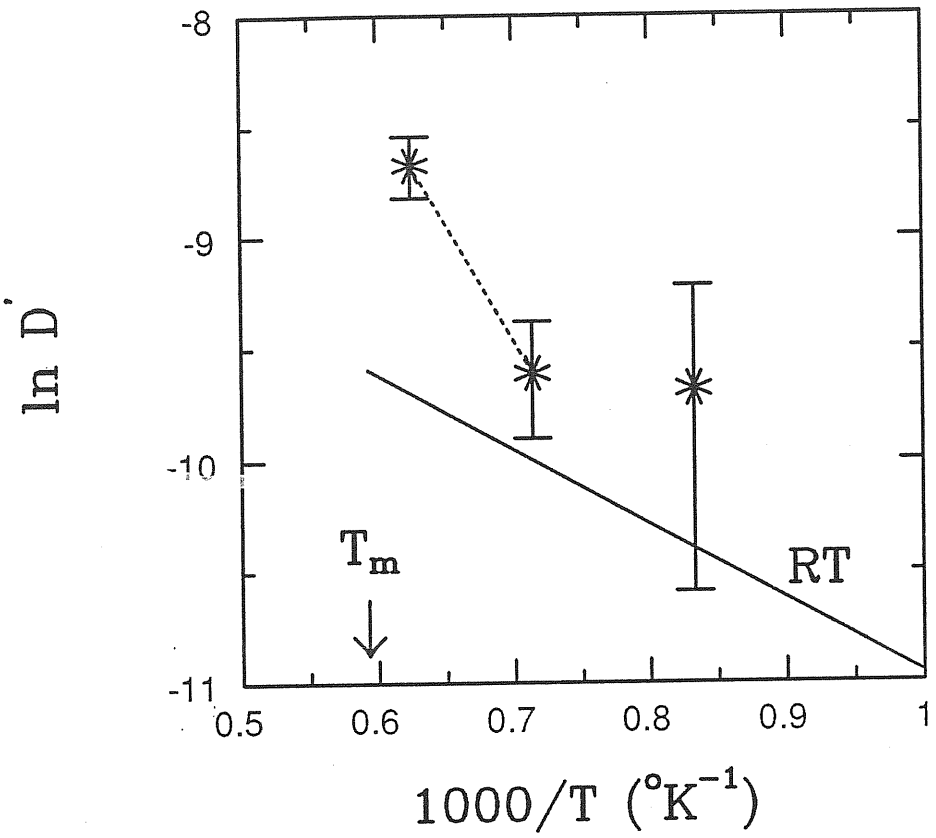

Figure 5.20: Computed self-diffusion coefficient D' (asterisks) as a function of temperature. The continuous line indicates the rate-theory values.

The values obtained for D' are shown in figure 5.20. We report here the numerical values (temperatures are in degrees Kelvin, diffusion coefficients in $10^{-5} \cdot cm^2/s$):

$$\begin{aligned} D'(1200) &= 7.2 \pm 4.3 \\ D'(1400) &= 7.8 \pm 2.0 \\ D'(1600) &= 19.8 \pm 2.8 \end{aligned}$$

The values yielded by the simple rate theory estimate discussed in section 5.2 are represented by the full line.

Once we have the values of D' at the three temperatures we may extract the corresponding values of the pre-exponential factor D'_0 and of the activation energy Q'. The calculated value of Q' between the two highest temperature points is $\sim 0.9 \pm 0.4\ eV$, a value definitely larger than the rate

theory value of 0.3 eV, which is expected to be correct at low temperatures. If we suppose that H_f does not depend on temperature, this difference is also the difference between the high and low temperature values of Q. Our value compares well with the experimentally observed difference (see section 2.2.1) of $0.5 - 1$ eV.

The ratio between the high temperature values of D'_0 and the rate theory value is 200, but it is subject to a large uncertainty. It might range from 10 to 5000. It can be noted however that this uncertainty is roughly the same observed in the experiments.

We computed the self-diffusion coefficient as $D = \mathrm{e}^{-F_f/k_B T} D'$, assuming S_f to be equal to the value obtained at 1000 $^\circ K$. The results are shown in figures 5.21 and 5.22 where they are compared with the LHA results and with the experimental data (see section 2.1.2). The error bars come from the uncertainties on E_f and D', and, for anharmonic results only, from the statistical uncertainty on S_f. The error bar for the LHA value is of the same order of the experimental error; for the anharmonic value is rather larger. Within the error bars, both results are in agreement with the experiments at all temperatures.

We conclude that the vacancy mechanism is a likely candidate for explaining the self-diffusion in silicon; it could explain also the curvature in the Arrhenius plot of D.

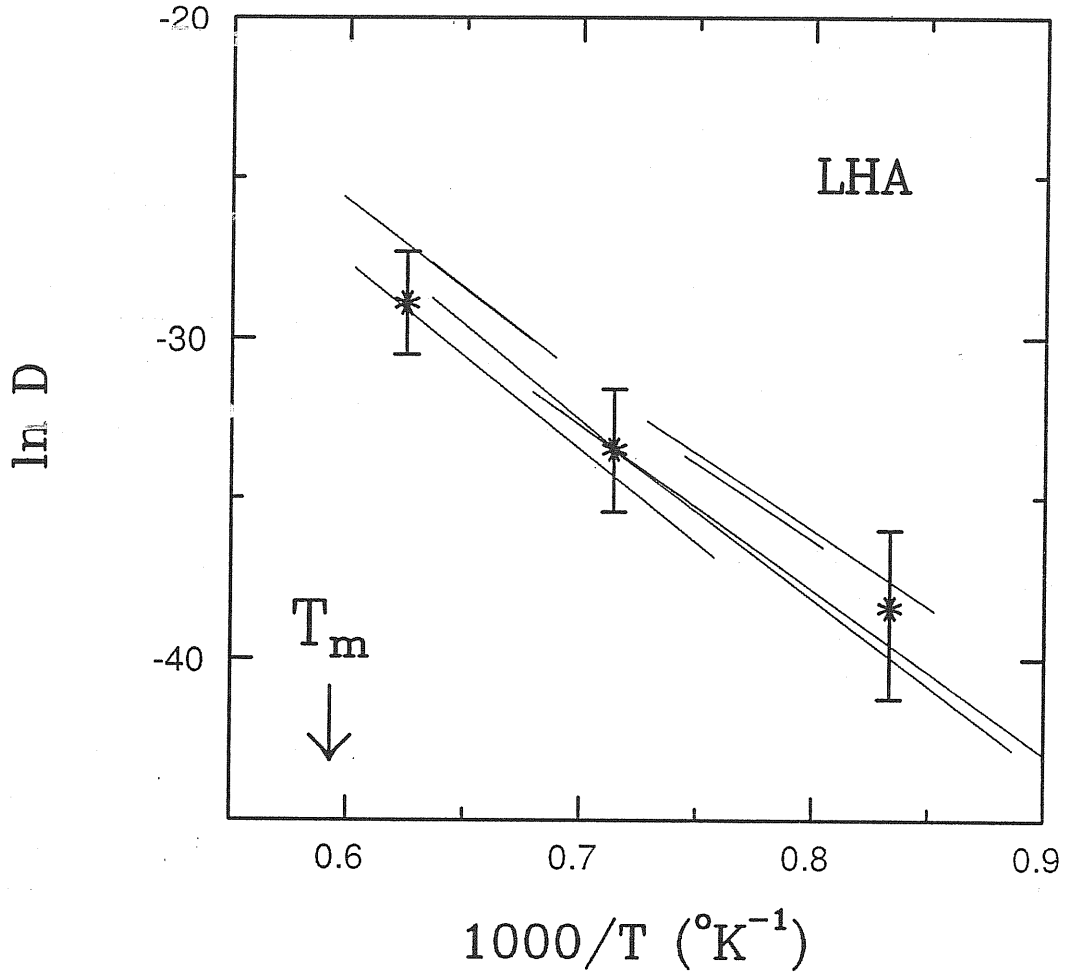

Figure 5.21: Self diffusion coefficient as a function of the temperature. Asterisks: LHA results. Lines: experimental data.

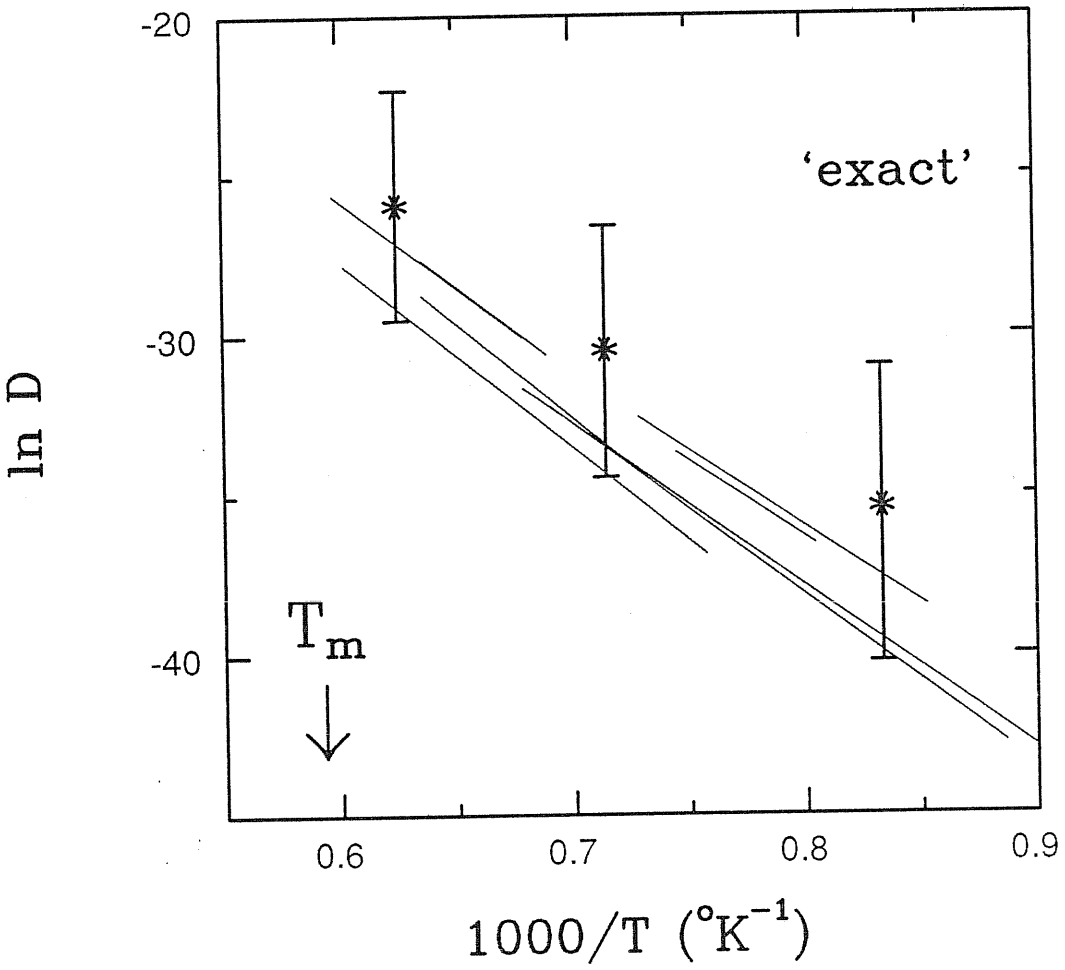

Figure 5.22: Self diffusion coefficient as a function of the temperature. Asterisks: "exact" results. Lines: experimental data.

Chapter 6

Conclusions

In this thesis we have presented the first finite temperature, ab-initio study of the lattice vacancy in silicon, with particular reference to its relevance to silicon self-diffusion. We showed that both the problem of computing the vacancy concentration and of studying the diffusion behaviour can be successfully tackled by the FPMD method, although the computational effort required is quite substantial.

We showed that the formation entropy turns out to be large ($\sim 5\ k_B$) if we compute it within the LHA, and even larger ($\sim 8\ k_B$) if we perform a fully anharmonic calculation. In the latter case, however, we need to improve the statistics in order to have more accurate results. Both results are in agreement with the available experimental data and indicate that in silicon the vacancy induced softening of the lattice is substantially larger than in metals.

The results for the vacancy migration show that that in the temperature range that we studied, which coincides with the experimentally accessed range for the silicon self-diffusion, dynamical effects are important. We found that the vacancy tends to move in a correlated way at $T = 1200,\ 1400\ ^\circ K$. At $1600\ ^\circ K$, near to the melting point, we showed that the usual picture of the vacancy needs to be modified. New migration processes appear, such as second-neighbour jumps; the vacancy itself may be found to be "spread" over a number of different lattice sites. These high temperature effects lead to a substantial increase in the computed self-diffusion coefficient.

Our computed self-diffusion coefficient is in agreement with the experi-

mental data. Moreover, it shows a steeper slope at high temperature similar to that observed in the Arrhenius plot of D. We conclude that the vacancy mechanism alone could explain the self-diffusion behaviour of silicon.

On the basis of our calculations we cannot exclude, however, that other mechanisms may be effective simultaneously. One of these mechanisms is the interstitial one. Recent calculations on the interstitial induced self-diffusion [10] show indeed that this mechanism give a self-diffusion coefficient comparable to the coefficient we obtained for the vacancy. To discriminate between these two possibility more work would be necessary, in particular an accurate evaluation of the formation entropy, including anharmonic terms, for both mechanisms should be done.

Appendix A

Constants of Motion

In this appendix, we are interested in explicitly finding the constants of motion of the FPMD system. There are two main reasons for doing that. First, the knowledge of these quantities allows to characterize the dynamics of the system and the statistical ensemble to which it belongs. Second, the numerical check of the constancy of the theoretical first integrals can be used as a test of the numerical accuracy of the integration scheme.

In addition to the total energy 3.18, the system studied in numerical calculations may have additional constants of motions corresponding to additional symmetries of the FPMD lagrangian.

In the case of a complex representation of the KS orbitals $|\psi_i\rangle$, an almost trivial symmetry is related to the invariance of the lagrangian with respect to multiplication of the orbitals by a (constant) phase factor.

As a consequence of this invariance, we have the following constants of motion:

$$J_{ii} = \langle\dot{\psi}_i|\psi_i\rangle - \langle\psi_i|\dot{\psi}_i\rangle \tag{A.1}$$

Another obvious but more interesting conservation law is the generalization of the linear momentum of the center of mass conservation law in isolated classical systems.

It is evident that, in the absence of external forces, the lagrangian 3.12 is invariant under a global displacement of the ionic coordinates by an arbitrary vector a :

$$\mathbf{R}'_I = \mathbf{R}_I + \mathrm{a} \tag{A.2}$$

if the electronic terms are invariant under a change of the KS orbitals

$$\psi'(\mathbf{r}) = \psi(\mathbf{r} - \mathbf{a}) \tag{A.3}$$

$$\dot{\psi}'(\mathbf{r}) = \dot{\psi}(\mathbf{r} - \mathbf{a}) \tag{A.4}$$

The explicit form for the associate constant of motion depends on the basis on which the wavefunctions are expressed. Within a plane-wave expansion scheme [79,80] ($c_{\mathbf{G}}$ being the Fourier coefficients of the wavefunctions) it is straightforward to verify that the following expression for the component ℓ of the (vector) generator induces the infinitesimal global translation of the system along the ℓ spatial direction:

$$\mathcal{P}^{\ell} = \sum_{I} M_I \dot{R}_I^{\ell} + \sum_{\mathbf{G},i} i\mu_i G^{\ell} \dot{c}_{\mathbf{G}}^{i*} c_{\mathbf{G}}^{i} \tag{A.5}$$

Whenever the second term in this expression can be neglected we recover the usual conservation of the center of mass momentum components. In the usual application of the FPMD method this is the case when averaging the expressions over few electronic periods.

In general, due to the periodic boundary conditions, the rotations are not symmetry operations for the FPMD lagrangian.

Finally, there is another set of constants of motion whenever a subset of the KS orbitals have the same occupation numbers. Indeed, it is easy to show that every unitary transformation of a set of occupied orbitals with the same occupation numbers f_i is a (continuous) symmetry group for the KS functional. Moreover, also the fake kinetic energy term is invariant, provided the unitary transformation is time independent.

This symmetry is reflected in the presence of additional constants of motion only referring to the occupied orbitals. The explicit expressions of such constants are the off-diagonal elements of the antihermitian (antisymmetric in the case of real ψ_i) matrix

$$J_{ij} = \langle \dot{\psi}_i | \psi_j \rangle - \langle \psi_i | \dot{\psi}_j \rangle \tag{A.6}$$

Their constance over the trajectories when $f_i = f_j$ can be easily checked by deriving J_{ij} with respect to the time and using the equations of motion as well as the fact that Λ is hermitian.

We explicitly note that, if the initial conditions for the electronic degrees of freedom are such that not all the J_{ij} are zero, there will be a "rotation" of the KS orbitals superimposed to the coupled ionic and electronic dynamics, and *completely decoupled from the ionic dynamics.*

Appendix B

Formulae used in the FPMD Program

The formulae shown here are expressed either in the real or in the K-space according to their actual use in the program. The K-point is supposed to be Γ. The meaning of the symbols is the following:
M = number of electron states = $2 \times N_s$ for silicon. N_s = number of atom of the species s. σ = number of atomic species. $\mathbf{R}_I^s$ = position vector of the I-th ion of the species s. Ω = cell volume. The primed sums do not include the term with $\mathbf{G} = 0$.

Electron charge density

$$\rho_e(\mathbf{r}) = \sum_{i=1}^{M} f_i |\psi_i(\mathbf{r})|^2$$

Ion charge density

$$\rho_I(\mathbf{r}) = \sum_{s=1}^{\sigma} \sum_{I=1}^{N_s} \frac{Z^s}{\pi^{3/2} R_c^{s\,3}} \exp\left[\frac{(\mathbf{r} - \mathbf{R}_I^s)^2}{R_c^{s\,2}}\right]$$

Total charge density

$$\rho_T(\mathbf{r}) = \rho_e(\mathbf{r}) + \rho_I(\mathbf{r})$$

We decompose the energy as $E_{KS} = E_{ke} + E_H + E_{xc} + E_{ps}^{loc} + E_{ps}^{nl} + E_{self} - E_{sr}$. These terms are respectively the electronic kinetic energy, the electrostatic Hartree term, the exchange and correlation energy, the electron-ion pseudopotential interaction and the purely ionic interaction potential energy. The pseudopotential interaction is written as a sum of a "local" part (the

non local component corresponding to $l = l_c$) plus a non-local part. The purely ionic interaction potential energy is written as the interaction between two broadened ionic charge distribution of radius $\mathbf{R}_c^s$ - minus the self-interaction of these charges.

We call $v_c = -1/r$ the Coulomb potential, $\delta v_l = v_l - v_{l_c}$, $\alpha_{l,m} = 1/\langle\varphi_{l,m}|\delta v_l|\varphi_{l,m}\rangle$, and define $F_{I,i}^{l,m}$ as

$$F_{I,i}^{l,m} = \sum_{\mathbf{G}} c_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{R}_I} \langle \delta v_l \varphi_{l,m} \hat{P}_{l,m} | \mathbf{G} \rangle$$

Here $\hat{P}_{l,m}$ is the projector over the spherical harmonics l, m, and $|\varphi_{l,m}\rangle$ are the pseudowavefunctions for which the pseudopotential v_l was calculated.

$$\begin{aligned}
E_{ke} &= -\frac{1}{2}\sum_{i=1}^{M} f_i \sum_{\mathbf{G}} G^2 \psi_i^*(\mathbf{G})\psi_i(\mathbf{G}) \\
E_H &= \frac{4\pi\Omega}{2}{\sum_{\mathbf{G}}}' \rho_T^*(\mathbf{G})\, v_c(\mathbf{G})\, \rho_T(\mathbf{G}) \\
E_{xc} &= \int \mathrm{d}\mathbf{r}\, \epsilon_{xc}^h(\rho_e(\mathbf{r}))\, \rho_e(\mathbf{r}) \\
E_{loc}^{ps} &= \Omega \sum_{s=1}^{\sigma}\sum_{i=1}^{M}\sum_{\mathbf{G}} \rho_e^*(\mathbf{G}) e^{-i\mathbf{G}\cdot\mathbf{R}_I^s} v_{l_c}(\mathbf{G}) \\
E_{nl}^{ps} &= \frac{(4\pi)^2}{\Omega}\sum_{s=1}^{\sigma}\sum_{I=1}^{N_s}\sum_{i=1}^{M}\sum_{l,m} f_i \alpha_{l,m} |F_{I,i}^{l,m}|^2 \\
E_{self} &= \frac{1}{\sqrt{2\pi}}\sum_{s=1}^{\sigma}\frac{N_s Z^{s\,2}}{R_c^s} \\
E_{sr} &= \frac{1}{2}\sum_{s=1}^{\sigma}\sum_{s'=1}^{\sigma}\sum_{I=1}^{N_s}\sum_{J=1}^{N_s'} \frac{Z_I^s Z_J^{s'}}{|\mathbf{R}_I^s - \mathbf{R}_J^{s'}|} \mathrm{erfc}\left(\frac{\mathbf{R}_I^s - \mathbf{R}_J^{s'}}{\sqrt{R_c^{s\,2} + R_c^{s'\,2}}}\right) \\
&\quad \text{excluding the terms with } s = s' \text{ and } I = J
\end{aligned}$$

Appendix C

Formulae for $\frac{\partial E_{KS}}{\partial \lambda}$

These formulae are written for the specific case of 1 atom ($I = 0$) being decoupled with his M_λ electron states from an elemental crystal. Therefore $\sigma = 2,\ N_2 = 1,\ M_\lambda = Z/2$.

Electron charge density

$$\frac{\partial \rho_e(\mathbf{r})}{\partial \lambda} = \sum_{\substack{i = M - \\ M_\lambda + 1}}^{M} \frac{f_i}{\lambda} |\psi_i(\mathbf{r})|^2$$

Ion charge density

$$\frac{\partial \rho_I(\mathbf{r})}{\partial \lambda} = \frac{Z}{\lambda \pi^{3/2} R_c^3} \exp\left[\frac{(\mathbf{r} - \mathbf{R}_0)^2}{R_c^2}\right]$$

Total charge density

$$\frac{\partial \rho_T(\mathbf{r})}{\partial \lambda} = \frac{\partial \rho_e(\mathbf{r})}{\partial \lambda} + \frac{\partial \rho_I(\mathbf{r})}{\partial \lambda}$$

$$\frac{\partial E_{ke}}{\partial \lambda} = -\frac{1}{2} \sum_{\substack{i = M - \\ M_\lambda + 1}}^{M} \frac{f_i}{\lambda} \sum_{\mathbf{G}} G^2 \psi_i^*(\mathbf{G}) \psi_i(\mathbf{G})$$

$$\frac{\partial E_H}{\partial \lambda} = 4\pi\Omega \sum_{\mathbf{G}} \frac{\partial \rho_T^*}{\partial \lambda}(\mathbf{G})\, v_c(\mathbf{G})\, \rho_T(\mathbf{G})$$

$$\frac{\partial E_{xc}}{\partial \lambda} = \int \mathrm{d}\mathbf{r}\, \frac{\partial \rho_e(\mathbf{r})}{\partial \lambda} \left(\epsilon_{xc}^h(\rho_e(\mathbf{r})) + \frac{\partial \epsilon_{xc}^h(\rho_e)}{\partial \rho_e} \rho_e(\mathbf{r}) \right)$$

$$
\begin{aligned}
\frac{\partial E_{loc}^{ps}}{\partial \lambda} &= \Omega \sum_{s=1}^{\sigma} \sum_{I=1}^{N_s} \sum_{\mathbf{G}} \left[\frac{\partial v_{l_c}(\mathbf{G})}{\partial \lambda} \rho_e^* + \frac{\partial \rho_e^*(\mathbf{G})}{\partial \lambda} v_{l_c}(\mathbf{G}) \right] \\
\frac{\partial E_{nl}}{\partial \lambda} &= \frac{(4\pi)^2}{\Omega} \sum_{s=1}^{\sigma} \sum_{I=1}^{N_s} \sum_{i=1}^{M} \sum_{l,m} \alpha_{l,m} \left(\frac{\partial f_i}{\partial \lambda} |F_{I,i}^{l,m}|^2 + f_i \frac{\partial |F_{I,i}^{l,m}|^2}{\partial \lambda} \right) \\
\frac{\partial E_{self}}{\partial \lambda} &= \frac{2}{\sqrt{2\pi}} \frac{\lambda Z^2}{R_c} \\
\frac{\partial E_{sr}}{\partial \lambda} &= \sum_{I=1}^{N-1} \frac{Z^2}{\lambda} |\mathbf{R}_I - \mathbf{R}_0| \mathrm{erfc} \left(\frac{\mathbf{R}_I - \mathbf{R}_0}{\sqrt{2} R_c} \right)
\end{aligned}
$$

Appendix D

Orthogonalization in the general Case

Define

$$\begin{aligned} X_{ij} &= (\Delta t)^2 \Lambda^{\dagger}_{ij} \\ A_{ij} &= \langle \overline{\psi}_i(t+\Delta t) | \overline{\psi}_j(t+\Delta t) \rangle \\ B_{ij} &= \langle \psi_i(t) | \overline{\psi}_j(t) \rangle \end{aligned}$$

where $\{\overline{\psi}_i(t+\Delta t)\}$ are the wavefunctions at time $(t+\Delta t)$ as evoluted from $\{\psi_i(t)\}$ in the absence of constraints. Then it can be shown, in analogy with ref. [69] (sum over repeated indexes is implied):

$$\frac{1}{\mu_i \mu_j} X_{ik} X^{\dagger}_{kj} + \frac{1}{\mu_i} X_{ik} B_{kj} + \frac{1}{\mu_j} B^{\dagger}_{il} X^{\dagger}_{lj} = \delta_{ij} - A_{ij}$$

which can be solved iteratively, starting from the zero-order solution $X^{(0)}_{ij} = \frac{\mu_i \mu_j}{\mu_i + \mu_j} (\delta_{ij} - A_{ij})$ using the recursion formula

$$X^{(n)}_{ij} = \frac{\mu_i \mu_j}{\mu_i + \mu_j} \left[(\delta_{ij} - A_{ij}) + \frac{1}{\mu_i} X^{(n-1)}_{ik} (\delta_{kj} - B_{kj}) + \frac{1}{\mu_j} \left(\delta_{il} - B^{\dagger}_{il} \right) X^{(n-1)}_{lj} - \frac{1}{\mu_i \mu_j} X^{(n-1)}_{ik} X^{(n-1)}_{kj} \right]$$

Appendix E

Checking the Convergence

E.1 Generalities

Here we expose the results of an extensive series of tests that we have performed at $T = 0$.

In the majority of our computations, we get the initial trial wavefunctions by diagonalizing a small Hamiltonian matrix constructed by using the local empirical pseudopotential of Louie *et al.* [76]. This gives a reasonable starting approximation for the wavefunctions, usually with the right energy ordering of the single particle states. Convergence to the correct ground-state was checked by acting on the electronic states with small random perturbations.

Since the lowest energy atomic structure of the vacancy is not the "ideal" one, e. g. the one with the atoms fixed at the perfect crystal positions, the atoms must be allowed to relax. This was done with the steepest descent procedure described at page 26.

In the computations shown in this appendix we are interested in a number of different quantities, namely:

- The formation energy $E_f^N = E_V^{N-1} - \frac{N-1}{N} E_B^N$, ($E_V$ and E_B being the energy of the system with and without the vacancy, respectively) both in the ideal and in the relaxed case.

- The migration energy E_m, which corresponds to the energy difference between the split vacancy and the ground state vacancy configuration. This energy difference is believed to correspond to the energy barrier

for the simple vacancy jump. This energy is always computed between relaxed configurations.

- The energy of the occupied gap level E_{level} induced by the vacancy, measured from the top of the valence band. In the unrelaxed silicon vacancy a threefold degenerate gap level exists. Therefore a Jahn-Teller relaxation is expected [49]. This effect leads to the splitting of the degenerate level into non degenerate levels. We explicitly verified that the gap level having the lowest energy is not degenerate, as previously obtained [50].

- The atomic relaxation of the first neighbors of the vacancy. This can be written in terms of two normal coordinates per atom, since only a symmetric and a tetragonal distortion have been experimentally detected [47,48]. We call Q_b the component of the position vector of these atoms along the line joining the position they have in the ideal situation with the vacancy ("breathing" mode: symmetric), and Q_p ("pairing" mode: tetragonal) the component of the vector position in one of the directions orthogonal to the previous one.

  To be more explicit, if we choose Cartesian axes such that the coordinates of the first neighbours of the vacancy are (in units of the lattice parameter)

$$\begin{aligned} \mathbf{r}_1 &= \frac{1}{4}(+1,+1,+1)/\sqrt{3} \\ \mathbf{r}_2 &= \frac{1}{4}(+1,-1,-1)/\sqrt{3} \\ \mathbf{r}_3 &= \frac{1}{4}(-1,+1,-1)/\sqrt{3} \\ \mathbf{r}_4 &= \frac{1}{4}(-1,-1,+1)/\sqrt{3} \end{aligned}$$

  then $Q_b = \sum_{i=1}^{4} \Delta\mathbf{r}_i \cdot \mathbf{r}$ ($\Delta\mathbf{r}_i$ being the displacements from the ideal configuration) and a possible choice for Q_p is

$$\begin{aligned} Q_p = {} & [\Delta\mathbf{r}_1 \cdot (-1,-1,+2) + \Delta\mathbf{r}_2 \cdot (-1,+1,-2) + \\ & \Delta\mathbf{r}_3 \cdot (+1,-1,-2) + \Delta\mathbf{r}_4 \cdot (+1,+1,2)]/4\sqrt{6}. \end{aligned}$$

k-point / N	Γ	Baldereschi [77]	2 Chadi-Cohen points [78]
8		2.94	3.80*
16		3.89	3.66
32	3.62	4.72	4.52
54	3.47	4.35	4.28
64	4.16	4.44	

* 4 points

Table E.1: Formation energy in eV of the unrelaxed vacancy as a function of the special points used and of the cell size N. $E_{cut} = 6Ryd, l_c = 1$ and $\gamma = 4$ are used.

E.2 Results

E.2.1 Convergence in the cell size and in the number of k-points

Since the crystal symmetry of silicon is T_d, the choice of the cells is limited to the ones having the following number of atoms (L being an integer):

$$N = \begin{cases} 2, 16, 54, 128, \ldots, 2 \times L^3 & \text{(FCC cells)} \\ 8, 64, \ldots, 8 \times L^3 & \text{(SC cells)} \\ 32, \ldots, 32 \times L^3 & \text{(BCC cells)} \end{cases} \tag{E.1}$$

From now on, the cells will be labeled by the number N of atoms they have when the crystal has no vacancies.

The formation energy of the unrelaxed vacancy is shown in table E.1. The cutoff energy, here and later (unless otherwise stated), was set to 6 Ryd, deferring the study of the convergence in E_{cut}.

From table E.1 we see that the convergence in the number of k-points is reached when the cell size reaches 64 atoms. In the case of 54 atoms, the Baldereschi point gives a value within 0.1 eV to the converged one, while using 64 atoms we obtain a difference in energy between the use of that point and of Γ of about 0.3 eV, that is with an error less than 10%.

In table E.2 the convergence in size up to 128 atoms is examined, using

N	54	64	128
E_f	3.78	4.41	4.57
E_{level}	1.14	1.11	0.93

Table E.2: Formation energy and gap-level position (measured from the top of the valence band), in eV, of the unrelaxed vacancy as a function of the cell size N. The Γ point, $E_{cut} = 6\,Ryd$, $l_c = 1$, and $\gamma = 1$ are used.

the Γ point and $\gamma = 1$. Because of the use of a different γ, the energies are shifted. However, it is sufficient to see that the variation in E_f while passing from 64 to 128 atoms is only of $0.16\,eV$. This is confirmed by the data shown in the table E.4. The migration energy in this latter case rises slightly, getting closer to the experimental value. It is noteworthy that the normal coordinates do not change.

In figure E.1 the absolute values of the displacements of the atoms from the crystalline positions in a 63 atom+vacancy cell are shown. They are compared with the curve α/r^2, which is the correct result at large distances according to the theory of elasticity. α was fitted here to pass through the dot indicating the relaxation of the first shell. Although the distances shown in the plot was far from being large, the curve fits fairly well the dots.

E.2.2 Convergence with respect to the plane-waves cutoff

Since a 64-atom cell with only Γ seems to be a reasonable choice, we studied the effect of the cutoff E_{cut} using that cell (Tables E.3 and E.4). We see that the calculation starts to be well converged (within $\sim 0.1\ eV$) at a cutoff of 12 Ryd. However the energy gain due to the relaxation process seems to be less sensitive to E_{cut}, being about $0.4 - 0.5\ eV$ already with an 8 Ryd cutoff.

Due to the presence of (Jahn-Teller) relaxation both the normal coordinates differ from zero. The values obtained for Q_b, Q_p are not too sensitive to the cutoff. Note that breathing relaxation turns out to be inward, in contrast to earlier assumptions [50,76], however in agreement with other recent self-consistent calculations [61].

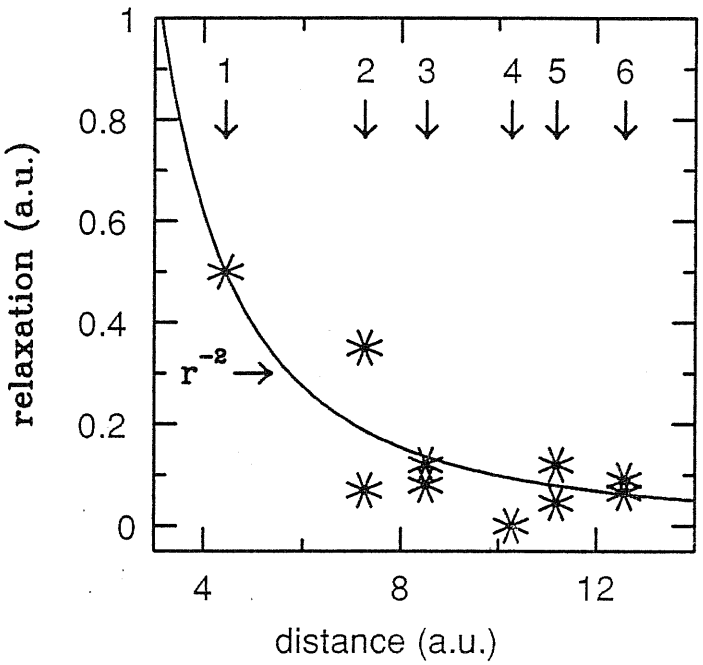

Figure E.1: Absolute values of the atomic displacements of the ions from the perfect crystalline positions in a 64 atom cell, compared with the prediction of the theory of elasticity (see text). The numbers from 1 to 6 on the top indicates that the atoms under the arrows are first, ..., sixth neighbours of the vacancy.

E_{cut}	6	8	10	12	14
E_f	4.16	3.73	3.42	3.22	3.14
E_{level}	1.11	1.05	1.01	0.99	1.00
E_f	3.88	3.28	2.96	2.69	
E_m	0.56	0.40	0.33	0.30	
Q_b	-0.31	-0.48	-0.53	-0.58	
Q_p	$+0.39$	$+0.45$	$+0.45$	$+0.61$	
E_{level}	0.76	0.66	0.66	0.66	

Table E.3: Formation and migration energy and gap-level position (measured from the top of the valence band), in eV, of the unrelaxed (above) and relaxed (below) vacancy as a function of the cutoff energy E_{cut} (in Rydberg), in a 64 atoms cell; the normal coordinates of the first shell are also shown. The Γ point, $l_c = 1$, and $\gamma = 4$ are used.

E_{cut}	10	12	10
E_f	2.98	2.69	3.14
E_m	0.33*	0.30	0.37
Q_b	-0.49	-0.58	-0.48
Q_p	$+0.38$	$+0.61$	$+0.38$
E_{level}	0.71	0.67	0.45

* $\gamma = 4$

Table E.4: Formation and migration energy and gap-level position (measured from the top of the valence band), in eV, of the relaxed vacancy as a function of the cutoff energy E_{cut} (in Rydberg), in a 64 (left columns) and 128 atoms cell (right column); the normal coordinates of the first shell are also shown. The Γ point, $l_c = 1$ and $\gamma = 1.5$ are used.

γ	1	1.5	2	4
E_f	3.37	3.42	3.42	3.42
E_{level}	1.02	1.02	1.02	1.01

Table E.5: Formation energy and gap-level position (measured from the top of the valence band), in eV, of the neutral, unrelaxed vacancy as a function of γ, in a 64 atoms cell. $E_{cut} = 10\,Ryd$, $l_c = 1$, and the Γ point are used.

Finally, not only the value of E_m shows a good convergence (see also Tab. E.6, next section), but the agreement with experimental data (see sect. 2.2.1) is fairly good.

E.2.3 Convergence in γ

The results of the test over γ in the ideal vacancy configuration are reported in Table E.5. As immediately seen, at least for this cutoff ($10\,Ryd$) the use of $\gamma = 1$ yields a very small error ($\sim 0.05\,eV$), while greater values offer a very small error, below the limits of accuracy of the method. The use of a greater E_{cut} would further improve the situation. Note that this choice of γ allows to save about 50% in computer resources.

The effect of γ on the relaxation has been tested allowing the neutral vacancy to relax with $\gamma = 1.5$. Even in this case the differences are very small (see table E.3 and compare with table E.4).

E.2.4 Effect of the d-component of the pseudopotential

Inclusion of an accurate representation of the $l = 2$ component of the pseudopotential by using $l_c = 2$ has a non negligible effect on the formation energy. Table E.6 shows that the introduction of the $l = 2$ term makes rise E_f and substantially reduces the gain due to the ionic relaxation. Note that the ionic positions (monitored by the Q's) are almost the same.

l_c	1	2	1	2
E_f	3.42	3.54	3.22	
E_{level}	1.01	0.92	0.99	
E_f	2.96	3.39	2.69	3.31
E_m	0.33	0.33	0.30	
Q_b	−0.53	−0.41	−0.58	
Q_p	+0.45	+0.44	+0.61	
E_{level}	0.66	0.64	0.66	

Table E.6: Relevant energies, in eV, and deformation parameters, in a.u., of the neutral, unrelaxed (above) and relaxed (below) neutral vacancy as a function of l_c, in a 64 atoms cell. The normal coordinates of the first shell are also shown. $\gamma = 4$ and the Γ point are used. Left: $E_{cut} = 10\,Ryd.$ Right: $E_{cut} = 12\,Ryd.$ The result in the rightmost column was obtained assuming the same ionic coordinates obtained with $l_c = 1$.

Acknowledgements

It is a pleasure to acknowledge Dr. Guido L. Chiarotti for his continuous help during the making of this work, in particular for the invaluable help during the computations. Dr. Giorgio Pastore and Dr. Franco Buda, in addition to help me in several ways, made a great deal of the work presented in chapter 3.

The computational work was performed under the SISSA-CINECA (Centro di Calcolo Interuniversitario dell' Italia Nord Orientale) collaborative project. Dr. Furio Ercolessi patiently took upon himself the nasty and time-consuming, but absolutely indispensable, job of managing the supercomputing resources of the School, in addition to his own research activity.

It is important to have technical help, but in my opinion it is equally important to work in a friendly and hospitable environment. I cannot mention here all the people who made my permanence in SISSA enjoyable because it is impossible to thank so many people in such a small space. To them my most grateful thought is addressed now and will be in the future.

Bibliography

[1] M. Lannoo and J. C. Bourgoin, *Point Defects in Semiconductors - vol. I*, Springer-Verlag, Berlin (1981).

[2] R. A. Swalin, Journal of Physics and Chemistry of Solids **18**, 290 (1961).

[3] J. Van Vechten, Physical Review B **33**, 2674 (1986).

[4] H. C. Casey and G. L. Pearson, in *Point Defects in Solids, Vol. 2*, edited by J. H. Crawford Jr. and L. M. Sifkin (Plenum Publishing Corp., 1975).

[5] A. Seeger and M. L. Swanson, in *Lattice Defects in Semiconductors*, edited by R. R. Hasiguti (Tokio and Pennsylvania Univ. Press, 1968).

[6] K. C. Pandey, Physical Review Letters **57**, 2287 (1986).

[7] A. M. Stoneham, V. T. B. Torres, P. M. Masri, and H. R. Schober, Philosophical Magazine A **58**, 93 (1988).

[8] R. Car and M. Parrinello, Physical Review Letters **55**, 2471 (1985).

[9] R. LeSar, R. R. Najafabadi, and D. Srolovitz, Physical Review Letters **63**, 624 (1989).

[10] P. Blöchl, D. B. Laks, S. T. Pantelides, E. Smargiassi, R. Car, W. Andreoni, and M. Parrinello, in *Proceedings of the 20th International Conference on Physics of Semiconductors*, 1990, to be published.

[11] S. Dannefaer, P. Mascher, and D. Kerr, Physical Review Letters **56**, 2195 (1986).

[12] A. Einstein, Annalen der Physik **17**, 549 (1905).

[13] C. P. Flynn, *Point Defects and Diffusion*, Clarendon Press, Oxford (1972).

[14] A. D. LeClaire, in *Physical Chemistry: an Advanced Treatise, vol. X*, edited by W. Jost (Academic Press, 1970).

[15] C. H. Bennett, in *Diffusion in Solids*, edited by A. S.Nowick and J. J. Burton (Academic Press, 1975).

[16] D. Shaw, in *Atomic Diffusion in Semiconductors*, edited by D. Shaw (Plenum Publishing Corp., 1973).

[17] W. Jost, *Diffusion in Solids, Liquids, Gases (3_{rd} printing)*, Academic Press, New York (1960).

[18] T. H. Yeh, in *Atomic Diffusion in Semiconductors*, edited by D. Shaw (Plenum Publishing Corp., 1973).

[19] R. F. Peart, Physica Status Solidi **15**, K119 (1966).

[20] B. J. Masters and J. M. Fairfield, Applied Physics Letters **8**, 280 (1966).

[21] J. M. Fairfield and B. J. Masters, Journal of Applied Physics **38**, 3148 (1967).

[22] R. N. Ghoshtagore, Physical Review Letters **16**, 890 (1966).

[23] J. Hirvonen and A. Anttila, Applied Physics Letters **35**, 703 (1979).

[24] L. Kalinowski and R. Seguin, Applied Physics Letters **35**, 211 (1979).

[25] F. J. Demond, S. Kalbitzer, H. Mannsperger, and H. Damjantschitsch, Physics Letters **93A**, 503 (1983).

[26] H. J. Mayer, H. Mehrer, and K. Maier, in *Radiation Effect in Semiconductors*, edited by N. B. Urli and J. W. Corbett (The Institute of Physics, 1977).

[27] J. R. Sanders and P. S. Dobson, Journal of Material Science **9**, 1987 (1974).

[28] L. Kalinowski and R. Seguin, Applied Physics Letters **36**, 171 (1980).

[29] G. L. McVay and A. R. Du Charme, Journal of Applied Physics **44**, 1409 (1973).

[30] G. L. McVay and A. R. Du Charme, Physical Review B **9**, 627 (1974).

[31] G. L. McVay and A. R. Du Charme, in *Institute of Physics Conference Serial vol. 23* (The Institute of Physics, 1975) p. 91.

[32] D. A. Petrov, Yu. M. Shashkov, and I. P. Akimchenko, in *Vop. Met. Fiz. Poluprov. Akad. Nauk SSSR* (Tr. Vtorogo Soveshch., 1957).

[33] G. Hettich, H. Mehrer, and K. Maier, in *Institute of Physics Conference Serial vol. 46* (The Institute of Physics, 1979) p. 500.

[34] P. Varotsos and K. Alexopoulos, in *Thermodynamics of Point Defects and their Relation with the Bulk Properties*, edited by S. Amelinckx, R. Gevers, and J. Nihoul (North-Holland, 1986).

[35] S. M. Hu, in *Atomic Diffusion in Semiconductors*, edited by D. Shaw (Plenum Publishing Corp., 1973).

[36] A. Seeger and K. P. Chik, Physica Status Solidi **29**, 455 (1968).

[37] W. Frank, U. Gösele, H. Mehrer, and A. Seeger, in *Diffusion in Crystalline Solids*, edited by G. E. Murch and A. S. Nowick (Academic Press, 1984).

[38] P. M. Fahey, P. B. Griffin, and J. D. Plummer, Reviews of Modern Physics **61**,(2), 289 (1989).

[39] U. M. Gösele, in *Festkörperprobleme vol. 26*, edited by P. Grosse (F. Vieweg & Sohn, 1986).

[40] M. Lannoo and G. Allan, Physical Review B **25**, 4089 (1982).

[41] M. Lannoo and G. Allan, Physical Review B **33**, 8789 (1986).

[42] G. H. Vineyard, Journal of Physics and Chemistry of Solids **3**, 121 (1957).

[43] G. Jacucci, in *Diffusion in Crystalline Solids*, edited by G. E. Murch and A. S. Nowick (Academic Press, 1984).

[44] F. Buda, G. L. Chiarotti, R. Car, and M. Parrinello, Physical Review Letters **63**, 294 (1989).

[45] G. Jacucci, in *Mass Transport in Solids*, edited by F. Benière and C. R. A. Catlow (Plenum Publishing Corp., 1983).

[46] J. Tersoff, Physical Review B **38**, 9902 (1988).

[47] G. D. Watkins, Journal of the Physical Society of Japan **18** *suppl.2*, 22 (1963).

[48] G. D. Watkins, in *Proceedings of the 7° International Conference on the Physics of Semiconductors — 3*, edited by P. Baruch (Dunod, 1965).

[49] J. C. Bourgoin and M. Lannoo, *Point Defects in Semiconductors - vol. II*, Springer-Verlag, Berlin (1981).

[50] G. A. Baraff, E. O. Kane, and M. Schlüter, Physical Review B **21**, 5662 (1980).

[51] G. D. Watkins and J. R. Troxell, Physical Review Letters **44**, 593 (1980).

[52] J. R. Troxell and G. D. Watkins, Physical Review B **22**, 921 (1980).

[53] W. Fuhs, V. Holtzhaur, S. Mantl, F. W. Richter, and R. Sturm, Physica Status Solidi **89**, 69 (1978).

[54] L. Elstner and W. Kamprath, Physica Status Solidi **22**, 541 (1967).

[55] J. C. Bourgoin, Physics Letters **106A**, 140 (1984).

[56] P. Varotsos, K. Eftaxias, and V. Hadjicontis, Physical Review B **38**, 6328 (1988).

[57] Y. Bar-Yam and J. D. Joannopoulos, in *Proceedings of the 17th International Conference on Physics of Semiconductors*, edited by J. D. Chadi and W. A. Harrison (Springer-Verlag, 1984).

[58] C. S. Nichols, C. G. van de Walle, and S. T. Pantelides, Physical Review B **40**, 5484 (1989).

[59] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, Physical Review Letters **52**, 1814 (1984).

[60] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, Physical Review Letters **54**, 360 (1985).

[61] P. J. Kelly, R. Car, and S. T. Pantelides, in *Proceedings of the 14th International Conference on Defects in Semiconductors, Paris* (who-knows, 1986).

[62] P. N. Keating, Physical Review **145**, 637 (1966).

[63] R. P. Feynman, Physical Review **56**, 340 (1939).

[64] P. C. Hohenberg and W. Kohn, Physical Review **136**, B864 (1964).

[65] W. Kohn and P. Vashishta, in *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum Publishing Corp., 1983).

[66] J. C. Perdew and A. Zunger, Physical Review B **23**, 5048 (1981).

[67] D. M. Ceperley and B. K. Alder, Physical Review Letters **45**, 566 (1980).

[68] W. Kohn and L. J. Sham, Physical Review **140**, A1133 (1965).

[69] R. Car and M. Parrinello, in *Simple Molecular Systems at Very High Density* (Plenum Publishing Corp., 1989) p. 455.

[70] J. P. Ryckaert, G. Ciccotti, and H. J. Berendsen, Journal of Computational Physics **23**, 327 (1977).

[71] D. K. Remler and P. A. Madden, Molecular Physics **70**, 921 (1990).

[72] H. C. Andersen, Journal of Chemical Physics **74**, 2384 (1980).

[73] S. Nosé, Journal of Chemical Physics **81**, 511 (1984).

[74] D. J. Evans and B. L. Holian, Journal of Chemical Physics **83**,(8), 4069 (1985).

[75] B. L. Holian, in *Proceedings of the International School of Physics "Enrico Fermi" — Course XCVII*, edited by G. Ciccotti and W. G. Hoover (North-Holland, 1986).

[76] S. G. Louie, M. Schlüter, J. R. Chelikowski, and M. L. Cohen, Physical Review B **13**, 1654 (1976).

[77] A. Baldereschi, Physical Review B **7**, 5212 (1973).

[78] D. J. Chadi and M. L. Cohen, Physical Review B **8**, 5747 (1973).

[79] M. T. Yin and M. L. Cohen, Physical Review B **26**, 5668 (1982).

[80] J. Ihm, A. Zunger, and M. L. Cohen, Journal of Physics C **12**, 4409 (1979).

[81] D. R. Hamann, M. Schlüter, and C. Chiang, Physical Review Letters **43**, 1494 (1979).

[82] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Physical Review B **26**, 4199 (1982).

[83] L. Kleinman and D. M. Bylander, Physical Review Letters **48**, 1425 (1982).

[84] M. Born, *The Mechanics of the Atom*, G. Bell and Sons, Ltd., London (1927).

[85] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, Berlin (1988).

[86] P. Lochak and C. Meunier, *Multiphase Averaging for Classical Systems*, Springer-Verlag, Berlin (1989).

[87] N. N. Bogoljubov, *Asymptotic methods in the theory of Non-linear oscillations*, Hindustan Publishing Corporation, Dehli (1961).

[88] G. Benettin, in *Nonlinear Evolution and Chaotic Phenomena* (Plenum Publishing Corp., 1987) p. 121.

[89] G. Pastore, E. Smargiassi, and F. Buda, to be published.

[90] G. B. Bachelet, G. Jacucci, R. Car, and M. Parrinello, in *Proceedings of the* 18° *International Conference on the Physics of Semiconductors, Stockholm 1986*, edited by O. Engström (World Scientific, 1987).

[91] G. B. Bachelet and G. De Lorenzi, Physica Scripta **T19A**, 311 (1987).

[92] C. H. Bennett, Journal of Computational Physics **22**,(2), 245 (1976).

[93] D. R. Squire and W. G. Hoover, Journal of Chemical Physics **50**,(2), 701 (1969).

[94] J. Q. Broughton and G. H. Gilmer, Journal of Chemical Physics **79**, 5095 (1983).

[95] P. Giannozzi, S. De Gironcoli, P. Pavone, and S. Baroni, to be published.

[96] M. Cardona, G. Harbeke, O. Madelung, and U. Rössler, in *Landolt-Börnstein, vol. 17 - Semiconductors* (Springer-Verlag, 1982).

[97] A. Da Fano and G. Jacucci, Physical Review Letters **39**,(15), 950 (1977).

[98] E. A. Stern and Ke Zhang, Physical Review Letters **60**,(18), 1872 (1988).

[99] H. Shechter, E. A. Stern, Y. Yacoby, R. Brener, and Zhe Zhang, Physical Review Letters **63**,(13), 1400 (1989).

[100] I. Stich, R. Car, and M. Parrinello, to be published.

[101] G. Jacucci and A. Rahman, Il Nuovo Cimento **4D**, 341 (1984).