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

ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Vacancy Migration and Self-Diffusion in Silicon via First-Principles Molecular Dynamics

Thesis submitted for the degree
of Magister Philosophiæ

Candidate: Enrico Smargiassi

Supervisor: Prof. Roberto Car

Academic Year 1987–88

TRIESTE

Vacancy Migration and Self-Diffusion in Silicon via First-Principles Molecular Dynamics

Thesis submitted for the degree
of Magister Philosophiæ

Candidate: Enrico Smargiassi

Supervisor: Prof. Roberto Car

Academic Year 1987–88

Introduction

The study of the diffusion processes in solids and particularly 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 has also a fundamental importance in the industrial processing of electronic devices (it is not by chance that the interest in the sixties turned from Germanium to Silicon and now towards the III-V compounds). In particular, self-diffusion controls the annealing processes.

A number of possible mechanisms may be considered for the self-diffusion: vacancy, in which atoms diffuse by means of jumps into an empty site, divacancy [1], in which two bounded vacancies are considered (more complex aggregates of defects have been also proposed [2]), interstitial and interstitialcy, in which an off-site atom wanders in the crystal eventually exchanging with an in-site, or even intrinsic (i. e., without defects) ones, such as the "concerted exchange" model [3] in which two atoms directly interchange their positions. Of course it is possible that more than one mechanism are effective simultaneously.

From an experimental point of view, the situation for Silicon is far from clear. In fact, due to some difficulties, e. g. that of discriminate with certainty one kind of defect from another, both the formation and migration enthalpy of the vacancy are still under debate. A considerable spread exists in the data. The same is true for the self-diffusion coefficient data where, in addition, a marked difference exists between high ($> 1300\ ^\circ K$) and low temperature data.

To estimate theoretically the diffusion coefficient at a given temperature of a thermally activated defect (we shall refer to the vacancy) it is necessary to evaluate:

a) their concentration, proportional to $e^{-\frac{G_f}{k_b T}}$, G_f being the Gibbs free energy of formation of the vacancy;

b) the vacancy jump frequency, proportional to $e^{-\frac{G_m}{k_b T}}$, G_m being the Gibbs free energy of migration.

The diffusion coefficient then will be proportional to the product of these 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 performed.

A possible way to solve the problem consists in performing Molecular Dynamics simulations, in which finite temperature effects are fully taken into account. However, MD relies on the assumption that a satisfactory ion-ion potential can be found. Usually an empirical potential fitted to some bulk properties is chosen, such as the Lennard-Jones for rare gases, but this way meets difficulties in the case of covalently bonded materials, in which many-body effects are relevant.

In fact, due to the presence of relatively delocalized valence electrons, the interaction between two atoms is affected by the presence of others, giving rise to three- (or more) body terms in the potential energy.

The recently developed Molecular Dynamics - Density Functional scheme [4] seems to be a promising way to tackle the problems illustrated above. In this Thesis, we illustrate the preliminary work done to apply this method to the case of the self-diffusion in Silicon due to the vacancy mechanism.

Both the problem of the vacancy formation and of its migration can be tackled by using first-principle Molecular Dynamics: the knowledge of the phonon spectrum allows one to get G_f [5], while a direct simulation of the jump process gives a quantitative value of its frequency, fully taking into account entropic, anharmonic and jump correlation effects.

The results of an extensive phase of testing show that a satisfactory description of the ideal Silicon vacancy can be achieved, yielding results in good agreement with other theoretical estimates. In particular, formation energies are obtained with an error of a few tenth of an eV.

A first simulation of the jump process then has been performed, following the time evolution of the system, heated to $\sim 1300\ ^\circ K$, for $\approx\ 2ps$. A first jump into the vacancy, followed by an immediate return and by some unsuccessful trials has been observed, showing the feasibility of the simulation. Higher-temperature runs, in which a sensible increase of the jump rate is expected, are in progress.

We have not attempted to perform a full calculation of the free energy of formation. Even if in principle a fully anharmonic treatment can be used, great difficulties are encountered [6] and it is more simple, but accurate enough, to do the job in the framework of the so-called quasi-harmonic approximation [7,8]. This requires the evaluation of the dynamical matrix; a possible computational scheme has been recently proposed by Bachelet and De Lorenzi [9].

Finally, a study of some alternative models (e. g. the Pandey model [3]) along the same lines is planned, allowing to compare these various mechanisms for self-diffusion.

This Thesis is organized as follows:

- In the first chapter, the theory and the experimental results on Silicon self-diffusion are shortly reviewed.
- In the second chapter the main aspects of the computational method are illustrated.
- In the third chapter the results of calculations are exposed and discussed.
- In the last chapter we draw some preliminary conclusions and outline the future work.

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge Dr. Guido L. Chiarotti for invaluable help during the computations, Prof. Michele Parrinello and Dr. Franco Buda for useful discussions and suggestions.
The computational work has been performed under the SISSA-CINECA collaborative project.

Contents

1 Self-Diffusion in Silicon 1

1.1 Thermodynamical Aspects 1

1.2 Experimental Results 3

1.2.1 Self-Diffusion Coefficient 3

1.2.2 Formation Enthalpy 5

1.3 Interpretation in Terms of Microscopic Mechanisms 7

1.4 Theoretical Studies 9

1.4.1 Statical Computations 9

1.4.2 Dynamical Simulations 12

2 The Computational Method 15

2.1 The Unified Molecular Dynamics - Density Functional Approach 15

2.1.1 Density Functional Theory and Local Density Approximation 15

2.1.2 The Unified MD-DF Scheme 17

2.2 Computing the Thermodynamical Properties 20

3 Results and Discussion 22

3.1 Checking the Convergence 22

3.1.1 Convergence in the Cell Size and in the Number of k-Points 23

3.1.2 Convergence Respect to the Plane-Waves Cutoff . . 25

3.1.3 The Relaxed Vacancy 25

3.1.4 Convergence in γ 28

3.1.5 Effect of the d-Component of the Pseudopotential . 28

3.2 The Jump Simulation 30

4 Conclusions 37

Chapter 1

Self-Diffusion in Silicon

1.1 Thermodynamical Aspects

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

$$\begin{aligned} \mathbf{j} &= -\boldsymbol{D} \cdot \nabla c \\ \frac{\partial c}{\partial t} &= -\nabla \cdot j = \nabla \cdot (\boldsymbol{D} \cdot \nabla c) = \boldsymbol{D} \nabla^2 c \end{aligned} \tag{1.1}$$

where c is the concentration of the diffusing particles and $\mathbf{j}$ their flux. The coefficient $\boldsymbol{D}$ 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 lattice materials $\boldsymbol{D}$ is a multiple of the identity and therefore we shall consider it as a scalar. The last equality in 1.1 is true only for homogeneous systems of this kind, which we are referring to.

An early analysis of Einstein using a random walk model links D to the mean square displacement of the particles:

$$D = \frac{< x^2(t) >}{6t} = \frac{\Gamma d^2}{6} \tag{1.2}$$

where the second equality refers to a lattice[1]: Γ is the jump frequency between a site to another and d its amplitude. If more kind of diffusion

[1]For lattices having less than cubic symmetry the factor 6 should be changed.

processes (supposed incorrelated) are present, D is the sum of all the individual coefficients.

The assumption of randomness of the jumps is not always justified: since the jumping particle has a peculiar dynamical situation, it probably has a different, generally higher, jump probability from 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. This is taken into account by a multiplicative factor f; for self diffusion in elemental materials it is (neglecting isotope effects) a mechanism dependent geometrical factor [10]. For interstitials and intrinsic mechanisms $f = 1$ (no effect); for the vacancy migration mechanism f has been found to be less than 1 [6,60] (1/2 for the diamond lattice).

For self-diffusion, we can write $\Gamma = \omega c$, with ω the frequency of the single process (e. g. the jump of an atom in a vacancy) and c the concentration of sites available for it. For intrinsic self-diffusion, $c = 1$; for defect-mediated processes, we know that the defect concentration is [11] $c = e^{-G_f/k_bT} = e^{S_f/k_b}e^{-H_f/k_bT}$ where G_f is the Gibbs' free energy of formation, that is the G of the defect system minus that of the perfect crystal (and similarly for H and S).

The ratio of atoms reaching the saddle-point for migration is [12] $x = e^{-G_m/k_bT} = e^{S_m/k_b}e^{-H_m/k_bT}$, H_m and S_m being the enthalpy and the entropy at the saddle point, measured respect to that of the defect system in the minimum-G state, and if ν is the frequency at which the saddle-point-atoms goes into the new site (an unknown quantity of the order of the highest vibrational frequency of the system), we get $\omega = x\nu$ and thus

$$\Gamma = \nu e^{-G_d/k_bT} = \nu e^{(S_f+S_m)/k_b}e^{-(H_f+H_m)/k_bT} \tag{1.3}$$

Putting this equation together with the eq. 1.2 we see that diffusion is a thermally activated process

$$D = D_0 e^{-Q/k_bT} \tag{1.4}$$

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

Such exponentially-behaved phenomena are at best described in terms of Arrhenius plots. They are graphs in which the ordinate axis, which in

this case refers to D, has a logarithmic scale while the abscissa refer to the inverse temperature. That is, what is plotted is $\log D$ versus $1/T$ and it is easily seen that the curve, if the values of Q and D_0 do not depend of T, is a straight line with a slope proportional to $-Q$ and an intercept with the vertical axis at the point $(1/T = 0, \log D_0)$.

Deviations from this shape must be attributed to the failure of the preceding assumptions, namely 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 which case two or more temperature ranges usually exist in which each single process is dominant. A knee in the curve indicates the temperature at which the cross-over between two of these regimes happens.

Even the presence of non-thermal processes may alter the graph.

1.2 Experimental Results

1.2.1 Self-Diffusion Coefficient

The experimental study of self-diffusion in semiconductors is not an easy task [13]. The main way to perform experiments is to diffuse radioactive isotopes in the sample and after some time to detect in some way the position they have assumed. The results are then fitted to an appropriate solution of eq. 1.1 to extract D at a specific temperature.

The first studies [14,15,16] were done by directly diffusing the only isotope disposable: ^{31}Si, with a half-life of 2.6 hours. It is clear that with such a short lifetime a very small distance can be covered, 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 [17] or exploiting its resonance in the proton scattering cross section [18,19,20], provided that the natural background of natural ^{30}Si (a concentration of about 3% in the natural Silicon) is properly subtracted.

The experimental results are shown in Table 1.2.1 and in figure 1.2.1. First, it must be stressed that an intrinsic uncertainity is inserted by the way in which the data are manipulated to extract physical quantities. For example, choosing for the data in ref. [14] a fitting value of 4.1 eV instead of 4.77 one get $D_0 \simeq 100\, cm^2 s^{-1}$ instead of 1800. The fitting procedure in fact, due to the relatively small temperature range in which measurements are done, is very delicate and so an error even of 100% in evaluating D_0 is

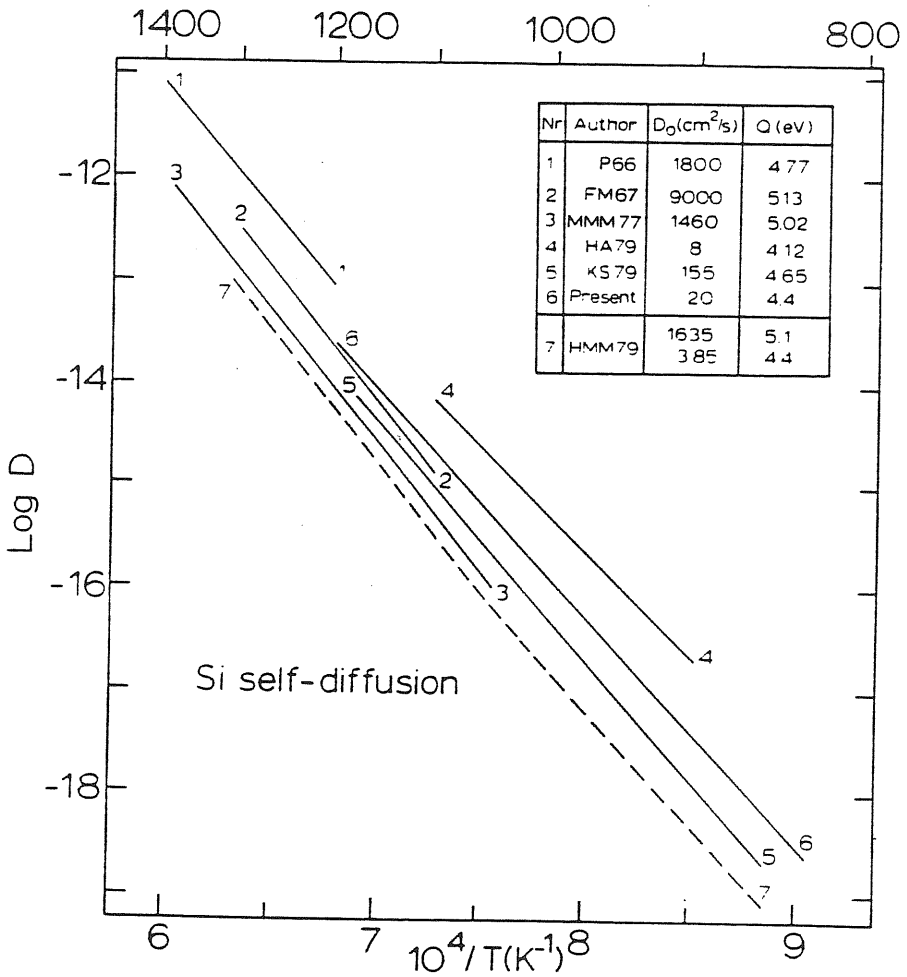

Figure 1.1:
Experimental data on self-diffusion coefficient in Silicon. Curve n.1: ref. [14]; n.2: ref. [15,16]; n.3: ref. [21]; n.4.: ref. [18]; n.5: ref. [19]; n.6: ref. [20]; n. 7: ref. [23] (^{71}Ge diffusion).

Reference	D_0 $(cm^2\ s^{-1})$	Q (eV)	Temperature range $(°K)$
Peart [14]	1800	4.77	1473–1673
Ghoshtagore [17]	1200	4.72	1451–1573
Mayer *et al* [21]	1460	5.02	1320–1660
Fairfield and Masters [15,16]	9000	5.13	1373–1573
Sanders and Dobson [22]	5.8	4.1	1243–1343
Kalinowski and Seguin [19,61]	154	4.65	1128–1448
Demond *et al* [20]	20	4.4	1103–1473
Hirvonen and Anttila [18]	8	4.1	1173–1373

Table 1.1:
Experimental enthalpy of migration and pre-exponential factor for Silicon.

not astonishing. This partly explains the large scatter in the values of D_0 at about the same T.

Since the properties of the defects are sensitive to the charge state, a doping dependence of D can be anticipated [24]. Table 1.2.1 reports the self-diffusion data over a range of high doping, both p and n. The quantity reported is the ratio between the self-diffusion coefficient in the doped samples and that in the undoped ones. Except for the data of ref. [17], which however have been questioned [25], all the data agree to give an appreciable increase of the diffusion coefficient. This however could be partially due to the onset of some impurity-mediated mechanism.

1.2.2 Formation Enthalpy

The direct detection of formation enthalpy and so of the equilibrium density of vacancies (and other thermally activated defects) is also a difficult task.

First, several kind of intrinsic defects may be created by thermal motion, and, at variance with the case of impurities, their concentration, either absolute or relative, cannot be varied as wanted. This implies that single phenomena cannot unambiguously be attributed to a single defect species, unless some additional information on the effect of specific defect mechanisms is available.

Reference	Doping type and dopant	Dopant conc. ($10^{19}\ cm^{-3}$)	$\frac{D_{dop}}{D_{und}}$	Temperature range ($°K$)
Fairfield	n: P, As	8–18.8	1.35–3.15	1363–1470
and Masters [16]	p: B	8–22	1.–1.75	
Ghoshtagore [17]	n: P	8.6	1.6	1451–1573
	p: B	10	0.72	
Hettich *et al* [23]	p: B	0.6–1.2	1.4–1.2	1318–1515

Table 1.2:
Doping dependence of the pre-exponential self-diffusion factor for Silicon. Tracer is ^{31}Si.

Second, due to the low concentration of intrinsic defects even at melting temperature[2], 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 orders of magnitude smaller than the concentration of thermally excited electrons.

The only way to proceed was for a long time to quench the crystal to a temperature low enough to allow a sensible detection [27]. This introduces a further source of error, because the Silicon vacancy, for example, is found to migrate even at very low temperature [28]: the typical clustering times being of the order of milliseconds, even cooling rates as high as $10^3\ °K/s$ do not exclude the possibility of substantial annealing and clustering of defects [29].

A few measurements of this kind have been performed, we refer here to the work of Estner and Kamprath [27], who detected the reduction of the free-carrier concentration in n-doped Silicon attributing it to the trapping by single vacancies, yielding $H_f = 2.4 \div 2.5\ eV$. Together with the value of $0.33\ eV$ for H_m obtained by Watkins [28] from low-temperature annealing data we obtain $Q \simeq 2.8\ eV$, far lower than reported in the previous section

[2]From self-diffusion data, under the assumption that the diffusion takes place via a defect mechanism, it is possible to bound the intrinsic defect concentration at melting [26] by $10^{-7} \div 10^{-10}$.

(Q about 4 eV). Since the data of ref. [28] are generally considered reliable, the probable explanation of the discrepancy relies in the difficulties mentioned above.

Recently however the refinement of the positron-annihilation technique [30] has allowed a set of measures [31] in a wide temperature range (300–1523 $^\circ K$), therefore avoiding the previous uncertainties. The reported value is $H_f = 3.6 \pm 0.2\ eV$, in good agreement with the estimates cited above. These authors also evaluate the formation entropy of the vacancy to be as high as $6 \div 10\ k_b$, but this value is subject to the uncertainity on the value of the dielectric constant in the neighbourhood of the vacancy (estimated to be ~ 6, but actually not known) and to the assumptions on the temperature dependence of the lifetime of positrons in the vacancy (a different choice would not even allow to estimate S).

1.3 Interpretation in Terms of Microscopic Mechanisms

A number of microscopic mechanisms has been proposed for the Silicon self-diffusion. Here we limit ourselves to a short survey, referring the reader for detailed discussions to the many reviews available [11,12,25,26,32].

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

i) A neat discontinuity appears at about 1300 $^\circ K$, for under that value the pre-exponential factor decreases, as jointly results from all the studies, of about two orders of magnitude. Q, over whose values a small error is expected, decreases from about 5 eV to about 4 eV. As previously said, this can be attributed to a variation of the migration mechanism or to a strong temperature dependence of the thermodynamical functions H and S.

ii) The high temperature values of D_0 are very high (so implying a very high diffusion entropy) compared not only with metals but even with Germanium.

A first possibility is that of an *intrinsic*, that is defect independent mechanism, which could be a direct exchange or a ring exchange one [11].

For a long time this has been regarded as scarcely reliable, due to the high energy required to break six covalent bonds. However, recently Pandey [3] has shown that a mechanism can be thought requiring the breaking of only two bonds at a time, bounding the need of energy within 5 eV, a value not far from experiments.

Other conceivable mechanisms are the *interstitial* or *interstitialcy*. Both involve off-site atoms, the difference being the way in which the migration happens. In the former the same ion wanders along the lattice hopping from one stable site to another, while in the latter it pushes away a "regular" atom which so becomes the interstitial one. What is the way in which interstitials move is decided of course by the enthalpy barriers it encounters.

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 are not at all obvious, so they need specific analyses. Recently some indirect experimental evidence for this mechanism has been found [32].

Following what is commonly believed about self-diffusion in Germanium, a *vacancy* mechanism got wide consideration[3]. Actually, in Ge the values of D_0 are roughly in agreement to what is found in Silicon at low T. This suggests that, at least in this range, this mechanism could be the dominant one.

Theoretical (see sect. 1.4) and experimental reasons lead one to believe that the vacancy has a low migration enthalpy of about $0.2 \div 0.4$ eV. Efforts have been produced [1,62,59] in order to show that the atomic relaxation and the phonon softening around the vacancy ("relaxion") can bring the formation entropy of the vacancy in agreement with evaluations made on the basis of *ii)*, which suggest $S/k_b \approx 8 \div 16$. It is worth that some experimental evidence has been produced supporting them [31].

A similar way to obtain an high S_m is to choose a mechanism which surely strongly perturbs the surrounding lattice. A *divacancy* mechanism is a possible one [1,12,14,17], but it is not clear if the formation enthalpy is low enough to allow a significant concentration of them to be present [25,31]. Further, it has yet to be shown that the migration may happen without quick dissociation into a pair of single vacancies [2,12].

[3]Even a *split-vacancy* mechanism has been proposed [33], suggesting that this could be the stable configuration. Experimental and recent theoretical studies, including ours, rule out this possibility.

The concept of strong relaxation has been extended [2] both for vacancy and interstitial up to suppose that some sort of *local melting* may take place around those defects. The onset of this phenomenon would have to be at the temperature at which the knee in the Arrhenius plots shows up; at lower temperatures "normal" vacancies and/or interstitials would be found. This would explain as the high entropy over that point as its sudden change (even H_f should be raised). Due to the complexity of this kind of model, however, it is difficult to assess its validity; further, it must be noted that positron-annihilation experiments [31] seem not to detect such a large modification of the structure at any temperature.

1.4 Theoretical Studies

To support and test the various views some theoretical computations have to be performed. They may be divided into two main categories: statical (absolute zero) and dynamical. We shall expose them separately.

1.4.1 Statical Computations

In these schemes 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 found by computations within a number of candidates. One then tries to get connections with the dynamical quantities of interest. The system is thought to be not only at thermodynamical equilibrium, but also at absolute zero.

The first results of this type, although several different attempts had been done [2,34,35], were unsatisfactory. As an example, we quote the work of Seeger [2] which in spite of the efforts spent to improve its model, based on semiempirical Morse potential, got a formation enthalpy of only 2.35 eV which, being in (probably fortuitous) agreement with the experimental data of Estner and Kamprath [27], is subject to the same strong criticisms.

Modern schemes are instead based over refined self-consistent methods and need large disposability of computer resources. The main results regarding the Silicon vacancy are listed in table 1.4.1. The results of Bar-Yam and Joannopoulos [36] were obtained by means of a supercell (up

Reference	V^0			V^{++}		
	H_f	H_m	Q	H_f	H_m	Q
Swalin [34]	2.32	1.06	3.38			
Benneman [35]	2.13	1.09	3.22			
Seeger [2]	2.35					
Bar-Yam *et al.* [36]	3.6	0.5°	4.1	4.6	0.3°	4.9
Car *et al.* [37,38]			4.2			5.1
Kelly *et al.* [39]	3.92	0.27	4.19	3.43	0.42	5.02

° empirical value

Table 1.3:
Theoretical enthalpy for the neutral and the doubly charged Silicon vacancy.

to 32 atoms) method, the others [37,38,39] using a self consistent Green's function method. All these authors included atomic relaxations. Note that while the formation enthalpy of the neutral vacancy is independent of the Fermi level μ_f, hence of doping, that of the charged ones does depend on it: the values reported in the table 1.4.1 are obtained supposing that μ_f lies at the middle of the gap for intrinsic Silicon.

Some of the data are directly comparable with experiments. For example, from section 1.1 we see that the empirical activation energy Q is the sum of the microscopic H_f and H_m. So we have listed even the sum of these enthalpies which have to be compared to the value of $4.1 \div 5.1$ cited in the section 1.2. The agreement is in general quite good, and similarly when the single migration enthalpy is referred to empirical values, that is the Watkins' data (0.33 eV for the double plus vacancy and 0.45 for the neutral).

We stress an important limitation of the static methods: they neglect finite temperature and dynamical effects. They neglect the possibility of immediate return jumps, just after the crossing of the barrier top, the possibility of multiple jumps, and are open to the criticisms that the true relaxation of the system during the jump is not well taken into account: for slow events a full relaxation is expected, but for fast ones its entity is not in this way estimable.

In other words, they simply disregard the dynamical, finite temperature

nature of the diffusion processes, that is one of their more prominent characteristics. This is disappointing since, for example, as seen in the previous sections there are experimental indications that temperature effects are not at all negligible.

What then cannot be done in a simple manner is the evaluation of the contribute of the single mechanisms to self-diffusion. In fact, such results do not carry sufficient informations to estimate quantitatively the major contribute given by the formation and migration entropy to the pre-exponential factor D_0.

A bridge towards these data must be supplied otherwise. A possible and largely exploited way is to use the so called "rate theory", or the "dynamical theory" (which however being physically and, up to the second order in the potential, numerically equivalent to the former will not be examined here).

The task of the rate theory is, if c is known, since $D \propto \Gamma = \omega c$, to calculate ω. If we assume that the system remains in the thermodynamical equilibrium before and throughout the jump and do not take into account the dynamical correlations we get [8,11]

$$\omega = K e^{-\Delta V/k_b T} \tag{1.5}$$

where K is given by

$$K = \frac{\prod_i^{3N-3} \omega_i}{\prod_i^{3N-4} \omega_i'} \tag{1.6}$$

where ω_i are the eigenfrequencies of the system around the minimum of the energy (here and later we shall subtract the center-of-mass coordinates from the degrees of freedom), and ω_i' the same but around the saddle point and excluding the imaginary one, i. e. that along the reaction coordinate.

Eq. 1.5 contains the difference in the potential energy between the saddle and the minimum point. Since the migration is a constant pressure process, it actually is a migration enthalpy. To get a migration free energy we should have to extract an entropy factor from K: if we put [8] $\overline{\omega}^{3N-3} = \prod_i^{3N-3} \omega_i$ we can write $S = k_b \sum_i^{3N-4} \log(\overline{\omega}/\omega_i')$ and so $K = \overline{\omega} e^{S/k_b}$. This choice however is by no means unique.

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 us the value of the diffusion coefficient for the considered mechanism.

1.4.2 Dynamical Simulations

A well known way to take fully into account the aspects overlooked by the previous methods is to simulate the true time evolution by means of computer experiments: this is the Molecular Dynamics (MD) method [6,8,40].

In principle, the method seems to be plainly applicable to our systems. It seems sufficient to start with a perfect crystal system, heat it and wait for the formation of a sufficient number of the the wanted defects; then their migration could be monitored, and the rates of creation and motion would directly give the physical quantities sought.

But this simple strategy cannot be practically used. Left apart considerations on the interatomic potential, on which we shall return later, the creation rate of defects is far lower than detectable from today's simulations. 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 the thermodynamical quantities of a previously created one and using $c = e^{-G_f/k_bT}$. Even in this case, the direct computation of the relevant thermodynamical quantities like F, the Helmholtz free energy, or S, 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 as configurational integrals which having quick varying integrands are very hard to solve [40].

This may instead be achieved in a satisfactory way using the quasi-harmonic approximation. Let us start with the expressions of F,U,S,P in the harmonic approximation [8]:

$$\begin{aligned} F &= \phi_0 + \frac{1}{2}\sum_i^{3N-3} \hbar\omega_i + k_bT \sum_i^{3N-3} \log\left(1 - e^{-\frac{\hbar\omega_i}{k_bT}}\right) \\ &\simeq \phi_0 + k_bT \sum_i^{3N-3} \log\left(\frac{\hbar\omega_i}{k_bT}\right) \qquad (1.7) \\ U &= \phi_0 + \frac{1}{2}\sum_i^{3N-3} \hbar\omega_i + \sum_i^{3N-3} \frac{\hbar\omega_i}{e^{\frac{\hbar\omega_i}{k_bT}} - 1} \simeq \phi_0 + 3Nk_bT \qquad (1.8) \end{aligned}$$

$$
\begin{aligned}
S &= -k_b \left[\sum_i^{3N-3} \log\left(1 - e^{-\frac{\hbar\omega_i}{k_b T}}\right) + \sum_i^{3N-3} \frac{\hbar\omega_i / k_b T}{1 - e^{\frac{\hbar\omega_i}{k_b T}}} \right] \\
&\simeq -k_b \left[\sum_i^{3N-3} \log\left(\frac{\hbar\omega_i}{k_b T}\right) - 1 \right] \qquad (1.9) \\
P &= P_\phi + P_\omega = -\frac{\partial \phi_0}{\partial V} - \left[\frac{\partial}{\partial V} (F - \phi_0) \right]_T \qquad (1.10)
\end{aligned}
$$

where ϕ_0 is the potential energy at the minimum energy configuration and the approximate equalities refer to the high-temperature case.

This approximation is not fully satisfactory. The anharmonic effects, which in the case of defect jumps are expected to be important since the particles sample unusually high energy regions, are responsible of a lot of important physical effects, such as the thermal expansion; and since in the eqs. 1.7 there is no volume dependence in the ω's we don't find any temperature dependence of the pressure.

In the Quasi-Harmonic (QH) approximation the anharmonicity is approximately introduced by a volume dependence of ϕ_0 and ω_i, the latter using Grünesein parameters, maybe empirical.

If we choose to compute them we have to perform a number of computations over different volumes (for each a new minimization of the energy respect to the ionic positions is required) of ϕ_0 and ω_i both for the bulk and the defect crystal, then to fit the results to some suitable function and at the end compute F_{bulk} and F_{vac}. The free energy of formation of the vacancy is given by [9]

$$
F_{CL} = F_{vac}(V_{bulk}, T) - \frac{N-1}{N} F_{bulk}(V_{bulk}, T) \qquad (1.11)
$$

if the lattice parameter is fixed,

$$
F_{CV} = F_{vac}\left(\frac{N-1}{N} V_{bulk}, T\right) - \frac{N-1}{N} F_{bulk}(V_{bulk}, T) \qquad (1.12)
$$

if the volume is fixed,

$$
G_{CP} = F_{vac}(V_{vac}, T) - \frac{N-1}{N} F_{bulk}(V_{bulk}, T) + P_{ext}\left(V_{vac} - \frac{N-1}{N} V_{bulk}\right) \qquad (1.13)
$$

if the pressure is fixed (in this case in fact the appropriate thermodynamical potential is G). Here V_{vac}, V_{bulk} are the volumes at which the internal pressure and the external are equal.

The approach described above has been successfully applied many times to simple cases, namely the ones in which a satisfactory potential function can be found: mainly noble gases (Lennard-Jones systems) ionic systems and some metals. For covalent semiconductors the problem is on the contrary hard: the presence of relevant classical and quantal many-body effects, sensitive to the chemical environment up to now ruled out for general purposes every proposed scheme. In fact, valence-force schemes [41] are fitted and useful only 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.

Regarding the transferability of the potential we note that, for example, the widely used Stillinger and Weber's form [42] has shown to be incapable of dealing correctly with all the three phases (solid, liquid and amorphous) of Silicon with the same parameters.

As a conclusion of the previous bird-eye view we can convince ourselves that for an accurate study of the self-diffusion problem a first-principle dynamical technique is particularly welcome. A strategy of this kind exists: it is the joint Molecular Dynamics - Density Functional scheme, which will be exposed and used in the next chapters.

Chapter 2

The Computational Method

2.1 The Unified Molecular Dynamics - Density Functional Approach

In this approach the interatomic forces are directly derived from the electronic ground state. This is treated in the framework of the Density Functional Theory (DFT) within the Local Density Approximation (LDA) for exchange and correlation effects.

In the present method the electronic wavefunctions are regarded as classical degrees of freedom of the system. The "motion" associated with them is a fictitious one, simply helping to keep always the electrons in their ground state, but the ground state properties of the electrons correspond to the "real" ones, and the study of quantities depending from the electrons (such as the electron density of states) are naturally feasible in this context. Note that this is of course not the case with empirical classical potential.

2.1.1 Density Functional Theory and Local Density Approximation

At the root of the method lies the well known Born-Oppenheimer adiabatic approximation (BO) that consists in considering decoupled the ionic and electronic motion. The justification for this is the very large difference in electronic and ionic masses; actually, such treatment can be shown to be correct to the first order in $\sqrt{\frac{m_{el}}{M_{ion}}}$ [43].

According the BO approximation the interatomic potential for a given nuclear configuration consists of the electronic ground-state energy plus the direct electrostatic interaction between the nuclei. Then at temperatures not too low, where quantum effects on the nuclear motion are negligible, the atomic dynamics is well described by Newton's equations using the BO interatomic potential energy. The problem is then to compute the electronic ground-state for a generic atomic configuration.

DFT provides a way of doing that, formally transforming the many body electronic problem into a self consistent single particle problem. Moreover, unlike Hartree-Fock, the resulting effective potential has a simple multiplicative form.

The theory is due to a very remarkable theorem demonstrated by P. Hohenberg and W. Kohn in 1964 [44,45]. According to that, the total energy of a system of density $n(r)$ in an external potential $v_{ext}(r)$ can be written as

$$E_{tot}[n(r)] = T[n(r)] + \int dr\, v_{ext}(r)n(r) + \int\int dr dr' \frac{n(r)n(r')}{|r-r'|} + E_{xc}[n(r)] \tag{2.1}$$

where $T[n]$ is the kinetic energy of a system of non interacting electrons with the same density $n(r)$ and $E_{xc}[n]$ is the remaining part (containing the so called exchange and correlation energy). The universal functional $E_{xc}[n]$ contains all the many-body information.

We are left with the problem of finding the density $n(r)$ corresponding to the minimum energy. It would be desirable to obtain a single-particle-like set of equations for some orbitals ϕ_i, to satisfy

$$E[n(r)] = \min_{\{\phi_i\}} \{E[\phi_i, R_I]\} \tag{2.2}$$

for a given configuration of ions.

Applying the stationary condition ($\delta E_{tot}/\delta n = 0$) to eq. 2.1 Kohn and Sham [46] were able to work out the so called KS equations for electrons acting under an effective potential v_{eff}

$$\left\{-\frac{1}{2}\nabla^2 + v_{eff}\right\}\phi_i = \varepsilon_i \phi_i \tag{2.3}$$

where

$$v_{eff}(r) = v_{ext}(r) + \int dr' \frac{n(r')}{|r-r'|} + \frac{\delta E_{xc}[n(r)]}{\delta n(r)} \tag{2.4}$$

and the density is given by

$$n(r) = \sum_i |\phi_i|^2 \tag{2.5}$$

The ϕ_i and ε_i are the so called KS orbitals and KS levels respectively.

The exact form of $E_{xc}[n]$ is of course unknown. The simplest and most used approximation is the so called "Local Density Approximation" (LDA) in which $E_{xc}[n]$ is supposed to be the same of an homogeneous electron gas having the same n. For this system, $E_{xc}[n]$ can be written as

$$E_{xc}[n(r)] = \int dr\, n(r)\varepsilon_{xc}\left(n(r)\right) \tag{2.6}$$

where $\varepsilon_{xc}(n)$ is the exchange and correlation energy per electron of an electron gas of uniform density n and eq. 2.4 becomes

$$v_{eff}(r) = v_{ext}(r) + \int dr' \frac{n(r')}{|r - r'|} + \frac{d}{dn}\left\{n(r)\varepsilon_{xc}\left(n(r)\right)\right\} \tag{2.7}$$

The function $\varepsilon_{xc}(n)$ has been extracted and fitted to an analytic expression [47] from accurate ground state Monte Carlo computations [48] of the homogeneous electron gas.

2.1.2 The Unified MD-DF Scheme

In the Unified Approach to DFT and MD [4] the electronic wave function and the ionic coordinates appearing in the 2.2 are regarded as dynamical variables entering a Lagrangian of the type:

$$\begin{aligned} L = &\frac{1}{2}\mu \sum_i \int_\Omega d^3r |\dot{\psi}_i|^2 + \sum_I \frac{1}{2} M_I \dot{R}_I{}^2 \\ &-U\left[\{\psi_i\}, \{R_I\}\right] + \sum_{ij} \Lambda_{ij} \left(\int_\Omega d^3\mathbf{r}\, \psi_i^* \psi_j - \delta_{ij} \right) \end{aligned} \tag{2.8}$$

The last term in 2.8 takes care of the orthogonality constraints for the variables ψ_i. The Hermitian matrix Λ_{ij} is the Lagrangian multipliers matrix introduced to impose the orthonormality condition. M_I are the ionic masses and μ is an adjustable parameter.

The Lagrangian of equation 2.8 gives the following differential equations for the variables ψ_i and R_I:

$$\mu\ddot{\psi}_i(\mathbf{r},t) = -\frac{\delta E}{\delta \psi_i^*} + \sum_k \Lambda_{ik}\psi_k \tag{2.9}$$

$$M_I\ddot{R}_I = -\nabla_{R_I} E \tag{2.10}$$

It can be shown that the equation 2.9, once the equilibrium condition for electrons is reached (i.e. $\ddot{\psi}_i = 0$ is satisfied) is equivalent within a unitary transformation to the KS equation 2.3. In this situation the eigenvalues of the Λ matrix coincide with the KS eigenvalues.

Even different procedures can be adopted to solve eq. 2.2: a steepest descent method to minimize with respect to the electronic (and ionic, if required) degrees of freedom may be used.

Under the electronic point of view, this implies that the initial values imposed to the ψ's must be such that the initial state is not orthogonal to the ground state. A smart choice of the initial values of the ψ's of course is preferable.

In the majority of the computations we diagonalize a small matrix built by using the empirical pseudopotential introduced by Louie *et al.* [49] which gives a reasonable starting approximation for the wavefunctions, usually with the right symmetry. Eventually a small random displacement was added to the electronic wavefunctions to give a missing component.

Once the minimization with respect to the $\{\psi_i\}$ is done the system lies in the BO surface; then the time evolution of the ionic position may start. The equation 2.9 has to be seen essentially as a tool of gaining the BO surface and to remain on it: no physical meaning has to be assigned to the dynamics of the electrons. The equation 2.10 instead represents the correct dynamics for the ionic system moving on the BO surface, under the effect of the instantaneous electronic ground state potential E_0.

In our method we have two requirements to fulfill: the total energy has to be conserved and the system should remain on the BO surfaces. This is possible by tuning the two adjustable parameter μ and Δ. The parameter μ has to be chosen such that the system remains as long as possible on the BO surface: the energy transfer between electronic and ionic subsystems has to be negligible.

A good compromise between the possibility of a faster integration of the phase space trajectories (that would require large μ) and this BO condition

has to be practically found. For silicon the value for the parameter μ was chosen to be 300–400 a.u. with an integration time step of the order of $10^{-16}s$. However, a good adiabaticity is difficult to achieve if the system has a metallic behaviour, i. e. small electronic excitation energies.

Why introduce such a machinery? The most important point is that the combined use of eqs. 2.9, 2.10 allows to obtain the BO nuclear dynamics while keeping the electrons very close to the instantaneous ground state. Once the electrons are initially brought into it the classical system described by the Lagrangian 2.8 evolves almost adiabatically and very few, if any, additional electronic minimizations are required to keep the system on the BO surface.

In addition, eqs. 2.9, 2.10 can be solved quite efficiently: suppose we want to solve eq. 2.3 for a system of N electrons; if we use M plane waves to represent a single wave function (the formal details may be found in refs. [50,51]), the solution of eq. 2.3, using standard methods, requires the diagonalization of a rank M matrix that give rise to $O(M^3)$ operations. Equation 2.9 instead requires $O(N \times M \times \ln M)$ and $O(N^2 \times M)$ operation (the last due to the orthonormalization process) that in case of $M \gg N$ is a considerable computational advantages making possible to treat larger systems.

The computation of the electronic density $n(\mathbf{r})$ requires a summation over the whole Brillouin Zone of the system:

$$n(\mathbf{r}) = \sum_{n,\mathbf{k}\in BZ} \psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}) \tag{2.11}$$

A MD cell is used on which periodic boundary conditions are imposed. If the system is disordered (that is always if $T \neq 0$) a spurious periodicity is then introduced which can be made harmless only by enlarging to a sufficient extent the cell itself.

Using special points in the BZ can reduce the infinite sum appearing in the 2.11 to a sum of only few terms [52,53] still having a good accuracy in representing the density. Making a larger MD cell the BZ becomes smaller and fewer k-points in eq. 2.11 are enough (for infinite cells only the $\mathbf{k} = 0$ point appears in the sum).

Then, approximating sum 2.11 with just the $\mathbf{k} = 0$ (Γ) point gives satisfactory results if a sufficiently large cell is chosen. Moreover, invariance under time-reversal allows to write $\varphi(\mathbf{k}) = \varphi^*(-\mathbf{k})$, and so $\varphi(\mathbf{0}) = \varphi^*(\mathbf{0})$;

therefore, using only the Γ point the periodic part of the wavefunctions can be made real, resulting in almost halving the need of independent plane waves coefficients. This leads to prefer large cells with respect to use many k-points.

A pseudopotential approach is used [50]. The most rigorous approach to this problem involves the so called first-principles norm-conserving pseudopotentials [54]. We use the form given by Bachelet *et al.* [55] and the Kleinman-Bylander method [56] for a fast evaluation of the integrals needed.

2.2 Computing the Thermodynamical Properties

The application of the MD-DF scheme to the defect problem does not imply, in principle, any difference from the usual MD methods. A preliminary calculation [5] shows the feasibility of the method for the vacancy.

A peculiar problem however arises in computing the phonon spectra, since the direct way of giving small displacements and to compute from the energy differences the force constants (frozen-phonon approach) leads one to a big number of self consistent calculations, wasting both human than computer resources.

An efficient and automatic method to generate dynamical matrices has been recently developed [9] with particular reference to the possibility of quick generation of dynamical trajectories given by the MD-DF method.

The basic idea is simple. For an harmonic system we can write the matrix equation

$$\boldsymbol{F} = -\boldsymbol{K}\boldsymbol{R} \tag{2.12}$$

where F_{ij} is the force referred to the i-th atomic coordinate and R_{ij} the coordinate itself, at $t = t_j$. The **K** matrix is then the dynamical matrix which now can be straightforwardly computed by matrix inversion (much less expensive than a diagonalization, used in standard methods).

The times $\{t_j\}$ can be chosen in an arbitrary way. A smart choice, expecially when using the MD-DF method, is to generate them such that the corresponding values of R lie on a MD trajectory, that is $t_j = t_0 + (j-1)\Delta t$, with small Δt. In this way a number of configurations can be produced in a cheap and automatic way.

For this method is effective two conditions have to be satisfied: the one physical, the other mathematical. The former is that the system must show small deviations from the harmonic behaviour along the whole trajectory; this can be obtained choosing initial conditions such that the resulting temperature is low enough. The latter derives from the need that the matrix R is non singular: the initial conditions thus must be such that all the normal modes are excited and the displacements are linearly independent. Further, it has been shown that the non-degeneracy can be reached only if the phonon spectrum is no more than twofold degenerate; the remedy is to use more than one MD trajectory, and more precisely at least as many as the greatest degeneracy divided by two. From this fact is seen that this method works at best for very low symmetry systems, for which a single MD trajectory are sufficient.

Chapter 3

Results and Discussion

3.1 Checking the Convergence

The joint MD-DF scheme allows, in principle, an exact evaluation of all the electronic (ground-state) and ionic properties, within the LDA- plus-pseudopotential scheme. Of course, the finite availability of computer memory and time forces one to resort to some approximations.

a) Finite cells are used. In particular, while studying a defect crystal it introduces an unphysical, large-scale ordered structure of defects. The interaction between the elements of such an array shows up, for example, in the broadening the flat band of the localized defects, which can be, even for quite big cells, of the order of about 1 eV [49].

   Since the crystal symmetry of Silicon is T_d, the choice of the cells is limited to the ones having (N being an integer):

$$\begin{cases} 2, 16, 54, 128, \ldots, 2 \times N^3 & \text{FCC cells} \\ 8, 64, \ldots, 8 \times N^3 & \text{SC cells} \\ 32, \ldots, 32 \times N^3 & \text{BCC cells} \end{cases} \tag{3.1}$$

b) The integral in K-space are performed using a finite, and small, number of special points.

c) A finite number of plane waves must be used in expanding the electronic wavefunctions. This is usually expressed by means of the cut-off energy E_c: $N_{pw} \simeq \frac{1}{6\pi}\Omega {E_c}^{3/2} = \frac{1}{6\pi}\Omega {G_w}^3$, Ω being the cell volume.

This implies that the density n and then all the integrals are cut off at $G_c = 2G_w$.

d) However it is possible to substitute G_c with some G_c' such that $\gamma = \frac{G_c'^2}{G_w^2} = \frac{G_c'^2}{E_c} < 4$. The validity of this new cutoff is as better as higher E_c is.

e) The non-local part of the pseudopotentials is taken exactly into account only up to a finite order in the angular momentum. That is, written $\hat{V}_{ps,nl} = \sum_l v_l \hat{P}_l$ the P's being projection operators, we state that $v_{l>l_c} \simeq v_{l_c}$, and using the completeness relation $\sum_l \hat{P}_l = \mathbf{1}$ we obtain $\hat{V} \simeq \sum_{l<l_c}(v_l - v_{l_c})\hat{P}_l + v_{l_c}$. This introduces a small error because high-l components in the wavefunctions do not enter appreciably in the core.

All these approximations can be quantitatively and independently controlled, permitting to achieve (if the computer is powerful enough) the desired degree of convergence, or at least to detect where the failure arises.

What values of the parameters have to be chosen must be discovered by trials for each different system which must be simulated. To this end, an extensive series of computations of the properties of the static (i. e. at absolute zero; ionic relaxation is eventually allowed) system has been done.

3.1.1 Convergence in the Cell Size and in the Number of k-Points

The formation energy of the unrelaxed vacancy, both neutral and doubly charged, is shown in table 3.1.1. The cutoff energy, here and later (unless otherwise stated), was set to 6 Ry, deferring the study of the convergence in E_c.

First, it must be stressed that the energy of formation for V^0 as a function of E_c has an unpleasant "sawtooth" behaviour, when computed using the Baldereschi point. In fact, E_{form} has a maximum for the 32-atoms cell. The possibility of a bad choice of the initial trial state has been excluded by repeating the runs with highly randomized electronic coordinates. This strangeness is probably due to the fact that for too small cells, at k = Baldereschi point at least, a conduction band state exists with

k-point / # of atoms	Γ	Baldereschi point	2 Chadi-Cohen points
8		2.94	3.80*
16		3.89	3.66
32		4.72	4.52
54	3.47	4.35	4.28
64	4.16	4.44	

* 4 points

Table 3.1:
Formation energy in eV of the unrelaxed vacancy as a function of the special points used and of the supercell size. The number of atoms refers to that of the perfect crystal. $E_c = 6Ry, l_c = 1$ and $\gamma = 4$.

energy less than the gap level. The former is filled and what one obtains is a situation physically very different from what was sought.

Again from table 3.1.1 we see that the convergence in the number (and kind) of k-points is reached when the cell size reaches 54–64 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 that and Γ of about 0.3 eV, that is with an error within 10%.

In table 3.1.1 the convergence in size up to 128 atoms is examined, using the Γ point and $\gamma = 1$. Because the use of a different γ, the energies

# of atoms	54	64	128
E_{form}	3.78	4.41	4.57
E_{level}	1.14	1.11	0.93

Table 3.2:
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 supercell size. $E_c = 6Ry, l_c = 1$. The Γ point and $\gamma = 1$ are used.

E_c	6	8	10	12	14
E_{form}	4.16	3.73	3.42	3.22	3.14
E_{level}	1.11	1.05	1.01	0.99	1.00

Table 3.3:
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 cutoff energy E_c(in Rydberg), in a 64 atoms – 63+vacancy supercell. Only the Γ point was used; $l_c = 1$ and $\gamma = 4$.

are shifted. However, to our end is sufficient to see that the variation in E_{form} while passing from 64 to 128 atoms is only of 0.16 eV. On the other hand, the differences encountered going from a 54 to a 64 cell are still quite substantial (0.63). This rules out the possibility to perform calculations using a 54-atom cell, using only Γ.

3.1.2 Convergence Respect to the Plane-Waves Cutoff

Once made sure that a 64-atoms cell with only Γ is a sensible choice, we studied the effect of the cutoff E_c using such a cell (Table 3.1.2). By extrapolation, we estimate the converged value of E_f, under the conditions illustrated, to be $\sim 3 \div 3.1$. Therefore, even with a cutoff as high as 10 Rydberg the error remains about $0.3 \div 0.4$ eV, while using 12 Rydberg it is at least halved.

3.1.3 The Relaxed Vacancy

Even at $T = 0$ the atomic structure of the vacancy is not the "ideal" one, but a relaxation of the nearest atoms at least must be allowed. In this case it is sufficient to allow the ions to move according 2.10 using a fictitious steepest descent dynamics.

Table 3.1.3 shows the results for a 64 atom cell, at different cutoff. The formation energy behaves in the same manner than in the unrelaxed case;

E_c	6	8	10	10	10
E_{form}	3.88	3.28	2.96	2.98	3.14
E_{migr}	0.56	0.40	0.33		0.37
ΔQ_b	-0.31	-0.48	-0.53	-0.49	-0.48
ΔQ_p	$+0.39$	$+0.45$	$+0.45$	$+0.38$	$+0.38$
E_{level}	0.76	0.66	0.66	0.71	0.45

Table 3.4:
Left: 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_c(in Rydberg), in a 64 atoms – 63+vacancy supercell. Even the relaxation of the first shell is shown. Only the Γ point was used; $l_c = 1$ and $\gamma = 4$. Right: $\gamma = 1.5$, 64 atoms cell; $\gamma = 1.5$, 128 atoms cell.

the energy gained in the relaxation process is scarcely sensitive to E_c, being about 0.45 eV.

The atomic relaxation of the first neighbors of the vacancy can be written in terms of two normal coordinates per atom, since only a symmetric and a tetragonal distortion are expected. 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), and Q_p ("pairing" mode) the component of the vector position in one of the directions orthogonal to the previous one. In particular, the second coordinate is chosen such that the plane defined by the Q_p and Q_b directions of one of these atoms contains one of the others (what other is chosen does not matter); that is, Q_p indicates some "pairing" (or "unpairing", if $Q_p < 0$).

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)

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

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

then Q_b will be given by $\mathbf{r}\cdot(1,1,1)/\sqrt{3}$ for the atom 1 ($\mathbf{r}$ being its position) and similarly for the others. The Q_p's will be

$$\begin{aligned} \mathbf{r}_1 \cdot \frac{1}{4}(-1,-1,2)/\sqrt{6} \\ \mathbf{r}_2 \cdot \frac{1}{4}(-1,1,-2)/\sqrt{6} \\ \mathbf{r}_3 \cdot \frac{1}{4}(1,-1,-2)/\sqrt{6} \\ \mathbf{r}_4 \cdot \frac{1}{4}(1,1,2)/\sqrt{6} \end{aligned}$$

instead of 0, as in the ideal geometry. By symmetry all these values must be the same for each atom, as explicitly verified in our computations.

In the presence of Jahn-Teller effect both the coordinates show a neat variation, not too sensitive to the cutoff. Note that breathing relaxation turns out to be inward, in contrast to earlier assumptions [49,63], however in agreement with some recent accurate computations [39]. It is noteworthy that using a 54 atoms cell (as used in ref. [49]) we also get Q_b with the positive sign, but Q_p with the negative one, which is almost certainly wrong.

Finally, we performed a series of calculations over the so called "split vacancy" configuration, an important issue since it is believed to represent the saddle point in the migration path. In fact, the difference in the values of energy between the split- and nonsplit-vacancy systems should be the activation energy for the migration via vacancy mechanism, and here is referred as E_{migr}.

When relaxing the system in this configuration care must be used to avoid that the atom on the saddle rolls downwards. We achieved this goal by imposing to the gap level wavefunction to have the degeneracy appropriate to the saddle point symmetry D_{3d}; this allows distortions only on the plane perpendicular to the migration path, which anyway did not appear.

As immediately seen, not only it shows a good convergence (see also Tab. 3.1.5, next section), but the agreement with experimental data [28] is

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

Table 3.5:
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 – 63+vacancy supercell. $E_c = 10\,Ry$ and $l_c = 1$. Only the Γ point was used.

quite good. This makes us quite trusting in the reliability of jump calculation.

3.1.4 Convergence in γ

The results of the test over γ in the ideal vacancy configuration are reported in Table 3.1.4. As immediately seen, at least for this cutoff ($10Ry$) the use of $\gamma = 1$ yields a very small error ($\sim 0.05\ eV$), while greater values offer an exceedingly good approximation, below the limits of accuracy of the method. The use of a greater E_c, as previously said, surely improves 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 letting the neutral vacancy relax with $\gamma = 1.5$, letting the other parameters as above. Even in this case the differences are very small.

3.1.5 Effect of the d-Component of the Pseudopotential

We have not included terms of order greater than 2 in the expansion of the pseudopotential (the only insertion of the $l = 2$ term causes a 30% increase in the computer time). Actually, the Silicon-Silicon bond is an s-p hybrid in character, and the range of these terms is small: so the chemical importance of their inclusion it is not likely to be substantial.

Table 3.1.5 shows that the introduction of the $l = 2$ term ($E_c = 10Ry$

l_c	1	2
E_{form}	3.42	3.54
E_{level}	1.01	0.92
E_{form}	2.96	3.39
E_{migr}	0.33	0.33
ΔQ_b	−0.53	−0.41
ΔQ_p	+0.45	+0.44
E_{lev}	0.66	0.64

Table 3.6:
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 − 63+vacancy supercell. $E_c = 10\,Ry$ and $\gamma = 4$. Only the Γ point was used.

was used) makes rise E_{form} and substantially reduces the gain due to the ionic relaxation. Note that the ionic positions (monitored by the Q's) are almost the same.

The variation in energy then seems to have a moderate influence on the ionic motion and, as seen by the fact that the migration energy is not changed by changing l_c, may probably be neglected while computing the dynamical properties of the vacancy.

As a conclusion, we conclude that the choice of a 64 atoms cell, $E_c = 10\ Ry$, Γ point, $\gamma = 1.5$, l_c=1 is a sensible one. Let us estimate the error expected in the formation energy: the cell size lead us to underestimate it of $\sim 0.2\ eV$, the Γ point of $\sim 0.3\ eV$, l_c of $\sim 0.4\ eV$, while $E_c = 10\ Ry$ leads to an overestimate of $\sim 0.2\ eV$. The algebraic sum yields $\sim 0.7\ eV$ over $\sim 3.5\ eV$, an error not greater than 20%.

For the migration energy the estimate is even better: the underestimate of E_m due to the cell size is $\sim 0.04\ eV$, while setting the cutoff to 10 Ry gives an overestimate of $\sim 0.06\ eV$, the other parameters giving negligible effects. The converged value then should be $\sim 0.31\ eV$, with an $\sim 0.02\ eV$ error.

3.2 The Jump Simulation

For the first dynamical simulation, following the suggestions reported in the previous paragraph, we choose a 63 atoms + vacancy cell, only the Γ point to sample its BZ, and $l_c = 1$; the cutoff was set to 12 Ry, since this choice implies a very small increase in the use of computer resources respect to $E_c = 10$ Ry. The parameters were chosen as follows: $\Delta t = 7a.\,u. = 7 \times 2.42\ 10^{-17}\ s = 1.69\ 10^{-16}\ s$; $\mu = 400\ m_0$, where m_0 is the true electron mass. The lattice parameter, and hence the volume, has been kept fixed to the empirical value of 10.263 $a.\,u.$.

An accurate minimization of the electronic energy has been first performed, using a steepest descent procedure and the algorithm proposed in ref. [57]. A randomization of the initial wavefunctions was done to make us sure that the found ψ's were the true ground states ones. Even the ionic coordinates has been slightly randomized in order to put the atoms in an off-equilibrium situation as the dynamics starts.

After leaving the system go, its temperature was monitored to be about 35 $^\circ K$. To raise it, we introduced a thermostat by rescaling the ionic velocities [58] in such a way that the average kinetic energy corresponds to a temperature of 500 $^\circ K$; this was done each time the temperature went out a 50 $^\circ K$ range about the desired value. After 100 time steps of this treatment 200 others followed in which the system was left to freely oscillate. Then this procedure was repeated for $T = 1000\ ^\circ K$ and $1500\ ^\circ K$. The final temperature, averaged over about 500 steps after reaching the final stage, was seen to be about $T = 1300\ ^\circ K$. This period was used even to allow the system to get as close as possible to the thermal equilibrium.

The figures 3.1, 3.2 respectively contain the square displacements of the 63 ions from the initial positions and the distance of the first four neighbours of the vacancy from its lattice site, for the whole run (10000 steps= 1.7 ps). If a jump into a new site (most likely, in the vacant site) takes place by one of the atoms it is immediately seen by the track it leaves in the upper part of the graphs. In particular, the figure 3.2 is intended to discriminate between the displacements of the vacancy neighbours towards or away from the vacancy; they are expected to have a big amplitude (soft modes [59], due to the lack of some bonds).

As immediately seen, in fact, the four first neighbours perform big oscillations around the equilibrium position, and one of them jumped into the

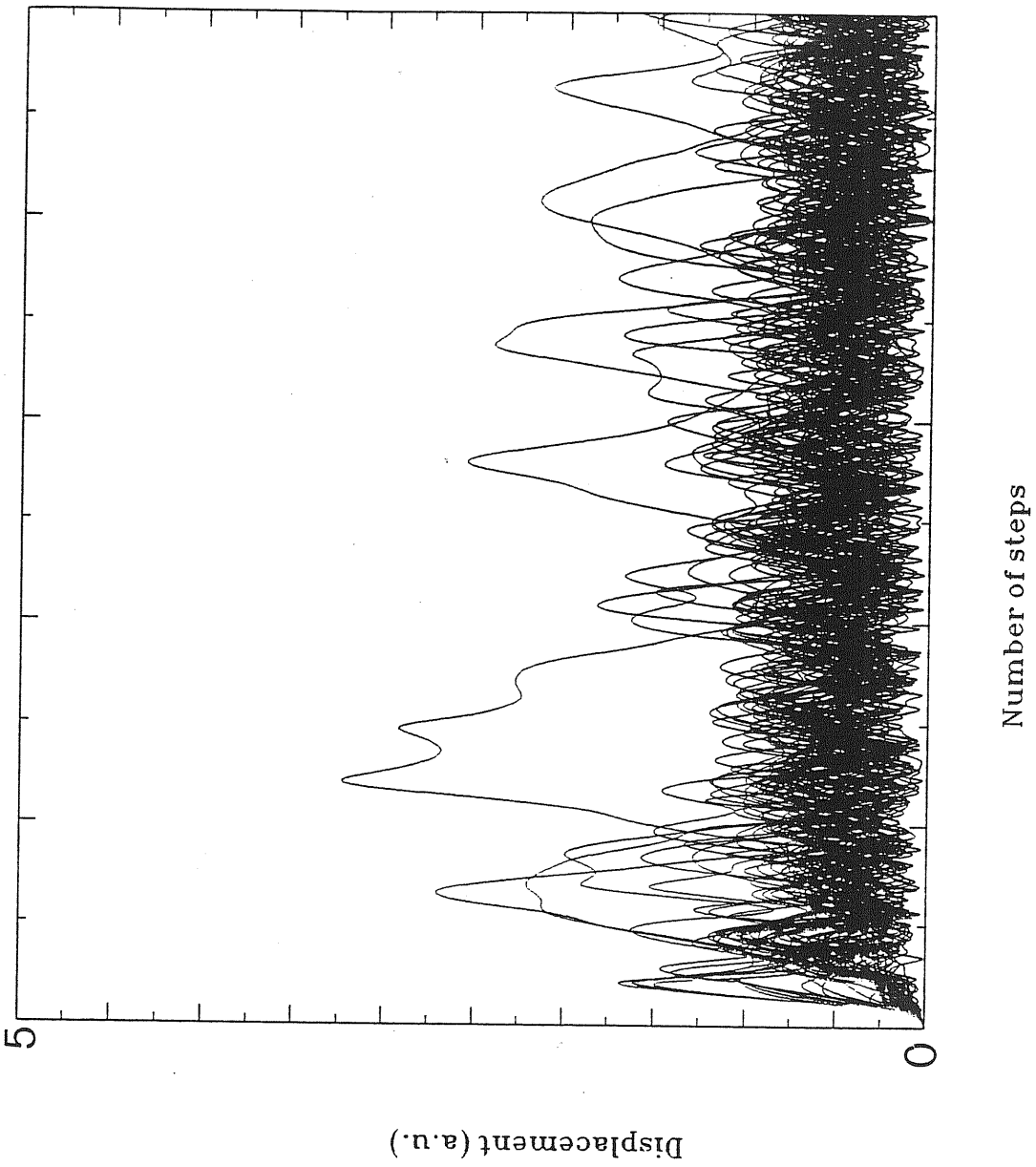

Figure 3.1:
$|\Delta \mathbf{r}|$ as a function of the number of steps

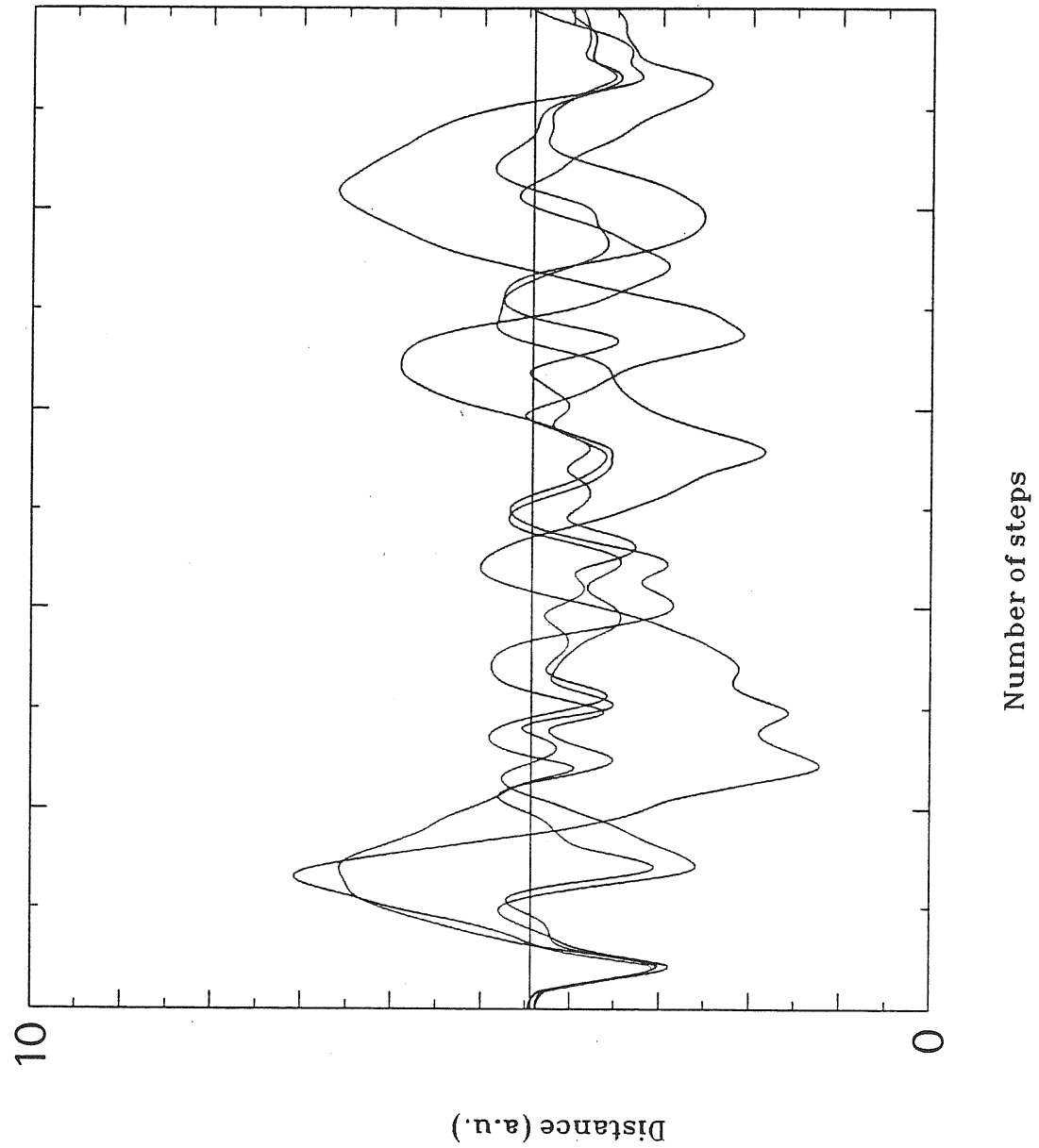

Figure 3.2:
$|r - r_{vac}|$ of the first shell as a function of the number of steps (see text)

vacancy in approximately 1900 steps after the end of heating. The distance from the old vacancy site is not zero because the inward relaxation towards the news vacancy, that is in the site just emptied.

After about one oscillation in the new site, it came back at the original position and no more hops are observed. Some trials instead can be seen, at about 5400 (the same atom just jumped) and 6800 steps, but none of them was successful, in spite of the fact that they have reached or maybe overcame the energy watershed lying between their original position and the vacant site; this distance is marked by the horizontal line. This shows that purely configurational arguments, i. e. based on neglecting the detail of the dynamics, such as those used by the rate theory, have to be taken with some care. Up to now nevertheless we cannot draw any firm conclusion under this respect.

The fact that the only jump observed happened at the beginning of the run leads to the suspect that it is due to the choice of the initial condition, and to an imperfect equilibration of the system. More statistics is needed to solve the problem.

Figure 3.3 shows the behaviour of the system temperature during the run. Of course the thermodynamical limit is far and the what we call "temperature", which actually is the instantaneous kinetic energy of the ionic system divided by $3Nk_b$, undergoes large fluctuations of amplitude $\sim 200\ ^\circ K$.

In addition, a systematical shift is observed. In fact, the final temperature results to be approximately $1180\ ^\circ K$, $120\ ^\circ K$ lower than at the beginning. This is the flag, rather than of the non-conservative nature of the discretized Newton equations which should have a minor role for these short times and high ionic masses, of the non perfect adiabaticity of the motion, which drains energy from the ions (hot) to give it to the electrons (cold).

To monitor in a more exact way this effect, we report in fig. 3.4 the values assumed by the total energy of the system minus the electron kinetic energy E_{cons}. In the case of perfectly adiabatic motion the curve should be a straight horizontal line, but we observe that apart small, short time scale oscillations a neat trend towards lower values. Evaluating the total energy loss of the ions and reporting it into temperature unit we have $\Delta E_{cons} \simeq 140\ ^\circ K$. Therefore the temperature lowering is almost entirely due to the gradual drift away from the BO surface.

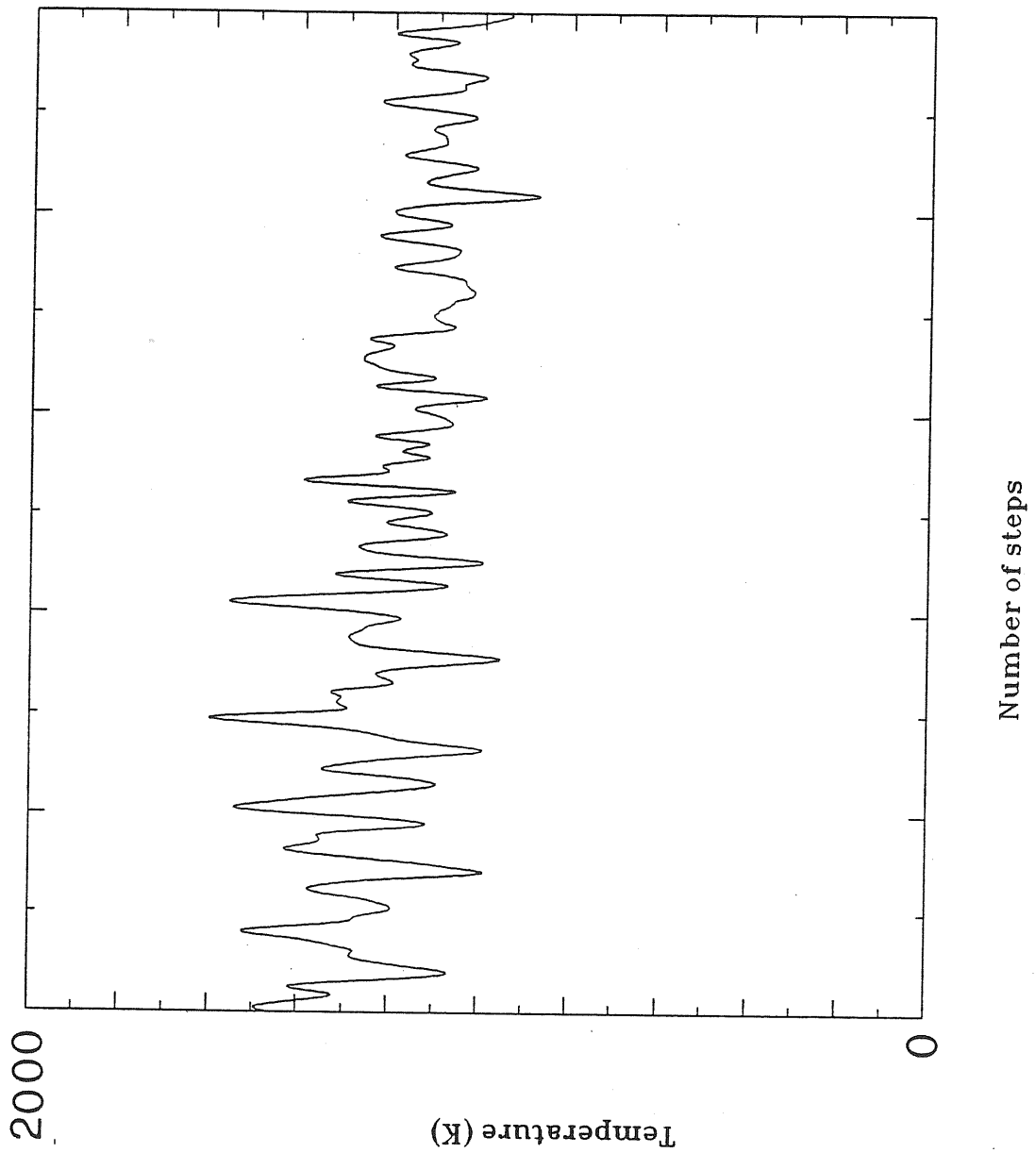

Figure 3.3:
Temperature as a function of the number of steps

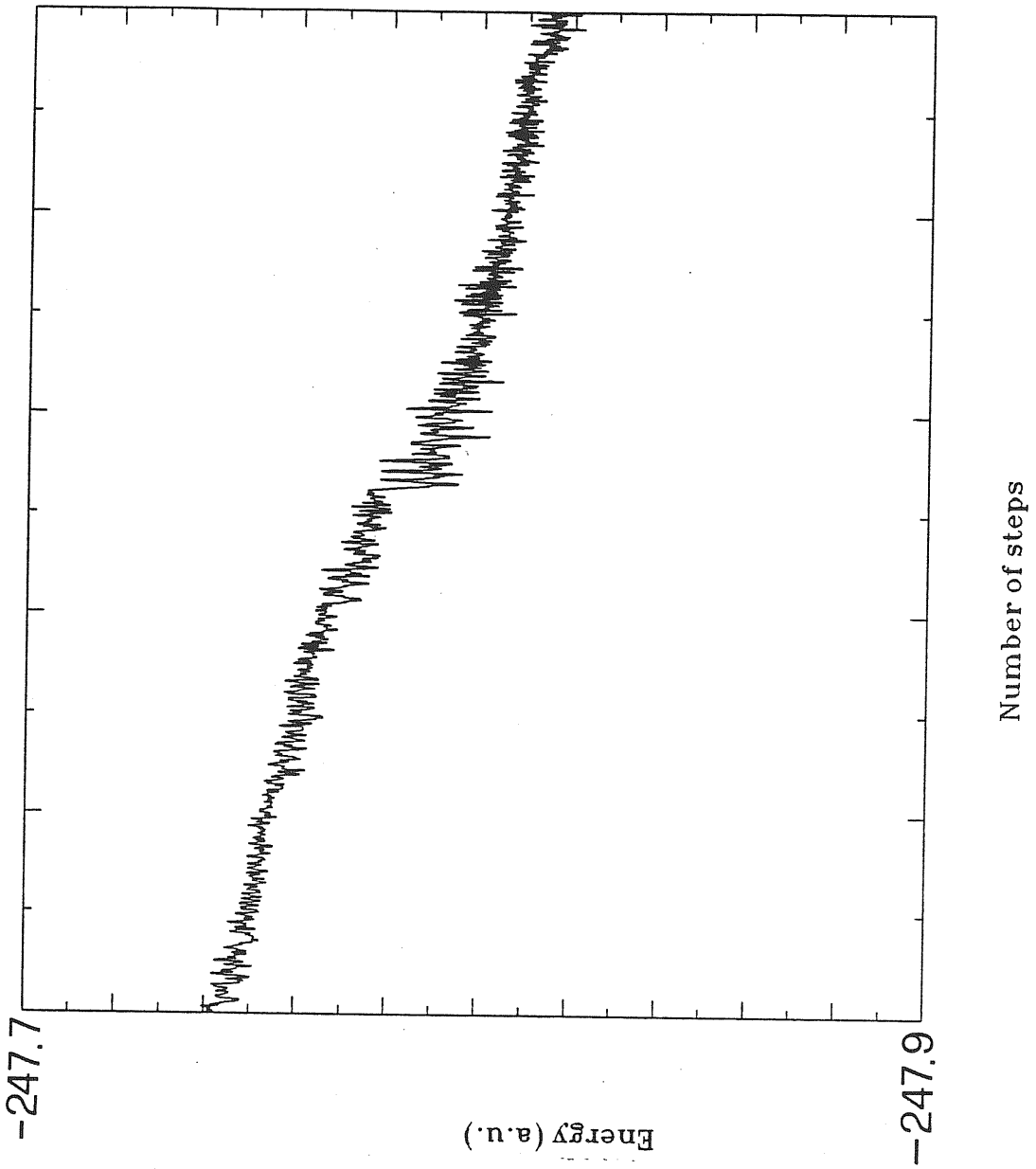

Figure 3.4:
E_{cons} as a function of the number of steps

The relatively high value of this effect is probably due to the presence of partially filled gap states which have among them an electron excitation energy much lesser than in a bulk semiconductor.

The possible remedy for this could be the introduction of a continuous thermostatting of the ionic system, such as the Nosé method [58], to preserve the original temperature, and of a periodic minimization of the electronic energy to regain the BO surface. For the run shown this latter operation would be sufficient to be performed only each 5000-10000 steps.

Chapter 4

Conclusions

The work here exposed has shown the feasibility of the MD-DF study of the Silicon vacancy. Its use in principle permits to overcome all the shortcomings affecting the other methods used, being both a first-principle (and, as such, capable to give the right ionic potential in any configuration) and a dynamical, finite temperature method (so exactly taking into account all the dynamical and thermodynamical effects).

The possible obstacles rise mainly from the necessity of small MD cells; the inclusion of the d-component of the pseudopotential does not seem to play a major role in the dynamics and even if it would result to be so it can be included without an untenable effort. The other approximations do not sensibly affect the results. Our statical computations show that the formation energy is estimated within a 20% error, an acceptable value.

The dynamical simulation performed has shown that we can proceed without any trouble at least up to $\sim 2\ ps$, and longer runs can be achieved by using the simple methods indicated at the end of the last section. Even if the meaning of the observed jump is not completely clear, leaving us doubtful that only a peculiar dynamical circumstance has led us to see it, the number of attempts seen indicates that a reasonable number of atomic jumps probably can be observed.

The next step in our research is, to have the maximum hopping probability, to do a new run at higher temperature; since the MD simulated materials have an higher melting point than the real ones (for Silicon is 1685 $^\circ K$) a temperature as high as 1800 $^\circ K$ can be reached. Since the migration barrier is about 0.3 eV high an increase in the jump rate of a

factor 3, respect to the run done, is expected.

Afterwards, the computations of the G-function and so the vacancy concentration will have to be done: the chosen method is that exposed in sect. 2.2 which was explicitly conceived for joint use with the MD-DF scheme and so it should be by far the best one for our purposes.

As a longer term project the study of other migration mechanisms is foreseen. An intriguing one is the Pandey concerted-exchange mechanism [3].

As the outlined job will be terminated, a cleaner theoretical framework on the Silicon self-diffusion will hopefully have been built up, and will allow to more deeply understand this controversial question.

Bibliography

[1] J. Van Vechten, *Phys. Rev. B* **33**, 2674 (1986)

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

[3] K. C. Pandey, *Phys. Rev. Lett.* **57**, 2287 (1986)

[4] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985)

[5] 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, Singapore (1987)

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

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

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

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

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

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

[12] 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., New York (1975)

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

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

[15] B. J. Masters and J. M. Fairfield, *Appl. Phys. Lett.* **8**, 280 (1966)

[16] J. M. Fairfield and B. J. Masters, *Journ. of Appl. Phys.* **38**, 3148 (1967)

[17] R. N. Ghoshtagore, *Phys. Rev. Lett.* **16**, 890 (1966)

[18] J. Hirvonen and A. Anttila, *Appl. Phys. Lett.* **35**, 703 (1979)

[19] L. Kalinowski and R. Seguin, *Appl. Phys. Lett.* **35**, 211 (1979)

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

[21] 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 Conf. Series **31**, London (1977)

[22] J. R. Sanders and P. S. Dobson, *Jou. Mater. Sci.* **9**, 1987 (1974)

[23] G. Hettich, H. Mehrer, and K. Maier, *Institute of Physics Conference Serial* **46**, 500 (1979)

[24] R. A. Swalin, in *Atomic Diffusion in Semiconductors*, edited by D. Shaw, Plenum Publishing Corp., New York (1973)

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

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

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

[28] G. D. Watkins, *Journal. of the Phys. Soc. of Japan* **18** *suppl.2*, 22 (1963)

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

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

[31] S. Dannefaer, P. Mascher, D. Kerr, *Phys. Rev. Lett.* **56**, 2195 (1986)

[32] 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, Orlando (1984)

[33] B. J. Masters, *Solid State Communications* **9**, 283 (1971)

[34] R. A. Swalin, *Jou. Phys. Chem. of Solids* **18**, 290 (1961)

[35] H. K. Benneman, *Phys. Rev. A* **137**, 1497 (1967)

[36] 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, Berlin, 1984

[37] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, *Phys. Rev. Lett.* **52**, 1814 (1984)

[38] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, *Phys. Rev. Lett.* **54**, 360 (1985)

[39] P. J. Kelly, R. Car, and S. T. Pantelides, in *Proceedings of the 14th International Conference on Defects in Semiconductors*, Paris 1986

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

[41] A. M. Stoneham, V. T. B. Torres, P. M. Masri, and H. R. Schober, *Phil. Mag. A* **58**, 93 (1988)

[42] F. H. Stillinger and T. A. Weber, *Phys. Rev. B* **31**, 5262 (1985)

[43] J. M. Ziman, *Electrons and Phonons*, Clarendon, Oxford (1960)

[44] P. C. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964)

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

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

[47] J. C. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981)

[48] D. M. Ceperley and B. K. Alder, *Phys. Rev. Lett.* **45**, 566 (1980)

[49] S. G. Louie, M. Schlüter, J. R. Chelikowski, and M. L. Cohen, *Phys. Rev. B* **13**, 1654 (1976)

[50] M. T. Yin and M. L. Cohen, *Phys. Rev. B* **26**, 5668 (1982)

[51] J. Ihm, A. Zunger, and M. L. Cohen, *Jou. of Phys. C* **12**, 4409 (1979)

[52] A. Baldereschi, *Phys. Rev. B* **7**, 5212 (1973)

[53] D. J. Chadi and M. L. Cohen, *Phys. Rev. B* **8**, 5747 (1973)

[54] D. R. Hamann, M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979)

[55] G. B. Bachelet, D. R. Hamann, and M. Schlüter, *Phys. Rev. B* **26**, 4199 (1982)

[56] L. Kleinman and D. M. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982)

[57] M. C. Payne, J. D. Joannopoulos, D. C. Allan, M. P. Teter, and D. H. Vanderbilt, *Phys. Rev. Lett.* **56**, 2656 (1986)

[58] S. Nosé, *Jou. Chem. Phys.* **81**, 511 (1984)

[59] M. Lannoo and G. Allan, *Phys. Rev. B* **33**, 8789 (1986)

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

[61] L. Kalinowski and R. Seguin, *Appl. Phys. Lett.* **36**, 171 (1980)

[62] M. Lannoo and G. Allan, *Phys. Rev. B* **25**, 4089 (1982)

[63] G. A. Baraff, E. O. Kane, and M. Schlüter, *Phys. Rev. B* **21**, 5662 (1980)