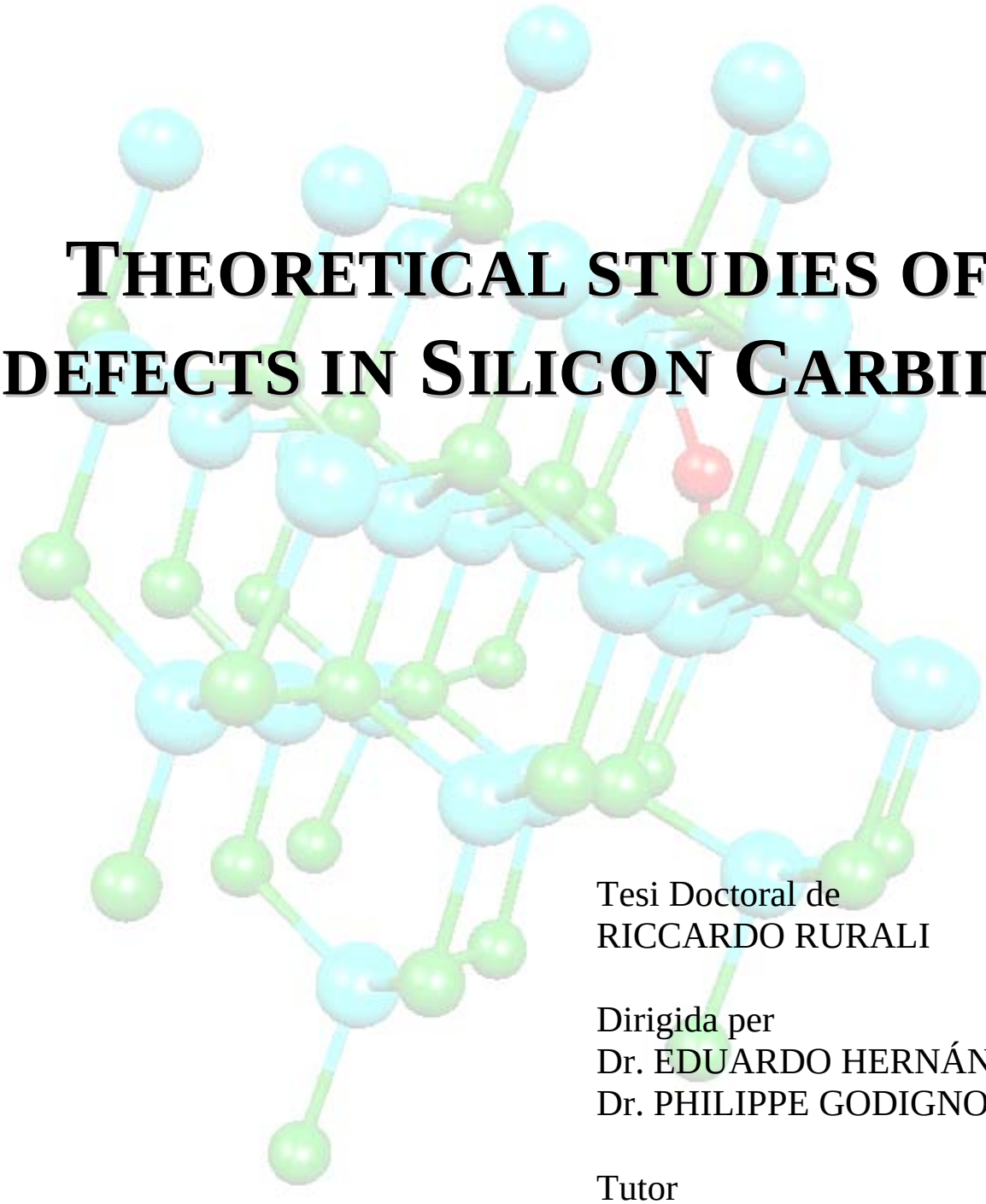
Universitat Autònoma de Barcelona

Facultat de Ciències
Departament de Física



THEORETICAL STUDIES OF DEFECTS IN SILICON CARBIDE

Tesi Doctoral de
RICCARDO RURALI

Dirigida per
Dr. EDUARDO HERNÁNDEZ
Dr. PHILIPPE GODIGNON

Tutor
Prof. JORDI PASCUAL GAINZA

El Professor Jordi Pascual i Gainza, Catedràtic de la Facultat de Ciències de la UAB i Tutor del present treball de recerca

CERTIFICA

Que aquesta memòria ha esta realitzada sota la direcció dels Drs. Eduardo Hernández i Philippe Godignon, Científics Titulars del CSIC, constituint el treball de Tesis del programa de Tercer Cicle de Ciències de Matreials.

Bellaterra, 6 de octubre de 2003.

Dr. Eduardo Hernández

Dr. Philippe Godignon

Prof. Jordi Pascual i Gainza

Considerate la vostra semenza:
fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza

Inferno XXVI

To José Luis

Considerate la vostra semenza:
fatti non foste a viver come bruti
ma per seguir virtute e canoscenza.

Inferno, XXVI

Acknowledgements

I would like to thank here all those persons who supported, helped and encouraged me during the development of this work.

Eduardo for greatly impulsing my scientific formation, for his permanent availability to discuss and clarify doubts and for his friendship.

Pablo for his continuous and patient assistance with SIESTA and for his high criticism, always showing me that what it looked to me so easy was not actually so.

José Luis for enlightening revelation about semiconductor physics at the coffee breaks and all the guys at the Electronic Structure Group at the ICMAB - CSIC.

Philippe for dragging me down to reality more than once with the typical question of a researcher in device fabrication: *What is this simulation useful for?*

Pepín for guiding me through the knots of EU networks.

Narcís as the reference experimentalist.

All the guys at the Power Device Group at the CNM - CSIC, especially Xavi and Pepe.

All the staff at CNM, particularly Montse, Quima and Lourdes.

Jordi Pascual for being a patient tutor at the Universitat Autònoma de Barcelona.

The EU that financed my PhD through ATOMCAD, a Research, Training and Mobility Network.

Per Hyldgaard, Elwira Wachovicz and Eleni Ziambaras at Chalmers University of Technology in Göteborg. Uwe Gerstmann at the University of Paderborn. Peter Déak, Balint Aradi and Adam Gali at Budapest University of Technology. Jerémie Baerdemaeker and Charles Dauwe at Ghent University. Giovanni Consolati at the Politecnico di Milano. Grzegorz Karwasz at the Universitá di Trento. Gilberto Teobaldi at the Universitá di Bologna. Marcin Kaczmarski at the ICMAB - CSIC.

The Centre de Supercomputació de Catalunya (CESCA) and the Centre Europeu per el Paral·lelisme de Barcelona (CEPBA) for the use of their computational facilities.

Mercè and Celia at the Departement de Física of the UAB.

Finally, I would like to thank Mònica, my friends and my parents for continuous encouragement.

Barcelona, October 7, 2003

Contents

Chapter 1

Introduction

Pluralitas non est ponenda sine necessitate
William of Ockham

Front-edge applications in power electronics require every day more semiconductor devices which are capable to offer high performances under extreme conditions, like very high voltages, high temperature or radiation contaminated environments.

Macroscopic experimental characterisations have been for many years the primary tools to understand the properties of materials. In spite of this, there are many evidences that some properties are intimately related to atomic scale processes, where the quantum mechanical nature of condensed matter shows up. However, the nanoscale is difficult to access experimentally and therefore atomic scale computer modelling of materials has become a field of growing interest for material design and it is especially appealing for power electronics.

In the latest decades, Silicon Carbide (SiC) has confirmed to be the most promising and realistic among the so-called wide band-gap semiconductors, which represent an attractive alternative to Silicon in a wide range of applications. Few things about this material are known, from the fundamental viewpoint and many elementary mechanisms - relevant for device fabrication - have to be fully understood yet. Until this task will not be accomplished, Silicon Carbide will not be a true competitor with respect to Silicon where long-standing know-how makes every technological step almost totally

controlled. In this sense, computer modelling is a privileged tool to gain a thorough understanding of those processes that take place at the atomic scale and that determine many of material properties.

In this work we will address the topic of theoretical characterisation of impurities in SiC by means of first-principles electronic structure calculations. Defects in a semiconductor is a primary importance issue, as they are at a time one of the element on which device design relies on (n-type and p-type dopants) and one of the main cause of electrical properties degradation (point defects, charge traps, stacking faults, ...). We will focus extensively on one hand on defect diffusion, both for what regards dopants (Chapter 6) and intrinsic defects (Chapter 7). This is a topic of the highest technological relevance, as the understanding of the atomic scale mechanisms that rule diffusion and the ability to control them could be turned into advantage from the viewpoint of optimal device design. On the other hand, we will discuss in some details the problems related to high-dose doping, dopant clustering and dopant passivation in the case of n-type impurities (Chapter 8).

We also talk briefly about how theoretical modelling can contribute in a valuable way to the study of bulk properties of the materials (Chapter 9) and in particular we will discuss some recently developed techniques to calculate the free energy of the different phases of a system, in order to evaluate the conditions of temperature and pressure for phase transitions.

Chapter 2

Silicon Carbide: a wide band-gap material for microelectronics

At present there is a considerable effort in the semiconductor industry to overcome the difficulties and limitations inherent in Si-based devices across the spectrum of industrial applications. Wide band-gap semiconductors, such as SiC, GaN and diamond offer the potential to overcome the difficulties of temperature and high voltage blocking that bedevil Si. SiC is proving to be the most attractive alternative at medium-term, offering significant advantages at both the high temperature and high voltage limits, whilst having the benefit of a tractable materials technology. In addition, present advances in the quality of samples and wafer size are on the brink of turning SiC from a long standing promise into a realistic material for device production. Its main advantages, in comparison to Si, are its high thermal conductivity, its high field breakdown strength, and a high saturation velocity, which make it ideally suited for high-temperature [1, 2, 3, 4, 5], -power, -frequency and radiation environments [6, 7, 8]. In this chapter we will outline the main characteristics of this material, with an emphasis on its crystalline structure and electronic properties which are ultimately responsible of the high interest that SiC has as a material for microelectronics.

2.1 Brief history

SiC does not exist in nature under the crystalline form and this contributed to the delay with which it called attention on its properties. At first, the main interest mankind saw in SiC consisted in the possibility to replace diamond

with it in a range of cutting and abrasive industrial tools. Shortly after, attention was moved to its electrical and electronic structure properties and it was realised that in that class of applications it was no more a second choice with respect to diamond, but under many aspects was a superior material. William Shockley, inventor of the bipolar junction transistor in 1947, recognised the potential of SiC, foreseeing the forthcoming success of this material, as shown in this quotation, dating back to 1959 [9]:

> *Now the big question is: how is the problem of high temperature going to be solved? What are the horses to put one's money on? [...] One approach is the logical sequence we see here: Ge, Si, SiC, C in that sequence[...]*
> *The SiC situation suffers from the very same thing that makes it so good. The bond is very strong and so all the processes go on at very high temperature [...]*
> *Another aspect of the SiC situation is similar to past situations in the semiconductor field. The lesson is that one should not give up too soon and one would not always look for gold at the ends of new rainbows [...]*

Nevertheless, for many years the real bottleneck had been the difficulty in growing good quality substrates. After the first pioneering results achieved by the Lely method [10], it was only in the late 70's, with the development of the physical vapour transport (PVT), also known as the modified-Lely method [11], that industrial quality mono-crystalline SiC started to be grown in a reliable and reproducible way. This stresses how the reasons of SiC success not only lie in its superior mechanical and electrical properties, but also in the ability of to develop and standardise all those industrial processes needed by the microelectronics industry aiming for device fabrication. This consideration, besides the substrate growth, applies also to the availability of a native oxide. Silicon dioxide - SiO_2 - has an excellent lattice parameter match with SiC and its long history of technological treatment, due to its use in the Si industry, gave a fundamental contribution to development of SiC industry. Even if the quality of the SiC-SiO_2 interface and the isolation property of the oxide layer still deserve a lot of work to be really competitive with Si, the availability of an industrial quality oxide was determinant to impulse the starting of device fabrication.

Nevertheless, together with all these physical and electrical considerations, it has to be taken into account that for many aspects the investments in SiC research and development have been for many years related with the degree of maturity of Si technology, so only when the exponential rate of growth of applications where Si could be used without significant problems

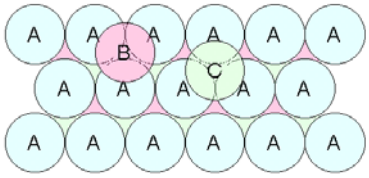


Figure 2.1: Different ways of piling up Si-C bilayers. If A is the first layer and B is the second layer, there are three possibilities for the third layer: it can be piled up in A, B or C stacking position, giving rise to different polytypes.

slowed down, the scientific interest in SiC was followed by an analogous industrial attractiveness.

2.2 Crystalline structure

Silicon carbide has the same structure as bulk silicon or diamond, but its *base* is made up of two different types of atoms: one Si and one C. The atoms are tetrahedrally co-ordinated by means of sp^3 hybridised bonds in such a way that every Si is bonded with four first-neighbour C and every C is bonded with four first-neighbour Si. The first-neighbour bond length is around 1.89 Å (longer than in diamond and shorter than in Si) and the bond angle is > 109 ° like for all the compounds with the same crystalline structure.

As we will see below, the multi-species nature, far from being a minor fact, implies many differences in the analogy with the corresponding one species compounds, like Si for instance. Not only the general electronic properties, as expected, but many structural and mechanical features are markedly different.

2.2.1 Polytypism

SiC has a hexagonal lattice where hexagonal planes of Si are alternated with hexagonal planes of C. To reproduce the correct periodicity of the base and to allow the tetrahedral bonding required by the sp^3 hybridisation, each plane must be properly shifted with respect to the plane below. The way successive planes are piled up is not unique, provided some constraints are respected, as Fig. (2.1) shows. Therefore different stacking sequences lead to different *polytypes*. SiC is often said to be polymorphic or to exhibit a one-dimensional polytypism.

The polytypes can be classified according to three basic crystallographic categories: cubic (C), hexagonal (H) and rhomboidal (R) for a total of more than 200 types [12]. Fig. (2.2) illustrates how different ways of piling up bilayers give rise to cubic or hexagonal polytypes.

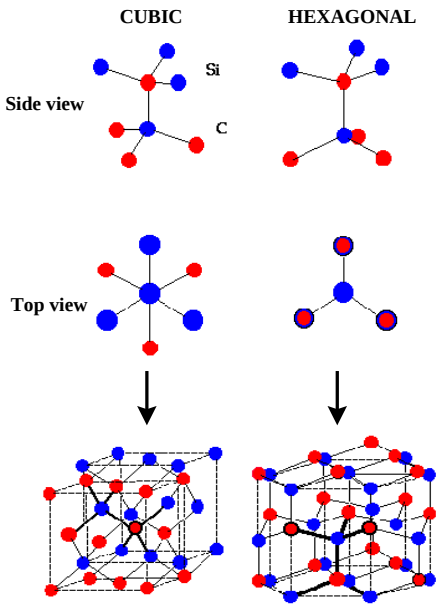


Figure 2.2: Side and top view of Si-C bilayer piling up in cubic and hexagonal SiC.

Cubic SiC

Cubic SiC has only one possible polytype, often indicated as 3C-SiC or β-SiC, and it is the most simple. Provided that each Si-C bilayer can be oriented into only three possible directions with respect to the lattice while the tetrahedral bonding is maintained and denoting them arbitrarily A, B and C [see Fig. (2.1)], the stacking sequence of cubic SiC is ABCABCABC... giving a cubic zinc blende structure. 3C-SiC has the smallest band-gap (~ 2.2 eV) and one of the highest electron mobilities (~ 1000 cm^2 V^{-1} s^{-1}). Notwithstanding the fact that 3C-SiC has been successfully grown in research environment, it is not currently available commercially in bulk form.

Cubic SiC possesses the highest symmetry and therefore it has also the smallest possible unit cell, made of eight atoms.

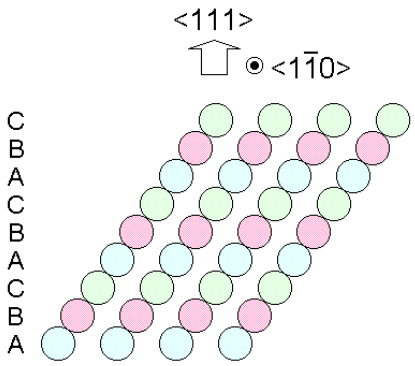


Figure 2.3: Stacking sequence of Si-C bilayer in 3C-SiC.

Hexagonal and rhomboidal SiC

When the stacking sequence is ABABAB... the symmetry is the purely hexagonal and the corresponding polytype is indicated as 2H-SiC.

All the other polytypes are a mixture of the cubic (zinc blende) and hexago-

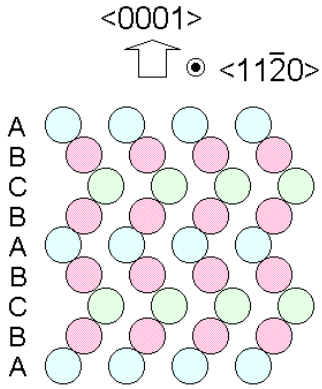


Figure 2.4: Stacking sequence of Si-C bilayer in 4H-SiC.

nal (wurtzite) bonding. 4H-SiC is made up of an equal number of hexagonal and cubic bonds, while 6H-SiC consists of two thirds of cubic bonds and one third of hexagonal bonds. It is noteworthy however that, despite the presence of cubic bonds, the overall symmetry is hexagonal for both these polytypes. All the hexagonal polytypes are denoted as α-SiC, but only 4H-SiC and 6H-SiC are available in bulk wafer form.

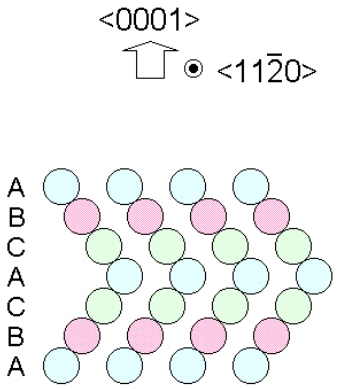


Figure 2.5: Stacking sequence of Si-C bilayer in 6H-SiC.

Property	3C-SiC	4H-SiC	6H-SiC
Band-gap (eV)	2.20	3.26	3.02
Electron mobility [$cm^2\ V^{-1}\ s^{-1}$]	1000	1000	400
Hole mobility [$cm^2\ V^{-1}\ s^{-1}$]	40	115	101
Saturation electron velocity [$cm\ s^{-1}$]	$2.5\ 10^7$	$2.7\ 10^7$	$2.7\ 10^7$

Table 2.1: Comparison of some relevant semiconductor proprieties among some of the most important SiC polytypes.

Among the few hexagonal polytypes here considered, 6H-SiC has the smaller degree of symmetry [as can be seen by its stacking sequence shown in Fig. (2.5)] and therefore has the largest unit cell. Polytypes with rhomboidal symmetry have also been found and 15R-SiC has been grown.

$\sim$

The differences among the polytypes go far deeper than what could be expected at a first sight. The wave function in each of them is different and so the band-gap energy, the carrier mobility and the breakdown field are affected, sometimes considerably, by the polytypism. In Table (2.1) some relevant parameters are given for the most typical polytypes.

As we have already anticipated, the polymorphism of SiC cannot even be neglected from the technological point of view and, on the contrary, the industrial development of SiC processing is profoundly indebted to the existence of many phases of the material. The most straightforward analogy with Si or diamond lattice, cubic SiC, is particularly difficult to grow and all

Property	Si	GaAs	GaP	3C-SiC	Diamond
Band-gap [eV]	1.10	1.40	2.30	2.20	5.50
Max. operating temp. [K]	600	760	1250	1200	1400
Melting point [K]	1690	1510	1740	2100	phase change
Physical stability	good	fair	fair	excellent	very good
Electron mobility [$cm^2\ V^{-1}\ s^{-1}$]	1400	8500	350	1000	2200
Hole mobility [$cm^2\ V^{-1}\ s^{-1}$]	600	400	100	40	1600
Breakdown voltage [$10^6 V\ cm^{-1}$]	0.3	0.4	-	4	10
Thermal conduct. [$W\ cm^{-1}\ K$]	1.5	0.5	0.8	5	20
Dielectric constant	11.8	12.8	11.1	9.7	5.5

Table 2.2: Comparison of some relevant properties between Si and some of the most important wide band-gap semiconductors.

cutting-edge devices are nowadays fabricated from 4H or 6H-SiC. This aspect is extremely important if we think that industrial tractability has been, in the end, the dominant reason to establish SiC as the most promising of wide band-gap semiconductors.

2.3 Electronic properties

One of the most interesting properties of a single-crystal of SiC that makes it extremely suitable for high temperature applications is its wide energy band-gap. Depending on the polytype, it can vary from ~ 2.2 eV to ~ 3.3 eV [see Tab. (2.1)]. The cubic phase, the simplest and most symmetric one which closely resembles Si, has a band-gap of 2.2 eV, while the two polytypes that nowadays are most currently used in industrial applications, 4H and 6H-SiC, two of the many possible hexagonal structures, have a band-gap of 3.26 eV and 3.02 eV respectively.

The band-gap width of a semiconductor is a crucial magnitude because it fixes the maximum operating temperature of the material. The maximum useful operating temperature is reached when the number of intrinsic carriers excited across the band-gap approaches the number of extrinsic carriers, those purposely added for doping the semiconductor. A wide band-gap, then, requires higher temperatures to achieve intrinsic conduction. In the case of SiC, for typical impurity doping levels they can easily approach 1200-1500 K: more than double than in the case of Si. These considerations of course hold only provided that the material is chemically stable below that temperature, i.e. it does not melt or undergo any other kind of phase transition.

This allows to understand better the outstanding performance of SiC with respect to its main competitors as wide band-gap semiconductors. GaP and diamond would have a comparable or even higher (in the case diamond) operating temperature, but they have a lower chemical stability or they undergo phase change that would make much more critical their employment in device fabrication. On the contrary, SiC is a very stable ceramic material up to 2050 K and therefore is a much more suitable candidate for industrial applications. Moreover, a side effect of a very high operating temperature is the immediate payoff of an improved reliability when operated at lower temperatures, which is an important condition in aerospace applications or in nuclear energy engineering.

A high breakdown field and a high thermal conductivity concur in making SiC definitely stand out among other wide band-gap semiconductors. A power device is ideally expected to block high voltages in the blocking mode and conduct high currents in the conduction mode. The maximum blocking voltage is roughly proportional to the product of the breakdown field and the depletion layer thickness W:

$$V_{max} = \frac{E_{max} W}{2} \tag{2.1}$$

where

$$W \propto \frac{1}{\sqrt{N_d}} \tag{2.2}$$

where N_d is the doping dose. SiC has a breakdown field E_{max}which is approximately one order of magnitude larger than in Si, therefore it can sustain the same voltage with a much smaller layer thickness [Eq. (2.1)], allowing to dope more heavily [Eq. (2.2)] and allowing so a higher operation temperature (see above). The higher thermal conductivity, on the other side, permits a more efficient thermal dissipation, which is a typically critical issue in power devices.

2.4 Doping

Doping for device fabrication is accomplished by epitaxially controlled doping and ion implantation. The temperature that would be required for doping via diffusion would be too high (> 1800 °C) and makes this technique almost useless. Either Al or B are currently used for p-type doping. Al is preferred because it has a shallower ionisation level of only 257 meV compared with 300 meV of B, but on the other side B is almost three times lighter, so it is preferred when ion implantation is the doping technique, because a less

extended damaged region will be created. N has been the most often used n-type dopant for quite a long time, due to its high solubility in SiC and an ionisation level of only 54 meV (in 3C-SiC), but P has been gaining increasing importance and recently it is chosen as often as N. A peculiar feature of these two dopants is that it has been shown that they prefer opposite lattice sites (N is more stable at C site and P at Si site) and this fact seems to open interesting opportunities to investigate the possible advantages of co-doping.

Epitaxial controlled doping is based on the so-called *site-competition epitaxy* [13] growth of SiC on bulk SiC wafers. The model relies upon N and C competition for a C site and Al and Si competition for a Si site in the SiC lattice and the technological process works on adjusting the Si/C gas flow rate in the growth reactor in order to control the amount of substitutional dopant incorporated in the lattice. Varying the Si and C gas flows, doses ranging from degenerately doped ($> 1 \times 10^{19}\text{cm}^{-3}$) to lightly doped ($1 \times 10^{14}\text{cm}^{-3}$) have been achieved. Epitaxial controlled doping is well developed when performed on Si-face substrates, but more work is needed to accomplish the same performance on C-face substrates, especially because the growth mechanisms have not yet been fully understood. This technique has also been used with B and P on Si-face substrates. The obvious limit of such a way of doping is that the grown layer will have the same uniform doping dose. The concentration could be possibly varied during the process, but the gradient slope would be limited by the response time of the gas flow injectors and it would be in any case absolutely unsuitable to create localised, sharp doped layers with the sub-micronic thickness tolerances required for pn-junctions. Therefore, the only reliable technique to realise localised and selective doping in SiC is ion implantation [14, 15, 16, 17].

Ion implantation consists of focusing a highly energetic ion beam on the substrate. The accelerated ions collide with the host lattice producing cascade displacements, in such a way that they can replace the hottest atoms at the highly stable lattice sites. In Sec. (2.2) we have noticed that SiC has a shorter bond length than Si. This does not only imply that the bond is harder to break, but also that SiC will offer a comparatively smaller interstitial volume to impurities, so that they will be much more stable. Although both materials are characterised by the same crystalline structure and the same type of covalent bond, annealing of damaged layers represents a much more serious problem in SiC. The higher mass density of SiC, for the reasons just discussed, requires ion energies 50 - 80 % higher than in Si to achieve comparable depths of implantation.

In contrast with Si, the ion distribution in SiC is scarcely influenced by the elevated implantation temperatures or extended annealing, due to the low diffusivity impurities have. Nevertheless, weird dependence of the

damaged layer characteristics on the implantation temperatures has been observed. Of course, over a certain dose threshold, the accumulation of radiation induced point defects degenerates in the formation of an amorphous layer, with the consequent loss of the high quality electronic properties of the material. The typical amorphisation energy density at room temperature is around $\sim 2 \times 10^{21}$keV cm^{-3}.

2.4.1 Selective doping

In all the main steps of technological processing necessary for device fabrication, SiC suffers of a disadvantage with respect to Si, because of the long industrial history of this pioneering semiconductor. The point is then being able to discriminate between those processing steps where Si exhibits a better tractability just because is a better know material - so the limit we have is *simply* that we do not know enough about SiC - and those other steps where SiC *does* inherently presents more problems than Si. In the first case a lot of work is needed, because all the basic knowledge already available for Si must be gained, but the latter is even more demanding, because it is not just a matter of filling a gap, but also of solving new problems.

Selective doping is an excellent example of this situation. For sure much more is known about Si doping, but it is also true that what is known is enough to let us say that SiC doping *is* a more difficult process. The ability to perform a selective doping as accurate as possible is the key for miniaturisation and high performance devices. In Si, doping may be performed by means of diffusion or ionic implantation. A quite clear trade-off between these techniques exists, because the more accurate ion implanted profiles imply an higher density of defects. In SiC, diffusion is not a practicable way of doping, because the diffusion constants of most of the electrically interesting impurities are much lower than in Si. This is the reason for which understanding what goes on at the microscopic level is of the highest interest and this is why we devote to this topic one of the main parts of this work [Chapter (6) and (7)].

The main critical secondary effect of ion implantation is damaging of the host lattice structure, deriving from the collision with the incoming, highly energetic ions. The resulting multiple atomic displacements give rise to a damaged layer, constituted to a large extent by point defects. Point defects modify the electrical properties of the crystal, introducing localised energy levels in the band-gap. Moreover, they can mediate dopant diffusion, at temperatures where ordinary thermal diffusion would be negligible, thus contributing to a degradation of the sharp edge of the doped region.

From these considerations it is clear that, if ion implantation is the only

way to perform selective doping in SiC, the possibility of recovering the lattice after the implantation is crucial.

In the already quoted speech of Shockley [9] this critical feature of SiC has been excellently outlined:

> *The SiC situation suffers from the very same thing that makes it so good. The bond is very strong and so all the processes go on at very high temperature [..]*

Then it is the rigidity of the bonds that causes many problems with SiC.

In other words, in SiC is more difficult to create defects, because bonds are stronger and therefore it is less easy to break them; at the same time, however, it is more difficult to anneal them out by thermal treatments: the damaged region will arrange itself in a configuration very far from the perfect lattice, but somewhat stable and will create again a local network made up of those strong bonds typical of SiC. Moreover, diffusion constants are very low for the Si and C atoms too, so this makes harder and harder the chance for the lattice to recover its original network. The trouble is that selective doping *does* require the creation of defects.

In other words, the physical properties of SiC lattice allow only the use of the doping technique that maximise the damage creates. Moreover, SiC it is more rigid that Si, making more difficult to recover such a damage.

Chapter 3

Solving the electronic structure problem

Since the advent of quantum mechanics, the possibility of solving computationally the equations that govern the microscopic interaction between electrons and nuclei was noticed, and this woke up the interest of many scientists, from solid state physicists to quantum chemists. Electrons are ultimately responsible for the binding of the atoms, and the structure of the electronic states in space and energy determines the physical and chemical properties of a material. Consequently, the ability to tackle the electronic structure calculation of a material would provide an outstanding insight on its more intimate nature. However, when considering the great complexity of the equations, it became clear that a solution was possible only for the simplest systems.

Our capability to solve the equations of quantum mechanics has increased tremendously in the last few decades, due to the continuous improvement of computer power together with the development of extremely powerful numerical methods. These advances made it possible to afford the exact solution of only one-electron systems [18], relevant to quantum chemistry, but the progress they introduce was too poor for typical solid state systems, where many atoms are necessary for a satisfactory description of the system. A breakthrough among methodological advances was represented by Density Functional Theory (DFT) [19, 20]. Since its first formulation, back in 1964, DFT encountered a growing interest and is nowadays a standard in computational materials science, especially in solid state physics. DFT provides a framework in which to understand the basic physics of many-electron systems without the complexity of the many-body wave function [21]. This can be

achieved by means of approximations to the exact theory [1], which have nevertheless been thoroughly tested since the development of the theory, giving us a large body of experience about which are the properties and the systems that can be accurately described.

The fundamental theory of condensed matter is well established and the electronic structure problem is very well defined. The interactions are electromagnetic and the equations are those of relativistic quantum mechanics. Even if the latter is the more complete solid state theory we have, for a very wide range of problems, relativistic effects can be neglected without significantly losing accuracy. The non-relativistic many-body Hamiltonian can be written in the following general form:

$$\hat{H} = -\sum_{I=1}^{N_n} \frac{\hbar^2}{2M_I}\nabla_I^2 \; - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m}\nabla_i^2 \; + \; \frac{e^2}{2}\sum_{I=1}^{N_n}\sum_{J\neq I}^{N_n} \frac{Z_I Z_J}{|R_I - R_J|} \; + \\ + \frac{e^2}{2}\sum_{i=1}^{N_e}\sum_{j\neq i}^{N_e} \frac{1}{|r_i - r_j|} \; - \; e^2 \sum_{I=1}^{N_n}\sum_{i=1}^{N_e} \frac{Z_I}{|R_I - r_i|} \qquad (3.1)$$

where R and M refer respectively to the position and the mass of the nuclei and r and m to the position and the mass of the electrons. All the properties of the system can be obtained by the solution of the Schrödinger equation [22]:

$$\hat{H}\Psi_i(r, R) = E_i \Psi_i(r, R) \qquad (3.2)$$

All the particles are coupled through the Coulomb interaction and moreover each particle statistic - may they be fermions, like the electrons, or either fermion or bosons like the nuclei - must be respected: the complexity of the problem is evident and not even a simple system like the He atom can be solved exactly. Hence the need for approximations.

3.1 Born-Oppenheimer approximation

The first approximation always introduced is the so-called Born-Oppenheimer approximation [23] or adiabatic approximation. It consists in assuming that electrons stay always in the same stationary state of the electronic Hamiltonian. If electrons move much faster than nuclei, we can assume that they instantaneously adapt to the adiabatically changing nuclear coordinates. This

[1] DFT provides an *exact* treatment if the E_{XC} functional, further on discussed, was known exactly; the approximation consists in relaying on approximated formulation of these functionals

allows the factorisation of the wave function as:

$$\Psi_{BO}(R,r,t) = \Theta(R,t)\Phi_n(R,r) \tag{3.3}$$

so, provided the system is in a stationary state, the electronic states are solutions of the time independent Schrödinger equation:

$$\hat{h}_e\Phi_n(R,r) = \epsilon_n(R)\Phi_n(R,r) \tag{3.4}$$

where $\hat{h}_e$ is the electronic Hamiltonian for a particular nuclear configuration:

$$\hat{h}_e = -\sum_{i=1}^{N_e}\frac{\hbar^2}{2m}\nabla_i^2 + \frac{e^2}{2}\sum_{i=1}^{N_e}\sum_{j\neq i}^{N_e}\frac{1}{|r_i - r_j|} - e^2\sum_{I=1}^{N_n}\sum_{i=1}^{N_e}\frac{Z_I}{|R_I - r_i|} \tag{3.5}$$

Once the electronic wave function is known for all the values of the nuclear positions, the substitution of the wave function in Eq. (3.1) in the Schrödinger equation for the whole system:

$$\frac{\partial\Psi_{BO}(R,r,t)}{\partial t} = \hat{H}\Psi_{BO}(R,r,t) \tag{3.6}$$

Although the basic adiabatic approximation retains the quantum description of both the electrons and the nuclei, the observation that the nuclear mass is rather large, suggests that it may be possible to introduce a further simplification, consisting in treating the nuclei as classical particles, while maintaining the quantum description of the electrons. This is particularly relevant when performing first principle molecular dynamics and one needs to integrate the atoms' equation of motion.

3.2 Total energy calculation in DFT

The main and most ambitious purpose of any solid state physics computational model is an estimation as accurate as possible of the total energy of the system.

By means of the Born-Oppenheimer approximation, we have simplified the problem and now the question is to solve Eq. (3.1). In this section we will discuss the basic ideas that represent the core of DFT, which allows to do that.

The most fundamental theorem on which DFT relies is that *the electron density $n(r)$ of a bound system of interacting electrons under some external potential $V(r)$ determines this potential uniquely.* An external potential $V(r)$ determines the ground state of an electronic density distribution. What DFT

relies on is that also the inverse is true. So, given an electronic density, there is a unique external potential $V(r)$ which can have determined it. In other words, once we know the electronic density of a system, we can deduce everything: the external potential, then the total energy and the atomic forces. Given its importance, let us rephrase it once more: since $n(r)$ determines the external potential $V(r)$ and the number of electrons N_e, it supplies the full Hamiltonian and any physical properties which we can extract from it.

Provided that the ground state energy is uniquely determined by $n(r)$, we can express this functional dependence as $E[n]$, separating the interaction of the electrons with the external potential $V(r)$ and the rest of the energy:

$$E[n] = \int V(r)n(r)dr \; + \; F[n], \tag{3.7}$$

where $F[n]$ is the sum of the kinetic energy of the interacting electrons and their mutual Coulomb interaction energy:

$$F[n] = T[n] + U[n] = \langle \Psi_0 | \hat{T} + \hat{V}_{ee} | \Psi_0 \rangle \tag{3.8}$$

Here we come to the second statement of Hohenberg and Kohn. If we now maintain $V(r)$ fixed in Eq. (3.7) and allow variations of $n(r)$, *the value of $n(r)$ which minimises the right hand side of Eq. (3.7) is the true ground state electron density of the system for that particular potential, and the energy is the ground state energy.* In other words, if we knew the functional form of F, we could get to the ground state by simply minimising the energy with respect to variations of the electron density, subject to the conditions $\int n(r)dr = N_e$ and $n(r) >= 0 \quad \forall r$. Kohn and Sham worked further in this direction, trying to figure out a satisfactory functional form of F and developed a formulation of DFT in terms of a self-consistent set of single-electron equations. The idea of self-consistency is analogous to that of standard Hartree-Fock theory, although here electronic correlation is included, and we will discuss it briefly afterwards [see Sec.(3.2.1)]. Nevertheless, before dwelling on self-consistency, the most urgent topic was finding out the functional form of F. At first, Kohn and Sham took into account the classical Coulomb-type potential created by the electronic distribution $n(r)$, defining a Hartree potential just like that of Hartree-Fock theory:

$$V_H(r) = \int \frac{n(r')}{|r - r'|} dr' \tag{3.9}$$

(the Coulomb potential of the nuclei is included in the external potential). Similarly, we can define the Hartree energy, as the purely classical energy deriving the interaction of the electronic density with itself.

$$E_H[n] = \int V_H(r)n(r)dr = \frac{1}{2} \int \frac{n(r)n(r')}{|r - r'|} dr dr' \tag{3.10}$$

At this point we should define the kinetic energy of the interacting system. The task is not straightforward, so what was proposed was to calculate the kinetic energy of a fictitious system of non-interacting electrons with the same density as the interacting one. Such a system is exactly described by Hartree-Fock theory and the corresponding many-body wave function is a Slater determinant made of one-electron states Ψ_i. This kinetic energy is very easy to calculate, but we must underline that it is not equal to the kinetic energy of the interacting, real system.

We can rewrite then the energy functional as:

$$E[n] = \int V(r)n(r)dr \; + \; T_s[n] \; + \; E_H[n] \; + \; E_{XC}[n]. \tag{3.11}$$

The first three terms are easily calculated, because they are the same ones we runs into in a normal independent particles problem. $E_{XC}[n]$ contains the exchange-correlation and the difference in kinetic energy between the real interacting ensemble of electrons and the fictitious non-interacting one. We can say that $E_{XC}[n]$ collects somehow the *exchange-correlation effects*, even though some authors like to underline its scarcely physical nature, treating it merely like a mathematical residual. There are two important points here: on one side, provided we know the exact functional form of $E_{XC}[n]$, we are doing here an exact treatment of the quantum mechanical problem, except for the Born-Oppenheimer approximation; on the other side, it has been shown that it is possible to work out excellent approximations of the exact functional that perform very well, predicting many important physical properties [we will come back on this crucial point in Sec. (3.3)].

Taking advantage of the variational nature of the energy functional with respect to the non-interacting electronic states Ψ_i and constraining them to be orthogonal, we obtain the self-consistent Kohn-Sham equations:

$$\hat{h}^{KS}\Psi_i(r) = \{-\frac{\hbar}{2m}\nabla^2 + V_{eff}\}\Psi_i(r) = \epsilon_i\Psi_i(r), \tag{3.12}$$

where V_{eff} is the effective potential, given by:

$$V_{eff}(r) = V(r) \; + \; V_H[n] \; + \; V_{XC}[n], \tag{3.13}$$

where $V_{XC}[n]$ is the exchange-correlation potential, defined as the partial derivative of $E_{XC}[n]$ with respect to $n(r)$. It is noteworthy that both the Hartree potential V_H and the exchange-correlation potential V_{XC} are non-local functionals of the charge density distribution, therefore the total effective potential V_{eff} at each point r depends on the electronic density in all points of space.

3.2.1 Solution of the Kohn-Sham equations

The solution of the Kohn-Sham equations is the core of a DFT problem and the ability of managing this task efficiently is what makes DFT such a useful theory. We will not go through it here, but we will spend a few lines in describing the procedure to solve self-consistently the electronic structure. The algorithm is well defined: (i) start with an initial guess for the electronic density $n(r)$; (ii) build the Kohn-Sham Hamiltonian and the overlap matrix of Eq. (3.2); (iii) solve the one-particle states; (iv) build a new charge density [2] from the occupied states, and (v) iterate until self-consistency is achieved, that is, until the charge density in two successive steps is equal within a certain tolerance.

The time consuming steps of this procedure are the construction of the KS Hamiltonian and its solution, from which the one-electron states are obtained. It is to perform these steps that a variety of methods, from standard direct diagonalisation to the Car-Parrinello [24] dynamical approach, exist, but the body of the self-consistency cycle is the same.

3.3 Exchange-correlation functionals

The piece missing in Eq. (3.11) is the functional of the exchange and correlation energy. Unluckily, this functional cannot be known exactly, however, the success of DFT relies just on the extremely good performance of its approximation. In this section we will give a brief outline of the most common flavours of exchange-correlation functionals.

3.3.1 The Local Density Approximation

From the formulation of the self-consistent Kohn-Sham equations it is clear that everything relies on the ability of including the exchange-correlation effects of $E_{XC}[n]$. Kohn and Sham [20] proposed a first approximation which is remarkably simple and which turned out to work embarrassingly well. It is the so-called *local density approximation (LDA)*:

$$E_{XC}^{LDA} = \int \epsilon_{xc}[n(r)]n(r)dr \tag{3.14}$$

where ϵ_{xc} is the exchange-correlation energy density of a uniform electron gas of density n. The good point about this approximation is that in such a

[2] This can be easily done from the relation: $n(r) = \sum_{i=1}^{N_e} |\Psi_i(r)|^2$

case the exchange part is know exactly and it is:

$$\epsilon_x(n) = -\frac{0.458}{r_s} \text{ (in a.u.)} \tag{3.15}$$

(where r_s is the radius of a sphere that contains one electron) and the correlation part is know analytically for some limiting cases and can be parametrised through numerical Monte Carlo simulations at intermediate densities. A widely used choice is the parametrisation of Perdew and Zunger [25] of the numerical results of Ceperley and Alder [26]. In other words, the *black box* we have nucleated under the name of $E_{XC}[n]$ can be known in this special case. The LDA approach amounts to assume that at every point, the electronic density has an exchange-correlation contribution equal to that of a homogeneous electron gas with the same density. LDA was expected to work very well for systems with a quite uniform electronic density, closely resembling that of an electron gas, however it turned out to work surprisingly well for most real systems, where those uniformity requirements are seldom met.

LDA provides excellent estimation of geometries and bond angles within 1 - 2 %. Ionisation energies, binding and dissociation energies have a typical accuracy of 10 - 20 % (typically giving overbinding). Where LDA has shown not to work very well is with weakly bonded systems (where H-bonds or Van Der Waals interactions are dominating).

Moreover, LDA has a marked tendency to seriously underestimate energy gaps, when its eigenvalues are interpreted as excitation energies. The ultimate cause of the gap error lays in the screening of the exchange hole when the electron is removed. A few gap correction techniques have been proposed, nevertheless this remains one of the most serious limitation of the LDA functional.

3.3.2 Generalised Gradient Approximation

The *generalised gradient approximation (GGA)* [27] represents a further refinement of the LDA functional. The idea is to treat LDA as the first term of a power series, which can then be further expanded. However, such an expansion does not converge monotonically and the first order correction worsens LDA performances. What is therefore implemented is a correction which sums the series to an infinite order and this purpose is achieved by the GGA:

$$E_{XC}^{GGA} = \int \epsilon_{xc}[n(r)]n(r)dr + \int f[n(r), |\nabla n(r)|]n(r)dr \tag{3.16}$$

where the second term is the correction to the LDA functional.

The use of GGA improves the energetics of the bonds, which are in general more accurate. The improvement is particularly drastic in the description of the H-bonds, even if the same does not happen with the Van Der Waals interactions. Unfortunately, it leaves almost unsolved the gap-error problem.

3.4 Basis sets

In Sec. (3.2) we have described the principles of DFT and how the Kohn-Sham equations reduce the many-body problem in terms of a problem of independent electrons which move in an effective, self-consistent potential. This approach represents an enormous simplification, because it allows us to forget about the many-body problem. However, the one-electron problem still has to be solved and to do this we need to numerically represent the one-electron wave functions in such a way that they can be manipulated computationally. This purpose is normally achieved expanding the electronic states on a basis set.

The basis set functions define a Hilbert space and the expansion that the eigenstates of the Kohn-Sham equation undergo is:

$$\Psi_i(r) = \sum_{\mu=1}^{N_b} C_{\mu i} \Phi_\mu(r) \tag{3.17}$$

where N_b is the number of the basis functions. Now the Schrödinger equation can be re-cast in matrix form:

$$h^{KS} C_i = s C_i \epsilon_i \tag{3.18}$$

in terms of the Kohn-Sham Hamiltonian h^{KS} and of the overlap matrix s

$$\begin{cases} h^{KS}_{\mu\nu} = \langle \Phi_\mu | h^{KS} | \Phi_\nu \rangle \\ s_{\mu\nu} = \langle \Phi_\mu | \Phi_\nu \rangle \end{cases} \tag{3.19}$$

Provided that the quality of the basis set is sufficiently good for the description of the system one wants to study, the problem is reduced from a set of coupled differential equations, to the diagonalisation of a $N_b \times N_b$ matrix.

3.4.1 Plane waves

The first basis set which was used in the case of solid state physics and which is still the most popular are plane waves (PW's) [28]. They are intimately linked to pseudopotentials [see further on Sec. (3.5)] and they are extremely familiar to the solid state community for their key role in the Bloch theorem.

A PW expansion has the form:

$$\Psi(r) = \sum_{\mathbf{G}=0}^{\mathbf{G_{max}}} C_k(\mathbf{G}) e^{i\mathbf{G}r} \tag{3.20}$$

where the sum is over reciprocal lattice vectors. The $|G_{max}|$ up to which the sum is extended represents the cutoff of the expansion. Carried to the real space, neglecting plane waves beyond a certain cutoff means loosing the resolution of the wave function for distances below a certain value.

There are many benefits connected with the use of PW's:

- It is a very systematic basis set and its quality can be systematically improved by simply adjusting one single parameter (the PW cutoff).
- It treats all points of space on the same way, making therefore no assumption on the system studied.
- The calculation of the Hamiltonian matrix elements is formally and computationally very simple.

On the other hand, the main drawbacks are:

- Many PW's per electron are needed to describe its orbital accurately, making PW's, despite their formal simplicity, an overall computational inefficient basis set.
- The unlocalised character of plane waves makes them unsuitable to describe efficiently localised electronic states (which would require an infinite number of PW's in the basis).
- Its use requires periodic boundary conditions.
- It is a a highly inefficient basis set every time extended empty regions are featured in the system (unavoidable when slabs or clusters need to be studied) [3].

[3] The problem with handling slabs consists in the constraint PW have to study periodic systems. If one needs to describe a system lacking of periodicity in one (like a slab), two (like a wire) or three (like a molecular cluster) dimensions a trick is normally used. It consists in using a very large simulation cell in the non-periodic direction, in such a way that the neighbouring images are not interacting. Hence the need to include a certain amount of empty space.

3.4.2 Localised orbitals

The main alternative to PW's consists in expanding the eigenstate wave functions on a basis set of localised orbitals. The quantum chemistry community have been using for many years basis sets made up of Gaussian-like orbitals, while among solid state physicists is becoming quite popular the use of optimised atomic orbitals like in the standard *LCAO (linear combination of atomic orbitals)* approach.

The advantages of such a basis set are:

- Very few atomic orbitals are needed and then the basis set is computationally highly efficient.
- The intrinsic atomic nature of the basis function makes quite easy to extract chemical information like charge transfer, Mulliken charges,... However, it should be underlined that such information is not always reliable.
- Localisation ideas make quite straightforward to apply the *nearsightedness* principle [29] to develop order-N algorithms [30, 31], i.e $O(N)$.
- No assumption of periodicity must be made and non-periodic systems can be easily studied.
- Handling large amounts of empty volume in the simulation cell is not especially critical, since the basis set is able to put less orbitals where they are not needed (in the empty regions).

Like PW's, localised orbitals basis sets too have some relevant negative aspects:

- The main inconvenient is the lack of a systematic way to optimise the basis set, because one must act on a wide range of parameters (number of orbitals, their extent, their shape,...) to increase its quality.
- The basis set is not orthogonal, which induces basis set superposition errors [32]. More importantly, using a larger number of basis functions does not guarantee better results.
- The calculation of the Hamiltonian matrix elements is formally complicated and in general expensive.

$\sim$

As we have discussed, PW and LCAO basis sets have their own advantages and drawbacks. Traditionally, PW has been the most typical choice in solid state physics, while LCAO has been preferred by quantum chemists. More recently, it seems that such a clear-cut separation is fading away, each day more. The solid state community, for instance, has relied for many years on the typical *condensed matter approach* of PWs, but recently it is moving more and more toward LCAO basis sets, even though PWs perhaps remain the most frequent choice. This is the case especially in those applications that require large supercells, and where, therefore, the efficiency of the basis set becomes crucial.

3.5 Pseudopotentials

An intrinsic problem connected to first principles electronic structure calculations is constituted by the representation of the wave functions $\Phi_i(r)$. One of the difficulties consists in the different behaviour $\Phi_i(r)$ has close to the atomic cores, where kinetic energies are very high, or in the space between the atoms, where kinetic energies vary much more slowly.

In many systems, however, there is a clear separation between the core and valence orbitals. The valence orbitals have the greater spatial extent and are the main responsibles for chemical bonds, while the core orbitals have energies far smaller and are confined around the nuclei. In other words, any perturbation experienced by an atom in condensed matter will induce only negligible response of the core orbitals. These considerations suggested the assumption of the so-called *frozen core* approximation, according to which the core orbitals are held fixed when total energy is minimised. Implementing the *frozen core* approximation has several advantages, as we will discuss briefly, but two features must be taken into account when *freezing* the core orbitals. Firstly, the core orbitals contribution to the Hartree and exchange-correlation energies felt by the valence orbitals has not be neglected. Secondly, the valence orbitals must remain orthogonal to the core orbitals, otherwise nothing would prevent them to collapse and become core orbitals themselves.

The potential felt by the valence electrons is due to: (a) the nuclei; (b) the core electrons; (c) the other valence electrons. The sum of the terms (a) and (b) represents what sometimes is called the true ionic potential. The pseudopotential method relies on replacing the true ionic potential with an ionic pseudopotential, constructed in such a way that it leaves the Kohn-Sham energies of the valence orbitals and the energetics of the whole system (in terms of relative energy differences) unaffected.

In other words, we want to be able to give an equivalent description of

the complex nucleus plus core electrons - as felt by the valence electrons - in such a way that above a certain cutoff radius the difference cannot be noted. Of course this prevents us form giving any physically reasonable description of what goes on below that cutoff radius, but as far as we are interested in a condensed matter system and then in how atoms chemically bind between each other, this is not a serious limitation.

This method brings along a remarkable improvement in terms of computational efficiency, because we do not need anymore to describe the many core electrons, modelling them with a unique pseudopotential, equal for all the atoms of the same species.

In 1982 Bachelet *et al.* [33] published a list of pseudopotential for all elements up to Pu, that has found widespread application. Improvements were introduced with the advent of the separable form of Kleinman-Bylander [34] and new pseudopotential lists have been proposed.

Of course one of the central issues with a pseudopotential is its *portability*. What should be questioned is to what degree the pseudopotential for a certain element can perform equally well, in different systems or environment. In other words, can a pseudopotential for H be used to describe the hydrogen molecule as well as a hydrogen impurity in a Si crystal? To be sure it does, pseudopotentials must be thoroughly checked and tested before using them.

3.6 A DFT code: SIESTA

Since the formulation of DFT, a number of attempts to code efficient programs to perform electronic structure calculations have been made and many have been remarkably successful.

In this section we will outline the main feature (an extended and detailed description can be found elsewhere [35]) of SIESTA (Spanish Initiative for Electronic Structure with Thousands of Atoms) [35] the code implementing DFT [19, 20] that we have used to obtain the results presented throughout this work. The code has been developed by a group of Spanish physicists in the late 90's. It is currently under continuous evolution and it has been recently efficiently parallelised.

SIESTA uses a basis set of localised atomic orbitals modelled on the classical linear combination of atomic orbitals (LCAO). Although the atomic orbitals decay quite rapidly, they have a virtually infinite extent. SIESTA uses a different approach, confining the atomic orbitals used to build the basis set inside a finite cutoff radius. The confinement is done taking care that the norm is conserved and providing a smooth truncation at the cutoff radius (first two derivatives continuous) The localised basis functions so

obtained are a modification of the classical LCAO scheme called *numerical atomic orbitals (NAO's)* [36, 37].

The advantages of using such basis functions are essentially two:

- the Hamiltonian matrix will be highly sparse because each atom will have a limited interaction radius with its neighbours. It can be shown that this makes the construction of H and S an $O(N)$ operation, where N is the number of the atoms in the system. This an important property because building up the Hamiltonian is a computationally demanding task.

- the smooth truncation of the basis functions allows to take advantage of the concept of localisation when implementing $O(N)$ [30, 31] algorithms [4].

The confinement of the orbitals is controlled through a parameter indicating the error committed in energy with the truncation. Besides adjusting the confinement in an easy way, the value of this parameter indicates at the same time how far one is from the standard LCAO case.

As mentioned above, SIESTA has the facility of solving the electronic structure problem in an $O(N)$ fashion. This represents one of the latest challenges in first-principles calculations, because the possibility to have a linear scaling algorithm as an alternative to standard $O(N^3)$ diagonalisation would allow to approach much larger systems and to deal with hundreds or a few thousands of atoms. Although this capability has not been used in the present work, many results can be found in literature obtained with the remarkable performances of this SIESTA algorithm.

[4] During the development of SIESTA, the focus has always been on coding an efficient, linear-scaling - order-N - DFT program, that is why a NAO basis set was chosen.

Chapter 4

Simulation techniques

In the previous chapter we have discussed how the total energy of a system can be calculated from first-principles, specifically according to density functional theory (DFT), which is nowadays the state-of-the-art *ab initio* approach in theoretical solid state physics and materials science. We have analysed DFT in detail because most of the calculations presented in this work were carried out in such a computational framework. However, a wide class of alternatives exists, and not necessarily from first-principles, like semi-empirical tight-binding methods or empirical potentials.

In this chapter we will briefly outline some of the most important techniques in materials science simulation (at least those relevant to this work), which are completely independent of the model used to calculate the total energy. The only assumption is that we are treating a system governed by a potential that can be derived, with respect to the position of the particles and lattice parameters. Such derivatives are the forces acting on each particle and stresses acting on the cell, as determined by the given potential.

Roughly speaking, there are three classes of simulation techniques: (i) static calculations, i.e. structural relaxations, (ii) dynamic calculations, i.e. molecular dynamics, and (iii) Monte Carlo simulations. As we have stressed above, each of these kinds of simulation can be in principle coupled with any computational model to calculate the total energy. Commonly, DFT calculations focus on structural relaxations, because dynamics require a comparatively larger number of evaluation of the energy and forces, in order to achieve statistical significance or to observe some kind of thermally activated transition. For these reasons, molecular dynamics is often erroneously associated to the use of empirical potentials, but as we will discuss in Chapter (6) efficient first-principles molecular dynamics can be performed as well. Monte Carlo resembles molecular dynamics in the sense that the effect of a non-zero temperature is also introduced, however the forces are not calculated and the

atoms are moved generating random configurations and accepting them on the basis of a probability acceptance rule: if the energy of the randomly generated configuration is lower than the previous one, the configuration is accepted; if it is higher, it is accepted with a probability dictated by the ratio of the Boltzmann factors. In this work no systematic use of Monte Carlo has been done, therefore we will not describe it (see for example Ref. [39]).

4.1 Structural relaxation

Structural relaxation takes advantage of the knowledge of the forces on the atoms to deduce the equilibrium geometry of a system. The principle is very simple: (i) the total energy and the forces are calculated and (ii) each atom is moved along the direction pointed by the resulting force on it. The process is iterated until the forces are reduced under a tolerance limit. With some more rigour, we should say that this problem is nothing more than the search for a minimum in a 3N-dimensional space (where N is the number of atoms). As can be appreciated in Fig. (4.1), the convergence of the maximum residual force - being slower - is a more restrictive criterion than than the convergence of the potential energy.

After defining the numerical problem, one can choose different algorithms solve it. The most obvious choice is the so-called *steepest descent* algorithm, where the atoms simply follow the minimum energy direction, according to the gradient they experience at each step. The line search is performed moving along the direction where the gradient decreases more rapidly, so the positions are updated according to the rule

$$x_i = x_{i-1} - \alpha f'(x_{i-1}). \tag{4.1}$$

This method works well for well-convex 3N-dimensional surfaces, but otherwise can be rather inefficient, as it requires a great many iterations for functions which have long, narrow valley structures.

The common alternative implemented in all the major simulation codes used today is the *conjugate gradient (CG)* algorithm. In the steepest descent method any new gradient is constrained to be perpendicular to the direction just traversed. However, what we really want to do is to proceed not down the new gradient, but rather in a direction that is constructed to be conjugate to the old gradient and, possibly, to all the previous ones [1]. In other words,

[1] We intend to provide here only a qualitative description of the method, without entering the mathematical details and proving the necessary theorems. However, an extensive description can be found, for instance, in *Numerical Recipes* [66].

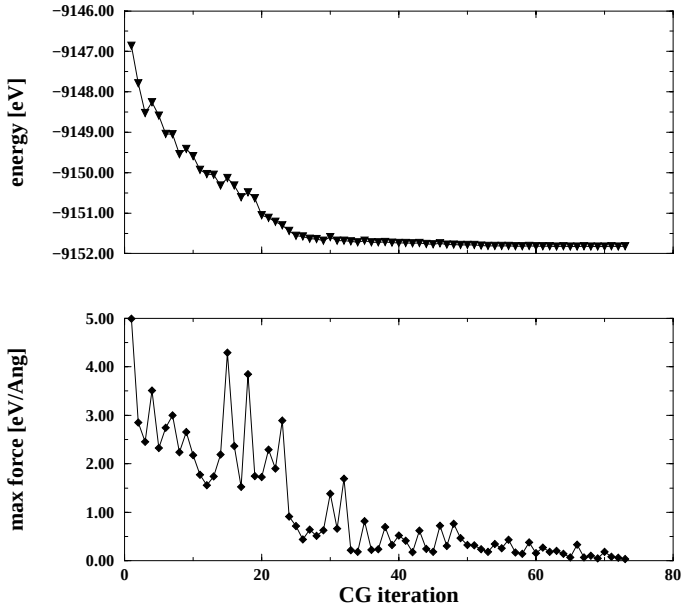


Figure 4.1: Convergence of the Kohn-Sham energy and of the maximum residual force in a structural relaxation (DFT minimisation with CG algorithm of 3C-SiC Si-face surface with an interstitial sub-surface oxygen). The forces converges more slowly and in a more noisy way than the energy.

conjugate gradient is an algorithm that has a longer *memory* than the single-step memory of the steepest descent method. Thus, in conjugate gradient minimisations the direction of the line search at step i is not only determined by the condition at steps $i-1$, but depend also on previous steps.

An intrinsic problem of minimisation algorithm is that, if they converge, it cannot be guaranteed that the absolute minimum has been reached. Depending on the starting guess for the minimum energy structure, it is possible to converge to one of the local minima that can in principle exist. There is no simple solution to this problem, unless starting from a physical sound geometry.

The fact that structural relaxation methods provide the equilibrium geometry should not be misunderstood. In some cases, typically in surface physics, the geometry of the system can be accessed experimentally, for instance by atomic force microscopy and comparison with structures obtained with a simulation can be of great help. On the other hand, in bulk defect studies, which is the main object of this work, the relaxed geometry induced by defect is not directly what we are interested in. However, if we want to in-

fer on defect stability or, for instance, tracing a diffusion barrier as a defect is moved throughout the lattice, the energies must be correctly calculated and they must therefore refer to the corresponding relaxed structures. Obtaining the equilibrium structure is of capital importance also in bulk studies, not so much for the geometry itself, but rather for the associated energy and self-consistent electronic structure.

4.2 Molecular dynamics

In structural relaxation the kinetic energy is not taken into account and what is obtained as a result of the minimisation process is the zero-temperature ground state structure or metastable structure if a local minimum is found. The main idea underlying molecular dynamics (MD) is to explicitly introduce a non zero-temperature and to observe the dynamic evolution of the system. In such conditions, the initial *thermal* velocities are obtained sampling them from a Maxwell-Boltzmann distribution corresponding to the simulation temperature and they are summed to the velocities obtained integrating the forces determined by the potential [2].

If in structural relaxation the forces are parameters of the minimisation algorithm, here they are used to let the system evolve recreating the conditions of a real experiment. The main steps of a typical MD algorithm are the following: (i) velocities of the atoms are randomly sampled from a Maxwell-Boltzmann distribution (this step is done only once, before beginning the dynamics to obtain the initial velocities); (ii) the potential is calculated and derived, (iii) Newton's equation of motion are integrated and the velocities and the new positions of the atoms are calculated.

In other words, we calculate the velocity of each atom and we let it move according to the laws of classical mechanics for a fixed time step Δt. Usually, the set of particles in an MD is treated as a classical many-body system, i.e. following the laws of Newton's classical mechanics. This is a very good approximation for a wide range of materials and conditions and it becomes critical only when dealing with the vibrational and rotational properties of light atoms and molecules (He, H_2), where the quantum effects are important.

Of course, it is very important to have a good algorithm for integrating Newton's equations of motion, numerically stable and possibly fast [3]. One

[2] The second order differential equation that is integrated is Newton's second principle of dynamics $F = m\frac{d^2r}{dt}$. Velocities and positions can therefore be directly deduced from it.

[3] Actually, the speed is is not a tight requirement, because even when the energy and

of the simplest and best alternatives to integrate the equations of motion is the so-called Verlet algorithm [39]. The new position can be estimated from

$$r(t+\Delta t) \sim 2r(t) - r(t-\Delta t) + \frac{f(t)}{m}\Delta t^2, \tag{4.2}$$

where Δt is the time step of the dynamics, $f(t)$ the force at instant t and m is the mass.

An important point to be made is that the choice of Δt must be done carefully. Larger time steps allow to span longer simulation times with a few evaluation of energy and forces, however when Δt is too long problems both on the numerical and physics side may arise. If the time step is too long, the configurations and the energies corresponding to two nearby snapshots of the dynamics will be, in general, very different. However, to conserve satisfactorily the energy, it has to vary smoothly and therefore Δt has to be conveniently small.

Typically, the main applications of MD are related to the extraction of the average statistical properties of a system. An example could be the estimation of the radial distribution function of a liquid or the diffusion coefficient of an impurity in a solid. These are cases in which many configurations have to be explored in the dynamics, thus requiring long simulation times. In such applications, first-principles molecular dynamics (FPMD) implies a very high computational load, but nothing prevents its use. However, especially when observation of very rare events is required, MD is still more commonly used in conjunction with tight-binding methods or empirical potentials.

A less performance-demanding application of FPMD will be discussed in Chapter 6. Therein, we will analyse the case of the diffusion of a B impurity in a solid. The hops that the impurity does from one minimum energy configuration to the other are not very likely events compared to the time scale that can be spanned by FPMD. What we have done (once again, all the details will be given in Chapter 6) was running a FPMD until one of this hop was observed. In this way, that transition can be characterised not only by the height of the corresponding barrier, but observed dynamically while it occurs.

Another interesting feature of MD in general (and particularly of FPMD, due to the high accuracy of energy and forces evaluation) is that it is the computational approach that more closely resemble a real experiment. For this reason, the dynamical simulation can be designed in a such a way that

---

the forces are calculated on the basis of an empirical potential, the time consumed there is normally much longer than what it takes to integrate the equations of motion, dominating the overall run time.

it is not *polarised* by the user's choices. This is very useful when it is difficult to formulate *a priori* hypothesis on the physics of a process, e.g. a molecule adsorption at a reconstructed surface. Therefore, MD can indicate which are the relevant structural relaxations to perform. On the other way around, MD can validate the results of a series of structural relaxations, to be sure that no important process has been neglected. After elucidating, for example a diffusion path, observing the mechanism really happen in a MD [4] is a strong confirmation that we have not caught the relevant physics. MD is thus by all means a computational experiment that can be used to validate what was deduced by static calculations or to suggest new ones. The most serious problem is that in many cases the time scale that should be spanned is prohibitive for the accurate first-principle approach and to move to semi-empirical or empirical methods reliable parametrisation are not always available.

4.2.1 Molecular dynamics in different ensembles

The MD algorithm discussed so far is a scheme to study the time evolution of a classical many-body system of N particles confined in a volume V and isolated from the rest of the universe. In such condition the total energy E is a constant of motion.

This is the so-called *micro-canonical ensemble* or simply NVE ensemble. Its implementation, as we have seen, is relatively straightforward and the principles underlying it are rather simple, as it is essentially only a matter of performing classical Newtonian dynamics.

However, most of the experimental conditions do not fit this somewhat restrictive constraint. In a laboratory, experiments are typically performed maintaining constant the temperature and fixing the volume of the specimen together with the pressure; alternatively, the volume may me left free to expand or contract and it is the temperature and the pressure that are fixed. What is certainly difficult to reproduce are the conditions of the micro-canonical ensemble: constant volume and energy.

For these reasons, a variety of more sophisticated algorithms were developed to perform simulations where the temperature on one side, and the volume or the pressure on the other, are conserved, respectively the NVT and the NPT ensembles. We will not give details on these algorithms, but excellent descriptions can be found elsewhere [38, 39, 40, 95]. The basic ideas of these algorithms consist in using an extended Hamiltonian, where fictitious variables are introduced, i.e. a thermostat to conserve the temperature or a

[4] See Chapter 6

barostat to conserve the pressure. These variables have their own (fictitious) mass and take part to the MD like any other particle, but have the duty of adsorbing or releasing energy, to conserve a specific observable quantity like pressure or temperature. It is important to note that what is conserved is the average of these magnitudes. Thus, in a NVT simulation, for instance, the thermostat will conserve the average temperature, but its instantaneous value will oscillate around it.

Chapter 5

Defects and defect diffusion in semiconductors

5.1 The formation energy

The study of defects in semiconductors is a very wide field, both for the intrinsic fundamental questions related to it and for the relevant implications from the application viewpoint. *Defect* is a fairly generic denomination when referred to semiconductors. In fact, it refers really to undesired defects as well as to dopants; in both cases a *defect* is a perturbation of the original, perfect crystalline structure of the host lattice, but in one case they are the random product of an imperfect technological step which lower the conduction properties of the material, while in the other they are purposely introduced to provide the material with special electrical features to be exploited later in device design.

The perturbation of a perfect, periodic lattice is therefore the common characteristic of what we will generally refer to as a *defect*, it being an intrinsic defect of the crystal or a dopant atom. Many of the properties that are normally of interest depend on this local perturbation and these aspects can be fruitfully approached by means of theoretical calculations.

The relaxation of the system, searching the minimum in the phase space, i.e. the optimum structure, reproduces the equilibrium configuration around the defect, thus giving an estimation of the extent of the local perturbation. However, the relaxed geometry, which most of the times is difficult to access experimentally, has itself only a limited interest. What is most interesting is the electronic structure, the projected and total density of states or the charge density associated to the relaxed geometry, because all the application relevant properties of the material ultimately depend on them.

All this amount of information comes along with the structure relaxation, as at every algorithm step the electronic structure of the system is calculated self-consistently, so it is always available, therefore also when the minimum energy configuration is reached.

What has been described so far is, in summary, almost everything one needs to know on the physics of a defect in a semiconductor host. Thus, given that one is interested, for instance, in a substitutional hydrogen defect in silicon, the only thing to be done, simplifying slightly, is changing one silicon atom for a hydrogen, relaxing the system and analysing the resulting electronic structure. However, most of the times this simplified picture does not correspond to a realistic situation of interest. To stick to the previous example, if the physics of a hydrogen defect in silicon is the topic of a study, first it should be established if the hydrogen atom is likely to substitute at a silicon site and after that at least it should be elucidated if, doing, it favours some charge state over others. As it can be seen, the crucial, preliminary question to address is the stability degree of a defect configuration, both from an absolute point of view, and comparing it to other possible configurations. In other words, we should ask what is the amount of energy that has to be spent (or that is gained) to *sustain* the *creation* of a defect.

A useful theoretical tool is represented by the *formation energy* of a defect, which can be conveniently defined as:

$$E_F = E_D^{tot} - \sum_i n_i \mu_i + q(\mu_e + E_V) \tag{5.1}$$

where E_D^{tot} is the total energy of the system containing the defect, n_i is the number of atoms belonging to species i and μ_i its chemical potential; the index i runs over all the species present in the system. The net charge state of the system is q and the chemical potential of the electron μ_e is referred to the top of the valence band, E_V: in this way μ_e varies from 0 - at the top of the valence band - to E_{gap} - at the bottom of the conduction band -, thus spanning the whole range of doping conditions.

Northrup and Zhang [41] gave a formulation of Eq. 5.1 for multispecies compounds that, in its different flavours, has become a standard. Eq. 5.1 in the case of a SiC lattice with a defect belonging to the generic species X is simply:

$$E_F = E_D^{tot} - n_{Si}\mu_{Si} - n_C\mu_C - n_X\mu_X + q(\mu_e + E_V) \tag{5.2}$$

It should be noticed that μ_{Si} and μ_C are the chemical potential of Si and C *in* SiC. We underline this fact, because the definition of the Si chemical potential is as delicate in SiC as it is straightforward in its bulk state, i.e. in bulk Si. The same of course holds for C in SiC or in bulk diamond. The

latest equation can be rewritten as:

$$E_F = E_D^{tot} - \frac{1}{2}(n_{Si} + n_C)\mu_{SiC}^{bulk} - \frac{1}{2}(n_{Si} - n_C)({\mu_{Si}}^{bulk} - {\mu_C}^{bulk} + \Delta\mu) + q(\mu_e + E_V) - n_X(\mu_X - \mu_B^{bulk}) \quad (5.3)$$

The chemical potential of bulk SiC is $\mu_{SiC}^{bulk} = \mu_{Si}^{bulk} + \mu_C^{bulk} - \Delta H_f$, where ΔH_f is the formation heat of SiC. Now E_F is function of the bulk chemical potential of C and Si and of the parameter $\Delta\mu$ that accounts for the difference between the chemical potentials of C and Si in SiC and in their respective bulk state. This reformulation has the advantage of being expressed in terms of well defined quantities (the bulk chemical potentials) and with the parameter

$$\Delta\mu = (\mu_{Si} - \mu_C) - ({\mu_{Si}}^{bulk} - {\mu_C}^{bulk}) \quad (5.4)$$

which is easily associable to the macroscopic stoichiometry conditions of the material. $\Delta\mu$ can vary between $-\Delta H_f$, limit that corresponds to the C-rich condition, and ΔH_f, for the Si-rich material, condition fixed by the inequalities $\mu_{Si} \leq {\mu_{Si}}^{bulk}$ and $\mu_C \leq {\mu_C}^{bulk}$.

If the chemical potentials of all the species involved are well-defined, this general relation allows us to estimate quantitatively the energy cost of *forming* a defect in a given configuration and charge state. However, the chemical potential of the impurity is a magnitude whose definition is very difficult to handle, as in principle it depends on which was the phase state of the element prior to doping and on its ground state. An exhaustive discussion for the case of H in SiC can be found elsewhere [42].

Luckily, this problem can be obviated on most occasions, provided that we restrict ourselves to analysing system with the same number of impurities. This is a quite common case. Referring again to the example given above, when approaching the study of H in Si, the first thing to do would be studying different configurations of the H atom, in different charge states. Therefore, when comparing the relative stabilities of all the systems considered, the term which is function of the H chemical potential always cancels out. Thus, although Eq. 5.2 gives an absolute estimation of the formation energy of a defect, in most of the cases it is used in relative stability comparisons; in this way relying on the definition of the impurity chemical potential is safely avoided.

Looking again to Eq. 5.2, it should be noted that there is a linear dependence of E_F on the electron chemical potential, i.e. the position of the Fermi level, whose slope is the system's net charge. In other words, for a non-neutral system, the formation energy depends on the doping conditions of the host material, therefore, changing the value of μ_e, the most stable charge state of a defect in a given configuration will in general also change.

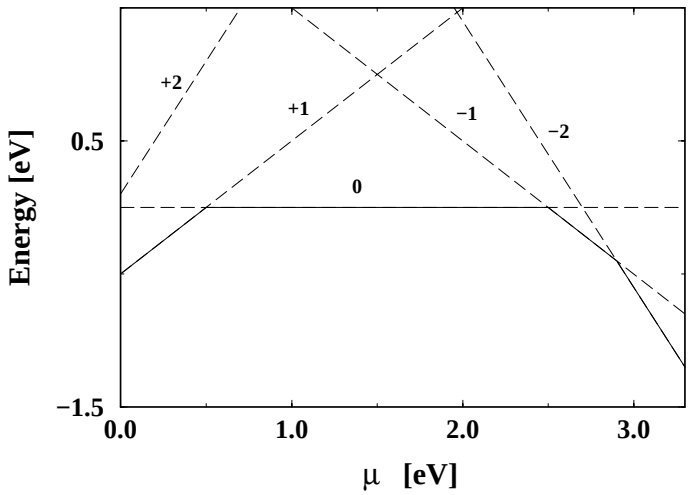


Figure 5.1: Formation energy vs chemical potential of the electron μ of a defect configuration for different charge states.

To conclude this section, we give a detailed example of how the former considerations are used in practise. Let us consider the prototypical example of an impurity in SiC [it could B, as in Chapter (6), N or P, as in Chapter (8)] that allows us to briefly refer to it when in the following sections we will face similar cases. Let us restrict ourselves to the case where only three configurations are relevant: substitution at Si site and two different kinds of interstitial position (it does not really matter which ones). We start considering one of the interstitial configurations in the stoichiometric case $\Delta\mu = 0$ and we perform a full structural relaxation for each of the relevant charge states (typically five: ± 2, ± 1 and 0). After that we calculate the formation energy in each of the five cases and we plot it as a function of μ_e in Fig (5.1). The plot indicates which is the charge state that the interstitial configuration considered favours for different positions of the Fermi level. From Fig (5.1) we can conclude, for instance, that for a Fermi level close to the conduction band edge (n-type doping condition) the impurity, in that interstitial position, is likely to be found in the -2 charge state, while for a Fermi level just above the valence band edge (p-type doping condition) that configuration would favour the $+1$ state. The procedure is repeated for the other interstitial position and for the substitutional configuration. Similar plots are obtained and finally overlapped in the same graph, as shown in Fig (5.2). Now we know, for a given doping condition of the host material, which is the most stable configuration of the impurity considered and which charge state it favours. We remark again that to carry out this study no

estimation of the chemical potential of the impurity is needed, because in all the plots discussed it would represents only a constant offset, thus we could assume it equal to zero (or any other arbitrary value).

So far we have been working in the stoichiometry hypothesis, i.e. $\Delta\mu = 0$, but at this point it might be interesting to study the difference that would be observed in a Si- or C-rich material. It can be seen from Eq. 5.2 that a $\Delta\mu \neq 0$ will affect the value of E_F only if the number of Si and C atoms is different, i.e. $n_{Si} \neq n_C$. In the simplified study case that we are discussing, this is the case only for the substitutional configuration. This means that the plot corresponding to the interstitials configuration will lay unmodified, but a Si-rich material will result in a formation energy of the substitutional higher of an amount equal to $\frac{1}{2}H_f$; in the same way, it will be $\frac{1}{2}H_f$ lower in the case on a C-rich material. As it is illustrated in Fig (5.2) this rather small shift can change the favoured configuration in determined region of the plot.

As a final remark, it should be noticed that we have shown how this kind of thorough analysis can be carried out without a well-defined value of the impurity chemical potential, μ. At the same time, however, we should underline how this lack of knowledge prevents us from comparing systems with different impurity species or number. Thus, nothing can be said about the relative stability of an interstitial P and an interstitial N, for instance, unless we know μ_P and μ_N. In the same way, we cannot directly compare the stability of one interstitial N with a pair of substitutional N atoms. The first case illustrated does demand a well-defined value of the chemical potentials, however in the other case a simplified, but effective analysis can be carried out most of the time, as we will discuss in the next section.

5.2 The aggregation energy

In the study of defects, an important issue is their tendency to aggregate or to dissociate, if aggregates form. The *aggregation* itself recalls the idea of comparing systems with a different number of impurities, i.e. how a system with n clustered impurities is favoured over a system with one single impurity. An intuitive approach to the analysis of the stability of an aggregate is the following: (i) calculating a supercell with the aggregate of interest formed, (ii) calculating the same supercell with the impurities in their most stable *isolated* arrangement, but spaced far enough from each other such that they are not interacting, (iii) evaluating the difference between the formation energies of the two cases. This procedure cannot be adopted in practise, because even referring to the simple case of a two-impurities aggregate, the super-

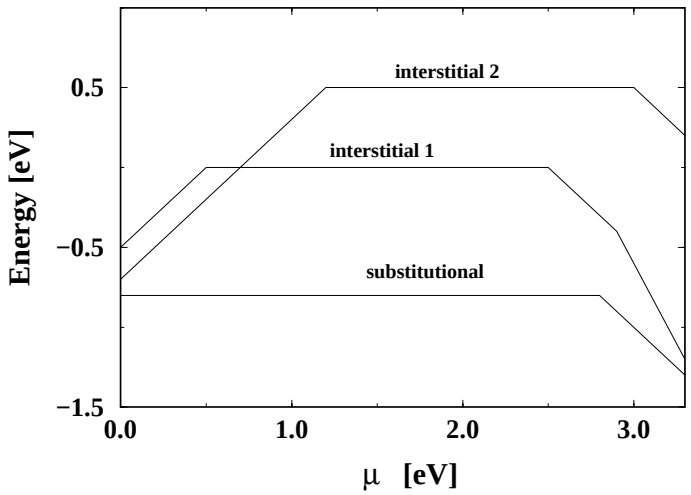


Figure 5.2: Formation energy vs chemical potential of the electron μ of three relevant defect configurations in their most stable charge state.

cell needed to achieve the non-interaction between them would be unusually large.

This apparent conceptual difficulty is easily overcome introducing the *aggregation energy*:

$$E_{agg} = E_F(\text{n-complex}) - nE_F(\text{isolated impurity}) \tag{5.5}$$

The idea underlying this relation is extremely simple and consists in using n times the formation energy of the isolated impurity in its most stable configuration. The obvious benefit of this approach is that the knowledge of the chemical potential of the impurity, once again, is not needed, as it cancels out. Moreover, the calculation does not need to be performed in a huge supercell, because we do not have anymore the problem of dealing with the issue of non-interacting isolated impurities.

This magnitude is very useful as it defines the *reaction heat* of the aggregation process, therefore how favoured is the formation of an aggregate under the conditions of thermal equilibrium. Unfortunately, it does not provide any information of the kinetics of the process of aggregation, so although an aggregation could be exothermic, high barriers could prevent it to happen easily.

5.3 Diffusion in semiconductors

Semiconductors have been one of the most important technological breakthroughs of 20^{th} century and, due to the increasing importance they have gained in a wide range of industrial applications, they have been impulsing both basic experimental research and theoretical modelling.

The necessity of performing feasible simulations, when the first principles approaches were still too demanding for the computational power available at the time, greatly encouraged the development of *ad hoc* empirical potentials that proved over the years to be extremely reliable if applied to a selected range of applications.

Modelling a material by an empirical potential brings along a trade off between accuracy and portability. The need to make an assumption on the spatial distribution of chemical bonds drastically limits the range of compounds to which good performances of the model can be expected. Many of the most well-known potentials have been developed for Si, of course, so they are expected to reproduce quite well ternary compounds like diamond, SiC, Ge, and it must be considered an outstanding success when a potential, like Tersoff potential, is able to accurately predict for C, together with diamond geometry, also the graphite structure. Another limiting issue for empirical potentials is that when they are flexible enough to span a certain amount of compounds, they need an accurate parametrisation of the atoms involved and of their mutual interactions. Such parametrisation can be sometimes difficult and in general needs a lot of experimental data.

The precision with which empirical potentials have been able to predict many equilibrium properties of semiconductors (lattice parameters, structural geometry) soon turned out to be not enough because all the topics related with the electronic structure of the solid remained uncovered. Then, as soon as the computational cost of *ab initio* methods became affordable, semiconductors became not only a ground to conquer, but also a crucial test field for DFT and its different functionals.

The size of the system still represents the most limiting factor of DFT performances, but the parallelisation of codes and the development of order-N algorithms highly contributed to make most of the problems approachable through first principles methods.

A peculiar limit of DFT - which is general, but especially annoying when the system studied is a semiconductor - is the systematic underestimation of the energy gap. The band diagram of a perfect semiconductor is reproduced with an excellent precision all over k space, except for a rigid downward shift of the conduction bands. The way to correct the gap error is then trivial in the case of a perfect crystal, because it is enough to make reference to

its experimental value. This issue becomes much less obvious when localised states of some kind of defect appear in the gap. A variety of techniques for gap correction have been developed, but a very wide class of problems - defect formation energy, transition barriers, relaxation of structures - can be successfully approached without depending on this point.

In semiconductors, even more than in other crystalline structures, it is definitely true that it is their defects that make them interesting. The intrinsic conduction, consisting of thermal excitation of electrons from the valence to the conduction band, is normally not achieved because it requires too high temperatures. In practise, semiconductors are doped and the electrons or holes of the shallow states induced are responsible of the conduction.

In semiconductors, then, the field of defects study is naturally wider than in other materials. Defects do not represent anymore only an element of degradation of the properties of the perfect crystal - like it is in the case of intrinsic defects or interstitial impurities - since a key feature such as conduction relies on the modification of the perfect lattice band structure induced by the presence of defects. Point defects are the defects which we typically address when we perform DFT calculations, because line defects like dislocations or stacking faults would require too large a cell. This notwithstanding, still we can address a number of interesting problems just by looking at point defects, like analysing the effect of dopants or characterising the diffusion path of an interstitial impurity. This makes clear enough why addressing the characterisation of defects, especially from the point of view of their electronic structure, may have a tremendous impact on materials technology.

In contrast to Si, which has been studied under the most different sides and which itself drove the development and testing of new theoretical models, SiC is a relatively virgin ground from the viewpoint of theoretical studies. Many processes that are very well know and for which simulations at every desirable order of accuracy are available in Si, have never been approached by theoretical modelling in the case of SiC. On one side this reveals a basic ignorance of the fundamental physics underlying many relevant processes, on the other it is simply a result of the higher industrial importance of Si, which made it more attractive in the last 40 years.

Things are now changing and there is a growing interest in theoretical modelling of SiC. The portability of some popular empirical potentials for Si has been checked and a few of them (the Tersoff potential and the Stillinger-Weber potential) are being successfully used with SiC. At the same time, the evolution of computation algorithms and the growing power of currently available computers (from massive parallel supercomputers to ordinary desktops) make more accessible tight-binding [43] and first-principles calculations and the use of empirical potentials remains confined to a narrower, but no

less important, range of applications, where long simulation times or very large supercells are required.

A particularly interesting field of application where atomic scale characterisation is strongly needed, is the modelling of diffusion in SiC. In Sec. (2.4) we have already briefly outlined the technological reasons that make it extremely interesting to gain a detailed insight into this process. In the next section we come back to it with some more details, while in Sec. (5.4) we will comment on how diffusive phenomena are approached computationally and modelled.

5.3.1 Diffusion in SiC

As mentioned previously, a considerable amount of experimental and theoretical work has been devoted to the study of native defects and dopant diffusion in semiconductors (see the review of Fahey and coworkers [44] and references therein). Most of this work has focused on the case of Si as host, and comparatively little effort has been devoted to SiC, in spite of its growing importance in the semiconductor industry.

In Section (2.4) it was outlined that the main way to perform selective doping and create pn-junction in SiC is ion implantation. The possibility to carry out this task by means of diffusive doping would be a great advantage, because the lattice damage resulting from the implantation process cannot, in general, be fully recovered by means of thermal annealing. Moreover, diffusive doping would permit to create deep junctions, whose depths could be reached only by means of high-energy implantators [1]. Lateral diffusion, is also a key parameter for reliable device design, for instance in the case of power VDMOS. However, notwithstanding recent progresses with this technique, the aim is almost hopeless for many front-edge technological applications, because diffusion constants in SiC for most of the dopant impurities are very low.

The problem is more general. The selective doping of a portion of semiconductor material is the tool to design localised pn-junctions on which semiconductor device operation is based. Technological advances push continuously toward miniaturisation and the nanoscale limit is getting closer and closer, so the tolerance on the shape and dimensions of the junctions becomes sharper every day. As a consequence of this, an increasing ability in

[1] Conventional implantators are not able to reach the implantation energies required to create deep junctions and one must recur to high-energy implantators. In high-energy implantation, the ions must be previously accelerated, resulting in a much more delicate and time-consuming process (normally a day must be scheduled for an implantation, while few minutes are necessary with conventional equipments.

controlling the profile and the depth of an implanted layer, throughout all the main steps of the device processing and then during its operation is required.

The aspects for which the impurities diffusion constitutes a challenging problem and is relevant for device applications are many. On one side we want to be able to *move* impurities in order to create a desired doped layer; then we want to be able to keep them confined there. The opposite applies, of course, for undesired impurities, like intrinsic defects - namely vacancies and native interstitials - where we simply want to eliminate them by annealing them out. This stresses that even if ideally it would be desirable to dope through diffusion, this does not eliminate at all the crucial role that diffusion plays in semiconductors, especially for device fabrication.

Whatever technique it is used, we must be able to control the stability of the doped region shape; at the same time we would like to be able to eliminate undesired defects by a proper thermal annealing process. This latter aspect is less relevant when the material is doped by diffusion, because no damaged region is created; on the other hand however, the skill required to dope by means of this technique and to control the shape of the doped layer do need a demanding knowledge of the physics underlying this process. As it can be seen, then, addressing the diffusion issue has multiple implications.

Throughout this work, we will address impurity diffusion in different frameworks. However, we will always deal with *bulk* diffusion, without taking explicitly into account *exodiffusion*, i.e. diffusion toward the surface. In such a case, the diffusion process is slightly more complicated and represents a further step in modelling, as the impurity migration does not occur in a perfect host crystal, but rather in a highly damaged region.

Understanding which are the mechanisms that at the atomic scale rule the diffusion is the first step to build a detailed knowledge. This information is often left out by macroscopic simulations and experimental measurements. Typically, what can be deduced by means of an appropriate series of calculations is the minimum energy path (MEP) and the barrier associated to it. Finding the MEP of a diffusion problem can be a difficult problem, directly brought along by the minimisation algorithm itself: when it converges a local minimum has been found, but nothing assures that it is the absolute one. A few tested methods have been proposed, like the nudged elastic bands method [45, 46] or the ridge method [47], for searching directly for the MEP. Whatever the employed method is, the results must be always handled with care and it must be checked that the resulting MEP is physically plausible and that it contains no suspect discontinuities.

5.4 Methods to obtain MEPs

The more general approach to find MEPs for a given impurity in a semiconductor system consists in mapping the energy surface it experiences all around the crystal lattice. From an operative point of view this is done through a series of structural relaxation calculations, with the impurity constrained at a fixed point. The structure around it is fully relaxed, and the extent and characteristics of the relaxation will be determined just by the perturbing presence of the impurity. The procedure is repeated many times in order to provide a 3D mapping of all the volume relevant to the diffusion process.

The problem is slightly subtler than what it may seem at a first sight. It is not just a question of moving the impurity in a pre-existent and independent force field, because such field is directly affected by the perturbation introduced by the presence of the impurity itself. The practical implication this observation has is that it is not just a matter of sampling the energy of the system, moving the impurity around, in a relaxed lattice. The relaxation depends on the defect itself and it will be influenced, sometimes heavily, by the different local configuration of the impurity, so one needs to perform many - as many as it is considered to be necessary for a satisfactory sampling - full relaxations, which are in general very expansive from the point of view of computational resources needed. It is noteworthy that the constraint on the position of the impurity is needed: the point sampled is not in general a local minimum and, in absence of a constraint, the relaxation would minimise the energy moving the impurity to the bottom of a valley of the energy surface; in this way all the points starting in a convex region around a local minimum would relax shifting there the impurity, thus preventing us to really sample the transition states. After noticing that every calculation is actually a full structure relaxation involving many atoms, it is quite straightforward to realise that the 3D sampling of even a limited portion of material is rather unpractical: too many heavy calculations would be required. Our alternative consists in working out a different kind of constraint and relying on some reasonable assumption about the physical process that is to be investigated.

5.4.1 One-atom diffusion process

Let us suppose that we are interested in detailing the MEP and the associated barrier of an atom diffusing from initial point (A) to an end point (B) in a given system. This simple picture models satisfactorily all those processes where it can be supposed that one atom at a time displaces. After defining a set of equidistant planes perpendicular to the vector AB, we can perform

a constrained minimisation for each plane, with the impurity constrained to lie within the plane, as illustrated in Fig. (5.3). In this way the number of structural relaxations needed decreases considerably without loss of generality. The impurity, in fact, is still free to relax wherever it finds energetically more convenient along that planes, so the 3D sampling of the region that could be possibly involved in the diffusion process is preserved.

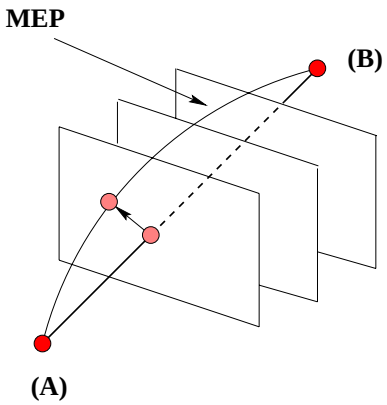


Figure 5.3: Constrained relaxation approach for one-atom diffusion processes.

We have already mentioned that the strategy of constrained relaxation calculations followed here is not guaranteed to converge to a minimum energy path (MEP) [48]. In particular, it can deviate from the MEP in regions close to saddle points, especially in the case of highly curved MEPs, resulting in a discontinuous path. Therefore, as noted above, the physical feasibility and plausibility of a deduced MEP must always be carefully checked.

Most of the diffusion processes of interest can be reduced to the problem of the motion of one atom that interacts only weakly with the surrounding lattice. With *weakly* we mean that the main effect of the motion of the impurity throughout the crystal is its local distortion, without requiring more complex exchange mechanisms [the latter case will be clearer in the discussion in Sec. (7.2.1)]. The diffusion problem of an impurity can be approached assuming that it will hop between minimum energy configurations. So, after identifying them, the above described method can be suitably employed. More in general, anyway, this procedure can be applied whenever one-atom diffusion path mus be characterised: when this is the full mechanism as well as when it is just a single step of a more complex ring mechanism, as will be discussed below.

5.4.2 Two-atom coordinated motion

The method we have just discussed, employed with the required care, is extremely useful to characterise a MEP and the associated barrier height and profile when the motion of only one atom is involved. Most of the diffusion events can be satisfactorily modelled with this simple approach. Despite the high relevance of such processes, there is a number of more complex cases where possible paths, for which feasibility and features must be checked, require to consider the concerted motion of two atoms. This can be viewed, in turn, as a co-ordinated motion of the two atoms or as a dimer rotation.

As we have discussed for the simple one-atom case, the motion of the impurity perturbs the systems, which relaxes around it. This consideration allows us to gain more insight on the real trouble we must face if two atoms are involved. Let us imagine the case of the position exchange of atom (A) and atom (B). One can think, at first, to move atom (A) toward the position of atom (B) and to analyse its MEP according to the algorithm described in Sec. (5.4.1). However, atom (A) and atom (B) must exchange their positions, so at a certain point atom (B) must leave its original position and start migrating toward the position of atom (A). It is evident that the energy surface atom (B) will experience is not independent on the position that atom (A) has in that moment and the same holds for all the rest of the exchange process.

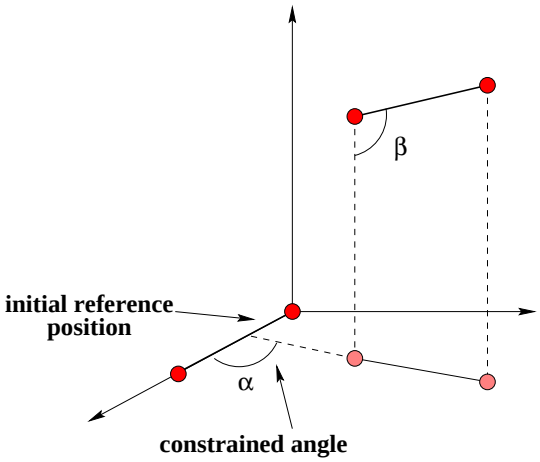


Figure 5.4: Constrained relaxation approach for dimer rotation processes.

To solve this problem we use a simple and powerful approach which obviates many of the mentioned inconveniences. Thinking of the exchange process as a dimer rotation it is easy to work out a constraint that leaves as

many degrees of freedom as we want to the system and at the same time allows us to map the total energy of the system all along the rotational space. Just as in the one-atom motion we were constraining one co-ordinate and leaving the impurity free to relax in what was left of the 3D space, here [see Fig. (5.4)] we constrain one angle of the dimer, allowing it to rotate along the other, to stretch or contract and to rigidly translate. In this scheme varying α in Fig. (5.4) from zero to π, we map out the two-atom exchange process. It is noteworthy that in this way we are not forcing the system toward any preferred directions, the only constraint is that α goes from zero to π which is a necessary condition, according to the definition of Fig. (5.4), for the dimer to rotate.

Just as in the case of the simple one-atom diffusion, the obtained MEP should be submitted to a careful validation analysis, searching for possible clues of unphysical or discontinuous diffusion path; in one way this case is even more critical than the former, simply because the atoms involved are two instead of one, on the other side the approach presented is possibly the more general, so if a plausible path is obtained, then it is likely that it is good candidate to be a true MEP.

5.5 Conclusions

In this chapter we have presented a brief review of the main theoretical tools that we will use throughout this work to extract quantitative information from the calculations (formation energy, aggregation energy) or to simulate special processes (methods to obtain MEPs). In the following chapters we will refer to the concepts and the definitions discussed in some more details here.

It is noteworthy that the methodology presented in this chapters are absolutely independent from the theoretical framework - i.e. DFT - in which we have carried out most of our calculations and they can be extended to any total energy calculation method, like tight-binding or empirical potentials.

Chapter 6

Boron diffusion in SiC

According to basic semiconductor theory, the purpose of introducing a dopant consists in using the localised, shallow level it should introduce when substituting a lattice atom. The states of the dopant can be easily thermally activated and participate in the conduction (of electrons or holes, according to the nature of the dopant). The so-called intrinsic conduction - the direct excitation of electrons from the valence band and holes from the conduction band - can be seldom achieved because of the extremely large band-gap of SiC and the considerably large temperature consequently required.

From a technological point of view, p-type doping is usually obtained by ionic implantation of Al or B species. B has proved to be an extremely interesting dopant for obtaining deep junctions in SiC, which are necessary for power device and MOSFET fabrication. Moreover, B doping results in higher quality MOS interfaces than are obtained with Al when the latter is used for semiconductor surface doping. B is the lightest among p-type dopants suitable for SiC, and this lightness results in comparatively low damage when the doping is produced with ion implantation, as is most frequently the case.

An issue that must be taken into account in binary compound semiconductors is that a dopant can interact with two different sub-lattices - in the case of SiC, Si sub-lattice and C sub-lattice - and can then substitute at two different lattice sites. The two configurations - differing for the substituted atom and the neighbouring environment - in general give rise to different local electronic structures, so that the perfect lattice band structure is modified in a different way. This aspect of B in SiC has been studied by Choyke *et al.* [49]. They demonstrated, by means of temperature dependent Hall effect admittance spectroscopy and deep level transient spectroscopy measurements, that B at a Si site gives a shallow level $\sim$ 300 meV above the valence band, while a B at a C site is responsible for a quite deeper level $\sim$ 600 meV above the valence band.

This has interesting consequences because, as we will discuss further on, the C site is energetically more convenient for a B impurity, so substitution at a C site is more stable than substitution at Si site. Thus the doping level that can be most easily activated is not the energetically favoured. This has obviously important consequences from the viewpoint of the activation efficiency.

In spite of the high interest of B as a p-type dopant, its diffusion features in SiC are still not well understood, and present some peculiarities, especially in the presence of an excess of Si. In addition, the diffusion coefficient of B in SiC is higher than that of all other dopant species (Al, N, P) currently used in SiC technology. This high diffusivity of B could be turned to advantage in the optimal design of devices, if it could be better understood.

B in Si has been studied theoretically in numerous occasions, both using first-principles electronic structure calculations [50, 51, 52, 53, 54] and semi-empirical tight-binding methods [55, 56]. Nichols *et al.* [50, 51] performed an extensive study of B, P and As, considering different configurations of the substitutional impurities interacting with interstitial Si, concluding that the diffusion of these dopants is primarily mediated by interstitials. Later, Zhu and coworkers [52] extended these results for the particular case of B, considering also the mutual interaction of two impurity atoms, and predicting that B diffusion activation occurs via a *kick-out* mechanism, through which a Si interstitial kicks out a substitutional impurity into the interstitial region, forcing it to diffuse through a series of hops from one hexagonal interstitial site to the next via a tetrahedral site. These results have been recently revised by Sadigh *et al.* [53] and by Windl *et al.* [54], who find that a substitutional B impurity is capable of capturing a Si interstitial, forming a fast diffusing complex, with no need to invoke a *kick-out* mechanism, as also confirmed by the tight-binding molecular dynamics results of Alippi *et al.* [56]. To our knowledge, no similar studies have been yet reported for the case of dopant diffusion in SiC.

6.1 Methodology

All the calculations presented in this chapter have been performed with the SIESTA code [see Sec. (3.6)], which implements Density Functional Theory (DFT) [19, 20] combined with the pseudopotential approximation to eliminate the core electrons, and uses a basis set consisting of numerical orbitals centred on the atoms (Numeric Atomic Orbitals, or NAO's).

We have approached the diffusion of B in SiC from different viewpoints, focusing each time on a particular aspect of the process. The required ac-

curacy of the calculation we set up has been adapted from time to time to their specific purpose.

As we will discuss in more details in Sec. (6.2), firstly we have studied the energetics of the different configurations: these calculations demand high accuracy, in order to carefully evaluate the relative degree of stability of the possible configurations a B impurity may take. This first set of structural relaxations set up the scene for the following calculations - the search for the Minimum Energy Path (MEP) as well as for the molecular dynamics - and therefore they must be as precise as possible, compatible with the computational loads. After that, we have deduced the MEP and its barrier through a series of constrained structural relaxations. To do this, we will have to evaluate quite many samples along a candidate diffusion path, so even if a high precision would be desirable, some high accuracy constraints can be relaxed. Finally, we have performed an extended first principles molecular dynamics. For this analysis, the energy and the forces must be evaluated many times, so it is compulsory to introduce some simplifications, to maintain the computational cost at a manageable level.

However some features are common to all the calculations presented, like pseudopotentials, mesh grid and functional parametrisation. The pseudopotentials employed are norm-conserving, and have been generated according to the Troullier-Martins [57] scheme, in the Kleinman-Bylander [34] separable form. For the B species we have used a core radius of 1.80 bohr, while for C we used 1.25 bohr, and 1.90 bohr for Si. These values assure a good transferability of the corresponding pseudopotentials. The program requires the use of a grid to compute some of the contributions to the matrix elements and total energy (in particular, those related to the exchange-correlation functional) and also for performing the Fourier transforms needed to evaluate the Hartree potential and energy by solving Poisson's equation in reciprocal space. We have used a grid fine enough to represent plane waves with kinetic energy up to 90 Ry. All our calculations have been carried out using the Local Spin Density Approximation (LSDA) with the Ceperley-Alder [26] functional as parametrised by Perdew and Zunger [25], though for a more accurate estimation of energy barriers (see below) we have also employed the Generalised Gradients Approximation (GGA) due to Perdew, Burke and Ernzerhof [58]. Only in the case of the dynamical simulations and in the tracing out of the B diffusion path was the unpolarised form of the LDA used, due to the large computational cost involved in such simulations. However, for an accurate estimation of energy barriers along the B diffusion path, we used spin polarised calculations.

The calculations described below were performed on 3C zinc-blende SiC supercells containing 64 atoms, plus the B impurity. A set of four k-points

generated according to the Monkhorst-Pack [59] scheme was used for the Brillouin zone sampling. We have performed numerical tests indicating that this set is sufficient to converge the total energy to within 0.25 meV/atom in this supercell size. The tests were performed checking the convergence of the energy difference between two representative configurations of a B impurity with respect to the number of k-points. They showed quite clearly the inadequacy of solely using the Γ point, so that at least four k-points were needed.

In order to maintain the computational costs at a manageable level, we have adapted the quality of the basis set employed to each type of calculation. For structural relaxation calculations, in which the number of energy and force evaluations is relatively small, a high quality basis set was used, consisting of double-ζ plus polarisation functions for the valence electrons of all atom types. The maximum extent of these functions was 5.965 Å, determined by an energy shift [35] parameter of 0.025 Ry. We will refer to this basis as the *full basis*. For the Molecular Dynamics (MD) simulations, where the energy and forces need to be evaluated thousands of times, we have restricted ourselves to a single-ζ basis, but retaining the polarisation orbital on both Si and B atoms. Eliminating the polarisation orbital from C atoms had only a small impact on the total energy, and thus it was not employed in these calculations. This basis will be referred to as the *reduced basis*. We have striven to design the reduced basis in such a way that, though its computational demands are significantly smaller than those of the full basis, it is still capable of providing good accuracy. We have performed tests indicating that the energy differences between two given structures as calculated with the full and reduced basis sets differ by 0.2 eV at most.

$\sim$

Using the highest level of accuracy described above (full basis, 90 Ry integration grid and four k-points) we have performed calculations on a 64-atom supercell of 3C-SiC. We obtain an equilibrium lattice parameter of 4.37 Å, and a bulk modulus of 232 GPa, which are in good agreement with previous DFT calculations using plane-waves [60, 61]. Chang and Cohen [60] obtained 4.361 Å and and 212 GPa for the lattice parameter and bulk modulus, respectively, using a plane-wave cutoff of 60 Ry, while Wang *et al.* [61] obtained 4.36 Å and 235 GPa. The experimentally measured values [62] for these quantities are 4.36 Å and 224 GPa, respectively.

6.2 Results

Recently, Bracht *et al.* [63] have performed B diffusion experiments in 4H- and 6H-SiC in which they suppress the transient enhanced diffusion of B due to the implantation damage by annealing the B implanted samples at 900 °C before subjecting them to thermal treatment at which diffusion does occur (ca. 1700-1800 °C). The measured depth B concentration profiles were interpreted by fitting to the solution of a kinetic model which contemplated either a *kick-out* mechanism or a vacancy mediated mechanism. An accurate fit to the experimental data was possible only if the *kick-out* process was assumed to be the mechanism underlying B diffusion. On the other hand, a vacancy-assisted mechanism, the other process considered, must be discarded. In that experimental framework, it was also shown that the temperatures required to activate B diffusion, typical of thermal annealing of SiC, were sufficiently high not to consider the effect of doping condition of the material and therefore of the charge state of defect.

In this Chapter, we attempt to provide a microscopic picture of the *kick-out* process on one hand, and of the interstitial diffusion of the B atom, once it has been removed from a lattice site, on the other. In the latest case, we have generalised the search for the MEP for different doping condition of the host material and therefore for different charge states of the B atom, because the interstitial diffusion can also takes place at temperatures at which the charge state of the B dopant might play a significant role. At first, we analyse in Section 6.2.1 the energetics and geometry of those interstitial B defects we are interested in; in this part we restrict our study to the neutral charge state. A study of the energetics of charged B defects has been previously presented [64, 65]. In Section 6.2.2 we discuss the diffusion of interstitial B for different doping condition of the material. Finally, in Section 6.2.3 we present a dynamical simulation of the *kick-out* mechanism.

6.2.1 Defect energetics

Firstly we have studied the substitutional configurations that B may assume, due the the high interest that such positions and their stability have from the applications viewpoint. In fact, generally speaking, it is when a dopant is substituting a lattice atom that it is more likely that it provides doping, shallow states. Coherently with experimental observations, we have found that substitutional B is highly stable. The substitution at C site - B_C - is energetically favoured to the substitution at Si site - B_{Si} - but both are far more stable than any other of the configurations we have investigated and we are to discuss in this section. As we have anticipated in Sec. (6.1), the

slightly higher stability of B_C is not the most desirable situation, because it is B_{Si} that is providing the doping state that is more easy to activate, being shallower. The reverse situation would imply a higher dopant activation rate, because most of the B will be at the lattice site giving the less deep level.

As the main goal of this study is to understand the mechanisms that rule the diffusion process at the atomic scale, from now on we will focus our attention on the configuration the impurity takes when it leaves a substitutional site and in the processes, configuration and complexes that may propitiate such an event.

In a pure and undistorted 3C-SiC network there are three types of high symmetry interstitial positions. Firstly, the hexagonal position, equivalent to that in pure Si or diamond, in which the coordination is 6; strictly speaking, in the case of SiC, the symmetry of this position is not hexagonal, as of the 6 nearest neighbour lattice atoms 3 are C atoms, while the remaining 3 are Si. As will be discussed below, this site further reduces its symmetry upon relaxation when an impurity is placed at this position. Secondly, there are two different positions with tetrahedral symmetry: one with all-Si first neighbours, and one with all-C. Furthermore, in semiconductor defect studies it has been found that *dumbbell* or split-interstitial structures, *i.e.* configurations in which a lattice atom and an interstitial share the same lattice position, are good candidates for low energy structures. Due to the stoichiometry of the lattice considered in this work, dumbbells need to be taken into account both at C and Si lattice sites. In particular, we have generated structures for $<100>$ and $<110>$ oriented dumbbells. Finally, to fully understand the interaction of B impurities with the SiC lattice, we have also considered substitutional B configurations [1], both at C and Si sites, interacting with the displaced atom (as an interstitial) at a neighbouring tetrahedral and hexagonal site. It will be useful at this stage to define a convention for labelling these different structures. The configuration obtained by placing the B impurity at the centre of a (3C 3Si) hexagon will be labelled as B_H (and H will indicate the hexagonal position itself). The tetrahedral structures with C and Si nearest neighbours will be labelled by B_{T_C} and $B_{T_{Si}}$, respectively. *Dumbbell* structures at a C lattice site will be denoted by $B_{C<100>}$ and $B_{C<110>}$, and for those at Si sites the notation will be the same, changing the sub-index accordingly. Finally, the substitutional defects will be labelled B_{Si}-Si_T for B at a Si lattice site with the displaced Si atom at a tetrahedral interstitial site, and B_{Si}-Si_H when the Si atom is at

[1] We have already clarified that *pure* substitutional B is a highly stable configuration; form now on we will refer as *B substitutional* to a B substitutional plus the substitute native atom in a nearby interstitial position, as our purpose is to study the way they interact.

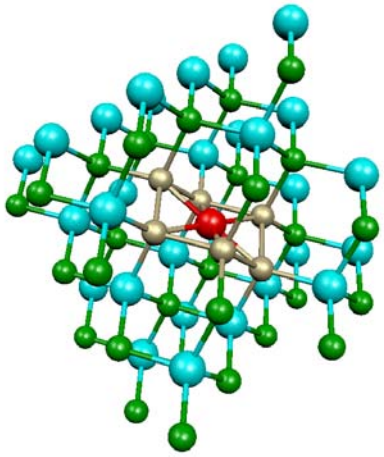


Figure 6.1: Hexagonal interstitial position (the six first-neighbours in gold colour). C atoms in green and Si atoms in cyan.

a hexagonal site; the same notation is used for B at C lattice sites, simply by changing Si for C. To distinguish between unrelaxed and relaxed defect structures, a r super-index will be used to denote a relaxed structure; for example B_H^r refers to the structure resulting after the relaxation starting from a B_H structure.

We have constructed models of all of these configurations in a 64 atom 3C-SiC supercell, plus a B interstitial, relaxing them using a conjugate gradients minimisation algorithm [66] until the forces were reduced below 0.04 eV/Å. The relaxed energies [2] obtained are given in Table (6.1). As can be seen there, the most stable configuration obtained results from relaxing the structure with the B impurity at a hexagonal site, namely the B_H^r structure. This structure is illustrated in Fig. (6.1). Although the starting H structure has six equidistant first neighbours (three C and three Si, as pointed out earlier), after the relaxation is completed, the B_H^r structure that results has the B interstitial slightly displaced from the centre of the hexagon; the three C atoms get closer to the impurity, while the Si atoms displace away from it. The distance from the impurity to the set of C atoms in the distorted hexagon is 1.68 Å, while the distance to the Si atoms is 2.00 Å. Concerning the tetrahedral structures, they all lead to local minima, but they lie much

[2] When comparing neutral systems with the same number of atoms of each species, comparing the total energies is the same than comparing the formation energies, as can be seen from Eq. (5.2)

Defect	Energy (eV)
B^r_H	0.00
$B^r_{T_C}$	0.82
$B^r_{T_Si}$	2.20
$B^r_{C<100>}$	0.51
$B^r_{C<110>}$	0.40
$B^r_{Si<100>}$	0.51
$B^r_{Si<110>}$	0.51
B_C-C^r_T	0.54
B_{Si}-Si^r_T	0.46
B_C-C^r_H	0.50
B_{Si}-Si^r_H	0.39

Table 6.1: Relative energies of the relaxed defect structures considered in this work. Energies are given with respect to the most stable structure found, namely the B^r_H structure, and were obtained using the *full basis* set and spin-polarised calculations.

higher in energy. In particular, the $B^r_{T_{Si}}$ is 2.2 eV higher in energy than the B^r_H structure, while the $B^r_{T_C}$ is 0.82 eV above. These are the least stable structures found in our calculations, and therefore it is unlikely that they play any significant role in the diffusion of interstitial B impurities in SiC lattices.

The remaining relaxed structures, both *dumbbell* and substitutional-interstitial complexes, are rather similar in energy, lying approximately 0.5 eV above the B^r_H structure. Of these, $B^r_{Si<100>}$ and B_{Si}-Si^r_H lie somewhat lower in energy, at 0.40 and 0.39 eV respectively. In fact, although we have used a different notation to label *dumbbell* and substitutional initial structures, the relaxed configurations are very similar, and it is therefore not surprising that their energies fall in the same range. This will have consequences for our dynamical studies reported below (see 6.2.3). We note in passing that the relaxed structures obtained starting from *dumbbell* configurations, in fact differ somewhat from the starting structures. In the case of the $B^r_{C<100>}$ and $B^r_{C<110>}$ structures, the C atom of the *dumbbell* continues being the first neighbour of the impurity after the relaxation, although there is significant distortion of the surrounding lattice. For the $B^r_{Si<100>}$ and $B^r_{Si<110>}$ structures, however, the impurity ends up being closer to the C lattice atoms than to the Si originally forming the *dumbbell*. With regard to the orientations, only $B^r_{C<110>}$ retains its initial orientation; in all other cases the relaxation is accompanied

by a slight rotation of the *dumbbell* axis.

Let us draw attention to the fact that the ordering of configuration energies we find here is at variance with that found in Si (see Ref. [52]). There, the most stable structure appears to be substitutional B with a nearby Si interstitial at a tetrahedral site, i.e. the structure we label B_{Si}-Si_T. It is also noteworthy that all relaxed structures turned out to have no spin polarisation.

In this study we have not considered charged configurations of the impurity or interstitials, though it is known from other studies that charged states may be important. The energetics of B in different charge states in 3C-SiC has been described elsewhere [64, 65] and the dependence of the formation energy of defects is given as a function of the chemical potential of the electron. Despite the relevance that charge state may assume under some circumstances, in this work we were mainly concerned with diffusion and we restricted ourselves to neutral defects. Bracht *et al.* [63] demonstrated that this is the relevant charge state at temperatures at which the diffusion of B and the *kick-out* process [described in Sec. (6.2.3)] is observed.

6.2.2 Boron diffusion path

Neutral B impurity

Since interstitial B is most stable at a H position, it is most natural to consider diffusion having one such configurations as the starting point. We shall consider the diffusion path between two B_H^r configurations. In order to determine the minimum energy path between start and end points, one could map the potential on a grid covering a certain volume around the path, performing constrained minimisations with the B atom fixed at each grid point. This, however, would require a large number of calculations, and we have adopted a different approach, consisting of the following. Having chosen initial (A) and end (B) points, we define a set of equidistant planes perpendicular to the vector AB, and we perform a constrained minimisation for each plane, with the B atom being constrained to lie within the plane [see Fig. (5.3)]. We have chosen a path starting at an H site, reaching another H position, second nearest neighbour of the starting one.

It is well known that the strategy of constrained relaxation calculations followed here is not guaranteed to converge to a minimum energy path (MEP)[48]. In particular, it can deviate from the MEP in regions close to saddle points, especially in the case of highly curved MEP's, resulting in a discontinuous path. This, however, did not happen in our case, indicating that the strategy adopted worked well for this system.

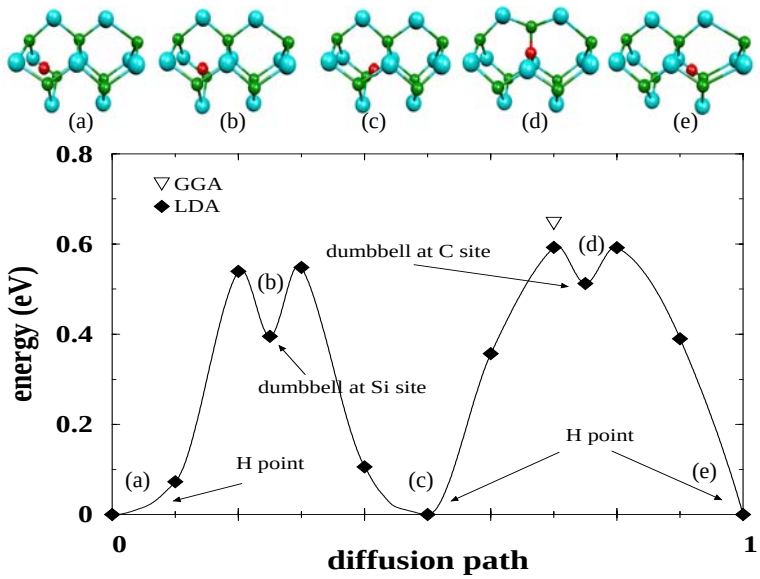


Figure 6.2: Barrier of neutral B diffusion and snapshots of some local configurations of the B impurity along the path (labels of the snapshots correspond to labels in the barrier). C atoms in green and Si atoms in cyan.

In view of the results concerning the defect energetics discussed above, where it is found that the relaxed structures have no spin polarisation, we have performed these constrained relaxations using non-spin polarised calculations. Also, we have used the LDA for these calculations, although in order to obtain a more accurate estimation of the barrier height, the lowest and highest energy points resulting from the above calculations have been relaxed again using the GGA functional due to Perdew, Burke and Ernzerhof [58] and including spin-polarisation.

Our calculations provide the energy profile along a candidate minimum energy path, which is illustrated in both panels of Fig. (6.2) with a dashed line (see also Ref. [67]), together with some configurations of the B atom and its closest lattice neighbours at specific points of the path.

As can be seen from Fig. (6.2), the energy profile consists of two barriers of slightly different heights, with a small dip at the cusp. Both barriers are situated on either side of the central point, which happens to be an H^r position. It is not surprising that the two barriers are different, given the stoichiometry of the lattice; in either pure Si or diamond they would have been the same, but here, due to the reduced lattice symmetry, they differ. In order to understand this difference, let us consider the starting structures (i.e.

previous to the relaxation). At the initial, end and middle configurations, the B interstitial has six nearest neighbours (3C, 3Si), at a distance of 1.81 Å, while at the remaining starting positions the nearest neighbours fall into two groups of 3 nearest neighbours and 3 second nearest neighbours. These intermediate positions can be classified according to the chemical identity of the three nearest neighbours. We shall label these structures using the following notation: $I_i(nSi,(3-n)C)$, which indicates a starting structure with nSi nearest neighbours and $(3-n)C$ nearest neighbours; the subindex is required to specify identical structures but with different nearest neighbour distance. Using this notation, the initial path can be described as:

$$\begin{aligned} H \rightarrow I_1(2Si,C) \rightarrow I_2(2Si,C) \rightarrow I_2(2Si,C) \rightarrow I_1(2Si,C) \rightarrow \\ H \rightarrow I_1(Si,2C) \rightarrow I_2(Si,2C) \rightarrow I_2(Si,2C) \rightarrow I_1(Si,2C) \rightarrow H \end{aligned} \qquad (6.1)$$

Note that in the first section of the path the relative abundance of Si in the first shell of neighbours of the B interstitial is higher (2 Si against 1 C), while in the second section of the path it is C that dominates. This difference turns out to be crucial for the details of the diffusion barrier profile. Configurations of type I_1 are very similar to the unrelaxed hexagonal B interstitial (B_H), while structures of type I_2 present a larger distortion with respect to the B_H configuration. Therefore, in I_1 structures the relative abundance of each type of chemical species has only a minor effect, and they relax to approximately the same energy (very similar to that of B_H^r) and to similar structures, maintaining the clear-cut separation between 1^{st} and 2^{nd} nearest neighbours of the B interstitial. On the other hand, in I_2 type structures, the relative abundance of the chemical species in the 1st shell of neighbours plays a more significant role. $I_2(Si,2C)$ structures relax by a rather small amount, and lead to stable I_2^r structures, having an energy around 0.54 eV above the H^r structure itself. $I_2(2Si,C)$ structures, on the other hand, distort significantly upon relaxation, and lead to slightly higher energy structures, in which the clear-cut separation between first neighbours and second neighbours breaks down. Rather, the B interstitial pairs with a C lattice atom, with the two Si atoms of the shell of nearest neighbours further away.

In order to better characterise the shape and height of the barriers, we have performed additional constrained relaxations to achieve a finer sampling in the region of the estimated barrier peaks. The results of these calculations lead to configurations which correspond to $B^r_{C<100>}$ and $B^r_{Si<100>}$ structures, slightly lower in energy than their neighbouring points, thus appearing as small dips at the top of the barriers. Let us emphasise at this point that the energy profile has been obtained using spin-unpolarised LDA calculations; nevertheless, in order to estimate the error in the barrier height, we have

also used spin-polarised GGA calculations at the barriers maxima, and also for the minima (namely, the B_H^r configuration), calculations which lead to a barrier height of 0.65 eV, compared to 0.59 eV with LDA. The small change is entirely due to the different functional, as once again no spin polarisation is present in the two configurations analysed.

It is interesting to note that, although the starting path has several tetrahedral interstitial sites in its vicinity, the system never relaxes to those configurations, as can be understood due to their high energy [see Table (6.1)]. Therefore, the diffusion mechanism often postulated for B in d-Si [52], and ruled out by the recent work of Windl *et al.* [54], Sadigh *et al.* [53] and Alippi *et al.* [56]. in which the impurity moves from a hexagonal configuration to another one move via a tetrahedral site, does not apply to the case of SiC either.

Charged B impurity

Up to this point, we have neglected the effect of the charge state of the impurity and we have performed all the calculations for the neutral defect. This assumption is justified on the basis of the experiments on B diffusion in 4H-SiC performed by Bracht *et al.* [63] as we have already reported [67] and discussed in this chapter's introduction. For this reason we have not carried out a detailed study of the energetics of charged B impurities, which can be found elsewhere [64, 65]. However, although the diffusion of the neutral B impurity is the relevant process in what concerns transient enhanced diffusion (TED), for the sake of generality we have investigated the effect of the charge state on the shape and height of the diffusion barrier discussed so far. Our purpose is indicating qualitatively how the doping condition of the material affects the barrier. Interstitial B does not need the high temperature typical of thermal annealing to diffuse, therefore the effect of its charge state should be taken into account. We have calculated the formation energy of all the configurations along the diffusion path in different charge states, - namely ± 2, ± 1 and 0 - according to Eq. (5.2). The total energy of the charged system has been corrected with the Madelung term [68] needed when periodic boundary conditions are employed.

Given a certain configuration of the B impurity, its most stable charge state depends on the value of the electron chemical potential μ_e, i.e. the position of the Fermi level, fixed by the doping condition of the material. This means that, given a value of μ_e, the different configurations can be found, in principle, in different charge states. In other words, as the B atom covers the diffusion path, a charge transfer might take place:

$$B_{pos\,1}^{+n} \Rightarrow B_{pos\,2}^{+n\,+m} + me^- \tag{6.2}$$

where m electrons are exchanged with the rest of the system, which can act as a charge reservoir. We have found out that for p-type doping conditions almost all the configurations are found in the 2+ state, except the two saddle points around the C *dumbbell*, which find their most stable arrangement in the 1+ charge state. For n-type doping conditions, on the other hand, the dominant charge state in the neutral, with the relevant exceptions of both the *dumbbells*, which are found in the 1− charge state. Of course the interest in B diffusion exists when B is the p-type impurity selected to dope the material, so the Fermi level will be fixed by the (shallow) acceptor level of substitutional B. From this point of view, the low μ_e case is the most relevant. In Fig. (6.3-i) we show the diffusion barrier for the case of p-type doping and in Fig. (6.3-ii) for n-type doping with a dashed line (μ_e at the valence band top and conduction band bottom, respectively). In both cases the neutral B diffusion barrier has been shown for comparison.

The interstitial diffusion barrier does not change significantly in the case of n-type doping [see Fig. (6.3-ii)] with respect to the neutral case, previously analysed. All the saddle points favour the neutral charge state and therefore the total barrier height does not change. The differences are limited to the higher stability experienced by the *dumbbells*, in the 1−, with respect to the reference B^r_H configuration. Nonetheless, this fact affects only marginally the diffusion kinetics [see Fig. (6.3-ii)], as the height of the barrier is fixed by the saddle points. For p-type doping condition, on the other hand, the most significant changes are detected just in the saddle points, thus, though the energy of the *dumbbells* themselves is not so much higher (and in one of the two cases is even lower), the nearby saddle points are. As a consequence of this fact, as can be seen in Fig. (6.3-i), both the barrier mediated by $B^r_{C\langle 100\rangle}$ and the barrier mediated by $B^r_{Si\langle 100\rangle}$ are appreciably higher. Contrary to the saddle points, the stable configurations have comparable energies with the corresponding configurations in the neutral state and one of them, $B^r_{C\langle 100\rangle}$, is even found to be almost as stable as the reference hexagonal configuration. The bottleneck is constituted by the barrier around the $B^r_{C\langle 100\rangle}$ configuration, and this grows up to 2.65 eV, then approximately 2 eV higher than for the neutral impurity. To better understand how can the barrier height be affected by the chemical potential of the electron, i.e. the doping condition of the material, we have plotted in Fig. (6.4), the dependence on μ_e of the formation energy of two representative configurations: the hexagonal and the dominant saddle point of the diffusion barrier. From that figure it is clear how the growth of the diffusion barrier in the case p-type material originates from the increased stability of the B^r_H structure for low values of μ_e. More specifically, the hexagonal configuration favours positive charge states in a much wider range than the saddle point considered, and this fact allows it

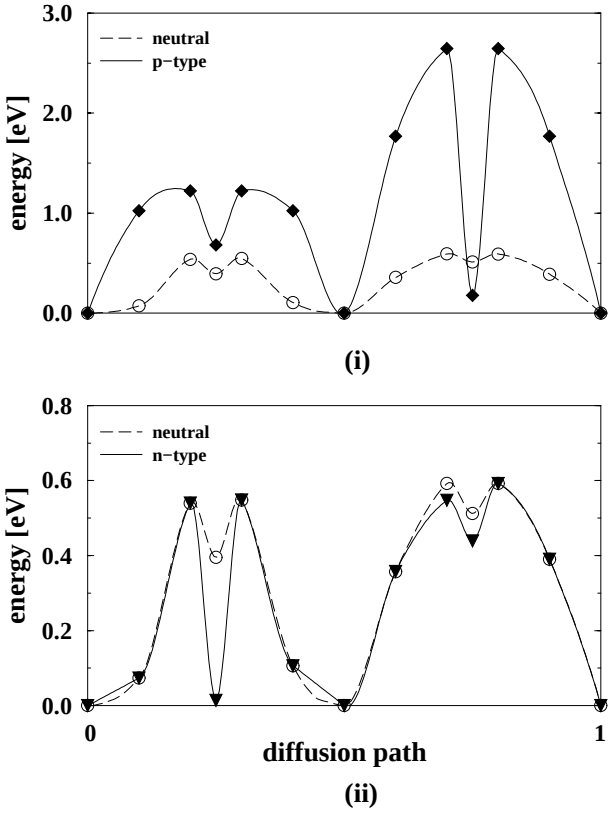


Figure 6.3: Minimum energy profile for the diffusion of a charged B impurity between two equivalent hexagonal sites under (i) p-type and (ii) n-type doping conditions. In both panels the diffusion barrier of a neutral B atom of Fig. (6.3) is shown for comparison with a dashed line.

to lower more its energy for low μ_e. On the other hand, for μ_e close to the conduction band edge, both configurations favours the neutral charge state and their difference is reduced up to approximately 0.6 eV.

The results presented so far concern the interstitial diffusion of a B impurity, but nothing has been said yet about the mechanism that activates such process, and that we are about to discuss in the next section. In doing so, we will restrict ourselves to the case of the neutral impurity, for the previously mentioned reasons [63, 67]. It should be pointed out, however, that for sufficiently low temperatures the doping condition of the material, and therefore the charge state of the defect, cannot be neglected and the barriers which are relevant in what concerns interstitial diffusion of B impurities are those we have discussed in this section.

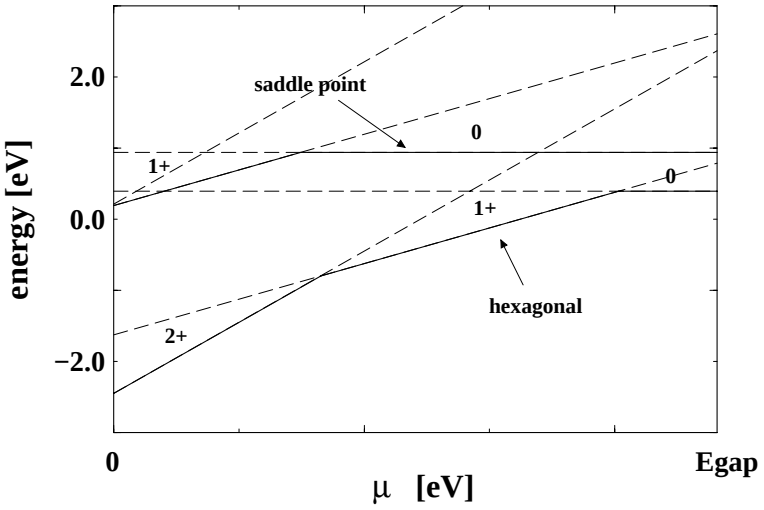


Figure 6.4: Formation energy of the hexagonal configuration and of the dominant saddle point in Fig. 6.2 as a function of the chemical potential of the electron, μ_e.

6.2.3 Molecular dynamics

The next stage of our study has consisted of an MD simulation of the dynamics of B in the SiC lattice. The aim of this study is twofold: firstly, we would like to test the concept of the *kick-out* mechanism, which has often been discussed in the context of B diffusion in d-Si [50, 51], and for which there is some experimental evidence in SiC [63]; secondly, we would like to interpret the dynamical behaviour of B as it diffuses through the lattice in the light of our static calculations of defect energies, structures and diffusion path barrier.

The equations of motion have been integrated numerically using the Verlet [38] algorithm with a time step of 2 fs, which is sufficiently small to provide good energy conservation. The total length of the run was 10 ps, and the starting configuration consisted of a B impurity at a Si lattice site, with the displaced Si atom at a neighbouring tetrahedral interstitial site, i.e. the B_{Si}-Si_T^r configuration. The initial velocities were sampled from the Maxwell-Boltzmann distribution at a temperature of 1000 K.

In this section's introduction we have already underlined that the most interesting configuration from the device operation viewpoint is B_{Si}, because this is the position where the impurity gives rise to the most shallow state, so that en electron of an occupied state in the valence band can be easily excited thermally, creating a hole. In order to focus our attention on the mechanisms

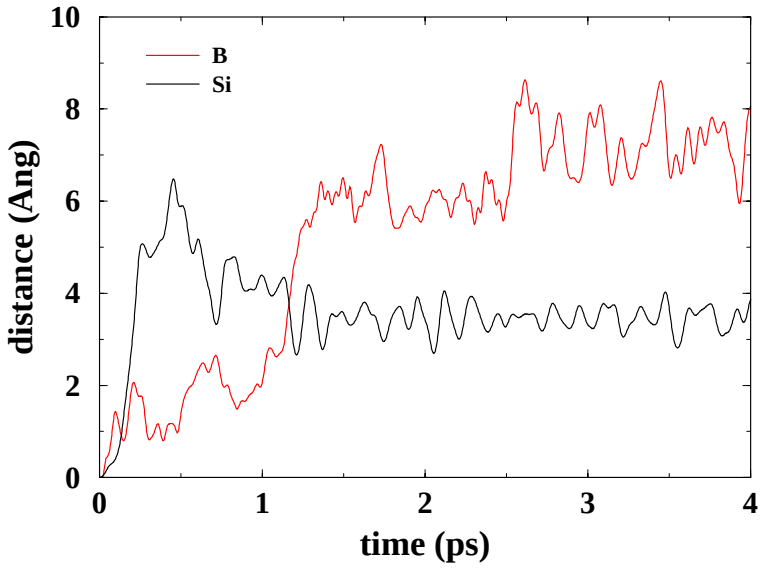


Figure 6.5: Displacement of B and the originally interstitial Si from their respective initial positions.

that may break this configuration, we have started the simulation with a B_{Si} and the corresponding native Si in a nearby interstitial position: this configuration had already been suggested by the static relaxations to be a good candidate to perturb the stability of the isolated B_{Si}. In other words, the energetics of the different configuration have indicated that a Si interstitial could rapidly destabilise a B_{Si}.

In Fig. (6.5) the distance of the B impurity and the displaced Si atom from their initial positions are plotted as a function of time. As can be seen from this figure, as soon as the simulation starts, the B atom is displaced from the substitutional site by the Si interstitial, which takes the lattice site. This is a manifestation of the *kick-out* mechanism. The first stage of the process consists of the structural relaxation to a configuration close to that of B_{Si}-Si_T, in which the Si atom has nearly returned to the lattice site, displacing the B atom from it, though it nevertheless remains attached to the Si atom, in a configuration resembling a dumbbell structure. This stage of the process occurs within 300 fs (0.3 ps) from the start of the simulation. During the remaining of the first ps of dynamics, the B atom remains loosely rotating around the Si atom, which occupies on average its equilibrium lattice site. After this time, the B atom abandons this coupling completely, and diffuses to a nearby hexagonal interstitial site, adopting, on average, a B^r_H

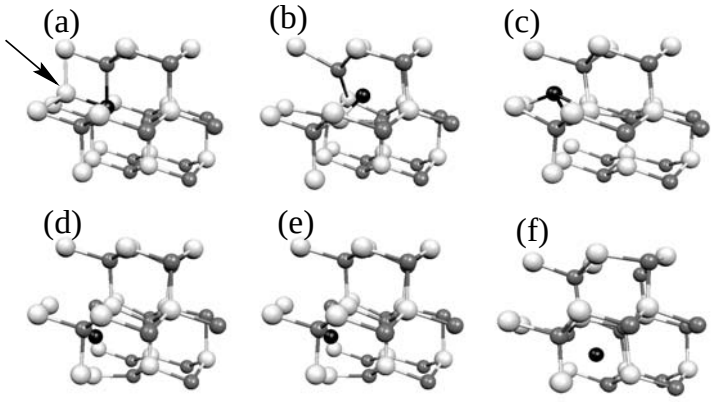


Figure 6.6: Snapshots of the dynamics of a substitutional boron impurity, illustrating the occurrence of the *kick-out* mechanism. Dark grey atoms are C atoms, and lighter grey Si atoms. Snapshot (a) corresponds to the starting configuration, with the B impurity at a Si lattice site, with the displaced Si atom at a neighbouring T site. In snapshot (b) the Si interstitial pushes the substitutional impurity out of the lattice site and into the interstitial region, taking its place. In snapshots (c) and (d) the impurity remains in close contact with the Si atom that enacted the *kick-out*, but by snapshot (e) it has moved away from it. Snapshot (f) illustrates the final configuration, with the impurity at a hexagonal interstitial site, and the Si atom occupying its lattice site.

configuration, where it remains for the rest of the simulation. This particular dynamical trajectory fits with the results obtained from the static calculations, which clearly indicate that, in the presence of a nearby Si atom, a B substitutional impurity at a T_{Si} site is unstable, and with the observation that B_H^r is the most stable configuration of a B impurity that we have found. By monitoring the time evolution of the Kohn-Sham energy, which plays the role of potential energy in the dynamics, we can confirm that there appears to be no significant energy barrier to the *kick-out* process. This can also be deduced from our earlier observation [see (6.2.1)] that the B_{Si}-Si_T structure relaxes to a configuration resembling a split interstitial in which the B atom is displaced from the Si sub-lattice position by the Si interstitial. Our results are also in agreement with the conclusion derived from the kinetic simulations and experimental work of Bracht and coworkers [63], which indicate that indeed the *kick-out* mechanism is the main responsible for the B

diffusion activation in SiC.

6.3 Conclusions

We have performed both static and dynamical first principles simulations of B impurities in the 3C-SiC zinc blende lattice. The most stable configuration we have found for the B impurity is similar to a hexagonal configuration, but with a lattice distortion which causes the B impurity to have three C atoms as first neighbours, at a distance of 1.68 Å, and three Si atoms as next nearest neighbours, at 2.01 Å. Our subsequent dynamical simulation also encounters this site, which lends weight to our conclusion that this is indeed the most stable configuration. Although our calculations have been performed for 3C-SiC, we expect that our results apply also for other phases (4H-, 6H-SiC), in view of the structural similarities between these phases.

The energies found for the T_i sites rule out the conventionally postulated mechanism for B diffusion in Si in the case of SiC, as they lie well above the energies of other possible configurations.

We have also obtained by means of constrained relaxation calculations a minimum energy path for B diffusion through the SiC lattice. The calculations indicate that the diffusion has an activation energy barrier of 0.65 eV. This path does not go through the tetrahedral interstitial sites, which are too high in energy. Due to the lower symmetry of the lattice as compared to Si or diamond, the path found here has two different barriers, of slightly different heights, which correspond to a structure close to a split interstitial with a C atom (higher barrier, ca. 0.65 eV), and a structure similar to a split interstitial with a Si atom (lower barrier of 0.55 eV). Introducing the effect of the charge state of the defect we were also able to consider how interstitial diffusion is affected under different doping conditions, finding that in p-type material the barrier is appreciably risen up to more than 2.5 eV.

Finally, we have also performed a molecular dynamics simulation illustrating how the presence of a nearby Si interstitial can cause a substitutional B impurity at a Si sub-lattice site to be displaced into the interstitial region, where it rapidly adopts the hexagonal configuration, while the Si interstitial moves to occupy the lattice site.

Chapter 7

Vacancy diffusion in SiC

One of the main drawbacks of SiC with respect to Si is that localised doping, necessary to create *pn*-junctions, can be achieved essentially only by means of ion implantation. Notwithstanding recent isolated successes [69], doping by diffusion, a quite popular technique in Si, cannot be extended to SiC because the diffusion coefficients of most of the impurities are normally very low. The obvious problem, especially if deep junctions and therefore high implantation energies are required, is that the crystal lattice is heavily damaged [see Fig. (7.1)] and thermal annealing is not always able to recover it satisfactorily.

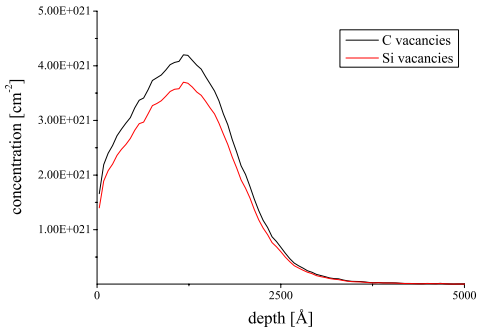


Figure 7.1: Simulation [70] of the vacancies created after an Al implantation in 4H-SiC at 150 keV with a fluency of 10^{14} cm^{-2}.

This makes clear enough why understanding the formation of defects and their kinetics in SiC is of capital importance, and a special emphasis must be put on intrinsic defects, because, regardless of the implanted species, they will always be the main result of an imperfect lattice recovery after the annealing.

In this chapter we present a theoretical study, based on first principles total energy calculations, of the diffusion of the Si and C vacancies (V_{Si} and V_C respectively) in 3C-SiC. Our calculations employed Density Functional Theory (DFT) combined with the pseudopotential approximation and they are aimed at identifying the mechanisms that are relevant to the migration of V_{Si} and V_C, characterising the minimum energy paths (MEP) and finding the corresponding barriers.

7.1 Methodology

Even though the system studied here is apparently simpler than the one of Chapter 6, lacking external impurities, the defect it features - the vacancy - is particularly difficult to handle computationally.

Vacancies and vacancy-related complexes relevant for vacancy diffusion have been studied in 128-atom 3C-SiC supercells. The larger cell size is dictated by the larger extent of the relaxation induced by a vacancy. Nieminen and coworkers [71] showed that a 512-atom supercell is required to converge the formation energy of the vacancy. However, the magnitude of this effect in SiC is lower than in Si, due to its stronger covalent bond. Moreover, for the purpose of this work, we limit ourselves to comparing the energies of very similar configurations and such relative differences have proved to converge much faster. The calculations have been performed in this case too by means of the SIESTA code [see Sec. (3.6)].

As we have discussed in some detail in Chapter 3, the size issue is particularly delicate in electronic structure calculations because, unless an order-N algorithm is used, the computational time scales with the cube of the number of atoms. The necessity to double the volume of the supercell, with respect to the system studied in Chapter 6, implies a strong worsening of the computational performance of the code and then the need for some simplifications in the calculations.

We have used the basis to which we referred as *full basis* in Sec. (6.1), but we restrict our analyses to a Γ point sampling of the Brillouin zone and no spin polarisation. However, as these factors are thought to be potentially quite important, we double checked the stable configurations by recalculating the total energy of stable relaxed configurations using *full basis* and 4 k-point according to the Monkhorst-Pack [59] scheme. The pseudopotentials and the grid as well as the form of LDA are the same to those employed in Sec. (6.1).

Structural relaxations have been performed by means of the conjugated gradient algorithm and convergence is achieved when the maximum force is smaller than 0.04 eV/Å. Except where noted, all the calculations consist of

full structural relaxations.

7.2 Results

The energetics of the vacancies has been recently studied for 3C-SiC [72] and 4H-SiC [72, 73]. In the paper by Torpo *et al.* [73] a detailed analysis of the structure of the vacancies and of the distortion induced by the Jahn-Teller effect can also be found. Here we are mainly concerned with diffusion, but nevertheless the formation energy of the two vacancies is needed. We have calculated the formation energies according to the relation given in Eq. 5.2 in the case of perfect stoichiometry and neutral charge state. We have obtained a formation energy of 3.30 eV for V_C and 7.88 eV for V_{Si} which are in good agreement with the results of Mattausch *et al.* [74]. The calculations have been checked also including a set of four k-points generated according to the Monkhorst-Pack scheme [59] for the Brillouin zone sampling. The error committed restricting the analysis to the Γ point is reasonably small for V_{Si}, where it is around the 5%, while for V_C it is 20%. In spite of this, as it is discussed later, the relative energy differences between states that are relevant to diffusion are much better converged and our conclusions are not invalidated.

In semiconductors, defect stability is generally influenced (and sometimes even strongly) by their charge state, so what holds for the neutral state does not necessarily do for positive or negative net charges. For this reason, after giving an overview of the possible diffusion mechanisms, firstly we will treat the case of neutral vacancy diffusion and secondly we will introduce the charge state of the defects, discussing their stability and if its effect on the diffusion paths is limited to modify the barriers' heights or if it acts on their shapes too.

7.2.1 Diffusion mechanisms

In a mono-species compound the mechanism of vacancy diffusion, at least in principle, is quite straightforward: a first neighbour of the vacant site can migrate and recombine with it. In the case of SiC, even for its most elemental polytype, the simplest mechanism is more complicated, because the atom which is required to recombine with the vacancy is one of its second neighbours (first neighbours belonging to the complementary species). Nevertheless, the displacement to the vacant site of one of its first neighbours is still possible and cannot be neglected. Therefore, the diffusion of one of the second neighbours of the vacant site will produce directly as a net result

the vacancy migration, while the diffusion of one of its first neighbours will create an antisite-vacancy pair. This means that we will also have to take into account that the vacancy diffusion could occur via the formation of such intermediate complexes. In other words, given that the formation of such an antisite-vacancy pair occurs, it should be investigated how it can mediate the vacancy migration process. The way the vacancy diffusion can proceed through the creation of such an antisite-vacancy pair will be discussed in more details for the case of the Si vacancy diffusion, in Sec. (7.2.3).

All our calculations consist of full structural relaxations and due to the high number of configurations analysed, the reciprocal space has been sampled only by the Γ point. In order to corroborate that this did not represent too serious an approximation, we have double checked a few representative configurations, sampling the Brillouin zone with a set of four k-points [59] and performing single configuration calculations of the optimised geometries obtained with the Γ point. In this different framework the structures are not relaxed, but the faster convergence of the total energy with respect to the forces assures us that we are not committing a large error. The barriers calculated in this way differ [discussed in Sec. (7.2.2) and (7.2.3)] at most of 0.5 eV from our results, what can be considered an acceptable error, compared with their magnitudes.

7.2.2 Diffusion of the C vacancy

Neutral C vacancy

In the case of V_C, the formation of the vacancy-antisite pair described above is not viable. The displacement of a Si to the neighbouring C vacant site would give rise to a Si vacancy / C antisite pair, i.e. $V_{Si} - \mathrm{Si}_C$, which turned out to be unstable. V_{Si} has a formation energy approximately 4.5 eV higher than V_C, therefore if the system has a recombination channel consisting in eliminating a Si_C antisite and recuperating the V_{Si}, provided that there is no barrier, it will follow it, thus reducing its energy. For these reasons, the V_C can only interact with the C sublattice. Therefore the only way for it to migrate is when one of the C second neighbours moves and recombines with it.

We have characterised the diffusion of a C atom to the vacant site by means of a series of constrained structural relaxations, adopting the strategy described in Section 6.2.2, calculating the formation energy of each configuration according to Eq. (5.3). The resulting barrier, shown in Fig. (7.2) with a continuous line, has a height of 4.08 eV and is symmetric, as expected. The trajectory is rather smooth and suffers only a slight deflection when

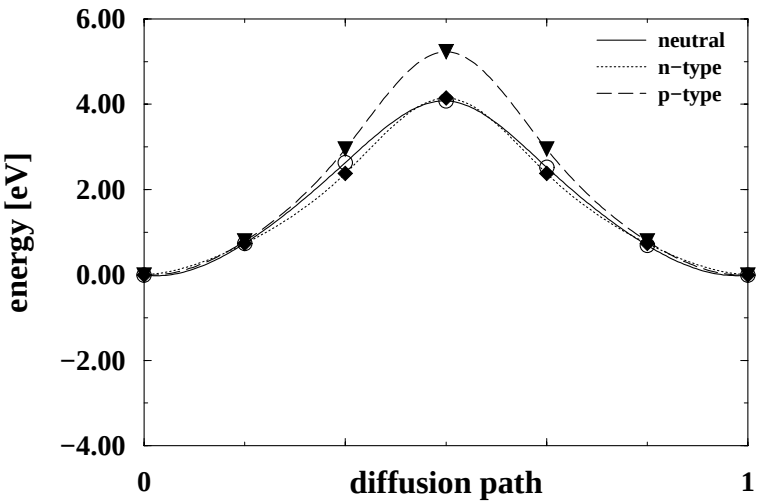


Figure 7.2: Diffusion barrier of V_C by means of a second neighbour diffusion mechanism. Charged vacancy in the case of p- and n-type doping (dashed and dotted line, respectively) and neutral vacancy (continuous line).

the C atom crosses the first neighbour shell, where it passes close to a Si, configuration to which corresponds the highest energy along the path.

Charged C vacancy

Like we did for B diffusion, after calculating the barrier for the neutral system, we have considered the dependence of the barrier on the charge state of the vacancy. The results are shown in Fig. (7.2). In p-type doping condition (dashed line), the system is always found in the $+2$ charge state and the diffusion barrier is 1.15 eV higher (5.23 eV instead of 4.08 eV in the neutral case). No significant changes in the barrier height have been detected for the n-type doping case (dotted line), where the system is always found in the -2 charge state.

7.2.3 Diffusion of the Si vacancy

Neutral Si vacancy

In contrast to the case of V_C, we have found that a V_{Si} can combine with one of its first neighbours, i.e. a C atom, leading to the formation of a $V_C - \mathrm{C}_{Si}$ pair. This configuration turns out to be stable for all the charge states analysed, because the *energetic cost* of creating an antisite is compensated

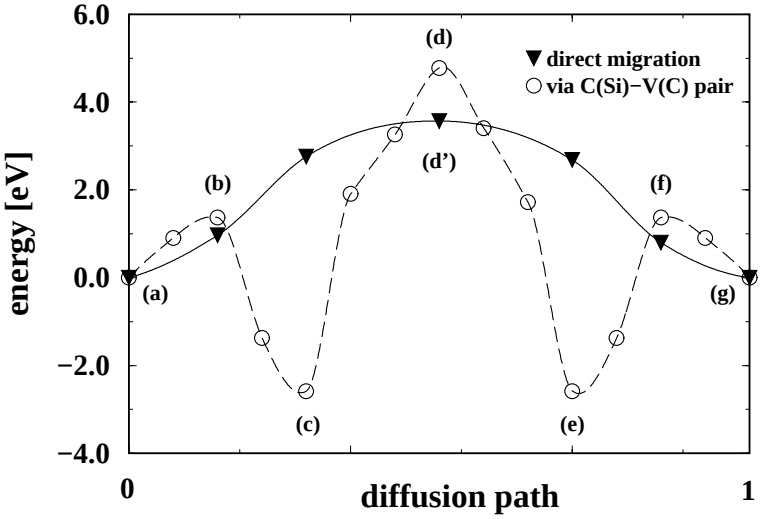


Figure 7.3: Diffusion barrier of a neutral V_{Si}. The second neighbour diffusion mechanism is represented by the continuous line. The dashed line indicates the different steps of the mediated diffusion mechanism where both the C and the Si sublattice take part to the process. (a) isolated V_{Si}; (b) barrier peak of the migration of a C first neighbour; (c) consequent formation of a $V_C - C_{Si}$ pair; (d) barrier peak of the coordinated position exchange of a Si atom and the C_{Si}; (e) a new $V_C - C_{Si}$ pair is formed; from (e) to (f) and to (g) recombination of the residual vacancy-antisite pair.

by the much lower formation energy of the V_C, compared to that of the V_{Si}. In fact, the $V_C - C_{Si}$ pair is not only a local minimum, but it is even more stable than the isolated V_{Si}. The meta-stability of the V_{Si} with respect to the $V_C - C_{Si}$ pair is a well-known fact in 3C-SiC [75, 72]. We have calculated the barrier of this transition and found that the displacing C atom has to face a relatively low barrier of 1.37 eV [as can be seen in Fig. (7.3)], and after crossing it, can occupy the vacant site, lowering the energy of the system of around 2.58 eV with respect to the isolated V_{Si}. This is in qualitative good agreement with TB calculation performed in 4H-SiC [75]. The energetics of this transition clearly indicates that it is likely to occur. From the viewpoint of V_{Si} diffusion mechanisms, it is a matter of speculating on how the resulting configuration can possibly evolve to finally produce a V_{Si} migration, with no spurious residual defects (antisites, other vacancies,...).

The first possibility that we have considered is a Si migrating and occupying the C vacant site. If this process occurred, a V_{Si} would appear back in the system, together with the neighbouring antisite pair $Si_C - C_{Si}$ that could

eventually recombine. Such a possibility can be readily discarded, because the first required transition would end up in an unstable configuration. This is not surprising at all: apart from the presence of a C_{Si}, it is the same reaction ($V_C \Rightarrow V_{Si} + Si_C$) whose feasibility was investigated (and discarded) in Section 7.2.2. It should be noted that this predicted instability has been confirmed by the calculations, as, at least in principle, one could also expect that the additional presence of a C_{Si} lowered the energy of the Si_C, providing more space.

The other possibility that we have taken into account is that a Si second neighbour of the C_{Si} could exchange its position with it. Provided that this process occurs, the V_C and the C_{Si}, could recombine, recovering the V_{Si}. The exchange process should be considered, because the resulting configuration is equivalent by symmetry to the starting one (the only difference is that the antisite C_{Si} is at a different Si site) that we already know to be stable. To characterise this process, we have adopted a constrained minimisation strategy, where the approach described in Section 6.2.2 has been generalised to the case of a 3D dimer rotation. Only one of the components of the solid angle defined by the dimer is constrained, in order to map all the rotational space spanned in the exchange process. Moreover, the dimer is allowed to stretch and compress or to rigidly translate [76]. We have obtained a barrier whose height, referred to the $V_C - C_{Si}$ configuration, is approximately 7.3 eV. Further details about the transition structures and the general geometry of the rotation were given elsewhere [76]. The transition needed to complete the V_{Si} migration is $V_C - C_{Si} \Rightarrow V_{Si}$. It does not need to be analysed, because it is the dual of the $V_C - C_{Si}$ pair creation described above. The V_{Si} diffusion mechanism mediated by the creation of a $V_C - C_{Si}$ pair that we have just discussed is sketched in Fig. (7.4) The barrier of the corresponding global process is illustrated in Fig. (7.3).

After this fairly complex case, where both the C and Si sublattices play a role in the diffusion process, we have analysed the simpler case in which the V_{Si} only interacts with the Si sublattice, in analogy to what we have done previously with the V_C. Proceeding like we did in Section 7.2.2, we have obtained a 3.57 eV symmetric barrier, with a trajectory very similar to what we have found in the case of V_C (where the diffusing second neighbour was a C atom). This plain second neighbour migration mechanism is showed in Fig. (7.5).

As it can be seen in Fig. (7.3) where the two barriers are compared, the diffusion mechanism that involves only the Si sublattice has a lower barrier. In spite of this, the $V_{Si} \Rightarrow V_C + C_{Si}$ transition is much more likely to occur and, if it takes place, it will force the diffusion process toward a higher energy path, where the $C_{Si} - Si$ exchange is required. In other words, the

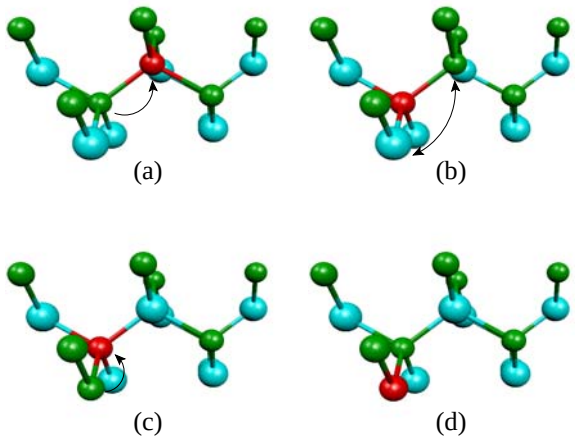


Figure 7.4: Diffusion of a V_{Si} by means the intermediate formation of a $V_C - C_{Si}$ pair [from (a) to (b)]. This mechanism requires that a Si atom and the C_{Si} exchange their position [from (b) to (c)] and that finally the residual C_{Si} and the V_C recombine [from (c) to (d)]. Grey atoms are C atoms, and cyan Si atoms; the vacancy is represented by a small red atom.

formation of the $V_C - C_{Si}$ pair will decrease the efficiency of the V_{Si} migration, preventing it from interacting only with the Si sublattice, mechanism to which corresponds the lower energy barrier.

It should be also pointed out that, given the appreciable height of the *rotation barrier*, another viable mechanism to achieve V_{Si} migration is that the process *steps back*, taking finally the direct and lower diffusion barrier. Looking at Fig. (7.3), this means avoiding the $(c) \rightarrow (d) \rightarrow (e)$ path and going through the energetically more convenient path $(c) \rightarrow (b) \rightarrow (a) \rightarrow (d^{'}) \rightarrow (g)$. Although this path could be more favoured, it is still energetically more costly than diffusion through the plain $(a) \rightarrow (d^{'}) \rightarrow (g)$ path.

Charged Si vacancy

The effect of the charge state has been studied performing full structural relaxations of all the configurations studied formerly, in the $\pm 2, \pm 1, 0$ charge state and calculating the corresponding formation energies [see Eq. (5.3)].

Let us first examine the case of p-type doping conditions of the host material. The barrier that corresponds to the direct migration path, where only the Si sublattice is involved in the process, does not suffer an important change with respect to the neutral case and we have found it to be only

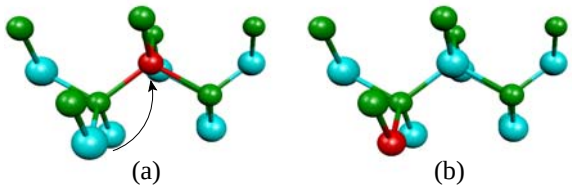


Figure 7.5: A Si second neighbour of the vacant site migrates and recombine with it. Grey atoms are C atoms, and cyan Si atoms; the vacancy is represented by a small red atom.

slightly higher, i.e. 0.32 eV higher with charge state +1 [dashed line in Fig. (7.6)] On the other hand, it is very interesting to note what happens with the mechanism involving both C and Si sublattices, mediated by the creation of a vacancy-antisite pair. The meta-stability of the isolated V_{Si} with respect to the $V_C - C_{Si}$ pair is strongly accentuated and the low reaction barrier corresponding to $V_{Si} \Rightarrow V_C - C_{Si}$ almost vanishes for $V_{Si}^{2+} \Rightarrow (V_C - C_{Si})^{2+}$, charge state in which such configurations are found in these doping conditions. Moreover, the energy difference between the resulting vacancy-

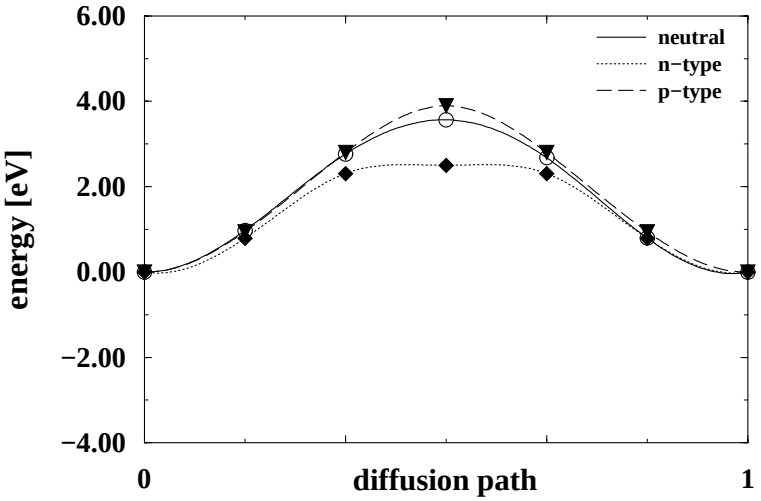


Figure 7.6: Diffusion barrier of V_{Si} by means of a second neighbour diffusion mechanism. Charged vacancy in the case of *p*- and *n*-type doping (dashed and dotted line, respectively). Neutral vacancy is shown for comparison (continuous line).

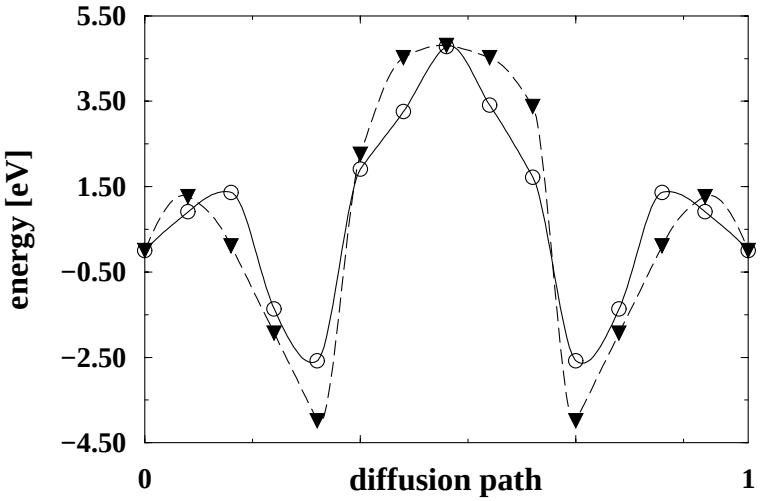


Figure 7.7: Diffusion barrier of a charged V_{Si} in the case of p-type doping (dashed line) when both the C and Si sublattices are involved. The diffusion barrier of the neutral vacancy is shown for comparison (continuous line). The vacancy-antisite pair can be much more likely formed and it is appreciably more stable. The coordinated exchange of the positions of the Si antisite pair has a higher energy barrier.

antisite pair and the isolated Si vacancy is much larger than how it was in the neutral case. A direct consequence of this fact is that the barrier for the rotation process that should follow the creation of $(V_C - \mathrm{C}_{Si})^{2+}$ grows larger and amounts to 8.78 eV (this barrier is represented by the dashed line in Fig. (7.7), where also the barrier corresponding to the neutral charge state case is shown for comparison).

When we have discussed the neutral system we have proposed the $V_{Si} \Rightarrow V_C - \mathrm{C}_{Si}$ as a channel that decreases the efficiency of the energetically more favoured diffusion path, the one that involves only the Si sublattice. Such a situation seems more likely here, because the barrier for the vacancy-antisite pair creation is lower too. In addition, when this transition takes place, the barrier required for the diffusion process to proceed (the C_{Si}-Si exchange) is now 1.42 eV higher, due to the increased stability of the vacancy-antisite pair. We can therefore conclude that V_{Si} will be a less mobile defect in p-type doping conditions.

We move now to discuss the situation of n-type doping conditions, where we have observed the opposite trend. The possibility that a vacancy-antisite pair is created, forcing the process to evolve along a higher energy path, is

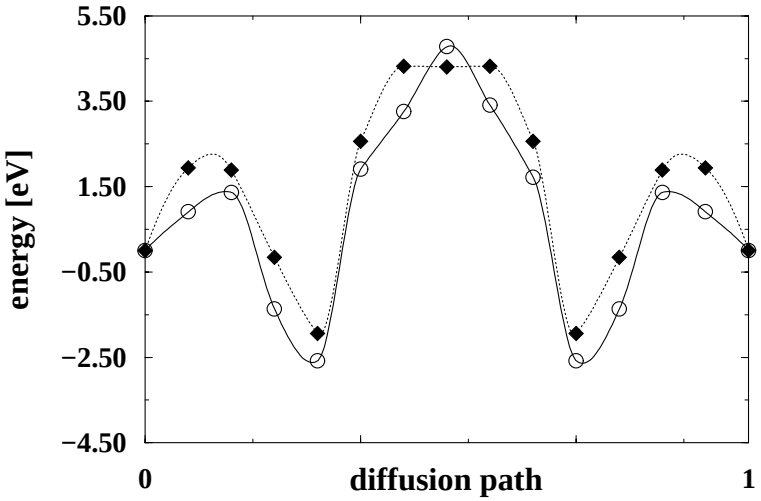


Figure 7.8: Diffusion barrier of a charged V_{Si} in the case of n-type doping (dotted line) when both the C and Si sublattices are involved. Diffusion barrier of the neutral vacancy is shown for comparison (continuous line). The barrier corresponding to the vacancy-antisite pair formation have now a height comparable to the second neighbour diffusion [see Fig. (7.6)].

still present. However the barrier of $V_{Si}^{2-} \Rightarrow (V_C - \mathrm{C}_{Si})^{2-}$ amounts to 1.89 eV and the energy gain in forming such a pair is 0.51 eV lower than in the neutral case. Finally, the rotation barrier is 1.11 eV lower [dotted line in Fig. (7.8)].

The direct migration path has a barrier which was found to be approximately 1.07 eV lower than in the neutral case, 2.50 eV instead of 3.57 eV [dotted line in Fig. (7.6)]. Such a barrier height is now comparable with the $V_{Si}^{2-} \Rightarrow (V_C - \mathrm{C}_{Si})^{2-}$ reaction barrier, therefore, not only the process is more energetically likely (lower barrier for the direct path), but it is at the same time less probable for the vacancy-antisite pair formation to act as a parallel channel. If we also consider that, even when the $(V_C - \mathrm{C}_{Si})^{2-}$ pair is formed, the following barrier is lower, we can conclude that the V_{Si} is a more mobile defect in n-type doping conditions.

7.3 Conclusions

We have analysed all the possible mechanisms of V_{Si} and V_C diffusion in cubic SiC. V_C can diffuse only by means of a direct migration of one of its second neighbours. On the contrary, for V_{Si} two mechanisms are possible. We have shown that even if V_{Si} encounters the lowest barrier in correspondence of the

direct path, the first step of the indirect path is highly probable and will force most of the times the system to evolve through it. In conclusion, even if the most favourable MEP corresponds to the diffusion of V_{Si}, its overall efficiency will be lowered by the parallel, indirect diffusion channel and the easiest diffusion will correspond to V_C.

Generalising our study the case of charged defects, we have been able to find the relevant diffusion barriers for different doping conditions. V_{Si} has resulted to be a more mobile defect in n-type material than in the prototypical neutral case study. On the other hand, in p-type material it was found to present higher barriers, being therefore less mobile. The same can be said about V_C, where n-type doping conditions seem to favour the mobility of the defect.

Chapter 8

N and P high-dose doping

In a growing wide class of device applications, achieving very high dopant densities is becoming a more and more important issue. As we discussed briefly in Section 2.3, a high doping dose results in a thinner depletion layer thicknesses (see Eq. 2.2), allowing miniaturisation of the devices. However, probably the most important application-driven reasons for which high-dose doping is extremely interesting is the realisation of high quality ohmic contacts and low resistive sources and drains in FET based devices. In spite of the interest that high-dose doping has, the solubility limit of the chosen dopant cannot always be approached without problems. In this chapter we discuss the effects of n-type high-dose doping in SiC, proposing an explanation of the experimental results based on the local electronic structure of the dopants and of dopant-complexes.

Nitrogen and phosphorus are the most commonly used n-type dopants in SiC applications. Since ion implantation is the only technologically viable way to selectively dope SiC, N has the considerable advantage of a low mass, which makes it the most typical choice as a donor species. On the other hand, P has a higher solubility limit in the SiC lattice and, as some experimental results seem to indicate, a higher electron mobility at high doses with respect to N.

As we have already recalled at the beginning of Chapter 6, the basic principles of semiconductor doping rely on the high stability of some impurity atoms when they substitute at a crystal lattice site and - especially - on their electronic properties, i.e. the donor or acceptor shallow level they provide. In a multispecies compound as SiC, then, it should be first of all asked if a given dopant is significantly most stable when it substitutes at one or the other of the two sites available, i.e. C or Si. Electron spin resonance experiments indicate that N is most favourably substituted at a C sub-lattice site [77], while P is most likely to substitute at a Si sub-lattice site [78]. This fact,

Implanted dose	Better activated dopant
$\sim 3.2 \times 10^{18}\text{cm}^{-3}$	N
$\sim 6.2 \times 10^{19}\text{cm}^{-3}$	N / P
$\sim 2.4 \times 10^{20}\text{cm}^{-3}$	P

Table 8.1: Results of the Hall effect measurements of Laube and coworkers [81].

together with the observation that N and P donors have similar ionisation energies [79, 80], suggested that co-doping of SiC with both N and P could maximise the doping efficiency, resulting in an overall increase in the dopant activation rate. The idea is that, if the N and P ions favour substitution of different host crystal species, they should be in principle able to occupy the lattice sites more efficiently, resulting in a higher average activation rate. Comparable ionisation energies of N_C and P_{Si} are then required in order to make the shallow levels they introduce indistinguishable.

In spite of the fact that the starting situation seems to be fairly clear, things turned out not to be so straightforward when the first experimental verifications were carried out. Laube *et al.* [81] have reported comparative Hall effect measurements implanting N and P, and co-implanting the two together at three different doses. They found that co-implantation did not provide systematically the higher activation rate, as expected, and rather its efficiency depended the doping dose. Their results are summarised in Table 8.1 At the highest dose, implanting P alone provided the most effective doping, followed by co-implanted N and P and then by N alone. The situation was reversed at the lowest dose, where the samples implanted with N alone were more effectively activated. It is only for the intermediate doping doses, that co-implanting appeared to be a convenient alternative to N or P doping.

The observed behaviour at low doses is most probably due to the fact that the benefit deriving from implanting a species with a lower mass are greater than those brought by a higher activation of the dopants. In other words, when the dopant concentration is far enough from the solubility limit (which is certainly true in the present case), each N atom has a huge number of C sites where to substitute. Therefore no advantage can be expected from implanting a fraction of P, whose purpose would only be to avoid such a saturation process. On the other hand, P is a heavier atom, thus the overall damage produced by the implantation would be more extended and difficult to anneal.

Less intuitive are the results that Laube and coworkers [81] obtained in

the high-dose region. It was suggested that a key role could be played by dopant-complexes whose formation becomes more and more likely as the total implanted dose grows. On this point, however, no agreement has been reached and no definitive proofs have been yet produced. For this reason, we have investigated theoretically, by means of electronic structure DFT calculations, the possible formation of passivated dopant-complexes

Further studies on N/P co-implantation have been carried out by other groups [82, 83], focusing on the benefit of co-doping from the viewpoint of the annealing of the damaged layer. On the theoretical side, N in SiC has been studied by means of first principles calculations by different groups [84, 85]; P related complexes and the deep levels they induce have been analysed by Gali *et al.* [86, 87], but a comprehensive study of P stability at different lattice sites in 3C and 4H-SiC has not yet been carried out. For these reasons the stability of N and P at the Si and C lattice site is the first issue that we will approach and only after that we will consider the possible formation of dopant complexes and the analysis of the corresponding electronic structures.

8.1 Methodology

The electronic structure calculations described in this chapter have been carried out by means of the SIESTA code. To evaluate the relative stabilities of the dopants at Si or C sites and their tendency to form aggregates, we need to employ a high-quality basis set, therefore we have used once again the *full basis* that we described in Chapter 6. An accurate representation of the charge density distribution is also required in order to describe the interaction between the dopants in the complexes that we will analyse.

The mesh grid and the pseudopotential for Si and C are the same than those described in Chapter 6 and 7. For N and P species we have used pseudopotentials with a core radius of 1.25 bohr and 1.85 bohr respectively. All the calculations were performed within the Local Density Approximation (LDA) with the parametrisation of Perdew and Zunger [25] of the Ceperley-Alder [26] functional. All the configurations, both for what concerns isolated dopant and dopant complexes, have been studied in the ± 2, ± 1 and neutral charge states. Spin-polarisation has not been systematically introduced. In the case of dopant complexes this was done because it is not strictly necessary, as the total number of electrons in the system is even. For the isolated substitutional configurations the effect of the spin polarisation was selectively checked only in the neutral charge state systems. In those test cases, the Local Spin Density Approximation (LSDA) framework was adopted.

The calculations in 3C-SiC were carried out in 64-atom supercells, while for those in 4H-SiC 96-atom supercells were used. In all the cases a set of four k-points generated according to the Monkhorst -Pack [59] scheme was used to sample the Brillouin zone.

8.2 Results and discussion

8.2.1 Isolated N and P dopants

Firstly, we have studied the electronic structure of isolated substitutional N in SiC, both at the C and Si lattice sites, namely N_C and N_{Si}, calculating the formation energies in the two cases. Our results indicate that N is much more stable at a C lattice site, along all the range of doping conditions of the material and both in the case of C-rich, Si-rich and stoichiometric SiC, as can be seen in Fig. (8.2)-a.

Concerning the ground state geometry, we have found that N_{Si} exhibits an unconventional relaxation for a substitutional, which is however well-known [88] and which is also a peculiar feature of substitutional N in diamond. What occurs is that while the most stable N_C retains the tetragonal symmetry, N_{Si} relaxes to an off-centre position, leading to a three-fold coordination with three C atoms. The off-centre relaxation has a limited impact with respect to the formation energy of the substitutional, being N_C^{off} is only $\sim 0.2\ eV$ more stable than N_C. However, the change in the electronic structure is much more relevant. As it can be seen in Fig. (8.1)-b, the off-centre geometry results in a intrinsic self-passivation of the originally shallow level that would correspond to the unrelaxed N_C.

When the concentration of N dopants is not extremely high, these considerations are not very important, because most of the N atoms will be found as N_C, as the difference in stability with N_{Si}^{off} amounts to few tenths of an eV. As we will discuss further on, the role played by N_{Si}^{off} is more important when a high doping dose might force the formation of N aggregates.

We predict that a N substitutional is most stable by far at a N_C site, and the energy difference between the N_{Si} (in the off-centre geometry) and N_C is approximately 6 eV throughout the range of doping conditions. That the N_C configuration is energetically more favourable is in agreement with experimental observations, but the difference of formation energies for N_{Si} and N_C configurations is larger than that obtained in the previous theoretical study of Miyajima *et al.* [84]. The origin for this discrepancy may be attributed to the different approximations used in Ref. [84], although a rigorous analysis is not possible since details of their calculations were not given. As expected,

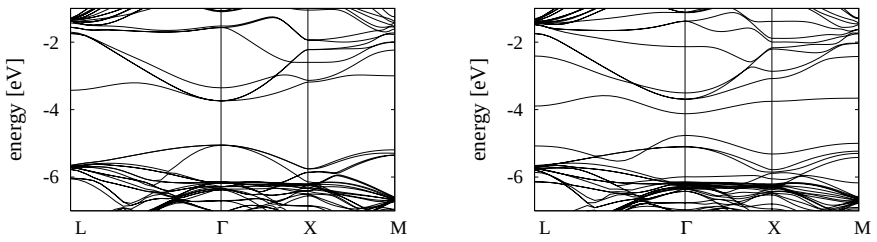


Figure 8.1: Band structure of N_{Si} (a) in the tetragonal symmetry (b) with an off-centre geometry. The N donor level that would be shallow in the case of the symmetric four-fold arrangement, falls down inside the band-gap as a results of the off-centre relaxation.

the most stable charge state of the dopant at zero temperature depends on the position of the electron chemical potential μ_e. When μ_e is below the doping level, the charge state is $+1$, while n-doping conditions lead first to a neutral state and finally to -1 state.

The case of P deserves closer attention, due to the fact that the experimental situation is far less clear than that of N. In the case of 3C-SiC, the results of our study indicate that P is most favourably located at a Si sublattice site: P_{Si}. In both substitutional sites the P atom retains the perfect lattice symmetry, with no off-centre relaxation of the dopant. The formation energy difference between P_C and P_{Si} is significantly smaller than in the case of N, and its value depends on the which is the most stable charge state of the dopant, in the different doping condition of the material.

For the $+1$ state , the energy difference is 0.68 eV. For the neutral state, the difference is reduced to 0.24 eV. The energy differences between different dopant configurations of P in 3C-SiC are much smaller than for N, as can be appreciated in Fig. (8.2) where we have plotted the formation energy of N_C and N_{Si} in panel (a) and that of P_C and P_{Si} in panel (b) vs. chemical potential of the electron. Given the reduced energy difference between the two substitutional, we have also considered the stability behaviour of P in 4H-SiC. We have not carried out the same analysis for N, as the energy difference between the two substitutionals is so large that no significant change can be reasonably expected. At least not large enough to modify the identity of the favoured substitutional. Polytypism in SiC affects the long-range periodicity of the lattice, without modifying the short-range neighbouring environment, therefore it reflects much more on the bulk properties of the material (band-gap, melting temperature,...) than on the local electronic structure of a

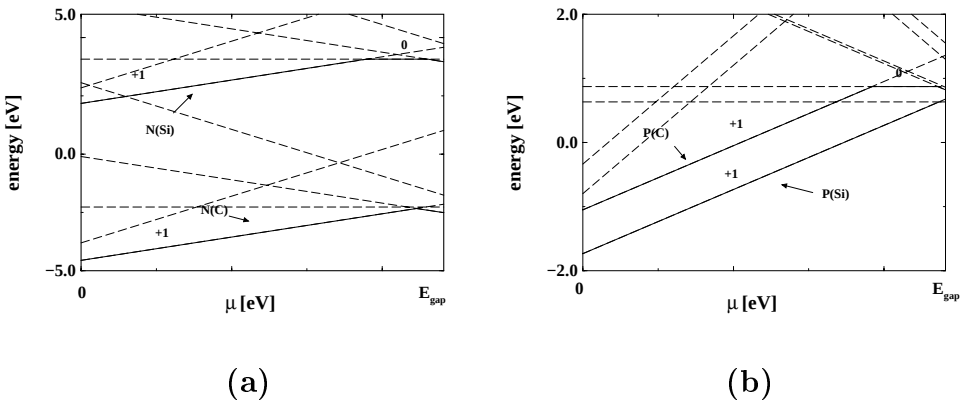


Figure 8.2: Formation energy of isolated (a) N and (b) substitutional at C and Si site in 3C-SiC. It can be seen that N favours substitution at C site, while P prefers Si site.

defect.

In the 4H polytype there are four inequivalent positions (since both Si and C can take either cubic or hexagonal lattice sites) where P can substitute. However, the difference of stability for P substituting a given host species at a cubic or hexagonal site is only of the order of hundredths of an eV. Once again, P_{Si} is more stable than P_C for all the doping conditions of the host material, the energy difference varying from 0.55 eV to 0.73 eV, depending on the value of μ_e. As in the case of 3C-SiC, at zero temperature P is found in the +1 (0) charge state for μ_e below (above) the dopant state.

The results discussed so far have been obtained assuming the condition of perfect stoichiometry, *i.e.* $\Delta\mu = 0$. We have also considered the two limiting cases of Si-rich and C-rich material. In the Si-rich case the substitution of a dopant at a C site is $\frac{1}{2}H_f$ more stable with respect to the stoichiometric case, where $H_f \approx 0.4$ eV is the SiC formation heat; in the same way, substitution at the Si site is $\frac{1}{2}H_f$ less stable, so the overall result is that the difference in stability between a dopant at a Si site and the same atom at a C site is lowered by an amount of H_f. For C-rich material the opposite is the case, so the substitution at the Si site is made more stable. These considerations are not very important for N, where the difference between the formation energy of N_{Si} and N_C is one order of magnitude larger than H_f. However, they become more relevant in the case of P, where P_C can even be made slightly more stable than P_{Si} in 3C-SiC (but not in 4H-SiC), when conditions of n-type doping and of Si-rich material hold simultaneously.

In agreement with published experimental results [77, 78], our calculations

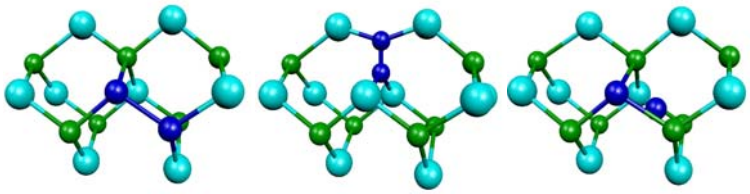


Figure 8.3: Impurity aggregates (before relaxation) selected to model the high-dose condition. Cyan balls are Si atoms, green balls are C atoms, and blue balls represent the dopant atoms. (a) substitutional dimer at nearest neighbour lattice sites; (b) split interstitial (in this particular case at a C site); (c) substitutional plus interstitial.

allow us to conclude that, in a SiC lattice, a N atom will substitute most favourably at a C site, while a P atom will do so at a Si lattice site, although in the latter case the preference can be reversed in 3C-SiC in the conditions discussed above.

8.2.2 Aggregates of N dopants

In order to focus our attention on the problems that may arise when the concentration of N dopants is very high, we have analysed the stability and the electronic structure of different N aggregates. We have considered the following cases, illustrated schematically in Fig. (8.3): (a) first neighbour substitutional, $N_C N_{Si}$ (b) split interstitial, $N_2^{C<100>}$ and (c) substitutional plus interstitial, $N_C N_H$. These complexes are intended to model some typically interesting situations from the viewpoint of dopant passivation. Assuming that a N preferably arranges into its most stable configuration, i.e. substituting at a C site, we have considered the most obvious possibilities for a further incoming N, thus supposing that it might substitute too at a lattice site [configuration (a)], it might share the same site with the other dopant, competing for it [configuration (b)], it might remain confined in an interstitial surrounding position [configuration (c)]. Malhan *et al.* [89] have used a similar approach to investigate second-nearest neighbour aggregates.

Firstly, we have calculated the tendency to aggregate, calculating the aggregation energy of these compounds. This magnitude was introduced in Chapter 5 and in the present case can be defined as follows:

$$E_{agg} = E_{complex} - nE[N_C] \tag{8.1}$$

Complex	E_{agg}
$N_C N_{Si}$	$3.74 + \Delta\mu$
$N_2^{C<100>}$	$3.08 + \frac{1}{2}\Delta\mu$
$N_C + N_i$	$3.71 + \frac{1}{2}\Delta\mu$

Table 8.2: Aggregation energies of N complexes in 3C-SiC.

where n is the number of N atoms atoms involved in the complex (so $n = 2$ in the cases we are considering). Eq. (8.1) expresses the energy gain (or loss) in gathering n dopants with respect to their most stable configuration, i.e. N_C. None of the complexes that we have taken into account exhibits a spontaneous tendency to aggregation, as can be seen in Table (8.2) where all the aggregation energy are positive. However, all the aggregates that we have considered are stable, therefore, if they form as a results of the non-thermal equilibrium conditions that the material experiences during the implantation cascade, they must overcome a barrier to dissolve. Although our simple model cannot provide information about the kinetic of the complexes formation, we think it is accurate enough to infer on their stability and their electronic structure.

We have calculated the formation energies of the considered complexes to find out which one is the most stable under the different conditions of doping of the material. The results are plotted in Fig. (8.4). For very low values of μ_e the substitutional plus interstitial configuration is dominating, but except for this short interval, the split interstitial aggregate in the neutral charge state is the most stable. All the complexes, however, fall in a quite narrow energy range, thus studying all of them is necessary to provide a thorough

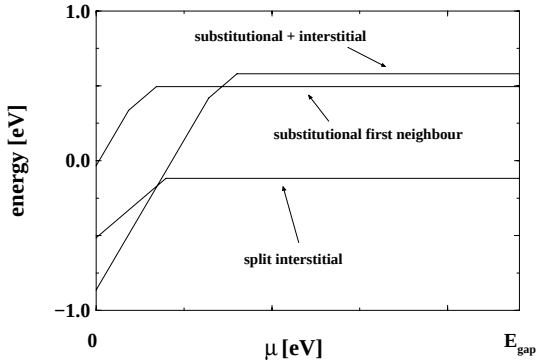


Figure 8.4: Formation energies of N aggregates as a function of μ_e.

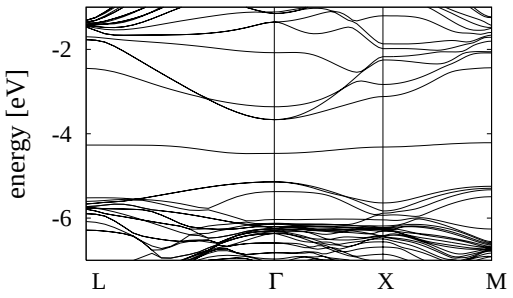


Figure 8.5: Band structure of the $N_C N_{Si}$ first neighbour substitutional.

picture of the passivation mechanisms.

First neighbour substitutional

This configuration models one of the most typical condition to which high-dose doping may lead: the efficient occupation of neighbouring lattice sites [see Fig. (8.3)-a]. We have already discussed that N favours substitution at C sites, but in the high-dose limit many of the C sites will already be occupied and there will be a growing probability that substitution at Si sites may also occur. Although the most stable configuration for a N substitutional at a Si site is N_{Si}^{off}, the plain N_{Si} is also a local minimum. To achieve the off-centre geometry, the symmetry must be *artificially* broken to observe a transition that would spontaneously occur at $T > 0$. In the case of a first neighbour substitutional dimer, the N_{Si} relaxes automatically to the off-centre geometry, because, being the N_C one of its neighbours, the perfect tetragonal symmetry is already perturbed. The interatomic distance between the two N in the relaxed geometry is quite large and it amounts to approximately 2.7 Å. As we have already discussed in Sec. (8.2.1), the N_{Si}^{off} is an intrinsically passivated configuration [see Fig. (8.1)-b], as its asymmetric geometry leads to a deep level inside the band-gap. However, in this aggregated configuration it has the significant side effect of passivating also the otherwise shallow level of the N_C. The deep N_{Si}^{off} state is only half occupied and can therefore be further populated by the loosely bound electron of N_C, leading to a full passivation of the two N dopants, as can be clearly seen in Fig (8.5).

Split interstitial

The split interstitial configuration is intended to model the competition for the same site of two N atoms, which ends up sharing it. An isolated N atom favours substitution at a C site, thus it was natural to study a split interstitial at a C site too. In such a configuration, the N atoms were found to relax to an interatomic distance of around 1.5 Å. The N_2 molecule in vacuum was calculated, finding a N-to-N length of 1.11 Å, therefore the effect of the crystal host consists mainly in an elongation of the molecule.

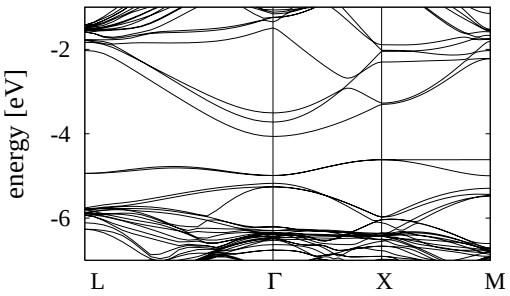


Figure 8.6: Band structure of the $N_2^{C<100>}$ split interstitial.

The band diagram of this kind of N complex also showed a deep, passivated state inside the band-gap [see Fig (8.6)], however the reduced distance between the dopants and the resemblance with the N_2 molecule suggests a different interaction. What we have here, is an ordinary plain N_2 molecule, slightly distorted by the crystal. The increased separation between the two N will lead to more loosely bound molecule, but the passivated state will essentially generate from the π states of the molecular orbitals.

Substitutional plus interstitial

The last representative configuration that we have taken into account derives from considering the perturbation effect that a N interstitial may exert on a nearby substitutional. In the high-dose limit this could be a relatively frequent case, given the rather unlikely occupation of a Si site. The interstitial N, originally located in a prefect hexagonal position, exhibits an appreciable outward relaxation, inducing a notable distortion in the surrounding host lattice. The final N-N distance amounts to 2.57 Å.

The passivation mechanism of this case closely resembles that of the first neighbour substitutional. The interstitial N_H contribute with a half populated deep level and therefore passivate the shallow level corresponding to N_C.

$\sim$

Summarising, all the N-complexes that we have considered self-passivate. However, the atomic-scale mechanisms that are responsible for the passivation are different in each case, following two patterns: the formation of a bound molecule or the charge trapping of a deep half-populated level. In this latest case, the passivating level is due in turn to the off-centre relaxation and to a N interstitial, depending on the geometry.

8.2.3 Aggregates of P dopants

We have carried out for P dopant-complexes a similar analysis than in the case of N. The complexes that we have chosen to model the high-dose condition are once again: (a) first neighbour substitutional, $P_{Si}P_C$ (b) split interstitial, $P_2^{Si<100>}$ and (c) substitutional plus interstitial, $P_{Si}P_H$ interstitial and the substitutional plus interstitial. The only difference with respect with the former case is that the split interstitial is now at a Si site and in the case of the substitutional plus interstitial configuration the substitutional is at a Si site. The reason is, of course, that if before we wanted to investigate the mechanisms that could lead to passivation of the actively doping N_C, while here the focus is on P_{Si}. The aggregation energy, defined with respect to the reference non-aggregated configuration, is now defined as:

$$E_{agg} = E_{complex} - nE[P_{Si}] \tag{8.2}$$

Once again, no aggregation seems to spontaneously occur, so the formation of these complexes must be the result of the non-thermal equilibrium condition of the material during the implantation.

A notable difference with N, is that in the case of P there is an aggregate, namely the first neighbour substitutional, which is strongly favoured with respect to the others. Another feature of P aggregates that is worth noting is that the substitutional plus interstitial configuration relaxes to a split interstitial arrangement.

In Fig. (8.7) the formation energy of these two relevant P aggregates (the first neighbour substitutional and the split interstitial) are shown. The energy difference is huge all over the electron chemical potential spectrum and amounts to approximately 3 eV, regardless of the value of μ_e and $\Delta\mu$,

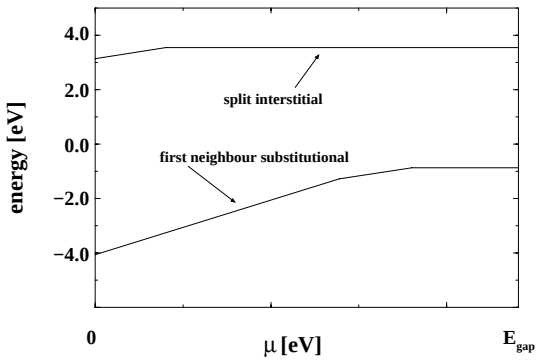


Figure 8.7: Formation energies of P aggregates as a function of μ_e.

therefore the higher stability of this complex over the others should not be questioned.

First neighbour substitutional

We have already discussed that isolated P substitutional does not suffer an off-centre relaxation like N_{Si} does. This results in a rather symmetric relaxation of the P first neighbour substitutional. The band structure of the system in the presence of this aggregate is illustrated in Fig. (8.9) and it can be clearly seen that there is no evidence of any passivating effect, as can be appreciated comparing it with the band structure of the isolated P_{Si} in Fig. (8.8). The absence of an intrinsically self-passivated level, as a result of an off-centre relaxation, is not enough to explain this situation. The fact that no deep level shows up means also that the two substitutional P rather ig-

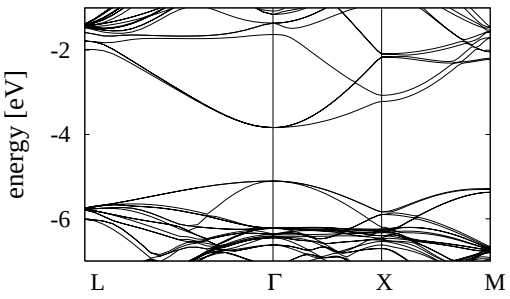


Figure 8.8: Band structure of isolated P_{Si}.

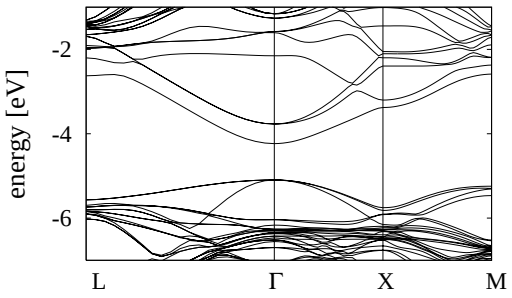


Figure 8.9: Band structure of the $P_{Si}P_C$ first neighbour substitutional.

nore each other and this is easily understood considering that both P slightly relaxes outward the perfect lattice sites, ending up with an interatomic P-P separation of about 2.0 Å.

This result has a capital importance from the viewpoint of high-dose doping. What we can conclude so far is that the P aggregate that is most likely to form in the high concentration limit, the first neighbour substitutional, is an active dopant centre and therefore the electron mobility of the material is not affected by the formation of this kind of complex.

Split interstitial

The geometrical and electronic structure features of this complex are quite similar to its N homologous. In this case a deep level inside the band-gap appears as a results of the rather short interatomic P-P distance to which the relaxation leads.

The fact that the substitutional plus interstitial configuration relaxes to the split interstitial arrangement has probably to be explained with the larger size of P with respect to N, what makes more difficult to accommodate in an interstitial position. Thus, the symmetry breaking induced by the presence of a P_{Si} favours the transition to the split interstitial configuration.

However, as we have briefly commented in the former subsection, the passivated nature of this complex does not worry us significantly, as its relevance should be by all means marginal, considering the high difference in stability with the first neighbour substitutional that makes the formation of the split interstitial rather unlikely.

8.3 Conclusions

We have shown that in the high-dose limit, when the concentration of dopants can favour their aggregation during the implantation cascade, self-passivation mechanisms can take place and significantly affect the achievable activation rate. This problem seems to be especially serious with N, where all the complexes that we have investigated exhibit self-passivation, giving rise to a deep level inside the band-gap. The passivation mechanisms range from pseudo-molecule bonding formation to the passivation induced by the off-centre relation of the isolated P_{Si}, but they all end up with a non-active doping centre. On the other hand, P seems to be almost immune by this problem, as the complex that is by far the most stable, and therefore the most likely to form, does not suffer of self-passivation and is a normal, active doping centre. The other P aggregates considered do show self-passivation, but their relevance is definitely marginal, due to the much higher stability of the first neighbour substitutional.

Therefore, we can conclude that, when approaching the solubility limit of the dopant or, more in general, for high-dose conditions, P is a more indicated n-type dopant for SiC than N and we think that this explains the experimental results of Laube and coworkers [81], according to which N and P cannot be systematically considered as a valid alternative to implanting N or P alone. In particular, N and P co-doping and N doping must be avoided when implanting high-doses.

Chapter 9

Phase coexistence in Si

Throughout this work, we have focused our attention on the study of defects. As we have remarked in Chapter 5, defects are among the most interesting features of semiconductors, as long as we consider *defects* all the perturbation of the ideal periodic crystal lattice, thus including dopants too. In the analysis of the structural and electronic properties of defects, theoretical modelling has represented a valuable support for experimental investigation, either answering or asking relevant questions, but many features of the bulk material can be approached as well from a simulation perspective.

In the early age of materials simulations, the first challenge of the different models was an accurate prediction of the bulk properties, such as the lattice parameter, the bulk modulus or the bond length. Nowadays, such analysis belongs more to standard model validation than to front edge research. In spite of this, the field related to phase transitions or, more in general, phase coexistence, cannot yet be trivially managed and has remained relatively unexplored for a long time.

The most typical example is estimating the melting temperature of a material at a given pressure. The attempt of solving this problem by running an extended molecular dynamics at constant pressure is doomed to failure. Melting, like every transition, has an associated free energy barrier [see Fig. (9.1)] and therefore, a phase transition is a thermally activated process. For these reasons, super-heating of the solid is required before the melting will be observed. Of course, if the purpose is evaluating the melting temperature, this is definitely not the most adequate technique. It can be better understood looking at Fig. (9.1) where we have pictorially represented the situation in which (a) $T < T_{melt}$, the solid state is more stable, (b) $T = T_{melt}$, the two phases are in equilibrium, i.e. we have reached the melting temperature and finally (c) $T > T_{melt}$, the liquid phase is the more stable. If we try to estimate T_{melt} by running a MD starting with the solid phase (Fig. (9.1)-a) and we

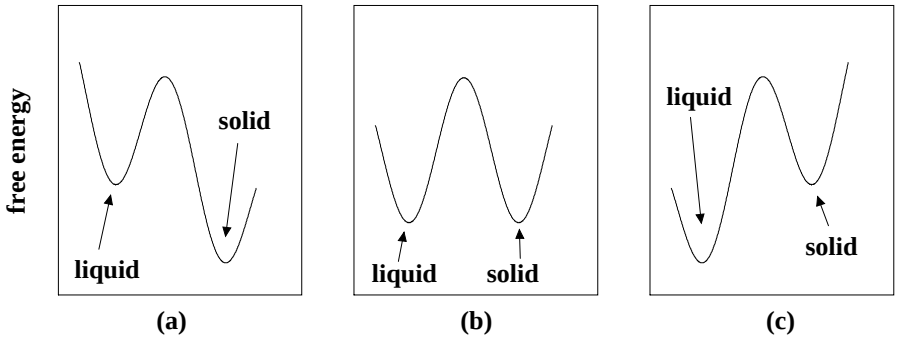


Figure 9.1: (a) $T < T_{melt}$: the solid phase is more stable; (b) $T = T_{melt}$: the solid and the liquid phase are in equilibrium; (c) $T > T_{melt}$: the liquid phase is more stable.

progressively raise the temperature of the simulation, we will not be able to do it. Even when we actually reached T_{melt}, we still would be in the situation of Fig. (9.1)-b: the two phases are in equilibrium, but the barrier between them prevents us to observe the transition in a finite simulation time. If we continue heating the system, we will progressively make the liquid phase more and more stable, in effect increasing the probability of transition, like can be appreciated in Fig. (9.1)-c.

A more accurate and reliable method consists in tracing separately the free energy of the two phases at a given pressure as a function of the temperature. The point where the two phases have the same free energy indicates by definition the temperature of phase coexistence, i.e. melting, at that pressure. The trouble with free energy is that it is an intrinsically difficult magnitude to calculate theoretically. As we will discuss in the next sections, free energy is defined as an integral over the phase space (positions and momenta of all the particles). This is at variance with energy (both potential and kinetic), temperature, volume and pressure that can be defined instantaneously and therefore averaged along a sufficiently long MD.

A well-established method to calculate the free energy of a system at a given temperature and pressure is the so-called *thermodynamic integration (TI)*. In the following section we provide a brief description of this technique, discussing benefits and trade-offs in its use for the estimation of a material melting temperature or any other phase transition.

The ambitious long term aim of this project is to track a complete phase diagram of the various phases of SiC: the liquid phase and all the stable solid phases, i.e. all the different polytypes. Despite this, the method is quite new and it needs to be extensively tested and checked, therefore we will present

here our results on silicon, which is a simpler and better defined system, and for which many reliable empirical potentials are available.

9.1 Thermodynamic integration

The Helmholtz free energy F is related to the canonical partition function as

$$F = -k_B T \log \int_V e^{-\beta U(\mathbf{r})} d\mathbf{r} + 3N k_B T \log \Lambda(T) \tag{9.1}$$

where k_B is the Boltzmann constant, T is the temperature, $\Lambda(T) = \left(\frac{h^2}{2\pi m k_B T}\right)^{\frac{1}{2}}$ is the thermal de Broglie wavelength and $\beta = \frac{1}{k_B T}$. The partition function is $Q = \int_V e^{-\beta U(\mathbf{r})} d\mathbf{r}$.

Let us assume that the internal energy U is a function of a generic parameter λ. How does the free energy change with a variation of λ? Deriving with respect to λ Eq. (9.1) we obtain

$$\frac{\partial F}{\partial \lambda} = \frac{\int \frac{\partial U_\lambda(\mathbf{r})}{\partial \lambda} e^{-\beta U_\lambda(\mathbf{r})} d\mathbf{r}}{\int e^{-\beta U_\lambda(\mathbf{r})} d\mathbf{r}}. \tag{9.2}$$

The apparently complicated expression in Eq. (9.2) is nothing more than the thermal average of $\frac{\partial U_\lambda}{\partial \lambda}$. Therefore

$$\frac{\partial F}{\partial \lambda} = \left\langle \frac{\partial U_\lambda}{\partial \lambda} \right\rangle. \tag{9.3}$$

Thus, Eq. (9.3) tells us how a variation of λ induces a change in the free energy. Thermodynamic integration (TI), taking advantage of Eq. (9.3), evaluates the free energy difference between the physical system of interest and a convenient reference system. Let us assume that U is built as a mixing of the potential of a reference system U_A and the potential of the physical system U_B:

$$U(\lambda) = (1 - \lambda) U_A + \lambda U_B. \tag{9.4}$$

As can be easily seen from Eq. (9.4), varying λ from 0 to 1, the reference known system A slowly turns into the system of interest B. At the same time, Eq. (9.4) is a generic relation of the kind $U = U(\lambda)$ assumed in the derivation of Eq. (9.3), therefore we can generalise the result, obtaining the free energy difference between the reference system A and the physical system B as:

$$F_B - F_A = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda. \tag{9.5}$$

A first limitation of this method is that there must be an available reference system for which the free energy can be tabulated or calculated analytically. However, as we will discuss further on, this problem can be overcome in practise due to the existence of appropriate systems that can serve as references (the Lennard-Jones fluid [90], the 12^{th} power potential [91], the Einstein solid). The real problem with thermodynamic integration is that it is scarcely efficient from the point of view of the required computational load. The switching from the reference system A to the system of interest B has to be conveniently slow, therefore the parameter λ in Eq. (9.4) has to vary slowly as well and, for each of its selected values, a sufficiently long MD simulation must be carried out, in order to achieve statistical significance of the results. If we thus let λ vary between 0 to 1 with a step $\Delta\lambda = 0.1$ (which may be a typical choice), 11 MD of, say, 10 ps, have to be performed to obtain an estimation of the free energy of the system of interest at *that* temperature and at *that* pressure. The process described so far should thus be iterated, varying the temperature, in order to generate a fine enough sampling along the T axis, allowing to track the $F = F(T)$ curve at the pressure of interest. It should be clear that the cost of such an iterative process makes most of the attempts to evaluate the free energy of a system along an extended range of temperatures problematic when first-principles methods are invoked.

A generalisation of TI is constituted by the *adiabatic switching (AS)* technique [92]. Like TI, this method also relies on the existence of a reversible path between the state of interest and a reference system. According to a theorem formulated back in 1910 by Hertz [93], when the Hamiltonian of a closed system is changed sufficiently slowly, the entropy of the system does not change. If the variation of the Hamiltonian is truly adiabatic, the system can be described at every time step t by a set of parameters N, V and $E(t)$ [1].

AS, like TI, exploits the slow variation of a parameter that turns the physical system of interest into a reference system for which the free energy is known. However, AS is notably more efficient than TI, because the *switching* is carried out within one single simulation.

The Helmholtz free energy as a function a parameter λ is defined as

$$F(\lambda) = U(\lambda) - TS(\lambda). \tag{9.6}$$

According to the Hertz theorem, if λ varies sufficiently slow, the process can be then considered adiabatic and therefore the entropy is conserved along the transition. The variation of the free energy is thus determined by the

[1] N number of particles, V volume and E energy.

variation of the internal energy only. Therefore

$$\Delta F = \int_0^t dt' \frac{d\lambda}{dt'} \frac{\partial U(\lambda)}{\partial \lambda} \tag{9.7}$$

Still, if we want to track the free energy dependence over a wide range of temperatures, we should iterate this process as much as necessary to achieve a satisfactory sampling of the temperature axis, but one single simulation is enough to estimate the free energy at a given temperature and pressure.

$\sim$

We have tackled the problem of free energy calculation by means of a recently proposed method [94], which, given a reliable reference for the free energy at a temperature and a pressure, allows an estimation of the free energy all over a wide range of temperatures with one single simulation. The method will be described in Section 9.2.

At variance with all the results presented in Chapter (6), (7) and (8), the calculations presented here have not be performed with SIESTA, which is a pure DFT code, but rather with the multi-algorithm multi-model package TROCADERO [95]. This package currently allows to perform different flavours of MD simulations, like SIESTA, in combination with various models, while SIESTA is build around a robust DFT core. The models implemented in TROCADERO range from different tight-binding models [96, 97, 98] to empirical potentials for semiconductors [99, 100] or metals [101, 102], and each of them can be combined with any of the available MD simulation algorithm.

9.2 Free energy calculation using Reversible Scaling simulation

The basic ideas underlying the *reversible scaling* technique [94] are as simple as highly efficient and provide the free energy over a wide temperature range without relying over uncontrolled approximations and - especially - with *one single* simulation.

Let us consider a system of N particles described by the Hamiltonian

$$H_0 = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + U_0(r_1, ..., r_N), \tag{9.8}$$

where r_i and p_i are the position and momentum vector of particle i, and U_0 is the potential energy function. We assume throughout this section the

system is in thermal equilibrium at temperature T and confined to volume V, the equilibrium volume at temperature T for the pressure of interest. The Helmholtz free energy is given by

$$F_0(T) = -k_B T \log \int_V d\mathbf{r}\, e^{-\frac{U_0}{k_B T}} \; + \; 3Nk_B T \log \Lambda(T). \tag{9.9}$$

Let us now consider the *scaled system* described by H_1, constructed from H_0 by scaling of a factor $0 < \lambda < 1$ the potential energy U_0

$$H_1(\lambda) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \lambda U_0(r_1, ..., r_N). \tag{9.10}$$

The free energy of the *scaled* system is

$$F_1(T_0, \lambda) = -k_B T_0 \log \int_V d\mathbf{r}\, e^{-\frac{U_0}{k_B T_0}} + 3Nk_B T_0 \log \Lambda(T_0), \tag{9.11}$$

where T is related to T_0 through the relation

$$T = \frac{T_0}{\lambda}. \tag{9.12}$$

Considering now Eq. (9.9), (9.11) and (9.12), it follows that the temperature dependence of $F_0(T)$ is related to the λ dependence of $F_1(T_0, \lambda)$ at temperature T_0:

$$\frac{F_0(T)}{T} = \frac{F_1(T_0, \lambda)}{T_0} + \frac{3}{2} N k_B \log \frac{T_0}{T}. \tag{9.13}$$

Therefore, the problem of calculating $F_0(T)$ for system H_0 is strictly equivalent to the problem of evaluating $F_1(T_0, \lambda)$ as a function of λ in the scaled system H_1. In other words, the calculation of the free energy of the system of interest at *various temperatures* can be transfered to the calculation of the free energy of the scaled system at *various* λ.

A problem of this kind can be approached using the standard adiabatic switching technique [92, 94] discussed in Section 9.1, which is easily implemented in molecular dynamics or Monte Carlo codes. Let us remark once more that, by means of one single simulation at T_0 of the scaled system where we vary λ, we obtain the free energy all over the range of equivalent temperatures $T = \frac{T_0}{\lambda}$, as spanned by λ. The prescriptions of AS, as we have discussed above, consist in running an MD for a time t_S, switching continuously λ from 0 to 1. If the simulated process is ideally adiabatic, the accumulated free energy difference

$$\Delta F_1(\lambda(t), \lambda(0)) = F_1(T_0, \lambda(t)) - F_1(T_0, \lambda(0)) \tag{9.14}$$

is equal to the reversible work $W(t)$ done by the generalised force $\frac{\partial H_1}{\partial \lambda}$, where

$$\Delta F_1(\lambda(t), \lambda(0)) = \int_0^t dt' \frac{d\lambda}{dt} U_0(r_1(t'), ..., r_N(t')) = W(t). \tag{9.15}$$

Combining Eq. (9.12), (9.13) and (9.14), the time dependence of the reversible work can be expressed as follows:

$$\frac{F_0(T(t))}{T(t)} = \frac{F_0(T(0))}{T(0)} + \frac{W(t)}{T_0} - \frac{3}{2} N k_B \log \frac{T(t)}{T(0)}, \tag{9.16}$$

where $T(t) = \frac{T_0}{\lambda(t)}$ and $T(0) = \frac{T_0}{\lambda(0)}$. This what de Koning, Antonelli and Yip [94] call the reversible scaling approach: each instant t along the AS simulation of system $H_1(\lambda)$ corresponds to the physical system of interest H_0 at a unique temperature $T = \frac{T_0}{\lambda}$.

In the following Section we will discuss the estimation of the melting temperature of Si. Therein we will evaluate the dependence of the free energy on the temperature, both for the solid and the liquid phase, searching for the intersection point between the two curves. That point will define the phase coexistence temperature, and therefore the melting, at the pressure at which the simulations will be carried out.

As we have said in this chapter's introduction, due to the novelty of the method employed - the reversible scaling technique - we have restricted our study to a simple and well-known material as Si. This choice allows us to use one of the many empirical potentials carefully parametrised for Si, avoiding thus the need to recur to the computationally more expensive first-principles techniques.

9.2.1 Computational methods

The potential that we have chosen is the *environment-dependent interatomic potential (EDIP)* [100], an empirical potential for Si developed by of Sydney Yip and coworkers, at the Massachusetts Institute of Technology. The EDIP potential is an evolution of the well-known Stillinger-Weber (SW) [103] potential and was provided with a higher flexibility in terms of favoured coordination. Thanks to these features, it proved to be especially accurate in describing different phases of Si, thus it is particularly suited to our purposes.

The transferability of potentials to configurations far from those used to build the parametrisation is a very delicate issue to which many efforts have been devoted in the most recent years. The EDIP potential is a rather successful attempt to achieve this aim in the case of silicon. The analogous

of the two-body and three-body terms of the SW potential here are *modulated* by a factor Z, which takes into account the local environment and the coordination of each atom. The total energy is given by the functional form:

$$\mathcal{U} = \sum_i \sum_{j \neq i} V_2(r_{ij}, Z_i) + \sum_{j \neq i} \sum_{k \neq i, k > j} V_3(\vec{r}_{ij}, \vec{r}_{ik}, Z_i), \tag{9.17}$$

where

$$V_2(r_{ij}, Z_i) = A \left[\left(\frac{B}{r_{ij}} \right)^{\rho} - e^{\beta Z_i^2} \right] e^{\frac{\sigma}{r_{ij} - a}}, \tag{9.18}$$

$$V_3(\vec{r}_{ij}, \vec{r}_{ik}, Z_i) = g(r_{ij})\ g(r_{ik})\ h(cos\theta_{ijk}, Z_i). \tag{9.19}$$

The factor $g(r) = e^{\frac{\gamma}{r-a}}$ is a pair interaction term, while the three-body term role is played by $h(cos\theta_{ijk}, Z_i)$, which is function of the angle θ_{ijk} formed by one atom and two of its neighbours. However, the distinction between two-body and three-body terms is only formal, because the EDIP potential is an intrinsically many-body potential, due to the nature of Z. The *environment dependent* term Z is defined as:

$$Z_i = \sum_{m \neq i} f(r_{im}) \tag{9.20}$$

where

$$f(r_{im}) = \begin{cases} 1 & \text{if } r_{im} < c, \\ e^{\frac{\alpha}{1 - x^{-3}}} & \text{if } c < r_{im} < a, \\ 0 & \text{if } r_{im} > a. \end{cases} \tag{9.21}$$

What Z does is *counting* the neighbours of each atom in such a way that atoms closer than c are *full* neighbours and atoms between c and a will give a fractional contribution to its coordination.

Although obvious from the equations, let us state explicitly that the major difference with the DFT computational framework, used throughout this work, is that here the electronic structure is not considered [2], making the calculations order of magnitudes computationally lighter.

Empirical potentials are often regarded as low level approximations to the real problem, but this interpretation is at least unfair. They surely are not as general as calculations from first-principles, but they are often the only viable way to approach system sizes that are relevant in experiments. Moreover, close to the specific conditions for which the parametrisation was performed, they can provide even more accurate predictions that DFT.

[2] More rigorously, we should say that it was taken into account when the parametrisation was carried out.

In this specific example, the EDIP potential represents an ideal tool for our purpose: it describes very well different phases of Si (characterised by different coordination and aggregation) and it allows us to test the novel method of reversible scaling free energy calculation requiring a limited computational load.

9.3 Results

9.3.1 Reference free energy for the liquid phase

As we have stressed in Section (9.2), the RS method needs a reliable reference free energy of the system of interest. For the liquid phase of Si, we have used conventional thermodynamic integration (see above). An adequate reference system for liquid Si where to start the TI from is the *Lennard-Jones fluid*. This system is simply a collection of Si atoms governed by a conventional Lennard-Jones potential:

$$V_{LJ} = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{9.22}$$

In our implementation we have tuned the parameter ϵ and σ so that the Lennard-Jones fluid had the same average interatomic distance than that given by the EDIP potential for liquid Si. This allows the integration between the two models to be as smooth as possible. The Helmholtz free energy of the Lennard-Jones fluid is known and is tabulated (see for instance Ref. [90]).

We have decided to perform TI of liquid Si at 2000 K, to obtain the free energy at that temperature. First we have run a 10 ps MD at constant temperature and constant pressure (NPT ensemble [3]) with the EDIP model, in order to evaluate the equilibrium volume V_{eq} of liquid Si with that potential and at that temperature. After that, we have launched an equilibration run of the system at constant volume V_{eq} and constant energy (NVE ensemble) with the Lennard-Jones potential. We have obtained in this way a set of equilibrated coordinates for the Lennard-Jones fluid, even though we have constricted the system in the equilibrium volume of the EDIP potential.

[3] In the straighter implementation of a MD algorithm, the energy E and the volume of the simulation cell V (together with the number of particles N) are conserved. This is the so-called *micro-canonical ensemble* or simply NVE ensemble. However, most of the experimental conditions do not fit this somewhat restrictive constraint, thus a variety of more sophisticated algorithms were developed to perform simulations where the temperature on one side, and the volume or the pressure on the other, are conserved, respectively the NVT and the NPT ensembles. We will not give details on these algorithms, but an excellent description can be found elsewhere [38, 39]

We have performed thermodynamic integration at 2000 K, with a step $\Delta\lambda = 0.1$. This results in 11 MD at constant volume and constant temperature (NVT ensemble). The total length of each MD was 10 ps with a time step of 1 fs. Equilibration plays a rather crucial role in these NVT calculations, that is why we have carefully carried out this task before approaching the TI MD themselves. However, this should in principle be done for every different value of λ, because a different λ gives a different *mixing* between the two potentials involved, generating by all means a different effective potential in each MD. To reduce the number of calculations required, we have followed this approach. The coordinates equilibrated for the Lennard-Jones potential have been used for the TI MD with $\lambda = 0$. This is correct, because when $\lambda = 0$ the effective potential that we are using is $U(\lambda) = \lambda U^{EDIP} + (1-\lambda)U^{LJ} = U^{LJ}$. Instead of running an equilibration MD for $\lambda = 0.1$, we have used the last set of coordinates from the $\lambda = 0$ dynamics. Thus we have proceeded iteratively in this way, using the last set of coordinates of TI MD with λ_n as a starting set of coordinates for TI MD with λ_{n+1}. Strictly speaking, we should have carried out a separate equilibration run for each λ, but we can reasonably assume that the equilibrated coordinates with λ_n would not differ significantly from those corresponding to λ_{n+1}.

From Eq. (9.5), we have evaluated $\Delta F = F^{EDIP} - F^{LJ}$ with the help of the tabulated value of F_{LJ} [90], obtaining a free energy of -5.50342 eV/atom. This is the reference value required for the RS simulation of the liquid phase.

9.3.2 Reference free energy for the solid phase

The solid phase reference can be calculated in different ways. The simplest, probably, is the so-called *quasi-harmonic approximation*, consisting in approximating the system as a collection of harmonic oscillators.

If the temperature is low enough, the approximation provides an accurate estimation of the free energy of the system, which can thus be calculated via integration of the the vibrational density of states (VDOS). In Fig. (9.2) we have plotted the phonon density of states of the Si unit cell obtained with the EDIP potential. The free energy at a conveniently low temperature can be obtained summing the energy per atom of the system in its equilibrium structure to the vibrational contribution of the free energy, which can be deduced from the calculated VDOS [104] as

$$F_{vib} = \sum_{k,j} \{ \frac{1}{2} h\nu_j(k) + k_B T \log[1 - e^{-\frac{h\nu_j}{k_B T}}] \} \tag{9.23}$$

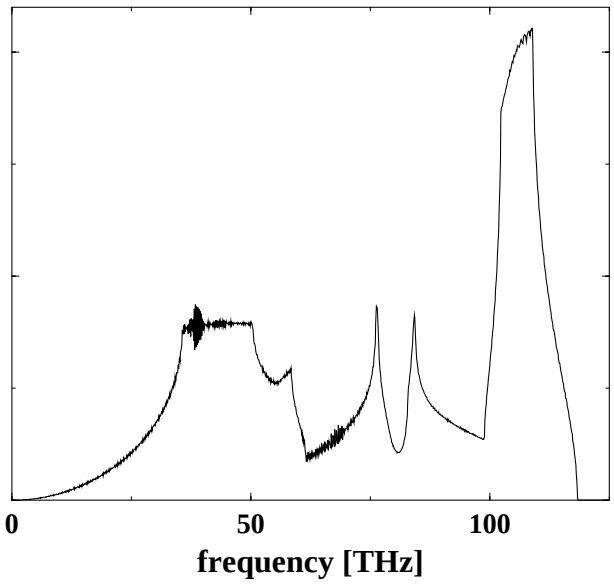


Figure 9.2: Vibrational density of states (VDOS) of Si according to the EDIP potential. We have constructed the constant force matrix by displacing in turn the two atoms in the unit cell and observing the forces induced on the atoms of a supercell. The vibrational states have been obtained calculating the eigenvalues of the dynamical matrix, which is the constant force matrix transformed to the reciprocal space. The supercell consisted of the unit cell plus one image along each direction of the lattice vectors, amounting to a total of 54 atoms (the relatively small size of the supercell is allowed by the first-neighbour extent of the potential).

where $\nu_j(k)$ are the frequencies of the vibrational modes corresponding to wave vector k and are obtained by diagonalisation of the dynamical matrix. The first term is the vibrational zero point energy.

The quasi-harmonic approximation gives a free energy per atom of -4.59972 eV at 200 K. One might think to deduce the free energy of the solid phase from Eq. (9.23) in a wider range of temperatures, but the more the temperature grows, the worse is the approximation provided by considering the crystal lattice as a collection of harmonic oscillators, so this alternative must be discarded [see also Fig. (9.3)]. However, as an alternative method, thermodynamic integration can still be used, just like in the case of the liquid system. For this reason, in order to check the reliability of the estimation of the free energy given by the quasi-harmonic approximation, we have also calculated it by means of thermodynamic integration.

In the case of a crystalline solid, a good reference is the *Einstein solid*. The Einstein solid is a system in which the atoms occupy the perfect lattice positions and are bound to them through a spring providing the necessary elastic force. Like in the case of the liquid, we have chosen a step $\Delta\lambda = 0.1$ for the thermodynamic integration and we have followed an analogous strategy for equilibration, including the optimisation of the equilibrium volume, in this case simply represented by optimisation of the lattice parameter.

The 11 NVT molecular dynamics have been performed at 200 K to compare with the results obtained with the quasi-harmonic approximation. The total length of the MD was 10 ps with a time step of 1 fs. We have obtained a free energy per atom of -4.6038 eV, indicating that the result provided by the quasi-harmonic approximation was highly accurate. We also did this for various temperatures, as can be seen in Fig. (9.3), to validate the predictions of the RS method discussed in the following section.

9.3.3 Reversible scaling simulation of the liquid and of the solid phase

So far, we have obtained good estimations of the free energy at 200 K of solid Si and at 2000 K of liquid Si. If we proceeded in the melting point evaluation in a conventional way, we should repeat the whole process for both the solid and the liquid phase at different temperatures, providing a fine enough sampling of the T axis which allowed a reliable interpolation of the data. With the RS technique, we essentially only need to carry out one MD for the solid phase and one for the liquid phase.

Before doing so, however, we have carefully equilibrated the solid system at 200 K and the liquid system at 2000 K. Free energy is a very sensible magnitude, so we should work safely far from any kind of numerical poor stability. After that, we have performed two NPT MD: one for the solid and one for the liquid of the scaled system [see Eq. (9.10) and (9.11)].

The λ parameter, in our implementation, has been scaled in such a way that during the required simulation time the temperature T [see Eq. (9.12)] was varying from 200 K to 2000 K for the solid and from 2000 K to 1000 K for the liquid. As in the case of the solid phase we are spanning a wider temperature interval, the length of the MD runs were chosen accordingly, in order to provide an equivalent sample. We ran MD of 18 ps for the solid, where 1800 K are covered, and 10 ps in the case of the liquid, to track the free energy along 10 ps.

Our results are shown in Fig. (9.3). We have estimated a melting point of Si of 1571 K, in excellent agreement with the results obtained by other groups

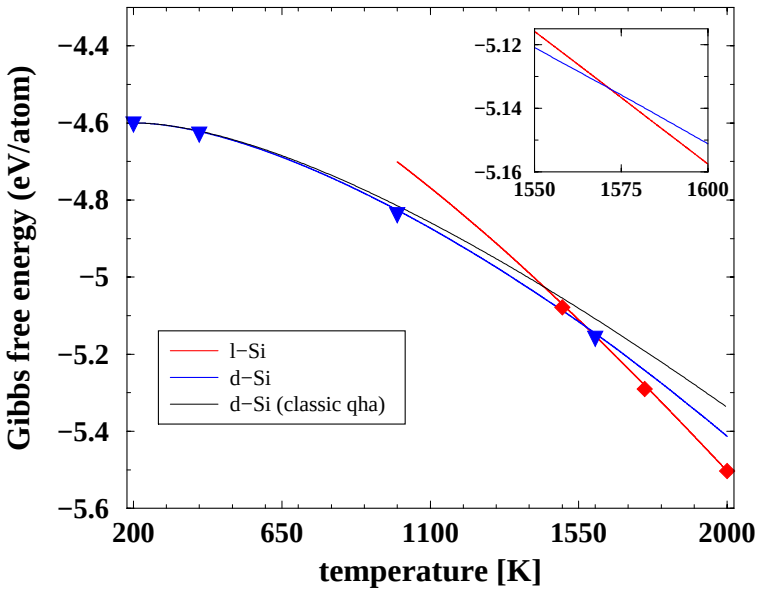


Figure 9.3: Free energy vs temperature of liquid (red curve) and solid Si (blue curve). Black line shows the temperature dependence of the free energy of solid Si according with the quasi-harmonic approximation, which proved to be very good up to approximately 750 K. The value at 200 K was used as a reference for the RS calculation of the solid phase. Diamonds and triangles show free energy estimation of the liquid and solid phase by means of thermodynamical integration. The value at 2000 K was used as a reference for the RS calculation of the liquid phase, while the value at 200 K was used to double check the estimation of the reference free energy of the solid phase obtained with the quasi-harmonic approximation.

based on the EDIP potential. Miranda *et al.* obtained ~ 1582 K with the RS technique [105] while Keblinski *et al.* obtained ~ 1550 K with an alternative MD-based approach [106]. The agreement with the experimental value of 1683 K is also considerbaly good. This powerful method has demonstrated a very high accuracy in the determination of such a subtle value, like free energy is. It is especially interesting for its efficiency, as we have frequently remarked all over this chapter, what makes perfectly reasonable to approach the phase coexistence problem from first-principles, removing the sometimes uncomfortable limitation of having to restrict ourselves to systems for which efficient empirical potentials or tight-binding parametrisation are available.

9.4 Conclusions and future developments

We have shown that the RS method can provide the free energy of liquid and solid Si with a high computational efficiency, predicting the melting temperature with a quite impressive accuracy. The immediate next step we will undertake is testing the described approach with a different model. Namely, we plan to check the melting temperature with a class of various tight-binding models [96, 97, 98]. This further step will serve to conclude the strictly methodological part of this project. This part has already been started and we can preliminarily report that non-orthogonal tight-binding [96] gives a melting temperature of Si of $\sim$ 1700.3 K, highly improving the accuracy of the EDIP potential..

The long-term application-oriented objective that we intend to investigate the phase coexistence in SiC, extracting an as accurate as possible phase digram. We previously discussed in Section 2.2 the polymorphic nature of SiC, whose solid state can be found under different polytypes. In that case, we will track the free energy dependence on temperature of the liquid phase, like in the case of Si, and in at least three different polytype of the solid phase: 3C-, 4H- and 6H-SiC. Besides evaluating the melting temperatures of the three polytypes, we will also investigate the phase coexistence condition between the three solid phases.

Cubic 3C-SiC is still difficult to grow as to produce industrial quality material, therefore any insight on the stability condition of this phase, as a function of temperature and pressure, could in principle have a high impact on the application viewpoint.

Chapter 10

The nature of D_I and D_{II} defects

Throughout this work, we have insisted several times on the fact that simulations should be carried out with the aim of explaining some experimental evidences or to help designing new relevant experiments. In this chapter we will present one of these cases, where, although experimental investigations represent the more prominent results, simulations support them, suggesting further directions to give to the experiments.

The subject of this chapter are the so-called D_I and D_{II} photoluminescence peaks, unidentified defects that bedevil SiC technology in many applications, as they affect in a systematic way the quality of the ion-implanted material. After presenting a brief introduction on these interesting defects, we will discuss some recent results of an unconventional way to perform depth-resolved photoluminescence measurements, we will analyse the insights given on the nature of the D_I and D_{II} defects [1] and the further experiments we decided to undertake in an attempt to provide a clearer picture.

$\sim$

A couple of words should be spent as a supplementary introduction to this chapter.

Due to my past experience with positron annihilation spectroscopy with polymers [107] I realised that it could be an effective experimental technique to better understand the nature of the D_I and D_{II} defects. The idea goes back to the International Conference on Silicon Carbide held in Tsukuba (Japan) in 2001, where the paper of Koshka and Melnychuck, reported in this

[1] In the following sections of this chapter, *D_I and D_{II} defects* will be slightly improperly used for brevity, meaning *defects whose signature in PL spectra are the D_I and D_{II} peaks.*

chapter, was presented. The deep-resolved photoluminescence experiments that they performed suggested me the idea of carrying out complementary experiments with a slow positron beam. I proposed the project to the group at the University of Ghent (Belgium) and we started a collaboration on this topic.

10.1 The D_I and D_{II} defects in SiC

Photoluminescence (PL) [108] is among the most widely used experimental techniques for the physical characterisation of semiconductors. It studies the luminescent radiation emitted from a sample that has been optically excited, typically with a laser beam.

When taking a PL spectrum of an ion-implanted SiC sample a huge variety of peaks appear, signatures of the different defects that have been created. A high-temperature thermal annealing ($1500-1650$ °C) is always carried out, in order to recover as much as possible the host crystal lattice and its effect is reflected in PL spectrum, where most of the peaks disappear.

However, since the first pioneering experiments of Choyke and Patrick [109, 110], the persistence of two groups of PL lines was noticed. The solid state community currently refers to them as D_I and D_{II} peaks and it has not yet been possible to provide a clear and reliable identification of the defects responsible for them. What makes the D_I and D_{II} peaks extremely interesting is that (i) they have been reported to persist up to 1700 °C [111, 112]; (ii) they appear as a results of the ion-implantation process, but this happens regardless of the type of the implanted species, which indicates that they should be attributed to some intrinsic defect-related complex. In addition, it should be underlined that D_I and D_{II} defects generate deep levels inside the band-gap, therefore being tremendously efficient charge traps. Many efforts have been devoted to their identification, but there is not yet an agreement on their nature.

It should be sufficiently clear that providing a definitive identification of these defects would have a high technological impact, as knowing their structure would probably suggests efficient annealing cycles or passivation strategies.

Recently, Koshka and Melnychuck [113] carried out a PL investigation of the D_I and D_{II} peaks on N-implanted 4H-SiC sample. They performed depth-resolved PL, combining conventional high quality PL measurements with a controlled etching technique, which allowed them to selectively access different depths of the sample. Let us describe their experimental procedure in some more detail. Firstly, they implanted with N their samples and they

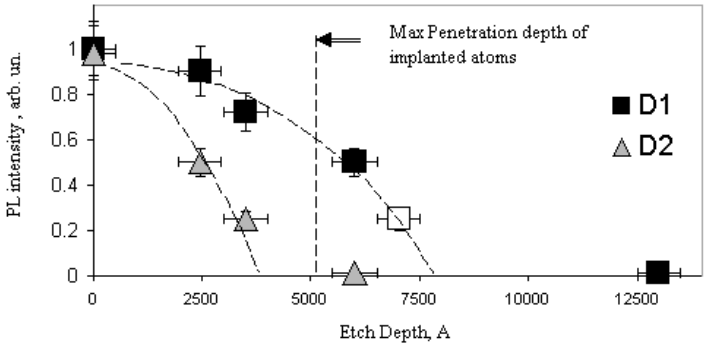


Figure 10.1: Results of depth-resolved PL [113] (figure reproduced with the permission of the authors).

performed a PL measurement. The investigated samples were 8-10 μm thick and they came from 6H-SiC epitaxial layers grown on commercial 6H-SiC substrate by chemical vapour deposition (CVD) at Mississippi State University. They were double-implanted with N at 600 °C with energies of 137 and 390 keV and doses of 3.3×10^{12} and 6×10^{12} cm^{-2}. A typical PL spectrum, where several different peaks showed up, indicates the large amount of defect formation induced by the implantation. After that, they subjected to thermal annealing the samples at 1550 °C for 30 minutes and they repeated the PL measurement, observing the well-known persistence of the D_I and D_{II} peaks. The PL technique itself cannot resolve different depths of the sample and the resulting spectra accounts for the average properties of the material. To get valuable information of the depth distribution of the defects which are the source of the D_I and D_{II} peaks, they performed a controlled, high-precision etching of the samples, removing a layer of an approximate thickness ranging from 1250 Å and 2500 Å and, after that, they repeated the PL measurement. Still the resulting spectra will show the average *defect population* of the sample, but the defects that were originally located in the removed layer obviously will not contribute anymore to the spectra. Iterating this procedure for a few times, they were able to track the intensities of the D_I and D_{II} peaks as a function of the depth. Their main findings are illustrated in Fig. (10.1), where the normalised intensities of the D_I and D_{II} PL lines vs. depth are plotted. In the same figure, also the estimated maximum penetration depth of the implanted N atoms is shown.

What is especially interesting in this figure is that the defects which are responsible for the D_I peak were found well beyond the extent of the implanted profile, while the sources of the D_{II} seem to vanish with it. Thus,

even when all the implanted layer was removed, a D_I PL peak was still clearly detectable. In other words, this experiment provided an evidence that D_I defects penetrate deeper than the implanted profile. At a first sight, this is not surprising, because, as we have already commented, D_I and D_{II} are always detected, whatever was the implanted species, thus it is not very strange that their depth distribution is not compatible with that of the implanted species. However, this still constitutes a valuable hint to identify the D_I defects, because if they were found that deep, it is likely that they penetrated, diffusing, during the annealing and therefore they should be complexes with a reasonably good mobility.

The identification of the D_I and D_{II} defects has been the subject of many investigations, theoretical as well as experimental. Although convincing proofs of it have not been presented so far, there is a general belief that D_I defects should be due to di-vacancies [110, 114] or, more in general, to vacancy-related complexes, while D_{II} peaks should be due to some kind of C di-interstitial [115].

In Chapter (7) we have discussed the mechanisms underlying vacancy diffusion in SiC, finding out that the C and the Si vacancy are not highly mobile defects, as their migration barrier are relatively high. On the other hand, the depth-resolved PL spectra of Koshka and Melnychuck indicate that, if the conventional attribution of the D_I peaks is correct, then vacancy-related complexes do diffuse, because they are found far beyond the mean penetration of the implanted ions. Unfortunately, we are not yet in a position to discard this possibility on purely theoretical grounds: we have studied the diffusion of vacancies, but we cannot exclude the formation of some fast diffusing vacancy-related complex, process to which could correspond, for not yet understood reasons, a lower barrier. On the other hand, the theoretical characterisation of vacancy diffusion suggests that it is certainly worth while further investigating the conventional attribution of the D_I defects and its compatibility with the results of Ref. [113].

10.2 Vacancy-related defect depth distribution as probed by positron annihilation spectroscopy

Positron annihilation spectroscopy (PAS) has become one of the most accurate techniques to detect vacancy-related defects in solid state matter. The basic idea underlying PAS consists in injecting positrons in the sample and studying the annihilation with their antiparticles, i.e. the electrons. The

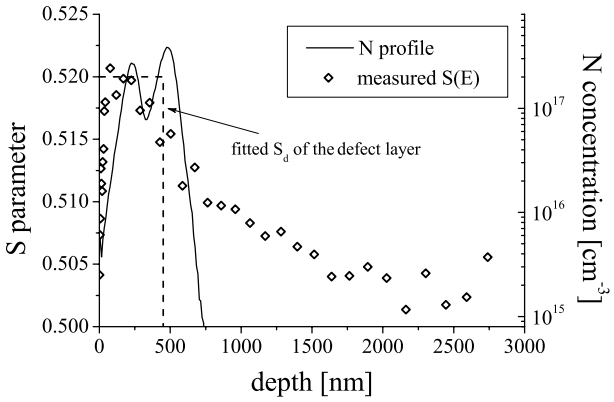


Figure 10.2: Measured $S(E)$ and fitted S_d. The simulated N implanted profile is also shown (continuous line).

study of the annihilation photopeak or of the mean lifetime of the positrons gives a valuable insight into the microscopic structure of the host material, providing information on the electronic environment around the annihilation region. The features of the electron-positron pair annihilation change significantly in the free-volume region of the material, what in solids reduces essentially to vacancies and vacancy-related complexes. On one hand, the average lifetime of the positron grows in a lower electron density; on the other hand, the annihilation peak gets sharper due to annihilation with low-momentum valence electrons, i.e. the positron annihilated in a vacancy. Therefore, PAS, in all its flavour, has become a privileged non-destructive experimental technique to detect free volume in materials and, in particular, vacancies in crystalline solids.

The great progresses made in recent years with slow positron beams [116] allowed performing depth selective analysis of the electronic and structural properties of the defects, by injecting monoenergetic positrons in the material with a well controlled mean implantation depth. Without this facility, PAS only provides information on the average free-volume of the material, what is good enough for bulk characterisation, but is certainly not sufficient for surface and sub-surface studies.

We have decided to perform PAS measurements by means of a slow positron beam in an attempt to correlate the depth distribution of vacancy-related defects to the depth distribution of the D_I or D_{II} defects as detected in Ref. [113].

Positron annihilation spectroscopy was performed at the slow positron beam at Ghent University by J. Baerdemaeker, under the direction of Prof. C. Dauwe. Measurements were taken with $\sim 3.5 \times 10^5$ counts in the photopeak. The resolution of the detector was 1.17 keV (FWHM) at 514 keV. Monoenergetic positrons with energy ranging from 100 eV to 29.1 keV were implanted in the samples, in order to probe layers up to few μm deep. The momentum of the positron-electron pair prior to annihilation induces a Doppler broadening in the annihilation photopeak. This effect is characterised by the line-shape parameter S, which is the ratio of the counts in the central region of the annihilation peak to the total number of counts in the peak.

We have simulated the N implantation by means of the Monte Carlo code *i2SiC* [70]. This code, in addition to other benefits in comparison to conventional packages like TRIM, proved to be very reliable in reproducing the tails of the profile, if any, due to ion channelling. The resulting profile well compares with what Koshka *et al.* presented originally [113], presumably obtained with TRIM [2], and that we have reproduced in Fig. (10.2). The N distribution exhibits two peaks, one at 235 nm and another at 475 nm, with a small dip between them. It vanishes rapidly around 600-700 nm and it is practically extinguished at 750 nm, so that no N impurity is likely to be present beyond that depth. Unfortunately, the detection limit of SIMS for N does not allow to experimentally confirm the absence of a N tail beyond 750 nm. Nitrogen, like many other dopants, has a very low diffusion coefficient in SiC, therefore the shape of the implanted profile does not change significantly after the annealing and the results of our simulation, obtained for the as-implanted sample, hold equally well also for the annealed sample.

To obtain a depth dependent profile of the lattice damage defects, the measured value of the line shape parameter $S(E)$ has to be deconvoluted into its elementary components. We have done it with the VEPFIT package [117], the most common code to solve the diffusion equation of the positrons. VEPFIT assumes a box-shaped layer structure for the defects, which means that it fits one value S_i of the S parameter for each defect layer i, so that S varies as a function of depth as a step-like function [see Fig. (10.3)]. In other words, VEPFIT associates to each layer a characteristic defect structure, represented by the fitted S_i.

[2] TRIM is a well-known successful package that was developed and parametrised for ion implantation in Si. Its *extension* to SiC does not rely on a new parametrisation for this material, but it only takes into account the higher density with respect to Si. In other words, when applied to SiC, what TRIM actually simulates are implantation in a *denser* Si. Under some conditions, this rough approximation is sufficiently good, but under some other - like when ion-channelling plays a prominent role - it is not.

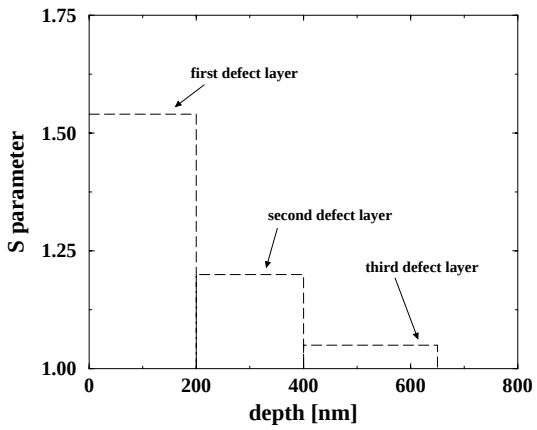


Figure 10.3: Box-shaped layer structure for the defects assumed by the VEPFIT code. In this fictitious example, three defect layers have been fitted, each one characterised by an S parameter. As a general rule, fits with a growing number of layers should be tried to be sure that the layer structure is fine enough to reproduce the variation of the S parameter. In spite of this, in conventionally implanted-semiconductors from one to three layer are sufficient to obtain a satisfactory fit.

The VEPFIT analysis allows us to compare the extent of the damaged layer firstly with the implanted species profile, and secondly with the depth distribution of the D_I and D_{II} PL intensities [113]. The measured parameter is given by:

$$S(E) = S_s F_s(E) + S_e F_e(E) + \sum_i^N S_{di} F_{di}(E) + S_b F_b(E) \qquad (10.1)$$

where S_s, S_{di}, and S_b are the S parameters associated with annihilation of positrons with electrons at the surface, in layers containing a typical defect structure and in the bulk material respectively; S_e is the contribution of the annihilation of the epithermal positrons. $F(E)$ indicates the fraction of positrons annihilating in each state. The analysis was carried out assuming different numbers of defect layers - *i.e.* N in Eq. (10.1) -, but the most satisfactory fit was obtained assuming $N = 1$, i.e. one single damage layer. The VEPFIT analysis needs the diffusion length L_b of the positrons in defect-free SiC in order to properly resolve the different components of $S(E)$. L_b is the average distance that a thermalised positron can diffuse into the bulk material before annihilating. Usually, for semiconductors this value ranges between 200 nm and 250 nm (240 nm is the value for Si). In order not to

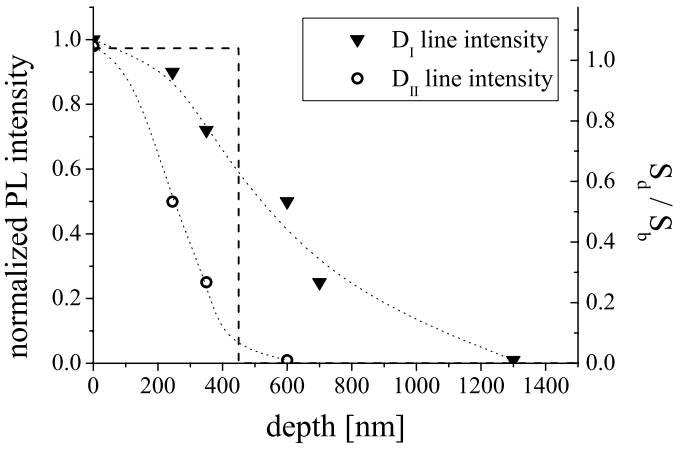


Figure 10.4: Depth distribution of D_I and D_{II} normalised PL intensities [113] and fitted layer thickness of the vacancy-related defects.

L_b [nm]	S_d	L_d [nm]	d_d [nm]	S_b
200	0.520	25 ± 2	470 ± 20	0.502
250	0.520	26 ± 3	430 ± 20	0.502

Table 10.1: Results of the VEPFIT analysis for different values of the defect-free diffusion length L_b. S_d is the fitted S parameter, d_d is the damaged layer thickness and L_d the diffusion length that the positrons have in it.

impose too restrictive assumptions on it, we have analysed the results in both the limiting cases, leaving all the other parameters free. In Table (10.1) we have summarised the results obtained. The analysis clearly indicates the existence of one single damaged layer whose thickness ranges from 430 to 470 nm. In Fig. (10.2) we show the extent of the damaged layer in the case of $L_b = 250$ nm, together with the implanted N distribution for comparison.

The value of S_d (the fitted value of the measured S parameter in the damaged layer) must be normalised to its bulk value in order to allow us to gain an insight into the nature of the defect structure. One bulk value that can be used is the one resulting from the fit itself (0.502), however we have also measured a virgin reference sample obtaining a value of S_b for SiC of 0.500 ± 0.001. Using the fitted value we have a S_d/S_b ratio of 1.035, which grows to 1.04 if we use the S_b of the reference sample. Though a di-vacancy in GaAs and in Si has typically such a signature [118], most authors suggest

that the $V_{Si}V_C$ di-vacancy in SiC gives rise to a ratio of 1.05 - 1.06 [119]. On the other hand, mono-vacancies have been recently reported to give rise to a S_d/S_b ratio of 1.033 - 1.04 [120, 121, 122], in better agreement with our results.

However, experiments indicate that after a $\sim$ 1550 °C annealing of samples implanted with a fluency comparable to the one used in this work, isolated point defects vanish [123], although not all the damage can be removed [124, 125, 126]. Anwand *et al.* [125, 126] measured a S_d/S_b ratio similar to ours and proposed that it might be due to dislocation loops. Nonetheless, more experimental work should be carried out to provide further evidences supporting this possibility.

It must be remembered that the VEPFIT analysis assumes box-shaped layer structures for the defects, therefore the S_d value that we have obtained for the only damaged layer fitted by the program is an average of the defects present therein. In other words, we cannot state that all di-vacancies have been annealed, but at least they are not the dominant type of defect, as indicated by a too low value of S_d/S_b.

Although the identification of the defects in the damage layer is not completely clear, our main concern in this work was to see if the depth distribution of the D_I and D_{II} PL intensity was compatible with the extent of the positron detectable defect layer. In Fig. (10.4) we have illustrated the depth distribution of the normalised D_I and D_{II} PL lines [113], together with the fitted S_d/S_b ratio of the damaged layer detected by slow positron beam measurements. It can be observed that the depth distribution of the D_I PL line is significantly different from the depth distribution of the vacancy-related complexes, including di-vacancies that are often proposed to be the origin of the D_I emission. It is possible that not all the di-vacancies have been annealed and the S_d value that we observe in the damaged layer comes from a mixture of di-vacancies and dislocation loops (or other not yet better identified defects), but it is clearly demonstrated that vacancy-related defects (including di-vacancies) are not found deeper than 430-470 nm, in contrast to the detected behaviour of the D_I defects. On the other hand, the D_{II} line, that is normally attributed to some kind of di-interstitial, exhibits a depth distribution that has good correlation with the fitted depth of the vacancy-defect layer [see Fig. (10.4)]. Consequently, the origin of the D_{II} defect is not inconsistent with a vacancy-related complex, which is also a conclusion entailed by Monte Carlo simulations of ion implantation in 6H-SiC [127], that indicated an accumulation of Si and C interstitials deeper in the material and an excess of vacancies closer to the surface.

10.3 Conclusions

Concluding, the comparison between the PL results of Ref. [113] and our PAS measurements allows us to suggest a less conventional attribution of the nature of these centres. According to our results, D_I defects are not consistent with di-vacancies, due to the absolute lack of correlation between the relatively deep penetration depth of those centres and the relatively shallow nature of the layer where positron detectable defects concentrate. On the other hand, D_{II} centres has a good correlation with the depth distribution of the vacancy-related defects, even though it is not enough to propose a certain attribution. The dominant defects in the damaged layer might be related to dislocation loops, however a complete annealing of di-vacancies in that shallow region cannot be ruled out. The D_I lines could be related to some interstitial complex. Unfortunately, PAS seems not to be sensitive to this kind of defects, so further investigations are needed to provide a reliable identification of these centres.

Vacancies and di-vacancies are expected to be scarcely mobile defects, as we have discussed in Chapter (7) and it is discussed in Ref. [128], therefore it appears reasonable - as the experimental results presented in this chapter have shown - that they are not the source of the deep D_I PL peaks.

Chapter 11

Conclusions

In this work we have discussed several aspects of defect physics in silicon carbide. In particular, we have concentrated on some of the most interesting features that defects in a semiconductor have from the technological viewpoint. Certainly defect diffusion is one of these topics and it is where we have focused on at first in Chapters 6 and 7. As we have discussed in the corresponding chapters, diffusion is an especially challenging subject and it is highly relevant for dopants as well as for intrinsic defects. In both cases there is prominent need to gain the ability of controlling defect distribution, even though with dopants we want to be able to confine them inside a region, i.e. the junction, while with intrinsic defects or undesired impurities what we want is simply to anneal them out.

The activation of dopants in the high-dose regime has been the subject of the study we have presented in Chapter 8. Dopants are intentionally introduced in the host crystal to modify its conductive properties, but this can happen only if they are electrically active.

In chapter 9 we have shown the performances of a brand new method to calculate the free energy of a system and then to evaluate melting points and phase coexistence. So far we have restricted to silicon, to thoroughly test the method. However, our mid-term aim is to study the phase transitions between different polytypes of silicon carbide.

Finally, chapter 10 has shown how experimental measurements can actually constitute a virtuous circle, giving crucial feedback to theoretical simulations.

Below we intend to perform a brief review of the most important results that we have obtained. Our calculations revealed that:

1. Implanted boron diffusion is activated via a *kick-out* mechanism. Interstitial silicon atoms makes substitutional boron unstable and remove

it from the lattice site. The silicon interstitial atom thus *takes back* its own site and the boron impurity diffuses interstitially with a quite low barrier of $\sim 0.65eV$. When the charge state of the impurity is taken into account, we have found that the interstitial diffusion barrier can grow up to more than 2.5 eV in p-type doping conditions, although in the high-temperature conditions that induce the *kick-out* it is the neutral charge state of the impurity which is relevant.

2. Carbon vacancy diffuses by means of a second-neighbour hopping mechanism. On the other hand, silicon vacancy favours the formation of a carbon vacancy − antisite pair and therefore two migration mechanisms are possible: the direct path already observed for the carbon vacancy and a mediated diffusion path. Both vacancies have been found to be more mobile under n-type doping conditions.

3. High-dose implantation of nitrogen may produce the formation of aggregates that self-passivate. On the other hand, the most stable phosphorus complex that can form is still characterised by a shallow level and is thus still an active dopant. For this reason, high-dose n-type doping is recommended to be carried out implanting phosphorus rather than nitrogen.

4. The melting temperature of crystalline silicon has been estimated by means of the reversible scaling method. A value of 1571.5 K by means of the EDIP potential and of 1700.3 K by means of non-orthogonal tight-binding, in very good agreement with the experimental value of 1683 K has been obtained.

5. Positron annihilation spectroscopy measurements has indicated that the conventional attribution of the D_I defects to di-vacancies is not feasible with the depth-distribution of the corresponding PL signal and that of the vacancy-related defects.

Silicon carbide is definitely an extremely interesting material for power electronics applications. However, there are still a considerable amount of unanswered questions on the atomic scale mechanisms that rule many technological processes. In this work we have demonstrated that theoretical modelling is a powerful tool to provide a microscopic picture of the material and to help improving its quality for industrial applications.

Bibliography

[1] D. E. Cusack, W. M. Glasheen, and H. R. Steglich, Trans. of 2nd Int. High Temprature Electronics Conf. (HiTEC), Session III, p.17 (1994)

[2] P. L. Dreike, D. M. Fleetwood, D. B. King, D. C. Sprauer, and T. E. Zipperian, IEEE Trans. Comput. Hybrids, Man. Tech. A **17**, 594 (1994).

[3] *Materials for High-Temperature Semiconductor Devices*, National Materials Advisory Board (NMAB-474), National Academy Press, Washington DC (1995).

[4] M. Tajima, Trans. 2nd Int. High Temprature Electronics Conf. (HiTEC), Session I, p.29 (1994).

[5] D. M. Fleetwood, F. V. Thome, S. S. Tsao, P. V. Dressendorfer, V. J. Dandini, and J. R. Schwank, IEEE Trans. Nucl. Sci. **NS-35**, 1099 (1998).

[6] J. W. Palmour and H. C. Carter Jr., Proc. 1993 Int. Semiconductor Device Research Symp., p.695 (1993).

[7] R. J. Trew, J. B. Yan, and P. M. Mock, Proc. IEEE **79**, 598 (1991).

[8] J. M. McGarrity, F. B. McLean, W. M. DeLancey, J. W. Palmour, H. C. Carter Jr., and J. A. Edmond, IEEE Trans. Nucl. Sci. **39**, 1974 (1992).

[9] W. Shockley, Proc. of the First Int. Conf. on SiC, Boston, MA , edited by J. R. O'Connor and J. Smiltens, Pergamon, New York (1960).

[10] Yu. M. Tairov and V. F. Tsvetkov, Journal of Crystal Growth **43**, 209 (1978).

[11] R. C. Glass, D. Henshall, V. F. Tsvetkov, and C. H. Carter, phys .stat .sol. (b) **202(1)**, 147 (1997).

[12] G. Pensl and W. J. Choyke, Phys. Rev. B **185**, 264 (1993).

[13] D. J. Larkin, Inst. Phys. Conf. Ser., No. 142, p.23, IOP Publishing (1996).

[14] S. Ahmed, C. J. Barbero, T. W. Sigmon, and J. W. Erickson, Appl. Phys. Lett., **65**, 1 (1994).

[15] M. Ghezzo, D. M. Brown, E. Downey, J. Kretchmer, and J. J. Kopanski, Appl. Phys. Lett., **63**, 1206 (1993).

[16] H. Du, M. Libera, Z. Yang, P. J. Lai, D. C. Jacobson, Yu. C. Wang, and R. F. Davis, Appl. Phys. Lett., **62**, 423 (1993).

[17] T. Dalibor, C. Peppermüller, G. Pensl, S. Sridhara, R. P. Devaty, W. J. Choyke, A. Itoh, T. Kimoto, and H. Matsunami, Inst. Phys. Conf. Ser., No. 142, p.517, IOP Publishing (1996).

[18] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: introduction to advanced electronic structure theory*, McGraw-Hill, New York (1989).

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

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

[21] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1988).

[22] A. Messiah, *Quantum Mechanics*, North-Holland, Amsterdam (1961).

[23] M. Born and J. R. Oppenheimer, Ann. der Phys. **84**, 457 (1927).

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

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

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

[27] J. P. Perdew and S. Kurth, *Density Functional Theory of Atoms and Molecules*, Oxford (1989).

[28] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[29] S. F. Boys in *Quantum Theory of Atoms, Molecules and the Solid State*, edited by P. O. Lödwin, Academic Press, New York.

[30] P. Ordejón, Comp. Mat. Sci. **12**, 157 (1998).

[31] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).

[32] S. F. Boys and F. Bernardi, J. Mol. Phys. **19**, 553 (1970).

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

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

[35] J. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter, **14**, 2745 (2002).

[36] B. Delley, J. Chem. Phys. **17**, 197 (1969).

[37] D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. Soler, Int. J. Quant. Chem. **65**, 453 (1997).

[38] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, (Clarendon Press, Oxford 1987).

[39] D. Frenkel and B. Smit, *Understanding Computer Simulation* (Academic Press, New York 1996).

[40] J. Chem. Phys. **115**, 10282 (2001).

[41] J. E. Northrup and S. B. Zhang, Phys. Rev. B **47**, 6791 (1992).

[42] B. Aradi, A. Gali, P. Deák, J. E. Lowther, N. T. Son, E. Janzén and W. J. Choyke, Phys. Rev. B **63**, 245201 (2001).

[43] C. M. Goringe, D. R. Bowler, and E. Hernández, Rep. Prog. Phys. **60**, 1447 (1997).

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

[45] M. Villarba and H. Jónsson, Surf. Sci., **317**, 15 (1994).

[46] G. Mills, H. Jónsson, G. K. Schenter, Surf. Sci., **324**, 305 (1995).

[47] I. V. Ionova and E. A. Carter, J. Chem. Phys., **98**, 6377 (1993).

[48] H. Jónsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, B. J. Berne, G. Ciccotti, and D. F. Coker (Eds.), (World Scientific, Singapore 1998).

[49] S. G. Sridhara, L. L. Clemen, R. P. Devaty, W. J. Choyke, D. J. Larkin, T. Troffer, and G. Pensl, J. Appl. Phys. **83**, 7909 (1998).

[50] C. S. Nichols, C. G. Van de Walle, and S. T. Pantelides, Phys. Rev. Lett. **62**, 1049 (1989).

[51] C. S. Nichols, C. G. Van de Walle, and S. T. Pantelides, Phys. Rev. B **40**, 5484 (1989).

[52] J. Zhu, T. Diaz de la Rubia, L. H. Yang, C. Mailhiot, and G. H. Gilmer, Phys. Rev. B **54**, 4741 (1996).

[53] B. Sadigh, T. J. Lenosky, S. K. Theiss, M. J. Caturla, T. Diaz de la Rubia, and M. A. Foad, Phys. Rev. Lett. **83**, 4341 (1999).

[54] W. Windl, M. M. Bunea, R. Stumpf, S. T. Dunham, and M. P. Masquelier, Phys. Rev. Lett. **83**, 4345 (1999).

[55] W. Luo, P. B. Rasband, P. Clancy, and B. W. Roberts, J. Appl. Phys. **84**, 2476 (1998).

[56] P. Alippi, L. Colombo, P. Ruggerone, A. Sieck, G. Seifert, and Th. Frauenheim, Phys. Rev. B **64**, 075207 (2001).

[57] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[58] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[59] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **8**, 5747 (1973).

[60] K. J. Chang and M. L. Cohen, Phys. Rev. B **35**, 8196 (1987).

[61] C. Wang, J. Bernholc, and R. F. Davis, Phys. Rev. B **38**, 12752 (1988).

[62] *Physics of Group IV Elements and III-V Compounds*, Vol. 17a of Landolt-Börnstein Tables, O. Madelung, M. Schulz, and H. Weiss (Eds.), (Springer-Verlag, Berlin 1982).

[63] H. Bracht, N. A. Stolwijk, M. Laube, and G. Pensl, Appl. Phys. Lett. **77**, 3188 (2000).

[64] M. Bockstedte and O. Pankratov, Mater. Sci. Forum **338-342**, 949 (2000).

[65] M. Bockstedte, A. Mattausch, and O. Pankratov, Mater. Sci. Forum **353-356**, 447 (2001).

[66] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN, The Art of Scientific Computing 2nd Edition*, (Cambridge University Press, Cambridge 1992).

[67] R. Rurali, P. Godignon, J. Rebollo, P. Ordejón, and E. Hernández, Appl. Phys. Letters, **81**, 2989 (2002).

[68] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).

[69] S. I. Soloviev, Y. Gao, and T. S. Sudarshan, Appl. Phys. Letters, **77**, 4004 (2000).

[70] E. Morvan, P. Godignon, M. Vellvehi, A. Hallen, M. Linnarsson and A. Yu. Kuznetsov, Appl. Phys. Letters, **74**, 3990 (1999).

[71] M. J. Puska, S. Poykkp, and R. M. Nieminen, Phys. Rev. B, **58**, 1318 (1998).

[72] M. Bockstedte, Mater. Sci. Forum **389-393**, 471 (2002).

[73] L. Torpo, V. Marlo, T. E. M. Staab, and R. M. Nieminen, J. Phys.: Condens. Matter, **13**, 6203 (2001).

[74] A. Mattausch, M. Bockstedte, and O. Pankratov, Mater. Sci. Forum **353-356**, 323 (2001).

[75] E. Rauls, Th. Lingner, Z. Hajnal, S. Greulich-Weber, Th. Frauenheim, and J. M. Spaeth, phys .stat .sol. (b) **217**, R1 (2000).

[76] R. Rurali, E. Hernández, P. Godignon, J. Rebollo, and P. Ordejón, Comp. Mat. Sci., **27/1-2**, 36 (2003).

[77] H. H. Woodbury and G. W. Ludwig, Phys. Rev. **124**, 1083 (1961).

[78] S. Greulich-Weber, M. Feege, J. M. Spaeth, E. N. Kalabukhova, S. N. Lukin, and E. N. Mokhov, Solid State Commun. **93**, 393 (1995).

[79] T. Troffer, C. Peppermüller,G. Pensl, and A. Schöner, J. Appl. Physics, **80**, 3739 (1996).

[80] M. A. Capano, J. A. Cooper Jr., M. R. Melloch, A. Saxler, and W. C. Mitchel, J. Appl. Physics, **87**, 8773 (2001).

[81] M. Laube, F. Schmid, G. Pensl, and G. Wagner, Mater. Sci. Forum, **389-303**, 791 (2002).

[82] J. A. Gardner, A. Edwards, M. V. Rao, N. Papanicolaou, G. Kelmer, M. A. Capano, M. Ghezzo, and J. Kretchmer, J. Appl. Phys. **83**, 5118 (1999).

[83] S. Blanque, R. Pérez, M. Zielinski, J. Pernot, N. Mestres, J. Pascual, P. Godignon, and J. Camassel, Mat. Sci. Forum, **433-436**, 653 (2003).

[84] T. Miyajima, N. Tokura, A. Fukumoto, H. Hayashi, and K. Hara, Jpn. J. Appl. Phys., **35**, 1231 (1996).

[85] P. Deák, A. Gali, J. Miro, R. Gutierrez, A. Sieck, and Th. Frauenheim, Mat. Sci. Forum **264-268**, 279 (1998).

[86] A. Gali, P. Deák, R. P. Devaty, and W. J. Choyke, Mat. Sci. Forum **338-342**, 795 (2000).

[87] A. Gali, P. Deák, P. R. Briddon, R. P. Devaty, and W. J. Choyke, Phys. Rev. B, **61**, 12602 (2000).

[88] S. A. Kajihara, A. Antonelli, J. Bernholc, and R. Car, Phys. Rev. Lett. **66**, 2010 (1991).

[89] R. K. Malhan, J. Kozima, T. Yamamoto, and A. Fukumoto, Mater. Sci. Forum, **389-303**, 541 (2002).

[90] J. K. Johnson, J. A. Zollweg and K. E. Gubbins, Mol. Physics, **78**, 591, (1993).

[91] D. A. Young and F. J. Rogers, J. Chem. Phys., **81**, 2789 (1984).

[92] M. Watanabe and W. P. Reinhardt, Phy. Rev. Letters, **65**, 3301 (1990).

[93] P. Hertz, Ann. Phys. (Leipzig) **33**, 537 (1910).

[94] M. de Koning, A. Antonelli and S. Yip, Phys. Rev. Letters, **83**, 3973 (1999).

[95] R. Rurali and E. Hernández, Comp. Mat. Sci., accepted.

[96] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert and R. Kashner, Phys. Rev. B **51**, 12947 (1995).

[97] I. Kwon, R. Biswas, C. Z. Wang, K. M. Ho and C. M. Soukoulis, Phys. Rev. B **49**, 7242 (1994).

[98] T. J. Lenosky, J. F. Kress, I. Kwon and A. F. Voter, Pys. Rev. B, **55**, 1528 (1997).

[99] J. Tersoff, Phys. Rev. Lett., **56**, 632 (1986).

[100] J. F. Justo, M. Z. Bazant, E. Kaxiras, V. V. Bulatov and S. Yip, Phys. Rev. B, **58**, 2539 (1998).

[101] F. Ercolessi, M. Parrinello and E. Tosatti, Phil. Mag. A **58**, 213 (1988).

[102] F. Cleri and V. Rosato, Phys. Rev. B **48**, 22 (1993).

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

[104] N. L. Allan, G. D. Barrera, C. E. Sims and M. B. Taylor, Phys. Chem. Chem. Phys., **2**, 1099 (2000).

[105] C. R. Miranda and A. Antonelli, private communication.

[106] P. Keblinski, M. Z. Bazant, R. K. Dash, and M. M. Treacy, Phys. Rev. B **66**, 064104 (2002)

[107] G. Consolati, R. Rurali and M. Stefanetti, Chem. Phys., **237**, 493 (1998).

[108] J. I. Pankove, *Optical Processes in Semiconductors* (Dover Pubblications Inc., New York 1971).

[109] W. J. Choyke and L. Patrick, Phys. Rev. B, **4**, 1843 (1971).

[110] L. Patrick and W. J. Choyke, PH. Rev. B, **5**, 3253 (1972).

[111] S. G. Sridhara, D. G. Nizhner, R. P. Devaty, W. J. Choyke, T. Dalibor, G. Pensl and T. Kimoto, Mater. Sci. Forum, **264-268**, 494 (1998).

[112] S. G. Sridhara, F. H. C. Carlsson, J. P. Bergman, A. Henry and E. Janzén, Mater. Sci. Forum, **353-356**, 377 (2001).

[113] Y. Koshka and G. Melnychuk, Mater. Sci. Forum, **389-393**, 513 (2002).

[114] W. J. Choyke, Z. C. Feng and J. A. Powell, J. Appl. Phys., **64**, 3163 (1988).

[115] A. Mattausch, M. Bockstedte and O. Pankratov, Mater. Sci. Forum, **389-393**, 481 (2002).

[116] P. J. Schultz and K. G. Lynn, Rev. Mod. Phys., **60**, 701 (1998).

[117] A. van Veen, H. Schut, J. de Vries, R. A. Haakvoort and M. R. Ijpma, Positron beams for solids and surfaces, in P. J. Schultz, G. R. Massoumi, P. J. Simpson (Eds.), Proceedings of the AIP Conference, **218**, 171 (1990).

[118] G. Brauer, W. Anwand , P. G. Coleman, J. Störmer, F. Plazaola, J. M. Campillo, Y. Pacaud and W. Skorupa, J. Phys.: Condens. Matter, **10**, 1147 (1998).

[119] G. Brauer, W. Anwand, Y. Pacaud, W. Skorupa, F. Plazaola, P. G. Coleman, A. P. Knights, J. Störmer and P. Willutzki, Phys. Rev. B **54**, 3084 (1996).

[120] A. Kawasuso, F. Redmann, R. Krause-Rehberg, T. Frank, M. Weidner, G. Pensl, P. Sperr and H. Itoh, J. Appl. Phys., **90**, 3377 (2001).

[121] A. Kawasuso, F. Redmann, R. Krause-Rehberg, M. Weidner, T. Frank, G. Pensl, P. Sperr, W. Triftshaüser and H. Itoh, Appl. Phys. Letters, **79**, 3950 (2001).

[122] M. -F. Barthe, P. Desgardin, L. Henry, C. Corbel, D. T. Britton, G. Kögel, P. Sperr, W. Triftshaüser, P. Vicente and L. diCioccio, Mater. Sci. Forum, 389-393, 493 (2002).

[123] C. C. Ling, C. D. Beling and S. Fung, Phys. Rev. B, **62**, 8016 (2000).

[124] Y. Pacaud, J. Stoemenos, G. Baruer, R. A. Yankov, V. Heera, M. Voelskow, R. Kögler and W. Skorupa, Nucl. Instrum. Methods B, **120**, 177 (1996).

[125] W. Anwand, G. Brauer and W. Skorupa, Mater. Sci. Forum, **363-365**, 442 (2001).

[126] W. Anwand, G. Brauer, H. Wirth, W. Skorupa, P. G. Coleman, Appl. Surf. Sci., **194**, 127 (2002).

[127] E. Morvan, J. Montserrat, J. Rebollo, D. Flores, X. Jordà, M. L. Locatelli, Mater. Sci. Forum, **264-286**, 737 (1998).

[128] L. Torpo , T. E. M. Staab and R. M. Nieminen, Phys. Rev. B **65**, 85202 (2002).