POINT AND EXTENDED DEFECTS IN GROUP IV SEMICONDUCTORS

Submitted by Thomas Andreas Georg Eberlein to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Physics.

March, 2004

I certify that all material in this thesis which is not my own work has been identified and that no material is included for which a degree has previously been conferred upon me.

ABSTRACT

First principles calculations of point and extended defects in silicon and 4*H*-SiC are reported.

In Si the self-interstitial is discussed in relation to the AA12 electron paramagnetic resonance (EPR) centre. Its electrical level, annealing behaviour and spin density are compared with experiment. Also in Si, so called rod-like defects which are nucleation centres for interstitials are reported. Results on the structure, energetics and electrical properties of extended interstitial defects with {113}, {111} and {001} habit planes are described and compared with experiment.

In SiC atomic models for the most common PL signals the Alphabet lines and the D_I line are presented. The calculations investigate antisite pairs and their donor levels and gap vibrational modes including isotopic shifts. Furthermore a model for the Z_1/Z_2 centres, the most common DLTS centre in 4*H*-SiC, is presented. An interstitial nitrogen - interstitial carbon defect is shown to be exceptionally thermally stable, bistable and has negative-U character with donor and acceptor levels close to those attributed to the defect.

Finally vanadium and erbium centres in 4*H*-SiC are modelled. They both lie at Si sites and can exist in three charge states with deep donor and acceptor levels. While isolated V_{Si} possesses intra-*d* and ionisation induced optical transitions around 0.94 eV and 2.9 eV respectively, it is shown that the intense and temperature stable intra-*f* optical transitions of Er are unlikely to be due to an *isolated* Er defect.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Bob Jones for his guidance, daily enthusiasm and constant search for progress.

I also thank all members of the AIMPRO group, Dr. P. R. Briddon for his support on code development, and obviously my colleagues José, Ben, Caspar, Chris, Deborah, James, Jon, Markus and Steve for all friendship and help throughout these three years.

Prof. J. Steeds is much thanked for discussions and criticisms. I also would like to thank Prof. D. W. Palmer for his useful comments and careful reading of parts of this thesis and Prof. S. Öberg for additional computation time.

CONTENTS

II Defects in 4*H*-SiC 53

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS

1. T.A.G. Eberlein, N. Pinho, R. Jones, B.J. Coomer, J.P. Goss, P.R. Briddon, and S. Oberg, *Self-interstitial clusters in silicon*, Physica B **308**, 454–457, 2001.
2. T.A.G. Eberlein, C.J. Fall, R. Jones, P.R. Briddon, and S. Oberg, *Alphabet luminescence lines in 4H-SiC*, Phys. Rev. B **65**, art. no. 184108, 2002.
3. J.P. Goss, T.A.G. Eberlein, R. Jones, N. Pinho, A.T. Blumenau, T. Frauenheim, P.R. Briddon, and S. Oberg, *Planar interstitial aggregates in Si*, J. Phys.-Condes. Matter **14**, 12843–12853, 2002.
4. T.A.G. Eberlein, R. Jones, and P.R. Briddon, *Z_1/Z_2 Defects in 4H-SiC*, Phys. Rev. Lett. **90**, art. no. 225502, 2003.
5. T.A.G. Eberlein, L. Huggett, R. Jones, and P.R. Briddon, *Hydrogen molecules in 4H-SiC and 2H-GaN*, J. Phys.-Condes. Matter **15**, S2897–S2902, 2003.
6. D. Prezzi, T.A.G. Eberlein, J.-S. Filhol, R. Jones, M.J. Shaw, P.R. Briddon, and S. Oberg, *Optical and electrical properties of Vanadium and Erbium in 4H-SiC*, submitted to Phys. Rev.
7. J.P. Goss, P.R. Briddon, T.A.G. Eberlein, R. Jones, N. Pinho, A.T. Blumenau, and S. Oberg, *Planar interstitial aggregates in Si and Ge*, submitted to APL.

INTRODUCTION

Silicon is the most common material used for semiconductors: more than 90% of electronic components are produced from it. Most electronic properties are influenced if not determined by the kind and amount of defects in the material. Advanced semiconductor devices are sub-micron devices requiring high spatial control of dopants. These dopants are however subject of diffusion and the diffusion of dopants in semiconductors is strongly influenced by the thermal behaviour of self-interstitial clusters, nucleated as a result of nonequilibrium processing steps such as ion implantation. During ion implantation, vacancies and interstitials are created as lattice damage. The formation energies of these defects are of the order of 3–4 eV. Although many of them recombine, a fraction stays either as single interstitial or diffuses through the semiconductor to build larger complexes with each other and with impurities like dopants.

One of the main *p*-type dopants of Si is boron. Cowern *et al.* (1990) proposed that the diffusion of B occurs through a kick-out mechanism. Excess self-interstitials kick B atoms out of their substitutional positions, enabling them to diffuse until a reverse kick-out occurs. This is known as transient enhanced diffusion (TED), caused by the self-interstitial with its surprisingly high mobility. In 1998 Mukashev *et al.* claimed to have directly observed the isolated self-interstitial. Previously, only defects which are identified as interstitials trapped at impurities had been observed. Chapter 3 discusses this fundamental and technologically important defect in silicon.

Extended defects can form from the aggregation of self-interstitials. In silicon, two different types of extended defects are observed. At high interstitial densities or high annealing temperatures dislocation loops form. However, if the density of interstitials or the annealing temperature is lower, so called rod like defects (RLDs) form instead. These defects are made of interstitial chains extended along $\langle 110 \rangle$. The most common form of this complex is the $\{311\}$ aggregate: a RLD in the $\{311\}$ habit plane. The RLDs are up to a micron long and 0.1 μm wide. Processing stages in wafer production after ion implantation require anneals around 800°C. At these temperatures the clusters and RLDs dissolve and flood the wafer with interstitials. This can promote TED which causes the loss of accurate profiling. TED persists as long as there is a source of interstitials in the form of clusters and RLDs. Chapter 4 discusses RLDs and their preferential choice of habit plane.

In contrast to silicon, SiC is a material that seems ideal for high-power, high frequency and high temperature electronic devices. It does not suffer from large reverse recovery inefficiencies typical for silicon when switching. Recent advances in crystal growth make it possible now to grow epitaxial and bulk SiC material of high quality. It is now also possible to grow semi-insulating material without adding impurities compensating the excess of unwanted nitrogen. However, as in silicon, point defects that degrade device performance are present. In SiC, apart from dislocations and stacking faults which play a major role in device degradation, point defects are important. As in silicon TED of boron and aluminium have been observed (Laube *et al.*, 1999; Janson *et al.*, 2000; Usov *et al.*, 1999). Point defects are formed during growth but are also generated during ion implantation: the only way to achieve selective area doping in SiC. Zhang *et al.* (2003) correlated the carrier lifetime with the Z_1/Z_2 centres and concluded that they may be the lifetime limiting defects in their samples. While Z_1/Z_2 is the most prominent deep level transient spectroscopic (DLTS) signal in 4*H*-SiC, the most prominent photoluminescence (PL) signals are D_1 and the related Alphabet lines. The latter can be seen directly after low energy irradiation while D_1 usually dominates the PL spectrum of implanted and irradiated SiC *after* annealing. Chapter 5 discusses the relation of the Alphabet lines and D_1 with antisites while chapter 6 establishes a tentative atomistic model for the Z_1/Z_2 centres.

In addition to intrinsic defects, transition metal impurities in semiconductors are important because they can possess multiple charge states and are efficient carrier traps. Moreover, they often give rise to internal intra-*d* optical transitions. Rare-earth dopants on the other hand are of increasing interest because their intra-*f* luminescence is stable to high temperatures in wide band gap materials (Zanatta, 2003). They are not noted for introducing deep levels.

Finally a brief analysis of vanadium, historically used to make semi-insulating SiC, and Er, important for its intra-f transitions which emit visible light, is presented in chapter 7.

CHAPTER 1

The Method

1.1 The Many-Body Problem

Any problem concerning the electronic structure of matter is described within wavefunction theory. The Schrödinger equation of the system is

$$\mathcal{H}\Psi = E\Psi, \tag{1.1}$$

where E is the total energy of the system. Often the time-dependence of the problem can be neglected and the wavefunction Ψ is a function of N_n nuclei coordinates $\mathbf{R}_i$ and N_e electron coordinates $\mathbf{x}_i$. Each coordinate $\mathbf{x}_i$ comprises space coordinates $\mathbf{r}_i$ and spin coordinates s_i,

$$\Psi \equiv \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}; \mathbf{R}_1, \ldots, \mathbf{R}_{N_\mathrm{n}}) \equiv \Psi(\mathbf{r}_1, s_1, \ldots, \mathbf{r}_{N_\mathrm{e}}, s_{N_\mathrm{e}}; \mathbf{R}_1, \ldots, \mathbf{R}_{N_\mathrm{n}}). \tag{1.2}$$

All quantities in this chapter are in atomic units. In this system of units, $\hbar$, e, m and $4\pi\epsilon_0$ are taken to be unity, where e, m and ϵ_0 are the electron charge, electron mass, and the permittivity of vacuum, respectively.

The non-relativistic Hamiltonian containing the usual kinetic and potential terms can be written as

$$\begin{aligned} \mathcal{H} = &-\frac{1}{2}\sum_{i}^{N_\mathrm{e}} \nabla_i^2 - \sum_{\alpha}^{N_\mathrm{n}} \frac{1}{2M_\alpha}\nabla_\alpha^2 + \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N_\mathrm{e}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \\ &-\sum_{i,\alpha=1}^{N_\mathrm{e},N_\mathrm{n}} \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \frac{1}{2}\sum_{\substack{\alpha,\beta=1\\ \alpha\neq\beta}}^{N_\mathrm{n}} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}, \end{aligned} \tag{1.3}$$

where M_α, Z_α and $\mathbf{R}_\alpha$ represent the mass, charge and location of the α-th nucleus, and $\mathbf{r}_i$ the i-th electron coordinate.

A short version of this Hamiltonian where the summands of Eqn. 1.3 are replaced by kinetic ($\mathcal{T}$) and potential operators ($\mathcal{V}$) is

$$\mathcal{H} = \mathcal{T}_\mathrm{e} + \mathcal{T}_\mathrm{n} + \mathcal{V}_\mathrm{ee} + \mathcal{V}_\mathrm{en} + \mathcal{V}_\mathrm{nn}. \tag{1.4}$$

In general even small problems with only a few atoms are intractable and approximations have to be made to simplify the problem.

1.2 Born-Oppenheimer Approximation

Since the electron mass is three orders of magnitude smaller than the mass of a proton and subsequently yet smaller than the mass of a nucleus, nuclear motion is much slower than electron motion. Electrons therefore respond instantaneously to atomic motion and both motions can be separated. The total wavefunction Ψ can be written as a product of the electronic wavefunction ψ and the nuclear wavefunction ϕ,

$$\Psi(\mathbf{x};\mathbf{R}) = \psi_{\mathbf{R}}(\mathbf{x})\,\phi(\mathbf{R}). \tag{1.5}$$

This was first proposed by Born and Oppenheimer (1925). The variables $\mathbf{x}$ and $\mathbf{R}$ represent now all electronic and nuclear degrees of freedom. The subscript on ψ stresses the fact that the electronic wavefunction depends on $\mathbf{R}$ in a parametric way.

The separated electronic problem is

$$\left(\mathcal{T}_{\mathrm{e}} + \mathcal{V}_{\mathrm{ee}} + \mathcal{V}_{\mathrm{en}}(\mathbf{R})\right)\psi_{\mathbf{R}} = E_{\mathrm{e}}(\mathbf{R})\psi_{\mathbf{R}}. \tag{1.6}$$

with the corresponding eigenvalue E_{e} which depends on the atomic positions $\mathbf{R}$. With this the Hamilton equation becomes

$$\begin{aligned}\mathcal{H}\psi_{\mathbf{R}}(\mathbf{x})\phi(\mathbf{R}) =& E_{\mathrm{e}}(\mathbf{R})\,\psi_{\mathbf{R}}(\mathbf{x})\phi(\mathbf{R}) + \psi_{\mathbf{R}}(\mathbf{x})\left(\mathcal{T}_{\mathrm{n}} + \mathcal{V}_{\mathrm{nn}}\right)\phi(\mathbf{R}) - \\ &- \sum_{\alpha}^{N_{\mathrm{n}}} \frac{1}{2M_\alpha}\left[\phi(\mathbf{R})\nabla_\alpha^2\psi_{\mathbf{R}}(\mathbf{x}) + 2\boldsymbol{\nabla}_\alpha\psi_{\mathbf{R}}(\mathbf{x})\cdot\boldsymbol{\nabla}_\alpha\phi(\mathbf{R})\right].\end{aligned} \tag{1.7}$$

Multiplying the equation by $\psi^*_{\mathbf{R}}(\mathbf{x})$ and integrating over $\mathbf{x}$ gives

$$\begin{aligned}E\phi(\mathbf{R}) =& (\mathcal{T}_{\mathrm{n}} + \mathcal{V}_{\mathrm{nn}} + E_{\mathrm{e}}(\mathbf{R}))\phi(\mathbf{R}) + \\ &+ W(\mathbf{R})\phi(\mathbf{R}) - \sum_{\alpha}^{N_{\mathrm{n}}} \frac{1}{M_\alpha}\int \psi^*_{\mathbf{R}}(\mathbf{x})\boldsymbol{\nabla}_\alpha\psi_{\mathbf{R}}(\mathbf{x})d\mathbf{x}\cdot\boldsymbol{\nabla}_\alpha\phi(\mathbf{R})\end{aligned} \tag{1.8}$$

with

$$W(\mathbf{R}) = -\sum_{\alpha}^{N_{\mathrm{n}}} \frac{1}{2M_\alpha}\int \psi^*_{\mathbf{R}}(\mathbf{x})\nabla_\alpha^2\psi_{\mathbf{R}}(\mathbf{x})d\mathbf{x}. \tag{1.9}$$

Integration over $\mathbf{x}$ means integration over the spacial coordinates $\mathbf{r}$ and summation over the spin states, i.e.

$$\int d\mathbf{x} = \sum_s \int d\mathbf{r}. \tag{1.10}$$

$W(\mathbf{R})$ is negligibly small and can be omitted mainly because of the nuclear mass in the denominator but also because the electron wavefunctions are often insensitive to nuclear displacement. The last term of Eqn. 1.8 is zero for a non-degenerate ground state, and subsequently a real Ψ. For degenerate ground states it can be important and lead to symmetry breaking as in the Jahn-Teller effect (Stoneham, 1975). If both terms are neglected the nuclear and electron motions are decoupled and the Hamiltonian is separated into Eqn. 1.6 and

$$(\mathcal{T}_{\mathrm{n}} + \mathcal{V}_{\mathrm{nn}} + E_{\mathrm{e}}(\mathbf{R}))\phi(\mathbf{R}) = E\phi(\mathbf{R}). \tag{1.11}$$

1.3 Hartree-Fock Method

From now on we concentrate on solving the Schrödinger equation for the electron (Eqn. 1.6). One way to construct a wavefunction for this is the Hartree-Fock (HF) method (Fock, 1930; Roothaan, 1951; Slater, 1965; Stoneham, 1975). Here the wavefunction of the many-electron system is constructed from an anti-symmetrised product of one-electron wavefunctions, thereby obeying the Pauli exclusion principle. It can be written in the form of Slater determinants [1] (Slater, 1929),

$$\Psi_{\mathrm{HF}}(\mathbf{x}_1,\ldots,\mathbf{x}_N)=\frac{1}{\sqrt{N!}}\det\begin{bmatrix}\psi_1(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1)\\ \vdots & \ddots & \vdots\\ \psi_1(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N)\end{bmatrix}, \tag{1.12}$$

where $\mathbf{x}_i$ includes spatial and spin coordinates of the i-th electron, and $\psi_j(\mathbf{x}_i)$ are orthonormal one-electron *spin-orbitals*, which are a product of a spatial orbital $\phi_j(\mathbf{r}_i)$ and a spin function $\sigma_j(s_i)=\alpha(s_i)$ or $\beta(s_i)$. Orthonormality then means

$$\int \phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)d\mathbf{r}_1=\delta_{ij} \quad \text{and} \sum_s \sigma_i^*(s)\sigma_j(s)=\delta_{\sigma_i\sigma_j}. \tag{1.13}$$

with the sum over up and down spins. With Ψ_{HF} as a solution of the Hamiltonian in 1.6 , one arrives at an expectation value for the total energy,

$$E_{HF}=\sum_i^N H_i+\frac{1}{2}\sum_{i,j=1}^N\left(J_{ij}-K_{ij}\right), \tag{1.14}$$

where

$$H_i=\int \psi_i^*(\mathbf{x}_1)\left[-\frac{1}{2}\nabla_i^2-\sum_{\alpha=1}^{N_{\mathrm{n}}}\frac{Z_\alpha}{|\mathbf{r}_i-\mathbf{R}_\alpha|}\right]\psi_i(\mathbf{x}_1)\,d\mathbf{x}_1 = \langle i\,|\mathcal{T}_{\mathrm{e}}+\mathcal{V}_{\mathrm{en}}|\,i\rangle\,, \tag{1.15}$$

$$J_{ij}=\iint \psi_i^*(\mathbf{x}_1)\psi_j^*(\mathbf{x}_2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\psi_i(\mathbf{x}_1)\psi_j(\mathbf{x}_2)\,d\mathbf{x}_1\,d\mathbf{x}_2 = \langle ij\,|\mathcal{V}_{\mathrm{ee}}|\,ij\rangle\,, \tag{1.16}$$

$$K_{ij}=\iint \psi_i^*(\mathbf{x}_1)\psi_j^*(\mathbf{x}_2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\psi_j(\mathbf{x}_1)\psi_i(\mathbf{x}_2)\,d\mathbf{x}_1\,d\mathbf{x}_2 = \langle ij\,|\mathcal{V}_{\mathrm{ee}}|\,ji\rangle\,. \tag{1.17}$$

H_i is the one-electron integral and J_{ij} and K_{ij} are called *Coulomb* or *Hartree integrals* and *exchange integrals*. Note that J_{ij} and K_{ij} are *four-centre* integrals and represent an electron-electron repulsion term and an exchange component representing spin-correlation effects, respectively. Integration over $\mathbf{x}$ means spatial integration over $\mathbf{r}$ and spin summation over s. Both four-centre integrals are real and $J_{ij}\geq K_{ij}\geq 0$. Since $J_{ii}=K_{ii}$ the two-electron summation in 1.14 can include the $i=j$ term.

Minimisation of 1.14 subject to the orthonormalization condition 1.13, gives the Hartree-Fock equations (McWeeny, 1989; Thijssen, 1999),

$$\mathcal{F}\,\psi_i(\mathbf{x}_1)=\sum_{j=1}^N\epsilon_{ij}\,\psi_j(\mathbf{x}_1), \tag{1.18}$$

with $\mathcal{F}$ known as Fock operator, and defined as

$$\mathcal{F}=\widehat{h}+\widehat{j}-\widehat{k} \tag{1.19}$$

with $\langle\psi_i|\widehat{h}|\psi_i\rangle=H_i$ (see Equation 1.15), and

$$\widehat{j}(\mathbf{x}_1)f(\mathbf{x}_1)=\sum_{k=1}^N\int \psi_k^*(\mathbf{x}_2)\psi_k(\mathbf{x}_2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\,d\mathbf{x}_2 f(\mathbf{x}_1) \tag{1.20}$$

$$\widehat{k}(\mathbf{x}_1)f(\mathbf{x}_1)=\sum_{k=1}^N\int \psi_k^*(\mathbf{x}_2)f(\mathbf{x}_2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\,d\mathbf{x}_2\psi_k(\mathbf{x}_1) \tag{1.21}$$

for $f(\mathbf{x}_1)$ an arbitrary function.

---

[1] The determinantal form ensures its anti-symmetry as exchanging two electron coordinates affects Ψ_{HF} by a factor of -1, and $\Psi_{\mathrm{HF}}=0$ if two electrons with the same coordinates are present.

The matrix $\boldsymbol{\epsilon}$ consists of Lagrange multipliers from the minimisation procedure. Its diagonal elements are

$$\epsilon_i \equiv \epsilon_{ii} = \langle \psi_i | \mathcal{F} | \psi_i \rangle = H_i + \sum_{j=1}^{N} (J_{ij} - K_{ij}) \, , \tag{1.22}$$

which compared with Equation 1.14, gives the total energy as

$$E_{HF} = \sum_{i=1}^{N} \epsilon_i - \frac{1}{2} \sum_{i,j=1}^{N} (J_{ij} - K_{ij}) \, . \tag{1.23}$$

The method described above is the *unrestricted open-shell Hartree-Fock method*. In the case of an even number of electrons the *restricted Hartree-Fock method* (RHF) can be used. Here the N orthonormal spin orbitals with spin functions $\sigma_i(s)$ are replaced by $N/2$ orbitals of form $\phi_i(\mathbf{r}_1)\alpha(s)$ and $N/2$ orbitals of form $\phi_i(\mathbf{r}_1)\beta(s)$. Eqn. 1.14 becomes

$$E_{RHF} = 2 \sum_{i}^{N/2} H_i + \sum_{i,j=1}^{N/2} (2J_{ij} - K_{ij}) \, , \tag{1.24}$$

where the spin dependence in H_i, J_{ij} and K_{ij} can now be ignored and the integration goes over $\mathbf{r}_1$ instead of $\mathbf{x}_1$. Note that the sums only go to $N/2$ as opposed to N before and hence the matrices in RHF are only 1/4 of the size of the matrices in the open-shell method.

The physical meaning of the eigenvalues ϵ_j was developed by Koopmans (1934). An important consequence of the Koopmans theorem is the ability to calculate excited states within the HF method.

Theorem 1 (Koopmans (1934)) *Assuming that the eigenstates do not vary after removal of one electron from the system, the ionisation energy I_m of the m-th electron is given by $I_m = -\epsilon_m$.*

When HF theory was applied to the homogeneous electron gas it was found that it predicts zero density of states at the Fermi energy. This is a major short-coming of the theory and due to the lack of electron-correlation, i.e., the spin-orbitals do not have a functional dependence on $\mathbf{r}_i - \mathbf{r}_j$. This problem can be solved by expressing Ψ as a linear combination of Slater determinants, known as Configuration Interaction (CI) method. For a given HF wavefunction Ψ_0, there are always other HF wavefunctions Ψ_a^r with the same symmetry and spin. The one-electron orbitals have different occupancies and Ψ_a^r means the Slater determinant is formed by replacing spin-orbital a in Ψ_0 with spin orbital r. In CI the new wavefunction is a sum of appropriately-weighted determinants

$$\Psi = c_0 \Psi_0 + \sum_{ra} c_a^r \Psi_a^r + \sum_{a<b, r<s} c_{ab}^{rs} \Psi_{ab}^{rs} + \ldots \tag{1.25}$$

In principle CI gives exact many-electron eigenstates and eigenvalues, however the computational requirements are even higher than for HF making this method impractical. The energy reduction obtained with CI is called correlation energy (E_{corr}). The exact energy of the many-electron system is

$$E_{exact} = E_{HF} + E_{corr}. \tag{1.26}$$

1.4 Variational Principle

In a computational calculation of the energy of a system the wavefunction has to be approximated either by evaluating it on a grid or by expanding it in terms of basis functions. The first method uses discretisation whereas the second makes use of the variational principle.

A wavefunction Ψ_0 is approximated by a sum of basis functions ϕ of a subspace of a Hilbert space,

$$\Psi_0 \approx \Psi_{\text{app}} = \sum_{i}^{M} c_i \phi_i(\mathbf{x}). \tag{1.27}$$

The expectation value of the Hamiltonian for the approximated wavefunction is

$$E_{\text{app}} = \frac{\langle \Psi_{\text{app}} | \mathcal{H} | \Psi_{\text{app}} \rangle}{\langle \Psi_{\text{app}} \mid \Psi_{\text{app}} \rangle}. \tag{1.28}$$

The variational principle states that this approximate value for the energy is always bigger than the true ground-state energy.

Theorem 2 (Variational principle) *The energy E_{app} computed from an approximate Ψ_{app} wavefunction is an upper bound to the true ground-state energy E_0. Full minimisation of the functional $E[\Psi]$ with respect to all allowed basis functions gives the true ground state Ψ_0 and energy $E_0 = E[\Psi_0]$; i.e.*

$$E_0 = \min_{\Psi} E[\Psi]. \tag{1.29}$$

This means the calculated energy is an upper bound for the ground-state energy and if the basis is chosen carefully should converge against it when the basis is increased.

For the ground-state the derivative of the energy

$$E_{\text{app}} = \frac{\sum_{i,j=1}^{M} c_i^* c_j H_{ij}}{\sum_{i,j=1}^{M} c_i^* c_j S_{ij}}, \tag{1.30}$$

where $H_{ij} = \langle \phi_i | \mathcal{H} | \phi_j \rangle$ and $S_{ij} = \langle \phi_i \mid \phi_j \rangle$, with respect to c_i has to vanish. This gives

$$\sum_{j=1}^{M} \left(H_{ij} - E_{\text{app}} S_{ij}\right) c_j = 0, \text{ for } i = 1, \dots, M. \tag{1.31}$$

This equation is a eigenvalue problem

$$\mathsf{S}^{-1} \cdot \mathsf{H} \cdot \mathbf{c} = E_{\text{app}} \mathbf{c}. \tag{1.32}$$

The matrix H is called the Hamiltonian and S is called overlap matrix. In the case of HF the spin-orbitals ψ_i are expanded as a linear combination of basis functions, usually atomic orbitals,

$$\psi_j(\mathbf{x}_1) = \sum_{i}^{M} c_{ij} \phi_i(\mathbf{x}_1), \tag{1.33}$$

and the Fock equation takes a matrix form. This generalised eigenvalue problem is known as *Roothaan equation* (Roothaan, 1951) and given by,

$$\mathsf{F} \cdot \mathsf{c} = \mathsf{S} \cdot \mathsf{c} \cdot \epsilon, \tag{1.34}$$

where S is again the overlap matrix $S_{ij} = \langle \phi_i \mid \phi_j \rangle$. The Fock Matrix F is

$$F_{ij} = H_{ij} + \sum_{\nu\mu} \left(\sum_{a} c_{\nu a} c_{\mu a} \right) \left(2 \langle \phi_i \phi_j | \mathcal{V}_{\text{ee}} | \phi_\nu \phi_\mu \rangle - \langle \phi_i \phi_\nu | \mathcal{V}_{\text{ee}} | \phi_\mu \phi_j \rangle \right) \tag{1.35}$$

and

$$H_{ij} = \langle \phi_i | \mathcal{T}_{\text{e}} + \mathcal{V}_{\text{en}} | \phi_j \rangle. \tag{1.36}$$

The equation is solved *self-consistently* by making a sensible initial guess at the set of $\psi_i(\mathbf{x}_1)$ and calculating the Coulomb and exchange potentials. These potentials are then fed back into the HF equations to calculate a new set of $\psi_i(\mathbf{x}_1)$. This cycle is repeated until the input and output potentials are sufficiently close. The big problem of the HF method is the huge computational effort, due to the four-centre integrals, for systems with more then a few atoms which makes this method impractical even without CI.

1.5 Density-Functional Theory

A computational far less demanding technique for electronic structure calculations in solids is Density-Functional Theory (DFT) (Hohenberg and Kohn, 1964; Kohn and Sham, 1965). For extensive reviews see Lundqvist (1983); Parr and Yang (1989); Jones and Gunnarsson (1989). Here the total energy E is a functional of the electronic charge density

$$n(\mathbf{r}_1) = N \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 \, ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N, \tag{1.37}$$

as opposed to the wavefunction in HF. And since in DFT the charge density and the Kohn-Sham one-electron orbitals only depend on the coordinates of one electron, $\mathbf{r}$ and $\mathbf{x}$ denote the coordinates of one particle from now on and not the coordinates of a set of N particles as before. N is simply obtained by quadrature of the charge density,

$$N = \int n(\mathbf{r}) d\mathbf{r}. \tag{1.38}$$

In HF the Hamiltonian of an N-electron system was completely fixed by an external potential v_{ext}, i.e. N and v_{ext} determine all properties for the ground state. Hohenberg and Kohn (1964) legitimised the use of the charge density instead of N and v_{ext}.

Theorem 3 (First Hohenberg-Kohn theorem) *The external potential is determined, within a trivial additive constant, by the electron density* $n(\mathbf{r})$.

The theorem states that the external potential and the number of electrons are uniquely defined by the charge density, hence the ground-state wavefunction is too.

The second Hohenberg-Kohn theorem states that for all densities n

$$E[n] = F\,[n] + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \, d\mathbf{r} \geq E[n_0], \tag{1.39}$$

where F is an universal (system-independent) functional, and accounts for electronic kinetic energy, electron-correlation and exchange-correlation, and n_0 is the ground-state density.

Theorem 4 (Second Hohenberg-Kohn theorem) *For a trial density* $\tilde{n}(\mathbf{r})$, *such that* $\tilde{n}(\mathbf{r}) \geq 0$ *and* $\int \tilde{n}(\mathbf{r}) \, d\mathbf{r} = N$ *(Hohenberg and Kohn, 1964),*

$$E_0 \leq E\,[\tilde{n}]\,. \tag{1.40}$$

This expression is analogous to Equation 1.29, and is the basis of a variational principle. Hence in this case

$$E_0 = \min_{\tilde{n}} E\,[\tilde{n}]\,. \tag{1.41}$$

The theorem was initially shown to be true for v-representable densities, but later Gilbert (1975) showed that it is true for all N-representable charge densities. A density $n(\mathbf{r})$ is N-representable if

$$n(\mathbf{r}) \geq 0, \int n(\mathbf{r}) d\mathbf{r} = N, \text{ and} \int \left| \nabla n^{1/2}(\mathbf{r}) \right|^2 d\mathbf{r} < \infty. \tag{1.42}$$

Despite the huge simplification of using the charge density as the variational variable, no approximations are made for the ground state.

If orthonormal orbitals $\psi_\lambda(\mathbf{x})$ are defined so that the charge density is

$$n(\mathbf{r}) = \sum_{\lambda,s} |\psi_\lambda(\mathbf{r}, s)|^2\,, \tag{1.43}$$

the total energy for the many-electron system as a functional of the charge density is given by (Parr and Yang, 1989)

$$E\,[n] = T_{\text{e}}\,[\psi_\lambda] + V_{\text{ee}}\,[n] + V_{\text{en}}\,[n] + E_{xc}\,[n]\,. \tag{1.44}$$

with

$$T_{\mathrm{e}}\left[\psi_\lambda\right] = -\frac{1}{2}\sum_{\lambda,s}\int \psi_\lambda^*(\mathbf{r},s)\nabla^2\psi_\lambda(\mathbf{r},s)\,d\mathbf{r} \tag{1.45}$$

$$V_{\mathrm{ee}}\left[n\right] = \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}' \tag{1.46}$$

$$V_{\mathrm{en}}\left[n\right] = -\int n(\mathbf{r})\sum \frac{Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha|}\,d\mathbf{r}. \tag{1.47}$$

The big problem with DFT is that the exact form of E_{xc} is unknown, but some standard approximations are currently used, see section 1.5.2.

1.5.1 Kohn-Sham (KS) equations

Using the variational principle and the fact that the orbitals ψ_λ are orthonormal the quantity

$$E - \sum_{\lambda,s}\epsilon_{\lambda,s}\left(\int |\psi_\lambda(\mathbf{r},s)|^2\,d\mathbf{x} - 1\right) \tag{1.48}$$

is minimised with respect to ψ_λ^* (and $\epsilon_{\lambda,s}$). This leads to the following set of one-electron Schrödinger equations (Kohn and Sham, 1965):

$$\left[-\frac{1}{2}\nabla^2 + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' - \sum_\alpha \frac{Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha|} + \frac{\delta E_{\mathrm{xc}}\left[n\right]}{\delta n(\mathbf{r})}\right]\psi_\lambda(\mathbf{r},s) = \epsilon_{\lambda,s}\psi_\lambda(\mathbf{r},s). \tag{1.49}$$

with the density n obtained by summing up all N occupied spin-orbital states,

$$n(\mathbf{r}) = \sum_{\lambda=1}^{N}\sum_s |\psi_\lambda(\mathbf{r},s)|^2\,. \tag{1.50}$$

The first three terms of Equation 1.49 are the kinetic energy, the Hartree energy and the external potential imposed by the ions, respectively.The fourth term lumps together all remaining many-body effects in the form of an exchange-correlation functional. One main pillar of DFT is the existence of an universal E_{xc} density functional (depending only on the electron density n), leading to the knowledge of the *exact* ground-state total energy and density.

Equations 1.49 and 1.50 – known as *Kohn-Sham equations* (Kohn and Sham, 1965) – are solved in a self-consistent loop as described in section 1.3.

1.5.2 The exchange-correlation functional

The most common approach to overcome the problem of the unknown functional for the exchange-correlation energy is the local density approximation (LDA), or local spin-density approximation (LSDA) for non-zero spin systems (Kohn and Sham, 1965; von Barth and Hedin, 1972; Perdew and Zunger, 1981). Here, it is assumed that the exchange-correlation energy is local. Normally this energy is separated into exchange and correlation parts, where in LSDA notation,

$$E_{\mathrm{xc}}[n_\uparrow,n_\downarrow] = E_{\mathrm{x}}[n_\uparrow,n_\downarrow] + E_{\mathrm{c}}[n_\uparrow,n_\downarrow]. \tag{1.51}$$

The ’up’ and ’down’ spin densities are defined by

$$n_s(\mathbf{r}) = \sum_{\lambda=1}^{N}|\psi_\lambda(\mathbf{r},s)|^2\,, \tag{1.52}$$

where s is either ↑ or ↓.

The exchange functional of a homogeneous electron gas is linear with respect to n, and its analytic form is given by (von Barth and Hedin, 1972),

$$E_{\mathrm{x}}[n_\uparrow,n_\downarrow] = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\left(n_\uparrow^{4/3} + n_\downarrow^{4/3}\right). \tag{1.53}$$

The correlation part is however more complicated. In the high-density regime, perturbation theory yields one expression (Perdew and Zunger, 1981), while in the low-density regime, a Green function quantum Monte Carlo method yields another (Ceperley, 1978; Ceperley and Alder, 1980). A simple parameterised functional form is then fitted to the numerical results. Several parameterisations are available, namely those from Perdew and Zunger (1981) (PZ), Vosko *et al.* (1980) (VWN) and Perdew and Wang (1992) (PW).

The generalised gradient approximation (GGA), a first-order expansions of E_{xc} in the density, includes terms depending on $\boldsymbol{\nabla} n$ (Perdew, 1991; Perdew *et al.*, 1996*a*,*b*).

1.6 Pseudopotentials

A known fact about atoms is, that the electronic core states are highly localised and have almost no effect on the chemical properties of atoms. Not only would the incorporation of core states in a calculation mean that the number of basis functions required to fit all states would become prohibitively large, but they also pose the problem that to satisfy the orthogonality condition the wavefunctions close to the core oscillate strongly. This makes them extremely difficult to model. Furthermore a full electron potential implies large total energies, therefore leading to large errors when comparing similar systems.

The core electrons are therefore removed by replacing the atomic potential by a *pseudopotential* where the core electrons create a screening effective potential $\mathcal{V}^{\text{ps}}(\mathbf{r})$. This pseudopotential has the disadvantage that it includes the potential from the valence states and is therefore not *transferable* . To construct a transferable pseudopotential an ionic potential must be constructed by subtracting the potential from the valence states,

$$\mathcal{V}^{\text{ps}}_{\text{ion}}(\mathbf{r}) = \mathcal{V}^{\text{ps}}(\mathbf{r}) - \int \frac{n^{\text{ps}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' - E_{\text{xc}}[n^{\text{ps}}(\mathbf{r})], \tag{1.54}$$

where

$$n^{\text{ps}}(\mathbf{r}) = \sum_{\lambda} \left|\psi^{\text{ps}}_{\lambda}(\mathbf{r})\right|^2 . \tag{1.55}$$

and the sum is over the occupied valence eigenstates of the pseudopotential $\mathcal{V}^{\text{ps}}(\mathbf{r})$.

The approximations made when using pseudopotentials are:

1. A one-electron picture is used to divide electrons into core and valence sets.
2. The *frozen core approximation.* The one-electron states in the core do not change when transfered from one environment to another but remain frozen.
3. The *small core approximation* assumes that the core and valence states do not overlap significantly and hence the approximation $E_{\text{xc}}(n_{core} + n_{valence}) = E_{\text{xc}}(n_{core}) + E_{\text{xc}}(n_{valence})$ is made. This is not always the case and a technique called *non-linear core corrections* (Louie *et al.*, 1982) is sometimes used to improve results.

Another important aspect about pseudopotentials is norm conservation. It means that outside a certain cutoff radius the pseudo-wavefunctions are *identical* to the real all-electron wavefunction. One type of pseudopotentials that are norm conserving are so called BHS-pseudopotentials of Bachelet, Hamann and Schlüter (Hamann *et al.*, 1982). These were tabulated by Bachelet *et al.* (1982) for elements between H and Pu. The calculations in this thesis make use of these pseudopotentials.

Detailed information on the pseudopotential method has been extensively reported (Harrison, 1966; Brust, 1968; Stoneham, 1975; Heine, 1970; Pickett, 1989).

1.7 Brillouin-Zone Sampling

To obtain physical quantities like the charge density 1.62 in the supercell method the Brillouin-zone (BZ) has to be integrated. The integrand function $f(\mathbf{k})$ is periodic in reciprocal space, and has no simple analytic form. To avoid numerical integration over a dense mesh, several schemes were developed. A clear scheme was proposed by Monkhorst and Pack (1976); Pack and Monkhorst (1977) (MP). The main advantages of their method are its simplicity, and convergence is verifiable.

The average $\bar{f}$, or the integral of f over the BZ with volume $(2\pi)^3/\Omega$, is simply given by

$$\bar{f} = \frac{\Omega}{(2\pi)^3} \int f(\mathbf{k})\, d\mathbf{k} \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{k}_n). \tag{1.56}$$

where the points in the reciprocal space of the $I \times J \times K$ grid where the function f is evaluated are given by

$$\mathbf{k}(i,j,k) = \frac{2i - I - 1}{2I}\mathbf{g}_1 + \frac{2j - J - 1}{2J}\mathbf{g}_2 + \frac{2k - K - 1}{2K}\mathbf{g}_3. \tag{1.57}$$

with integers I, J and $K \geq 1$.[2]

If the supercell has a high enough symmetry some of the **k**-points become equivalent and the redundant **k**-points can be removed when the remaining **k**-points are weighted by a factor w_n which is the number of equivalent **k**-points to **k**-point n.

Finally, the average $\bar{f}$, or the integral of f over the BZ with volume $(2\pi)^3/\Omega$, is simply given by

$$\bar{f} = \frac{\Omega}{(2\pi)^3} \int f(\mathbf{k})\, d\mathbf{k} \approx \frac{\sum_n w_n f(\mathbf{k}_n)}{\sum_n w_n} \tag{1.58}$$

where the sum now only goes over the remaining inequivalent **k**-points.

1.8 AIMPRO Method

AIMPRO is the name of the DFT code implementation used in this work. It uses a real-space Gaussian type basis set (Jones and Briddon, 1999) within a supercell geometry. This choice of boundary conditions has the advantage that surface effects are completely avoided. However, one problem of the supercell approach is that the energies have to be obtained by Brillouin-zone (BZ) integration which makes an expansion of the wavefunction in reciprocal-space necessary.

1.8.1 Basis functions

The supercell uses a set of Bloch basis functions $B_{\mathbf{k}i}(\mathbf{r})$ that are built from a set of Cartesian-Gaussian functions ϕ_j centred at atomic site locations $\mathbf{R}_i$ with N_L lattice vectors $\mathbf{L}_n$,

$$B_{\mathbf{k}i}(\mathbf{r}) = \frac{1}{\sqrt{N_L}} \sum_n \sum_j \phi_j(\mathbf{r} - \mathbf{R}_i - \mathbf{L}_n)\, e^{\mathrm{i}\mathbf{k}\,\mathbf{L}_n}. \tag{1.59}$$

with $\mathbf{k}$ as a reciprocal-space vector. The localised orbitals are given by [3]

$$\phi_j(\mathbf{r} - \mathbf{R}_i) = \phi_\alpha^l(\mathbf{r} - \mathbf{R}_i) = (x - R_{i,x})^{l_1}(y - R_{i,y})^{l_2}(z - R_{i,z})^{l_3} e^{-\alpha(\mathbf{r}-\mathbf{R}_i)^2}, \tag{1.60}$$

where each l_i is an integer ≥ 0.

The Kohn-Sham orbitals $\psi_{\mathbf{k}\lambda}$ are then expanded in terms of all $B_{\mathbf{k}i}$ basis functions,

$$\psi_{\mathbf{k}\lambda}(\mathbf{r}) = \sum_i c^{\mathbf{k}}_{\lambda\, i} B_{\mathbf{k}i}(\mathbf{r}), \tag{1.61}$$

The charge-density is then

$$n(\mathbf{r}) = \sum_{i,j,\mathbf{k}} b_{ij}(\mathbf{k}) B^*_{\mathbf{k}i}(\mathbf{r}) B_{\mathbf{k}j}(\mathbf{r}), \tag{1.62}$$

$$b_{ij}(\mathbf{k}) = \sum_\lambda f_{\mathbf{k}\lambda} (c^{\mathbf{k}}_{\lambda\, i})^* c^{\mathbf{k}}_{\lambda\, j}, \tag{1.63}$$

where $f_{\mathbf{k}\lambda}$ is the occupancy of the $\mathbf{k}\lambda$ state. This should be 2 for a filled level and 0 for an empty one in the case of a spin averaged calculation. For a spin polarised calculation it would be 1 and 0, respectively. If a Fermi smearing is used each level can have a partial occupancy between 0 and either 1 or 2 for spin polarised and spin averaged calculations,

---

[2] In the original paper, $I = J = K$ is imposed, and for a particular I the scheme is normally referred as MP-I^3.

[3] The ϕ_α^l denote the fact that there are several basis functions with different angular momentum l but the same exponent α in AIMPRO.

respectively. Furthermore the filling can be chosen to be either metallic or non-metallic. In the latter case the $f_{\mathbf{k}\lambda}$ are independent of $\mathbf{k}$.

The $B_{\mathbf{k}i}(\mathbf{r})$ satisfy Bloch's theorem as

$$B_{\mathbf{k}i}(\mathbf{r}+\mathbf{L}_n) = e^{\mathrm{i}\mathbf{k}\mathbf{L}_n} B_{\mathbf{k}i}(\mathbf{r}). \tag{1.64}$$

Thus $B_{\mathbf{k}i}$ and $B_{\mathbf{k}'j}$ are orthogonal and the Hamiltonian, and other translational symmetric functions, have matrix elements diagonal in $\mathbf{k}$-space. However, the basis functions for the same $\mathbf{k}$ are not orthogonal, the overlap between them is $S^0_{ij}(\mathbf{k})$, where

$$\begin{aligned} S^{\mathbf{g}}_{ij}(\mathbf{k}) &= \int B^*_{\mathbf{k}i}(\mathbf{r}) e^{-\mathrm{i}\mathbf{g}\mathbf{r}} B_{\mathbf{k}j}(\mathbf{r}) d\mathbf{r} \\ &= \sum_{ijn} e^{\mathrm{i}\mathbf{k}\mathbf{L}_n} \int \phi_i(\mathbf{r}-\mathbf{R}_i) e^{-\mathrm{i}\mathbf{g}\mathbf{r}} \phi_j(\mathbf{r}-\mathbf{R}_j-\mathbf{L}_n) d\mathbf{r}, \end{aligned} \tag{1.65}$$

and $\mathbf{g}$ is a reciprocal lattice vector. To normalise the KS orbitals the $c^{\mathbf{k}}_{\lambda i}$ coefficients satisfy

$$\sum_{ij} (c^{\mathbf{k}}_{\lambda i})^* c^{\mathbf{k}}_{\lambda j} S^0_{ij}(\mathbf{k}) = 1. \tag{1.66}$$

The use of localised basis sets can lead to over-completeness which induces numerical instabilities. The advantage of Gaussian-like functions however is that their related integrals can be found analytically.

1.8.2 Split BHS-pseudopotentials

To calculate Coulombic and dipolar interactions (like in the Madelung energy) slowly converging sums have to be evaluated. The reason for this is that series like $1/r^p$ (with $p \leq 2$) converge very slowly if at all as r goes to infinity. To solve this problem, Ewald (1921) proposed a method in which the series is split into a fast-converging and a slowly-converging part,

$$\sum_L \frac{1}{r_L^p} = \sum_L \frac{\mathrm{erfc}(\alpha r_L)}{r_L^p} + \sum_L \frac{\mathrm{erf}(\alpha r_L)}{r_L^p}. \tag{1.67}$$

Evaluation of the first term is straightforward as erfc(x) converges to zero when $x \to \infty$. The last term, slowly converging in real-space, is Fourier transformed, and consequently its terms are now *short-ranged* in reciprocal space. The parameter α controls the transition between the real-space and reciprocal-space summations. Although the analytical result does not depend on the choice of α, in practise this is not true as a consequence of the summations being taken over a finite set of lattice vectors. An extensive treatment of this method was given by Leeuw *et al.* (1980).

In AIMPRO the BHS-pseudopotentials are therefore split into a local term and a semi non-local term

$$V^{\mathrm{ps}}_{\mathrm{en}}(\mathbf{r}) = V^{loc}_{\mathrm{en}}(\mathbf{r}) + V^{nonl}_{\mathrm{en}}(\mathbf{r}). \tag{1.68}$$

The semi non-local part can be written as

$$V^{nonl}_{\mathrm{en}}(\mathbf{r}) = \sum_{lm} \left|Y_{lm}\right\rangle V^{nonl}_l(r) \left\langle Y_{lm}\right|, \tag{1.69}$$

where Y_{lm} is a spherical harmonic of angular momentum l. The local term is of the form

$$V^{loc}_{\mathrm{en}}(\mathbf{r}) = -\sum_{an} \frac{\mathbf{Z}_a}{|\mathbf{r}-\mathbf{R}_a-\mathbf{L}_n|} \mathrm{erf}(c_a\,|\mathbf{r}-\mathbf{R}_a-\mathbf{L}_n|), \tag{1.70}$$

where the c_a are coefficients obtained during the fitting of the pseudopotential. This part of the pseudopotential arises from a Gaussian charge density

$$\rho(r) = \sum_{an} Z_a \left(\sqrt{c_a/\pi}\right)^{3/2} e^{-c_a|\mathbf{r}-\mathbf{R}_a-\mathbf{L}_n|^2}. \tag{1.71}$$

It is included in the electrostatic energy and evaluated in $\mathbf{k}$-space whereas the non-local term is included in the electron-nuclear energy and evaluated in real-space.

1.8.3 Reciprocal-space basis

As the Hartree (including the local pseudopotential) and exchange-correlation energies are evaluated in $\mathbf{k}$-space the charge density has to be Fourier transformed. The quality of the transformation is dependent on the grid of $\mathbf{g}$-vectors. In practise a uniform grid of vectors inside a sphere of radius g_{cut} defining a *cut-off energy* is used,

$$E_{\mathrm{cut}} = \frac{1}{2} g^2_{\mathrm{cut}}. \tag{1.72}$$

This cut-off energy is usually increased until the total energy has attained a converged value.

1.8.4 The total energy

The total energy is the energy per supercell and includes the following terms:

$$E = T_{\mathrm{e}}\left[\psi_\lambda\right] + V_{\mathrm{en}}^{\mathrm{ps}}\left[n\right] + V_{\mathrm{ee}}\left[n\right] + E_{xc}\left[n\right] + E_{ion}, \tag{1.73}$$

here E_{ion} is the Madelung energy of point ions corrected by the pseudopotential term. After the pseudopotential is split in local and non-local parts the contributions to the total energy are

$$E = \frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k})\{T_{ij}(\mathbf{k}) + V_{ij}^{nonl}(\mathbf{k})\} + E_e + E_{XC} + E_{ion}. \tag{1.74}$$

The individual contributions to the energy are described below:

Kinetic energy T_e

The kinetic energy can easily be expressed by the coefficients

$$T_e = \frac{1}{N_L}\sum_{\lambda\mathbf{k}} f_{\lambda\mathbf{k}} \int \psi_{\mathbf{k}\lambda}^*(\mathbf{r}) \left(\frac{-1}{2}\right) \nabla^2 \psi_{\mathbf{k}\lambda}(\mathbf{r})\, d\mathbf{r} \tag{1.75}$$

$$= \frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) T_{ij}(\mathbf{k}) \tag{1.76}$$

where the kinetic matrix is

$$T_{ij}(\mathbf{k}) = \int B_{\mathbf{k}i}^*(\mathbf{r})\Big(\frac{-1}{2}\Big)\nabla^2 B_{\mathbf{k}j}(\mathbf{r}) d\mathbf{r} \tag{1.77}$$

$$= \sum_n e^{\mathrm{i}\mathbf{k}\mathbf{L}_n} \int \phi_i(\mathbf{r}-\mathbf{R}_i)\Big(\frac{-1}{2}\Big)\nabla^2\phi_j(\mathbf{r}-\mathbf{R}_j-\mathbf{L}_n)\, d\mathbf{r} \tag{1.78}$$

Electron-nuclear potential V_{en}^{nonl}

The non-local energy contribution of the pseudoatoms is

$$V_{\mathrm{en}}^{nonl} = \frac{1}{N_L}\sum_{\lambda\mathbf{k}} f_{\lambda\mathbf{k}} \int \psi_{\mathbf{k}\lambda}^*(\mathbf{r}) V_{\mathrm{en}}^{nonl}(\mathbf{r}) \psi_{\mathbf{k}\lambda}(\mathbf{r}) d\mathbf{r} \tag{1.79}$$

$$= \frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) V_{ij}^{nonl}(\mathbf{k}). \tag{1.80}$$

Electrostatic energy E_e

The electrostatic energy term is combined with the local pseudopotential term as

$$E_e = V_{\mathrm{ee}} + V_{\mathrm{en}}^{loc} = \frac{1}{2N_L}\int \frac{(n(\mathbf{r}_1)-\rho(\mathbf{r}_1))(n(\mathbf{r}_2)-\rho(\mathbf{r}_2))}{|\mathbf{r}_1-\mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - E_{lr}. \tag{1.81}$$

E_{lr} represents a long-range electrostatic energy arising from dipole moments and needs separate discussion. E_e can be evaluated by Fourier transformation using

$$\frac{1}{r} = \int \frac{4\pi e^{\mathrm{i}\mathbf{q}\,\mathbf{r}}}{\mathbf{q}^2}\frac{d\mathbf{q}}{8\pi^3} \tag{1.82}$$

$$E_e = 2\pi N_L \Omega^2 \int \frac{(n_{\mathbf{q}}-\rho_{\mathbf{q}})(n_{-\mathbf{q}}-\rho_{-\mathbf{q}})}{\mathbf{q}^2}\frac{d\mathbf{q}}{8\pi^3} - E_{lr} \tag{1.83}$$

Here $n_{\mathbf{q}},\ \rho_{\mathbf{q}}$ are transforms :

$$n_{\mathbf{q}} = \frac{1}{N_L\Omega}\int n(\mathbf{r}) e^{-\mathrm{i}\mathbf{q}\mathbf{r}} d\mathbf{r} = \frac{1}{N_L\Omega}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) S_{ij}^{\mathbf{q}}(\mathbf{k}) \tag{1.84}$$

$$\rho_{\mathbf{q}} = \frac{1}{N_L\Omega}\sum_{an} Z_a e^{-\mathbf{q}^2/4c_a} e^{-\mathrm{i}\mathbf{q}(\mathbf{R}_a+\mathbf{L}_n)} \tag{1.85}$$

For a lattice, $\mathbf{q}$ must be equal to a reciprocal lattice vector $\mathbf{g}$ if the function to be transformed is periodic. Then using

$$\int \frac{d\mathbf{q}}{8\pi^3} = \frac{1}{N_L \Omega} \sum_{\mathbf{g}} \tag{1.86}$$

n and ρ can be written as

$$n(\mathbf{r}) = \sum_{\mathbf{g}} n_{\mathbf{g}} e^{\mathrm{i}\mathbf{g}\mathbf{r}} \tag{1.87}$$

$$\rho_{\mathbf{g}} = \frac{1}{\Omega} \sum_a Z_a e^{-\mathbf{g}^2/4c_a} e^{-\mathrm{i}\mathbf{g}\mathbf{R}_a}. \tag{1.88}$$

Now the long-range electrostatic energy can be defined as

$$E_{lr} = \lim_{\mathbf{g}\to 0} 2\pi\Omega \frac{(n_{\mathbf{g}} - \rho_{\mathbf{g}})(n_{-\mathbf{g}} - \rho_{-\mathbf{g}})}{\mathbf{g}^2} \tag{1.89}$$

and the electrostatic energy finally expressed as

$$E_e = 2\pi\Omega \sum_{\mathbf{g}'} \frac{(n_{\mathbf{g}} - \rho_{\mathbf{g}})(n_{-\mathbf{g}} - \rho_{-\mathbf{g}})}{\mathbf{g}^2}. \tag{1.90}$$

The sum over $\mathbf{g}'$ excludes the $\mathbf{g} = 0$ singular contribution. This can be argued away as it is a long range effect and such macroscopic electrical fields are outside the theory. However, the $\mathbf{g} = 0$ contribution to the exchange-correlation term is retained. As $\mathbf{g} \to 0$,

$$\begin{aligned}
n_{\mathbf{g}} \to & \frac{1}{\Omega} \left[N_e - \frac{1}{N_L} \sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) \int B^*_{\mathbf{k}i} \mathrm{i}\mathbf{g}\mathbf{r} B_{\mathbf{k}j} d\mathbf{r} - \frac{1}{2N_L} \sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) \int B^*_{\mathbf{k}i} (\mathrm{i}\mathbf{g}\mathbf{r})^2 B_{\mathbf{k}j} d\mathbf{r} + \dots \right] \\
= & \frac{1}{\Omega} \left[N_e - \frac{1}{N_L} \sum_{ijn\mathbf{k}} b_{ij}(\mathbf{k})\, e^{\mathrm{i}\mathbf{k}\mathbf{L}_n} \int \phi_i(\mathbf{r} - \mathbf{R}_i)(\mathbf{r}\mathbf{g}\mathrm{i}) \phi_j(\mathbf{r} - \mathbf{R}_j - \mathbf{L}_n) d\mathbf{r} - \right. \\
& \left. - \frac{1}{2N_L} \sum_{ij\mathbf{k}} b_{ij}(\mathbf{k})\, e^{\mathrm{i}\mathbf{k}\mathbf{L}_n} \int \phi_i(\mathbf{r} - \mathbf{R}_i)(\mathbf{r}\mathbf{g})^2 \phi_j(\mathbf{r} - \mathbf{R}_j - \mathbf{L}_n) d\mathbf{r} + \dots \right]
\end{aligned} \tag{1.91}$$

$$\begin{aligned}
\rho_{\mathbf{g}} \to & \frac{1}{\Omega} \sum_a Z_a e^{-\mathrm{i}\mathbf{g}\mathbf{R}_a - \mathbf{g}^2/4c_a} \\
= & \frac{1}{\Omega} \left[\sum_a Z_a - Z_a(\mathrm{i}\mathbf{g}\mathbf{R}_a) - \frac{1}{2} Z_a \left\{ (\mathbf{g}\mathbf{R}_a)^2 + \mathbf{g}^2/2c_a \right\} + \dots \right]
\end{aligned} \tag{1.92}$$

Thus $(n_{\mathbf{g}} - \rho_{\mathbf{g}})/\mathbf{g}^2$ is well-defined when $\mathbf{g} \to 0$ if the cell is neutral and has no dipole or quadrupole moment, while E_{lr} which can be written as

$$E_{lr} = \lim_{\mathbf{g}\to 0} -2\pi\Omega \left(\frac{n_{\mathbf{g}} - \rho_{\mathbf{g}}}{g} \right)^2 \tag{1.93}$$

is well-defined when $\mathbf{g} \to 0$ if the cell is neutral and has no dipole moment.

Exchange-correlation energy E_{XC}

The exchange-correlation term is also evaluated by Fourier expansion as the exchange-correlation density is periodic.

$$\epsilon^{XC}(\mathbf{r}) = \sum_{\mathbf{g}} \epsilon^{XC}_{\mathbf{g}} e^{\mathrm{i}\mathbf{g}\mathbf{r}} \tag{1.94}$$

where

$$\epsilon^{XC}_{\mathbf{g}} = \frac{1}{N_L \Omega} \int \epsilon^{XC}(\mathbf{r}) e^{-\mathrm{i}\mathbf{g}\mathbf{r}} d\mathbf{r} \tag{1.95}$$

$$E_{XC} = \frac{1}{N_L} \int \epsilon_{XC}(n(\mathbf{r}))\, n(\mathbf{r})\, d\mathbf{r} = \Omega \sum_{\mathbf{g}} \epsilon^{XC}_{-\mathbf{g}} n_{\mathbf{g}}. \tag{1.96}$$

Define the exchange-correlation potential by

$$\mu^{XC}(\mathbf{r}) = \frac{d}{dn}(n\,\epsilon^{XC}(n)) \tag{1.97}$$

$$\mu^{XC}_{\mathbf{g}} = \frac{1}{N_L\Omega}\int \mu^{XC}(\mathbf{r})e^{-\mathrm{i}\mathbf{g}\mathbf{r}}\,d\mathbf{r}. \tag{1.98}$$

If any change is made to the charge density, then

$$\Delta E_{XC} = \frac{1}{N_L}\int \mu^{XC}(\mathbf{r})\Delta n d\mathbf{r} \tag{1.99}$$

$$\Delta E_e = \frac{1}{N_L}\int \frac{(n(\mathbf{r}_1)-\rho(\mathbf{r}_1))(\Delta n(\mathbf{r}_2)-\Delta\rho(\mathbf{r}_2))}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1\,d\mathbf{r}_2 - \Delta E_{lr} \tag{1.100}$$

The ion-ion interaction

The ion-term includes the Madelung energy of point ions corrected by the pseudopotential term which cancels the long range Coulomb field,

$$E_{ion} = \frac{1}{2}\sum_{(b\,a\,n)'}\frac{Z_bZ_a}{|\mathbf{R}_b-\mathbf{R}_a-\mathbf{L}_n|} - \frac{1}{2N_L}\int\frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1d\mathbf{r}_2 \tag{1.101}$$

$$\begin{aligned} &= \frac{1}{2}\sum_{(b\,a\,n)'}\frac{Z_bZ_a}{|\mathbf{R}_b-\mathbf{R}_a-\mathbf{L}_n|}\,\mathrm{erfc}\left(c_ac_b\frac{|\mathbf{R}_b-\mathbf{R}_a-\mathbf{L}_n|}{c_a+c_b}\right) - \\ &-\sum_a Z_a^2\frac{c_ac_b}{(c_a+c_b)\sqrt{\pi}}, \end{aligned} \tag{1.102}$$

where the sum goes over all a, b, n so that $|\mathbf{R}_b - \mathbf{R}_a - \mathbf{L}_n| \neq 0$.

1.8.5 Self-consistency cycle

In a calculation the total energy E has to be minimised with respect to $c^{\mathbf{k}}_{\lambda j}$ subject to orthonormality. As in Eqn. 1.48 this can be accomplished by minimising E' with respect to both $(c^{\mathbf{k}}_{\lambda i})^*$ and $E_{\lambda\mathbf{k}}$. Here,

$$E' = E - \frac{1}{N_L}\sum_{\lambda\mathbf{k}} f_{\lambda\mathbf{k}}E_{\lambda\mathbf{k}}\{\sum_{ij}(c^{\mathbf{k}}_{\lambda i})^*c^{\mathbf{k}}_{\lambda j}S^0_{ij}(\mathbf{k}) - 1\}. \tag{1.103}$$

Minimising E' with respect to $E_{\lambda\mathbf{k}}$ gives the normalisation condition. Orthogonality is imposed as $\psi_{\mathbf{k}\lambda}$ are eigenfunctions. A change in $(c^{\mathbf{k}}_{\lambda i})^*$ results in a change of $b_{ij}(\mathbf{k})$ as

$$\Delta b_{ij}(\mathbf{k}) = f_{\lambda\mathbf{k}}(\Delta c^{\mathbf{k}}_{\lambda i})^*c^{\mathbf{k}}_{\lambda j}. \tag{1.104}$$

The change in E when $b_{ij}(\mathbf{k})$ is varied is

$$\Delta E = \frac{1}{N_L}\sum_{ij\mathbf{k}}\{T_{ij}(\mathbf{k}) + V^{nonl}_{ij}(\mathbf{k}) + V^H_{ij}(\mathbf{k}) + V^{XC}_{ij}(\mathbf{k})\}\Delta b_{ij}(\mathbf{k}). \tag{1.105}$$

V^H and V^{XC} arise from changes to E_e and E_{XC}, respectively. Thus

$$\begin{aligned} V^H_{ij}(\mathbf{k}) &= N_L\frac{\partial E_e}{\partial b_{ij}(\mathbf{k})} = 4\pi N_L\Omega\sum_{\mathbf{g}'}\frac{(n_{-\mathbf{g}}-\rho_{-\mathbf{g}})}{\mathbf{g}^2}\frac{\partial n_{\mathbf{g}}}{\partial b_{ij}(\mathbf{k})} \\ &= 4\pi\sum_{\mathbf{g}'}\frac{(n_{-\mathbf{g}}-\rho_{-\mathbf{g}})}{\mathbf{g}^2}S^{\mathbf{g}}_{ij}(\mathbf{k}) \end{aligned} \tag{1.106}$$

$$\begin{aligned} V^{XC}_{ij}(\mathbf{k}) &= N_L\frac{\partial E_{XC}}{\partial b_{ij}(\mathbf{k})} = N_L\Omega\sum_{\mathbf{g}}\mu^{XC}_{-\mathbf{g}}\frac{\partial n_{\mathbf{g}}}{\partial b_{ij}(\mathbf{k})} \\ &= \sum_{\mathbf{g}}\mu^{XC}_{-\mathbf{g}}S^{\mathbf{g}}_{ij}(\mathbf{k}). \end{aligned} \tag{1.107}$$

The first is not really a Hartree term as it involves a local pseudopotential term as well. There is a complication regarding the average potential. A contribution to $V^H_{ij}(\mathbf{k})$ comes from the limiting value as $\mathbf{g} \to 0$ of

$$\lim_{\mathbf{g}\to 0} 4\pi\frac{(n_{-\mathbf{g}}-\rho_{-\mathbf{g}})}{\mathbf{g}^2}S^{\mathbf{g}}_{ij}(\mathbf{k}) = V_{av}S^0_{ij}(\mathbf{k}) \tag{1.108}$$

However, if the average potential is included here, then it must be included in E_e, and should cancel out. However AIMPRO does not include this term in $V_{ij}^H(\mathbf{k})$, i.e. it does not shift the eigenvalues.

Putting these changes into E' gives the usual eigenvalue equation,

$$\sum_j \{T_{ij}(\mathbf{k}) + V_{ij}^{nonl}(\mathbf{k}) + V_{ij}^H(\mathbf{k}) + V_{ij}^{XC}(\mathbf{k}) - E_{\lambda\mathbf{k}} S_{ij}^0(\mathbf{k})\} c_{\lambda j}^{\mathbf{k}} = 0. \tag{1.109}$$

In the same way as for Eqn. 1.34 the solution is found in a self-consistency cycle. The matrix elements are evaluated from an initial guess for the wavefunction coefficients $(c_{\lambda j}^{\mathbf{k}})^0$ and new eigenvalues $(c_{\lambda j}^{\mathbf{k}})'$ are calculated. Then a weighted average of the input and output coefficients is taken,

$$(c_{\lambda j}^{\mathbf{k}})^1 = (1-w)(c_{\lambda j}^{\mathbf{k}})' + w(c_{\lambda j}^{\mathbf{k}})^0. \tag{1.110}$$

The matrix elements are now calculated from the new coefficients $(c_{\lambda j}^{\mathbf{k}})^1$ and the procedure is repeated till the difference between input and output coefficients and hence charge density is below a defined threshold.

The total energy can be written, after multiplying Eqn. 1.109 by $f_{\mathbf{k}\lambda}(c_{\lambda i}^{\mathbf{k}})^*/N_L$ and summing, as

$$E = \frac{1}{N_L}\sum_{\lambda\mathbf{k}} E_{\lambda\mathbf{k}} f_{\mathbf{k}\lambda} + E_{ion} + E_e - \frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) V_{ij}^H(\mathbf{k}) + E_{XC} - \frac{1}{N_L}\sum_{ij\mathbf{k}} V_{ij}^{XC}(\mathbf{k}) b_{ij}(\mathbf{k}) \tag{1.111}$$

The first term is the one-electron term (band structure) and the others have obvious meaning. If the average potential is included in the band structure it should also be included in V^H. Thus it makes no net contribution. Both series in E_e and E_{ion} converge and there is no need to worry about their cancellation. One has however to worry about stability of this expression for large cells where g can be very small and E_{lr} significant.

1.8.6 Atomic forces

The force on an atom is the change of total energy with respect to the change in ion position. A change in the ion positions $\mathbf{R}_j$ introduces a change to the electron density in two ways, firstly through a change in $b_{ij}(k)$ and secondly through a change in the basis functions B_{ki}. Thus

$$\Delta n = \sum_{ij\mathbf{k}} b_{ij}(\mathbf{k}) \Delta\{B_{\mathbf{k}i}^*(\mathbf{r}) B_{\mathbf{k}j}(\mathbf{r})\} + \sum_{ij\mathbf{k}} B_{\mathbf{k}i}^*(\mathbf{r}) B_{\mathbf{k}j}(\mathbf{r}) \Delta b_{ij}(\mathbf{k}) \tag{1.112}$$

The changes to E_e and E_{XC} are

$$\Delta E_e = \frac{1}{N_L}\int \frac{(n(\mathbf{r}_1) - \rho(\mathbf{r}_1))(\Delta n(\mathbf{r}_2) - \Delta\rho(\mathbf{r}_2))}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - \Delta E_{lr} \tag{1.113}$$

$$\Delta E_{XC} = \frac{1}{N_L}\int \mu^{XC}(\mathbf{r}) \Delta n d\mathbf{r} \tag{1.114}$$

With the Fourier expansions of $1/r$ and $\mu^{XC}(\mathbf{r})$ the individual components become

$$\Delta E_e = 4\pi\Omega \sum_{\mathbf{g}'} \frac{(n_{-\mathbf{g}} - \rho_{-\mathbf{g}})((\Delta n)_{\mathbf{g}} - (\Delta\rho)_{\mathbf{g}})}{\mathbf{g}^2}, \tag{1.115}$$

$$(\Delta n)_{\mathbf{g}} = \frac{1}{N_L\Omega}\int e^{-i\mathbf{g}\mathbf{r}} \Delta n(\mathbf{r}) d\mathbf{r} = \frac{1}{N_L\Omega}\sum_{ij\mathbf{k}} \{S_{ij}^{\mathbf{g}}(\mathbf{k}) \Delta b_{ij}(\mathbf{k}) + b_{ij}(\mathbf{k}) \Delta S_{ij}^{\mathbf{g}}(\mathbf{k})\}, \tag{1.116}$$

$$\Delta S_{ij}^{\mathbf{g}} = \int e^{-i\mathbf{g}.\mathbf{r}} \Delta\{B_{\mathbf{k}i}^*(\mathbf{r}) B_{\mathbf{k}j}(\mathbf{r})\}, \tag{1.117}$$

$$\begin{aligned} \Delta E_e = & \frac{1}{N_L}\sum_{ij\mathbf{k}} V_{ij}^H(\mathbf{k}) \Delta b_{ij}(\mathbf{k}) + \\ & + 4\pi \sum_{\mathbf{g}'} \frac{(n_{-\mathbf{g}} - \rho_{-\mathbf{g}})}{\mathbf{g}^2} \left(\frac{1}{N_L}\sum_{\mathbf{k}} b_{ij}(\mathbf{k}) \Delta S_{ij}^{\mathbf{g}}(\mathbf{k}) - \Omega(\Delta\rho)_{\mathbf{g}} \right), \end{aligned} \tag{1.118}$$

$$\begin{aligned} (\Delta\rho)_{\mathbf{g}} = & \frac{1}{N_L\Omega}\int e^{-i\mathbf{g}\mathbf{r}} \Delta\rho(\mathbf{r}) d\mathbf{r} \\ = & \frac{1}{\Omega}\sum_a \Delta \int d\mathbf{r} e^{-i\mathbf{g}\mathbf{r}} Z_a \left(\sqrt{c_a/\pi}\right)^{3/2} e^{-c_a|\mathbf{r} - \mathbf{R}_a - \mathbf{R}_L|^2} \\ = & \frac{1}{\Omega}\sum_a Z_a e^{-\mathbf{g}^2/4c_a} \Delta e^{i\mathbf{g}\mathbf{R}_a} \end{aligned} \tag{1.119}$$

whereas

$$\Delta E_{XC} = \frac{1}{N_L}\sum_{ij\mathbf{k}} V_{ij}^{XC}(\mathbf{k})\Delta b_{ij}(\mathbf{k}) + \frac{1}{N}\sum_{ij\mathbf{gk}} b_{ij}(\mathbf{k})\mu_{-\mathbf{g}}^{XC}\Delta S_{ij}^{\mathbf{g}}(\mathbf{k}). \tag{1.120}$$

The other terms are

$$\frac{1}{N_L}\sum_{ij\mathbf{k}}\{T_{ij}(\mathbf{k}) + V_{ij}^{nonl}(\mathbf{k})\}\Delta b_{ij}(\mathbf{k}) + \frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k})\Delta\{T_{ij}(\mathbf{k}) + V_{ij}^{nonl}(\mathbf{k})\} + \Delta E_{ion}. \tag{1.121}$$

Gathering together terms in $\Delta b_{ij}(\mathbf{k})$ and using the KS equations and normalisation of $c_{\lambda i}^{\mathbf{k}}$ one obtains

$$\frac{1}{N_L}\sum_{ij\mathbf{k}}\{T_{ij}(\mathbf{k}) + V_{ij}^{nonl}(\mathbf{k}) + V_{ij}^{H}(\mathbf{k}) + V_{ij}^{XC}(\mathbf{k})\}\Delta b_{ij}(\mathbf{k}) = -\frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}^{e}(\mathbf{k})\Delta S_{ij}^{0}(\mathbf{k}), \tag{1.122}$$

where

$$b_{ij}^{e}(\mathbf{k}) = \sum_{\lambda} f_{\lambda\mathbf{k}}(c_{\lambda i}^{\mathbf{k}})^{*} c_{\lambda j}^{\mathbf{k}} E_{\lambda\mathbf{k}}. \tag{1.123}$$

Hence

$$\begin{aligned}\Delta E = &-\frac{1}{N_L}\sum_{ij\mathbf{k}} b_{ij}^{e}(\mathbf{k})\Delta S_{ij}^{0}(\mathbf{k}) + \Delta E_{ion} + \frac{1}{N}\sum_{ij\mathbf{k}} b_{ij}(\mathbf{k})\Delta\{T_{ij}(\mathbf{k}) + V_{ij}^{nonl}(\mathbf{k})\}+\\ &+\frac{4\pi}{N}{\sum_{ij\mathbf{gk}\text{‘}}}' b_{ij}(\mathbf{k})\frac{(n_{-\mathbf{g}} - \rho_{-\mathbf{g}})}{\mathbf{g}^2}\Delta S_{ij}^{\mathbf{g}}(\mathbf{k}) + \frac{1}{N}\sum_{\mathbf{g}}\mu_{-\mathbf{g}}^{XC} b_{ij}(\mathbf{k})\Delta S_{ij}^{\mathbf{g}}(\mathbf{k})-\\ &-4\pi\Omega{\sum_{\mathbf{g}}}'\frac{(n_{-\mathbf{g}} - \rho_{-\mathbf{g}})}{\mathbf{g}^2}(\Delta\rho)_{\mathbf{g}},\end{aligned} \tag{1.124}$$

$$\begin{aligned}\Delta T_{ij}(\mathbf{k}) =&\Delta\int B_{\mathbf{k}i}^{*}(\mathbf{r})\left(\frac{-1}{2}\right)\nabla^2 B_{\mathbf{k}j}(\mathbf{r})d\mathbf{r} &(1.125)\\ =&\sum_{L} e^{\mathbf{ikR}_L}\Delta\int \phi_i(\mathbf{r}-\mathbf{R}_i)\left(\frac{-1}{2}\right)\nabla^2\phi_j(\mathbf{r}-\mathbf{R}_j-\mathbf{R}_L)d\mathbf{r}. &(1.126)\end{aligned}$$

The ΔE_{ion} term is given by

$$\Delta E_{ion} = \frac{1}{2}{\sum_{ban}}'\Delta\frac{Z_b Z_a}{|\mathbf{R}_b - \mathbf{R}_a - \mathbf{L}_n|}\mathrm{erfc}(c_a c_b|\mathbf{R}_b - \mathbf{R}_a - \mathbf{L}_n|/(c_a + c_b)). \tag{1.127}$$

These are terms very similar to those needed in cluster AIMPRO (Briddon, 1999).

CHAPTER 2

Experimental Techniques to Investigate Defects

2.1 Deep Level Transient Spectroscopy

Lang (1974) first introduced a technique called deep level transient spectroscopy (DLTS) which is described in this section. DLTS allows the determination of the electrical properties of deep level electron or hole traps with concentrations as low as 0.1 A big advantage of DLTS is the fairly undemanding requirements toward the equipment. A disadvantage of the technique is that it reveals no structural or composition information making the identification of a trap very difficult. One can however combine DLTS experiments with other methods like EPR to get more informations about the defect. In practise this is however difficult since EPR requires a larger defect concentration such as 10^{16} cm^{-3}.

Defects with deep levels have an electron or hole wavefunction that is strongly localised around the defect. This is a result of the deep potential valley confining the electron or hole. The strong localisation in real space corresponds to a delocalisation in k-space with the result that deep states tend to be nonradiative recombination centres.

Defects are normally classified as traps, generation or recombination centres, depending on the relative capture processes. In DLTS all defects contributing to the DLTS signal are in the depletion region of a diode with only a small leakage current. The carrier concentrations in the depletion region are zero and recombination cannot occur, hence all defects in DLTS act as traps.

2.1.1 Capture and emission rates of deep states

A deep trap has a capture rate for electrons and holes, c_e, c_h and an emission rate for electrons and holes, e_e, e_h. The capture rates are given by

$$c_e = \sigma_e \langle v_e \rangle n_e \tag{2.1}$$
$$c_h = \sigma_h \langle v_h \rangle n_h \tag{2.2}$$

where σ_e and σ_h are the capture cross sections, n_e, n_h are the concentrations and $\langle v_e \rangle$, $\langle v_h \rangle$ are the average thermal velocities for electrons and holes, respectively.

The thermal emission rates can be described with Boltzmann statistics and since at thermal equilibrium the capture and emission rates of neutral material must be equal, are given by

$$e_e = \frac{\sigma_e \langle v_e \rangle N_c}{g} \exp\left(-\Delta E_e / k_B T\right) \tag{2.3}$$
$$e_h = \frac{\sigma_h \langle v_h \rangle N_v}{g} \exp\left(-\Delta E_h / k_B T\right) \tag{2.4}$$

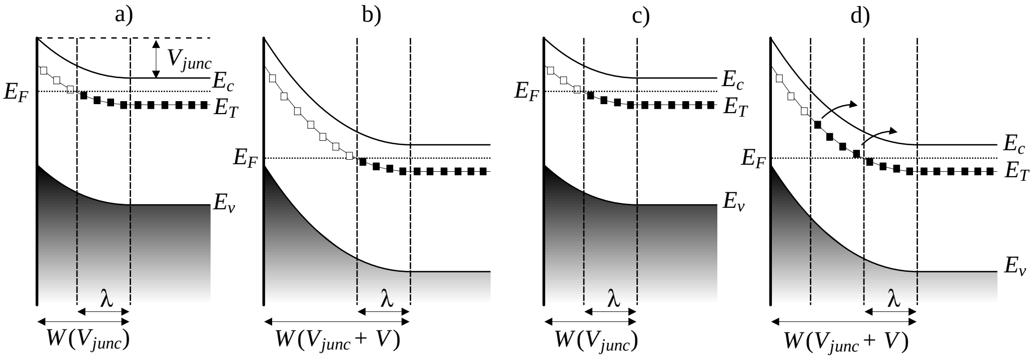


Figure 2.1: Band-structure diagram for the n-type side of a p$^+$-n junction under a majority-carrier pulse. Steps shown are a) junction without reverse-bias, b) before the pulse under reverse-bias V, c) during pulse with zero voltage, and d) after the pulse under reverse-bias.

where N_c and N_v are the effective densities of states in the conduction and valence band, respectively, and g is the degeneracy of the defect level. The parameter ΔE is the energy separation of the trap from the corresponding band edge, i.e. $\Delta E_e = E_c - E_T$ and $\Delta E_h = E_T - E_v$, where E_c, E_v and E_T are the energies of the conduction band, the valance band and the trap, respectively. Note that ΔE is the free energy that can be expressed in terms of the enthalpy ΔH and entropy ΔS by $\Delta E = \Delta H - T\Delta S$. Inserted in Equation 2.3 yields

$$e_e = \frac{\sigma_e \langle v_e \rangle N_c}{g} \exp\left(\Delta S_e / k_B\right) \exp\left(\Delta H_e / k_B T\right) \tag{2.5}$$

The slope of an Arrhenius plot therefore gives the enthalpy of the trap and not the free energy provided that $\sigma_{e,h}$ has little dependence on T. Note also that $N_{e,h} \propto T^{3/2}$ and the average velocities $\langle v_{e,h} \rangle \propto T^{1/2}$ and therefore $\langle v_{e,h} \rangle N_{e,h} \propto T^2$. In an Arrhenius plot e_e/T^2 vs. $1/T$ is plotted to find the activation enthalpy which is identical to the activation energy for $T \rightarrow 0$.

2.1.2 DLTS measurements

For a DLTS experiment a p$^+$-n or p-n$^+$ junction or a Schottky diode is used. These device structures have a carrier-depleted space charge region due to the electric dipole across the junction. An external potential difference (bias voltage) can increase or reduce this region dependent on the sign of the potential difference. The width of the space-charge region is given by

$$W = \sqrt{\frac{2\epsilon\left(V_{junc} + V\right)}{e N_D}} \tag{2.6}$$

where ϵ is the permittivity of the junction, V_{junc} and V the built-in and external potentials, e the electron charge and N_D the density of ionised donor dopants in the lightly doped side of the junction. With A being the junction area the capacitance can be written as

$$C = \frac{\epsilon\, A}{W} = \sqrt{\frac{e \epsilon N_D\, A^2}{2(V_{junc} + V)}}. \tag{2.7}$$

The last equal sign is only valid if N_D is uniform in the depletion region. A very important feature of these junctions is that an electron or hole emitted to the respective band edge is expelled out of the depletion region on a time scale of 10^{-12} to 10^{-10} s. This prevents re-trapping by impurities or other defects. Therefore the region closest to the junction is depleted of all free carriers. Beyond this is a transition region. This region starts where the deep level crosses the Fermi level as can be seen in Fig. 2.1. The width of this region is given by

$$\lambda = \sqrt{2\epsilon \frac{E_F - E_T}{e^2\, N_D}}, \tag{2.8}$$

where E_F is the Fermi level. This transition region, where the traps are occupied, may give rise to significant underestimations of the trap concentrations of up to 50 %, especially at low reverse bias voltages, but it can be neglected for the measurement of the trap levels. Usually almost all of the depletion region is essentially fully depleted, i.e. $\lambda \ll W$.

In an experiment first the depletion region is increased by applying a reverse bias as seen in Fig. 2.1 b. Then a filling pulse causes the density of filled traps $n_T(t)$ in the depletion region to increase with time according to

$$n_T(t) = N_T\left(1 - e^{-c_e t}\right) \tag{2.9}$$

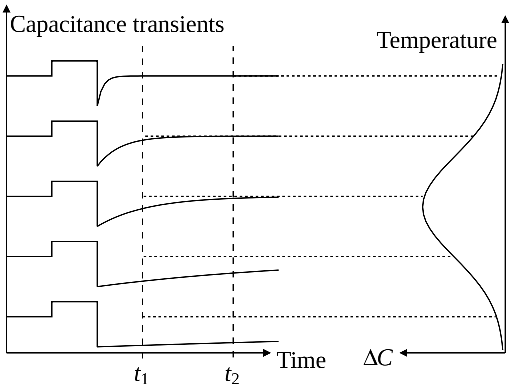


Figure 2.2: DLTS spectrum (right) for a majority carrier trap at a given rate window. On the left are capacitance transients at different temperatures.

where N_T is the density of traps in the junction. For a long enough filling pulse all traps are filled (Fig 2.1 c). After the filling pulse the depletion region is increased again to its size before the pulse and the traps empty (Fig 2.1 d) as described by

$$n_T(t) = N_T e^{-e_e t}. \tag{2.10}$$

This causes the capacity to increase and a capacitance transient

$$\frac{\Delta C}{C_0} = -\frac{N_T}{2\,N_D} \exp(-e_e\, t) \tag{2.11}$$

can be measured. The fraction in front of the exponent is only correct for $\lambda \ll W$.

The sign of the capacitance change is positive when majority carriers are emitted and negative for minority carrier emission. If the majority carrier is known it is hence easy to distinguish between emission and capture of electrons. This unfortunately does not tell whether a defect is a donor or acceptor, since both $0 \to +$ and $- \to 0$ transitions emit electrons.

The resulting output signals of an experiment are shown in Fig 2.2 on the left for measurements at different temperatures. A method called a boxcar differentiator is used for the measurement. The signal is measured at time t_1 and t_2 after the filling pulse. This measurements are repeated while the junction or diode is being slowly heated. The DLTS signal has a maximum for the time constant

$$\tau = \frac{1}{(e_e)_{max}} = (t_1 - t_2)\,\ln^{-1}\left(\frac{t_1}{t_2}\right) \tag{2.12}$$

By changing t_1 and t_2, the temperature at which the DLTS signal is maximal is changed. For determination of the activation energy only the temperature at which the maximum of the DLTS signal occurs has to be measured; the relative heights of the signal are needed only for the trap concentration measurement. Several pairs (T_{max}, $(e_e)_{max}$) are evaluated and plotted as $\ln((e_e)_{max}/T^2_{max})$ vs. $1/T_{max}$ to obtain ΔH and the capture cross sections of the defects. However, in practise $\sigma_{e,h}$ are difficult to obtain with an accuracy of more than an order of magnitude since the Arrhenius line has to be extrapolated to $T^{-1} = 0$.

2.2 Electron Paramagnetic Resonance

Electron paramagnetic resonance (EPR) is a powerful experimental tool to study point defects in semiconductors. From the magnetic resonance spectra one can obtain information about the symmetry and equivalent neighbours of a defect and hence its microscopic structure. Auxiliary techniques like uniaxial stress or optical illumination allow to perturb the defect which reveals dynamic information and also energetic features not available from the static spectra. One drawback of EPR is the fact that it can only observe paramagnetic defects. However since defects of interest are in most cases electronically active they possess more then one charge state in the gap of which at least one has an odd number of electrons and therefore is paramagnetic. Exceptions are so called negative-U defects (see chapter 6).

2.2.1 The Spin Hamiltonian

A paramagnetic defect in a magnetic field **B** has an electron spin **S**, a nuclear spin **I** and an orbital angular momentum **L**. All of these four quantities can interact which gives a contribution to the spin Hamiltonian. For many defects the orbital angular momentum is strongly quenched, therefore only spin-orbit interaction is taken into account as a perturbation of

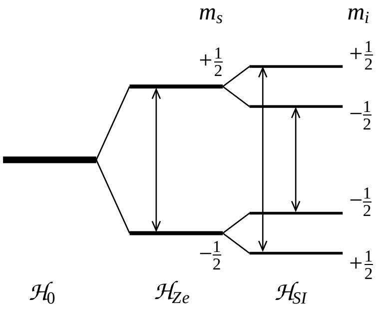


Figure 2.3: The splitting of energy levels of a $S = 1/2$ defect in a sample with $I = 1/2$ nuclei under the influence of a magnetic field. The first splitting ($\mathcal{H}_{Ze}$) is caused by the Zeeman effect and the second ($\mathcal{H}_{SI}$) by the hyperfine interaction. Allowed transitions between levels are shown by arrows.

$\mathcal{H}_{Ze}$ (see subsection 2.2.2). In a series expansion the spin Hamiltonian can then be written in the form $\mathbf{S}^i\mathbf{I}^j\mathbf{B}^k$, where i, j and k are integers. In second order, i.e. $i + j + k = 2$, the spin Hamiltonian becomes

$$\begin{aligned} \mathcal{H}_{spin} &= \mu_B\, \mathbf{B} \cdot \mathsf{g} \cdot \mathbf{S} \quad &+\mathbf{S} \cdot \mathsf{A} \cdot \mathbf{I} \quad &+\mathbf{S} \cdot \mathsf{D} \cdot \mathbf{S} \quad &+\mathbf{I} \cdot \mathsf{Q} \cdot \mathbf{I} \quad &+\mu_N\, \mathbf{B} \cdot \mathsf{g}_N \cdot \mathbf{I} \qquad (2.13) \\ &= \mathcal{H}_{Ze} &+\mathcal{H}_{SI} &+\mathcal{H}_{SS} &+\mathcal{H}_{II} &+\mathcal{H}_{Zn} \qquad (2.14) \end{aligned}$$

The following three subsections explain the first three terms. $\mathcal{H}_{II}$ and $\mathcal{H}_{Zn}$ are neglected in this overview.

2.2.2 Zeeman effect

A steady magnetic field $\mathbf{B}$ along an arbitrary direction couples with the electronic magnetic dipole moment $\boldsymbol{\mu}$ of a defect, resulting in a Hamiltonian

$$\mathcal{H}_{Ze} = -\boldsymbol{\mu}\,\mathbf{B} \tag{2.15}$$

The magnetic moment arises from both its electron spin $\mathbf{S}$ and orbital angular momentum $\mathbf{L}$ with their corresponding magnetic moments $\boldsymbol{\mu}_S$ and $\boldsymbol{\mu}_L$,

$$\boldsymbol{\mu} = \boldsymbol{\mu}_L + \boldsymbol{\mu}_S = -\mu_B\, g_L\, \mathbf{L} - \mu_B\, g_S\, \mathbf{S}, \tag{2.16}$$

where μ_B is a proportionality factor of $e\hbar/2m_e c$, known as the Bohr magneton, and $g_L = 1$ and $g_S = 2.0023$ are the so called *g-values*. The orbital angular momentum is strongly quenched for many defects and the total angular momentum and the magnetic dipole moment originate primarily from the electron spin. Spin-orbit interaction ($\mathbf{L} \cdot \mathbf{S}$) can still induce a small orbital magnetic moment. This is taken into account by expressing the total magnetic moment with an effective g-tensor instead of the previous g-values. Therefore

$$\boldsymbol{\mu} = -\mu_B\, \mathsf{g} \cdot \mathbf{S}, \tag{2.17}$$

where g is now a symmetric tensor $\mathsf{g} = g_s \cdot \mathbb{1} + \Delta\mathsf{g}$. $\Delta\mathsf{g}$ reflects the small orbital magnetic contribution induced by the spin. It can be evaluated by second-order perturbation theory, where a spin-orbit coupling term $\lambda\, \mathbf{L} \cdot \mathbf{S}$ is expanded. If $|n\rangle$ and $|0\rangle$ represent excited and ground states with energies E_n and E_0 respectively,

$$\Delta\mathsf{g}_{i\,j} = -2\lambda \sum_{n \neq 0} \frac{\langle 0\, |\mathbf{L}_i|\, n\rangle\, \langle n\, |\mathbf{L}_j|\, 0\rangle}{E_n - E_0}, \tag{2.18}$$

where the sum is over all excited states (Bourgoin and Lannoo, 1983).

The Hamiltonian can now be expressed as

$$\mathcal{H}_{Ze} = \mu_B\, \mathbf{B} \cdot \mathsf{g} \cdot \mathbf{S}. \tag{2.19}$$

The eigenvalues of this Hamiltonian can be written using the azimuthal spin quantum number m_s as follows,

$$\Delta E_{Ze} = \mu_B\, B\, g\, m_s, \tag{2.20}$$

where m_s takes on the $2S + 1$ values from $-S$ to $+S$. The effect of this additional Hamiltonian is the splitting of the degenerate defect level into $2S+1$ equally spaced levels, the so called Zeeman splitting (see Fig. 2.3). Allowed transitions between these levels can occur for $\Delta m_s = \pm 1$. In an EPR experiment the sample is placed in a microwave cavity which is inside a magnetic field. A microwave source with a frequency identical to the resonance frequency of the cavity ν_0 is used to induce transitions between the Zeeman levels. The magnetic field $\mathbf{B}$ is varied till the Zeeman splitting matches the resonance frequency at

$$B_{res} = h\,\nu_0 / g\,\mu_B. \tag{2.21}$$

This resulting in a loss of the quality factor Q of the cavity which is easily measured by the impedance of the cavity.

The value of g is angular dependent and is given by

$$g = \sqrt{g_1^2 n_1^2 + g_2^2 n_2^2 + g_3^2 n_3^2} \tag{2.22}$$

where $g_{1,2,3}$ are the principal values of g, and $n_{1,2,3}$ are the direction cosines of the corresponding principal axes of g with respect to $\mathbf{B}$. In the experiment not only the resonant magnetic field is measured but also its dependence on the field orientation. This gives characteristic angular rotation patterns for different symmetries of a defect. The Zeeman splitting therefore allows to determine the symmetry of a defect and the orientation of the principal values of g in the crystal.

2.2.3 Hyperfine interaction

A nucleus with a non-zero nuclear magnetic moment $\mathbf{I}$ can be thought of as source for a local magnetic field induced by the magnetic nucleus, $\mathbf{B}_{\text{loc}} \propto \mathsf{A} \cdot \mathbf{I}$. The spin-Hamiltonian can now be written as

$$\mathcal{H}_{SI} = \sum_k \mathbf{S} \cdot \mathsf{A}^{(k)} \cdot \mathbf{I}^{(k)}. \tag{2.23}$$

The sum goes over all atoms for which $I \neq 0$. A is called the hyperfine tensor, for a more precise description see Stoneham (1975).

The solution of the above equation in first-order perturbation theory gives changes to the defect level of

$$\Delta E_{SI} = \sum_k A^{(k)}\, m_i^{(k)}\, m_s, \tag{2.24}$$

where m_i is the nuclear azimuthal quantum number and

$$A = \sqrt{A_1^2 n_1^2 + A_2^2 n_2^2 + A_3^2 n_3^2}. \tag{2.25}$$

In analogy to g (see Eqn. 2.22) $A_{1,2,3}$ are the principal values of A and $n_{1,2,3}$ are the direction cosines of the principal axes of A with respect to the direction of $\mathsf{g} \cdot \mathbf{B}$. Experimentally A is determined in the same way as g, but instead of the main peaks the weaker hyperfine satellites are used. An important fact of the hyperfine interaction is that the hyperfine tensors of two nuclei k and l are related by $\mathsf{A}^{(k)} = \mathsf{R} \cdot \mathsf{A}^{(l)} \cdot \mathsf{R}^{-1}$, where R relates the position of both nuclei. Therefore symmetry equivalent magnetic nuclei contribute to the same EPR satellite, i.e. each peak arises from a *shell* of symmetrically equivalent paramagnetic nuclei.

The result of this additional Hamiltonian is the further splitting of each degenerate Zeeman split level into further $2I+1$ levels. This is shown in Fig. 2.3 for $I = 1/2$.

The hyperfine tensor can be decomposed into an isotropic and an anisotropic part

$$\mathsf{A} = A_0 \cdot \mathsf{1} + \mathsf{T}. \tag{2.26}$$

For only one magnetic nucleus the electron-nuclear hyperfine contribution in 2.23 can now be described as dipole-dipole interaction between two magnetic dipole moments separated by $\mathbf{r}$. This gives rise to the anisotropic term of A.

$$\mathcal{H}_{SI,aniso} = g\, \mu_B\, g_N\, \mu_N \left(\frac{3\,(\mathbf{S}\cdot\mathbf{r})\,(\mathbf{I}\cdot\mathbf{r})}{r^5} - \frac{\mathbf{S}\cdot\mathbf{I}}{r^3} \right) = \mathbf{S} \cdot \mathsf{T} \cdot \mathbf{I}. \tag{2.27}$$

The isotropic part is determined by the Fermi contact interaction, it represents the interaction of the nuclear moment $\mathbf{I}$ and the magnetic field produced at the nucleus by the electron spin $\mathbf{S}$.

$$\mathcal{H}_{SI,iso} = \frac{8}{3}\pi\, g\, \mu_B\, g_N\, \mu_N\, \delta(r)\, \mathbf{S} \cdot \mathbf{I} = A_0\, \mathbf{S} \cdot \mathbf{I}. \tag{2.28}$$

The Fermi contact interaction is non-zero for spin-densities which have an s-like component at the nucleus.

To analyse the hyperfine spectra the wavefunctions associated with gap-states are approximated as linear combinations of atomic orbitals ϕ_i (LCAO) centred on atoms i near the defect Slater (1965),

$$\psi(\mathbf{r}) = \sum_i \eta_i \left(\alpha_i \phi_{s,i}(\mathbf{r}) + \beta_i \phi_{p,i}(\mathbf{r}) \right) \tag{2.29}$$

where $\sum_i \eta_i^2 = 1$ and $\alpha_i^2 + \beta_i^2 = 1$. Here the atomic orbitals are simplified to linear combinations of s- and p-functions, $\phi_{s,i}$ and $\phi_{p,i}$, respectively.

Watkins and Corbett (1961) showed that A can then be interpreted as

$$A_0 = \frac{8}{3}\pi\, g\, \mu_B\, g_N\, \mu_N\, \eta_i^2\, \alpha_i^2\, \phi_{s,i}^2(\mathbf{0}) \tag{2.30}$$

$$\mathsf{T}_{mn} = g\, \mu_B\, g_N\, \mu_N\, \eta_i^2\, \beta_i^2 \left\langle \phi_{p,i} \left| \frac{3 r_m r_n}{r^5} - \frac{\delta_{mn}}{r^3} \right| \phi_{p,i} \right\rangle . \tag{2.31}$$

The values for $\phi_{s,i}^2(\mathbf{0})$ and the integral in 2.31 are available for numerous atomic species from self-consistent Hartree-Fock calculations.

In summary the hyperfine tensor reveals the localisation of the defect wavefunction on the surrounding atoms (η_i^2) and the corresponding s to p ratio (α_i^2/β_i^2). This analysis breaks down when the defect wavefunction can not be approximated by hybridised sp-functions, like e.g. for shallow impurities.

2.2.4 Spin-Spin interaction. Fine structure

So far we only considered defects with one single unpaired electron. However there are defect with more then one unpaired electron. This is quite often the case for vacancy defects. If $\mathbf{S}$ is larger than 1/2 one has to include the spin-spin interaction in the Hamiltonian

$$\mathcal{H}_{SS} = \mathbf{S} \cdot \mathsf{D} \cdot \mathbf{S} \tag{2.32}$$

unless the defect has cubic symmetry. Then D is 0 and no fine structure is observed. In the case of two unpaired electrons, D originates from the interaction of the spins of the electrons $\mathbf{S}_1$ and $\mathbf{S}_2$ separated by $\mathbf{r}$. In analogy to 2.27 the Hamiltonian can be written as

$$\mathcal{H}_{SS} = g^2\, \mu_B^2 \left(\frac{\mathbf{S}_1 \cdot \mathbf{S}_2}{r^3} - \frac{3\,(\mathbf{S}_1 \cdot \mathbf{r})\,(\mathbf{S}_2 \cdot \mathbf{r})}{r^5} \right) . \tag{2.33}$$

The fine structure tensor can then be written as

$$\mathsf{D}_{mn} = \frac{1}{2}\, g^2\, \mu_B^2 \left\langle \Phi \left| \frac{\delta_{mn}}{r^3} - \frac{3 r_m r_n}{r^5} \right| \Phi \right\rangle , \tag{2.34}$$

where Φ is the wavefunction of the two unpaired electrons. The addition of the fine structure Hamiltonian produces a further splitting of the EPR spectrum. This splitting can easily be distinguished from the hyperfine splitting since it goes through zero when the sample is rotated in the magnetic field whereas the hyperfine splitting does not.

2.3 Photoluminescence

Luminescence occurs when a system in an excited electronic state relaxes to a lower energy state by a radiative process. Methods to produce luminescence can be accelerated electrons, electric fields, heat or photonic excitation, for cathode-, electro-, thermo-, and photo-luminescence (PL), respectively. Out of the above mentioned methods PL is the most commonly used for the characterisation of light emitting centres. For reviews of this technique, see Davies (1981); Lightowlers (1990) and Davies (1999).

In a PL experiment photons with an energy bigger then the bandgap are used to generate electron-hole pairs. These pairs thermalise rapidly at the band edges. There are several routes via which electrons and holes can recombine. The direct route is electron-hole recombination, another is the capture of one or both of the carriers by carrier traps (defects) generating what is called a bound exciton (BE), and the formation of free-excitons (FE) (Lightowlers, 1990; Dean and Herbert, 1979).

2.3.1 Transitions

The interaction of an electromagnetic (EM) wave with an electron can be described by a Hamiltonian

$$\mathcal{H}_{pl} = \frac{\hbar\, e}{\mathrm{i}\, mc} \mathbf{A} \cdot \nabla , \tag{2.35}$$

where e and m are the electron charge and mass, respectively, and

$$\mathbf{A} = \mathbf{A}_0 \cos\left(\mathbf{k} \cdot \mathbf{r} - 2\pi\nu\, t\right) \tag{2.36}$$

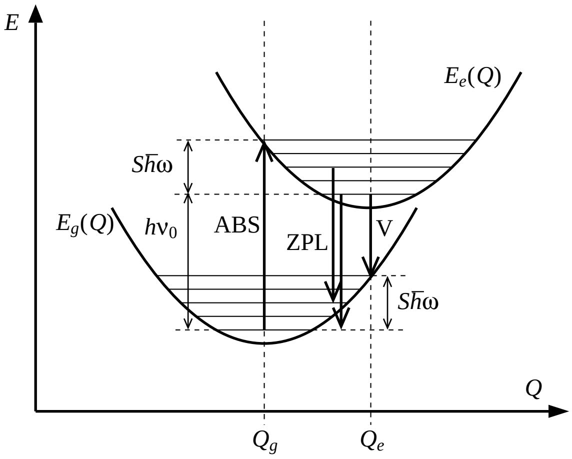


Figure 2.4: Configuration Coordinate Diagram. ABS denotes absorption from the ground state to an excited state, ZPL denotes the transitions giving rise the the zero-phonon line and V is the so called "vertical" transition. For explanations see the text.

is the vector potential of the EM wave. $\mathbf{k}$ is the wave vector, and ν the frequency. This Hamiltonian is a semiclassical approximation since the EM wave is unlike the electron not quantised. Furthermore it is only valid in first order of $\mathbf{A}$. The transition probability W between an initial state ϕ_i of energy E_i and a final state ϕ_f of energy E_f is described in time-dependent perturbation theory by Fermi's golden rule

$$W = \frac{2\pi}{\hbar}\delta\left(E_f - E_i - h\,\nu\right)\left|\left\langle \phi_f \left| \widehat{V} \right| \phi_i \right\rangle\right|^2 . \tag{2.37}$$

Using $\mathcal{H}_{pl}$ as the perturbation potential $\widehat{V}$ this becomes

$$W = \frac{2\pi}{\hbar}\delta\left(E_f - E_i - h\,\nu\right)\left|\left\langle \phi_f \left| \frac{\hbar e}{2imc}\mathbf{A}_0 \cdot e^{\mathbf{ik}\cdot\mathbf{r}}\,\boldsymbol{\nabla} \right| \phi_i \right\rangle\right|^2 . \tag{2.38}$$

For localised states the wavelength of an EM wave $2\pi/k$ is normally much larger than the extension of the wavefunction. This means that $\mathbf{k}\cdot\mathbf{r}$ is smaller than one and hence $e^{\mathbf{ik}\cdot\mathbf{r}}$ can be replaced by 1. If $\mathbf{A}$ is parallel to the z-axis Eqn. 2.38 can be written as

$$W = \frac{2\pi}{\hbar}\delta\left(E_f - E_i - h\,\nu\right)\left|\left\langle \phi_f \left| \frac{\hbar e}{2imc}A_0\frac{\partial}{\partial z} \right| \phi_i \right\rangle\right|^2 . \tag{2.39}$$

This can be further simplified since

$$[\mathcal{H}_0, z] = -\frac{\hbar^2}{2m}\frac{\partial}{\partial z} \tag{2.40}$$

to

$$W = \frac{2\pi}{\hbar}\delta\left(E_f - E_i - h\,\nu\right)\left|\frac{e\,A_0}{\hbar c}\left(E_f - E_i\right)\left\langle \phi_f \left| z \right| \phi_i \right\rangle\right|^2 . \tag{2.41}$$

From the matrix element one can see that allowed transitions are those for which the irreducible representation including z is contained in the product of the representations corresponding to the initial and final states.

Uniaxial stress can lower the defect symmetry and split the associated levels, giving information on the symmetry. Polarisation can also induce different transitions depending on whether the light is polarised parallel to the stress (so called π component) or perpendicular to it (σ component).

2.3.2 Configuration coordinate diagram

In an electronic transition not only the change of the electronic structure but also the change of the lattice has to be taken into account. This change can be described via vibrational states of the lattice. One therefore has coupled electronic-vibrational transitions, i.e. *vibronic transitions*, instead of only electronic transitions. This system is described by what is called a configuration coordinate diagram shown in Fig. 2.4. In this diagram it is assumed that the defects only interact with one lattice coordinate Q. This also means that the system has only one normal vibration mode "along Q". The parabola in the diagram are then simply the graphical representation of the total energy of a crystal with defects giving rise to PL. They are centred at the equilibrium atomic configurations for the ground (Q_g) and excited states (Q_e), respectively. The solid horizontal lines are the phonon states of each state separated by $\hbar\omega$. If the harmonic term ω of

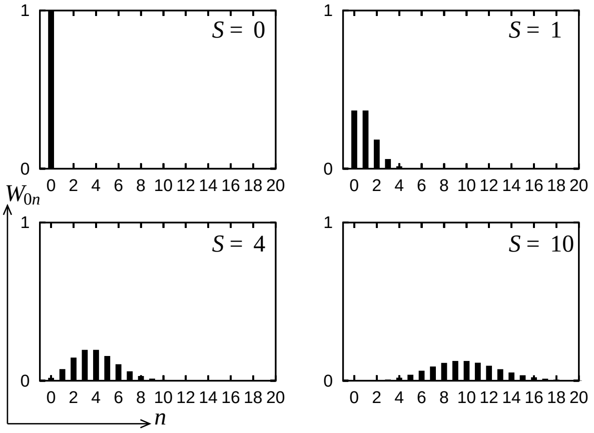


Figure 2.5: Transition probabilities W_{0n} from the no-phonon excited state to the ground state under the emission of n phonons for different coupling strength between the electronic structure and the lattice S.

the inter-atomic potential is the same in both states

$$E_g(Q) = \frac{1}{2}\mu\,\omega^2(Q - Q_g)^2 \tag{2.42}$$

$$E_e(Q) = h\,\nu_0 + \frac{1}{2}\mu\,\omega^2(Q - Q_e)^2, \tag{2.43}$$

where μ and ω are the reduced mass and frequency of an oscillator with a normal coordinate Q.

Illumination populates the excited state of the system (ABS in Fig. 2.4). At low temperatures the system rapidly thermalises towards the zero-phonon state of energy $h\,\nu_0 + \hbar\,\omega/2$. During light emission the electronic state is changed on a much smaller time scale than atomic relaxation and according to the Frank-Condon principle light emission corresponds to a vertical downward transition (see V in Fig. 2.4) followed by the relaxation of the system towards the minimum configuration. However, real emission spectra are normally composed of several peaks, corresponding to the number of phonons created during the process. The transition probabilities between the no-phonon excited state and ground states with n phonons are defined by Eqn. 2.41. This can be evaluated to get a simple relation between the transition probability and the number of phonons created,

$$W_{0n} = \frac{S^n}{n!}e^{-S}. \tag{2.44}$$

S is known as the Huang-Rhys factor. Its physical significance can be understood as follows. After a "vertical" emission, the energy given by

$$E_e(Q_e) - E_g(Q_e) = h\nu_0 - \frac{1}{2}\mu\,\omega(\Delta Q)^2 = h\nu_0 - S\hbar\omega, \tag{2.45}$$

where $\Delta Q = Q_e - Q_g$, is involved in photon and phonon creation, and a "vertical" relaxation will most probably emit a photon with energy $h\nu_0 - S\hbar\omega$ and S phonons with energy $\hbar\omega$. S is therefore the most probable number of phonons created during the luminescence process.

In Fig. 2.5 relative intensities for n-phonon transitions are shown for $S = 0, 1, 4, 10$. Note that the coupling of electronic and vibrational states does not affect the transition probability it only spreads the luminescence over a range of phonon energies. It can be seen that for weak coupling (small S) the zero-phonon-line (ZPL) has the highest transition probability. The ZPL is defined as the transition where no phonons are created, i.e. if n is the number of phonons in the ground state and m in the excited state $m - n = 0$ transitions give rise to the ZPL (see ZPL in Fig. 2.4).

2.3.3 Exciton binding energy

For a BE one charge carrier is normally strongly bound by a defect and the other is weakly bound in either the conduction or valence band. To estimate the binding energy of the weakly bound particle the solution of a hydrogenic problem can be used. The Coulomb potential in this case is

$$V(\rho) = -\frac{e^2}{4\pi\epsilon_r\epsilon_0\rho} \tag{2.46}$$

This potential gives rise to energies

$$E_x(n) = E_g - \frac{\mu\, e^4}{8h^2\epsilon_r^2\epsilon_0^2 n^2}, \tag{2.47}$$

where E_g, μ, ϵ_r and ϵ_0 are the band-gap, the reduced mass of the electron-hole pair, the dielectric constant and the relative permittivity, respectively. n is the principal quantum number of the system. For the estimate of the binding energy only the ground state $n = 1$ is of interest. The binding energy is therefore

$$E_b = \frac{\mu\, e^4}{8h^2\epsilon_r^2\epsilon_0^2}. \tag{2.48}$$

In $4H$-SiC the values for the electron effective mass tensor are $m_{ML} = 0.33m_0$, $m_{M\Gamma} = 0.58m_0$, and $m_{MK} = 0.31m_0$ (Volm *et al.*, 1996), where m_0 is the free electron mass. The effective mass can be estimated by $\mu = (m_{ML}m_{M\Gamma}m_{MK})^{1/3} = 0.39m_0$ if the three masses are equally important. With $\epsilon_r = 9.7$ for $4H$-SiC the exciton binding energy for a weakly bound electron is 56 meV.

Part I

Defects in Silicon

CHAPTER 3

The Self-Interstitial in Silicon

3.1 Introduction

Advanced semiconductor devices are sub-micron devices requiring high spatial control of dopants. These dopants are however subject to diffusion and the diffusion of dopants in semiconductors is strongly influenced by the thermal behaviour of self-interstitial clusters, nucleated as a result of nonequilibrium processing steps such as ion implantation. Complete understanding and detailed modelling of the removal of ion-implantation damage in Si and the accompanying TED caused by the interstitials are hampered by a lack of insight into the clustering phase in the defect evolution.

Ion implantation is a key processing step in the current manufacture of silicon based electronic devices. However, this method not only introduces the desired dopants, but also creates lattice damage in the form of vacancies and self-interstitials. More than 90% of these annihilate during a subsequent anneal but the remainder, arising from self-interstitials spatially separated from vacancies, do not annihilate but instead form stable aggregates. These may be small clusters containing just a few vacancies or self-interstitials, such as those proposed for the B^4_{80}/J- and W-lines (Kaminskii *et al.*, 1987; Hourahine *et al.*, 2000; Coomer *et al.*, 1999) photoluminescence centres or the B3 electron paramagnetic defect (Coomer *et al.*, 2001), and as such anneal around 500°C, but they also occur as extended defects.

The small interstitial clusters (I_1 to $\sim I_8$) are the missing link in understanding the nucleation of extended defects like the rod-like defects (RLDs) discussed in chapter 4 and dislocations.

A great many calculations using a variety of methods have already been carried out to try to determine the properties and the evolution of the clusters. But no calculations at the present time have explained all the “magic numbers” found by Cowern *et al.* (1999) for the stable interstitial clusters. Furthermore unlike in diamond (Hunt *et al.*, 2000; Goss *et al.*, 2001) the single interstitial has not yet been identified. Of the small interstitial clusters only I_3 and I_4 are believed to be identified. The I_2 interstitial either has a high-energy configuration trapped in a barrier, or has no stable configuration and diffuses to build bigger defects or recombines at the surface.

This chapter models the isolated self-interstitial and reviews an assignment of photo-EPR and DLTS signals with I_1 (Mukashev *et al.*, 1998) by comparing calculated with measured observables. Before an overview reviews the tri- and tetrainterstitial clusters I_3 and I_4 followed by the di-interstitial I_2 and finally the self-interstitial I_1.

3.2 Experimental Results and Atomic Models of Small Interstitial Clusters

3.2.1 The I_3 and I_4 defects

The W-optical centre

The W-optical centre is detected in Si which has been exposed to lattice damaging treatments and a heat treatment. It is seen in both *n*- and *p*-type silicon following Si implantation (Schultz *et al.*, 1992). The centre is characterised by

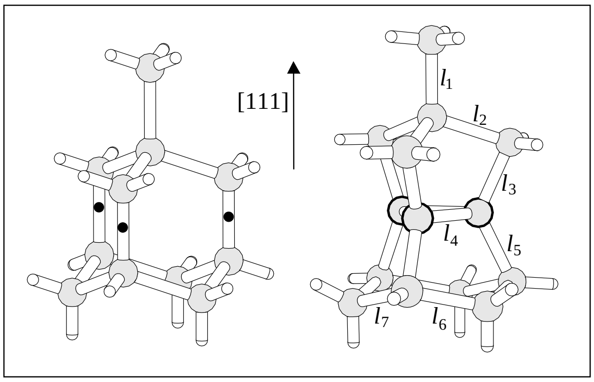


Figure 3.1: The high energy I_3 structure which is seen in the B3 centre

a sharp peak at 1.0182 eV and an associated phonon sideband due to a local vibrational mode (LVM). Uniaxial stress measurements reveal the symmetry of the defect to be trigonal.

It is generally accepted that the W-centre arises from the agglomeration of an intrinsic species, and although earlier work favoured a vacancy complex (Kirkpatrick *et al.*, 1976; Minaev *et al.*, 1981; Estreicher *et al.*, 1997), more recent experiments and theory (Schultz *et al.*, 1992; Davies *et al.*, 1987; Nakamura *et al.*, 1998; Coomer *et al.*, 1999) have supported the interstitial nature of the W-centre.

The LVM has an energy of 70.0 meV and lies just above the Raman frequency consistent with a strengthened Si–Si bond (Davies *et al.*, 1987). Calculations have shown that vacancy defects possess weak reconstructed Si–Si bonds and such defects could not give rise to local modes lying above the Raman frequency. Moreover, the stress response of the W-optical line is very weak and out of keeping with an assignment to a vacancy cluster: for example, the shift of the W-luminescence line with stress is very much smaller than that of V_6 (Kaminskii and Lavrov, 1998; Hourahine *et al.*, 2000). This suggests that the defect is interstitial related and its trigonal symmetry indicates I_3 as a likely candidate.

A trigonal I_3 centre is obtained by placing three interstitial atoms at the centres of adjacent $\langle 111 \rangle$ bonds (Fig. 3.2). The bonds of the interstitials can all be reconstructed (Coomer *et al.*, 1999). Although the defect is fully coordinated, a gap level is induced by the compressive stress. This suggests that the defect can bind excitons with a localised hole leading to optical lines around 1 eV. The I_3 defect possesses a number of LVMs. A symmetric vibrational mode lies at 74 meV, consistent with experiment (70.0 meV) (Davies *et al.*, 1987).

Recent work (Estreicher *et al.*, 2001*b*) has found that another form of I_3 has lower formation energy but is much more mobile. The suggestion is then that both forms arise in irradiated material but the immobile form of I_3 described here persists to 300°C while the mobile one is rapidly lost. Estreicher *et al.* (2001*a*) find that the three self-interstitials of the immobile form easily exchange sites, thus continuously turning around the trigonal axis.

The B3 EPR centre

The B3 EPR centre has been linked to I_4 (Coomer *et al.*, 2001). B3 is formed in neutron irradiated *p*-Si heated between $\sim$ 200–500°C. It is one of only eight defect centres observed in irradiated silicon which have been reported to possess D_{2d} symmetry and its stability to high temperatures suggests a simple secondary irradiation product of particularly low formation energy. The lack of low temperature stress response indicates that the D_{2d} symmetry does not result from

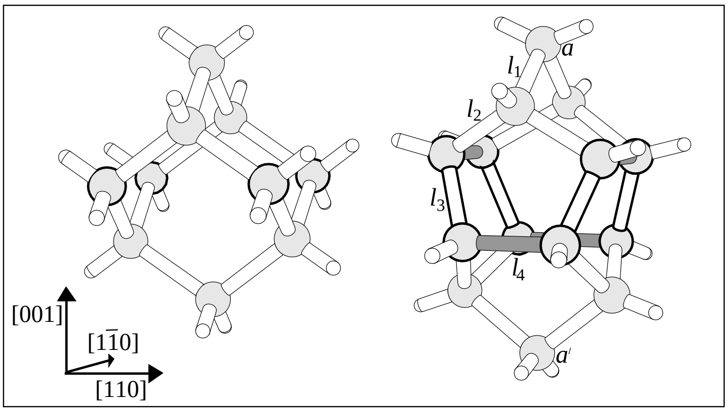


Figure 3.2: The I_4 structure

a Jahn-Teller distortion (Brower, 1976). Analysis of the hyperfine structure reveals further information: (i) The defect centre is probably vacant. (ii) Two equivalent Si atoms lie along the principal ⟨001⟩ axis. (iii) Only 9% of the unpaired electron is localised on each of these two atoms. (iv) The electronic wavefunction on these atoms is predominantly *p*-like with only 6% *s* character. Recent work (Mchedlidze and Suezawa, 2003) presented a more detailed analysis and found in addition to the 9% of the unpaired electron on two equivalent Si atoms, 2% on 8 equivalent and 0.5% on another 12 equivalent atoms. The thermal stability of the centre – discussed below – excludes an identification with the low energy form of I_2.

The structure of I_4 shown in Fig. (3.2) (Arai *et al.*, 1997; Coomer *et al.*, 2001) reveals almost perfect coordination. However, strain can induce a donor level to move into the band gap above E_v. The position of the donor level of I_4 was found by comparing the ionisation energy of I_4 with that of the carbon-interstitial which has an experimentally determined donor level at $E_v + 0.28$ eV (Song and Watkins, 1990). The I_4 ionisation energy was found to be $0.01 - 0.12$ eV higher than that of the carbon-interstitial placing the donor level of I_4 between $E_v + 0.16$ and $E_v + 0.27$ eV. This result is consistent with the tentative correlation of the $E_v + 0.29$ eV hole trap with the B3 centre (Mukashev *et al.*, 1982). No other electronic levels are calculated to lie within the band gap.

The unpaired wavefunction in I_4^+ is spread over a number of atoms, whilst the largest amplitude lie near atom sites labelled *a* in Fig. (3.2). Mulliken analysis shows that 6% of the unpaired wavefunction is localised on each atom along [001] in excellent agreement with the value obtained from ^{29}Si hyperfine measurements (9%). It is also found that the unpaired electronic wavefunction is strongly *p*-like in character consistent with the 94% anisotropic component observed. The good agreement with the spin density is a strong evidence that the model is correct and this is further supported by the agreement between the calculated and observed stress-tensor (Coomer *et al.*, 2001). Recently the assignment of B3 to I_4 was confirmed by EPR studies finding two equivalent Si sites in the first shell, a next interaction with four equivalent Si sites in the second shell, and a last interaction with eight equivalent Si sites in the third shell (Pierreux and Stesmans, 2003).

The formation energy of the I_4 defect, $E_f(I_4)$, relative to four bulk atoms was calculated using the supercell method to be 8.7 eV. Experimental determination of this energy gives a value of around $E_f(I_1) + 3.4$ eV (Cowern *et al.*, 1999). Later it will be shown that the calculated $E_f(I_1)$, the formation energy of the isolated ⟨110⟩ orientated interstitial, is around 3.4 eV. This brings the calculated formation energy of I_4 to around $E_f(I_1) + 5$ eV, in fair agreement with the experiment given the errors involved in both the theoretical and experimental determination of these values. The binding energy of I_4 relative to four separated [110] orientated interstitial atoms is calculated to be 6.9 eV.

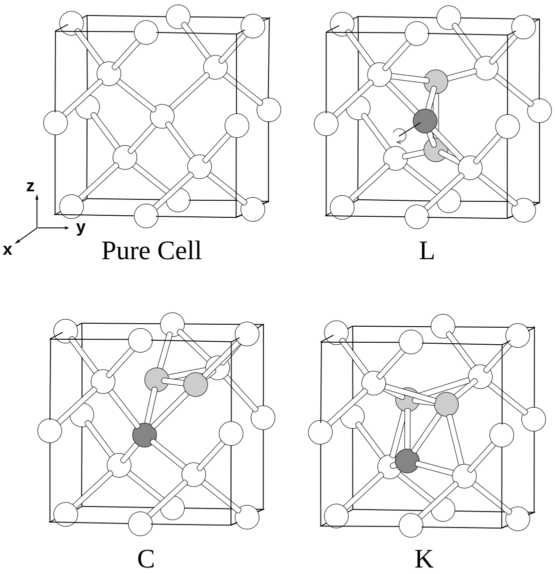
The high thermal stability of B3 is explained by the low formation energy calculated for I_4. This result is expected from the I_4 model because of the near ideal bonding arrangement of all atoms. The model also explains the finding that I_8 also possesses remarkably low formation energy (Cowern *et al.*, 1999). Clearly, eight [001] split-interstitial pairs can be sited on a (001) plane resulting in the formation of two neighbouring I_4 defects. It is likely that this $(I_4)_2$ defect will possess a lower formation energy per interstitial than two separated I_4 units due to the mutual strain relief interaction between the units. The large increases observed in the formation energies of the defects I_5, I_6, I_7 relative to I_4 and I_8 is also consequent from this model as these intermediate defects may be unable to achieve full coordination.

Effect of annealing

The effect of annealing irradiated or implanted Si and following the evolution of photoluminescent (PL) centres suggests a simple evolutionary path for the interstitials. Terashima *et al.* (1997) implanted CZ-Si with P and B ions and subsequently annealed the samples between 200 and 700°C. Around 200°C, the W-line increases in intensity (relative to the B-related bound exciton) until about 300°C, followed by a decrease and disappearance of the signal at 500°C. As the W signal diminished, the intensity of another line called ‘I1’ or X at 1.0398 eV increased. This was not present at 200°C but became apparent at 300°C, and its intensity increased until about 450°C. It subsequently decreased and vanished by 600°C. These two lines dominated the spectrum but there were less intense bands due to ‘I2’ at 1.080 eV– a trigonal defect containing two B atoms (Sauer and Weber, 1983) – and a line at 1.097 eV. The centres I1 and I2 should not be confused with the single and di-interstitial defects labelled I_1 and I_2 respectively.

Similar results were reported by Awadelkarim *et al.* (1990). Here, both CZ-Si and FZ-Si were implanted with C and O and annealed to 500°C. Besides the well known carbon and oxygen related lines: P (0.7671 eV), H (0.9254 eV), L (1.0038 eV), G (0.9694 eV), and C (0.7890 eV), the X and W lines were also detected. Moreover, in FZ-Si these, together with the Y-line at 1.08 eV, were the principal emissions. The W-line is detected up to about 350°C and vanished at 400°C, while the X-line is detected between about 250 and 500°C. The W line requires a donor level at $E_v + 0.1$ eV not at $E_v + 0.29$ eV. Therefore B3 can not be correlated with W.

In summary, the annealing data in irradiated FZ-Si indicate that the W-centre is formed at an earlier stage than X and is stable until about 350°C but as this disappears, the X-centre evolves and is stable until about 500°C. Thus it is inferred that the W and X lines are related to intrinsic defects and the X centre is a larger aggregate than W. Both optical signals are seen in absorption as well as luminescence possibly indicating transitions between gap states. The assignments of the W and X lines to I_3 and I_4 are entirely consistent with the annealing pattern.



Pure Cell L

C K

Figure 3.3: The structures of I_2 proposed by Lee (L), Coomer (C) and Kim (K)

3.2.2 Di-interstitial I_2

The di-interstitial has not yet been identified, but at least one electron paramagnetic resonance (EPR) spectrum has long been attributed to the di-interstitial. This centre, labelled P6, was studied in great detail by Lee *et al.* (1976); Lee (1998). The P6 EPR centre is formed in neutron, ion or MeV proton irradiated intrinsic Si. Its g-tensor reveals a C_2 axis along [100] at 200 K, but at 300 K, the symmetry becomes D_2 or D_{2d}. The activation energy for this change is only 0.2 eV and represents a thermally activated reorientation between equivalent sites.

P6 anneals at 170°C and has been correlated with a DLTS donor level at $E_v + 0.4$ eV. ^{29}Si hyperfine interactions have been resolved for two shells. The first contains a pair of equivalent Si atoms with an isotropic hyperfine tensor. The second contains a single Si atom with an anisotropic hyperfine tensor. 23% of the spin-density resides on the unique atom largely in a p-orbital and only 3% on the two equivalent Si atoms. Uniaxial stress studies showed that the defect could reorientate at 100°C with an activation energy of 0.6 eV and with a low pre-exponential factor of 10^8 s^{-1}. This activation energy is similar to the anneal of the 0.4 eV DLTS level found in irradiated p-Si. The energy-stress tensor of P6 is consistent with a defect of D_{2d} symmetry with principal values and directions $B_1 = B_2 = 15.4$ eV along [100] and [010], and $B_3 = -30.8$ eV along [001]. Here the C_2 axis in the low temperature form points along [010] through the unique Si atom.

Lee *et al.* (1976) suggest that P6 is due to I_2^+ with three atoms sharing a lattice site resulting in a C_2 axis along [010]. Recent *ab initio* calculations (Kim *et al.*, 1999) indicate that this structure is meta-stable and I_2 possesses C_{1h} symmetry at low temperatures (Fig. 3.3 K), and C_{2v} when some distortion is allowed. In spite of this difference in symmetry, the defect was identified with P6. This cannot be correct and the true nature of P6 is at present unresolved. Lee (1998) proposed a structure about 1 eV higher in energy but with C_2 symmetry (Fig. 3.3 L).

Calculations from the AIMPRO group (Coomer, 2000; Eberlein *et al.*, 2001) find a structure degenerate in energy with Kim's structure and with the same symmetry (Fig. 3.3 C). They also reveal that the L and C structure (see Fig. 3.3) have a low diffusion barrier of around 0.5 eV and therefore can not be seen in EPR experiments at room temperature. Furthermore their symmetry does not match the one for the P6 centre. The L structure has the right symmetry but the calculated B-tensor does not agree with the measured one. The previous assignments of structures to the P6 centre must be wrong and the correct structure remains unknown.

3.2.3 Single interstitial I_1

Experiments at low temperatures (∼ 20 K) using electronic irradiation of doped Si show that self-interstitials possess a high mobility even at these low temperatures (Watkins, 1964) and are therefore hard to observe. Only one group (Mukashev *et al.*, 1998, 1999, 2000) claimed to have observed the isolated interstitial. They used a photo-EPR experiment and a DLTS experiment and assigned the two signals to I_1 at the T site in the +2 charge state. The DLTS experiment showed an electrical level at $E_c - 0.39$ eV as well in pulled (Cz) samples as in floating zone (FZ) silicon samples irrespective of acceptor impurities. This indicates that it arises from an intrinsic defect. This defect labelled E1 is above mid-gap and therefore only observed under minority carrier injection, it can not be seen in n-Si. Its cross section for trapping electrons is many times that for holes, i.e. the centre is positively charged. Below 77 K the defect anneals out under minority carrier injection and at about 350 K without. The rate of annealing depends on the minority carrier injection level and

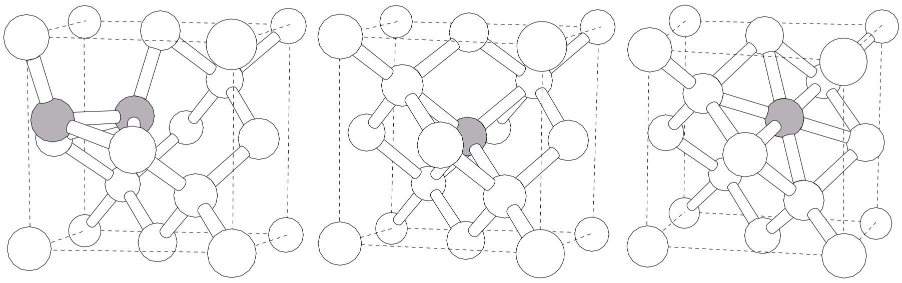


Figure 3.4: The ⟨110⟩ split, tetrahedral and hexagonal configuration of the I_1

on the concentration of carbon and oxygen impurities. The EPR spectrum labelled AA12 has the same temperature range of stability, injection annealing behaviour and electron level position, which suggests that it arises from the same defect. In addition to this the spectrum anneals out at around 160 K due to a change of charge state (Mukashev *et al.*, 1998). An unresolved hyperfine structure (hfs) due to ^{29}Si at 4 to 6 equivalent sites indicates that the defect is either at a tetrahedral (T) or at a hexagonal (H) site. The frequency of around 9.5 MHz corresponds to a localisation of the paramagnetic wavefunction of around 1.5% (Bekman *et al.*, 1989). The growth of other interstitial defects such as C_i, Al_i, and Si_i-O_i is detected after the anneal. The implication is that the (+/++) level of I_1 lies at $E_c - 0.39$ eV and the defect is unstable when neutral.

On the other hand, low energy irradiation does not lead to the production of AA12 even at 4.2 K. This may be because of the greater degree of ionisation consequent upon *e*-irradiation thus causing athermal recombination enhanced motion (Watkins, 1964). It appears that rapid diffusion of interstitials occurs after 1.3 MeV proton implantation of *n*-Si, where I^{++} could not form (Hallen *et al.*, 1999). Even in *p*-type Si an extremely effective recombination-enhanced migration is expected (Watkins, 2001). This questions whether the isolated interstitial is observable in Si.

Before the measurement of AA12 there had been no EPR evidence of an isolated silicon atom in p-type silicon after in situ electron irradiation at cryogenic temperatures. Only defects which are identified as interstitials trapped at impurities were observed in $\sim 1:1$ concentrations to that of the isolated vacancy. Although defect production is observed to be much lower at cryogenic temperatures in n-type silicon, migration may only occur after annealing to 140–175 K, where some of the trapped interstitial configurations are observed to emerge (Watkins, 2001).

The different candidates for the single interstitial are the ⟨110⟩ and ⟨100⟩ split sites, the bond centred, hexagonal (H) and tetrahedral (T) site. Two of these configurations, the bond centred and the ⟨100⟩ split interstitial, are about 1 eV higher in energy than the remaining three (Lee *et al.*, 1998) and are no longer considered as low energy structures in recent literature. The ⟨110⟩ split interstitial has a *dumbbell* orientated along the ⟨110⟩ direction that substitutes a lattice atom. This structure has C_{2v} symmetry (Fig. 3.4 left).

The first LDA calculation proposing that the ⟨110⟩ split interstitial is the most stable defect in the neutral charge state was by Baryam and Joannopoulos (1984). Later calculations by Blöchl *et al.* (1993) and Needs (1999) using the LDA scheme also found this configuration to be the most stable with a formation energy of 3.3 eV. In contrast to this Clark and Ackland (1997) obtained 2.16 eV and Lee *et al.* (1998) 4.19 eV in their LDA calculations. Chadi (1992) also found that the ⟨110⟩ split is the lowest energy configuration, but not only in the neutral, also in the charge states from -1 to +2. In contrast to this Lee *et al.* (1998) and Baryam and Joannopoulos (1984) clearly find the tetrahedral site to be the most stable configuration for the positively charged defects.

3.3 Calculation Method

A number of possible locations for the interstitial were examined in supercells with 64 and 216 host atoms. The 64 atom supercell was made up from 2 x 2 x 2 conventional unit cells while the 216 atom supercell was made up from 3 x 3 x 3 conventional unit cells. The examined locations for the interstitial were the T site with T_d symmetry, the H site with D_{3d} symmetry and the ⟨110⟩-split interstitial with C_{2v} symmetry. The bond-centred interstitial was examined in the 64 atom cell in the positive charge states +1 and +2 but found to be unstable. It relaxed towards a T site. All calculations were done with a MP-2^3 sampling of the Brillouin zone. The basis was increased from pdpp to dddd (see Table 3.1) in the 64 atom cell while in the 216 atom cell a ddpp basis was used for the calculations. Due to the fact that the interstitial at the T site has a t_2 level, the calculations in the 216 atom cell were repeated with Γ-point sampling only, to avoid using the

Table 3.1: Number of different basis functions ϕ_j (see Eqn. 1.60) on each silicon atom for different basis sizes.

Basis	dddd	ddpp	pdpp
Number of localised orbitals	40	28	22

Table 3.2: Formation energies (eV) of the ⟨110⟩ interstitial in the neutral charge state. The values under "Needs" and "Lee" are the results of Needs (1999) and Lee *et al.* (1998), respectively.

basis	dddd	ddpp	pdpp	TM-pdpp	"Needs"	"Lee"	ddpp
	65 atoms						217 atoms
E_f	3.34	3.41	3.59	3.72	3.31	4.19	3.37

split triplet level at the MP-2^3 k-points. A Fermi smearing of 0.02 eV was used to fill the one-electron levels to support numerical stability of the code. All atoms where allowed to relax.

The formation energy (E_f) of a defect is defined as:

$$E^f(\mathrm{X}) = E(\mathrm{X}) - n\,\mu + q\,E_F, \tag{3.1}$$

where $E(\mathrm{X})$ is the energy of the cell containing the defect, n is the number of atoms in the supercell, μ is the chemical potential of Si taken from the energy per atom of bulk material, q is the net charge of the supercell (without the background charge) and E_F the Fermi energy.

The electrical level of a defect $E(q/q+1)$ is calculated from the difference in total energies between the two charge states q and $q+1$ and is $\Delta E_\mathrm{D}(q/q+1) = E^f_\mathrm{D}(q) - E^f_\mathrm{D}(q+1)$. This value is then compared to the analogous quantity found for a *reference* defect with a known electrical level from experiment ($\mathrm{E_m}$), i.e.

$$\begin{aligned} E(q/q+1) - \mathrm{E_m} =& \Delta E_\mathrm{D}(q/q+1) - \Delta E_\mathrm{M}(q/q+1) \\ =& \left(E^f_\mathrm{D}(q) - E^f_\mathrm{M}(q)\right) - \left(E^f_\mathrm{D}(q+1) - E^f_\mathrm{M}(q+1)\right) \end{aligned} \tag{3.2}$$

The idea behind this is that only energies of supercells with the same net charge are compared and errors due to charged supercells should at least partially cancel out. This method requires that the defect and reference are treated in the same sized cell with the same basis and furthermore they should have similar levels (Resende *et al.*, 1999; Jeong *et al.*, 2002). This method when applied to the donor levels of chalcogen centres in Si gave accurately *relative* donor levels for S, Se and Te by comparing their ionisation energies (Coutinho *et al.*, 2003). This method is also referred to as marker method and the reference defect is called marker.

3.4 Results

The formation energy for the interstitial at the ⟨110⟩ site in the neutral charge state is found to be 3.41 eV for the 65 atom cell with the ddpp basis (Table 3.2), in good agreement with previous calculations of 3.31 eV (Needs, 1999) and 3.30 eV (Hakala *et al.*, 2000). Increasing the size of the supercell with the same basis to 217 atoms decreases the formation energy by only 0.04 eV. If a 4^3 MP sampling scheme is used, this energy increased by 0.01 eV for the 65 atom supercell. The calculations are therefore converged with respect to the size of the supercell and the k-point sampling. The convergence with basis size can easily be seen from Table 3.3 showing relative total energies for the interstitial in charge states from +2 to -1 for different basis sizes. In the neutral charge state the ⟨110⟩-split interstitial is the structure with the lowest formation energy but the H interstitial is only slightly higher while the T site is ∼0.2 eV higher in energy. If the H interstitial is not constrained to have D_{2d} symmetry the interstitial moves along ⟨111⟩ by ∼0.5 Å out of the hexagon. This new configuration is ∼0.05 eV lower than the H configuration but ∼0.07 eV higher than the ⟨110⟩ split interstitial in a calculation with ddpp basis. In the following this configuration will be referred to as H'. With the biggest basis, H' is again ∼0.05 eV lower then H and now energetically degenerate with the ⟨110⟩ configuration. Due to the tiny energy difference of the ⟨110⟩ and H' interstitial in the neutral charge state a high mobility can be expected.

To analyse the artificial effect of the splitting of the t_2 manifold due to the MP-2^3 sampling the calculations of charge states with an occupied t_2 manifold have been repeated with Γ-point sampling in the 217 atom supercell where qualitatively the same results were found (see Table 3.4).

In the negative charge state the ⟨110⟩ oriented interstitial remains the structure with the lowest formation energy. The acceptor level calculated with the bulk as marker was however found to be only 0.09 eV below $\mathrm{E_c}$ and could within the error of the method also be resonant with the conduction band. To investigate the acceptor level the C interstitial in the 216 atom cell was used as a marker. The error of the method is about 0.2 eV. This defect has a well known acceptor level at $\mathrm{E_c} - 0.10$ eV. The marker method changed the acceptor level only slightly up to 0.03 eV below $\mathrm{E_c}$. The ⟨110⟩ interstitial has therefore either a very shallow or no acceptor level.

In the positive charge states +1 and +2 the T interstitial has the lowest formation energy. In the +2 state, the T site is stable against small perturbations, i.e. when moved slightly from its position the atom returns during the relaxation. The T interstitial possesses a t_2-level in the upper part of the gap which is occupied by one electron in the single positive

Table 3.3: Relative total energies (eV) for I_1 in a 65 atom supercell with different basis sets. Apart from the values under “Lee” the configurations are all high symmetry forms.

charge state	T	⟨110⟩-split	H
		dddd basis	
-1	0.15	0.00	0.16
0	0.22	0.00	0.06
+1	0.00	0.62	0.48
+2	0.00	1.40	1.09
		ddpp basis	
-1	0.29	0.00	0.24
0	0.21	0.00	0.12
+1	0.00	0.56	0.47
+2	0.00	1.41	1.14
		pdpp basis	
-1	0.34	0.00	0.25
0	0.25	0.00	0.12
+1	0.00	0.56	0.46
+2	0.00	1.44	1.17
		TM-pdpp basis	
-1	0.41	0.00	0.28
0	0.31	0.00	0.13
+1	0.00	0.53	0.44
+2	0.00	1.44	1.18
		“Lee”	
-1	0.38	0.00	0.19
0	0.16	0.00	0.04
+1	0.00	0.59	0.47
+2	0.00	1.32	1.15

charge state. In this case a Jahn-Teller distortion splitting the t_2-manifold would be expected to occur. This is the origin of a C_{3v} distortion. However, it seems that the distortion is small and only lowers the energy by 0.07 eV compared to the T_d structure. To obtain the Jahn-Teller distortion the calculations have to be performed with a Γ-point sampling, since at other k-points the t_2-manifold is already split for the T_d configuration. The distortion lowers the symmetry from T_d with 4 bonds of 2.45 Å to C_{3v} with 3 bonds of 2.39 Å and one of 2.58 Å. The t_2-manifold would be empty in the +2 charge state and a distortion would no longer be expected.

In the positive charge state, a Mulliken analysis shows that only 7% of the spin density is localised on the interstitial in the three p-orbitals (Fig. 3.5). The isotropic hyperfine interaction with the nucleus of the unique Si atom is then almost zero consistent with the 1.5% low value seen experimentally for AA12.

To find its diffusion barriers the interstitial was displaced in small steps along a vector, defined by the initial and final positions of the diffusion step. The interstitial was then allowed to relax only in directions perpendicular to the vector while the other atoms were allowed to relax without constraints. This method is called ’ridge-walking’. The calculations were performed in the 65 atom cell with the pdpp basis. In the double positive charge state the H-site was found to be a saddle point on the diffusion path of I_1^{++}. The barrier in this charge state is then about 1.1 eV. Note that the calculated barriers are always upper limits for real barriers since it can not be excluded that the path with the lowest barrier was missed in the calculation. Using the Arrhenius diffusion equation $D = D_0 \exp(-Q/k_B T)$ with $D_0 = \nu s^2$ where s is the distance between T-sites $(a_0/\sqrt{2})$ and $\nu = 10^{13}$ s^{-1}, T^{++} should anneal around 430 K. The same path was found for the

Table 3.4: Relative total energies (eV) for I_1 in a 217 atom cell with ddpp basis and MP-2^3 and Γ-point sampling of the Brillouin zone.

charge state	T	⟨110⟩-split	H
		MP-2^3	
0	0.18	0.00	0.08
+1	0.00	0.51	0.51
+2	0.00		1.2
		Γ-point	
0	0.28	0.00	0.04
+1	0.00	0.25	0.30

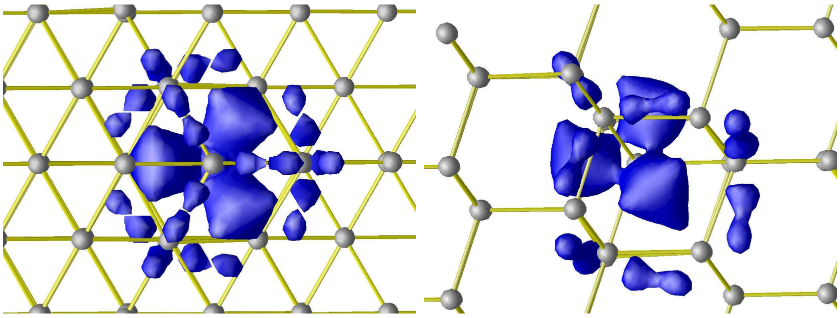


Figure 3.5: The wavefunction of the t_2 level of the T interstitial. The left figure shows a $\langle 111\rangle$ view and the right a view along $\langle 110\rangle$ at an isosurface of the anti-bonding sp^3 orbital of the triplet level.

single positive interstitial. The barrier here is ~0.5 eV corresponding to an annealing temperature of around 200 K.

Turning to the neutral charge state. When an extra electron is added to the Jahn-Teller distorted T$^+$ interstitial it relaxes to a H' site. The reason for the relaxation to H' and not $\langle 110\rangle$ is a barrier between $\langle 110\rangle$ and H. This barrier was found to be around 0.35 eV (see Fig. 3.6) and higher than the 0.15 eV found by Needs (1999) and Lee *et al.* (1998). To eliminate the possibility that this high value comes due to the relatively small basis the runs were repeated with the biggest basis set used and no significant differences were found. In addition to this another diffusion method was used. It is called 'Dimer Hybrid' method (Henkelman and Jonsson, 1999) and is described in Papagiannidis (2003). This method found a saddle point at around 0.33 eV. Although the structure of the saddle-point is very similar to the one found by Needs (1999) the energy is significantly higher. The neutral self-interstitial is therefore found to be less mobile as thought before and should only anneal at ~130 K. This is in aggreement with the observation of interstitial movement between 140–175 K (Watkins, 2001). Under ionising conditions, however, the interstitial can change from a T to a H site and back to a T site in a neighbouring cage by trapping holes and electrons in exchange, i.e.

$$\begin{aligned} &T^+ + e \rightarrow H^0 \\ &H^0 + h \rightarrow T^+. \end{aligned} \tag{3.3}$$

In this way the interstitial can diffuse almost without any barrier under electron irradiation.

To calculate the donor levels S and S$_2$ were used as markers. They possess second donor levels at E$_c - 0.6$ and E$_c - 0.39$ eV, respectively (Carlson *et al.*, 1959) – the latter quite close to the AA12 centre. Table 3.5 gives the calculated ionisation energies and hence the separation of the donor levels to the conduction band for the T structure. The (0/+) level of I$_1$ *at the T site* is found to be slightly below the conduction band ($\sim$ E$_c - 0.1$ eV) if the Jahn-Teller distortion for T$^+$ and the change from T^0 to H^0 is ignored. Including these structural changes gives the single donor level of I$_1$ at $\sim$ E$_c - 0.2$ eV. The second donor level of I$_1$ lies at $\sim$ E$_c - 0.25$ eV when the Jahn-Teller lowering of T$^+$ is ignored and $\sim$ E$_c - 0.35$ eV when it is included. This value is close to the level attributed to AA12 and updates the published one (Eberlein *et al.*, 2001) where a less accurate method was used. A plot of the formation energies including energies of charged defects derived from the calculated level positions is shown in Fig. 3.7.

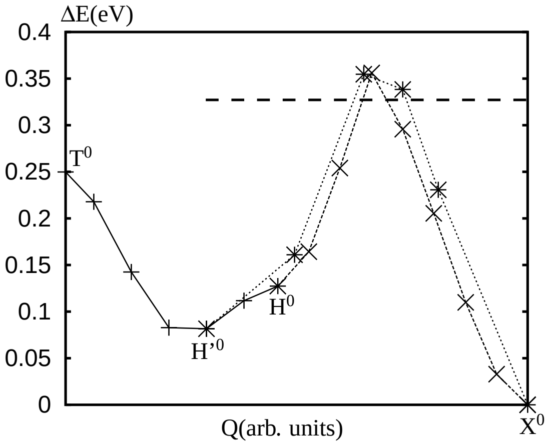


Figure 3.6: Diffusion barrier for the neutral self-interstitial. The horizontal broken line is the diffusion barrier between the $\langle 110\rangle$ site (denoted by X) and the H site found by the 'Dimer Hybrid' method. The other values were calculated using the 'ridge-walking' method.

Table 3.5: Electrical levels (eV) of S, S_2 and T (with T_d symmetry) below the conduction band in 65 and 217 atom cells. M denotes the defect that is used as marker to calculate the remaining two.

	S		S_2		T	
	(0/+)	(+/++)	(0/+)	(+/++)	(0/+)	(+/++)
exp.	0.3	0.6	0.19	0.37		
65 atom	M	M	0.17	0.41	0.06	0.31
65 atom	0.33	0.57	M	M	0.08	0.27
217 atom	M	M	0.21	0.43	0.11	0.28
217 atom	0.28	0.54	M	M	0.09	0.23

3.5 Conclusions

In summary, the calculations demonstrate that the silicon self-interstitial is stable at the T-site in the double ionised state for a Fermi level below $\sim E_c - 0.35$ eV in agreement with the double donor level of AA12 at $E_c - 0.4$ eV. Above this the T interstitial is in the +1 charge state. The $\langle 110 \rangle$ and H' structure have the lowest formation energy for a Fermi level above $\sim E_c - 0.2$ eV (see Fig. 3.7). In n-type Si the interstitial is therefore neutral and its barrier to migration is about 0.35 eV in agreement with the observation of interstitial movement at around 160 K. Radiation enhanced migration and the low barriers explain the fact that the isolated interstitial is not observable in n-type Si. In p-type Si the barrier to migration is about 1 eV since the charge state is now +2 implying a thermal stability to around 100°C. The isotropic spin density at the T interstitial was found to be very small.

These properties are qualitatively consistent with those of the AA12 centre and support its identification with the self-interstitial.

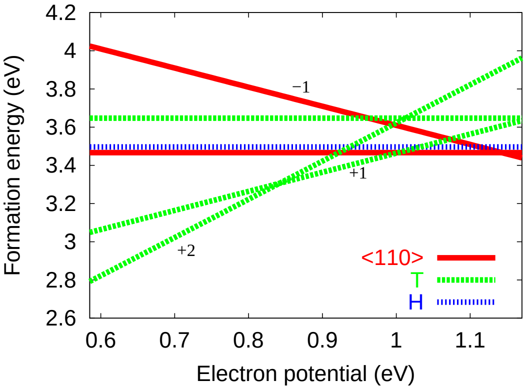


Figure 3.7: Formation energies of the interstitial in the upper half of the band gap for the T, H and $\langle 110 \rangle$ form, taking into account the symmetry distortions of T^+ and H^0. T^0 is unstable and relaxes towards H^0.

CHAPTER 4

Rod-like Defects in Silicon

4.1 Introduction

Rod-like defects (RLDs) are elongated structures with different habit planes. Most of them lie on a {113} habit planes. RLDs have been observed using high resolution transmission electron microscopy (HRTEM) in silicon and germanium following electron-irradiation, ion-implantation, or thermal annealing of oxygen rich Si (Muto and Takeda, 1995). An anneal around 450°C is usually necessary to break up smaller interstitial clusters to form the RLDs. The RLDs are metastable and transform to perfect dislocation loops lying on {111} planes over extended anneals. The atomic structure of RLDs was deduced in a remarkable HRTEM study (Takeda, 1991) and supported by subsequent atomistic modelling (Kohyama and Takeda, 1992, 1995; Kim *et al.*, 1997, 2000; Goss *et al.*, 2002). The RLDs are up to a micron long and perhaps 0.1 μm wide. They are only periodic along $[1\bar{1}0]$.

Their evolution has been investigated using HRTEM. p-CZ Si was electron irradiated at 450°C. Only after about an hour were any defects detected. A line of about 100 interstitials formed a $[1\bar{1}0]$ chain through the thin foil. Observations were made on as few as three such chains which occupied three $[1\bar{1}0]$ channels, separated by $[33\bar{2}]a_0/2$ with a (113) habit plane (Takeda and Kamino, 1995). The insertion of $[1\bar{1}0]$ chains leads to characteristic 5 and 7 fold rings when viewed along $[1\bar{1}0]$ (see Fig. 4.1). One of the bonds of each 5-fold ring is then switched with a neighbouring 6-fold ring to separate the 5-fold rings and make a 5–6–7 ring sequence (see Fig. 4.1(c)). This is called an I-defect and is separated from normal 6-fold rings by a 5–7 ring pair. A set of I-defects can then be inserted along $[33\bar{2}]$. The arrangements of I defects is not periodic but interrupted by O-defects at various distances. These involve a bond switch at the terminating 7-rings and turn these 7-rings into an 8-ring adjacent to another 5-ring (see Fig. 4.2(b)). The actual structure of the {113} defect is quite complex and its elucidation by HRTEM represents a remarkable achievement.

Nevertheless, it is also possible to stack $\langle 110 \rangle$ chains onto {111} habit planes, and typically 10% of the RLDs have this habit plane (Fedina *et al.*, 1998, 1999). In Ge (Muto and Takeda, 1995), and possibly in silicon (Takeda, 1997; Claverie, 2003), a third structure lying in {001} planes has been detected in samples also containing RLDs. Such {001} oriented structures are common in heat-treated nitrogen-containing diamond. These defects are composed of $[1\bar{1}0]$ chains of bonded tetra-interstitials (Humble *et al.*, 1985; Goss *et al.*, 2000, 2001, 2003), and possess a completely distinct structure from the {113} and {111} RLDs. There is also evidence that vacancies can form planar defects, with HRTEM revealing a planar vacancy-related structure inhabiting a {113} plane which may subsequently play a role in the formation of the interstitial defect (Fedina *et al.*, 1999).

RLDs are not stable above about 900°C. They evolve firstly in stacking faults bordered by Frank partial dislocations and ultimately to perfect {111} dislocation loops in both Si and Ge. HRTEM studies have shown that on average 75% of the RLDs transform into dislocation loops (Robertson *et al.*, 2000). In fact the remaining fraction may also convert to loops but simply be unresolved in the experiment, or some dislocation loops are nucleated from self-interstitial clusters too small to be detected.

Dislocation loops are both a source and sink for self-interstitials (Claverie *et al.*, 1999), and the size distribution of faulted loops shows a growth broadly consistent with an Oswald ripening process. For short anneals one finds a population of

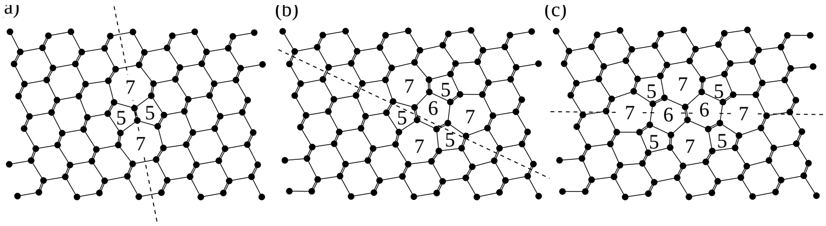


Figure 4.1: Interstitial $[1\bar{1}0]$ chains projected onto the $(1\bar{1}0)$ plane. The vertical and horizontal axes are [113] and $[\bar{3}\bar{3}2]$ respectively. (a) shows the structure generated by the insertion of a $[1\bar{1}0]$ interstitial chain, (b) after one bond-switch described in text, and (c) after a second bond switch. Five- and seven-member rings are indicated, as are 'hexagonal' six-member rings.

{113}, and both perfect dislocation loops (PDL) and faulted dislocation loops (FDL): the {113} defects dissolve first and for long anneal times or higher anneal temperatures only FDLs remain (Claverie *et al.*, 1999). Cristiano *et al.* (2000) examined the relative stability of faulted and perfect loops in silicon. For material implanted with high energy Ge ions followed by rapid thermal annealing at 900°C it was shown that the initial populations of faulted loops was higher than perfect loops. Faulted loops dominate over perfect loops for radii smaller than ∼40 nm.

It has been suggested that in Si RLDs possess characteristic photoluminescence (PL) and electrical levels determined from deep level transient spectroscopy (DLTS). Implantation and annealing produces a number of electrically active defects. Many of these are point defects which anneal below 600°C. For high dose implantations in *p*-Si followed by an anneal above 600°C, RLDs are formed for which a DLTS peak at $E_v + 0.50$ eV has been assigned (Benton *et al.*, 1997; Coffa *et al.*, 2000; Libertino *et al.*, 2001). Optical data closely follows the electrical characterisation (Coffa *et al.*, 2000; Libertino *et al.*, 2001), and for high dose implantations followed by annealing above 680°C, extended defects of the {113} type are formed, as characterised by the observation of the 0.9 eV (1376 nm) peak (Coffa *et al.*, 2000; Libertino *et al.*, 2001), and where the implantation damage is greatest a broad band centred around 0.8 eV is also seen (Coffa *et al.*, 2000). The 0.9 eV band, which is commonly referred to as the '9' or '903' line has been studied in some detail (Jeyanathan *et al.*, 1994; Lightowlers *et al.*, 1994). The association of the PL with the RLDs seen in HRTEM was made by observation of the etch pits under annealing. Furthermore, uniaxial stress measurements point to a defect with C_{1h} symmetry, i.e. a mirror plane, in which the dipole moment lies, pointing along close to [111]. The large magnitude of stress splitting and lack of vibronic coupling is thought to suggest rather delocalised electronic states are involved in the optical transition, and the lack of an effect under magnetic field suggests that the defect is neutral. The overall picture is consistent with the RLD trapping excitons and the radiative recombination occurring at some C_{1h} site associated with this defect. The same defect could not explain the $E_v + 0.5$ eV centre. Other transitions which are less well characterised are thought to arise at sites associated with the RLDs but which have a lower density.

The structure and energetics of the {113} RLDs have been studied extensively using a range of theoretical techniques. The first studies complemented the initial HRTEM studies, and took the form of simple 3-body inter-atomic potential calculations. This proved rather effective, generating structures that matched the HRTEM very closely (Parisini and Bourret, 1993; Kohyama and Takeda, 1992). Tight-binding calculations have also been used to predict a range of properties associated with the {113} RLDs (Kohyama and Takeda, 1995; Kim *et al.*, 1997). More recently density functional techniques predicted the energies of the various types of stacking of the {113} defects and showed that the role of the length of the chains and width of the rods was important (Kim *et al.*, 2000; Goss *et al.*, 2002). However many of these studies based on super-cell calculations appear to neglect the strain associated with the finite size of the cells. It has been shown from tight-binding calculations (Alippi and Colombo, 2000) that the strain fields associated with the [110]-chains is large and the driving force for the aggregation in the {113} planes. All atoms other than those at the chain ends are four-fold coordinated and possess bond length and angles reasonably close to bulk values (Kohyama and Takeda, 1995).

4.2 Atomic Models for Planar Interstitial Aggregates

The structure of the {113} defect is presented schematically in Figs. 4.1 and 4.2. The core of the defect, shown in Fig. 4.1, is a chain of self-interstitials inserted into a $[1\bar{1}0]$ channel. There are two self-interstitials per $a_0[1\bar{1}0]/2$ period which break bonds between Si atoms of the surrounding cage. Reconstruction then occurs involving the bonds of atoms of the inserted chain so that all atoms become fully coordinated except those at the chain ends. The isolated chain constructed in this way is shown in Fig. 4.1(a). There is evidence from HRTEM studies that defects involving as few as two or three isolated chains exist (Takeda *et al.*, 1994; Takeda and Kamino, 1995). Additional distortions involving bond switching between atoms surrounding the core must occur to account for the HRTEM results in RLDs (Takeda *et al.*, 1994). The structures following one and two bond-switching operations are shown in Fig. 4.1(b) and (c), respectively. The first bond-switch separates the 5-member rings and has been found to reduce the energy considerably (Kim *et al.*, 1997; Alippi and Colombo, 2000). This intermediate configuration plays a role in RLDs lying in more than one {113} plane.

Table 4.1: Previously published formation energies per self-interstitial (eV) for the structures depicted in Fig. 4.1.

(a)	(b)	(c)	Method	Ref.
-	-	1.6	Stillinger-Weber potential	Kohyama and Takeda (1992)
2.2	1.7	1.7	Tight-binding	Kim *et al.* (1997)
2.05	1.58	1.60	Tight-binding	Alippi and Colombo (2000)
-	-	1.02	LDA-DFT	Kim *et al.* (2000)
2.0	1.5	1.5	LDA-DFT	Goss *et al.* (2002)

A second bond-switch, symmetrically equivalent to the first and shown in Fig. 4.1(c), does not significantly lower the energy (Kim *et al.*, 1997; Alippi and Colombo, 2000; Goss *et al.*, 2002), but is implied by the HRTEM simulation (Takeda *et al.*, 1994).

Previous theoretical studies have produced a fairly narrow range of formation energies per interstitial for the three models of an isolated chain shown in Fig. 4.1. These are summarised in Table 4.1, which shows that the second bond-switch in some cases actually *increases* the formation energy. Although all silicon atoms in an *infinite* chain are fully four-fold coordinated, the strain associated with the inclusion of the interstitials may lead to gap states. This is believed to be the case for the isolated tetra-interstitial in silicon which is also fully-coordinated, but possesses a donor level lying near E_v associated with compressed bonds. In the ionised state the resulting defect is believed to be the B3 EPR centre (Coomer *et al.*, 2001). Although tight-binding calculations suggest that there are no localised gap levels associated with *any* of the chains of Fig. 4.1 (Alippi and Colombo, 2000), recent density-functional calculations have suggested that the structure in Fig. 4.1(a) possesses both occupied and empty levels in the band-gap. However, the bond switching which lowers the energy, apparently leads to their elimination although it is unclear whether residual defect bands lie near the band edges (Goss *et al.*, 2002).

To form an extended two dimensional platelet lying in a {113} habit plane, additional chains are placed parallel to those in Fig. 4.1(c). The final platelet depends on the separation of $[1\bar{1}0]$ chains lying in the (113) plane and Fig. 4.2 shows four different *periodic* models. The difference arises from the repeat frequency of the $[1\bar{1}0]$ chains in the $[\bar{3}\bar{3}2]$ direction. Following the notation of Kim *et al.* (1997) where I represents an interstitial chain, and O an 'empty' channel. Fig. 4.2 shows /I/, /IO/, $/\mathrm{I}_a\mathrm{I}_b\mathrm{O}/$ and $/\mathrm{I}_a\mathrm{I}_b\mathrm{I}_c\mathrm{O}/$, where the subscripts label the location of the chains in the periodic structure. Since the horizontal lattice vector for the host material is given by $a_0[\bar{3}\bar{3}2]/2$, the true periodic descriptions for structures (a) and (c) are $/\mathrm{I}_a\mathrm{I}_b/$ and $/\mathrm{I}_a\mathrm{I}_b\mathrm{O}_\alpha\mathrm{I}_c\mathrm{I}_d\mathrm{O}_\beta/$, respectively. For example, for structure (c) alternating chains are in anti-phase with each other, as shown in the close-up view in Fig. 4.2(e).

Previous calculations have predicted different types of stacking to have the lowest energy. Tight binding (Kim *et al.*, 1997) and LDA (Goss *et al.*, 2002) calculations yield /IIO/, with the latter suggesting a formation energy per interstitial of just 0.48 eV. Another LDA study (Kim *et al.*, 2000) preferring /IIIO/, with a formation energy per interstitial of 0.55 eV, although in this study no energy was given for the /IIO/ structure. The /IIO/ geometry has a density of self-interstitials broadly consistent with experiment (∼ 2/3 of the channels are filled with interstitial chains (Takeda, 1991; Kohyama and Takeda, 1992)). Many studies report formation energies per interstitial which are much higher than those quoted above, such as Alippi and Colombo (2000) where the minimum formation energy is 1.13 eV, while Kim *et al.* (1997) give formation energies greater than 1.16 eV. The source of the large variation in these energies may be in part due to the different computational schemes, but it seems more likely that the discrepancy is from the large elastic term due to artificial constraints on the volume of the computational cells.

The majority of RLDs inhabit a {113} plane, but a minority (∼10-15%) lie in a {111} plane (Fedina *et al.*, 1998, 1999; Chou *et al.*, 1995). The structure of these is not completely understood but they are thought to involve $[1\bar{1}0]$ chains lying in a (111) plane (Chou *et al.*, 1995) as shown in Fig. 4.3 where the packing density of chains is maximum. This structure can be generated from Fig. 4.1(a) by repeating the $[1\bar{1}0]$ chain in the (111) habit plane indicated by the dashed line. Empirical potential calculations of Chou *et al.* (1995) indicate that *infinite* {111} oriented RLDs were 20% *lower* in energy (per self-interstitial) than those with a {113} habit plane, which was confirmed by recent density functional calculations, although the magnitude of the formation energies were rather different (the empirical method and LDA calculations yield 1.31 eV and 0.46 eV respectively for the {113} RLD) (Goss *et al.*, 2002). The energy lowering effect of the bond switch does not appear to be as important in this case, probably due to the parallel chains cooperating to relieve the strain along [111].

The final class of interstitial platelets lie in {001} habit planes. These are the only form of self-interstitial planar aggregates seen to date in diamond. They also occur in Ge, where they co-exist with the RLDs, and there is evidence that they can exist in silicon (Takeda, 1997; Claverie, 2003). The structure is not completely defined, but models based on the tetra-interstitial exhibit saturated bonding (Ferreira-Lima, 1975; Humble *et al.*, 1985; Muto and Takeda, 1995; Goss *et al.*, 2000, 2001). Fig. 4.4 shows the simplest structure based on tetra-interstitials. In diamond this possesses a formation energy per interstitial just 8% that of an isolated self-interstitial (Goss *et al.*, 2001, 2003).

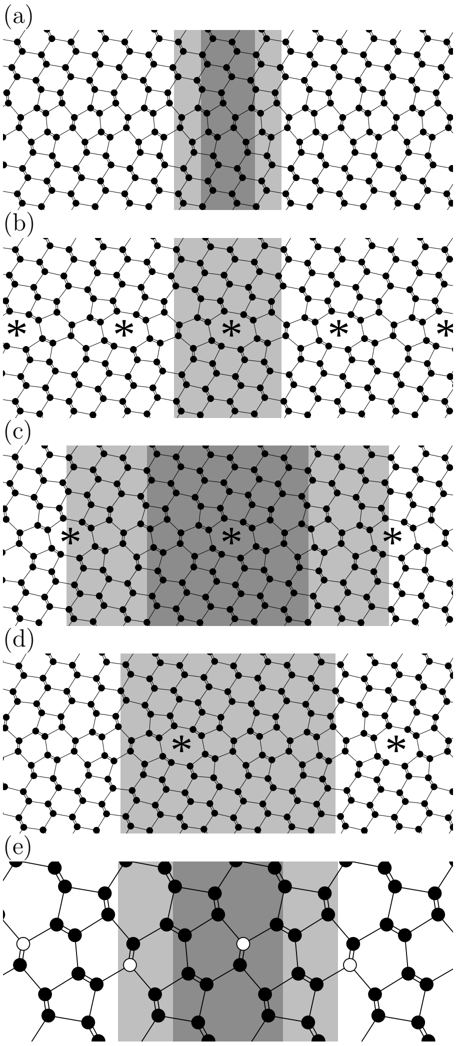


Figure 4.2: Schematics of four {113} interstitial platelets projected onto the $(1\bar{1}0)$ plane. The vertical and horizontal axes are [113] and $[\bar{3}\bar{3}2]$ respectively. The light-shaded areas indicate a unitcell of each structure, and, for structures (a) and (c) the darker shaded area shows the apparent periodicity where it differs from the actual period along $[\bar{3}\bar{3}2]$. The aggregates can be described as (a) /I/, (b) /IO/, (c) /IIO/ and (d) /IIIO/, where I represents the single chain shown in Fig. 4.1(c), and O an 'empty' $[1\bar{1}0]$ channel which forms an 8-member ring, indicated by an asterisk. (e) shows (a) close-up, with white circles indicating interstitial atoms in the chains that are deeper into the $(1\bar{1}0)$ plane than those shown in black, the difference in depth being $a_0[1\bar{1}0]/4$.

4.3 Calculation Details

Isolated $[1\bar{1}0]$ chains were simulated using various unit cell geometries. For {113} platelets, cells with lattice vectors $a_0[1\bar{1}0]/2$, $n\,a_0[113]/2$ and $m\,a_0[33\bar{2}]$ containing p chains of $[1\bar{1}0]$ interstitials were utilised. m, and p are chosen such to represent the models in Fig. 4.2 with the minimum number of atoms, and n is typically 1, but $n = 2$ was also examined to test the effects of the periodicity along this direction. Where the unit cells are allowed to relax completely, there is no significant difference between $n = 1$ and $n = 2$. {111} RLDs were analysed with cells whose vectors are $n\,a_0[111]/2$, $a_0[1\bar{1}0]/2$, and $a_0[11\bar{2}]/2$ where n was taken to be 9 and 12. {001} platelets were investigated using cells with vectors

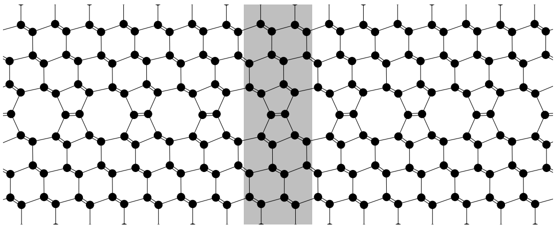


Figure 4.3: Schematic structure of the {111} interstitial platelet, projected onto the $(1\bar{1}0)$ plane. The vertical and horizontal axes are [111] and $[11\bar{2}]$, respectively. The periodicity is indicated by the grey band.

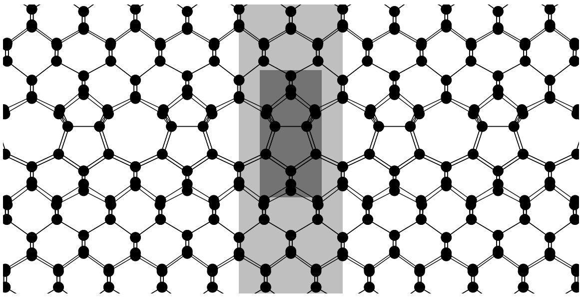


Figure 4.4: Schematic structure for the {001} self-interstitial platelet in group IV materials, projected onto the (110) plane. The vertical and horizontal axes are [001] and $[1\bar{1}0]$, respectively. The periodicity is indicated by the light grey band. The dark grey area indicates an individual tetra-interstitial unit.

$4\,a_0[001]$, $a_0[110]$, and $a_0[1\bar{1}0]$.

The relative stability of different platelets can be found from the formation energy, $E^f(\mathrm{X})$, defined by Eqn. 3.1. With the formation energies per interstitial defined in this way defects containing different numbers of interstitials can be compared.

For planar systems the volume of the unit cells has been relaxed, typically by varying the lattice vector normal to the plane of the defect to minimise the total energy. The change in the lattice constant found from the volume relaxation also gives an estimate of the Burgers vector for the RLDs and other planar interstitial aggregates. This procedure has been used previously for similar defects in diamond with close agreement with experiment (Goss *et al.*, 2000).

Electronic density of states (DoS) and the joint density of states (JDoS) are evaluated by calculating the Kohn-Sham spectra for the given system over a dense Monkhorst-Pack (Monkhorst and Pack, 1976) mesh of points within the Brillouin-zone such that the reciprocal-lattice volume represented by each point is approximately a cube of side length around 0.1 Å^{-1}. Each band is broadened using a Gaussian of width 0.1 eV.

4.4 Results

4.4.1 $[1\bar{1}0]$ interstitial chains

As shown from tight-binding calculations (Alippi and Colombo, 2000), one expects an 'isolated' $[1\bar{1}0]$ chain of self-interstitials to give rise to strain fields that may lead to interaction with neighbouring chains in adjacent cells. In order to minimise the impact of such interactions the defects have been relaxed in a range of unit cell orientations and geometries.

First the results from a unit cell made up from $6 \times 6 \times 1$ fcc unit cells containing 72 host atoms and having lattice vectors $a_0[30\bar{3}]$, $a_0[03\bar{3}]$, and $a_0[1\bar{1}0]/2$ are report. The three core structures are shown schematically in Fig. 4.1. As found previously (Kim *et al.*, 1997; Alippi and Colombo, 2000), the initial bond-switching procedure lowers the formation energy considerably. For structures (a), (b) and (c) the calculated formation energy per interstitial is 2.0, 1.5 and 1.5 eV, respectively. The interaction of the defects in different cells can be investigated by changing one of the lattice vectors, so that, for example, $a_0[03\bar{3}]/2$ becomes $a_0[36\bar{9}]/2$, and $a_0[1\bar{1}0]/2$. Such changes were found to affect the formation energy per interstitial by less than 0.1 eV (< 5%).

Additionally, the cross-sectional area has been relaxed by *isotropically* varying the two lattice vectors perpendicular to the chain direction in order to minimise the cell energy. This results in a net lowering of the formation energy for each structure to 1.7, 1.2 and 1.2 eV /I, for structures (a), (b) and (c), respectively. Note that the unrelaxed formation energies are very close to tight-binding values (Kim *et al.*, 1997; Alippi and Colombo, 2000), but higher than the more recent LDA results (Kim *et al.*, 2000). The in-plane lattice constants are increased by 1.5%, 1.6% and 1.6% for zero, one or two bond-switches, respectively.

To see explicitly the effect of cell size, a larger cell was constructed from $7 \times 7 \times 1$ fcc unit cells. The formation energies for the three structures are 1.8, 1.4 and 1.4 eV /I for the fixed volume case. Similarly, a supercell made up from $8 \times 6 \times 1$ fcc cells yields 1.9, 1.4 and 1.3 eV. The formation energy of structure (a) has been recalculated in a cell with lattice vectors $a_0[330]$, $a_0[1\bar{1}0]/2$ and $a_0[22\bar{4}]$, yielding 1.8 eV /I.

The different cell sizes and geometries model the strain interaction of the interstitial chains with the periodic images in different ways, but give rise to very similar formation energies. In particular, note that these energies lie in a range bounded by the fixed and relaxed volume calculations in the $6 \times 6 \times 1$ fcc cell, and it therefore seems reasonable to assume that the formation energies are given to within around 0.1 eV by 1.8, 1.3 and 1.3 eV /I for structures (a), (b) and (c), which fits in well with the previously published values summarised in Table 4.1.

The electronic DoS for the three structures of Fig. 4.1 are shown in Fig. 4.5. There are two conclusions to be drawn

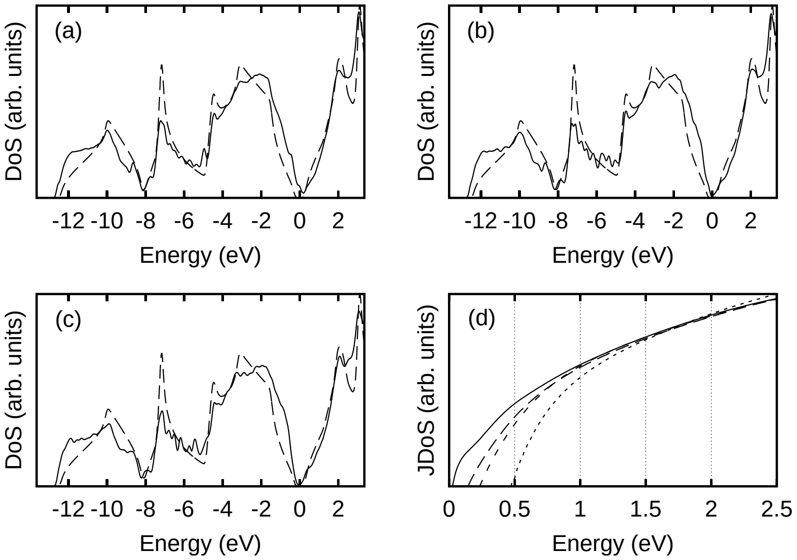


Figure 4.5: Electronic density of states for the [110]-chains in silicon. (a)–(c) show the defect DoS (solid lines) alongside those of bulk silicon (dashed lines) for the structures (a)–(c) in Fig. 4.1. (d) shows the JDoS for all three structures with the solid, long-dashed and short-dashed lines representing (a)–(c) in Fig. 4.1, respectively, with the dotted line indicating the JDoS for bulk silicon.

from these figures. The first, and most significant result is that *isolated* infinite [110]-chains give deep strain-induced gap levels. In fact, for some cell geometries the [110]-chain, before bond-switching, is nearly metallic, as can be seen from the JDoS in Fig. 4.5(d). Other cells lead to a small band-gap, but the general trend is for close-by empty and filled mid-gap strain-induced states. As generally known density functional theory underestimates the band-gap as can be seen in the JDoS for bulk silicon which has a gap of around 0.5 eV. This complicates the interpretation of the DoS and JDoS for the various structures. However, it seems reasonable to infer that there is a significant reduction in the band-gap for isolated [110]-chains. Such a conclusion is consistent with the view that the large strain caused by the presence of the interstitials leads to band gap narrowing.

The second conclusion is that the bond-switching which lowers the formation energy of the chains leads to an opening of a gap at the Fermi-energy corresponding to structure (a). One might therefore surmise that the bond-switching is an example of Peierls distortion, i.e. a structural reorganisation converting a one-dimensional metallic system into one which is semi-conducting. However, the magnitude of the splitting is insufficient to clear the gap and the gap levels persist for the bond-switched defect, in contradiction with the tight-binding results of Alippi and Colombo (2000), even though the defect level localisation follows very closely the strain field patterns that they obtained.

4.4.2 [110]-chain ends in silicon

To investigate the electrical properties of the chain ends a chain of six self-interstitials was inserted along $[1\bar{1}0]$ in 144 and 256 atom bulk cells made up from $3 \times 3 \times 2$ and $4 \times 4 \times 2$ conventional unit cells, respectively. One end of the chain was terminated with a hydrogen atom to remove one Si dangling bond preventing interaction with the second dangling bond at the other end. The resulting distance between dangling bonds in adjacent cells is ~10.7 Å.

With the change of charge state a qualitative structural change takes place. In the neutral charge state the chain end lies approximately in line with its 'ideal' site, i.e. where it would lie if the chain was continued. This is shown schematically in Fig. 4.6(a). However, when the chain end is positively charged the end atom moves about 0.85 Å along $[1\bar{1}0]$ and 0.3 Å along [001], yielding a three-member-ring similar to the tri-interstitial defect proposed for the W-line PL centre (Coomer *et al.*, 1999). The relaxed structure is depicted in Fig. 4.6(b). Therefore in the positive charge state all silicon atoms are *at least* four-fold coordinated. The donor level of this system can be estimated by comparing the ionisation energy of the defective cell with that of a bulk cell with the same geometry (see section 3.3). This gives a donor level at $\mathrm{E_v} + 1.05$ eV or $\mathrm{E_v} + 0.95$ eV for the 144 and 256 atom cells, respectively. The accuracy of such calculations is to within a few tenths of an eV, and the conclusion from these calculations is that the chain ends are likely to be donors with levels in the upper half of the band-gap, probably close to the conduction band minimum.

A Mulliken bond-population analysis reveals that only around 15% of the unpaired electron in the neutral charge state is localised on the end atom of the chain, perhaps suggesting that the donor level arises from the strain in a manner similar to the gap levels in the infinite chain (Sec. 4.4.1). To examine this possibility both ends of the 6 atom chain were terminated with hydrogen. The addition of the second H atom eliminates the shallow electrical level, but leaves a donor

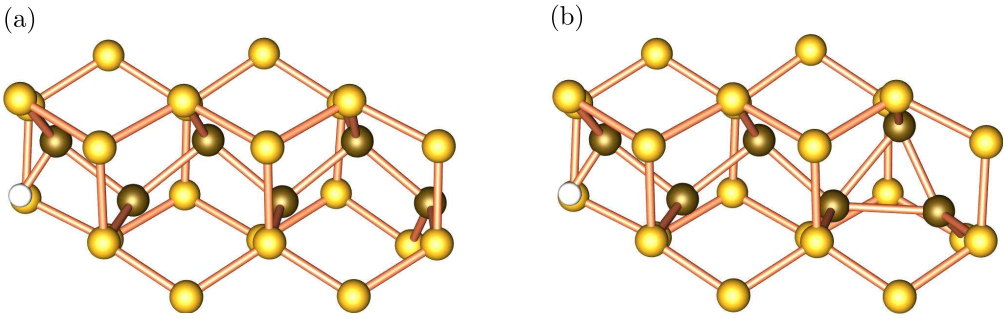


Figure 4.6: The H terminated 6 self-interstitial chain (black) viewed along [11$\bar{2}$]. The vertical and horizontal axes are [11$\bar{2}$] and [1$\bar{1}$0], respectively. (a) the neutral and (b) the positive charge states. Host silicon atoms are in grey, and H in white.

level at $E_v + 0.6$ eV. In fact this level arises from an occupied band that is also present in the partially hydrogenated case. However, in the partially hydrogenated case the deep band leads to a *double donor level* at $E_v + 0.6$ eV. This mid-gap level is consistent with the presence of deep levels in the infinite chain.

To conclude, the level close to the conduction band is dominated by the presence of a dangling bond but delocalised in the strain field of the chain. The level close to the conduction band agrees quite well with the electrical levels expected to participate in the luminescence signal at 903 meV (Lightowlers *et al.*, 1994; Jeyanathan *et al.*, 1994). Experiment shows that the defect responsible for this emission possesses C_{1h} symmetry and a dipole axis close to $\langle 111 \rangle$ and is attributed to RLDs or defect sites intimately associated with them. Note that in the limit of an infinite chain, the symmetry is C_{1h} with a strain field along $\langle 111 \rangle$. Also note that the donor level at $E_v + 0.6$ eV is in reasonable agreement with the level detected via DLTS at $E_v + 0.5$ eV (Coffa *et al.*, 2000; Libertino *et al.*, 2001).

In summary, it is tempting to correlate the 903 meV optical band with recombination processes at the periphery of the RLDs where there are dangling bonds, and the $E_v + 0.5$ eV DLTS defect with a strain induced level. However, given the uncertainties inherent in the calculation of electrical levels in silicon, as well as the limitation of the short chain examined in this study, such assignments can only be tentative.

4.4.3 {113} oriented planar aggregates

Various stacking sequences of the structure in Fig. 4.1(c) separated by 8-member rings can be constructed. As outlined in Sec. 4.2, Fig. 4.2 shows three types of stacking: /I/, /IO/, /IIO/ and /IIIO/. Table 4.2 lists the calculated formation energy per interstitial for various types of stacking. In each case the supercell volume has been relaxed by varying the lattice constant in the [113] direction, so that for the minimum in energy the lattice vector of the supercell is given as $(1 + \alpha)\, a_0[113]$. For each case the change in lattice parameter is given in Table 4.2, which can be interpreted in terms of a displacement measured in HRTEM experiments. The unit cells have also been allowed to relax in a manner similar to that described in previous studies (Kohyama and Takeda, 1995). The three lattice vectors are no longer constrained to be orthogonal allowing a shear component to the displacement vector, in line with the experimental observations. The three lattice vectors are then $a_0[1\bar{1}0]/2$, $m\, a_0[33\bar{2}]/2$, and $\{(1 + \alpha')[113] + \beta[33\bar{2}]\}\, a_0$, with the value of m appropriate to the periodicity of the model. The value of α is close to the values obtained without shear, and the additional energy saved by the inclusion of the shear component is relatively small, but significant. The calculation including the shear component requires considerable computational effort. The structures are therefore obtained approximately using a reduced k-point sampling scheme, and evaluate the converged formation energy using the relaxed geometry obtained in this way. To establish the accuracy of this approach the converged formation energy was calculated over the range of α' and β for the /IO/ structure, and the location of the minimum obtained with the minimal sampling scheme was found to be in agreement with that that yields the converged formation energy. The energy drop associated with the shear is very modest, and in particular they are only marginally significant with respect to the error in the formation energy with

Table 4.2: Formation energy per interstitial, E^f (eV) for the various periodic structures considered for the {113} oriented RLD. For each structure the displacement vectors are listed in terms of the parameters α, α' and β, as defined in the text, along with the drop in E^f associated with the inclusion of shear, ΔE^f (meV). m is the number of $a_0[33\bar{2}]/2$ periods in the model.

Model	m	E^f	$\alpha \times 10^2$	$\alpha' \times 10^2$	$\beta \times 10^3$	ΔE^f
/I/	1	0.69	7.3	7.4	7.6	3
/IO/	1	0.65	6.7	7.0	7.6	8
/IIO/	3	0.47	6.8	7.0	9.1	8
/IIIO/	2	0.48	7.1	7.0	9.1	9

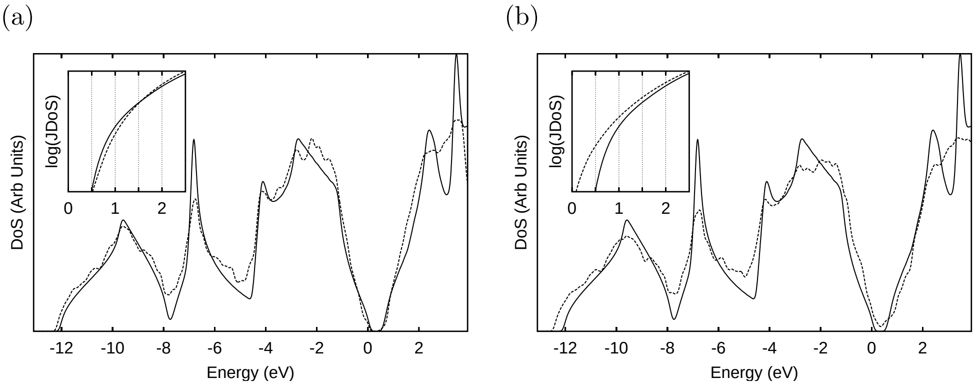


Figure 4.7: Electronic density of states (DoS) calculated for the /IIIO/ structure in fig. 4.2 in silicon (dashed line) plotted with that of bulk silicon (solid line). Inset shows the joint density of states for the same systems. (a) for the completely relaxed structure and (b) for the lattice constant in the [113] direction fixed to that of bulk silicon.

respect to the computational parameters is thought to be a few meV. However, the relatively flat energy surface as a function of the shear component is consistent with the wide range of displacement vectors quoted in the literature, which may vary simply due to local strains in the material dominating the distortion of the lattice around the RLD.

The magnitudes of the formation energies broadly agree with those previously published for silicon from tight-binding and DFT calculations. The /IIO/ and /IIIO/ types of stacking have similarly low energies, with reasonable agreement with previous LDA calculations (0.55 eV, Kim *et al.* (2000)).

The electronic density of states for all structures indicate that there are no *deep* electronic states. However, the strain associated with the presence of the planar defect could lead to a net band-gap narrowing, as suggested by previous calculations (Goss *et al.*, 2002). However, this is not unambiguously the case, and other studies have indicated that the {113} defects do not introduce levels into the band-gap (Kohyama and Takeda, 1995; Alippi and Colombo, 2000). Fig. 4.7(a) shows the density of states (DoS) and joint density of states (JDoS) obtained for the /IIIO/ relaxed structure along with that of a defect free cell. The indication is that if there is band gap narrowing the effect is likely to be extremely modest. However, if the DoS and JDoS are obtained for a constrained {113} defect then the increase in strain induces a strong band-gap narrowing, as indicated in Fig. 4.7(b). In this figure the DoS's have been plotted such that Fermi-levels are coincident, to facilitate comparison. This has no effect on the JDoS which shows a reduction in the band-gap almost to zero despite all silicon atoms being fully four-fold coordinated. Although the constraint of the lattice is somewhat artificial, the results suggest that in the vicinity of the edges of the RLDs where the strain is not able to relax in the same way as an infinite platelet, there would be the potential for strain induced band gap states. Consequentially optical and electrical activity may then occur. Unfortunately quantifying these effects is beyond the scope of the current study.

4.4.4 {111} oriented planar aggregates

Unlike the {113} planar defects, the self-interstitial aggregate formed by condensing $[1\bar{1}0]$ chains onto a {111} plane has only one low energy stacking sequence (Fig. 4.3). The cells were constructed to have lattice vectors $a_0[1\bar{1}0]/2$, $n\,a_0[111]$, and $a_0[11\bar{2}]/2$. The formation energies were calculated for n=3 and n=4, and it was found that subsequent to the relaxation normal to the plane of the defect the formation energies agreed to within around 1%. As with the {113} oriented RLDs, the volume of cells containing the {111} structures were relaxed by allowing the lattice constant normal to the plane of the defect to increase such that the formation energy per interstitial is minimised. The possibility of a shear component in the $[11\bar{2}]$ direction was also examined, but it was found that this is not energetically favoured.

The dilation of the lattice is estimated to be $\sim 0.21\,a_0[111]$, i.e. $\sim a_0[111]/5$. This is in agreement with the value quoted by Fedina *et al.* (1999), and twice that of Chou *et al.* (1995).

Perhaps surprisingly, the {111} structure is found to be *lower* in energy than any {113} structures that has been studied. The minimum formation energy per interstitial is calculated to be 0.36 eV with a dilation of the lattice of $0.21\,a_0\langle 111\rangle \approx a_0\langle 111\rangle/5$.

The DoS and JDoS for the {111} RLD is plotted in Fig. 4.8. In particular there is no evidence of any electronic states being introduced into the band-gap, either from the DoS or JDoS. One conclusion is that the two orientations of RLDs do not generate any deep levels arising from the fully four-fold-coordinated regions. However, one should bear in mind that the electronic levels associated with the chain-ends will occur in both structures, and the regions close to the edges of the RLDs are likely to also have an effect on the band-gap, as indicated for the {113} oriented structure (Sec. 4.4.3).

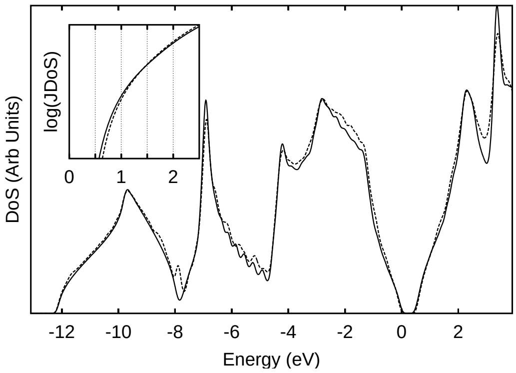


Figure 4.8: Electronic density of states (DoS) calculated for the {111} structure in fig. 4.3 in silicon (dashed line) plotted with that of bulk silicon (solid line). Inset shows the joint density of states for the same systems.

4.4.5 {001} oriented planar aggregates

As discussed in the introduction, for both diamond (Humble *et al.*, 1985) and germanium (Muto and Takeda, 1995), self-interstitials can aggregate on {001} planes. For diamond the model has been recently explored, (Goss *et al.*, 2003) with a structure based on the tetra-interstitial giving a particularly low formation energy. The structure of these platelets in diamond is shown schematically in Fig. 4.4. For diamond the difference in energy between the various reconstructions is relatively modest, and the calculations have been restricted to the model shown in Fig. 4.4.

As with previous structures, the volume is relaxed normal to the plane of the defect. The formation energy per interstitial for the volume relaxed cell is 0.45 eV. This value is very close to the /IIO/ and /IIIO/ types of stacking for the {113} RLD. Unlike the RLDs, the {001} planar aggregates do not *necessarily* give rise to any unsatisfied bonds, provided that the number of interstitials is a multiple of four. In a previous analysis of the RLDs on the {113} habit plane, the size dependence of the aggregates was argued based on a balance between the various component parts of the formation energy, including two dangling bonds per [110] chain (Kim *et al.*, 2000). If the geometry is thermodynamically driven, it is this contribution that gives rise to large aspect ratio, and rod-like nature of the [110]-chain agglomerates. However, it seems likely that kinetic factors will also be important. In contrast, the {001} oriented planar aggregate based on the tetra-interstitial, thermodynamically and kinetically there is no driving force to prefer, say, [110] over [$\bar{1}$10], and the defect is free to obtain an aspect ratio close to unity.

The DoS and JDoS for the {001} planar aggregate is shown in Fig. 4.9. The JDoS suggests an increase in the band-gap associated with the presence of a {001} self-interstitial platelet, and must be a consequence of the strained bonds in the defect. This could lead to possible optical activity via the recombination of excitons trapped in the strain field of such as defect.

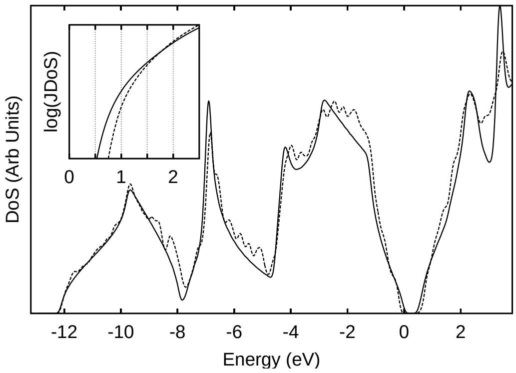


Figure 4.9: Electronic density of states (DoS) calculated for the {001} structure in Fig. 4.4 in silicon (dashed line) plotted with that of bulk silicon (solid line). Inset shows the joint density of states for the same systems.

Table 4.3: Formation energy per self-interstitial (eV) for Si and Ge in the three crystallographic planes considered. The results for diamond are from Goss *et al.* (2003) for comparison. The formation energy densities are in units of eV/a_0^2, n is the interstitial density (a_0^{-2}) which is independent of the material. The magnitude of the dilation normal to the habit plane as a measure of the Burgers vector, $|\vec{b}|$, is listed in units of a_0. The {113} structures are /IIO/.

	{113}	{111}	{001}
E^f	0.47	0.36	0.45
n	1.6	2.3	2.0
E^f/a_0^2	0.75	0.82	0.90
$\|\vec{b}\|$	0.23	0.36	0.35

4.5 Finite Aggregate Size Effects

In the preceding analysis the emphasis has been given to *infinite* aggregates of self interstitials. From this a result was obtained in agreement with previous low-level theory calculations that predicted the {111} RLDs to have a lower formation energy than the {113} oriented variety. This is in stark contrast to the experimental observation that {113} RLDs constitute more than 80% of the population. The formation energy per interstitial are summarised in Table 4.3.

However, such arguments can be misleading for a number of reasons. (i) The planar aggregates in reality are finite and edge effects such as dangling bonds and the elastic energy arising from a bounding dislocation will contribute to the formation of these defects. (ii) The nucleation of the planar aggregates is unclear. Some suggestion has been made that pre-existing vacancy or 'zero-defects' lying in the {113} plane form the basis of the RLDs, and it is the energetics of the formation of these centres that controls the populations of the two RLD orientations. (iii) There will be *kinetic* factors that affect the formation, particularly differentiating between the RLDs and the {001} defects. It is difficult to quantify (ii) and (iii), but one can estimate the contribution to the formation energy of the planar aggregates from the elastic energy.

To estimate the energy of a finite RLD elasticity theory was used with the results for the infinite RLDs. Since there is no simple formula for the elastic energy of the dislocation line of a rectangular platelet of the form of RLDs the finite RLDs are approximated by circular platelets. The total energy of a circular platelet can be written as the sum of the energy of the dislocation line and the platelet energy (Hirth and Lothe, 1968).

$$E = \frac{\mu b^2 r}{2(1-\nu)}\left(\ln\left(4\frac{r}{b}\right)-1\right) + \pi r^2\sigma, \qquad (4.1)$$

where μ is the shear modulus (7.55x10^{10} N/m^2), ν is the Poisson's ratio (0.3), σ is the platelet energy per unit area and b the modulus of the Burgers vector. The latter two are taken from the infinite RLDs. The radius for each platelet has to be chosen so that all platelets have the same number of interstitials N. With n being the density of interstitials per unit area the radius is

$$r = \sqrt{N/\pi n}.$$

Comparing the energies of the {113}, {001} and {111} platelets one finds that for N less than ~1000–3000 interstitials the {113} /IIO/ platelet is lowest in energy. For more than this number of interstitials the {111} platelet is favoured (see Fig. 4.10). This number of interstitials can of course not directly be related to the observed RLDs due to the different

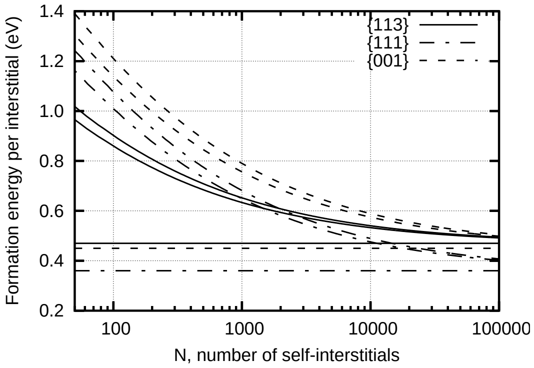


Figure 4.10: Formation energy per self-interstitial including the elastic energy of circular platelets in the three different habit planes according to Eqn. 4.1. For each orientation there are two lines representing representing a range corresponding to a ±0.3% error in the estimate of the Burgers vectors. The horizontal lines show the asymptotic limits for each case.

geometry, but since for a RLD the ratio of the dislocation line to the dislocation area is even bigger than for a circular

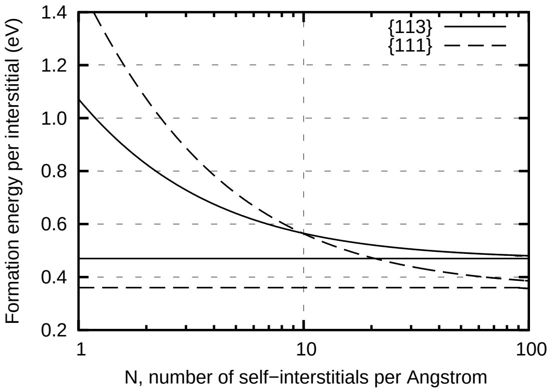


Figure 4.11: Formation energy per self-interstitial including the elastic energy of infinite dislocation dipoles according to Eqn. 4.2. The horizontal lines show the asymptotic limits for each case.

platelet the transition should occur at an even larger value. The energy of the {001} platelet is always higher than the energy of the {111} platelet and only becomes lower than the energy of the {113} for high N.

One can also simulate the elastic energy of an infinitely long dislocation dipole using the expression for the energy per unit length:

$$E/L = \frac{\mu\, b^2}{2\pi(1-\nu)}\left(\ln\left(\frac{r}{b}\right)+1\right)+r\sigma, \tag{4.2}$$

where the terms have the same meaning as in Eqn. 4.1, with the exception that r now represents the separation of the two dislocations. This may be a more reasonable representation of the RLDs given their large aspect ratio. Comparing the {113} and {111} orientations using Eqn. 4.2 yields the data in Fig. 4.11. The formation energies cross over for a dipole separation of 10.0 self-interstitials per Å, which amounts to around 19 [110]-chains. Allowing for an error in $|\vec{b}|$ in line with the circular loops (Fig. 4.10) yields an uncertainty in the cross-over points of the order of 2–4 [110]-chains. Thus the elastic energy of these infinitely long dipoles tends to support the formation of rather wide {113} RLDs in line with experimental observation.

A broad conclusion from this analysis is that the {113} habit plane might be favoured over {111} since it reduces the strain energy. This can mainly be attributed to the fact that it can introduce so-called O units which in turn is because of the bond-switching that already lowers the energy of an individual chain.

However, it is impossible to rule out a kinetic role in the preference for {113} defects over both the {111} and indeed the very stable {001} orientations. Furthermore the existence of other defects that either nucleate interstitial aggregation on particular planes is difficult to assess as is the presence large defects lying in the {113} plane such as the vacancy clusters and ‘zero-defects’ suggested previously (Fedina *et al.*, 2000).

4.6 Conclusions

In line with previous calculations it was found that a bond-switching for the [110] chains results in a structure apt for aggregation on a {113} plane. In contrast to previous calculations, however, the introduction of strain-related levels into the band-gaps is anticipated. Such activity is also likely to persist in the aggregated form at the periphery of the RLDs where the strain is less well accommodated by the relaxation of the surrounding lattice.

The 903 meV PL band and $E_v + 0.5$ eV DLTS peak are consistent with the properties of the chain ends and strain induced levels, respectively. The chains will contain no undercoordinated atoms in the positive charge state, where the chain end resembles the three-member-ring structure previously assigned to the W-line PL centre (Coomer *et al.*, 1999).

In line with preliminary calculations in Goss *et al.* (2002) and previous empirical calculations by Chou *et al.* (1995) it was found that the {111} RLDs are in fact significantly lower in energy than the more commonly observed {113} variety, and that the {001} oriented defect seen in HRTEM diamond and Ge is also competitive in terms of the formation energy of *infinite* systems. However, the elastic energy of a *finite* loop tends to lead to the formation of the {113} oriented planar defects, at least for modest numbers of self-interstitials. One should note, however, that although this is consistent with the observation, it is impossible to rule out other causes for the preferential uptake of interstitial aggregates in {113} planes, such as kinetic factors or pre-existing defects in the material.

Part II

Defects in $4H$-SiC

CHAPTER 5

PL Centres in 4H-SiC

5.1 Introduction

In contrast to silicon SiC is hard to dope by diffusion, instead ion-implantation is used to achieve selective area doping. The drawback of this technique is that irradiating the crystal with dopant atoms creates a great deal of lattice damage, like vacancies, interstitials, antisites and complexes. This becomes a problem if the created defects are electrically active, resulting in a reduction of the carrier lifetime. It is even possible that the irradiation reduces the electrically active dopant concentration. Many of the point defects can be eliminated through thermal annealing, there are however some which are stable to high temperatures, like the PL centre D_1.

This chapter is concerned with atomic models for antisites and self-interstitials and their connection with the most common PL centres in SiC introduced through irradiation. After a short introduction into polytypism of SiC the PL spectra are presented

5.1.1 Polytypes of SiC

SiC takes different crystal structures with the same chemical composition, so-called "polytypes", because of different stacking orders. More than 200 different polytypes are known in strong contrast to its constituents silicon and carbon. Three common polytypes of SiC are the cubic 3C and the hexagonal 4H and 6H. The difference between polytypes is the stacking of SiC-bilayers, i.e. {111} (cubic polytype) or {0001} planes (hexagonal polytypes). The 3C-polytype has a

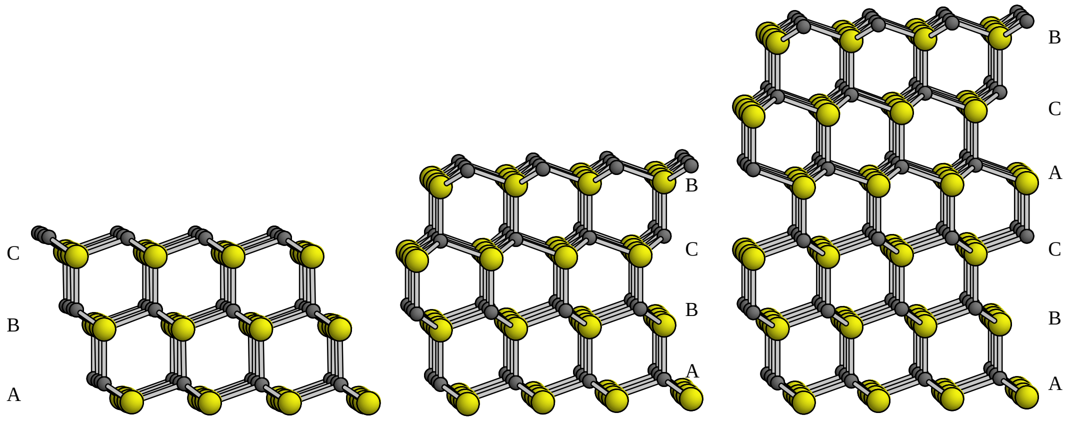


Figure 5.1: Three polytypes of SiC, 3C, 4H, and 6H, from left to right.

Table 5.1: Zero-phonon (ZP) energy (eV) and LVMs (cm^{-1}) for the P–U centres in 6*H*-SiC.

Centre	ZP energy	LVM$_1$	LVM$_2$
P–T	2.467–2.409	~1070	~1440
U	2.36	-	1989

zincblende structure, with stacking order ABC, like silicon or diamond. 4*H* has a ABCB stacking sequence and 6*H* has ABCACB (see Figure 5.1).

Locally every atom is fourfold coordinated with sp^3-bonded Si and C atoms. In the hexagonal polytypes the nearest neighbour (NN) surrounding of a lattice site is always the same, whereas there are two different surroundings for next nearest neighbours. One is identical to the surrounding of a cubic crystal, defining the (quasi-)cubic site or k site, and the other to a hexagonal crystal, defining the (quasi-)hexagonal site or h site. Needless to say that if third NNs are taken into account the lattice sites are neither cubic nor hexagonal.

5.2 PL Spectra

5.2.1 Alphabet lines

Room temperature electron-irradiation of 4*H*-SiC creates a series of around 40 sharp photoluminescence lines in a 0.1 eV interval around 2.8 eV (Egilsson *et al.*, 1999*b*). These are labelled $a_k, b_k, c_k \cdots l_k$, with increasing energy and $k \leq 4$. Selective photoexcitation spectroscopy demonstrates that the lines with different k are linked to excitations of the same defect and the whole set seems to come from about 12 centres. The lines are detected immediately after room temperature irradiation and disappear and/or transform among themselves or into a new band labelled D$_\text{I}$ around 1000–1300°C. A phonon replica of the b_1 no-phonon-line shows a gap vibrational mode at 629 cm^{-1} (78 meV) between the top of the acoustic branch at 610 cm^{-1} (75.6 meV) and bottom of the TO optic branch at 774 cm^{-1} (96 meV). Most of the Alphabet lines anneal out at about 900°C and the spectrum of the D$_\text{I}$ centre emerges. However some of the Alphabet lines only anneal at 1300°C.

Recent work on the Alphabet lines has indicated that the defects result from damage to the carbon sublattice as the PL lines are produced when the electron-irradiation energy lies below the threshold, ~250 keV, for damage on the Si sublattice (Steeds *et al.*, 2000). Moreover, some of the spectral lines are observed well outside the damaged region implying rapid defect mobility at room temperature under ionising conditions. This suggests that interstitials are involved in the formation of the defects.

The carbon interstitial (C$_i$) in both bulk Si (Watkins and Brower, 1976; Leary *et al.*, 1998) and diamond (Hunt *et al.*, 2000; Goss *et al.*, 2001) takes the form of a $\langle 100 \rangle$ split-interstitial where a pair of atoms share a substitutional site and are aligned along a cube edge. The self-interstitial in both materials is known to be mobile under ionising conditions (Watkins, 1964; Twitchen *et al.*, 1999).

5.2.2 D$_\text{I}$ line

Even though D$_\text{I}$ has been detected more than 30 years ago (Patrick and Choyke, 1972) and is probably the best studied luminescence centre in SiC, its structure is still unknown. This PL centre is thermally stable to high temperatures (above 1700°C) (Dalibor *et al.*, 1996; Egilsson *et al.*, 1999*a*). In 4*H*-SiC its zero phonon line is around 2.901 eV. The centre is observed in all polytypes after irradiation and is also observed in as-grown material after quenching from growth temperature (Haberstroh *et al.*, 1994; Kennedy *et al.*, 1990). Its appearance in irradiated and as-grown material suggests that the defect is intrinsic. The centre can be seen in n-type as well as p-type material (Egilsson *et al.*, 1999*a*). Recent studies (Storasta *et al.*, 2001) show that the D$_\text{I}$ luminescence arises from a bound exciton with a deep hole at $E_v + 0.34$ eV. In addition to the sharp zero phonon line two phonon gap modes around 83 meV (669.4 cm^{-1}) have been detected.

5.2.3 P–U lines in 6*H*

In 6*H* the carbon interstitial has been identified by LVM spectroscopy making use of C isotopic shifts (Evans *et al.*, 2002). Six carbon-related PL centres, called P, Q, R, S, T and U with decreasing zero-phonon line energy from 2.47 to 2.36 eV were detected. P to T have very similar LVMs – each centre has two modes – while U has only one mode that is higher in frequency (see Table 5.1). It has been speculated that it may arise from a Si–H bond, due to its high frequency.

For the experiment ^{13}C enriched 6*H*-SiC samples were irradiated by low-energy electrons with energies below the silicon displacement threshold. Annealing measurements showed that the P–T lines disappeared after 900°C in p-type but persist

Table 5.2: Formation energies (eV) of antisites in stoichiometric 4*H*-SiC. The values under Torpo are from Torpo *et al.* (2001, 1998); Torpo and Nieminen (1999).

Defect	ddpp	dddd	Torpo
C_{Si}	3.60	3.93	~3.5
Si_C	4.15	4.11	~4.2
C_{Si}-Si_C	5.41	5.61	5.8 ($3C$)

up to 1300°C in n-type samples. The intensity of the U line increased in both n- and p-type SiC after annealing.The $^{12}C -^{12} C/^{13}C -^{13} C$ ratio of the LVMs of the P–T lines were 1.033 and 1.04 for LVM_1 and LVM_2, respectively. The splitting of the LVMs due to the isotopes was 1:1 for the lower frequency mode and 1:2:1 for the higher.

5.3 Calculation Method

To investigate the antisite defects C_{Si}, Si_C, and close-by antisite pairs followed by interstitial defects C_i and Si_i in the following section, 72 and 128 atom supercells were used. The supercell was made up from 4x4x1 hexagonal unit cells, i.e. the unit cell was repeated 4x4 times along two basis vectors in the basal plane and once along the c axis. As in Eqn. 3.1 the formation energy (E_f) of a defect is defined as:

$$E^f = E_{tot} - n_C\mu_C - n_{Si}\mu_{Si} - qE_F, \tag{5.1}$$

where n_C and n_{Si} is the number of carbon and silicon atoms in the supercell respectively. Here μ_C and μ_{Si} are the chemical potentials of C and Si in SiC. The chemical potentials satisfy the conditions that they are bounded above by the energies of a C or Si atom in standard thermodynamic forms like pure diamond or Si, and $\mu_C + \mu_{Si} = \mu_{SiC}$, the energy of a Si-C pair in bulk 4*H*-SiC. The heat of formation (H_f) is the difference between the energies of a Si and a C atom, in their standard states, and a Si-C pair in 4*H*-SiC, i.e.

$$H_f = \mu_C^{bulk} + \mu_{Si}^{bulk} - \mu_{SiC}. \tag{5.2}$$

The calculated values of 0.59 eV for the ddpp basis and 0.715 eV for the dddd basis are in good agreement with earlier studies (Zywietz *et al.*, 1999; Torpo *et al.*, 2001) and an experimental value of 0.72 eV (Kubaschewski and Alcock, 1979). The chemical potentials are then

$$\mu_C = \mu_C^{bulk} - \frac{1}{2}H_f - \frac{1}{2}\Delta\mu, \tag{5.3}$$

$$\mu_{Si} = \mu_{Si}^{bulk} - \frac{1}{2}H_f + \frac{1}{2}\Delta\mu \tag{5.4}$$

with $\Delta\mu$ being $-H_f$, 0, and H_f for C-rich, stoichiometric and Si-rich material, respectively.

Vibrational modes were found by computing the energy second derivatives with respect to the displacement of the defect atoms and their neighbours. The dynamical matrix of the supercell is then found using these values together with those found from a Musgrave–Pople potential for the other atoms in the supercell (Jones and Briddon, 1998). The computed phonon gap is between 625 cm^{-1} (77.5 meV) and 765 cm^{-1} (94.8 meV) compared with experimental values of 610 and 774 cm^{-1}.

The calculated a and c cell dimensions of 4*H*-SiC are 3.054 and 9.999 Å compared with experimental values of 3.078 and 10.086 Å respectively (Madelung and Schulz, 1989). The bulk modulus was found to be 227 GPa for 4*H*-SiC in agreement with an experimental value of 224 GPa for 3*C*-SiC (Madelung and Schulz, 1989). When the basis for both Si and C was increased to d orbitals associated with each of the four exponents, the Si-C bond length decreased by only 0.5% and the bulk modulus increases by 0.4%.

5.4 Results

5.4.1 Antisites and antisite-pair

Due to the two inequivalent C and Si substitutional sites in 4*H*-SiC there are four possible antisite pairs separated by a single bond and a further four pairs at second neighbour separation although these involve the same atomic species such as Si_C–Si_C. Third neighbour defects involve C_{Si} and Si_C pairs. There are 6 of these and 6 at fifth neighbour separation. Hence the number of distinct pairs at third and fifth neighbour separation is sufficient to explain the 12 defects attributed

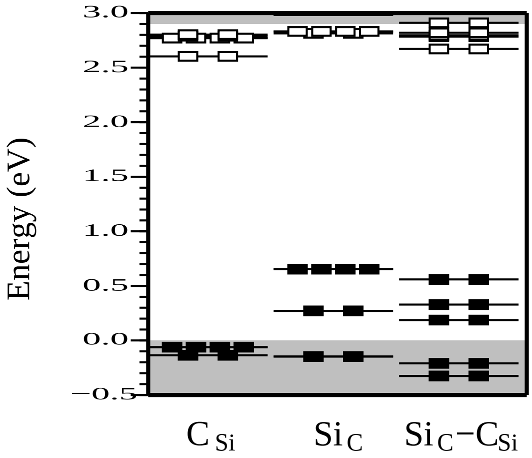


Figure 5.2: The Kohn-Sham levels at Γ of antisite defects in 4*H*-SiC evaluated in 72 atom supercells. Occupied (empty) levels are denoted by filled (empty) boxes. Grey regions correspond to bulk band positions at the zone-centre. The calculated band gap for 4*H*-SiC in this cell at Γ is 2.9 eV. Note C_{Si} has no gap levels whereas Si_C has a deep donor level which becomes shallower for the Si_C–C_{Si} pair.

to the Alphabet lines in 4*H*-SiC. Early calculations (Wang *et al.*, 1988) found Si_C to be electrically inactive in 3*C*-SiC but later work (Torpo *et al.*, 2001) suggests electrical activity and places the donor level at $E_v + 0.21$ eV in 4*H*-SiC.

The two inequivalent C in 4*H*-SiC lead to two possible C_{Si} and likewise two possible Si_C all of which possess trigonal symmetry. Each pair of antisites is found to be degenerate in energy. In the 72 atom cell the formation energies are 3.6 and 4.2 eV for C_{Si} and Si_C, respectively. Allowing an isotropic relaxation of the supercell volume or increasing the supercell to 128 atoms decreases the formation energy by 0.1 eV for both defects. These energies are sufficiently high that the defects would not normally be present in equilibrium. Similar energies were also found with a larger basis of four *d* orbitals on each atom. The energy of Si_C dropped by only 0.04 eV while that of C_{Si} increased by 0.33 eV. These changes are quite small, less than 10% of the formation energies, and the calculations are therefore believed to be converged with respect to the basis.

Fig. 5.2 shows the Kohn-Sham levels at Γ in a 72 atom supercell containing antisite defects. The electron bands of the bulk 4*H*-SiC were calculated in a supercell of the same size as the defect cells and are shown in grey. The calculated band gap in 4*H*-SiC is 2.18 eV compared with an experimental value of 3.26 eV. The underestimation of the gap is a well-known failing of density functional theory and should not affect the energies of neutral defects. C_{Si} is an electrically inert defect with C–C bond lengths of 1.660 and 1.664 Å (|| to **c**) at a *h* site or 1.659 and 1.645 Å (|| to **c**) at a *k* site, compared with 1.54 Å in diamond. Si_C at a *k* site has Si–Si bond lengths of 2.120 and 2.131 Å (|| to **c**) compared with 2.33 Å in bulk Si. At a *h* site the bond lengths are 2.116 and 2.143 Å (|| to **c**). For the neutral Si_C defect, the Kohn-Sham levels at the zone-centre indicate a fully-occupied doublet ~0.65 eV above E_v or, using the experimental value for the band gap, 2.61 eV below the conduction band minimum. An alternative way of estimating the (0/+) level is to compare the formation energies of the defect in the neutral and positive charge states with the energies of the neutral and charged host. This gives a level 2.73 eV below E_c for Si_C^k and 2.72 eV for Si_C^h. In the 128 atom supercell the level lowers to $E_c - 2.94$ eV. It is not possible to find accurate energy levels using density functional theory with periodic boundary conditions but the calculated ones are qualitatively consistent with a deep donor level around $E_c - 2.9$ eV associated with the Alphabet series, if an exciton binding energy of ~0.1 eV is assumed.

This suggest that a perturbed form of Si_C is responsible for the Alphabet series.

The wavefunction of the Si_C donor level is composed of *p*-orbitals with Mulliken bond populations giving 28% localisation on the unique Si atom. The Si_C defect has almost tetrahedral symmetry and is associated with a triplet mode (at 665 cm^{-1} in 3*C*) which is split into a doublet at 657.0 (*E* symmetry) and singlet at 655.4 cm^{-1} (A_1 sym.). Only the latter mode would be expected to accompany a PL transition and is in good agreement with the observed replica at 629 cm^{-1} (78 meV). The mode shifts by only 0.8 cm^{-1} with ^{13}C consistent with its origin as a compressed Si–Si bond and agrees with recent investigations showing that the phonon replica of b_1 hardly shifts in ^{13}C material (Steeds, 2001). In 3*C* the Si_C has also an A_1 mode close to the continuum, which disappears in 4*H*.

At the moment there is no EPR evidence for Si_C. Son *et al.* (2001) initially identified the EI6 EPR-centre with Si_C^+ in 4*H*. However later papers disputed the assignment (Konovalov *et al.*, 2003; Bockstedte *et al.*, 2003*a*; Bratus *et al.*, 2001; von Bardeleben *et al.*, 2002) and suggested V_C^+ at a *h* site to be responsible for the EPR signal.

Among the antisite pairs, the nearest neighbour pair is unique as it alone has one Si–C bond within the core. There are four inequivalent versions. Two are trigonal centres and two possess C_{1h} symmetry. In the 72 atom cell the energies of the trigonal defects are about 0.3 eV higher than the C_{1h} ones which differ by only 0.07 eV. In the 128 atom cell the trigonal defect is still 0.2 eV higher in contrast to other DFT calculations in 96 atom supercells who find no significant

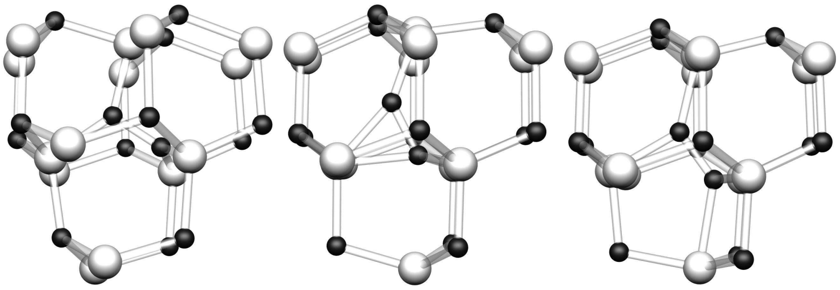


Figure 5.3: The neutral C interstitial in the C_i^{hC}, C_i^{kC-A} and C_i^{kC-B} structure from left to right. The C atoms are shown in black and the Si atoms in white.

energy difference (Gali *et al.*, 2003). The binding energy of the pair is ∼2.3 eV for both cell sizes. The (0/+) level of the antisite pair at the h site is calculated to be 2.86 eV and 3.00 eV below the conduction band in the 72 and 128 atom cell, respectively. Both times this is deeper than Si_C, which suggests that the pair may be responsible for the most energetic transition (the l-band at 2.8998 eV) or possibly the 2.901 eV (427.26 nm) D_I band.

The antisite pair gives a large number of gap vibrational modes but two Si related ones (with A_1 sym.) at 625 and at 647 cm^{-1} lie close to the observed replicas around 669.4 cm^{-1} (83 meV) of D_I. Consistent with experiment (Sadowski *et al.*, 2000), where a shift with ^{13}C is not detected, the calculated shift of the first mode is only 0.4 cm^{-1}. The second mode however shifts by 7 cm^{-1}.

5.4.2 Self-interstitials

C interstitial

The lowest energy structure for the C interstitial is a split form as for similar defects in Si or diamond. Out of the 2 inequivalent C and 2 inequivalent Si sites the C sites are lower in energy. The C interstitial at a h C-site (C_i^{hC}) has an e level in the middle of the bandgap with 2 electrons in the neutral charge state. An analysis to whether it is a spin singlet (S=0) or a triplet (S=1) defect supports the triplet state. In the 72 atom supercell the S=1 defect is only 0.02 eV lower than the S=0 one, in the 128 atom cell this difference increased to 0.17 eV. However, a many-body singlet state which is also possible cannot be examined with the theory used.

At the k C-site two configurations with only a small structural difference were found, C_i^{kC-B} and C_i^{kC-A}. Both are S=0 singlet states, the triplet states are ∼0.14 eV higher in energy in the 128 atom cell. In the 72 atom cell C_i^{kC-A} is 0.14 eV lower in energy than C_i^{kC-B} and in the 128 atom cell 0.08 eV. Under stoichiometric conditions the lowest energy structure has a formation energy of 7.0 eV and 6.9 eV in 72 and 128 atom cells, respectively.

For both sites, k and h, there is also the possibility that the C interstitial takes a [110]-split form (in cubic notation) like the silicon self-interstitial (see Chapter 3). Both structures, C_i^{hC-110} and C_i^{kC-110}, are roughly 0.25 eV higher in energy than the ground state structure.

There is also the possibility to form a dumbbell with a Si atom, these forms (see Fig. 5.4) are 0.62 and 0.83 eV higher in energy than the ground state for k and h sites, respectively. Furthermore they only have a tiny barrier and are effectively saddle points for a diffusion in the basal plane. The diffusion barrier for the C interstitial in the neutral charge state along

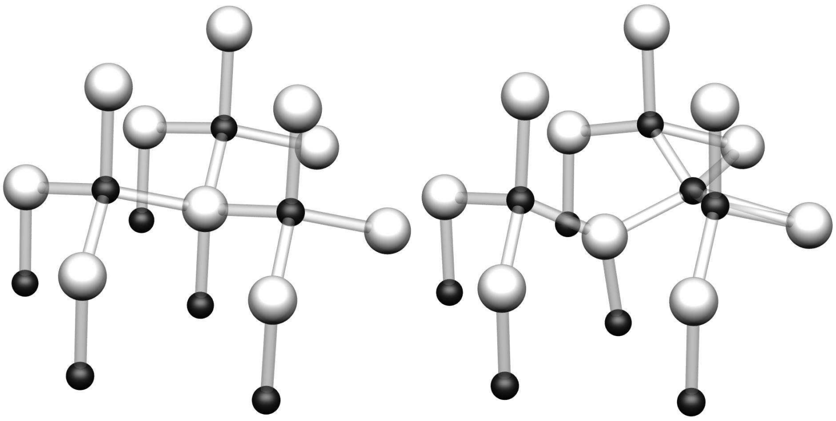


Figure 5.4: The neutral C_i defect trapped near a k Si site. The left figure shows bulk 4H-SiC for comparison. This structure is a saddle point in the diffusion path of C_i^{kC-A}

Table 5.3: Relative energies (eV) of neutral defects in 4*H*-SiC. The donor and acceptor levels above the valence band (eV) are calculated in the 72 atom cell with the bulk as marker.

Defect	E_{rel}^{72}	E_{rel}^{128}	(0/+)	(-/0)
C_i^{kC-A}	0.00	0.00	1.5	2.8
C_i^{kC-B}	0.14	0.08	1.6	2.8
C_i^{hC}	0.18	0.05	1.6	2.9
C_i^{hC-100}	0.23			
C_i^{kC-100}	0.27			
C_i^{kSi}	0.62	0.64		
C_i^{hSi}	0.83			

a $C_i^{kC-A} \rightarrow C_i^{kSi} \rightarrow C_i^{kC-A}$ path is subsequently only ~0.6 eV. In the positive charge state the C_i^{kSi} structure becomes 1.0 eV higher in energy than the C_i^{kC-A} form. This is in agreement with a calculated diffusion barrier of 0.9 eV for C_i^+ in 3*C*-SiC by Bockstedte *et al.* (2002), assuming the C_i^{kSi} structure is still a saddle point in the positive charge state. The relative energies for different structures can be found in Table 5.3. There is also experimental evidence for interstitial movement at low temperatures in 4*H*-SiC. von Bardeleben *et al.* (2000) observed only 30% of the expected vacancy concentration in proton irradiated 4*H*-SiC and concluded that a vacancy annihilation process has to take place already at room-temperature. Since the vacancies are immobile at this temperature interstitials are the obvious candidates for this process.

Table 5.4 shows the LVMs of the C interstitial. The wavenumbers, isotope shifts and splittings of the hexagonal C interstitial, C_i^{hC}, are in excellent agreement with the measured values for the P–T centres. However the calculated diffusion barrier of 0.6 eV is in strong contrast to the annealing temperatures (900°C in n-type and 1300°C in p-type) of the P–T lines. The high frequency LVMs of the other two configurations for C_i are also in agreement with experiment, however the low frequency modes show considerable differences. The calculated acceptor levels at $E_v + 2.8$ eV are in fair agreement with the zero-phonon energy of 2.4 eV measured in 6*H*, especially if the levels shift as the band off-set of the conduction bands of 4*H* and 6*H*.

From this results it can be concluded that the experiment measured LVMs of a defect involving C–C split configurations, the high annealing temperatures of the defect suggest though that these C interstitials are part of a bigger defect with higher thermal stability (like $(C\text{–}C)_{Si}$).

Si interstitial

Turning to the silicon interstitial, a bistable defect. Several different configurations both at the *k* and *h* site were considered, including ⟨100⟩ and ⟨110⟩ like dumbbell structures and caged configurations. The lowest energy configuration for the neutral defect is one where Si_i builds a dumbbell with a bulk Si atom at a *h* site. Each Si atom of the dumbbell has 3 C neighbours. Locally this structure is identical to the ⟨110⟩-split interstitial in Si (see chapter 3). The first Si dumbbell atom has one C neighbour at 1.76 Å and two at 1.9 Å while the second has one C neighbour at 1.76 and two at 2.0 Å. The two Si atoms of the dumbbell are 2.14 Å apart and the angle between the **c**-axis and the dumbbell is 38°. At the *k* site this angle is 41° and the separation of the two dumbbell atoms is 2.06 Å. Only one of the Si atoms has 3 C neighbours the other one has only 2. This configuration is almost 1 eV higher in energy than the *h* site. This can be attributed to the fact that there are different nearest neighbours for the interstitial in the two different sites. It is also possible to align the dumbbell along a lattice direction perpendicular to the *c*-axis. This structure is, for both *h* and *k* sites, around 0.6 eV higher in energy than the lowest energy structure. This means that at the *k* site the dumbbell structure aligned perpendicular to the c-axis is 0.4 eV lower while at the *h* site it is 0.6 eV higher in energy than the corresponding structures not aligned along lattice vectors.

Due to the higher energy at *k* sites only *h* site defects will be discussed in the following.

The formation energy of the lowest energy structure at the *h* site is 9.3 eV in stoichiometric material. This decreases to 9.0 eV for the dddd basis. This is in agreement with another DFT calculation in 3*C*-SiC finding a formation energy of

Table 5.4: LVMs (cm^{-1}) of the C interstitial. Under ^{13}C shift the ratio of the ^{12}C to ^{13}C dumbbell LVM is reported as well as the line splitting in parentheses.

Defect	LVM_1	^{13}C shift	LVM_2	^{13}C shift
C_i^{kC-A}	923	1.02 (1:1:1:1)	1470	1.04 (1:2:1)
C_i^{kC-B}	968	1.03 (1:1)	1428	1.04 (1:2:1)
C_i^{hC}	1070	1.03 (1:1)	1485	1.04 (1:2:1)

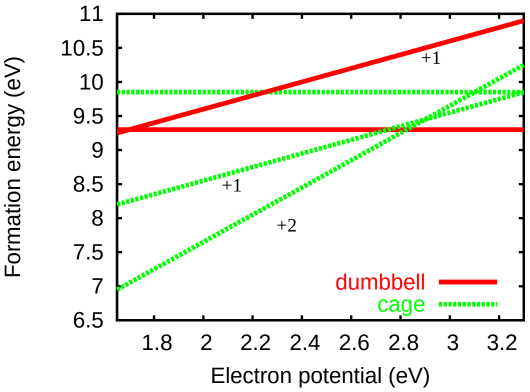


Figure 5.5: Formation energies (eV) of Si_i at *h* sites for the caged and dumbbell structure.

8.9 eV (Bockstedte *et al.*, 2003*b*). In the positive charge state, Si_i (at the *h* site) is located at a cage site surrounded by three C atoms at a distance ~1.85 Å. The dumbbell configuration is 0.47 eV higher in energy in this charge state, but in the neutral charge state it is 0.55 eV lower than the cage site. Whereas in the +1 charge state there is still a barrier between the cage and dumbbell site, in the +2 charge state the dumbbell interstitial relaxes immediately to a cage site. The Si interstitial at the *h* site has no acceptor level. In the +2 charge state the Si_i at the *k* site is still found to be higher in energy – by 1 eV – than at the *h* site. A look at the Kohn-Sham eigenvalues shows that the caged Si_i^{2+} has a mid-gap level occupied by two electrons. This means that further charge states – +3 and possibly +4 – are possible. However due to the uncertainties of energies and therefore electrical levels of highly charged supercells they are not considered here. Figure 5.5 shows the formation energies of the two different structures for different charge states in the upper part of the band-gap. The (+/0) level has been calculated with nitrogen as marker (see chapter 6 for explanations) and is subsequently found to be resonant with the conduction band. If compared with Fig. 3.7 one can easily see the similarity in behaviour of Si_i in Si and SiC. In the positive charge states it preferes a caged configuration while in the neutral it forms a dumbbell with overcoordinated Si atoms. Table 5.5 shows the electrical levels of dumbbell and cage structure provided a transformation between the two structures does not take place on the timescale of an electronic transition. Including the transformation between them the single donor is found around $E_v + 2.7$ eV and the double donor at $E_v + 2.9$ eV. This means the Si_i shows a negative-*U* behaviour with a (++/0) level at $\sim E_v + 2.8$ eV. However within the error of the method the double donor could also be below the single donor and the defect would be a positive-*U* one with close-by single and double donor.

5.4.3 Production mechanism for antisites

In this section a mechanism for the production of antisites during irradiation is discussed. The first stage involves the creation of a C vacancy and C_i pair. Since the defects were seen far outside the irradiated area and interstitials are mobile an interstitial mechanism will be discussed.

If C_i moves to a Si site, as shown in Fig. 5.6, then the reaction $C_i^{kSi} \rightarrow C_{Si} + Si_i$ is endothermic by 5.9 eV and will not occur. However, if the defect now traps two holes then the reaction $C_i^{kSi} + 2h \rightarrow C_{Si} + Si_i^{++}$ is endothermic by only 0.2 eV. If Si_i can trap more holes and become even more positively charged the reaction can easily become exothermic. This reaction results in the formation of C_{Si} and a silicon interstitial. Suppose that Si_i^{++} rapidly neutralises by trapping two electrons then it is able to move to a dumbbell configuration. In turn, the resulting Si_i defect can itself lead to the production of an antisite. $Si_i \rightarrow Si_C + C_i^C$ is endothermic by 1.9 eV but $Si_i + h \rightarrow Si_C + (C_i^C)^+$ is endothermic by only 0.4 eV. This strongly suggests that a supply of holes and electrons makes the formation of antisite pairs likely. However, here it is assumed that the barriers to the reactions and the pre-exponential factors in the rates are not important.

Table 5.5: Electrical levels (eV) of Si_i (at *h* sites) above the valence band, calculated in a 72 atom cell with the bulk as marker. The level in parentheses is found to coincide with the conduction band if calculated with N as marker. A transformation between the two structures is not taken into account.

Defect	(0/+)	(+/++)
dumbbell	1.7	
cage	(2.8)	~2.9

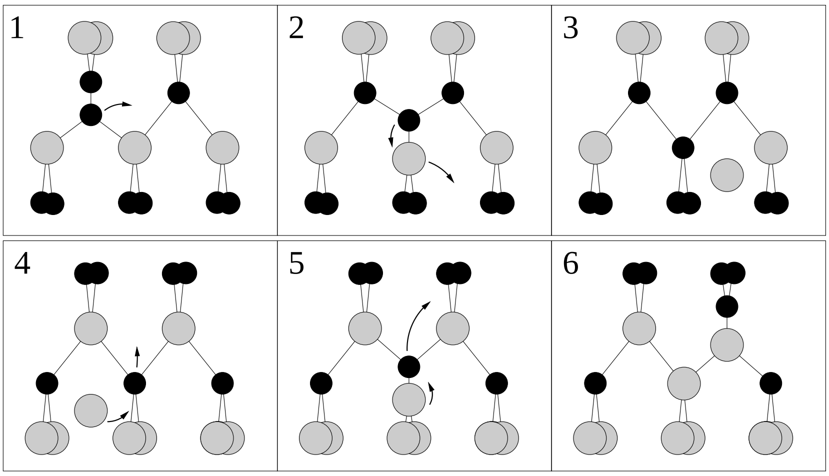


Figure 5.6: Schematic diagram of the processes leading to antisite pair formation. A carbon interstitial at a C site (1) moves to a neighbouring Si site (2) whereupon a Si interstitial defect (Si_i^{++}) is created along with $\mathrm{C_{Si}}$ (3). The cage sited Si interstitial (4) neutralises and forms a Si split-interstitial (5). A carbon atom is then ejected leading to the formation of a carbon interstitial and $\mathrm{Si_C}$ (6).

5.5 Conclusions

The antisite pairs possess properties in common with the centres giving rise to the Alphabet series of luminescent lines. The transition energy, local vibrational mode and its isotopic shift, thermal stability and the large number of defects are all consistent with experiment. Hole trapping can explain their formation in irradiated material. The most stable antisite pair may be responsible for the most energetic Alphabet line or the $\mathrm{D_I}$ band, which survives to higher temperatures.

Thus the Alphabet bands can be assigned to antisite pairs with the zero–phonon line of the exciton bound to $\mathrm{Si_C}$ being perturbed by a nearby $\mathrm{C_{Si}}$. These pairs are expected to be thermally stable and survive anneals to 1300°C. Nevertheless, the widely different annealing behaviour and the energy threshold for the creation of the individual lines suggests that some may be $\mathrm{Si_C}$ perturbed by other defects.

The low diffusion barrier for the C interstitial which enables diffusion at room temperature seems to have prevented its direct detection so far. It is unclear whether the C interstitial at a *h* site is a spin singlet or triplet, LSDA calculations support the latter. The P–U lines seem to arise from defects *involving* the C interstitial in a C–C split form, but these defects must be larger then the isolated interstitial to explain their thermal stability.

A summary of the electrical levels of defects in this chapter is shown in Fig. 5.7. The Si interstitial is a negative-U centre with a (++/0) transition ∼0.5 eV below the conduction band or has two close-by single and double donor levels. From the KS levels it looks as if it has a +4 charge state for a Fermi level close to the valence band.

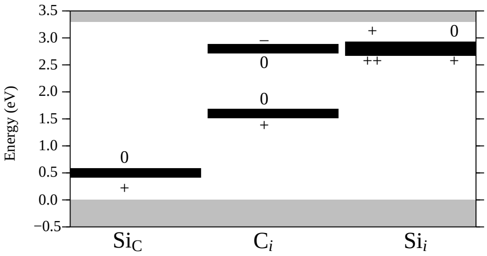


Figure 5.7: Positions of electrical levels (eV) in the band-gap for the defects investigated in this chapter. The Si interstitial either has close-by single and double donor or a negative-U (++/0) level.

CHAPTER 6

The Z_1/Z_2 Defects in $4H$-SiC

6.1 Introduction

Rapid growth of SiC by CVD methods leads to a deterioration in quality with an increase in electrically active grown in defects (Fujihira *et al.*, 2002; Pintilie *et al.*, 2002). Among these, the Z_1 and Z_2 defects (denoted by Z_1/Z_2) are dominant in n-type $4H$-SiC, as well as material that has been exposed to radiation (Dalibor *et al.*, 1996, 1997*a*; Hemmingsson *et al.*, 1997, 1998).

6.2 Experimental Results

Z_1/Z_2 are a pair of defects with closely spaced electrical levels that possess unusual thermal stability – surviving to 1700°C (Storasta *et al.*, 2001). In other polytypes like $6H$-SiC, the E_1 and E_2 defects have similar properties (Aboelfotoh and Doyle, 1999; Hemmingsson *et al.*, 1999; Gong *et al.*, 1999). These four centres are negative-U (Anderson, 1975) having inverted donor (0/+) and acceptor (-/0) levels. The donor and acceptor levels of Z_1 (Z_2) in $4H$-SiC occur at $E_c - 0.43$ (0.46) and $E_c - 0.67$ (0.71) eV, respectively, while the corresponding levels of E_1 and E_2 in $6H$-SiC are shallower by $\sim$0.2–0.3 eV.

In n-type as-grown $4H$ CVD epilayers Z_1/Z_2 is the only observable defect (Kimoto *et al.*, 1995) and is stable up to 2000°C (Dalibor *et al.*, 1997*b*). Ion implantation increases the defect concentration but also broadens the Z_1/Z_2 peak, indicating that another defect is created with a level overlapping with the one of Z_1/Z_2. This new defect disappears after annealing at around 1000°C (Dalibor *et al.*, 1997*b*; Hemmingsson *et al.*, 1998). One peculiar behaviour of Z_1/Z_2 is its strong decrease in concentration at 1700°C to the as-grown value which is stable to beyond 2000°C (Hemmingsson *et al.*, 1998; Dalibor *et al.*, 1997*b*). This suggests that the defect is not actually breaking up at 1700°C but another defect is becoming mobile at this temperature and is interacting with Z_1/Z_2.

There has been much speculation about their origin. They have been linked with the D_I photoluminescence band with zero phonon line at 2.901 eV in $4H$-SiC which also is thermally stable to high temperatures (see Section 5). However, the deep hole at $E_v + 0.34$ eV of D_I rules out a connection with Z_1/Z_2 having levels close to E_c.

More recently, Z_1/Z_2 has been linked with a silicon vacancy V_{Si} (Kawasuso *et al.*, 2001*b*). This connection was suggested when it was shown that the annealing behaviour of vacancy centres observed by positron annihilation is correlated with the annealing of Z_1/Z_2 in $4H$-SiC (Kawasuso *et al.*, 2001*b*) and E_1/E_2 in $6H$-SiC (Kawasuso *et al.*, 2001*a*). However, in both cases the maximum in the concentration of DLTS centres was not directly linked to that of the positron annihilation signal. This allows the alternative interpretation that the DLTS centres and vacancies annealed together by mutual annihilation.

Recent work has shown that the Z_1/Z_2 defect in as-grown material increases in concentration both with nitrogen doping and when growth proceeds under carbon rich growth conditions (Pintilie *et al.*, 2002). However more recently Storasta *et al.* (2004) found Z_1/Z_2 in phosphorus doped SiC at concentrations up to 10^{16} cm^{-3} while the N contamination was below 10^{15} cm^{-3}. The presence of a single N atom in the defect implies that the neutral defect would be paramagnetic in its neutral charge state and that the negative-U behaviour arises from a different structural form for the positive and negative charged defects. This is similar to many other negative-U defects such as interstitial boron (Troxell and Watkins, 1980) and hydrogen (Herring *et al.*, 2001) in Si, and the DX-centre in AlGaAs (Chadi and Chang, 1988).

The alternative possibility is a defect with an *even* number of electrons when neutral. There are very few examples of such centres which display negative-U properties. One example is the vacancy in Si whose (+/++) level lies above (0/+) (Watkins and Troxell, 1980). However, the experimental cross-sections of the Z_1/Z_2 centres rule out an assignment of Z_1/Z_2 to levels of a double donor and thus the most likely model for Z_1/Z_2 consists of a complex of a single N atom together with an intrinsic centre such as a carbon interstitial or silicon vacancy. The concentrations of both would then be expected to increase in C-rich material.

6.3 Method

The calculations in this section were carried out in large 128 atom unit-cells with the same shape as in the previous chapter. The Brillouin sampling was done using a Monkhorst-Pack (MP) 2^3 grid. Convergence tests were carried out using MP-4^3 sampling for some structures. Differences in total energy between the two schemes were generally less than 0.07 eV.

For Si and C atoms a ddpp basis has been used while for N a big dddd basis was employed. The chemical potentials for Si and C were taken as described in the previous chapter while the chemical potential for nitrogen was taken from the energy of a N atom in β-Si_3N_4. For this it was assumed that the chemical potential for Si is the same in SiC and Si_3N_4. This choice gives the formation energy of the nitrogen donor, N_C, in stoichiometric SiC in the neutral charge state to be 2.5 eV. If the chemical potential for nitrogen is taken from a nitrogen molecule N_2, $E^f(N_C)$ is around 0 eV. This is probably due to the bad description of small molecules in DFT.

Comparing the ionisation energies of N_C (Woodbury and Ludwig, 1961) and P_{Si} (Greulich-Weber *et al.*, 1995) at hexagonal sites with the ionisation energy of bulk 4*H*-SiC gave (0/+) levels at $E_c - 0.52$ eV and $E_c - 0.54$ eV, respectively. These require upward shifts of ~ 0.5 eV to bring them into agreement with experimental values of ~ 0.05 eV (Gotz *et al.*, 1993; Capano *et al.*, 2000). The donor levels of defects in the upper part of the gap and given below have therefore been shifted upwards by 0.5 eV, i.e. N has been chosen as marker for donor levels in the upper half of the bandgap. This shift is generally too much for levels just above mid-gap, i.e. the further the defect level is away from the marker the worse the method becomes.

6.4 Results

6.4.1 Substitutional nitrogen and C_{Si}

The first defect investigated is a nitrogen impurity at a Si site denoted by N_{Si}. Its concentration is expected to be enhanced under carbon rich conditions and the centre may be a negative-U one in analogy with the Si donor in AlGaAs which forms the well-known DX-centre. However, N_{Si} is found to be 5.9 eV higher in energy (in stoichiometric material) than N_C and shows no negative-U behaviour, furthermore it is less stable by 3.4 eV when N is interchanged with one of its C neighbours, resulting in a N_C-C_{Si} pair. This structure has a low formation energy of 5.0 eV and (0/+) and (-/0) levels at $E_c - 1.7$ and $E_c - 0.6$ eV, respectively. Thus this defect is also a positive-U one with deep levels and can be excluded as a model for Z_1/Z_2. The binding energy of the pair is 1.2 eV.

6.4.2 Silicon vacancy and nitrogen

Next it is investigated whether the Si vacancy, V_{Si}, can be discounted as a candidate for Z_1/Z_2. The calculations are restrict to V_{Si} at a h site since previous theory has shown that the different lattice sites give quite similar energies and level positions in the gap. In agreement with this theory (Torpo *et al.*, 2001, 1999) the V_{Si} is found to be stable with spin $S = 1$ when neutral, also consistent with experiment (von Bardeleben *et al.*, 2000), and $S = 3/2$ when negatively charged. The $S = 0$ neutral defect and the $S = 1/2$ negatively charged centres are 0.13 and 0.22 eV, respectively, higher in energy. The local symmetry, i.e. the symmetry of a tetrahedron made up from the four neighbours of the vacancy, is T_d in the neutral and negative charge state.

A Mulliken bond-population analysis shows that the spin density of the neutral defect is mainly composed of p-orbitals on the four C atoms forming the vacancy cage with little s-contamination. Each nearest neighbour C atom is found to have 13% of the total spin density compared with an experimental estimate of 12.8±4% (Wagner *et al.*, 2002). Its formation

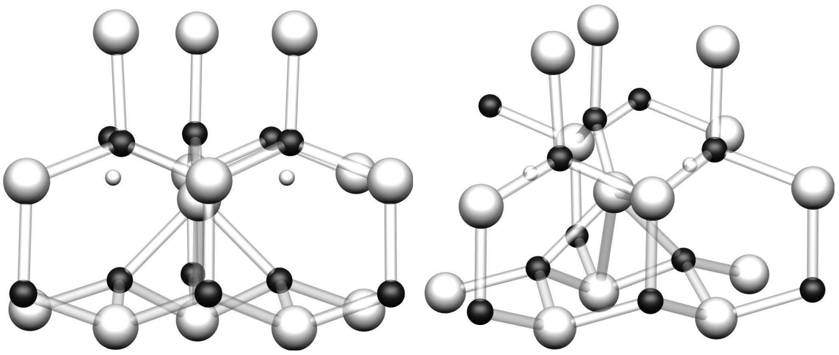


Figure 6.1: The saddle-point configuration for the $V_{\rm Si}$ diffusion. C (Si) atoms are shown in black (white). The isolated small white spheres indicate the starting and final position of the diffusion Si atom.

energy is 8.4 eV. It is found, also in agreement with previous work (Zywietz *et al.*, 1999), that it possesses a deep (-/0) level at $E_v + 1.1$ eV. The defect is in fact less stable than interchanging the vacancy and a C neighbour to form $V_{\rm C}$-$C_{\rm Si}$ (Rauls *et al.*, 2000). The resulting centre is 1.6 eV lower in energy and possesses shallower (-/0) and (0/+) levels at $E_c - 0.7$ and $E_c - 1.0$ eV (or $E_c - 1.5$ eV if calculated with the bulk as marker), respectively. The instability of $V_{\rm Si}$ may be connected with the disappearance of the EPR signal attributed to it above 750°C (Lingner *et al.*, 2001). The diffusion barrier and the saddle point (see Fig. 6.1) were found using the 'Dimer Hybrid' method (Papagiannidis, 2003). This gave a barrier of around 3.8 eV for the neutral defect, corresponding to an annealing temperature of ~1200°C. This supports the idea that the disappearance at 750°C is related to the transformation into a $V_{\rm C}$-$C_{\rm Si}$ pair rather then due to the onset of vacancy movement. It can however not be excluded that another defect becoming mobile at 750°C interacts with the vacancy.

$N_{\rm C}$ binds to $V_{\rm Si}$ with an energy of 3.3 eV and the resulting defect possesses (-/0) and (0/+) levels at $E_v + 1.5$ eV and around E_v. The structure with the N atom along the c axis is slightly lower in energy (≤ 0.1 eV) than the structure with the N atom and the vacancy in the basal plane. The N atom stabilises $V_{\rm Si}$ suppressing the transformation of the latter into $V_{\rm C}$-$C_{\rm Si}$ as this defect, with a neighbouring $N_{\rm C}$, is 0.3 eV higher in energy than $V_{\rm Si}$-$N_{\rm C}$. The defect where $N_{\rm C}$ neighbours $V_{\rm C}$-$C_{\rm Si}$ has acceptor and donor levels at $E_c - 0.9$ and $E_c - 1.9$ eV. None of these levels are close to those of Z_1/Z_2 and a connection between the centres can be ruled out. The same is true for the electrical levels of the divacancy as can be seen in Torpo *et al.* (2002).

6.4.3 C interstitial with N

Next models based on C_i are considered. Due to the low barrier C_i is expected to diffuse easily until trapped by a majority defect such as nitrogen. The $N_{\rm C}$-C_i pair, which can also be described as a nitrogen interstitial N_i, is also a split-interstitial pair where both N and C are located at a carbon site. The C–N bond length are 1.33, 1.37 and 1.42 Åfor the possitive, neutral and negativly charged defect, respectively. The formation energy of the neutral defect is 7.1 eV. However, the binding energy of the pair is only 2.3 eV and the complex would be expected to dissociate at a temperature given by $T \sim (W + E_B)/k_B \ln(\nu_0/\nu)$ or ~ 850°C when $\nu_0 \sim 10^{13}$ s^{-1} – a typical atomic jump frequency – and $\nu \sim 1$ s^{-1} the number of diffusion steps per second [1]. W is the diffusion barrier of C_i from chapter 5. The defect possesses (0/+) and (-/0) levels at $\sim E_c - 2.0$ and $E_c - 1.7$ eV which, together with its low thermal stability, exclude it as a model for Z_1/Z_2. The deep character of these levels prevents easy detection but the defect will compensate $N_{\rm C}$. The defect should, however, be IR-active and the negatively charged defect has a calculated local vibrational mode at 1324 cm^{-1}. The thermal stability and the acceptor level are however similar to the EH6/EH7 DLTS levels (Hemmingsson *et al.*, 1997; Zhang *et al.*, 2003) which disappear at around 750°C and have acceptor levels at $E_c - 1.65$ eV. The $N_{\rm C}$-C_i pair could therefore be one of the two overlapping DLTS levels of EH6/EH7.

6.4.4 Carbon di-interstitial C_2

Clearly, the Z_1/Z_2 defects are much more stable and there are unlikely to be any dangling bonds. Their presence would not only reduce the stability of the centre but lead to deep mid-gap levels. It is possible to form a defect without dangling bonds when C_i traps a second interstitial. This occurs when two C_i are placed at the centres of adjacent parallel bonds across one of the two types of hexagon found in hexagonal SiC (see Fig. 6.2). The interstitials can be bonded together not only as σ bonds but their p-orbitals, parallel to [0001] in Fig. 6.2, can form a π-bond. The bond length of this double bond is 1.36 Å. This results in a strong bond and the binding energy of the two interstitials is 5.7 eV. Such a defect would survive anneals of 2200°C according to the above equation for the dissociation temperature using $E_B = 5.7$ eV

---

[1] It might disappear at a lower temperature by trapping another mobile centre.

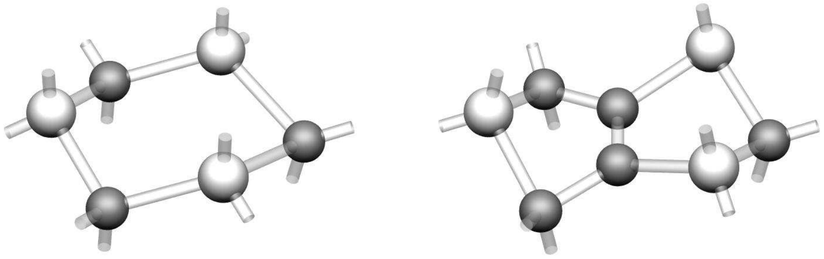


Figure 6.2: Left: the chair form of the basal plane hexagon. Right: the C_i-C_i π-bonded pair where two additional C atoms are added to the hexagon and bonded together. C (Si) atoms are shown in black (white). All atoms are fully coordinated.

and $W = 0.6$ eV. This defect has a donor level at E_v+0.5 eV and does not possess an acceptor level. This form of the di-carbon interstitial is also believed to occur in diamond (Goss *et al.*, 2001). The defect has LVMs at 1570 (A_1 sym.), 1360 (B_1), 1330 (A_1) and 1100 cm^{-1}(B_1).

6.4.5 Carbon di-interstitial next to a N donor N_C-C_2

The defect of interest here is a dimer composed of N_i and C_i. An alternative description is N_C-C_2, i.e. a di-interstitial trapped at a N donor N_C. The lowest energy form of the complex is shown in Fig. 6.3 a) in the positive charge state. The nitrogen atom is not part of the π-bond (this structure is ~ 0.25 eV higher in energy) but one of the neighbours to a π-bonded C-atom. Clearly, the defect is thermally stable as these bonds have to be broken in order for it to dissociate.

However, in the negative charge state a π^* orbital is occupied and the defect structure is unstable and changes spontaneously into the one shown in Fig. 6.3 b). Now the N atom and one of the C_i atoms become three-fold coordinated. On an electron counting argument, both of these possess a dangling bond containing *two* electrons. Thus the structural relaxation can be thought of as turning these into low energy lone pair orbitals which are filled. It is this rearrangement which stabilises the defect and leads to negative-U behaviour. The C–C double bond with a bond length of 1.36 Å in the positive charge state becomes a single bond in the negative charge state with bond length of 1.51 Å. The C–N bond changes only slightly from 1.42 Å to 1.45 Å if the charge state changes from positive to negative.

The calculated (0/+) and (-/0) levels lie at $E_c - 0.3$ eV and $E_c - 0.6$ eV, respectively, and are in good agreement with the observed levels of Z_1/Z_2. Thus the exceptional thermal stability of N_i-C_i, its bistability, electrical levels, and ease of formation all support the assignment to Z_1/Z_2. Also note that the two different types of hexagons found in 4*H*-SiC and 6*H*-SiC, namely the boat and chair configurations, would lead to two types of N_i-C_i defects consistent with the existence of Z_1 and Z_2 in 4*H*-SiC and E_1 and E_2 in 6*H*-SiC.

6.5 Conclusions

The most likely model found for the Z_1/Z_2 centres involves a thermally stable π-bonded di-carbon interstitial complex next to a nitrogen atom. The di-interstitial does not possess any unpaired electrons. The π-bond is unstable in the negative charge state but the presence of a single N atom in the defect allows the formation of two atoms with lone pairs. The calculated energy levels and thermal stability are in good agreement with observations.

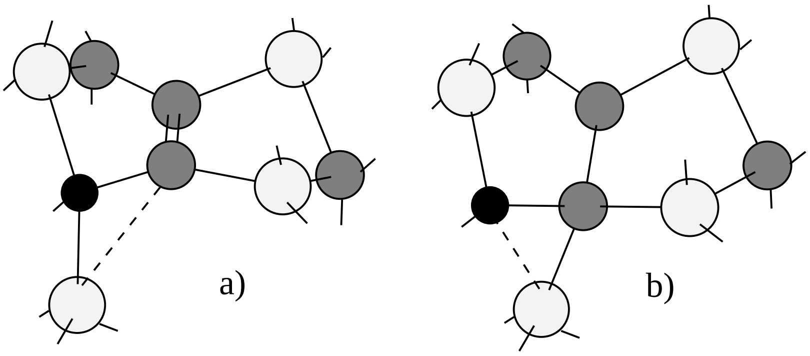


Figure 6.3: (a) The N_i-C_i defect in the positive charge state and (b) in the negative charge state. N, Si, C atoms are shown in black, white and grey. Note the π-bond in (a) is broken in (b) and N and one carbon atom possess a dangling bond.

The fact that the as-grown concentration is stable up to 2000°C while the concentration of irradiation induced defects decreases to the as-grown level could be explained by the fact that the defect complex is in fact more stable but another defect becomes mobile and reacts with $\mathrm{Z}_1/\mathrm{Z}_2$. Once the other defect is annealed out the $\mathrm{Z}_1/\mathrm{Z}_2$ concentration can not be lowered any further. A mobile vacancy offers therefore an explanation for the simultaneous disappearance of a positron-annihilation signal due to vacancies and $\mathrm{Z}_1/\mathrm{Z}_2$ defects. This does not require the vacancy to be part of $\mathrm{Z}_1/\mathrm{Z}_2$!

The calculated annealing temperature for the silicon vacancy supports the idea that it disappears at 750°C due to a transformation into V_C-C_Si.

CHAPTER 7

Optical and Electrical Properties of Vanadium and Erbium in 4*H*-SiC

7.1 Introduction

Transition metal impurities in semiconductors are important because they can possess multiple charge states and are efficient carrier traps. Moreover, they often give rise to internal intra-d optical transitions. Rare-earth dopants on the other hand are of increasing interest because their intra-f luminescence is stable to high temperatures in wide band gap materials (Zanatta, 2003). They are not noted for introducing deep levels. Recently, however, it has been reported that the intra-f luminescence of Er in SiC is stable to around 180°C and that the defect has a deep donor level (Choyke *et al.*, 1994; Shishkin *et al.*, 2000; Klettke *et al.*, 2001; Pasold *et al.*, 2003). It is then of interest to compare the properties of a transition metal like vanadium, whose internal transition is also stable at room temperature (Tajima *et al.*, 1997), with the rare-earth Er impurity.

While the d orbitals on V are known to interact strongly with surrounding atoms, the f-shell is almost decoupled. Russel-Saunders $L - S$ coupling within the f-shell leads to a $^4I_{15/2}$ ground state separated from an $^4I_{13/2}$ excited state by ~ 0.803 eV. The surrounding crystal field splits the ground state into a number of levels with separation ~ 2 meV (Choyke *et al.*, 1997) as well as introducing parity breaking terms leading to dipole allowed intra-f transitions. Similarly, crystal field and covalent effects split the d^1 manifold of V into a singly occupied e-level lying ~ 0.94 eV below empty t_2 levels. The latter are further split in hexagonal SiC into e and a_1 levels. Optical transitions seen in both absorption and luminescence occur between the e levels which are further split by ~ 1 meV mainly by spin-orbit coupling (Kaufmann *et al.*, 1997). The transition energy differs by ~ 20 meV in different polytypes (Mitchel *et al.*, 1999) and when V occupies different types of Si sites. The occupancy of the d-manifold can be changed by doping. Donor levels are reported at $E_v + 1.3$ and $E_v + 1.5$ eV for different sites in 6*H*-SiC (Baur *et al.*, 1997; Dalibor *et al.*, 1997*b*) and average values of $E_v + 1.53$ (Evwaraye *et al.*, 1996) or $E_v + 1.66$ eV (Mitchel *et al.*, 1999) have been reported for 4*H*-SiC. There is a continuing debate over the position of the acceptor level. In 4*H*-SiC, values at $E_c - 0.860$ and $E_c - 0.895$ eV have been assigned to V at the two Si sites (Dalibor *et al.*, 1997*b*), but other values, $E_c - 0.80$ eV (Jenny *et al.*, 1996) and even $E_c - 0.97$ eV (Achtziger and Witthuhn, 1998) have been reported although the last may be a V-complex. Early optical absorption investigations attributed a threshold at ~ 2.8 eV to the V acceptor level (Jenny *et al.*, 1995) but this transition will be reassigned below.

In contrast, the strong Coulomb potential in Er prevents a change in the occupancy of the f-shell whatever the Fermi-level. However, recent DLTS experiments on p-type 4*H*-SiC implanted with Er reveal a level at $E_v + 0.75$ eV (Pasold *et al.*, 2003) or $E_v + 0.68$ eV (Reshanov *et al.*, 2003). The identification was aided using both stable and radioactive ^{160}Er implants while the absence of a Poole-Frenkel effect suggests a donor level. Only about 10% of the implanted Er is electrically active and no Er related levels in the upper third of the gap were found from implants in n-type material. Although there have been reports of several EPR centres related to Er in as-grown n-type 6*H*-SiC (Baranov *et al.*, 1997), no detailed atomistic models were given.

Er related photoluminescence has been reported for several polytypes (Choyke *et al.*, 1994; Klettke *et al.*, 2001; Reshanov *et al.*, 2003; Uekusa *et al.*, 1999). The observation of the 2.36 eV $^2\mathrm{H}_{11/2} \rightarrow {}^4\mathrm{I}_{15/2}$ as well as the 0.803 eV (1.54μm) $^4\mathrm{I}_{13/2} \rightarrow {}^4\mathrm{I}_{15/2}$ transitions in n-type SiC (Choyke *et al.*, 1997) imply an $\mathrm{Er}^{3+}(4f^{11})$ electronic configuration. It might then be thought that substitutional Er can only behave as an acceptor like the isoelectronic Al substitutional defect. This, however, as shall be shown below, is not correct.

7.2 Method

The calculations are carried out in 72 atom supercells with ddpp basis for Si and C. Hartwigsen, Goedecker and Hutter (Hartwigsen *et al.*, 1998) pseudopotentials with frozen non-polarised $2p^6$ and $4f^{11}$ cores were used for V and Er respectively. A non-linear core correction was included for Er.

Calculations were carried out using MP-2^3 k-point sampling (Monkhorst and Pack, 1976) and the charge density was Fourier transformed with a maximum cut-off of 300 Ry. All atoms in the cell were allowed to relax.

The chemical potentials of Er and V are set to zero as their absolute solubilities are not of interest. The doping concentrations used in experiments of sometimes up to 10% are well above the equilibrium values.

The energy levels are found using the marker method. It is then natural to use the calculated and observed donor levels of V to improve the estimate of the Er level.

The calculated a and c lattice parameters for 4H-SiC are found to be 3.061 and 10.030 Å , and in good agreement with experimental values of 3.078 and 10.086 Å, respectively (Madelung and Schulz, 1989). The bulk modulus was found to be 224.2 GPa for 4H-SiC in agreement with an experimental value of 224 GPa, reported for 3C-SiC (Madelung and Schulz, 1989).

7.3 Results

7.3.1 Vanadium in 4H-SiC

The formation energy for substitutional V_{Si} at a cubic site is 4.3 eV lower than at a corresponding C site and 4.7 eV lower than at a pseudo-T_d interstitial site near C. For the substitutional impurity, V_{Si}, three of the V-C bonds are 1.94 Å while the fourth along [0001] is 1.96 Å. The Kohn-Sham levels show a gap related e level occupied by a single electron and unoccupied e and a_1 levels in agreement with the configuration required to explain the 0.94 eV intra-d absorption band (Kaufmann *et al.*, 1997). In the negative charge state, an additional electron occupies the lower e level leading to $S = 1$ or $S = 0$ spin states. The latter is found to be higher in energy by 0.5 eV. Thus $\mathrm{V}_{\mathrm{Si}}^-$ has $S = 1$ in agreement with EPR investigations (Baur *et al.*, 1997).

The donor level of V_{Si}, found from the comparison of the ionisation energies of V and of bulk material in the 72 atom cell, is $\mathrm{E_v} + 1.39$ eV. This is in good agreement with the reported values $\mathrm{E_v} + 1.53$ (Evwaraye *et al.*, 1996) and $\mathrm{E_v} + 1.66$ eV (Mitchel *et al.*, 1999). Turning to the acceptor level. Here, the electron affinity of V_{Si} is compared with that of bulk placing the level at $\mathrm{E_c} - 0.57$ eV. This is in fair agreement with the observed levels around $\mathrm{E_c} - 0.8$ to $\mathrm{E_c} - 0.9$ eV discussed above.

Previous radio-tracer investigations into the $(-/0)$ level of V found an erratic activation efficiency which varied inexplicably between samples (Achtziger and Witthuhn, 1998). This may be due to complexes with hydrogen. Hydrogen introduced during growth or by processing is known in the case of Si to strongly interact with transition metals and perturb their levels (Sachse *et al.*, 1997). It is found that H bonds to V_{Si} along [0001] with a binding energy of 2.2 eV. The neutral centre has $S = 0$ and would not be paramagnetic. Its acceptor level is deeper than that of isolated V by 0.13 eV while the donor level is shallower by 0.2 eV. This might mean that its acceptor level is not detected in the radio-tracing experiments (Achtziger and Witthuhn, 1998) which would then explain the low and sample dependent activation efficiency of V noted there. It would, however, be necessary to assume that Cr and Ti are not affected by hydrogen in the same way as no such erratic behaviour was observed for these impurities.

Finally, the intra-d $^2E \rightarrow {}^2E$ transition can be found from the energies of the two configurations found by promoting an electron from the lower to the upper e level. Similar calculations have been carried out for defects in diamond (Goss *et al.*, 1996). This gives an excitation energy of 0.87 eV in good agreement with the observed value of ~ 0.94 eV reported for 4H-SiC (Mitchel *et al.*, 1999).

7.3.2 Erbium in 4H-SiC

The formation energy for substitutional $\mathrm{Er}_{\mathrm{Si}}$ at a cubic site is 5.8 eV lower than at a corresponding C site and 8.3 eV lower than at the same interstitial site described above. Thus it is likely that Er only occupies the Si sub-lattice. In this case, three of the Er-C bonds are 2.16 Å and the unique one is 2.21 Å. The Kohn-Sham energy levels of the cell containing Er revealed an unexpected result. The gap contained a t_2 level occupied with a single hole and lying in the lower half of the gap. This was split into e lying below a_1 levels localised on the *C-neighbours* of Er. This demonstrates that the electrical activity of Er is to be understood in terms of a vacancy model where Er 'donates' its three valence electrons to the states formed by the dangling bonds on its four C neighbours similar to transition metal impurities like Pt in Si. This suggests that, unlike other group-III dopants like Al, $\mathrm{Er_{Si}}$ can act as both a donor and as an acceptor. The donor and acceptor levels of $\mathrm{Er_{Si}}$ are found to be $\mathrm{E_v} + 0.53$ eV and $\mathrm{E_c} - 1.90$ eV respectively. If the levels are corrected using the errors in the V donor level as described above, the donor level of Er is placed between $\mathrm{E_v} + 0.67$ and $\mathrm{E_v} + 0.74$ eV and in excellent agreement with experimental values around $\mathrm{E_v} + 0.7$ eV. There are no reports of the acceptor level. In fact, this is to be expected because it is so deep. DLTS experiments in 3*C*-SiC might reveal the $(-/0)$ level, shifted by the 0.9 eV conduction band offset with 4*H*-SiC, to around $\mathrm{E_c} - 1$ eV.

Clearly, the electropositive character of Er suggests an attraction with nitrogen or hydrogen. The trigonal form of the complexes have been investigated. They have either a substitutional N at a C neighbour of Er or a H atom directly bonded to the metal. The binding energy of the pairs is found to be 3.24 and 3.1 eV respectively and their donor levels lie 0.04 and 0.23 eV higher than the donor level of $\mathrm{Er_{Si}}$. The former is also close to the observed level. As expected from the vacancy model, the t_2 shell is now filled and there are no acceptor levels in the gap for both of these defects. Thus these impurities have passivated the acceptor activity of Er.

7.4 Conclusions

Both impurities are found to possess three charge states, with deep donor and acceptor levels. The calculated levels are in fair agreement with experimental values with errors about 0.2–0.3 eV. However, the mid-gap acceptor level of Er is too deep to be detected by DLTS, although the trivalent character of Er strongly implies that a deep acceptor level should exist. In contrast with V, its electrical activity arises from the dangling bonds on the four C neighbours and not from the $4f^{11}$ core which has been frozen in the calculation.

Hydrogen and nitrogen bind strongly to both V and Er and perturb the donor levels. H deepens the acceptor level of V but both H and N passivate the acceptor activity of Er. The presence of H may explain the erratic activation efficiency of implanted V which varies dramatically between samples (Achtziger and Witthuhn, 1998), while its impact on the excitation efficiency of Er is unknown. In Si, hydrogen increases the luminescence intensity (Kocher-Oberlehner *et al.*, 2003).

The calculated intra-*d* optical transition at 0.87 eV of V confirms the assignment of the ∼ 0.94 eV luminescence. Negatively charged $\mathrm{V^{3+}}$ has a $S = 1$ ground state, in agreement with EPR studies, but can exist in an $S = 0$ state which is placed 0.5 eV higher in energy. This state leads then to a trap whose $(-/0)$ level lies around $\mathrm{E_c} - 0.3$ to $\mathrm{E_c} - 0.4$ eV. This is in excellent agreement with the observed threshold at $\sim \mathrm{E_c} - 0.4$ eV but previously attributed to the *ground state* V acceptor level (Jenny *et al.*, 1995). As a consequence of the calculations, this acceptor level is attributed to the $S = 0$ excited $\mathrm{V^{3+}/V^{4+}}$ transition.

The EPR results show that more than one Er defect exists and the question arises whether isolated $\mathrm{Er_{Si}}$ or complexes, such as Er-N and Er-H, dominate the intra-*f* luminescence. The 1.54μm transition has only been detected in nitrogen doped material where $\mathrm{Er_{Si}^-}$ or $\mathrm{Er_{Si}}$-$\mathrm{N_C}$ can exist. It seems unlikely that isolated $\mathrm{Er_{Si}}$ is the optical centre. This is because a photo-generated hole would be captured by the negatively charged acceptor and the subsequent capture of an electron would release only ∼1.90 eV which is insufficient to excite the 2.36 eV $^2\mathrm{H}_{11/2} \rightarrow {}^4\mathrm{F}_{15/2}$ intra-*f* transition (Klettke *et al.*, 2001). Thus, according to the results, it is unlikely that isolated Er is responsible for the observed internal transition in contrast with V. However, Er-N or possibly Er-H complexes, having only donor levels below $\mathrm{E_v} + 0.9$ eV in 4*H*-SiC remain candidates for the optical centres.

CONCLUDING REMARKS

Although there have been many experiments on defects in silicon, the dominant semiconductor used up to now, there have been no *direct* observations of the self-interstitial. Only recently Mukashev *et al.* (1998) claimed to have detected it. And although there have been many calculations on the self-interstitial, few predictions of the electrical levels or spin densities exist which can be compared against experimental values. The calculations in this thesis show that the silicon self-interstitial is stable at the T-site in the double ionised charge state for a Fermi level below $\sim \mathrm{E_c} - 0.4$ eV in agreement with the observed double donor level of AA12. Furthermore the diffusion barrier and the spin-density of the paramagnetic singly charged defect are also qualitatively consistent with those of the AA12 centre and support its identification with the interstitial.

It is however surprising that this defect has not been detected by other groups. Furthermore there seem to be still a big discrepancy between the experimentally determined barrier for self-diffusion and the calculated values.

Turning to rod-like defects in Si. Many remarkable experiments like the HRTEM study by Takeda (1991) and theoretical studies have been carried out on them. There was however no explanation to why these defects aggregate on a {113} habit plane. It was shown here, that although the {111} RLDs are in fact significantly lower in energy than the more commonly observed {113} variety, and that the {001} oriented defect seen in HRTEM diamond and Ge is also competitive in terms of the formation energy of *infinite* systems, the elastic energy of a *finite* loop tends to lead to the formation of the {113} oriented planar defects. Furthermore, it was shown that the 903 meV PL band and $\mathrm{E_v} + 0.5$ eV DLTS peak are consistent with the properties of the chain ends and strain induced levels, respectively. Although a combination of first principles total energy calculation and elasticity theory could explain the {113} habit plane, the influence of kinetic factors could not be examined with the methods employed. The possibility that the choice of the habit plane is dictated by kinetic factors rather then formation energies can therefore not be excluded.

In contrast to silicon much less is known about defects in SiC. The lower space group symmetry of the crystal makes the analysis of EPR experiments much more difficult. The existence of inequivalent lattice sites adds further to the complexity in this material compared to silicon. There are therefore only very few known defects in this material. In addition to some PL experiments, DLTS is mainly used to look at electrically active defects. This unfortunately results in the fact that in addition to the annealing behaviour of a defect only its optical transition energies and electrical levels are known. Structural information like that found in Si is very limited. Despite this disadvantage atomistic models for the most common PL and DLTS centres were analysed.

It was shown that the antisite pairs possess properties in common with the centres giving rise to the Alphabet series of luminescent lines. Their transition energy, local vibrational mode and isotopic shift and thermal stability are all consistent with experiment. The most stable antisite pair may be responsible for the most energetic Alphabet line or the $\mathrm{D_I}$ band, which survives to higher temperatures. Thus the Alphabet bands can be assigned to antisite pairs with the zero–phonon line of the exciton bound to $\mathrm{Si_C}$ being perturbed by a nearby $\mathrm{C_{Si}}$. A model involving hole and electron trapping was presented to explain their formation in irradiated material. This model employed the high mobility of self-interstitials, already known from Si and Diamond.

The calculation of the diffusion barrier for the C interstitial in SiC finds indeed a low barrier enabling diffusion at room temperature. This seems to have prevented its direct detection so far. It was therefore concluded that the P–U PL lines seem to arise from defects *involving* the C interstitial in a C–C split form. However, these defects must be larger then the isolated interstitial to explain their thermal stability.

The most likely model found for the $\mathrm{Z_1/Z_2}$ centres involves a thermally stable π-bonded di-carbon interstitial complex next to a nitrogen atom. The di-interstitial does not possess any unpaired electrons. The π-bond is unstable in the negative charge state but the presence of a single N atom in the defect allows the formation of two atoms with lone pairs. The calculated energy levels and thermal stability are in good agreement with observations.

In the final chapter the transition metals erbium and vanadium in SiC were discussed. EPR results show that more than one Er defect exists and the question arises whether isolated $\mathrm{Er_{Si}}$ or complexes, such as Er-N and Er-H, dominate the intra-f luminescence. Since the 1.54μm transition has only been detected in nitrogen doped material where $\mathrm{Er_{Si}^{-}}$ or $\mathrm{Er_{Si}}$-$\mathrm{N_C}$ can exist, it was concluded that it is unlikely that isolated $\mathrm{Er_{Si}}$ is the optical centre. This is because a photo-generated hole would be captured by the negatively charged acceptor and the subsequent capture of an electron would release only ~1.90 eV which is insufficient to excite the 2.36 eV $^2\mathrm{H}_{11/2} \rightarrow {}^4\mathrm{F}_{15/2}$ intra-f transition. However, Er-N or possibly Er-H complexes, having only donor levels below $\mathrm{E_v} + 0.9$ eV in 4H-SiC remain candidates for the optical centres.

Although there seems to be experimental evidence now for the self-interstitial in Si, many questions remain open. The simulated diffusion barrier is much lower than barrier heights deduced from experiments. Also the activation energy for self diffusion is about 1 eV higher than predicted by LDA. It is also unclear how RLDs nucleate and the identity of their nucleation centre. Furthermore it is unkown how they transform into dislocation loops under further annealing. Turning to SiC, further research has to be done regarding the Si interstitial. It remains open whether it has a (++/0) negative-U transition in the upper part of the gap. The model for $\mathrm{Z_1/Z_2}$ has to be revised should the proposal that N is not part of the defect be verified.

BIBLIOGRAPHY

Aboelfotoh M and Doyle J, Phys. Rev. B **59**, 10823 (1999). 63

Achtziger N and Witthuhn W, Phys. Rev. B **57**, 12181 (1998). 69, 70, 71

Alippi P and Colombo L, Phys. Rev. B **62**, 1815 (2000). 42, 43, 45, 46, 48

Anderson P W, Phys. Rev. Lett. **35**, 953 (1975). 63

Arai N, Takeda S, and Kohyama M, Phys. Rev. Lett. **78**, 4265 (1997). 33

Awadelkarim O, Henry A, Monemar B, Lindstrom J, Zhang Y, and Corbett J, Phys. Rev. B **42**, 5635 (1990). 33

Bachelet G B, Hamann D R, and Schlüter M, Phys. Rev. B **26**, 4199 (1982). 10

Baranov P, Ilyin I, and Mokhov E, Solid State Commun. **103**, 291 (1997). 69

von Bardeleben H, Cantin J, Reshanov S, and Rastegaev V, Mater. Sci. Forum. **433-4**, 507 (2002). 58

von Bardeleben H, Cantin J, Vickridge I, and Battistig G, Phys. Rev. B **62**, 10126 (2000). 60, 64

von Barth U and Hedin L, J. Phys. C **5**, 1629 (1972). 9

Baryam Y and Joannopoulos J D, Phys. Rev. B **30**, 2216 (1984). 35

Baur J, Kunzer M, and Schneider J, Phys. Stat. Sol. (a) **162**, 153 (1997). 69, 70

Bekman H H P T, Gregorkiewicz T, and Ammerlaan C A J, Phys. Rev. B **39**, 1648 (1989). 35

Benton J L, Libertino S, Kringhøj D J, Eaglesham D J, M. P J, and Coffa S, J. Appl. Phys. **82**, 120 (1997). 42

Blöchl P E, Smargiassi E, Car R, Laks D B, Andreoni W, and Pantelides S T, Phys. Rev. Lett. **70**, 2435 (1993). 35

Bockstedte M, Heid M, Mattausch A, and Pankratov O, Mater. Sci. Forum. **389-3**, 471 (2002). 60

Bockstedte M, Heid M, and Pankratov O, Phys. Rev. B **67**, art. no. 193102 (2003*a*). 58

Bockstedte M, Mattausch A, and Pankratov O, Phys. Rev. B **68**, 205201 (2003*b*). 61

Born M and Oppenheimer R, Ann. Physik **87**, 457 (1925). 4

Bourgoin J and Lannoo M, *Point Defects in Semiconductors II: Experimental Aspects* (Springer–Verlag, Berlin, 1983). 22

Bratus V Y, Makeeva I N, Okulov S M, Petrenko T L, Petrenko T T, and von Bardeleben H J, Physica B **308**, 621 (2001). 58

Briddon P R, EMIS Data Rev. **20**, 330 (1999). 17

Brower K L, Phys. Rev. B **14**, 872 (1976). 33

Brust D, in *Methods in Computational Physics*, volume 8, p. 33 (Academic Press, 1968). 10

Capano M, Cooper J, Melloch M, Saxler A, and Mitchel W, J. Appl. Phys. **87**, 8773 (2000). 64

Carlson R O, Hall R N, and Pell E M, J. Phys. Chem. Solids **8**, 81 (1959). 38

Ceperley D M, Phys. Rev. B **18**, 3126 (1978). 10

Ceperley D M and Alder B J, Phys. Rev. Lett. **45**, 566 (1980). 10

Chadi D J, Phys. Rev. B **46**, 9400 (1992). 35

Chadi D J and Chang K J, Phys. Rev. Lett. **61**, 873 (1988). 64

Chou C T, Cockayne D J H, Zou J, Kringhøj P, and Jagadish C, Phys. Rev. B **52**, 17223 (1995). 43, 48, 51

Choyke W, Devaty R, Clemen L, Yoganathan M, Pensl G, and Hassler C, Appl. Phys. Lett. **65**, 1668 (1994). 69, 70

Choyke W, Devaty R, Yoganathan M, Pensl G, and Edmond J, in *Proceedings of the 7th International Conference on Shallow-Level Centers in Semiconductors*, p. 297 (World Scientic Publ. Co., Singapore, 1997). 69, 70

Clark S J and Ackland G J, Phys. Rev. B **56**, 47 (1997). 35

Claverie A, unplublished (2003). 41, 43

Claverie A, Giles L F, Omri M, de Mauduit B, Ben Assayag G, and Mathiot D, Nucl. Instrum. Methods B **147**, 1 (1999). 41, 42

Coffa S, Libertino S, and Spinella C, Appl. Phys. Lett. **76**, 321 (2000). 42, 47

Coomer B J, *A first principles study of radiation defects in semiconductors*, Ph.D. thesis, School of Physics, University of Exeter, United Kingdom (2000). 34

Coomer B J, Goss J P, Jones R, Öberg S, and Briddon P R, Physica B **273–274**, 505 (1999). 31, 32, 46, 51

Coomer B J, Goss J P, Jones R, Öberg S, and Briddon P R, J. Phys. Cond. Matter **13**, L1 (2001). 31, 32, 33, 43

Coutinho J, Torres V, Jones R, and Briddon P, Phys. Rev. B **67**, art. no. 035205 (2003). 36

Cowern N, Janssen K, Vandewalle G, and Gravesteijn D, Phys. Rev. Lett. **65**, 2434 (1990). 1

Cowern N E B, Mannino G, Stolk P A, Roozeboom F, Huizing H G A, van Berkum J G M, Cristiano F, Claverie A, and Jaraíz M, Phys. Rev. Lett. **82**, 4460 (1999). 31, 33

Cristiano F, Grisolia J, Colombeau B, Omri M, de Mauduit B, Claverie A, Giles L F, and Cowern N E B, J. Appl. Phys. **87**, 8420 (2000). 42

Dalibor T, Pensl G, Kimoto T, Matsunami H, Sridhara S, Devaty R, and Choyke W, Diam. Relat. Mat. **6**, 1333 (1997*a*). 63

Dalibor T, Pensl G, Matsunami H, Kimoto T, Choyke W, Schoner A, and Nordell N, Phys. Status Solidi A-Appl. Res. **162**, 199 (1997*b*). 63, 69

Dalibor T, Peppermuller C, Pensl G, Sridhara S, Devaty R, Choyke W, Itoh A, Kimoto T, and Matsunami H, Instit. Phys. Confer. Ser. **142**, 517 (1996). 56, 63

Davies G, Rep. Prog. Phys. **44**, 787 (1981). 24

Davies G, in *Identification of Defects in Semiconductors*, (edited by M Stavola), volume 51B of *Semiconductors and Semimetals*, p. 1 (Academic Press, 1999). 24

Davies G, Lightowlers E C, and Ciechanowska Z E, J. Phys. C **20**, 191 (1987). 32

Dean P J and Herbert D C, in *Excitons*, (edited by K Cho), number 14 in Topics in current physics, pp. 55–176 (Springer–Verlag, 1979). 24

Eberlein T A G, Pinho N, Jones R, Coomer B J, Goss J P, Briddon P R, and Öberg S, Physica B **308–310**, 454 (2001). 34, 38

Egilsson T, Bergman J, Ivanov I, Henry A, and Janzen E, Phys. Rev. B **59**, 1956 (1999*a*). 56

Egilsson T, Henry A, Ivanov I, Lindstrom J, and Janzen E, Phys. Rev. B **59**, 8008 (1999*b*). 56

Estreicher S K, Fedders P A, and Ordejón P, Physica B **308–310**, 1 (2001*a*). 32

Estreicher S K, Gharaibeh M, Fedders P A, and Ordejón P, Phys. Rev. Lett. **86**, 1247 (2001*b*). 32

Estreicher S K, Weber J, Derecskeikovacs A, and Marynick D S, Phys. Rev. B **55**, 5037 (1997). 32

Evans G, Steeds J, Ley L, Hundhausen M, Schulze N, and Pensl G, Phys. Rev. B **66**, art. no. 035204 (2002). 56

Evwaraye A, Smith S, and Mitchel W, J. Appl. Phys. **79**, 7726 (1996). 69, 70

Ewald P P, Ann. Phys. **64**, 253 (1921). 12

Fedina L, Gutakovskii A, and Aseev A, Cryst. Res. Technol. **35**, 775 (2000). 51

Fedina L, Gutakovskii A, Aseev A, Van Landuyt J, and Vanhellemont J, Phil. Mag. A **66**, 423 (1998). 41, 43

Fedina L, Gutakovskii A, Aseev A, Van Landuyt J, and Vanhellemont J, Phys. Status Solidi A **171**, 147 (1999). 41, 43, 48

Ferreira-Lima C A, *Defects in electron irradiated germanium*, Ph.D. thesis, University of Cambridge, U.K. (1975). 43

Fock V, Z. Phys. **61**, 126 (1930). 4

Fujihira K, Kimoto T, and Matsunami H, Appl. Phys. Lett. **80**, 1586 (2002). 63

Gali A, Deak P, Rauls E, Son N, Ivanov I, Carlsson F, Janzen E, and Choyke W, Phys. Rev. B **67**, art. no. 155203 (2003). 59

Gilbert T L, Phys. Rev. B **12**, 2111 (1975). 8

Gong M, Fung S, Beling C, and You Z, J. Appl. Phys. **85**, 7604 (1999). 63

Goss J P, Coomer B J, Jones R, Fall C J, Briddon P R, and Öberg S, Phys. Rev. B **67**, 165208 (2003). xi, 41, 43, 49, 50

Goss J P, Coomer B J, Jones R, Fall C J, Latham C D, Briddon P R, and Öberg S, J. Phys. Cond. Matter **12**, 10257 (2000). 41, 43, 45

Goss J P, Coomer B J, Jones R, Shaw T D, Briddon P R, Rayson M, and Öberg S, Phys. Rev. B **63**, 195208 (2001). 31, 41, 43, 56, 66

Goss J P, Eberlein T A G, Jones R, Pinho N, Blumenau A T, Frauenheim T, Briddon P R, and Öberg S, J. Phys.: Cond. Matter **14**, 12843 (2002). 41, 42, 43, 48, 51

Goss J P, Jones R, Breuer S J, Briddon P R, and Öberg S, Phys. Rev. Lett. **77**, 3041 (1996). 70

Gotz W, Schoner A, Pensl G, Suttrop W, Choyke W, Stein R, and Leibenzeder S, J. Appl. Phys. **73**, 3332 (1993). 64

Greulich-Weber S, Feege M, Spaeth J, Kalabukhova E, Lukin S, and Mokhov E, Solid State Commun. **93**, 393 (1995). 64

Haberstroh C, Helbig R, and Stein R, J. Appl. Phys. **76**, 509 (1994). 56

Hakala M, Puska M J, and Nieminen R M, Phys. Rev. B **61**, 8155 (2000). 36

Hallen A, Keskitalo N, Josyula L, and Svensson B, J. Appl. Phys. **86**, 214 (1999). 35

Hamann D R, Schlüter M, and Chiang C, Phys. Rev. Lett. **48**, 1425 (1982). 10

Harrison W A, *Pseudopotentials in the theory of metals* (W. A. Benjamim, New york, 1966). 10

Hartwigsen C, Goedecker S, and Hutter J, Phys. Rev. B **58**, 3641 (1998). 70

Heine V, Solid State Physics **24**, 1 (1970). 10

Hemmingsson C, Son N, Ellison A, Zhang J, and Janzen E, Phys. Rev. B **58**, 10119 (1998). 63

Hemmingsson C, Son N, and Janzen E, Appl. Phys. Lett. **74**, 839 (1999). 63

Hemmingsson C, Son N, Kordina O, Bergman J, Janzen E, Lindstrom J, Savage S, and Nordell N, J. Appl. Phys. **81**, 6155 (1997). 63, 65

Henkelman G and Jonsson H, J. Chem. Phys. **111**, 7010 (1999). 38

Herring C, Johnson N M, and Van de Walle C G, Phys. Rev. B **64**, 125209 (2001). 64

Hirth J P and Lothe J, *Theory of dislocations*, Materials Science and Engineering (McGraw-Hill, New York, 1968). 50

Hohenberg P and Kohn W, Phys. Rev. **136**, 864 (1964). 8

Hourahine B, Jones R, Safonov A N, Öberg S, Briddon P R, and Estreicher S K, Phys. Rev. B **61**, 12594 (2000). 31, 32

Humble P, KacKenzie J K, and Olsen A, Phil. Mag. A **52**, 605 (1985). 41, 43, 49

Hunt D C, Twitchen D J, Newton M E, Baker J M, Anthony T R, Banholzer W F, and Vagarali S S, Phys. Rev. B **61**, 3863 (2000). 31, 56

Janson M, Linnarsson M, Hallen A, Svensson B, Nordell N, and Bleichner H, Appl. Phys. Lett. **76**, 1434 (2000). 1

Jenny J, Skowronski J, Mitchel W, Hobgood H, Glass R, Augustine G, and Hopkins R, Appl. Phys. Lett. **68**, 1963 (1996). 69

Jenny J, Skowronski M, Mitchel W, Hobgood H, Glass R, Augustine G, and Hopkins R, J. Appl. Phys. **78**, 3160 (1995). 69, 71

Jeong S, Shim H, Yoon H, Cheong M, Choi R, Suh E, and Lee H, J. Appl. Phys. **91**, 9711 (2002). 36

Jeyanathan L, Lightowlers E C, Higgs V, and Davies G, Mater. Sci. Forum **143-147**, 1499 (1994). 42, 47

Jones R and Briddon P R, *The* ab initio *cluster method and the dynamics of defects in semiconductors*, volume 51A of *Semiconductors and Semimetals*, chapter 6 (Academic Press, Boston, 1998). 57

Jones R and Briddon P R, in *Identification of Defects in Semiconductors*, (edited by M Stavola), volume 51A of *Semiconductors and Semimetals*, p. 287 (Academic Press, 1999). 11

Jones R O and Gunnarsson O, Rev. Mod. Phys. **61**, 689 (1989). 8

Kaminskii A S and Lavrov E V, Solid State Commun. **106**, 751 (1998). 32

Kaminskii A S, Leiferov B M, and Safonov A N, Fiz. Tverd. Tela (Leningrad) [Sov. Phys. Solid State] **29**, 961 (1987). 31

Kaufmann B, Dornen A, and Ham F, Phys. Rev. B **55**, 13009 (1997). 69, 70

Kawasuso A, Redmann F, Krause-Rehberg R, Frank T, Weidner M, Pensl G, Sperr P, and Itoh H, J. Appl. Phys. **90**, 3377 (2001*a*). 63

Kawasuso A, Redmann F, Krause-Rehberg R, Weidner M, Frank T, Pensl G, Sperr P, Triftshauser W, and Itoh H, Appl. Phys. Lett. **79**, 3950 (2001*b*). 63

Kennedy T, Freitas J, and Bishop S, J. Appl. Phys. **68**, 6170 (1990). 56

Kim J, Kirchhoff F, Aulbur W G, Wilkins J W, Khan F S, and Kresse G, Phys. Rev. Lett. **83**, 1990 (1999). 34

Kim J, Kirchhoff F, Wilkins J W, and Khan F S, Phys. Rev. Lett. **84**, 503 (2000). 41, 42, 43, 45, 48, 49

Kim J, Wilkins J W, Khan F S, and Canning A, Phys. Rev. B **55**, 16186 (1997). 41, 42, 43, 45

Kimoto T, Itoh A, Matsunami H, Sridhara S, Clemen L, Devaty R, Choyke W, Dalibor T, Peppermuller C, and Pensl G, Appl. Phys. Lett. **67**, 2833 (1995). 63

Kirkpatrick C G, Noonan J R, and Streetman B G, Radiat. Eff. **30**, 97 (1976). 32

Klettke O, Reshanov S, Pensl G, Shishkin Y, Devaty R, and Choyke W, Physica B **308**, 687 (2001). 69, 70, 71

Kocher-Oberlehner G, Jantsch W, Palmetshofer L, and Ulyashin A, Appl. Phys. Lett. **83**, 623 (2003). 71

Kohn W and Sham L J, Phys. Rev. **140**, A1133 (1965). 8, 9

Kohyama M and Takeda S, Phys. Rev. B **46**, 12305 (1992). 41, 42, 43

Kohyama M and Takeda S, Phys. Rev. B **51**, 13111 (1995). 41, 42, 47, 48

Konovalov V, Zvanut M, and van Tol J, Phys. Rev. B **68**, art. no. 012102 (2003). 58

Koopmans T, Physica **1**, 104 (1934). 6

Kubaschewski O and Alcock C B, *Metallurgical thermochemistry* (Pergamon, Oxford, 1979). 57

Lang D V, J. Appl. Phys. **45**, 3014 (1974). 19

Laube M, Pensl G, and Itoh H, Appl. Phys. Lett. **74**, 2292 (1999). 1

Leary P, Jones R, and Öberg S, Phys. Rev. B **57**, 3887 (1998). 56

Lee W C, Lee S G, and Chang K J, J. Phys. Cond. Matter **10**, 995 (1998). xi, 35, 36, 38

Lee Y H, Appl. Phys. Lett. **73**, 1119 (1998). 34

Lee Y H, Gerasimenko N N, and Corbett J W, Phys. Rev. B **14**, 4506 (1976). 34

Leeuw S W D, Perram J W, and Smith E R, Proc. Roy. Soc. Lond. A **373**, 27 (1980). 12

Libertino S, Coffa S, and Benton J L, Phys. Rev. B **65**, 195206 (2001). 42, 47

Lightowlers E C, *Growth and Characterisation of Semiconductors*, chapter 5, p. 135 (Adam Hilger, Bristol, 1990). 24

Lightowlers E C, Jeyanathan L, Safonov A N, Higgs V, and Davies G, Mater. Sci. Eng. B **24**, 144 (1994). 42, 47

Lingner T, Greulich-Weber S, Spaeth J, Gerstmann U, Rauls E, and Overhof H, Physica B **308**, 625 (2001). 65

Louie S G, Froyen S, and Cohen M L, Phys. Rev. B **26**, 1738 (1982). 10

Lundqvist S, *Theory of the Inhomogeneous Electron Gas* (Plenum, New York, 1983). 8

Madelung O and Schulz M, editors, *Landolt-Börnstein, new series* (Springer, Berlin, 1989). 57, 70

Mchedlidze T and Suezawa M, J. Phys.: Cond. Matter **15**, 3683 (2003). 33

McWeeny R, *Methods of Molecular Quantum Mechanics* (Academic Press, New York, 1989). 5

Minaev N, Mudrii A, and Tkachev V, Phys. Status Solidi B-Basic Res. **108**, K89 (1981). 32

Mitchel W, Perrin R, Goldstein J, Saxler A, Roth M, Smith S, Solomon J, and Evwaraye A, J. Appl. Phys. **86**, 5040 (1999). 69, 70

Monkhorst H J and Pack J D, Phys. Rev. B **13**, 5188 (1976). 10, 45, 70

Mukashev B, Spitsyn A, Fukuoka N, and Saito H, Jpn. J. Appl. Phys. Part 1 - Regul. Pap. Short Notes Rev. Pap. **21**, 399 (1982). 33

Mukashev B N, Abdullin K A, and Gorelkinskii Y V, Phys. Status Solidi A **168**, 73 (1998). 1, 31, 34, 35, 73

Mukashev B N, Abdullin K A, and Gorelkinskii Y V, Uspekhi Fizicheskikh Nauk **170**, 143 (2000). 34

Mukashev B N, Abdullin K A, Gorelkinskii Y V, and Tokmoldin S Z, Mater. Sci. Eng. B **58**, 171 (1999). 34

Muto S and Takeda S, Phil. Mag. Lett. **72**, 99 (1995). 41, 43, 49

Nakamura M, Nagai S, Aoki Y, and Naramoto H, Appl. Phys. Lett. **72**, 1347 (1998). 32

Needs R J, J. Phys. Cond. Matter **11**, 10437 (1999). xi, 35, 36, 38

Pack J D and Monkhorst H J, Phys. Rev. B **16**, 1748 (1977). 10

Papagiannidis S, *Ab Initio Modelling of Defect Complexes in Semiconductors*, Ph.D. thesis, University of Newcastle upon Tyne (2003). 38, 65

Parisini A and Bourret A, Phil. Mag. A **67**, 605 (1993). 42

Parr R G and Yang W, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989). 8

Pasold G, Albrecht F, Grillenberger J, Grossner U, Hulsen C, Witthuhn W, and Sielemann R, J. Appl. Phys. **93**, 2289 (2003). 69

Patrick L and Choyke W, Phys. Rev. B **5**, 3253 (1972). 56

Perdew J P, in *Electronic Structure of Solids '91*, (edited by P Ziesche and H Eschrig) (Akademie Verlag, Berlin, 1991). 10

Perdew J P, Burke K, and Ernzerhof M, Phys. Rev. Lett. **77**, 3865 (1996*a*). 10

Perdew J P, Burke K, and Wang Y, Phys. Rev. B **54**, 16533 (1996*b*). 10

Perdew J P and Wang Y, Phys. Rev. B **45**, 13244 (1992). 10

Perdew J P and Zunger A, Phys. Rev. B **23**, 5048 (1981). 9, 10

Pickett W E, Comp. Phys. Rep. **9**, 115 (1989). 10

Pierreux D and Stesmans A, Phys. Rev. B **68**, 193208 (2003). 33

Pintilie I, Pintilie L, Irmscher K, and Thomas B, Appl. Phys. Lett. **81**, 4841 (2002). 63

Rauls E, Lingner T, Hajnal Z, Greulich-Weber S, Frauenheim T, and Spaeth J, Phys. Status Solidi B-Basic Res. **217**, R1 (2000). 65

Resende A, Jones R, Öberg S, and Briddon P R, Phys. Rev. Lett. **82**, 2111 (1999). 36

Reshanov S A, Klettke O, Pensl G, and Choyke W J, Mater. Sci. Forum **433**, 459 (2003). 69, 70

Robertson L S, Jones K S, Rubin L M, and Jackson J, J. Appl. Phys. **87**, 2910 (2000). 41

Roothaan C C J, Rev. Mod. Phys. **23**, 69 (1951). 4, 7

Sachse J U, Sveinbjörnsson E Ö, Jost W, and Weber J, Phys. Rev. B **55**, 16176 (1997). 70

Sadowski H, Schulze N, Frank T, Laube M, Pensl G, and Helbig R, Mater. Sci. Forum. **353-3**, 401 (2000). 59

Sauer R and Weber J, Physica. B. **116**, 195 (1983). 33

Schultz P J, Thompson T D, and Elliman R G, Appl. Phys. Lett. **60**, 59 (1992). 31, 32

Shishkin Y, Choyke W J, Devaty R P, Achtziger N, Opfermann T, and Witthuhn W, Mater. Sci. Forum **338**, 639 (2000). 69

Slater J C, Phys. Rev. **34**, 1293 (1929). 5

Slater J C, *Quantum Theory of Molecules and Solids*, volume 2 (McGraw-Hill, New York, 1965). 4, 23

Son N, Hai P, and Janzen E, Phys. Rev. Lett. **87**, art. no. 045502 (2001). 58

Song L W and Watkins G D, Phys. Rev. B **42**, 5759 (1990). 33

Steeds J, Carosella F, Evans G, Ismail M, Danks L, and Voegeli W, Mater. Sci. Forum. **353-3**, 381 (2000). 56

Steeds J W, Private communication (2001). 58

Stoneham A M, *Theory of Defects in Solids* (Oxford University Press, London, 1975). 4, 10, 23

Storasta L, Carlsson F, Sridhara S, Bergman J, Henry A, Egilsson T, Hallen A, and Janzen E, Appl. Phys. Lett. **78**, 46 (2001). 56, 63

Storasta L, Henry A, Bergman J, and Janzén E, in *Proceedings of the International Conference on Silicon Carbide and Related Materials 2003* (2004). 63

Tajima M, Kumagaya Y, Nakata T, Inoue M, and Nakamura A, Jpn. J. Appl. Phys. Part 2 - Lett. **36**, L1185 (1997). 69

Takeda S, Japan J. Appl. Phys. **30**, L639 (1991). 41, 43, 73

Takeda S, proceeding of the 10th Conference on Microscopy of Semiconducting Materials, IoP Conference Series **157**, 25 (1997). 41, 43

Takeda S and Kamino T, Phys. Rev. B **51**, 2148 (1995). 41, 42

Takeda S, Kohyama M, and Ibe K, Phil. Mag. A **70**, 287 (1994). 42, 43

Terashima K, Ikarashi T, Watanabe M, and Kitano T, Mater. Sci. Forum **258–263**, 587 (1997). 33

Thijssen J M, *Computational Physics* (Cambridge University Press, Cambridge, 1999). 5

Torpo L, Marlo M, Staab T E M, and Nieminen R M, J. Phys. Cond. Matter **13**, 6203 (2001). xi, 57, 58, 64

Torpo L and Nieminen R, Mater. Sci. Eng. B-Solid State Mater. Adv. Technol. **61-2**, 593 (1999). xi, 57

Torpo L, Nieminen R M, Laasonen K E, and Pöykkö S, Appl. Phys. Lett. **74**, 221 (1999). 64

Torpo L, Poykko S, and Nieminen R, Phys. Rev. B **57**, 6243 (1998). xi, 57

Torpo L, Staab T, and Nieminen R, Phys. Rev. B **65**, art. no. 085202 (2002). 65

Troxell J R and Watkins G D, Phys. Rev. B **22**, 921 (1980). 64

Twitchen D J, Hunt D C, Wade C, Newton M E, Baker J M, Anthony T R, and Banholzer W F, Physica B **273–274**, 644 (1999). 56

Uekusa S, Awahara K, and Kumagai M, IEEE Trans. Electron Devices **46**, 572 (1999). 70

Usov I, Suvorova A, Sokolov V, Kudryavtsev Y, and Suvorov A, J. Appl. Phys. **86**, 6039 (1999). 1

Volm D, Meyer B, Hofmann D, Chen W, Son N, Persson C, Lindefelt U, Kordina O, Sorman E, Konstantinov A, Monemar B, and Janzen E, Phys. Rev. B **53**, 15409 (1996). 27

Vosko S J, Wilk L, and Nusair M, Chinese J. Phys. **58**, 1200 (1980). 10

Wagner M, Thinh N, Son N, Chen W, Janzen E, Baranov P, Mokhov E, Hallin C, and Lindstrom J, Phys. Rev. B **66**, art. no. 155214 (2002). 64

Wang C, Bernholc J, and Davis R F, Phys. Rev. B **38**, 12752 (1988). 58

Watkins G, Phys. Status Solidi A-Appl. Res. **186**, 167 (2001). 35, 38

Watkins G D, *Radiation damage in semiconductors*, p. 97 (Dunod, Paris, 1964). 34, 35, 56

Watkins G D and Brower K L, Phys. Rev. Lett. **36**, 1329 (1976). 56

Watkins G D and Corbett J W, Phys. Rev. **121**, 1001 (1961). 24

Watkins G D and Troxell J R, Phys. Rev. Lett. **44**, 593 (1980). 64

Woodbury H and Ludwig G, Phys. Rev. **124**, 1083 (1961). 64

Zanatta A, Appl. Phys. Lett. **82**, 1395 (2003). 1, 69

Zhang J, Storasta L, Bergman J, Son N, and Janzen E, J. Appl. Phys. **93**, 4708 (2003). 1, 65

Zywietz A, Furthmuller J, and Bechstedt F, Phys. Rev. B **59**, 15166 (1999). 57, 65