A First Principles Study of Hydrogen Related Defects in Silicon

Submitted by Benjamin Hourahine to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Physics, March 2000.

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

The results of first principles calculations are presented for several types of point defect in crystalline silicon that contain hydrogen. The information presented in this work is derived using cluster and supercell geometries within the local density functional theory.

The properties of molecular hydrogen within bulk silicon are considered and compared with recent experimental observations. The interaction between hydrogen molecules and other defects such as chemically inert voids or oxygen impurities are also examined.

Several defects which have been suggested to form in proton-implanted silicon are examined. The behaviour of hypothetical complexes between silicon self-interstitials and several hydrogen atoms are examined in an attempt to resolve a long standing argument over the nature of two families of complexes between native defects and hydrogen. A new form of hydrogen dimer which is suggested to form in such material is simulated for the first time, and compared against both experiment and other dimer structures.

The interaction of hydrogen with multiple vacancies in the silicon lattice is simulated, leading to an assignment for a large family of optically active defects as complexes between hydrogen and the largest vacancy centre so far observed in silicon.

The behaviour of complexes between hydrogen and isolated carbon impurities, which substitute for silicon atoms in the lattice, is simulated, allowing a consistent model for two sets of apparently disparate experimental observations to be suggested.

Acknowledgements

The School of Physics at Exeter and the EPSRC are gratefully acknowledged for providing the necessary funding and resources over the course of this PhD.

Professor Bob Jones deserves special thanks for his enthusiastic support, patient supervision and infectious curiosity in this subject.

I would also like to thank the extended AIMPRO family, particularly Patrick, Jon, António, José, Paul, Joachim, James, Chris, Chris, Malc and Sven for the various and diverse help, assistance and collaboration. I'd also like to extend my regards to the members of the semiconductor physics group at Exeter.

I would also like to acknowledge the helpful comments, unpublished data, and discussions with and from the many collaborators and colleagues involved in this area, particularly Ron Newman, Rhian Pritchard, Alexi Safonov, Lone Hoffmann, Brian Bech Nielsen, Edward Lavrov, Stefan Estreicher, Emil Roduner and Mike Stavola.

I would also like to thank my family, to whom I dedicate this thesis.

Contents

List of publications

Some of the material presented within this thesis has been published over the course of the degree. The relevant publications are marked.*

1. Anomalous shift of the 1075 cm^{-1} oxygen-hydrogen defect in silicon.
   B. Hourahine, R. Jones, S. Öberg, and P. R. Briddon. *Materials Science Forum*, 258, 277 (1997).*

2. Local vibrational modes of weakly bound O–H complexes in Si.
   B. Bech Nielsen, K. Tanderup, M. Budde, K. Bond Nielsen, J. L. Lindstrom, R. Jones, S. Öberg, B. Hourahine, and P. R. Briddon. *Materials Science Forum*, 258, 391 (1997).

3. Hydrogen molecules in silicon located at interstitial sites and trapped in voids.
   B. Hourahine, R. Jones, S. Öberg, R. C. Newman, P. R. Briddon, and E. Roduner. *Physical Review B*, 57, 12666 (1998).*

4. A theoretical study of O chemisorption on GaN $(0001)/(000\bar{1})$ surfaces.
   J. Elsner, R. Gutierrez, B. Hourahine, R. Jones, M. Haugk, and Th. Frauenheim. *Solid State Communications*, 108, 953 (1998).

5. *Ab inito* studies of hydrogen molecules in silicon.
   B. Hourahine, R. Jones, S. Öberg, R. C. Newman, P. R. Briddon, and E. Roduner. In *High Performance Computing*, Ed. R. J. Allan, M. F. Guest editors, A. D. Simpson, D. S. Henty, and D. A. Nicole, Kluwer Academic, New York (1998).*

6. Molecular hydrogen traps within silicon.
   B. Hourahine, R. Jones, S. Öberg, and P. R. Briddon. *Materials Science and Engineering B*, 58, 24 (1999).*

7. Self-interstitial–hydrogen complexes in silicon.

   B. Hourahine, R. Jones, S. Öberg, and P. R. Briddon. *Physical Review B*, 59, 15729 (1999).*

8. Optically active hydrogen dimers in silicon.

   B. Hourahine, R. Jones, A. N. Safonov, S. Öberg, P. R. Briddon, and S. K. Estreicher. *Physica B*, 273, 176 (1999).*

9. The Interaction of Hydrogen with Deep Level Defects in Silicon.

   R. Jones, B. J. Coomer, J. P. Goss, B. Hourahine, and A. Resende. pp 173-248 in *Special defects in semiconducting materials*. Edited by R. P. Agarwala. Scitech Publications Ltd. Zuerich-Uetikon, Switzerland (2000).*

10. Identification of the Hexavacancy in Silicon with the B_{80}^{4} Optical Center.

    B. Hourahine, R. Jones, A. N. Safonov, S. Öberg, P. R. Briddon, and S. K. Estreicher. *Physical Review B* 61, 12594 (2000).*

11. Weakly bound carbon-hydrogen complex in silicon.

    L. Hoffmann, E. V. Lavrov, B. Bech Nielsen, B. Hourahine, R. Jones, S. Öberg, and P. R. Briddon. *Physical Review B*, 61, 16659 (2000).*

List of tables

List of figures

Common abbreviations and acronyms

AB	Anti-bonding site between the atom site and a T_d site
AIMPRO	*Ab initio* modeling program
B3LYP	3 parameter Becke gradient corrected Lee-Yang-Parr density functional
BC	Bond-centred site between two neighbouring atoms
BLYP	Becke gradient corrected Lee-Yang-Parr local density functional
BPW91	Becke gradient corrected 1991 Perdew-Wang local density functional
C	The site between two next-nearest neighbouring atoms
CI	Configuration interaction
CNDO	Complete neglect of diatomic overlap
DCI	Double-excited configuration interaction
DFT	Density functional theory
DLTS	Deep level transient spectroscopy
DMC	Diffusion Monte Carlo
ENDOR	Electron-nuclear double optical resonance
EPR	Electron paramagnetic resonance
E_{xc}	Exchange correlation energy
FCC	Face-centred cubic lattice
FFT	Fast Fourier transform
GGA	Generalised gradient approximation
HF	Hartree-Fock
INDO	Intermediate neglect of diatomic overlap
IR	Infrared
KS	Kohn-Sham
LDA	Local density approximation
LDF	Local density functional
MCTS	Minority carrier transient spectroscopy

MD	Molecular dynamics
MINDO	Modified intermediate neglect of diatomic overlap
MNDO	Modified neglect of diatomic overlap
MP-n	n-th level Møller–Plesset perturbation theory
MP-n^3	$n \times n \times n$ Monkhorst-Pack k-point sampling
NDDO	Neglect of differential diatomic overlap
NMR	Nuclear magnetic resonance
ODLTS	Optically detected deep level transient spectroscopy
PL	Photoluminescence
PRDDO	Partial retention of differential diatomic overlap
PW91	1991 Perdew-Wang local density functional
QMC	Quantum Monte Carlo
RHF	Spin restricted Hartree-Fock
SC	Simple cubic lattice
SDCI	Single and double-excited configuration interaction
SDTCI	Single, double and triple-excited configuration interaction
SIMS	Secondary ion mass spectrometry
SVWN	Spin-polarised Vosko-Wilk-Nusair local density functional
T_d	Tetrahedral interstitial site surrounded by four equivalent atoms
TEM	Transmission electron microscopy
TF	Thomas-Fermi
UHF	Spin unrestricted Hartree-Fock
VMC	Variational Monte Carlo
VWN	Vosko-Wilk-Nusair local density functional
ZDO	Zero diatomic overlap
ZPL	Zero phonon line

Chapter 1

Introduction

This thesis contains the results of a set of first-principles simulations for several hydrogen-related centres in silicon. The question of what "first-principles" or *ab initio*[1], actually means is debatable. For the purposes of this work it is taken to be that, in principle, only the chemical composition (the number and type of species in a system), needs to be known to simulate that system. In practice experimental information on the structure of the system is gratefully received, since this can greatly curtail the set of configurations to be investigated. This paradigm contrasts with the empirical or semi-empirical approach where a variable number of parameters are adjusted to fit some observed properties of a calibration system. These parameterised models are then used on other problems, but in practice the nagging worry that the parameterisation will fail for some cases remains.

The particular implementation used for these calculations is the AIMPRO code, in both its cluster and periodic supercell incarnations. The many-body behaviour of the electrons in the simulated systems is considered using the density functional theories of Hohenberg–Kohn and Kohn–Sham, within the local density approximation. The computation effort in calculating such systems is reduced by a combination of the use of pseudopotentials and a localised Gaussian basis with a novel fitting procedure.

Hydrogen is one of the most important and ubiquitous impurities in silicon, and is introduced at almost every stage and step of manufacture and processing of commercial material. Water impurities lead to incorporation of H during growth or processing steps such as plasma and chemical etching or the Smart Cut® process [1] rely on hydrogen for their effectiveness. Additionally, hydrogen is directly included into silicon by plasma based etching or passivation methods. Amorphous, polycrystalline material and CVD grown silicon contain substantial amounts of hydrogen

[1] From the beginning

due to the growth processes used, and for specific applications such as amorphous solar-cells further hydrogen is added to improve the material.

The surrounding lattice modifies the properties of this simplest of atoms tremendously, leading to somewhat different behaviour from when it is present in small molecules. In addition to the Si–H and H–H bonds familiar in chemistry, the lattice leads to new structures such as the H_2^* dimer (see chapter 7 for a description of this defect), or hydrogen bonded to pairs of other atoms.

Monoatomic hydrogen in silicon is a fast diffusing species and is also chemically active, so readily travels to and binds with many other defects. The resulting complexes are often electronically inactive, but this does not always follow. In some cases such as transition metals, new electronic levels are are introduced into the forbidden bandgap, or existing levels are displaced. The passivation of defects also leads to change in the electronic properties of the material, for example passivating part of the doping in the material, leading to a lower carrier concentration. Additionally hydrogen has a catalytic role in processes such as oxygen diffusion, leading to substantially increased mobility of this impurity. The presence of molecular hydrogen in silicon leads to a large reservoir of "hidden" hydrogen in the material which is mobile at room temperature and above. This can then act as an unexpected source of hydrogen atoms to processes in the material.

The wide ranging properties of hydrogen have led to a spectrum of experimental techniques being applied to the study of related defects. For example, one of the first applications of the recently developed Laplace deep level transient spectroscopy (LDLTS) method was to study transition-metal–hydrogen complexes where conventional DLTS could not resolve signals from many distinct defects. Both Fourier transform infrared-spectroscopy (FTIR) and electron paramagnetic resonance (EPR) have been particularly useful for the study of both proton implanted material and also silicon treated by indiffusion of hydrogen gas at high temperature, leading to an understanding of the bonding and symmetry of defects.

The wide variety of of chemistry involved in hydrogen defects has also provided an important proving ground for modern theoretical methods. To successfully describe the behaviour of a defect, not only must methods be able to simulate the chemical bonding between the atoms to sufficient accuracy, but also the calculations should be efficient enough to allow sufficient of the surrounding lattice to be modeled, allowing for effects such as long range strain fields and polarisation of the host. This has lead to a drive towards parallel computational codes which can treat several hundred atoms routinely to an accuracy of meV per atom, while also exposing some of the short comings of many methods. For example, as discussed in chapter 5, all of the theoretical methods so far applied to the interstitial hydrogen molecule in silicon have been unable to produce the fine details of the experimentally observed spectra, though whether this is a failing in the theory or the model for the defect remains unclear. On a more positive note, theory has been successful in helping to resolve experimental puzzles such as the structure of anomalous muonium in silicon, and continues to act as a test for defect models such as those discussed in chapter 8.

The layout of the thesis is as follows. In chapter 2 the problems of simulating quantum mechanically many-particle systems are discussed, and the customary solutions based on the Hartree independent electron model and the Hartree-Fock corrections are discussed. The alternative charge-density based approach of Hohenberg and Kohn is introduced and the Kohn-Sham equations are discussed. Chapter 3 presents the formalism used in this work as implemented in the AIMPRO code. Short descriptions of several experimental methods which have been used in studying the defects discussed in this thesis are presented in chapter 4. The succeeding chapters then present the results of simulations performed on selected defects.

Chapter 5 is concerned with the behaviour of the hydrogen molecule within the silicon lattice and its interaction with oxygen impurities. The recent experimental observation(s) of the molecule are discussed, and observed vibrational modes and the diffusion barrier are compared with the results of simulation. Chapter 6 presents results upon the properties of hydrogen pairs at open defects in the lattice. It is shown that the behaviour of hydrogen molecules within voids or platelet structures in silicon is consistent with the data which were used to provide the first spectroscopic observation of the molecule in silicon. The behaviour of hydrogen interacting with the stable ring hexavacancy in silicon is also considered in this chapter, leading to an assignment for a family of bound-exciton centres observed by photo-luminescence as complexes between hydrogen and V_6, this allows an identification of the B_{80}^{4} centre to be made as the hexavacancy. Chapter 7 suggests that in addition to the two known hydrogen dimers in silicon, the molecule and the H_2^* pair, a third dimeric structure, H_2^{**}, is present in silicon at low temperatures. Chapter 8 discusses the continuing debate as to whether a prominent family of infrared active centres are related to interstitial- or vacancy-hydrogen centres. The vacancy related forms of these defects has already been simulated and agree well with the observed spectra, but this has not prevented further argument, the interstitial *alternative* defects are therefore considered and shown to be inconsistent with the observed data. As a final set of results, the interaction of a single substitutional carbon with hydrogen is considered in chapter 9. There have been two distinct sets of experimental observations on this system, the electrical levels observed which are considered to be consistent with the hydrogen sitting next to the carbon, but the observed infrared modes match a defect where the hydrogen is further away. The properties of these systems are simulated, and the resulting vibrational and electronic data is used to suggest a reconciliation between the two sets of observations.

Chapter 2

The many body problem

To accurately simulate the properties of real molecules and solids at a level which both qualitative and quantitative predictions can be made with confidence, requires an excellent representation of the behaviour of the components of such systems. While classical approximations using simple parametrized models, such as interatomic potentials [2] or bond-charge models [3] have been employed in the past to study problems involving collections of atoms with some success[2], quantum mechanical methods have (fairly recently) become well established in the study of such systems.

Successful non-relativistic simulation of the electronic and structural properties of objects containing up to about 1,000 atoms [5, 6] have become almost routine in the last few years. Such systems require an accurate solution to a Schrödinger-like, many-particle equation.

The Schrödinger equation itself is easily constructed for arrangements of many particles, but is impossible to directly solve for anything beyond the simplest of systems without some approximations. A large range of choices for these necessary simplifications have been provided by various workers. Several of the most popular (and their antecedents) are briefly discussed in this chapter.

2.1 The many-particle Schrödinger equation

The non-relativistic, time independent, many-particle equation for a distribution of electrons and nuclei can be written in the general form

$$\{\mathscr{H} - E\}\,\Psi(\mathbf{r},\mathbf{R}) \;=\; \{T_{\mathrm{K.E.}} + V_{\mathrm{e-e}} + V_{\mathrm{e-n}} + V_{\mathrm{n-n}} - E\}\,\Psi(\mathbf{r},\mathbf{R}) = 0\ , \quad (2.1)$$

with E being the many-body energy of the system, $T_{\mathrm{K.E.}}$ the kinetic energy of both the electrons and nuclei in the system, $V_{\mathrm{e-e}}$ the potential energy due to electron-

[2]Systems consisting of millions of atoms [4] are currently simulated using such methods.

electron interactions, $V_{\mathrm{e-n}}$ the electron nuclear interactions, and $V_{\mathrm{n-n}}$ the nuclear-nuclear interactions. The many-body wavefunction, $\Psi(\mathbf{r}, \mathbf{R})$ then depends on both the positions and spins of all N nuclei and n electrons in the system.

Ignoring effects such as spin-orbit coupling, and spin–spin interactions, it is possible to explicitly write out all of the potentials in equation 2.1.

In atomic units, $\hbar = e = m_{\mathrm{e}} = 4\pi\epsilon_0 = 1$, for a normalised distribution of electrons at the set of coordinates $\mathbf{r}_i$ and nuclei of charges Z_a at coordinates $\mathbf{R}_a$, these terms become

$$\begin{aligned}
T_{\mathrm{K.E.}} &= -\frac{1}{2}\left\{\sum_{\mu} \nabla_\mu^2 + \sum_a \frac{1}{M_a}\nabla_a^2\right\} \\
V_{\mathrm{e-e}} &= +\frac{1}{2}\sum_{\mu,\nu(\neq\mu)} \frac{1}{|\mathbf{r}_\mu - \mathbf{r}_\nu|} \\
V_{\mathrm{e-n}} &= -\sum_{\mu,a} \frac{Z_a}{|\mathbf{r}_\mu - \mathbf{R}_a|} \\
V_{n-n} &= +\frac{1}{2}\sum_{a,b(\neq a)} \frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|} \ .
\end{aligned}$$

The kinetic energy ($T_{\mathrm{K.E}}$) is summed over all of the electrons ($\mu = 1 \rightarrow N$) and nuclei ($a = 1 \rightarrow n$). The electron–electron and nuclear–nuclear potentials are summed over each distinct pair combination, hence the $\frac{1}{2}$ to remove double counting.

2.2 The Born–Oppenheimer approximation

As written, the Schrödinger equation is intractable, but a simplification due to Born and Oppenheimer [7] has been used in almost all attacks on this problem. They noted that since the ratio of the masses of the electron and atomic nuclei is so large, the energy scales on which electronic and nuclear motion occurs are distinct[3].

The decoupling of the nuclear and electronic motion can be used to write an approximate Schrödinger equation for the motion of electrons in solids and molecules.

[3]It is interesting to note that caution should be exerted in using the Born–Oppenheimer approximation for vibro-rotary motion of molecules (see chapter 5), since molecular rotation causes a "centrifugal distortion" which manifests itself in the vibrational spectrum.

If $\Psi(\mathbf{r}, \mathbf{R})$ written as the product

$$\Psi(\mathbf{r}, \mathbf{R}) \quad = \quad \Psi_{\mathbf{R}}(\mathbf{r}) \cdot \chi(\mathbf{R}) \,, \tag{2.2}$$

where $\Psi_{\mathbf{R}}(\mathbf{r})$ is a function of the distribution of the electrons for a fixed nuclear arrangement and $\chi(\mathbf{R})$ is a function which modulates the wavefunction dependent on nuclear position, then $\Psi_{\mathbf{R}}(\mathbf{r})$ satisfies a wave-equation of the form

$$\left\{ -\frac{1}{2} \sum_{\mu} \nabla_{\mu}^2 + V_{\mathrm{e-e}} + V_{\mathrm{e-n}} + V_{\mathrm{n-n}} \right\} \Psi_{\mathbf{R}}(\mathbf{r}) \quad = \quad U(\mathbf{R}) \Psi_{\mathbf{R}}(\mathbf{r}) \,. \tag{2.3}$$

So for each nuclear distribution, $U(\mathbf{R})$ is part of a set of eigenvalues for different electronic states. $\Psi_{\mathbf{R}}$ and U must be continuous functions of $\mathbf{R}$, particularly for the case of degenerate systems.

Substituting equation 2.2 into the Schrödinger equation, and considering equation 2.3 leads to

$$\left\{ -\frac{1}{2} \sum_{n} \frac{1}{M_n} \nabla_n^2 + U(\mathbf{R}) \right\} \Psi(\mathbf{r}, \mathbf{R}) \quad = \quad E \Psi(\mathbf{r}, \mathbf{R}) \,, \tag{2.4}$$

which can then be written as

$$\Psi_{\mathbf{R}}(\mathbf{r}) \left\{ -\frac{1}{2} \sum_{n} \frac{1}{M_n} \nabla_n^2 + U(\mathbf{R}) - E \right\} \chi(\mathbf{R}) =$$
$$\sum_{n} \frac{1}{2M_n} \left\{ \chi(\mathbf{R}) \nabla_n^2 \Psi_{\mathbf{R}}(\mathbf{r}) + 2 \nabla_n \chi(\mathbf{R}) \cdot \nabla_n \Psi_{\mathbf{R}}(\mathbf{r}) \right\} \,, \tag{2.5}$$

where the right hand side of equation 2.5 is zero if $\Psi_{\mathbf{R}}(\mathbf{r})$ only depends parametrically on $\mathbf{R}$.

This will then leave an approximate Schrödinger equation for the motion of the nuclei, which is decoupled from the electron motion, of the form

$$\left\{ -\sum_{n} \frac{1}{2M_n} \nabla_n^2 + U(\mathbf{R}) \right\} \chi(\mathbf{R}) \quad = \quad E \chi(\mathbf{R}) \,. \tag{2.6}$$

Taken together with equation 2.3 this allows separate descriptions for the electronic and nuclear motion.

2.3 The Hartree approximation

Having simplified $\mathscr{H}$ by writing it as a function of the fixed distribution of nuclei, the Schrödinger equation is still intractable due to the presence of the $V_{\mathrm{e-e}}$ potential. Clearly some form of simplifying approximation is needed for this term.

Hartree [8] proposed that the electronic wavefunction could be written such that the electrons behaved as a collection of orthogonal single-particle states experiencing an interaction with both the fixed nuclei and an average of *all* of the other electrons in the system. These electronic states must also be self-consistent within this averaged field, i.e., the field due to all of the electrons and nuclei is identical to that acting to produce that distribution of electrons.

This approximation is tantamount to writing the many-body wavefunction as

$$\Psi_{\mathbf{R}}(\mathbf{r}) \quad = \quad \Pi_i \psi_i(\mathbf{r}_i)\ , \tag{2.7}$$

where $\psi(\mathbf{r}_i)$ are a set of independent single particle wavefunctions, which are functions of position and spin for the n electrons.

From this *ansatz* wavefunction it is then possible to write a Hamiltonian embodying this approximation. For a normalised set of wavefunctions the Hamiltonian for the energy of state ψ_i is

$$\left\{ -\frac{1}{2}\nabla_i^2 - \sum_n \frac{Z_n}{|\mathbf{r}_i - \mathbf{R}_n|} + \sum_{j(\neq i)} \int |\psi_j(\mathbf{r}_j)|^2 \frac{\mathrm{d}\mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|} \right\} \psi_i(\mathbf{r}_i) \quad = \quad \mathcal{E}_i \psi_i(\mathbf{r}_i). \tag{2.8}$$

This then gives a set of n simultaneous integro-differential equations to solve self-consistently for the set of n ψ_i states.

The customary method of solution to this problem is by an iterative approach in which the potential energy approximating the second and third terms in equation 2.8 is assumed, from which the single particle wavefunctions are then calculated. These eigenfunctions then give a new potential from the new total charge distribution, which is again used to re-evaluate yet another set of wavefunctions. The process is repeated until the functions remain unchanged between iterative cycles (or differ below a certain tolerance).

It is straightforward to show that the Hartree wavefunction is the lowest energy solution possible for the Hamiltonian described by using the variational principle (Raleigh ratio) [9], which states that for a wavefunction of energy

$$E \quad = \quad \frac{\langle \Psi | \mathscr{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{2.9}$$

the lowest possible energy is given when

$$
\begin{aligned}
\frac{\delta E}{\delta \Psi} &= 0 \\
\frac{\delta^2 E}{\delta^2 \Psi} &> 0
\end{aligned}
\tag{2.10}
$$

The solutions to the Hartree equation have one very obvious failing, they are not anti-symmetric on exchange of electrons, so they cannot represent Fermionic systems without some adjustment. The Pauli exclusion principle [10] can be tacked on *ad hoc* by forcing an occupancy of the one electron levels to give unique quantum numbers for each of the n electrons in the system, but this is clearly unsatisfactory.

There are several well known problems with this method for finding the electronic wavefunction. Calculation of excited states using this approach are fraught with problems. The naïve approach of taking the non-ground state functions of the n one electron equations literally as excited many-body states (Koopmans' theorem [11]) fails. This is due to the lack of a "relaxation" correction to the energy of all of the other states in the system, since they would re-adjust to the new Hartree potential due to this change in the electron distribution. This leads to the realization that the only meaningful states in the system are those which are occupied. A more accurate method of considering transitions between states is due to Slater [12]. If the energy of a system which is experiencing a transition from $|\cdots, 1, \cdots, 0, \cdots\rangle \rightarrow |\cdots, 0, \cdots, 1, \cdots\rangle$ is considered to be a continuous function of the change in occupancy between the two levels, then the energy of the system can be expressed as a Taylor expansion about $|\cdots, \frac{1}{2}, \cdots, \frac{1}{2}, \cdots\rangle$. The difference in the total energy between the two configurations is then, up to the third order, related to the derivative of the energy with respect to the occupancy of the two levels[4].

In the case of solids, if the Hartree approximation is applied to a jellium system then the Somerfield independent electron model for metals results. Jellium is a good approximation for conduction electrons in simple metals such as sodium, and consists of a uniform positive charge spread over the whole volume of the solid. This positive charge typically exactly balances the negative charge of the n electrons in

[4] As we shall see later, this can be linked directly to the energy of Kohn-Sham energy levels via Janak's theorem, as discussed in section 2.5.3.

the system. The jellium solid is usually taken to be infinite in extent, removing the requirement to consider surface effects.

For neutral jellium the Somerfield model gives no net cohesive energy for the electron, i.e., the electrons are not bound to the solid in this simple model. This lack of cohesive energy can be traced back to the independent electron approximation. More sophisticated approximations, which include more accurate representations of the collective electron behaviour, do not suffer from this effect.

2.4 The Hartree-Fock approximation

For Fermionic systems, the approximate wavefunction of Hartree (eqn. 2.7) lacks some relevant physics, namely the Pauli exclusion principle. Both Fock [13] and Slater [9] suggested a method of including this by writing the many-body wavefunction as a set of single-particle states that are explicitly anti-symmetric with respect to exchange of particles.

For a set of n particles this wavefunction can be written as a normalised Slater determinant [9] of the form

$$\Psi(\mathbf{r}_1 \cdots \mathbf{r}_n) = \frac{1}{\sqrt{n!}} \det |\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \cdots \psi_z(\mathbf{r}_n)| \ , \tag{2.11}$$

where $\psi_a(\mathbf{r}_i)$ is the a-th orthonormal single-particle spin-orbital at spatial and spin coordinates $\mathbf{r}_i$. This determinant then guarantees that the overall wavefunction is anti-symmetric with respect to exchange of the particles.

The single particle states are then orthogonal with respect to both spatial distribution (i, j) and spin (α, β), i.e.,

$$\langle \psi_i^\alpha | \psi_j^\beta \rangle = \delta_{ij}\delta_{\alpha\beta} \ , \tag{2.12}$$

with i being the spatial dependency, and α the spin, leading to the familiar Fermionic level filling.

By applying the variational principle to this wavefunction it is possible to derive an appropriate Hamiltonian for the Hartree-Fock approximation.

Starting from

$$E = \left\langle \Psi \left| \sum_i h_i + \frac{1}{2} \sum_{i,j(\neq i)} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right| \Psi \right\rangle \tag{2.13}$$

where

$$h_i \quad = \quad -\frac{1}{2}\nabla_i^2 - \sum_n \frac{Z_n}{|\mathbf{r}_i - \mathbf{R}_n|} \ . \tag{2.14}$$

Using the indistinguishability of the electrons in the determinant, this energy becomes

$$\left\langle \Psi \left| \sum_i^n h_i \right| \Psi \right\rangle \quad + \quad \frac{1}{2} n(n-1) \left\langle \Psi \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \Psi \right\rangle \ , \tag{2.15}$$

Expanding the second term in equation 2.15 as a set of spin-orbitals from the HF determinant gives

$$\frac{1}{2} \sum_{i,j(\neq i)} \left\langle \psi_i(1)\psi_j(2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \{\psi_i(1)\psi_j(2) - \psi_j(1)\psi_i(2)\} \right\rangle \tag{2.16}$$

which can be rewritten as

$$\begin{aligned} &\sum_i \langle \psi_i(1)|h_i|\psi_i(1)\rangle \quad + \\ \frac{1}{2} \sum_{i,j(\neq i)} \Bigg\{ &\left\langle \psi_i(1)^*\psi_i(1) \left| \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right| \psi_j(2)^*\psi_j(2) \right\rangle \quad - \\ &\left\langle \psi_i(1)^*\psi_j(1) \left| \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right| \psi_j(2)^*\psi_i(2) \right\rangle \Bigg\} \end{aligned} \tag{2.17}$$

If the energy is then minimized subject to the orthonormality of the single-particle states, then with the introduction of a set of undetermined multipliers this becomes

$$\delta E - \sum_{i,j} \lambda_{ij} \left\{ \langle \delta\psi_i|\psi_j\rangle + \langle \psi_i|\delta\psi_j\rangle \right\} \quad = \quad 0 \ . \tag{2.18}$$

This is satisfied when

$$\left[h_1 + \sum_j \{J_j(1) - K_j(1)\} \right] \psi_i(1) \quad = \quad \sum_j \lambda_{ji}\psi_j(1) \tag{2.19}$$

where the Coulomb operator, J_u, and the exchange operator, K_u, are defined as

$$J_u(1)\psi_a(1) \quad = \quad \left\langle \psi_u(2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \psi_u(2) \right\rangle \psi_a(1) \tag{2.20}$$

$$K_u(1)\psi_a(1) \quad = \quad \left\langle \psi_u(2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \psi_a(2) \right\rangle \psi_u(1) \ . \tag{2.21}$$

If we then define the Fock operator as

$$\mathscr{F}_1 = h_1 + \sum_u \{J_u(1) - K_u(1)\} \tag{2.22}$$

then

$$\mathscr{F}_1 \psi_i(1) = \sum_j \lambda_{ji} \psi_j(1). \tag{2.23}$$

If this equation is then transformed by the introduction of new set of canonical spin orbitals $\psi_i'(n)$, such that the form of $\mathscr{F}_1$ is unchanged and λ_{ji} becomes a diagonal matrix, then the familiar Hartree-Fock equation results

$$\mathscr{F}_1 \psi_i'(1) = \mathcal{E}_i' \psi_i'(1) \; . \tag{2.24}$$

Dropping the primes then gives

$$\mathscr{F}_1 \psi_i(1) = \mathcal{E}_i \psi_i(1) \; , \tag{2.25}$$

i.e., the conventional form for this expression.

The set of n spin-orbitals are then evaluated self-consistently, as with the Hartree wavefunctions, since from the definition of the Fock operator in equation 2.22, the solution for a particular spin-orbital is dependent on the other $n-1$ occupied states in the system.

The exchange anti-symmetry that the resulting wavefunctions possesses leads to an "exchange hole" of net unit charge around each electron in the system, where all of the other electrons of parallel spin are repelled. This exchange potential for the a-th electron in the system is of the form

$$V_{\mathrm{x}}^a(\mathbf{r}_1, \mathbf{r}_2) = \frac{-1}{|\mathbf{r}_1 - \mathbf{r}_2|} \sum_{b(\neq a)} \langle \psi_b(\mathbf{r}_1) | \psi_b(\mathbf{r}_2) \rangle \tag{2.26}$$

where the sum is over all of the other electrons of parallel spin. It is this exchange term which then will give a net cohesive energy for jellium.

From this exchange energy, the exchange energy E_x can be defined for a spin-average uniform electron gas as a function of the charge density such that

$$E_{\mathrm{xc}} = \int \rho(\mathbf{r}) \epsilon_{\mathrm{xc}}[\rho(\mathbf{r})] \mathrm{d}\mathbf{r} \; , \tag{2.27}$$

with

$$\epsilon_x(\rho) \;=\; -\frac{3}{2}\left(\frac{3\rho}{8\pi}\right)^{\frac{1}{3}} . \tag{2.28}$$

Similarly the spin polarised exchange energy is given by

$$E_x(\rho_\uparrow, \rho_\downarrow) \;=\; -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\left(\rho_\uparrow^{4/3} + \rho_\downarrow^{4/3}\right) . \tag{2.29}$$

2.4.1 The Roothaan equations

When applied to molecules or solids, in contrast to isolated atoms, numerical solution of the set of Hartree-Fock equations becomes infeasible. Roothaan and Hall suggested a reformulation of the HF equations by using of a set of basis functions to expand the spatial dependency of the spin-orbitals, thus allowing the problem to be numerically solved by matrix diagonalization [14].

If the spatial part of the orbitals is represented as a sum of M basis functions, i.e.,

$$\psi_i \;=\; \sum_j^M c_{ji}\theta_j , \tag{2.30}$$

where θ_j is one of the M orbitals weighted by c_{ji}, then the set of M basis functions allows the expression of up to M linearly independent spatial wavefunctions. Substituting this definition into equation 2.25 then gives

$$\mathscr{F}_1 \sum_j^M c_{ji}\theta_j(1) \;=\; \mathcal{E}_a \sum_j^M c_{ji}\theta_j(1) \tag{2.31}$$

which can be transformed into

$$\sum_j^M c_{ji}\left\langle \theta_i(1)|\mathscr{F}_1|\theta_j(1)\right\rangle \;=\; \mathcal{E}_a \sum_j^M c_{ji}\left\langle \theta_i(1)|\theta_j(1)\right\rangle . \tag{2.32}$$

If overlap ($\mathbf{S}$) and Fock ($\mathbf{F}$) matrices of the form

$$S_{ji} \;=\; \left\langle \theta_i(1)|\theta_j(1)\right\rangle \tag{2.33}$$

$$F_{ji} \;=\; \left\langle \theta_i(1)|\mathscr{F}_1|\theta_j(1)\right\rangle \tag{2.34}$$

are introduced, then the Roothaan matrix equation becomes

$$\mathbf{Fc} \;=\; \boldsymbol{\mathcal{E}}\mathbf{Sc} , \tag{2.35}$$

with $\mathbf{c}$ being an $M \times M$ matrix, and $\boldsymbol{\mathcal{E}}$ being a diagonal $M \times M$ matrix of the orbital energies. This equation possesses a non-trivial solution only in the case of

$$\det |\mathbf{F} - \mathcal{E}_a \mathbf{S}| \;=\; 0 \tag{2.36}$$

which again requires a self-consistent solution, but couched in terms of the adjustable mixing coefficients c_{ji}. The resulting set of $\boldsymbol{\mathcal{E}}$ energies consist of n occupied orbitals, and $(M - n)$ virtual orbitals. If M were an infinite set of basis functions, this set of virtual orbitals would give the full set of possible unoccupied Hartree-Fock wavefunctions.

If the form of the F_{ji} elements is then written in full, they become

$$F_{ji} \;=\; h_{ji} + \sum_{l,m} b_{lm} \left\{ 2(ij|lm) - (im|lj) \right\} \tag{2.37}$$

with the density matrix elements $\mathbf{b}$ being defined as

$$b_{lm} \;=\; \sum_{u} c^{*}_{lu} c_{mu} \;, \tag{2.38}$$

and the terms of the form $(ab|cd)$ being two electron, four centre, integrals over the basis functions, i.e.,

$$(ab|cd) \;=\; \int \theta^{*}_{a}(1)\theta_{b}(1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \theta^{*}_{c}(2)\theta_{d}(2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \;. \tag{2.39}$$

These equations have been extended to the spin polarized case by Pople and Nesbet [15].

The initial determination of the four electron integrals of the form $(ab|cd)$, combined with the requirement of re-evaluating the set of $\mathbf{b}$ density matrix elements, are the main computational bottle necks for this method, leading to a calculation that scales as $\mathcal{O}(n^4)$.

2.4.2 Approximations to the Hartree-Fock method

Due to this large computational effort, there is an incentive to simplify the HF calculations in some way. A large number of approximations have been suggested by various workers to simplify this problem in some way.

As described, the spin of all n electrons in the system are considered explicitly, this is referred to as an unrestricted open-shell Hartree-Fock (UHF) calculation.

In the case of "closed shell" systems where the electrons are all in fully occupied orbitals with a total many-body symmetry of Γ_1, if we were to assume that the behaviour of the two electrons in an orbital shared the same spatial distribution, then the calculation can be simplified by considering a system containing half the number of particles, each of which possesses a charge of -2. Such restricted HF (RHF) would be known as spin averaged to DFT practitioners (see section 2.5). Alternatively a hybrid of the two methods, namely restricted open-shell HF, where the full shells are treated as with RHF and any open shell states are treated as in UHF, is also used.

Many of the alternative methods to simplify the HF equations rely on some form of parameterisation to the four-centre overlap integrals, leading to a class of semi-empirical methods. The simplest parameterisation of this form is the zero differential overlap approximation (ZDO) [16] or complete neglect of differential overlap (CNDO), where the four centre overlap integrals $(ab|cd)$ are taken to be non-zero only in the case that $a = b$ and $c = d$. The integral $(aa|cc)$ is then usually taken to be an empirical parameter fitted to give agreement with a set of Hartree-Fock calculations.

A slightly more sophisticated version of this approximation is the intermediate neglect of differential diatomic overlap (INDO). In this approximation the four-centre exchange integrals are retained if they centred on the same atom and of the form $(ij|ij)$ (which is equal to $(ij|ji)$ for real basis functions). Pople [17] developed an improved form of INDO, known as neglect of diatomic differential overlap (NDDO). With NDDO, functions of the form $(im|li)$ are retained when they are centred on the same atom, and two centre-functions are retained in the case $(ab|cd)$ where a and b are centred on one atom and c and d on another atom.

Several variants of NDDO approximation have been developed by Dewar, leading to the modified intermediate neglect of differential diatomic overlap (MINDO) approximation, which increase in sophistication from MINDO/1 to MINDO/3 [18].

An alternative "neglect" approximation is the partial retention of differential diatomic overlap (PRDDO), where most of the overlap integrals are removed by a careful use of orthogonality between states [19].

2.4.3 Improvements on the Hartree-Fock approximation

For the jellium system mentioned in section 2.3, the HF approximation has several notable failings. Principle amongst these is the density of states at the Fermi-level. If we represent the electronic states as a set of plane-waves then the energy of jellium filled to the Fermi-wavevector k_{F} is [20]

$$E = \sum_{k}^{k_{\mathrm{F}}} k^2 - \frac{k_{\mathrm{F}}}{\pi} \sum_{k}^{k_{\mathrm{F}}} \left\{ 1 + \frac{k_{\mathrm{F}}^2 - k^2}{2k_{\mathrm{F}}k} \ln \left| \frac{k_{\mathrm{F}} + k}{k_{\mathrm{F}} - k} \right| \right\}. \tag{2.40}$$

The most obvious problem with this expression is that the density of states, which is given by

$$g(E) = \frac{\partial k}{\partial E},$$

is zero at the Fermi energy, leading to an unphysical model for simple metals [21].

Of more relevance to this thesis, structural parameters of defects in silicon such as $\mathrm{C_sH_i}$ have been shown to be incorrectly predicted by pure HF theory (see chapter 9).

The deficiencies of HF theory can be traced back to the assumption that each electron in the system experiences an average potential due to the other electrons in the system, i.e., the electronic behaviour lacks correlation. Since there is allowance for exchange in this method, the electrons with parallel spin are in some respects correlated, but the anti-parallel case is not accounted for.

Whilst it might be possible to introduce a correlation function into the many-body wavefunction in a similar manner to the exchange potential, there is no obvious form for this term(s) [22], instead the methods used to introduce correlation rely on considering the corrections due to "excited states" of the one-electron determinants.

2.4.3.1 Configuration interaction

If the ground state determinant from the Roothaan equations (section 2.4.1) is taken as a reference, adopting the notation

$$\Psi_0 = |\psi_1\psi_2 \cdots \psi_x \cdots \psi_{n-1}\psi_n| \tag{2.41}$$

for the "ground" state, and

$$\Psi_x^e = |\psi_1\psi_2 \cdots \psi_x^e \cdots \psi_{n-1}\psi_n| \tag{2.42}$$

for a determinant in which the x-th single-particle state has been excited into the e-th level. This allows excited determinants to be written as

$$\Psi_{xyz\cdots}^{efg\cdots}\,, \tag{2.43}$$

for a determinant in which the x-th, y-th, and z-th single-particle states have been excited into their e-th, f-th, and g-th states respectively.

Following the discussion reviewed in reference [23], since every normalizable antisymmetric wavefunction can be represented as a sum of Slater determinants built up from a complete set of one-electron functions, then the correlated ground state can be found by optimising the set of *all* Slater determinants. If the many-body wavefunction is written as the set of *all* of the possible excited and ground state determinants, i.e.,

$$\Psi_{\mathrm{CI}} \;=\; C_0\Psi_0 + \sum_{a,i} C_a^i \Psi_a^i \cdots + \sum_{ab,ij} C_{ab}^{ij}\Psi_{ab}^{ij}\cdots + \cdots \tag{2.44}$$

then by optimising both the one-electron and one-determinant parameters c and C it is possible to find the "exact" non-relativistic energy of the system.

The energy difference between this "exact" energy and that of the ground-state HF determinant is the correlation energy for the system. In practice including an infinite set of determinants each with an infinite basis set is impossible. The number of determinants to be included can be pruned back somewhat, since only the functions with the same symmetry as the ground state will form part of the ground state function, i.e., for the H_2 molecule only determinants of symmetry ${}^1\Sigma_g^+$ need to be included. This is due to the well known result that integrals of the form

$$\langle \Psi_i | \Psi_j \rangle \tag{2.45}$$

are only non-zero in the case that the products of the irreducible symmetries of the functions contains the totally symmetric representation, and since the Hamiltonian is always of symmetry Γ_1, this means that $\Gamma_{\Psi_i} = \Gamma_{\Psi_j}$ for the integral to be non-zero.

In practice, this still leaves an infinite set of functions to consider, therefore "full CI" is the best energy which can be evaluated for a given finite set of basis functions. Alternatively "limited CI" can be performed by considering only a subset of the possible determinants for a given basis set. Of these limited CI calculations the common types are the single, double, and triple excited CI (SDTCI), and similarly the SDCI and DCI forms. Such limited CI suffers from one major flaw, since it is not "size consistent", i.e., the energy of molecules A and B are different when calculated separately versus when in the same calculation but with an infinite separation between the molecules.

2.4.3.2 Møller–Plesset perturbation theory

Since typical limited CI calculations routinely involve 10^8 or more configurations even for small problems, the incentive to use a simplified method for including correlation corrections is present. If the HF wavefunction is similar to the "correct" correlated wavefunction then perturbation theory can be applied to improve the correspondence. In Møller–Plesset perturbation theory, the zeroth order Hamiltonian is taken to be

$$\mathscr{H}^{(0)} = \sum_{i=1}^{n} f_i \tag{2.46}$$

where f_i is the one-electron Fock operator. The energy is then

$$E_{\mathrm{SCF}} = E^{(0)} = \langle \Psi_0 | \mathscr{H}^{(0)} | \Psi_0 \rangle \tag{2.47}$$

The first order perturbation is then

$$\mathscr{H}^{(1)} = \mathscr{H} - \sum_{i=1}^{n} \mathscr{F}_i \tag{2.48}$$

with $\mathscr{H}$ being the Born-Oppenheimer electronic Hamiltonian, and $\mathscr{F}$ the Fock operator. This gives a total energy of

$$E^{(0)} + \langle \Psi_0 | \mathscr{H}^{(1)} | \Psi_0 \rangle \tag{2.49}$$

It is only for second order perturbation theory (or better) that correlation corrections apply to the wavefunction. The perturbation in the energy for the second order is

$$E^{(2)} = \frac{1}{4} \sum_{abij} \frac{\{(ia|jb) - (ib|ja)\}^2}{\mathcal{E}_a + \mathcal{E}_b - \mathcal{E}_i - \mathcal{E}_j} \tag{2.50}$$

leading to a total energy of

$$E^{(0)} + E^{(1)} + E^{(2)} \tag{2.51}$$

The standard notation for a given level of Møller–Plesset perturbation is MP^n where n is the order of perturbation. The scaling with level of perturbation is poor, since the computational effort is $\mathcal{O}(N^{n+3})$, but unlike limited CI this approach does have the advantage of size consistency.

2.5 Density functional theory

Instead of considering the many-body problem in terms of a set of wavefunctions, considerable simplification can be achieved if the distribution of charge can be taken to be the fundamental variable. Before the Hohenberg-Kohn theorem (section 2.5.2), all such attempts represented only model systems, with historically the Thomas-Fermi model being the first of these.

2.5.1 Thomas-Fermi theory

Thomas [24] and Fermi [25] assumed that "Electrons are distributed uniformly in the six-dimensional phase space for the motion of electrons at a rate of two for each h^3 of volume", and that there is an effective potential which "is itself determined by the nuclear charge and this distribution of electrons". From these assumptions, jellium with a range of electron momenta at zero temperature of up to the Fermi level, i.e., k_F, possess an energy and charge density of

$$E_\mathrm{F} = \frac{k_\mathrm{F}^2}{2} \tag{2.52}$$

$$\rho = \frac{N}{V} = \frac{k_\mathrm{F}^3}{3\pi^2}. \tag{2.53}$$

This leads to an expression for the Fermi wavevector k_F in terms of the charge density

$$k_\mathrm{F} = \left(3\pi^2\rho\right)^{\frac{1}{3}} , \tag{2.54}$$

which then allows the electronic kinetic energy to be defined in terms for the charge density as

$$T = \int_0^{k_\mathrm{F}} \frac{3k^4}{2k_\mathrm{F}^3}\rho \, \mathrm{d}k. \tag{2.55}$$

If the electron density is then taken to be slowly varying, the kinetic energy at a particular point can be taken to be locally equal to the uniform gas case, i.e.,

$$T(\mathbf{r}) = \frac{3}{10}\left(3\pi^2\right)^{\frac{2}{3}}\rho(\mathbf{r})^{\frac{5}{3}}. \tag{2.56}$$

The total energy for the non-uniform electrons can then be written as a function of the kinetic energy, external potential on the electrons $V(\mathbf{r})$, and the electron–electron repulsion, i.e.,

$$\begin{aligned} E &= T + V_{\mathrm{e-V}} + V_{\mathrm{e-e}} \\ &= \frac{3}{10}\left(3\pi^2\right)^{\frac{2}{3}}\int \rho(\mathbf{r})^{\frac{5}{3}}\,\mathrm{d}\mathbf{r} - \int \rho(\mathbf{r})V(\mathbf{r})\,\mathrm{d}\mathbf{r} + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 \end{aligned} \tag{2.57}$$

If this energy is minimised with respect to the charge distribution (while keeping the same total charge), then the Thomas–Fermi (TF) equation results. This minimisation can be cast in the form of

$$\delta\left(E - \mu N\right) = 0\,, \tag{2.58}$$

with μ being the chemical potential of the set of N electrons. The resulting form for the TF equation is

$$\mu = \frac{1}{2}\left\{3\pi^2\rho(\mathbf{r})\right\}^{\frac{2}{3}} - V(\mathbf{r}) + \int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r}_1-\mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_2 \tag{2.59}$$

which can then be solved self-consistently, allowing the charge-density to be used as a fundamental variable.

2.5.2 The Hohenberg–Kohn theorem

The TF theory has two major failings, in that there is no physical justification for taking the charge density as the fundamental variable and the theory itself does not provide energies which are accurate enough to predict molecular structures correctly [26].

The first of these problems is somewhat alleviated by the work of Hohenberg and Kohn [27], who demonstrated that the charge density, for the non-degenerate case at least, uniquely determines the energy of the system (the theorem has subsequently been extended to the N representable degenerate case by Levy [28], see reference [29] for a review of this).

If we assume, for the moment, that the ground-state charge density is not uniquely determined by the potential acting on it, then it is possible to show that this assumption leads to an inconsistency. If two distinct external potentials V_1 and V_2 lead to many-body wavefunctions Ψ_1 and Ψ_2, which both have the same charge density, n, then from the variational principle

$$E_1 = \langle \Psi_1 | \mathscr{H}_{V_1} | \Psi_1 \rangle \leqslant \langle \Psi_2 | \mathscr{H}_{V_1} | \Psi_2 \rangle = \langle \Psi_2 | \mathscr{H}_{V_2} + V_1 - V_2 | \Psi_2 \rangle \ .$$

This then gives

$$E_1 \ \leqslant \ E_2 + \int [V_1(\mathbf{r}) - V_2(\mathbf{r})] n(\mathbf{r}) \, \mathrm{d}\mathbf{r} \ . \tag{2.60}$$

But similarly,

$$E_2 \ \leqslant \ E_1 + \int [V_2(\mathbf{r}) - V_1(\mathbf{r})] n(\mathbf{r}) \, \mathrm{d}\mathbf{r} \ , \tag{2.61}$$

then, by adding 2.60 and 2.61 this would give

$$E_2 + E_1 \ \leqslant \ E_1 + E_2, \tag{2.62}$$

which for a non-degenerate ground state becomes

$$E_2 + E_1 \ < \ E_1 + E_2 \ , \tag{2.63}$$

which is clearly impossible. Therefore the charge density must uniquely determine the potential experienced by the system, and hence the self-energy of the electrons.

Hohenberg and Kohn then defined an energy functional of charge density

$$F_{\mathrm{HK}}[\rho] \ = \ T[\rho] + V_{\mathrm{ee}}[\rho] \tag{2.64}$$

where

$$V_{\mathrm{ee}}[\rho] \ = \ J[\rho] + \mathrm{non-classical\ term}, \tag{2.65}$$

with $J[\rho]$ being the classical electrostatic repulsion of a charge distribution, and $T[\rho]$ its kinetic energy.

This then allows the energy to be written as

$$E[\rho] \ = \ \int \rho(\mathbf{r}) V(\mathbf{r}) \, \mathrm{d}\mathbf{r} + F_{\mathrm{HK}}[\rho] \tag{2.66}$$

for which an energy variational principle of the form

$$E_0 \leqslant E[\tilde{\rho}] , \tag{2.67}$$

with $\tilde{\rho}$ being a trial charge density, can be constructed. This follows since

$$\langle\tilde{\Psi}|\mathscr{H}|\tilde{\Psi}\rangle = E[\tilde{\rho}] \geqslant E[\rho], \tag{2.68}$$

which from the stationary principle gives

$$\delta\left\{E[\rho] - \mu\left[\int \rho(\mathbf{r})\,\mathrm{d}\mathbf{r} - N\right]\right\} = 0 , \tag{2.69}$$

which gives an Euler-Lagrange equation of the form

$$\mu = \frac{\delta E[\rho]}{\delta\rho(\mathbf{r})} = V(\mathbf{r}) + \frac{\delta F_{\mathrm{HK}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{2.70}$$

If a new functional $G[\rho]$ is then defined such that

$$G[\rho] = F_{\mathrm{HK}}[\rho] - \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{dr}_1\mathrm{dr}_2, \tag{2.71}$$

i.e., the Hartree energy (section 2.3) is removed from $F_{\mathrm{HK}}[\rho]$, then the total energy can be written such that

$$E[\rho] = \int V(\mathbf{r})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{dr}_1\mathrm{dr}_2 + G[\rho]. \tag{2.72}$$

Hohenberg and Kohn then suggested forms for $G[\rho]$, but the use of this function was superseded by the work of Kohn and Sham the following year.

2.5.3 The Kohn-Sham equations

Following on from the definition for $G[\rho]$ given by Hohenberg and Kohn, Kohn and Sham [30] formulated a more tractable functional of the charge density, which could be expressed exactly for certain problems.

They defined a new exchange-correlation functional of the form

$$E_{\mathrm{xc}}[\rho] = G[\rho] - T[\rho] , \tag{2.73}$$

where $T[\rho]$ is the kinetic energy of a non-interacting electron gas. Kohn and Sham introduced a set of normalised orthogonal orbitals, ψ_i, which allow the kinetic energy be written as

$$T = -\frac{1}{2}\sum_i \langle\psi_i|\nabla^2|\psi_i\rangle , \tag{2.74}$$

and the total charge density as

$$\rho(\mathbf{r}) \;=\; \sum_i |\psi_i(\mathbf{r})|^2 . \qquad (2.75)$$

This then gives a total energy of

$$E[\rho] \;=\; T + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 + \int V_{\text{appl.}}(\mathbf{r})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} + E_{\text{xc}}[\rho]. \qquad (2.76)$$

If this energy is then minimised subject to the constraint that the Kohn-Sham orbitals remain normalised, then with the introduction of a set of Lagrange multipliers, $\mathcal{E}_i$,

$$E - \sum_i \mathcal{E}_i \left(\langle\psi_i|\psi_i\rangle - 1\right) , \qquad (2.77)$$

is stationary with respect to variation in $\langle\psi_i|$. This then leads to a set of wave equations of the form

$$\mathcal{E}_i\psi_i(\mathbf{r}) \;=\; \left[-\frac{1}{2}\nabla^2 + V_{\text{appl.}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|}\,\mathrm{d}\mathbf{r}_j + \frac{\delta E_{\text{xc}}}{\delta\rho}\right]\psi_i(\mathbf{r}) , \qquad (2.78)$$

where $V_{\text{appl.}}$ is due to external potentials, and $\frac{\delta E_{\text{xc}}}{\delta\rho}$ is the exchange-correlation potential. This equation is then solved by the usual self-consistent method.

This still leaves the problem of what form E_{xc} takes. As yet there is no exact form for this term, but a number of approximations to it have been suggested.

2.5.3.1 Janak's theorem

Janak demonstrated [31] that the variation of the total energy within density functional theory with respect to the occupancy of a Kohn-Sham orbital is equal to the energy of the many-body equivalent to that level. Hence the highest occupied Kohn-Sham level is seen to be at the same energy as the ionization energy of the system [31–33]. The proof of this theorem follows by considering the change in the energy of the system with the occupancy, f, of the i-th level

$$\frac{\mathrm{d}E}{\mathrm{d}f_i} \;=\; \frac{\partial E}{\partial f_i} + \left(\frac{\delta E}{\delta\psi_i}\right)\left(\frac{\delta\psi_i}{\delta f_i}\right) . \qquad (2.79)$$

The second of these two terms is zero, since

$$\begin{aligned}\left(\frac{\delta E}{\delta\psi_i}\right) &= E_i f_i \left(\frac{\delta\langle\psi_i|\psi_i\rangle}{\delta\psi_i}\right)\\ &= E_i f_i \cdot 0 ,\end{aligned}$$

as a result of equation 2.78. Similarly the first term of equation 2.79 is E_i.

This theorem has subsequently also been proved by an independent approach based on considering the exchange-correlation potential for a system intermediate to the states between which a transition occurs [33].

It is then possible to combine the transition state theory of Slater discussed in section 2.3 with Janak's theorem. This then allows the evaluation of the energy of electronic transitions from the ground to excited states within density-functional theory.

A variant of this method has recently been developed to study electronic transitions of carriers from the valence or conduction bands to defects in solids by Resende *et al.* [34, 35]. If the electronic affinity of the highest occupied state of the defect is calculated by considering

$$A_d \quad = \quad E_n(\mathbf{R}_n) - E_{n+1}(\mathbf{R}_{n+1}) \; , \tag{2.80}$$

i.e., the energy difference between the relaxed structures in the charge states with n and $n+1$ electrons, at their corresponding relaxed coordinates of $\mathbf{R}_n$ and $\mathbf{R}_{n+1}$. When expanded as a Taylor series about $E_{n+\frac{1}{2}}(\mathbf{R}_{n+\frac{1}{2}})$ this gives

$$-\frac{\partial E_{n+\frac{1}{2}}(\mathbf{R}_{n+\frac{1}{2}})}{\partial n} - \frac{\partial^2 E_{n+\frac{1}{2}}(\mathbf{R}_{n+\frac{1}{2}})}{\partial n \partial \mathbf{R}_{n+\frac{1}{2}}} \cdots \tag{2.81}$$

If the displacement of atomic coordinates is small during the transition, this gives $A_d = \mathcal{E}_{n+\frac{1}{2}}(\mathbf{R}_{n+\frac{1}{2}})$, i.e., the highest (partially) occupied KS level. This electron affinity can then be compared against that calculated for a known defect. The electron affinity of this standard defect can then be corrected by comparison with experiment to give the adjustment needed to produce the transition energy to the appropriate band-edge. If both the standard defect and also the defect of interest posseses similar many-body electronic states, this allows the correction to then be used for the defect of interest. This circumvents the necessity to calculate transitions to an electronically excited state since the experimentally measured level contains the energy of the electron or hole in the excited configuration in addition to the ground-state affinity. This has the advantage of avoiding having to calculate excited state configurations of the defect, which would be beyond the scope of the ground-state density functional theory.

2.5.3.2 The local density approximation

If we return to the question of the approximation used for E_{xc}, then the most common starting point is the local charge density.

This is very much in the spirit of the Thomas-Fermi method, which uses the kinetic energy of jellium as the local K.E. of a non-uniform electron gas. If $E_{\mathrm{xc}}[\rho(\mathbf{r})]$ is assumed to be locally the same as a uniform gas of density ρ, then the total exchange and correlation energy for a spin-polarised electron gas can be written as

$$E_{\mathrm{xc}}[\rho_{\uparrow}(\mathbf{r}),\rho_{\downarrow}(\mathbf{r})] = \int \mathcal{E}_{\mathrm{xc}}[\rho_{\uparrow}(\mathbf{r}),\rho_{\downarrow}(\mathbf{r})]\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}. \tag{2.82}$$

This of course still leaves the question of what the form of $\mathcal{E}_{\mathrm{xc}}[\rho]$ is. In practice this term is derived from parameterisation of results from diffusion or variational quantum Monte Carlo (see section 2.6) and formal perturbation theory. This then leads to a class of local density functional theories which are parametrized from different Monte Carlo calculations, for example the Ceperly-Alder QMC calculations [36], from which an LDA parameterisation for spin-polarised electron gases was developed by Vosko, Wilk, and Nusair (the VWN functional [37]). Details of the approximate Perdew-Zunger [38] functionals used in the cluster calculations in this thesis are given in chapter 3.

2.5.3.3 Beyond LDA

Several attempts have been made to improve upon the LDA by incorporating local gradient information, to give exchange correlation energies of the form

$$E_{\mathrm{xc}}[\rho_{\uparrow}(\mathbf{r}),\rho_{\downarrow}(\mathbf{r})] = \int \mathcal{E}[\rho_{\uparrow}(\mathbf{r}),\rho_{\downarrow}(\mathbf{r}),\nabla\rho_{\uparrow}(\mathbf{r}),\nabla\rho_{\downarrow}(\mathbf{r})]\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}, \tag{2.83}$$

which again are parameterised against accurate calculations. No single "best" choice for these functions has been suggested, but one of the most popular is that of Perdew and Wang (PW91 functional [39]).

Alternatively, empirical hybrid functionals such as that formed by a mixture of the Lee-Yang-Parr correlation functional [40], Becke's gradient correction [41] and Hartree-Fock exchange energy (the BLYP and B3LYP functionals) have been found to give good results for small molecules [42].

2.6 Quantum Monte Carlo

Applying Monte-Carlo methods directly to the many-particle wavefunction of a system to minimize the energy is now starting to become a feasible method of studying defects in solids [43]. Historically, stochastic methods have been used to find the correlated wavefunction of jellium and provide the exchange-correlation potential needed for LDF calculations. The two classes of method for such calculations are either variational (VMC) or projector Monte Carlo, the most common form of which is diffusion Monte Carlo (DMC).

In the first of these techniques the energy of the anti-symmetric many-body wavefunction (represented as a determinant) is minimized by considering the Raleigh ratio for a trial function, which is then modified to find the lowest energy state. The method usually chosen to perform the minimization is to generate a set of walkers on the multi-dimensional configurational surface of the problem, representing a population of possible solutions. The energy of each trial solution is then evaluated, the walkers are then displaced randomly on the energy surface and their energy re-evaluated. The choice of whether use any of the new set of displaced walkers as replacements for their predecessors in further steps is then made on basis of a selection criteria, commonly the Metropolis algorithm [44]. This is a probabilistic selection which based on a Boltzmann-like distribution (for such calculations, the variation in the energy of the walkers is usually minimized instead of the total energy since this gives equivalent results). The newly selected population is then again displaced and selections made, and this is iterated over the displacement-selection criteria. In principle the Metropolis algorithm asymptotically converges towards the ground state, but this convergence is not always realised since computational constraints limit the configuration space searched.

The second, and potentially more accurate method of projector Monte Carlo defines an operator which projects out the ground state wavefunction from a trial function. In the case of diffusion Monte Carlo this takes the form of a time-dependent Schrödinger equation, with an imaginary time direction, which over a sufficiently long time period causes any trial state to decay to the ground-state. The projection of the system forward in time is then achieved by representing a many-body

	γ	β_1	β_2	
Averaged	-0.1423	1.0529	0.3334	
Polarised	-0.0843	1.3981	0.2611	
	A	B	C	D
Averaged	0.0311	-0.0480	0.0020	-0.0116
Polarised	0.0155	-0.0269	0.0007	-0.0048

Table 2.1: The Perdew-Zunger parametrisation for the spin-averaged and fully spin-polarised electron gas.

wavefunction as a set of walkers which are randomly displaced, then duplicated or destroyed depending if the position displaced to lowers the energy. The major difficulty with this approach is the *Fermion problem*, in that the method will project out a Bosonic ground-state, this is usually overcome by forcing the system to possess a fixed $3N$ - 1 dimensional nodal surface running through the many-body wavefunction. Unfortunately finding the best choice of such a surface is then one of the limiting factors to projecting out the true Fermionic ground-state.

Ceperley and Alder [36] applied VMC to find the correlation energy of low-density jellium, for both spin averaged and polarised systems. By interpolating between these results and the high-density jellium case obtained by formal perturbation theory, Perdew and Zunger [38] parameterised $\mathcal{E}_c$ for unpolarised and fully polarised jellium as

$$\begin{aligned}\mathcal{E}_c &= \gamma\left(1+\beta_1\sqrt{r_s}+\beta_2 r_s\right)^{-1} \text{ (for } r_s \geqslant 1) , \\ \mathcal{E}_c &= B+(A+Cr_s)\ln(r_s)+Dr_s \text{ (for } r_s < 1) ,\end{aligned}$$

where the Weigner-Seitz radius of each electron is $r_s = (4\pi\rho/3)^{-1/3}$. The coefficients for the averaged and polarised electron gas is given in table 2.1.

For a system of intermediate polarisation ξ, where

$$\xi = \frac{\rho_\uparrow - \rho_\downarrow}{\rho_\uparrow + \rho_\downarrow} ,$$

the averaging procedure of von Barth and Hedin [45] can be applied to give

$$\begin{aligned}\mathcal{E}_c(\rho,\xi) &= \mathcal{E}_c(\rho,0)+f(\xi)(\mathcal{E}_c(\rho,0)-\mathcal{E}_c(\rho,1))\\ f(\xi) &= \frac{(1+\xi)^{4/3}+(1-\xi)^{4/3}-2}{2^{4/3}-2}.\end{aligned}$$

This then leads to functions which can be closely approximated by power series in ρ [21].

2.7 Pseudopotentials

Even with the computational speed of density-functional calculations, treating sufficiently many atoms to be able to simulate large systems such as defects in solids requires a method of reducing the number of states that must be considered explicitly.

The approximation of replacing the potential due to the atomic nuclei and "core" electrons in a group of atoms with some form of simplified approximate potential is known as a pseudopotential. As yet, there has not been any formal proof that it is possible to do this whilst retaining the correct chemical and physical properties for such pseudo-atoms. This of course has not prevented the widespread use of this approximation, since there are several major advantages to its use.

There are three main incentives to use pseudo-atoms in calculations. The first of these is due to the orthogonality of all of the states in the system. If the valence electrons in the problem are to be orthonormal to the core electrons, then these valence states each require a large number of nodal surfaces. For example, the valence s orbitals of isolated silicon atoms require 2 nodes each. This means that the basis used for an "all electron" calculation with every electron included explicitly should be able to represent very rapid changes in the wavefunctions, so contain many functions. This can then lead to a tendency towards numerical instability.

The second reason is accuracy, since as the energy of the core levels is appreciably lower than the valence states, if there is the same fractional error in calculating both, the effect from the core states is disproportionately larger. Hence if the core states can be removed, then calculations become more sensitive to energy changes in the valence states.

The final reason is speed, since the number of electronic states considered is reduced dramatically. In addition, for heavy atoms where the core levels require relativistic treatment, as the pseudopotential need only be calculated from one relativistic simulation of the atom, then subsequent calculations using the pseudo-atoms can treat the valence states as non-relativistic wavefunctions, leading to simplification of many further calculations.

In practice, pseudopotentials have been developed in the context of solid state physics, hence are often used within density-functional calculations. Hartree-Fock pseudopotentials are less well-developed, due partly to the complexity of dealing with the four-centre terms from the Coulomb and exchange terms.

The history of pseudopotentials in calculations on solids dates back to the 1950s [46], where empirical functions were fitted to produce some observable(s) of the system by varying $V^{\text{pseud.}}$ in

$$\left(T + V^{\text{pseud.}}\right)\psi_i \quad = \quad E_i\psi_i. \tag{2.84}$$

The fitted pseudopotential is then used to model some related system. There is an issue in the transferability of such pseudopotentials, since there will be a component in the potential due to the exact form of the valence states in the system which the parameterisation is carried out. One solution to this is to adjust the form of the pseudopotential to leave a transferable "ionic" potential,

$$V^{\text{pseud.}}_{\text{ionic}}(\mathbf{r}) \quad = \quad V^{\text{pseud.}}(\mathbf{r}) - \int \frac{\rho^{\text{pseud.}}(\mathbf{r}_1)\,\mathrm{d}\mathbf{r}_1}{|\mathbf{r}-\mathbf{r}_1|} - V_{\text{xc}}\left[\rho^{\text{pseud.}}(\mathbf{r})\right], \tag{2.85}$$

with $\rho^{\text{pseud.}}$ being the total charge density of the single particle pseudo-wavefunctions.

A substantial improvement occurred over these ionic pseudopotentials with the modern norm-conserving pseudopotentials [47], which possess the same charge distribution beyond a defined radius from the core as an all electron calculation.

Bachelet, Hamann and Schlüter [48] developed transferable norm-conserving pseudopotentials for all elements between H and Pu. Their prescription is now described.

The first stage is to solve the Kohn-Sham equations for the isolated, spherically symmetric, atom of interest, allowing the states to be labeled in terms of the angular momentum, l, (or $j = l \pm \frac{1}{2}$ for heavier atoms using the Dirac equation). All of the atomic states necessary for bonding in the solid are used as part of the basis for this calculation, such as additional d or f functions. In order that these functions should actually be used as part of the basis, the atom is calculated in a slightly ionised state with a small occupancy of these functions. This ionization is tabulated for each atom, and for relevant angular momentum states, by the originators of the pseudo-potential used. It is this ionization which lead to the construction of *ab initio* pseudo-potential being considered a "black art" by some.

The set of all-electron wavefunctions and eigenvalues from the atomic calculation are then used to construct an all electron potential $V^\nu(\mathbf{r})$, which contains contributions from nuclear, Hartree and exchange-correlation potentials. This potential then possesses a singularity at $\mathbf{r} = 0$. Hence a new potential, the first-stage pseudo-potential, is constructed to the form

$$V_l(\mathbf{r}) \quad = \quad V^\nu(\mathbf{r})\,(1 - f(r/r_{c,l})) - c_l^\nu(r/r_{c,l}). \tag{2.86}$$

The function $f(x)$ is chosen such that as $x \to 0$, then $f(x) \to 1$, a common choice for this function is something like $\exp(-x^{3.5})$. This then leads to $V_l \to c_l^\nu$ at $x = 0$, with c_l^ν chosen such that the lowest energy eigenvalue generated by using the potential V^ν is the same as that from using V_l. This leads to the pseudo-wavefunction, ψ_{1l}^ν, for this potential being the same as the all-electron wavefunction at large radii, apart from a Normalisation constant. This naturally follows since the potential V_l is the same as V^ν at large distance. The value of $r_{c,l}$, termed the core-radius, determines the distance at which the true- and pseudo-wavefunction match. Care must be exerted in the choice of $r_{c,l}$ to remove nodes from the pseudo-wavefunction, but still leave the pseudo-wavefunction capable of describing the chemical nature of the atom. The usual choice of $r_{c,l}$ is halfway between the last node and the last extremum of the all-electron wavefunction.

The pseudo-wavefunction, ψ_{1l}^ν, is now modified so that it normalised to match the all-electron function. This is done by defining

$$\psi_{2l}^\nu \quad = \quad \gamma_l^\nu\left[\psi_{1l}^\nu(\mathbf{r}) + \delta_l^\nu r^{l+1} f(r/r_{c,l})\right]\;, \tag{2.87}$$

with γ and δ being normalisation coefficients.

The pseudo-potential which would generate ψ_{2l}^{ν} can be found by inverting the Schrödinger equation using the eigenvalues of the all-electron system. This leads to a new pseudo-potential which will generate the ψ_{2l}^{ν} eigenfunction, but still contains Hartree and exchange-correlation potentials. The Hartree potential can be removed exactly by subtraction, but the exchange-correlation is somewhat more difficult to deal with, since it can only be approximately separated from the pseudo-potential. This is done by subtracting the all-electron exchange-correlation potential due to the charge density and any spin-polarisation from the pseudo-potential. This then leaves the bare electron-ion pseudo-potential $V_l(\mathbf{r})$. For heavier atoms, a scalar relativistic pseudopotential can be defined which contains a spin-orbit contribution, but for the light atoms in this work this is not considered to be required.

The pseudo-potential is then fitted to a simple functional form of

$$V_l(r) \quad = \quad -\frac{Z_v}{r}\left[\sum_{i=1}^{2} c_i^c \operatorname{erf}\left(\sqrt{\alpha_i^c r}\right)\right] + \sum_{i=1}^{3}\left[A_{i,l} + r^2 A_{i+3,l}\right]\exp\left(-\alpha_{i,l} r^2\right) \quad (2.88)$$

with Z_v and α_i^c being the charge of the valence electrons and the inverse of the extent of the core charge. The first term in the new potential is strictly local, since it is independent of angular momentum, and appears only as the Coulombic charge due to the valence electrons at $\alpha_i^c r \gg 1$. The second term does depend on the angular momentum of the state, and so is non-local. It is possible to account for additional spin-orbit interactions by using a further, non-local, angular dependent potential [48], but this has not been used in the current work.

Chapter 3

The AIMPRO methodology

The particular methods used in this work, and embodied in the AIMPRO code, are now described. There are two forms of the code. The first is at heart a cluster-based, local density functional approach (using a modified form of the Perdew–Zunger [38] exchange-correlation energy), with a localised basis of Cartesian Gaussian and Gaussian-type functions and the pseudopotentials of Bachlet *et al.* [48]. The supercell code is also a Gaussian basis code, and uses the exact Perdew–Wang [39] local exchange-correlation.

There are several approximations and methods specific to these implementation which merit some additional description.

Matrix formulation

As with the Hartree-Fock equations, the Kohn-Sham equations can be cast in the form of a matrix problem similar to the Roothaan equations (see section 2.4.1), i.e.,

$$(\mathbf{H} - \boldsymbol{\mathcal{E}}\mathbf{S})\mathbf{c} \quad = \quad 0 \,, \tag{3.1}$$

with two set of matrices, one of each spin. For the cluster code, the overlap matrix $\mathbf{S}$ takes the customary form, with the Hamiltonian matrix for this system being

$$H_{ij} \quad = \quad T_{ij} + V_{ij}^{\text{pseud.}} + V_{ij}^{\text{H}} + V_{ij,s}^{xc} \,, \tag{3.2}$$

i.e., the kinetic, pseudopotential, Hartree and exchange potentials. In the case of spin-polarized calculations, there will be a set of matrix problems for each of the two spin components of the system.

3.1 Basis

The Gaussian functions used as a basis are commonly of either s or p type, with their form being

$$\phi_i(\mathbf{r}) = F \exp\left\{-\alpha_i(\mathbf{r}-\mathbf{R})^2\right\} \tag{3.3}$$

at site $\mathbf{R}$ [49]. The function F is either 1 for s functions, or $(v - \mathbf{R}_{iv})$, with $v = x$, y or z for p-type orbitals.

In addition to the more conventional atom centred Gaussian basis used in other codes, further s-type functions are commonly placed at bond-centred sites between atoms. These additional functions are particularly useful for materials such as silicon, where there is a d component required in the basis, which is then provided by these additional Gaussian functions, without the use of computationally expensive atom centred polarisation functions.

Using such Gaussians, the charge density then becomes

$$\rho(\mathbf{r}) = \sum_{\lambda=1}^{\text{occ.}} f_\lambda |\psi^\lambda|^2 = \sum_{\lambda=1}^{\text{occ.}} f_\lambda \sum_i \left|c_i^\lambda \phi_i\right|^2 , \tag{3.4}$$

where f_λ is the occupancy of the λ-th single particle state, ψ^λ, in the system.

If the density matrix is

$$b_{ij}^s = \sum_\lambda f_\lambda^s c_{i,s}^\lambda c_{j,s}^\lambda , \tag{3.5}$$

(note that this is real for these functions) then

$$\rho(\mathbf{r}) = \sum_{ij,s} b_{ij}^s \phi_i(\mathbf{r})\phi_j(\mathbf{r}). \tag{3.6}$$

The adoption of a Gaussian basis set also leads to particularly simple forms for the kinetic and pseudopotential matrices, i.e.,

$$T_{ij} = -\frac{1}{2}\int \phi_i \nabla^2 \phi_j \, \mathrm{d}\mathbf{r} \tag{3.7}$$

$$V_{ij}^{\text{pseud.}} = \int \phi_i \sum_a V_a^{\text{pseud.}} \phi_j \, \mathrm{d}\mathbf{r} , \tag{3.8}$$

both of which are readily evaluated analytically.

3.2 Accelerated evaluation of the Hartree terms

As described in section 2.3, the Hartree energy is given by

$$U^{\mathrm{H}} = \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2\ .$$

Since the evaluation of this term using localised basis functions scales as $\mathcal{O}(N^4)$, there is a very strong incentive to approximate this part of the calculation.

Sambe and Felton [50] introduced the concept of an "intermediate fit" to the charge density, using a second set of basis functions which were more convenient for the evaluation of the Hartree energy. This intermediate charge density takes the form of

$$\tilde{\rho}(\mathbf{r}) = \sum_k c_k g_k(\mathbf{r}). \tag{3.9}$$

This then leads to the approximate Hartree energy being expressed in terms of a set of two-centre integrals, taking the form

$$\tilde{E}_H = \int \frac{\rho(\mathbf{r}_1)\tilde{\rho}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 - \frac{1}{2}\int \frac{\tilde{\rho}(\mathbf{r}_1)\tilde{\rho}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{3.10}$$

which is exact when $\tilde{\rho} = \rho$. If the difference between the Hartree energy and this "fitted" energy is then written as

$$U^{\mathrm{H}} - \tilde{U}^{\mathrm{H}} = \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 - \sum_{ijk} c_k b_{ij} t_{ijk} - \frac{1}{2}\sum_{kl} c_k c_l g_{kl}\ , \tag{3.11}$$

which can then be minimised with respect to the fitting coefficients of the intermediate charge density to give

$$\sum_{ij} t_{ijk} b_{ij} = \sum_l g_{kl} c_l\ , \tag{3.12}$$

where

$$t_{ijk} = \int \frac{\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_1)g_k(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 \tag{3.13}$$

$$g_{kl} = \int \frac{g_k(\mathbf{r}_1)g_l(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2. \tag{3.14}$$

The solution to equation 3.12 gives the coefficients to the intermediate fit, and also allows that the expression for the fitted Hartree energy as

$$\tilde{U}^{\mathrm{H}} = \frac{1}{2}\sum_{kl} c_k c_l g_{kl}. \tag{3.15}$$

This energy is alway bounded by $\tilde{U}^{\mathrm{H}} \leqslant U^{\mathrm{H}}$, allowing an assessment of the quality of the intermediate fit by the increase in $\tilde{U}^{\mathrm{H}}$ with a larger basis.

There are a number of possible choices for the set of g_k functions used in this intermediate basis. For this work, the form of the intermediate basis used for the cluster calculation is a set of simple s-type Gaussians [51]

$$g_k(\mathbf{r}) = \exp\left\{-\alpha(\mathbf{r}-\mathbf{R})^2)\right\}. \tag{3.16}$$

This then leads to a fitted Hartree potential

$$\begin{aligned} \tilde{V}^{\mathrm{H}}_{ij} &= \frac{\partial \tilde{U}^{\mathrm{H}}}{\partial b_{ij}} && (3.17)\\ &= \sum_k c_k t_{ijk}. && (3.18)\end{aligned}$$

It is important to note at this point that the potential should be calculated from a fitted charge density which has not been forced to be normalised [52], since this would lead to an unphysical shift in the energy of the occupied eigenvalues.

For the supercell code the alternative choice of planewaves is taken [52]. The charge density

$$\rho(\mathbf{r}) = \sum_{ij} b_{ij}\phi_i\phi_j$$

is Fourier transformed (using a FFT in practice) to give the intermediate fit (in terms of a set of plane-waves, which are an exact fit for a sufficiently large cut-off) to the charge density, where the ϕ's are the original Cartesian Gaussian basis functions. The charge density in reciprocal space, $n(\mathbf{G})$, that comes out of this FFT is then used to calculated the Hartree energy from $n(\mathbf{G})$, this is done using the method of Ewald [53].

3.3 Evaluation of exchange-correlation

Similarly, a second basis set can be introduced to calculate the other "difficult" term in the energy, namely the exchange-correlation [21].

For the supercell, since the intermediate fit is the set of planewaves already produced for the Hartree energy, the standard method for supercell codes can be applied. The exchange-correlation potential, $\epsilon_{\rm xc}(\mathbf{r})$, can be constructed in real space and then $\epsilon_{\rm xc}(\mathbf{r})\rho(\mathbf{r})$ is Fourier transformed to give $E_{\rm xc}$.

For the cluster code, if a fitted charge density $\tilde{\rho}$ is introduced such that

$$\tilde{E}_{\rm xc} = \int \mathcal{E}_{\rm xc}\left(\tilde{\rho}\uparrow, \tilde{\rho}\downarrow\right)\tilde{\rho}\,{\rm d}\mathbf{r}, \tag{3.19}$$

then the evaluation of this energy can be simplified. This is an alternative to the other common method of evaluating this term for real-space calculations, which is to numerically integrate on a mesh of points (see for example Ref. [54] for a discussion of this method within linear scaling methodology).

While there is a natural choice of fitting the intermediate Hartree basis to minimise the electrostatic energy, there is no similarly obvious criterion for fitting the exchange–correlation basis. For the cluster code, a basis of s-Gaussian functions are fitted by minimising

$$\int \left[\rho_s(\mathbf{r}) - \tilde{\rho}_s(\mathbf{r})\right]^2\,{\rm d}\mathbf{r}, \tag{3.20}$$

to give

$$\tilde{\rho}_s(\mathbf{r}) = \sum_k d_k^s h_k(\mathbf{r})\;, \tag{3.21}$$

with h_k being the set of new basis functions. Hence

$$\sum_l H_{kl} d_l^s = \sum_{ij} u_{ijk} b_{ij}^s\;, \tag{3.22}$$

with

$$H_{kl} = \int h_k(\mathbf{r}) h_l(\mathbf{r})\,{\rm d}\mathbf{r} \tag{3.23}$$

$$u_{ijk} = \int \phi_i(\mathbf{r}-\mathbf{R}_i)\phi_j(\mathbf{r}-\mathbf{R}_j) h_k(\mathbf{r})\,{\rm d}\mathbf{r}. \tag{3.24}$$

The spin-averaged exchange energy then becomes (after dropping the spin label s)

$$\tilde{E}_{\rm xc} = \sum_k d_k \int h_k(\mathbf{r})\mathcal{E}_{\rm xc}(\tilde{\rho})\,{\rm d}\mathbf{r}. \tag{3.25}$$

For a set of s-Gaussians, these integrals are proportional to the average exchange-correlation density for the charge represented by the basis h_k, which for a slowly

i	A_i	p_i	q_i
1	-0.9305	0.3333	0
2	-0.0361	0.0	0
3	0.2327	0.4830	1
4	-0.2324	0.0	1

Table 3.1: Parameters for the spin-polarised exchange-correlation potential in cluster AIMPRO.

varying charge density will be approximately the exchange-correlation energy for the average charge density. Hence

$$\left(\tilde{\mathcal{E}}_{\mathrm{xc}}\right)_k \approx \mathcal{E}_{\mathrm{xc}} \left(\frac{\sum_l d_l \int h_l h_k \, \mathrm{d}\mathbf{r}}{\int h_k \, \mathrm{d}\mathbf{r}} \right) . \tag{3.26}$$

Now since $\mathcal{E}_{\mathrm{xc}}(\rho) \simeq A\rho^n$, with $n = 0.30917$ (see section 2.6), the expression for $\mathcal{E}_{\mathrm{xc}}$ can be further approximated. If a function is set up to interpolate between $n = 0 \to 2$ of the form

$$f(n) = \ln \left(\frac{\langle \tilde{\rho}^{\,n} \rangle_k}{\langle \tilde{\rho} \rangle_k^n} \right) , \tag{3.27}$$

then around $n = 0.3$, $f(n)$ is approximated by interpolating between the exact results at n =0, 1, and 2 to give

$$f(n) = \frac{n(n-1)f(2)}{2} , \tag{3.28}$$

where $f(2)$ can then be found from $\left\langle \tilde{\rho}^{\,2} \right\rangle_k$.

This formalism can then be readily extended to the spin-polarised case [21], with

$$\tilde{E}_{\mathrm{xc}} \left(\rho_\uparrow, \rho_\downarrow \right) = \sum_{i,s} A_i \int \rho_s^{p_i+1} \rho_{-s}^{q_i} , \tag{3.29}$$

with the coefficients for A_i, p_i and q_i given in table 3.1. Then, replacing ρ with $\tilde{\rho}$ gives

$$\tilde{E}_{\mathrm{xc}} = \sum_{k,s} d_k^s \mathcal{E}_{k,s} , \tag{3.30}$$

with

$$\mathcal{E}_{k,s} = \sum_i A_i \langle \tilde{\rho}_s^{p_i} \tilde{\rho}_{-s}^{q_i} \rangle_k \int h_k \, \mathrm{d}\mathbf{r}. \tag{3.31}$$

As with the spin averaged case, if $f(p, q)$ is then defined to interpolate between the known integer cases, such that

$$\langle \tilde{\rho}\,_s^p \tilde{\rho}\,_{-s}^q \rangle_k = \langle \tilde{\rho}_s \rangle_k^p \, \langle \tilde{\rho}_{-s} \rangle_k^q \exp[f(p, q)] \,, \tag{3.32}$$

with

$$f(p, q) = \ln \left(\frac{\langle \tilde{\rho}\,_s^p \tilde{\rho}\,_{-s}^q \rangle_k}{\langle \tilde{\rho}_s \rangle_k^p \, \langle \tilde{\rho}_{-s} \rangle_k^q} \right) . \tag{3.33}$$

$f(p, q)$ can then be approximated by

$$f(p, q) = \frac{1}{2} p(p-1) f(2, 0) + \frac{1}{2} q(q-1) f(0, 2) + pq f(1, 1) \,. \tag{3.34}$$

3.4 Self-consistency

The customary method of solution to the Kohn–Sham equations is to use a self consistency cycle, ideally with the minimum number of iterations. The initial charge density in AIMPRO is taken either from the same atomic calculations used to produce the pseudopotentials (as in the cluster code), or the charge density produced during optimisation of the basis set for a particular material (as in the supercell code). This then leads to an "initial guess" for the k coefficients $\mathbf{c}_{\mathrm{in}}^0$ of the basis sets used. After the first evaluation of eigenvalues, the resulting eigenvectors give a new set of expansion coefficients $\mathbf{c}_{\mathrm{out}}^1$. If it is assumed that improved coefficients for the next step in the self-consistency cycle can be constructed from a weighted combination of the new coefficients, $\mathbf{c}_{\mathrm{out}}^1$, and the originals, $\mathbf{c}_{\mathrm{in}}^0$, then

$$\mathbf{c}_{\mathrm{in}}^1 = (1 - w)\mathbf{c}_{\mathrm{in}}^0 + w\mathbf{c}_{\mathrm{out}}^1 \,. \tag{3.35}$$

This weighting factor w can then be chosen such that the difference between $\mathbf{c}_{\mathrm{in}}^n$ and the new $\mathbf{c}_{\mathrm{out}}^{n+1}$ which is derived from it rapidly become identical between iterations.

If the operation of determining the new charge density is denoted as

$$\mathbf{c}_{\mathrm{out}}^{n+1} = \hat{L}[\mathbf{c}_{\mathrm{in}}^n] \,, \tag{3.36}$$

then assuming that w is small this can be linearised to give

$$\mathbf{c}_{\text{out}}^{n+2} = \mathbf{c}_{\text{out}}^{n+1} + w\hat{D}\left[\mathbf{c}_{\text{out}}^{n+1} - \mathbf{c}_{\text{in}}^{n}\right] , \tag{3.37}$$

with

$$\hat{D} = \frac{\delta \hat{L}}{\delta \mathbf{c}}. \tag{3.38}$$

This then gives the condition for self-consistency as $\left[\mathbf{c}_{\text{out}}^{n+1} - \mathbf{c}_{\text{in}}^{n}\right] = 0$, which in turn gives

$$(1-w)(\mathbf{c}_{\text{in}}^{n} - \mathbf{c}_{\text{out}}^{n+1}) - w\hat{D}\left[\mathbf{c}_{\text{out}}^{n+1} - \mathbf{c}_{\text{in}}^{n}\right] = 0 , \tag{3.39}$$

i.e., since the input and output are identical

$$(1-w)(\mathbf{c}_{\text{in}} - \mathbf{c}_{\text{out}}) - w\hat{D}\left[\mathbf{c}_{\text{out}} - \mathbf{c}_{\text{in}}\right] = 0 \tag{3.40}$$

If w is a small value, say $\overline{w}$, then the output density $\overline{\mathbf{c}}_{\text{out}}$ will be

$$\overline{\mathbf{c}}_{\text{out}} = \mathbf{c}_{\text{in}} + \overline{w}\hat{D}\left[\mathbf{c}_{\text{out}} - \mathbf{c}_{\text{in}}\right] , \tag{3.41}$$

and hence

$$(1-w)(\mathbf{c}_{\text{in}} - \mathbf{c}_{\text{out}}) - w(\overline{\mathbf{c}}_{\text{out}} - \mathbf{c}_{\text{in}})/\overline{w} = 0 \tag{3.42}$$

$$= \mathbf{e}. \tag{3.43}$$

The vector $\mathbf{e}$ can then be minimised by considering the Coulomb energy of this difference in charge density, i.e,

$$\frac{1}{2} e_k e_l G_{kl}. \tag{3.44}$$

This procedure can be extended to include a greater history of previous charge densities, allowing a rapid evaluation of the self-consistent charge density.

3.4.1 Finite temperature

The smooth convergence of energy can only occur if the charge density changes without major discontinuities between iterations. One of the situations where such discontinuous changes regularly occur is that of a previously unfilled level sinking below the highest occupied state of the system, which causes a sharp redistribution of charge. A possible solution to this, as used in both forms of AIMPRO is to apply a finite electron temperature, not in a attempt to physically model excited electrons, but instead to allow a smooth reordering of levels. As the two levels become closer, charge will gradually transfer between the lower an upper levels without a discontinuous jump.

In practice this finite temperature (typically of around $k_\mathrm{B}T = 0.04$ eV) is implemented by using a Fermi distribution for the occupancy (f_i) of each of the i levels. The obvious method of implementing this is to consider the electronic free energy F of the N electrons, which would be

$$F = E + k_\mathrm{B}T \sum_i \{f_i \ln f_i + (1 - f_i) \ln(1 - f_i)\} - \mu \left\{ \sum_i f_i - N \right\} . \quad (3.45)$$

Minimising this expression with respect to f_i and μ give

$$f_i = \frac{1}{\exp(E_i - \mu)/k_\mathrm{B}T + 1} \quad (3.46)$$

$$\sum_i f_i = N . \quad (3.47)$$

3.5 Forces and second derivatives

After the self-consistent energy is generated, the question of the minimum energy of the atomic positions then arises. There are a large variety of possible methods for optimising structures, but the particular choice made for this work is to use analytical forces and a conjugate gradient algorithm to determine the position of a nearby (local) minimum.

Unlike plane-wave calculations, the Hellmann-Feynman theorem [55, 56] cannot be directly used for fixed, atom centred basis functions, due to the presence of addition Pulay terms [57] arising from the force due to moving the basis functions. Instead the individual terms of the Hamiltonian must be separately differentiated.

3.5.1 Forces

Starting from the energy for cluster, using the *fitted* Hartree and exchange-correlation terms

$$E[\rho] = T_s[\rho] + \int V^{\text{ext.}}(\mathbf{r})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 + E_{\text{xc}}[\rho] + E_{i-i} \ ,$$

from which the forces can be explicitly differentiated in respect of displacements positions of the ion cores, $\Delta\mathbf{R}$, i.e.,

$$\begin{aligned} \Delta E &= \sum_{ij} b_{ij}\Delta\left(T_{ij} + V_{ij}^{\text{pseud.}}\right) + \sum_{ij}\left(T_{ij} + V_{ij}^{\text{pseud.}}\right)\Delta b_{ij} \\ &\quad + \Delta\tilde{E}_H + \Delta\tilde{E}_{\text{xc}} + \Delta E_{i-i} \ , \end{aligned}$$

with

$$\begin{aligned} \Delta\tilde{E}_H &= \sum_{kl} c_k G_{kl}\Delta c_l + \frac{1}{2}\sum_{kl} c_k c_l \Delta G_{kl} \\ \Delta\tilde{E}_{\text{xc}} &= \sum_{ks} \mathcal{E}_{ks}\Delta d_{ks} + \sum_{ks} d_{ks}\Delta\mathcal{E}_{ks} \end{aligned}$$

The changes in the weighting vectors, $\Delta\mathbf{c}$ and $\Delta\mathbf{d}$, for the two fitted basis sets can be evaluated by considering

$$\sum_k G_{kl}\Delta c_l \ , \tag{3.48}$$

which equals

$$\sum_{ij}(t_{ijk}\Delta b_{ij} + b_{ij}\Delta t_{ijk}) - \sum_l c_l \Delta G_{kl} \ , \tag{3.49}$$

and similarly

$$\sum_k H_{kl}\Delta d_l^s = \sum_{ij}\left(u_{ijk}\Delta b_{ij,ds} + b_{ij}^s \Delta u_{ijk}\right) - \sum_l d_l^s \Delta H_{kl} \ . \tag{3.50}$$

These terms can be combined [21] to give

$$\sum_\lambda E_\lambda \Delta\left(\sum_{ij} c_i^\lambda c_j^\lambda S_{ij}\right) - \sum_{ij\lambda} E_\lambda c_i^\lambda c_j^\lambda \Delta S_{ij} \ , \tag{3.51}$$

and since

$$\sum_{ij} c_i^\lambda c_j^\lambda S_{ij} = N \ , \tag{3.52}$$

this becomes

$$-\sum_{ij\lambda} E_\lambda c_i^\lambda c_j^\lambda \Delta S_{ij} \ . \tag{3.53}$$

$\Delta\mathcal{E}_{ks}$ is dealt with by considering $\Delta\langle\tilde{\rho}_s\rangle_k$ and $\Delta\langle\tilde{\rho_s^2}\rangle_k$, similarly T_{ij} and S_{ij} depend only on the form of the basis functions for their dependence on $\mathbf{R}$ so are easily differentiated. The energy due to the pseudopotententials has a dependency on $(\mathbf{r}-\mathbf{R}_a)$, so is evaluated from integrating

$$\int \phi_i \Delta V_a^{\text{pseud.}} \phi_j \ \mathrm{d}\mathbf{r} \tag{3.54}$$

by parts.

3.5.2 Second derivatives

Once the structure has been relaxed, using the derived forces, the dynamical-matrix for the system can be calculated. In AIMPRO the second derivatives of the energy are evaluated by a finite-difference method where atoms in the cluster (usually only a subset of interest) are displaced in turn from their equilibrium, and a new set of forces from a separate self-consistency cycle is calculated.

If the force on atom b, in the Cartesian direction m, due to the atom a having been displaced a distance $+\epsilon$ along the direction $\mathbf{l}$ is $f_{mb}^{+\epsilon}(\mathbf{l},a)$, then the second derivative matrix is

$$D_{\mathbf{l}a,mb} \quad = \quad \left[f_{mb}^{+\epsilon}(\mathbf{l},a) - f_{mb}^{-\epsilon}(\mathbf{l},a)\right]/2\epsilon \ . \tag{3.55}$$

An advantage of this approach is that the resulting matrix contains contributions from the even *anharmonic* terms of the potential energy surface [58], so reaches beyond the harmonic approximation to some extent.

Since only part of the dynamical matrix is explicitly calculated, the rest is commonly filled with approximate derivatives for the rest of the cluster or supercell. The usual choice for this is the inter-atomic potential of Musgrave and Pople [59], which considers only interaction between nearest neighbours.

The vibrational modes of the system can then be calculated from the second derivatives matrix by starting from the set of equations for motion of the atoms.

The equation for the l-th atom of mass M_l is

$$M_l \frac{\partial^2 R_{mb}}{\partial t^2} = -\sum_{la} D_{la,mb} R_{mb} . \tag{3.56}$$

If a trial solution to eqn. 3.56 of the form

$$R_{mb}(t) = \frac{1}{\sqrt{M_l}} A_{mb} \exp(-i\omega t) \tag{3.57}$$

is introduced, then an easily soluble matrix equation

$$\sum_{la} \left[\frac{D_{la,mb}}{\sqrt{M_l M_m}} - \delta_{ab}\delta_{lm}\boldsymbol{\omega}^2 \right] A_{la} = 0 , \tag{3.58}$$

gives the frequencies, $\boldsymbol{\omega}$, and normal modes, $\mathbf{Q}$. In practice the terminating atoms on the surface of clusters are usually chosen to have an artificially high mass ($>$ 1000 a.m.u.) to remove their local modes.

Unfortunately, the frequencies derived by this method are only accurate for those modes localised on the defect or within the phonon-gap. For modes which are in resonance with the lattice, other methods such as Green's functions [21] must be considered.

Once the normal coordinates of the modes have been obtained, the infrared intensity of a particular vibrational transition can be found by considering the net change in dipole moment over the course of the mode. The total dipole moment of a cluster is

$$\mathbf{P} = \sum_a Z_a \mathbf{R}_a - \int \mathbf{r}\rho(\mathbf{r})\, \mathrm{d}\mathbf{r} , \tag{3.59}$$

and the atoms can be displaced along the normal mode to give $\Delta\mathbf{P}$, which allows the square of the effective charge of the i-th mode to be calculated from

$$\eta_i^2 = M_i \sum \left(\frac{\partial \mathbf{P}_i}{\partial \mathbf{Q}_i} \right)^2 , \tag{3.60}$$

where the sum is over any degenerate components of the mode.

Chapter 4

Experimental techniques

Several common experimental techniques which are of relevance to the defects studied in this thesis are now briefly described. Since I cannot claim to be anything approaching expert in any of the techniques outlined here, the reader is referred for further information on all of these methods to the recent review edited by Stavola [60], and references therein.

For electrically active defects, such as the carbon-hydrogen defects studied in chapter 9, where the electrical states of the defect are deep within the band-gap of the host semiconductor[5] the usual technique of choice to characterise the electronic levels defects is Deep Level Transient Capacitance Spectroscopy (DLTS). Such levels in the forbidden gap of the semiconductor can be detected by other techniques, such as photoluminescence. Luminescence comes into its own for excited states of defects such as excitons bound to a defect, as with the B centres described in chapter 6. While the vibrational modes of a system can be studied by luminescence, the more common method is to make use of absorption or scattering of light by the vibrations of a defect, as in infrared and Raman spectroscopy. Such techniques are readily applied to both molecular hydrogen (chapter 5), other hydrogen dimers (chapter 7), and complexes between hydrogen and native defects (chapter 8). In addition to the techniques directly applied to the defects discussed in this thesis, the versatile method of electron paramagnetic resonance is also described. All of the above techniques can also be allied to the application of uniaxial stress on the crystal, allowing the symmetry of the defect to, usually, be unambiguously assigned.

[5]Deep is a relative term, for silicon this is typically taken to be at least 0.1 eV away from the band edges, where the hydrogenic model [61] is inapplicable

4.1 Vibrational spectroscopy

In a one-phonon transition between two of a defect's vibrational levels there are two main optical processes that can occur. Either a photon is absorbed (or emitted, but we will consider the first process only in this section), or alternatively scattered by the defect. The first process is used in infrared absorption spectroscopy and the second in Raman spectroscopy.

Since the total crystal momentum of the process must be conserved, this gives rise to an additional constraint on the absorption process. So for example one phonon processes are restricted to vertical transitions in the dispersion space, and two phonon processes must produce both phonons with equal but opposite momenta.

4.1.1 Infrared absorption

For a photon to be absorbed during a vibrational transition between states of a defect $\langle\chi_i|\boldsymbol{\mu}|\chi_f\rangle$ must be non-zero, where $\boldsymbol{\mu}$ is the electric dipole alowing the transition between the initial, χ_i, and final , χ_f, states. This naturally leads to the additional selection criteria that for a transition to be symmetry allowed, $\Gamma_{\chi_i}\otimes\Gamma_{\boldsymbol{\mu}}\otimes\Gamma_{\chi_f}\supset\Gamma_1$, i.e., the product of the symmetries contains the totally symmetric irreducible representation.

The integrated intensity of a particular absorption line can be derived [62] by considering the absorption coefficient as a function of frequency, hence

$$I \quad = \quad \int \alpha(\omega)\,\mathrm{d}\omega = \frac{2\pi^2\rho\nu^2}{ncM}\ , \tag{4.1}$$

with ρ and M being the concentration and mass of the oscillating part of the impurity, inside a material of refractive index n (c is of course the speed of light). The apparent charge, ν, is related to the effective charge, e, of the defect by a local field correction which is material dependent. A further consideration arises from orientational degeneracy of the defect, since the observed effective charge will be reduced by a factor related to this degeneracy [63].

4.1.2 Raman scattering

If intense light (typically from a laser source) is shone onto a sample containing defects, then light scattered from such centres can give information on the vibrational states of the defects. The main component of the scattered light is that due to elastic Raleigh scattering at the same energy as the incident light, which contains no information about the vibrational properties of the defects. On either side of the main Raleigh line are the Stokes (at lower energy) and anti-Stokes lines. Of these two weak bands, the Stokes lines are the stronger components (except in the case of a population inversion) and are due to energy transfer from the incoming photons to a vibrational state(s) of the defects in the sample.

The dipole moment, $\boldsymbol{\mu}$, induced on a defect by an electrical field $\mathbf{E}$ will be $\boldsymbol{\mu} = \underline{\underline{\alpha}}\mathbf{E}$, where the polarisability tensor $\underline{\underline{\alpha}}$ is of the same irreducible representation as the product functions (x^2, xy, ...) for the symmetry, if any, of the defect considered. Whether any energy can be transfered depends on the magnitude of $\boldsymbol{\mu}$, i.e.,

$$|\boldsymbol{\mu}| = |\mathbf{E}| \left\langle \chi_i | \underline{\underline{\alpha}} | \chi_f \right\rangle . \tag{4.2}$$

The information from Raman and infrared absorption spectroscopy is often complimentary, since due to the different selection rules one technique can often observe modes that the other cannot, for example in the case of $C_{As}H$ complexes in GaAs, the bend-related mode at 739 cm^{-1} which should be observable by infrared spectroscopy has so far only been seen by Raman spectroscopy [64], presumably in part through a small effective charge for the mode.

4.2 Photoluminescence spectroscopy

There are a variety of luminescence techniques, the name of the method usually giving the source of the energy which excites the emitted light, such as electroluminescence or triboluminescence. In the case of photoluminescence, the light is produced by the relaxation of the system after it is pumped by previous illumination. Above band-gap ($h\nu > E_g$) illumination creates electron-hole pairs which then recombine, often mediated by trapping at a defect site (this is distinct from below band-gap illumination, where internal states of a defect are excited directly).

This recombination need not necessarily be radiative, but in the cases where at least part of the process emits light this can give information about the trapping defect. For example in the case of bound excitons, such as the B family, both the electron and hole are trapped as a pair, with one of the components of the exciton localised on the defect, binding the other via a Coulombic interaction. The recombination then gives a measure of the binding energy to the defects, since this decreases the energy of the free exciton recombination at 1.1551 eV by this binding (note that this free exciton recombination is 14.3$\pm$0.5 eV lower than the band-gap of silicon due to the binding between the electron and hole [65]).

The recombination probability of an electron-hole pair, as with the vibrational techniques discussed so far, is controlled by an overlap integral of the form $\langle\psi_i\chi_i|\boldsymbol{\mu}|\psi_f\chi_f\rangle$, which again gives rise to a symmetry condition of $\Gamma_{\psi_i\chi_i} \otimes \Gamma_{\boldsymbol{\mu}} \otimes \Gamma_{\psi_f\chi_i} \supset \Gamma_1$, where $\boldsymbol{\mu}$ is the dipole. Unlike the purely vibrational transitions in the previous technique, the initial and final states are labeled by both their electronic and vibrational states. This leads to the possibility of measuring information on both the electronic and vibrational states of the defect simultaneously. It is worth noting at this point that since electron and hole spin are important in these transitions, the symmetry labels for states should come from the full double group for the centre, and not merely the spatial symmetry [66].

The electron-phonon coupling between the electronic and vibrational states of the defect appears in the intensity of the transitions, which can be expressed as

$$I_{\mathrm{e,v}} = \langle\psi_i(\mathbf{r})|\mathbf{r}|\psi_f(\mathbf{r})\rangle\langle\chi_i(\mathbf{q})|\mathbf{q}|\chi_f(\mathbf{q})\rangle \; , \tag{4.3}$$

for the transition from the initial (i) to the final (f) states, with ψ and χ are the electronic and vibrational parts of the wavefunction, dependent on the electronic ($\mathbf{r}$) and normal-mode ($\mathbf{q}$) coordinates. This allows the, usually unknown, electronic states and, often harmonic, vibrational states to be analysed separately. As with Raman spectroscopy, this leads to a series of Stokes and anti-Stokes lines flanking a central zero-phonon line (ZPL), but unlike Raman spectroscopy the ZPL is specific to the defect and depends on the electronic properties of the state as well as its vibrational ground state.

This dependence on vibrational state also means that the mass of the isotopes in the defect affect the spectra, as with the hydrogenated and deuterated B lines (chapter 6). This is most clearly seen in the sharp zero-phonon band, which shifts with isotope due to the difference between the zero point energy of the electronic ground and excited states, which is in turn due to the difference in their vibrational anharmonicity. Care must be exerted in analysing such spectra, since if the ground and excited states do not involve appreciable amplitude on one of the defect atoms, no isotope shift from that atom will be observed. This means that PL can sometimes be used to obtain the lower bound on the number of atoms of a particular chemical type in the defect, but more may be present. For example, it was initially suggested that one of the B centres contained three hydrogen atoms, but this has been revised to at least four [67].

Application of an electrical field to the sample during photoluminescence gives a method to determine whether the defect possesses inversional symmetry, since an alternating field will cause centro-asymmetric defects to display a line shape of derivative form [68, 69], due to a first order Stark effect. The use of a magnetic field also adds further information through the Zeeman effect, on any unpaired spin in the defect, since the field will split the degeneracy between formerly equivalent m states, leading to a spin Hamiltonian with components from the spin-orbit coupling of both the hole and the electron, and their mutual interaction. For the case of bound excitons, the response of the hole and electron components will be different, due to the local screening effects of the defect, and hence this gives a technique to discriminate the properties due to the hole and electron within the exciton. This discrimination can also be achieved through studying the temperature dependence of the luminescence, since the hole and the electron will be unequally bound to the centre, and hence ionise at different rates with temperature, which allows the intensity to be fitted to a simple equation [70] of the form

$$I(T) = \frac{I_0}{(1+f)(1+p)} , \tag{4.4}$$

with f and p describing the shallow and deeply trapped parts of the centre. f and p are Boltzmann like distributions of the form

$$\begin{aligned} f &= \frac{a}{1 + bT^{3/2}\exp(-E_s/kT)} \\ p &= gT^{3/2}\exp(-E_d/kT) \ . \end{aligned}$$

E_s and E_d are then the ionisation energies for the shallow and deeply trapped parts of the exciton, which are fitted using the variables a, b and g.

One of the most prominent photoluminescence systems in silicon is the I_1 or W line luminescence at 1018.2 meV which is produced in silicon by irradiation or ion implantation [71, 72]. There is a well resolved local vibrational mode at 70.0 meV observed as a phonon replica of the main line, which is considered to be due to a strong Si–Si bond. Uniaxial stress studies of this centre have been used to assign a trigonal symmetry, but the actual structure of the defect remains contentious [73, 74], since models based on both vacancy and self-interstitial related centres have been put forward to explain the observed spectra. The most recent observations which combine luminescence with techniques such as stress [75] or correlation with growth of other, interstitial related, defects [76] suggest that the centre is due to an interstitial aggregate, with the most recent theoretical model [73] proposing that it consists of a ring of three interstitials lying in a (111) plane around the centre of a T_d site.

4.3 Capacitance spectroscopy

The usual definition of the depths of electronic levels within the band gap are with respect to the valence and conduction band edges. If the energy of the system when a carrier is trapped at the defect is compared against that of an un-trapped carrier in a band-state, then the position of the state with respect to the conduction band for emission of an electron can be defined as

$$E|e_d, 0_c\rangle - E|0_d, e_c\rangle \ ,$$

and similarly for hole emission to the valence band as

$$E|0_d, h_v\rangle - E|h_d, 0_v\rangle \ ,$$

with the electron or hole occupancies given for the defect and the appropriate band edge.

For shallow states this gives an easy classification into donor and acceptor states. If the energy of the state is deeper, then the level can instead be better thought of as an electron or hole trap, depending of which process is the faster to change the charge state. For example given a defect with a single state deep in the gap that when filled gives as neutral centre, and when empty a positive defect[6], this will be a hole trap if the $(+^n) \rightarrow (+^{n-1})$ process is more rapid, or an electron trap if the reverse process were faster. The capture rate of a centre is given by [77, 78]

$$C_c = \sigma_c[c]\sqrt{\frac{3kT}{m_c^*}} \, , \tag{4.5}$$

with $[c]$ being the concentration of the carrier of interest, which possesses an effective mass of m^*, and would be captured by the centre with an effective cross-section of σ_c. Similarly the thermal emission rate from a centre is given by a Boltzmann like relation

$$e_c = \sigma_e \exp(-\Delta E/kT) \, . \tag{4.6}$$

Whether a state is filled for a given Fermi-level and temperature then depends on the relative rates of capture and emission for electrons and holes. Using the principle of detailed balance [79] leads to the occupancy of an electron trap being

$$\text{Occ.} = \frac{\sigma\langle\nu\rangle g_c}{g} \exp\left(\frac{-\Delta E}{kT}\right) \, , \tag{4.7}$$

with g_c and g being the degeneracy of the conduction band and the defect, ΔE the trapping energy, and

$$\langle\nu\rangle = \sqrt{\frac{3kT}{m_c^*}} \, ,$$

with a similar relation holding for hole traps. This can then be used to measure the position of many-body electronic transitions in the band gap by producing a Arrhenius plot of the log of emission rate over the temperature squared, $\ln(e_c T^{-2})$, against $1/T$, but the "energy" obtained is the change in the total enthalpy, since ΔE is a change in the Gibbs free energy.

The measurement of the transition position is usually performed by deep level transient capacitance spectroscopy (DLTS) [80]. This method makes use of a Schottky barrier or p-n junction diode, within which the space-charge formed by the junction can be probed. The innate electrical field within the diode acts to sweep free

[6] i.e., a 0/+ many-body state, where the standard notation of i/j is used for a state which is of charge i when it is filled and j when empty.

carriers out of the space-charge region, leaving an insulator-like region, the width of which can be varied by applied voltage. In the case of a *p*-*n* junction the width of this depletion region is then given by

$$W = \sqrt{\frac{2\epsilon(V_{\text{int.}} + V_{\text{appl.}})}{q\,[I]}} ,$$

where ϵ is the dielectric constant of the junction, $V_{\text{int.}}$ and $V_{\text{appl.}}$ are the internal and applied voltages across the junction, q the charge on the defect, and $[I]$ is the concentration of the defects. This leads to the junction possessing a capacitance of $\epsilon A/W$, where A is the area of the junction. This then gives a technique to measure the position of the levels by applying a transient voltage pulse to the diode, and then watching the exponential decay in the capacitance change this induces[7].

The capacitance change from the pulse arises since as the width of the depletion zone is changed by the applied pulse, defect centres which were at equilibrium inside the edge of the depletion layer are now outside the depletion zone during the pulse. In this position, the centres rapidly reach equilibrium by capture of the abundant carriers present outside of the depletion zone, at a rate given by [81]

$$[d] - [d^c(t)] = ([d] - [d^c(0)]) \exp(-C_c t) , \tag{4.8}$$

where the concentration of the defect $[d]$ after it has captured a carrier $[d^c]$ changes in proportion to the capture rate of that process, C_c. If the pulse lasts sufficiently long enough, all of the defects will then reach equilibrium in the region outside the depletion zone, at which point if the pulse ends, they are returned to the depletion region and again find themselves in a non-equilibrium charge state. This additional space charge within the depletion region then manifests in the capacitance of the junction as a decaying transient as the defects relax back into equilibrium at a rate of

$$[d^c(t)] = [d] \exp(-E_c t) \tag{4.9}$$

[7]The nature of the thermal decay of the capacitance is more complicated for situations such as very high defect concentrations, traps that change their behaviour strongly with strength of the applied field (such as the Poole-Frenkel effect), and multiple centres with similar electronic levels, but other than to note that these effects exist, they will not be discussed further here.

by emission of trapped carriers. This relaxation can be followed by capacitance measurement, typically at frequencies of > 1 MHz, i.e., too rapidly for the traps to respond to any changes in the width of the depletion region.

In DLTS, the filling pulse is repeated at a constant frequency while the temperature of the sample is slowly varied. The original method of analysing the capacitance transient introduced by Lang [80], was to make use of box-car integration between times t_1 and t_2 after the filling pulse at t_0, with the capacitance change then defined as

$$\Delta C = C(t_1) - C(t_2) , \tag{4.10}$$

which then gives the pulse rate at which the maximum capacitance change would occur as

$$e_c = \frac{1}{t_1 - t_2} \ln(t_1/t_2) , \tag{4.11}$$

with e_c being the carrier emission rate for the appropriate process (hole or electron emission).

The interval $t_1 - t_2$ is known as the “rate window” of the system. The capacitance-change signal is then maximised when the rate of thermal emission by the defects matches the rate window of integrator. The resulting signal for a fixed rate window as a function of temperature is shown pictorially in figure 4.1 If the rate window is then changed this signal maximum shifts to a new temperature, and hence the change in emission rate can be found as a function of temperature, which from an Arrhenius plot of equation 4.7 gives the thermal emission energy, and also the capture cross-section, σ_∞, at $1/T = 0$.

The most recent method developed to analyse the capacitance transients is “Laplacian” DLTS [82], where the recorded transient at a particular temperature is processed to give a spectral-density function by an finding the inverse Laplace transform for the transient, with the activation energy and cross-section then obtained by repeating the experiment at several temperatures as with conventional DLTS.

There are several variants of the basic DLTS technique, such as double-DLTS where two filling pulses at different voltages are used to map the depth distribution

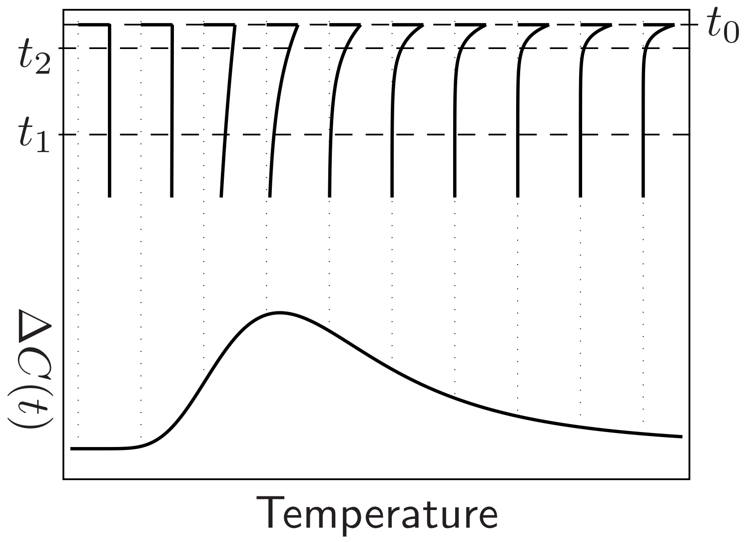


Figure 4.1: Schematic of a deep level transient spectra for a fixed rate window, i.e, the change in capacitance, ΔC, over the time period between t_1 and t_2 after the filling pulse at t_0.

of defects, since each pulse fills a different region of the space-charge layer. In addition other methods of inducing a capacitance transient such as optical excitation (ODLTS) can be applied [83]. Minority carrier (MCTS) traps can be examined in addition to the majority carriers [84], since flood-lighting the sample fills minority traps and hence causes a capacitance change.

4.4 Electron paramagnetic resonance

For defects with unpaired electronic spin, the technique of EPR can be applied to study the structure and localisation of the spins in the system and also the nuclear spins of surrounding atoms. When the defect centres are at a sufficiently low concentration that there is no appreciable interaction between the spins of different centres, then absorption of radiation can be described in terms of transitions between the Zeeman split spin states of the electronic spin. In practice, the transition energies of the states of a defect are studied as a function of applied magnetic field within a resonant cavity operating at a specific frequency. The most common method of detecting the position of a resonance peak is to apply a modulation to the applied magnetic field. Typically the field is varied at a frequency of around 100 KHz by an amount which is small in comparison to the line width. The component of the EPR absorption signal at 100 KHz is then amplified via a lock-in system, leading to a derivative signal which clearly shows the peak position. More advanced methods using field modulations at more than one frequency have also been developed [85].

The magnetic dipole of a paramagnetic centre is given by

$$\boldsymbol{\mu} \;=\; -\underline{\underline{g}}\mathbf{J}\mu_B\,, \tag{4.12}$$

in terms of the Bohr magneton, $\mu_B = eh/2mc$. The total angular momentum of the spin consists of the orbital and electron spin momenta, $\mathbf{J} = \mathbf{L}+\mathbf{S}$. In the case of light elements possessing only s and p valence orbitals, as with the defects considered in this thesis, the orbital angular momentum is strongly quenched, leading to $\mathbf{J} = \mathbf{S}$. This then allows the g factor for the system, due to the anisotropic environment of a crystal, to be written as two components

$$\underline{\underline{g}} \;=\; g\underline{\underline{1}} + \underline{\underline{g^a}} \tag{4.13}$$

where the isotropic g term is the Landé g-factor (=2.0023), and the anisotropic tensor is due to the local environment. This anisotropic term can be considered as due to a small perturbation from spin-orbit coupling of the form [86]

$$g^a_{ij} \;=\; -2\lambda \sum_{n>0} \frac{\langle\Psi_0|L_i|\Psi_n\rangle\langle\Psi_n|L_j|\Psi_0\rangle}{E_n - E_0}\,, \tag{4.14}$$

where the interaction of states Ψ_n, of energies E_n, with the ground state are governed by the their orbital momentum and the spin-orbit coupling strength of λ.

In an applied magnetic field, $\mathbf{B}$, this then leads to a spin Hamiltonian of the form

$$\mathscr{H}_Z \;=\; \mu_B \mathbf{S}\underline{\underline{g}}\mathbf{B}_0\,. \tag{4.15}$$

The $(2S+1)$ eigenstates of such a Hamiltonian can then be labeled by the azimuthal quantum number m_z, where the energy of the m_z state is then $\mu_B g B_0 m_z$, with g having an angular dependence on the applied magnetic field of the form

$$g^2 \;=\; \sum_i g^2_{ii}\cos^2\theta_i\,, \tag{4.16}$$

where the angles θ_i are between the applied magnetic field and the principle axes of the diagonal form of $\underline{\underline{g}}$. This then allows the principle values of $\underline{\underline{g}}$ to be obtained for tetrahedral semiconductors by rotating $\mathbf{B_0}$ within a (110) plane, giving the change in the resonant magnetic field between the major cubic symmetry axes $\langle 110\rangle$, $\langle 100\rangle$ and $\langle 111\rangle$. This then allows the symmetry of the centre to be deduced.

Additionally for paramagnetic centres, there are further hyperfine splittings of the Zeeman levels due to interactions with the spins of nearby nuclei. The additional term this introduces into the Hamiltonian is then of the form

$$\mathscr{H}_{\text{Hyp.F.}} = \mathbf{S}\underline{\underline{A}}\mathbf{I} \,, \tag{4.17}$$

with the interaction, $\underline{\underline{A}}$, between the electronic spin, $\mathbf{S}$, and the nuclear spins $\mathbf{I}$ taking the general form [87]

$$\mathscr{H}_{\text{Hyp.F.}} = -\frac{8\pi}{3} g\mu_B g_n \mu_N |\psi_s(0)|^2 \mathbf{I}\cdot\mathbf{S} - g\mu_B g_n \mu_N \left(\frac{\mathbf{I}\cdot\mathbf{S}}{r^3} - \frac{3[\mathbf{I}\cdot\mathbf{r}][\mathbf{S}\cdot\mathbf{r}]}{r^5}\right) \,,$$

with g_n being the nuclear g value and μ_n the nuclear magneton. The wavefunction of the unpaired spin, ψ_s, is considered at the nuclear core with which the hyperfine-splitting arises. The first term in this correction is the isotropic Fermi-contact term (correcting for the point-dipole approximation at the nuclei), and the second anisotropic term is due to the types of nuclear spin present near the electron, and also the character of the electronic spin, i.e., s, p, d, etc. The additional energy splitting arising from these terms then takes the form

$$\Delta E = A m_s m_i \,, \tag{4.18}$$

with

$$A^2 = \sum_i A_{ii}^2 \cos^2\theta_i \tag{4.19}$$

as with the g factors of the normal Zeeman splitting. The $(2I+1)$ levels are then superimposed over the Zeeman split electron resonance.

The double-resonance condition of both the electronic and nuclear Zeeman splittings can be then probed via the ENDOR method. The typical continuous-wave form of the technique uses a microwave source to saturate the electronic Zeeman transition at the resonant energy. The nuclear transition is then selectively excited at radio frequency, leading to a measurable change in the electronic resonance condition. There are also alternative techniques based on pulsed excitation of the nuclear transitions [88, 89].

4.5 Uniaxial stress

The preceding techniques can all be coupled with uniaxial stress to determine the symmetry and reorientation barriers of defects. If stress is applied along a crystallographic direction of a particular class such as [100], [110] or [111], this will have the effect of lowering the symmetry of defects in the material.

This lowering of symmetry is particularly noticeable in the case of orientational degeneracies of non-cubic defects, leading to distinct populations of defects oriented amongst what become separable crystallographic directions. Kaplyanskii [90] tabulated the stress induced shifts between non-degenerate vibrational levels of defects of the seven crystalographically allowed types of symmetry for cubic crystals. The non-degenerate modes shift according to the applied stress σ_{ij}, giving a second-rank piezospectroscopic tensor, $\underline{\underline{A}}$, such that

$$\Delta\nu = \sum_{ij} A_{ij}\sigma_{ij} \; . \tag{4.20}$$

The symmetry of the defect under stress can reduce the number of independent components of the tensor. For the case of degenerate modes, the stress induced splittings also can contain contributions from the breaking of the mode degeneracies. The splittings of such modes have been tabulated by Hughes and Runciman [91].

Similarly the perturbation of the 3942 cm^{-1} silicon zero-phonon photoluminescence line has been tabulated by Davies *et al.* [92] for small stresses on defects of various symmetries, where the resulting spectral splitting are obtained by considering the transitions between all of the allowed combinations of split ground and excited states.

In addition to revealing symmetry, photoluminescence measurements which are made while under [110] stress using plane-polarised light can show whether a transition is due to an electrical or magnetic dipole. This occurs since spectra taken along $[1\bar{1}0]$ and [001] are identical for [110] plane-polarised light if the transition occurs through an electric dipole transition. Alternatively, if the signal is the same for the two orientations for light polarised perpendicular to [110], the transition must be magnetic-dipole allowed [90, 93]. This occurs since the [110] stress creates compressive and sheer stresses that couple distinctly to magnetic and electrical dipoles.

Stress also gives a method of determining reorientation barriers and rates. Since applied stress splits the degeneracy of sites for a defect, given the opportunity these centres will reorient into the alignment that is the lowest energy. The population of the higher energy forms of the defect will then follow a Boltzmann distribution if given sufficient time to relax, hence if the defects are cooled down under stress (or indeed formed whilst under stress as with the vacancy defects in chapter 6) then there will be a large population in the preferred alignment as compared to the others. If the fraction of the defects in this alignment is then monitored by a method sensitive to dichroism, such as optical spectroscopy with polarised light, then if the sample is re-heated without applied stress the dichroism will decay by the defects thermally reorienting. Alternatively, if the relaxed system is stressed, the growth in the population along the low energy orientation also gives the reorientation rate, and hence the activation energy when this is considered as a function of temperature.

Chapter 5

Calculations on $H_{2\,i}$ in crystalline silicon

The existence of interstitial molecular hydrogen in silicon was first proposed by theory in 1983 [94, 95]. Since then a large number of calculations on this system have been published employing a diverse selection of methods, initially just on the structure and energetics of the molecule, and its relative energy with respect to other hydrogen species in silicon. Later calculations, in addition to these properties, also obtained the bond–stretching frequency for the interstitial molecule.

The general consensus of the various authors is that the equilibrium position of the single molecule within the lattice is at the tetrahedral site. As to the alignment of the molecule at that site, the majority view (particularly in the later papers) is that the molecule lies with its H–H bond aligned along a [100] direction, but some authors state that a [111] alignment is preferred. Where calculated, the barrier to rotation of the molecule is always found to be less than 0.1 eV in later calculations this is typically > 0.01 eV.

The calculated vibrational modes for the molecule range from 3396 to 4849 cm^{-1} between authors. Observation of two *distinct* molecular forms of hydrogen in silicon (see section 5.1.2), with frequencies differing by 500 cm^{-1} led to two different views amongst theoreticians, who favoured one type or the other of the experimental data as being due to interstitial molecules, but this has now been largely resolved.

Unfortunately a further problem subsequently arose, in that while all of the calculations find that the molecule should be free to rotate within the crystal, much like in the case if H_2 in GaAs, the experimental evidence is that the molecular motion is strongly restricted. This discrepancy has not been resolved, and remains an area of active study in several groups.

5.1 Previous work

5.1.1 Theoretical calculations

Mainwood and Stoneham [94, 96] used CNDO in a cluster based set of calculations, with a static lattice. They found the molecule to be most stable at the $T_{d\,i}$ site aligned along [100] (they also considered [110] and [111] alignments), with a barrier of 0.093 eV to rotation about the [110] axis. The barrier to molecular diffusion, via $T_{d\,i} \rightarrow \mathrm{Hex}_i \rightarrow T_{d\,i}$, was calculated to be 0.95 eV, and the relative stability of the molecule with respect to $2 \times \mathrm{H}^0$ atoms in the lattice was found to be ~ 2 eV/atom.

Corbett *et al.* [95] performed MNDO based calculations using a $Si_{14}H_{20}$ cluster. The molecule was found to be most stable at the $T_{d\,i}$ site, and relative stability of H_2 with respect to two H^0 atoms at M sites in the lattice was found to be 1.6 eV, with a barrier to diffusion through the lattice of 2.7 eV. The molecule was thought to be [111] aligned.

Deák *et al.* [97, 98] using MINDO/3 cyclic cluster calculations on Si_{32} predicted that both H_2^* (see chapter 7) and also isolated H^0 were more stable than H_2 by 0.49 and 0.24 eV respectively. The molecule was found to be nearly degenerate in both the [100] and [111] alignments, with a 0.56 eV barrier to diffusion.

Chang *et al.* [99, 100] using a LDA-DFT Si_8 supercell, preferred the [111] alignment of the molecule at $T_{d\,i}$, and found the barrier to rotation to be negligible. In both $p-$ and $n-$type silicon, the molecule was found to be 2.7 eV more stable than isolated H^0_{BC} and 0.4 eV more stable than H_2^*. The molecular bond length was calculated to be 0.88 Å in the earlier paper and 0.85 Å in the later.

Briddon *et al.* [101] and Jones [102] using an LDA-DFT approach (AIMPRO), in $Si_{51}H_{52}$ clusters, found the molecule possesses a bond length of 0.85 Å when at the $T_{d\,i}$ site. The vibrational frequency of the molecule at this site was calculated to be 3561 cm^{-1}. The most stable alignment of the molecule was found to be along [100] with a diffusion barrier of greater than 1 eV. The molecule was calculated to be 3.3 eV more stable than separate H^0_{BC} atoms.

Maric *et al.* [103] and Estreicher *et al.* [104] using both PRDDO in 14 and 44 atom clusters, and also Hartree–Fock calculations in 14 atom clusters, found the molecule at the tetrahedral site to have a bond length of 0.70 Å and a barrier to rotation of less than 0.1 eV. The preferred alignment was along [111], and the barrier to diffusion was calculated to be less than 2 eV. The molecule was found to be higher in energy than H_2^* in both calculations, by 0.82 eV for the PRDDO and 0.34 eV for the HF.

Van de Walle *et al.* [105–110] using a 32 atom supercell LDA-DFT method, found the molecule at the $T_{d\,i}$ site to favour [100] over [111] by 0.01 eV, with a 0.86 Å bond length, which gave rise to a molecular frequency of 3396 cm^{-1}, based on a polynomial fit of the molecular bond-length vs. energy [111]. Based on an examination of trends in the molecular bond length in other semiconductors [109], a correlation between the interstitial charge density and the downward shift of the molecular frequency from that observed in free H_2 molecules was suggested (similar to the downward shift in molecular stretch frequency noted by Nørskov for H_2 in jellium [112]). The barrier to migration of the molecule was found to be 1.1 eV and the binding energy of the molecule was calculated to be 2 ± 0.5 eV (i.e., 0.54 and 1.74 eV lower in energy than H_2^* and H_{BC}^0).

Kim *et al.* [113], using empirical tight-binding calculations in 54 atom supercells found the H–H distance to be 0.74 Å with a molecular stretching frequency of 4396 cm^{-1} at $T_{d\,i}$.

Nakamura *et al.* [114–118] using a combination of HF, MP2–HF and DFT with both GGA and LDA in $Si_{10}H_{16}$ cluster calculations found the molecule to preferential align along [100], with a rotational barrier of the order $\sim$ 30 meV. The frequencies of the molecule obtained using different calculational schemes are summarised in table 5.1. The earlier work [115–118] attempted to correct the harmonic frequencies by scaling by 0.89 to obtain an estimate of the anharmonic frequency. The frequencies obtained agree approximately with those of Murakami *et al.*, but contrast with those of Pritchard *et al.* and Leitch *et al.* (see section 5.1.2).

Okamoto *et al.* [119] using GGA and LDA-DFT supercell calculations with Si_{32} and Si_{64} found a preferred molecular alignment along [100], with a barrier to rotation of 0.003 (0.002) eV for LDA (GGA) calculations. The barrier for diffusion of the molecule was calculated to be 0.91 (0.89) eV for LDA (GGA), with a frequency of the molecule in the [100] and [111] alignments of 3020 (3310) cm^{-1} and 3078

Method	Basis	H–H (Å)	ω_0 cm^{-1}
Restricted-HF	3-21G	0.708	4849
	6-31G*	0.709	4969
MP2–HF	3-21G	0.727	4649
LDA-VWN	6-31G*	0.789	3773
GGA-BPW91	6-31G*	0.764	4020
GGA-BLYP	6-31G*	0.747	4380
GGA-B3LYP	6-31G*	0.739	4468
	6-31+G*	0.740	4470
HF [116]	3-21G	0.702	4470
HF [118]	3-21G	0.709	4423

Table 5.1: Calculated frequencies for a H_2 molecule in $Si_{10}H_{16}$ clusters after Nakamura *et al.* [114] (note: The work in earlier papers [116, 118] presents a frequency scaled by 0.89 in an attempt to allow for anharmonicity.)

(3363) cm^{-1} for LDA (GGA).

Further calculations by Okamoto *et al.* [120] using Hartree–Fock, HF with MP2 corrections, DFT, and hybrid DFT, with a variety of basis sets and cluster sizes in the GAUSSIAN 94 code, showed that pure HF calculations on the hydrogen molecule at the tetrahedral site overestimate the molecular stretch frequency, typically by several hundred cm^{-1}. This error in frequency was ascribed to a combination of the lack of electron correlation in HF, and the small clusters used by Nakamura *et al.* In addition, they found that including Si 3*d* orbital functions in the basis caused a large downward shift in the molecular stretch frequency, showing the importance of a sufficiently large basis set. Frequencies were calculated for H_{2i} using a polynomial fit to the energy as a function of bond length [111], which included a correction for anharmonicity. Unfortunately, due to computing constraints, the largest relaxed cluster that was used in this work ($Si_{60}H_{56}$) was not sufficient to ensure that the molecular vibrational frequency was converged with respect to cluster size, but they estimated the frequency, with a correction for anharmonicity, of the molecule to be of the order of 275 cm^{-1} lower than of free hydrogen, when the molecule is encapsulated in the silicon lattice (see table 5.2 for a summary of their calculations).

Cluster	Method	Basis	$\omega_{\text{anharmonic}}$ cm^{-1}
$Si_{10}H_{16}$	B3LYP	6-31G	4162
		6-31G(d)	4002
		6-31G(2d)	4001
		6-311G(d)	3999
		6-311G(2d)	4025
		6-311G(2d, p)	4011
	HF	6-31G(2d)	4568
	LDA(SVWN)		3436
	GGA(BLYP)		3859
	B3LYP		4001
	MP2-HF		3992
$Si_{60}H_{56}$ relaxed	B3LYP	6-31G(d on central Si_{26})	3893
$Si_{10}H_{16}$ fixed		6-31G(d)	3945
$Si_{60}H_{56}$ fixed		6-31G(d on central Si_{26})	3770
$Si_{84}H_{64}$ fixed		6-31G(d on central Si_{26})	3693

Table 5.2: Calculated frequencies for a H_2 molecule in silicon clusters after Okamoto *et al.* [120]. Clusters labeled fixed did not did not have the atoms surrounding the hydrogen molecule relaxed.

LDA and GGA supercell calculations by Kim *et al.* [121, 122] found broad agreement with the work of Okamoto, and also suggested the presence of an additional localised mode at around 650–700 cm^{-1}.

5.1.2 Experimental work

Markevich *et al.* [123–125] after heating n-type float-zone (fz) and Czochralski (Cz) silicon to 1200 °C in hydrogen or deuterium atmospheres of 1.5 atm, and then annealing at a temperature between 40 and 80 °C, produced an infrared active defect with a line at 1075.1 cm^{-1} in H_2 treated samples (consisting of two closely spaced peaks), which shifted *upwards* to 1076.3 cm^{-1} in deuterated samples. The defect was initially thought to be a $O_{BC}H_i$ pair similar to those observed by Bech Nielsen *et al.* [126], but in the light of the later work by Pritchard *et al.* [127], was instead re-identified as an interstitial H_2 molecule sited near a bond-centred oxygen atom. Using the absorption due to the nearby $^{18}O_{BC}$ peak (which accounts for 0.2% of natural oxygen), and assuming that the oscillator strength of the two complexes were similar, the concentration of the OH_2 complex was estimated to be 7×10^{14} cm^{-3}, i.e., about 10% of the solubility limit of atomic H at 1200 °C [128]. The presence of the defect was found to introduce no new DLTS levels in the sample [124]. The annealing properties of the defect [125] were found to be unusual, since the defect dissociated on annealing above 80 °C, but reformed upon annealing at lower temperatures (typically around 40 °C), whilst heating above 100 °C irreversibly destroyed the complex. The binding energy of the complex was measured to be 0.28 ± 0.02 eV and the barrier to diffusion of the dissociated components was found to be 0.78 ± 0.05 eV [125].

Murakami *et al.* [129–132, 116, 133–135] after treating $n-$type fz-Si with a remote hydrogen (deuterium) plasma observed a Raman line at 4158 ± 3 (2990 ± 2) cm^{-1}, with a line–width at half maximum intensity, $\Delta\omega$, of 34 (20) cm^{-1}. An additional mode at 590 cm^{-1} in the H treated samples was also observed. These lines were interpreted to be due to the high frequency stretch mode of *interstitial* H_2 (which was not resolved into the expected ortho- and para-lines) and an associated rotational line of ortho–H_2 (see section 5.4 for a discussion of the ortho/para forms of hydrogen). The width of the H_2 related stretch mode decreased with increasing substrate temperature during hydrogenation.

The intensities of the observed lines reached a maximum for hydrogenation of the Si-substrate at a temperature of 400 °C. Due to the lack of a change in the Raman Si longitudinal-optical phonon mode at 521cm^{-1} on hydrogenation, and the narrowing of an EPR H-related line for hydrogenation above 300 °C, Murakami *et al.* suggested that the hydrogen was in the form of molecules sited in the undamaged lattice, and not trapped inside extended defects.

Modes characteristic of platelets were observed at 2100 cm^{-1} with an additional mode at around 2140 cm^{-1} in the hydrogen treated samples. Murakami *et al.* claimed that these features were due to a distinct, damaged, surface layer based on SIMS profiling and Raman spectroscopy of samples with varying thicknesses of surface material etched away [116, 134]. These measurements showed a peak in hydrogen concentration corresponding to the defects responsible for the ~ 2100 cm^{-1} platelet-like modes at ~ 250 Å, and a slightly deeper peak for the H_2 concentration at ~ 750 Å (with measurable penetration of H_2 to between 1000 and 3000 Å). In addition the peak concentrations for the two species occurred at different temperatures (500 K for the SiH modes, and 650 K for the H_2 modes). Isothermal annealing [132] gave an activation energy for the dissociation the H_2 species as 1.6 ± 0.4 eV, with the H_2 related line annealing out at about 500 °C [132]. The SiH related modes annealed out over the range 250–500 °C, with a small peak at 2120 cm^{-1} remaining at 450 °C [132], this peak has been implicated as being important precursor in exfoliation of silicon [136].

Studies using crystalline, micro-crystalline (with ~ 7 nm grain size) and amorphous silicon [116, 134], gave a ratio for the intensity of the H_2 related modes of $1 : \frac{1}{5} : 0$ for the three materials treated with a 1×10^{16} cm^{-2} area doses of hydrogen [129].

Treatment with deuterium plasma followed by hydrogen plasma gave rise to no mixed HD stretch mode [131], suggesting that desorption of hydrogen from platelet structures did not form part of the mechanism for the formation of H_2.

Leitch *et al.* [137–139] also produced the H_2 and D_2 related complex in $n-$ and $p-$type fz silicon, with an additional HD form with a Raman active stretch mode at 3629 ± 2 cm^{-1}. $\Delta\omega$ for the lines was somewhat smaller than those of Murakami *et al.* at ~ 25 cm^{-1}. This contrasts with the much sharper and lower frequency H_2 line previously observed in GaAs by Vetterhöffer *et al.* [140] (see table 5.3), in which the ortho and para forms of H_{2i} were resolved, which coupled with uniaxial stress measurements [141] demonstrates that the molecule freely rotates in GaAs. The mode at 590 cm^{-1} attributed to a rotation of o-H_2 by Murakami *et al.* was not observed.

As with Murakami *et al.* the line width of the H_2 related mode decreased as the substrate temperature used during hydrogenation was increased. SiH related modes were always seen with the H_2 related modes, and the SiH related mode was always intense if the H_2 mode was also intense. Based on attenuation of the laser intensity within the sample, the peak H_2 depth was estimated as ~ 250 nm, and the SiH complex depth as $100 - 200$ nm. Etching away 25 nm of the surface reduced the intensity of both the H_2 and SiH peaks to 85%. Annealing of the samples gave a dissociation energy of for the H_2 complex as 2.3 eV (the defect disappeared at about 400 °C), which was comparable to the 2.2 eV free energy of platelets [142].

The peak at 4157 cm^{-1} was observed to have a shoulder at 4129 cm^{-1} which gained intensity from the main, higher frequency peak, to become the dominant line after annealing at 475 °C. During annealing , the SiH related mode at 2132 cm^{-1} dropped in both intensity and frequency, falling to 2112 cm^{-1}.

	H_2	HD	D_2
ω (cm^{-1})	3925.9 (J=1)	3446.5	2842.6
	3934.1 (J=0)		
$\Delta\omega$ (cm^{-1})	3	8	7

Table 5.3: Experimental frequencies of H_2 molecules in crystaline GaAs after Vetterhöffer *et al.* [140]

The intensity of the SiH and H_2 related peaks as a function of sample temperature show two overlapping profiles [139] (see figure 5.1 for the profiles measured by Leitch *et al.*), and based on this, Leitch *et al.* suggest a two stage process leading to the formation of the H_2–related complex. First, formation of platelets occur in the silicon at around 150 °C, then these defects behave as trapping sites for any further hydrogen introduced into the material. This leads to a build-up of H_2 within the open space inside the platelets, and hence a gas–like environment. This is unlike the situation in GaAs, where platelets are not formed, and so *true* interstitial, H_2 is produced directly. Based on this model, Leitch *et al.* attempted hydrogenation of silicon at below 150 °C (and hence did not form the platelets necessary for defects of the type observed by Murakami *et al.*) and produced defects that at room temperature possessed Raman lines at 3601 and 2622 cm^{-1} for H and D plasma. On cooling these samples to 10 K, these lines shifted upwards to the frequencies observed by Pritchard *et al.* for isolated H_{2i} [143], with line widths of ~ 6 cm^{-1}. Like Pritchard *et al.*, no o-/p- splitting was observed in these lines.

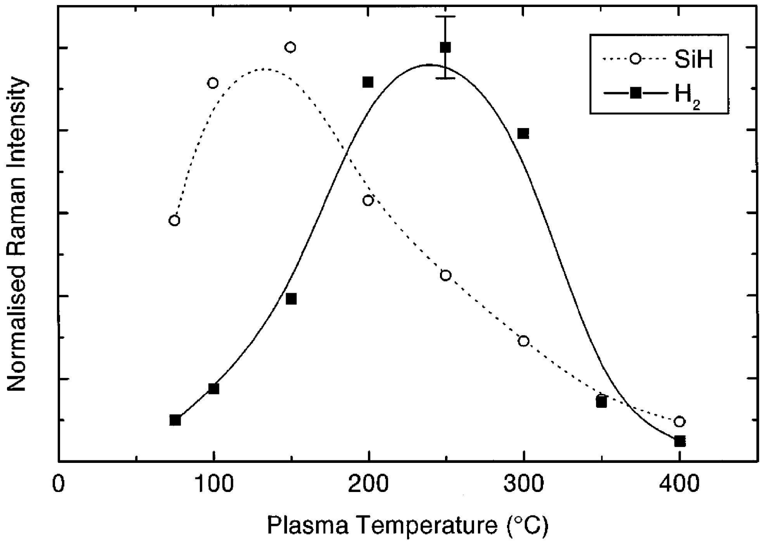


Figure 5.1: Overlapping Raman profiles for intensities of the $Si - H$ and H_2 related modes in silicon as a function of sample temperature during plasma hydrogenation of the material, after Leitch *et al.* [139].

Pritchard *et al.* [144, 143], performed hydrogenation of fz and also Cz silicon similar to that of Markevitch *et al.* in H_2, D_2 and mixed $H_2 + D_2$ atmospheres, for $n-$ and $p-$type Cz and fz silicon, at between 1000 and 1300 °C. Infrared spectra were then taken at 10 K. The line observed by Markevitch *et al.* was resolved into two components, a line at 1075.1 cm^{-1} with a shoulder at 1075.8 cm^{-1}. As with the previous samples, these lines shifted up to 1076.6 and $\sim$ 1076 cm^{-1} in the deuterated material. Satellite lines were detected around the modes, in a pattern consistent with the line being due to a defect containing bond-centred oxygen, and so having neighbouring silicon atoms of several isotopic combinations as found in the sample, i.e., ^{28}Si, ^{29}Si, ^{30}Si (this type of satellite lines were also observed for the nearby $^{18}O_{BC}$ absorption line). Based on the intensity of the $^{18}O_{BC}$ line, it was estimated that $\frac{1}{2}$ of the hydrogen present at the solubility limit [128] was involved in complexes of this type (assuming a single H atom was present in each complex).

The oxygen related line around 1075 cm^{-1} was decomposed into two Voight–type profiles with widths of $\sim$ 0.9 cm^{-1} and a separation of 1.1 cm^{-1} in the case of the H_2 treated material, and widths of $\sim$ 0.6 cm^{-1} and separation of 0.5 cm^{-1} in the D_2 treated material. The relative intensities of the two profiles were 3 : 1

for the 1075.1 cm^{-1} and 1075.8 cm^{-1} H–related lines and 2 : 1 for the 1076.6 and 1076 cm^{-1} D–related lines. In the case of mixed $H_2 + D_2$ treated material, the oxygen–related line fell at 1076.3 cm^{-1} with a width of 0.7 cm^{-1}.

In addition a new set of higher frequency lines, which were very sharp ($\Delta\omega =$ 0.1 – 0.2 cm^{-1}) were also observed in thicker ($\geqslant$ 5 mm) samples (see table 5.4 for a full list of the observed lines). The intensities of ν_1 and ν_2 were found to be correlated with that of the oxygen–related line, but ν_3 was anti-correlated to it, ν_3 was also found to be present in low oxygen fz material. Annealing the Cz samples at between 70 – 110 °C, resulted in a transfer of intensity from ν_1 and ν_2 to ν_3, whilst from 110 – 320 °C the intensity of ν_3 remained unchanged, but that of ν_1 and ν_2 continued to decrease.

	H_2	HD	D_2
ν_1	3788.9	3304.3	2775.4
ν_2	3730.8	3285.3	2715,2716.0
ν_3	3618.3	3264.8	2642.5
Main O line	1075.1	1076.3	1076.6

Table 5.4: Local vibrational modes, cm^{-1}, of H related modes observed in silicon by Pritchard *et al.* [144, 143] after in-diffusion of hydrogen into material containing oxygen.

The possibility that the high frequency ν_1 and ν_2 lines were due to interstitial water molecules was ruled out, since water would be expected to have an intense molecular wag mode at $\sim$ 1595 cm^{-1}. Also, mixed HDO would be expected to have a high frequency mode at $\sim$ 3707 cm^{-1} little shifted from H_2O and a low frequency mode at $\sim$ 2727 cm^{-1} similar to that of D_2O, unlike the observed lines mid way between the H_2 and D_2 cases at 3304.3 and 3285.3 cm^{-1}. Instead it was suggested that the modes were due to a single hydrogen molecule trapped at one of two sites near O_i in the lattice (the possibility of the two lines being due to ortho/para states of the molecule was discounted, due to the large size of the splitting).

Based on the assumption that the effective charge of the oxygen related mode in the complexes responsible for ν_1 and ν_2 was little changed from that of isolated

O_{BC}, i.e., 3.5 e [145] in the H_2 treated material, this then gave the effective charges for the ν_1 and ν_2 lines associated with the two different oxygen peaks as $\eta_{\nu_1} = 0.07\ e$ and $\eta_{\nu_2} = 0.14\ e$, assuming that the main O peak was associated with ν_1 and the shoulder with ν_2. The alternative assignments gave $\eta_{\nu_1} = 0.12\ e$ and $\eta_{\nu_2} = 0.08\ e$. In the case of the D_2 treated material, the first set of assignments gave $\eta_{\nu_1} = 0.06\ e$ and $\eta_{\nu_2} = 0.17\ e$, which was considered to be more consistent with the H_2 samples than the alternative alignment which gave $\eta_{\nu_1} = 0.09\ e$ and $\eta_{\nu_2} = 0.12\ e$. Using this preferred assignments, this gave η_{ν_3} to be $\sim 0.1\ e$.

The realisation that the O–related complex contained *two* H atoms, then changed the concentration of hydrogen in these defects to 1×10^{16} cm^{-3} for hydrogenation at 1300 °C, and 4×10^{15} cm^{-3} for 1200 °C, which was considered to be in good agreement with previous SIMS profiling of similar material [146].

In the earlier paper, the structure responsible for ν_3 was not identified, but later, more detailed annealing studies [143] using sequential annealing for 30 – 60 minutes at temperatures of between 35 – 130 °C followed by an immediate cool-down to 10 K for measurement, found 1.3 – 2.6 $\times 10^{23}$ cm^{-3} sites that the complex responsible for ν_3 could be present at, i.e., the concentration of $T_{d\,i}$. The samples used were 17 mm thick Cz silicon crystals, in which no extra modes around 1800 – 2300 cm^{-1} (see chapter 8) had been observed after hydrogenation (i.e., little lattice damage). Based upon a two centre statistical model for the process $I_{(\nu_1+\nu_2)} \rightarrow I_{\nu_3}$ (if further trapping sites were present, then they would have to possess similar properties to those of ν_3 for the reversible annealing behaviour to occur, but also be IR inactive) and assuming the structural degeneracy of the two O–H_2 complexes to be 6 (i.e., the number of equivalent $T_{d\,i}$ sites around O_{BC}), then fitting the annealing data to

$$\frac{[O_i\text{–}H_2]_{(T)}}{[\nu_3]_{(T)}} = \left(\frac{g_1}{g_2}\right) \exp\frac{\Delta E}{kT} \tag{5.1}$$

gave $\Delta E = 0.26 \pm 0.02$ eV, in good agreement with the binding energy found by Markevitch *et al.* If the value of g_1 is 6×2 (six cage sites for two complexes), then g_2 is 1.3 – 2.6 $\times 10^{23}$ cm^{-3}, which is approximately equal to the number of sites an isolated molecule could occupy times the number of orientations at an isolated site, which would be 4 for [111], 6 for [110], or 3 for [100]. The gives the number density of sites for the isolated molecule to be between 6×10^{22} and 1×10^{22} cm^{-3}, i.e., comparable to the concentration of $T_{d\,i}$.

Zhou and Stavola [147, 148] studied the stress response of the 3618.4 cm^{-1} line assigned to the isolated H_2 molecule. Using samples with $[H_2]\sim 5\times 10^{15}$ cm^{-3} they obtained a nearly diagonal piezospectroscopic tensor $\underline{\underline{A}}$, in terms of the positive compressive stress and the resulting stress-split shifts of the component of the line in cm^{-1}/GPa, of the form

$$\begin{pmatrix} 9.6 & & \\ -1.8 & -15.6 & \\ 1.4 & -2.8 & 1.1 \end{pmatrix}. \tag{5.2}$$

The shifts obtained were characteristic of a triclinic centre, which was taken to probably be of C_1 symmetry, which is in strong contradiction with the theoretical predictions of D_{2d} or C_{3v} symmetry. The tensor is rotated less than 14° from the $\langle 100\rangle$ cube axes, and the large effect caused by $\langle 100\rangle$ stresses suggest the molecule is aligned approximately along a $\langle 100\rangle$ direction, but the exact orientation and placement of the molecule could not be determined in this study.

The molecule was found not to reorient under stress at 4.2 K. Furthermore based on the effect of cooling down the sample from room temperature under stress, Zhou and Stavola concluded that there must be a barrier of at least 60 K to molecular rotation.

Despite the low symmetry of the centre, as with previous studies, no splitting in the molecular stretch mode for HD molecules was observed. The suggestion was made that only one of two possible forms of the molecule is observed, since if there is sufficient difference in energy between the two forms then the higher energy form will reorient. The anomalous frequency of the HD line when compared with H_2 and D_2 in silicon and GaAs was taken to support this model.

5.2 Calculations on the O_i–H_{2i} complex

5.2.1 Interstitial water in silicon

To investigate the structure and vibrational properties of the water molecule in silicon both as an isolated molecule and bound to O_i, calculations using a 148 atom

H-terminated cluster centred on the tetrahedral interstitial site were performed. The basis consisted of independent *s* and *p* Gaussian orbitals, with 5 (4) different charge (wavefunction) Gaussians with differing exponents sited at each Si atom, and three at each H atom of the molecule. An additional 6 charge and wavefunction Gaussians were also placed on the oxygen atom(s). A fixed linear combination of two Gaussian orbitals was sited on the terminating H atoms. In addition another set of Gaussian *s* orbitals were placed at each Si–Si bond centre in both the charge and wavefunction basis sets. All atoms, except the terminating H ones, were allowed to relax. The second derivatives of the energy were found for the water molecule, and also the O_i (and it's neighbouring Si atoms), where present.

The O-H bonds of the isolated molecule in free space are calculated to be 0.960 Å and the H-Ô-H angle equal to 107.9°. These are close to the experimental values of 0.957 Å and 104.5°. The vibrational modes are given in table 5.5 and compared with experimental values for free H_2O [149].

It is noticeable that although the two O-H stretch modes at 3521.6, 3352.1 (H_2) and 2571.8, 2432.7 (D_2) cm^{-1} lie close to those observed in Si, there is no mode in the HDO case lying in the region mid-way between these values. This is because there is little direct coupling between the H atoms. Thus the H-D modes observed by Pritchard *et al.* around 3300 cm^{-1} cannot be explained. This argument is shown to be unaffected by inserting the molecule into a cluster containing O_i (see table 5.5).

The O-H lengths and angle of the water molecule placed near an interstitial T_d site are 0.99 Å and 109°. The effect of the cluster is then to weaken the O -H bonds. This molecular species is probably metastable and would in practice dissociate forming H_2 and O_i defects. The local modes are given in table 5.5 and demonstrate the absence of any H-D mode around 3000 cm^{-1}.

The mode at 1100 cm^{-1} is the antisymmetric stretch mode of O_i. This drops by 54 cm^{-1} for $^{18}O_i$ — the oxygen mass on the water remaining ^{16}O. In the D_2 case, the 1512 cm^{-1} bend mode drops and couples strongly with the 1100 cm^{-1} O_i mode. In fact the two modes at 1095 and 1115 cm^{-1} involve substantial amplitude on O_i. This can be seen as these modes shift to 1111 and 1044 cm^{-1} for $^{18}O_i$. For a fictional mass on the hydrogen atoms of 2.5 a.m.u, the bend mode of water lies at 1011.7 cm^{-1} and below that of the O_i mode which has been pushed *upwards* to 1100.6 cm^{-1}.

Free Water

(H_2O)	(HDO)	(DHO)	(D_2O)
Experiment			
3756	3707		2788
3657	2727		2671
1595	1402		1178
This work			
3755	3710	3708	2752
3659	2691	2692	2635
1574	1380	1379	1153

Water in Si

$(H_2O)_i^{16}O_{BC}$	$(HDO)_i^{16}O_{BC}$	$(DHO)_i^{16}O_{BC}$	$(D_2O)_i^{16}O_{BC}$	$(H_2O)_i^{18}O_{BC}$
3521.6	3354.9	3519.2	2571.8	3521.6
3352.1	2560.3	2444.2	2432.7	3352.1
1512.3	1373.0	1309.6	1114.7	1512.3
1100.3	1100.1	1099.2	1095.3	1046.4
723.5	689.4	697.6	682.5	723.0
708.6	681.6	653.2	651.6	708.5
659.6	616.0	617.0	594.8	659.6
591.2	576.4	580.9	576.4	591.1

Table 5.5: Calculated local vibrational modes (cm^{-1}) for the $(H_2\,^{16}O)_i$ molecule near O_{BC} in silicon or for isolated water molecules.

This demonstrates that a molecule of water can cause an upward displacement of the local mode due to O_i in favourable cases. However, as emphasised above, the observed H-D vibrational modes rule out this molecule as responsible for the effect.

5.2.2 $O_i + H_{2i}$

There are several obvious locations for the molecule close to O_i in the silicon lattice but since the binding between the molecule and the O atom is so weak, these are close in energy and difficult to distinguish. Furthermore, it seems that there are a large number of configurations differing in the orientation of the molecule and all of these have to be considered as candidate structures.

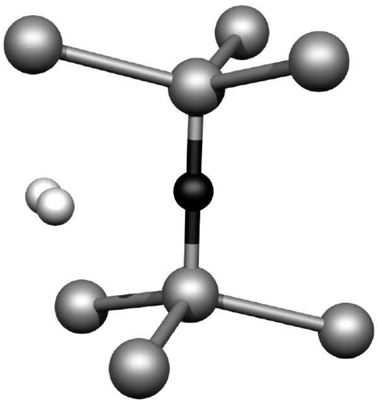


Figure 5.2: Schematic illustration of the O_iH_2 defect suggested to have modes similar to one of the observed O–H defects.

It is supposed that the molecule resides near a T_d cage site and oxygen decorates one of the nearby Si-Si bonds. The case of the O atom bridging a $[11\bar{1}]$ Si-Si bond which forms part of the cage surrounding the T_d site as shown in figure 5.2 is first considered. The molecule in the relaxed configuration then is orientated almost perpendicular to the Si-O-Si bond, and nearly along $[1\bar{1}0]$. The H-H bond length is 0.77 Å and each H atom is about 2.37 Å from O. The molecule has been pushed slightly away from the T_d site as the two nearest Si atoms (which are themselves not bonded to O) are 2.19 Å away. The vibrational modes of the cluster are given in table 5.6.

The H_2 stretch mode at 3855 cm^{-1} is again lower than that calculated for an isolated molecule. The molecular frequency depends on the size of the surrounding cage. The 3855 cm^{-1} mode lies close to the experimental modes around 3750 cm^{-1} associated with oxygen. The two H atoms are almost equivalent and this explains why only one H-D mode is present although the calculations show a 10 cm^{-1} split-

$H_{2i} + {}^{16}O_{BC}$	$HD_i + {}^{16}O_{BC}$	$DH_i + {}^{16}O_{BC}$	$D_{2i} + {}^{16}O_{BC}$	$H_{2i} + {}^{18}O_{BC}$
3855.6	3339.5	3350.1	2726.7	3855.6
1129.8	1129.7	1129.6	1129.5	1078.1
733.3	685.6	690.6	678.6	731.4
696.2	663.2	654.6	653.7	685.3
655.3	650.1	645.0	630.9	649.0
643.6	627.6	627.4	536.3	638.9
630.6	539.5	542.6	533.6	622.6
576.7	536.0	536.1	532.8	576.1

Table 5.6: Local vibrational modes (cm^{-1}) of the $H_{2i} + O_{BC}$ complex in silicon shown in figure 5.2.

ting caused by deviations from ideal σ_h symmetry.

The mode at 1129.8 cm^{-1} is clearly due to O_i as it shifts downwards by only 0.3 cm^{-1} when D replaces H but by 52 cm^{-1} with ^{18}O. This is comparable with the 51 cm^{-1} shift observed in the 1136 cm^{-1} mode of O_i when ^{18}O replaces ^{16}O. The small downward shifts on replacement of H with D demonstrate that there is very little direct coupling between O_i and the molecule and this is essential to these arguments, based on anharmonicity, if an upward shift is to be seen in the deuterated case. It is not clear why the experimental O mode is shifted downwards by as much as 65 cm^{-1} from that of isolated O_i.

The modes at 733 to 576 all involve the movement of H, and have not been detected so far. They represent H_2 bend and librational modes. Their shifts with ^{18}O given in table 5.6 demonstrate that many of them also involve the movement of O.

Of particular interest here is the mode at 577 cm^{-1}. This represents a librational mode as the two H atoms are displaced almost parallel to their bond. It can be argued that the anomalous frequency shift of the 1075 cm^{-1} band is to be understood through an anharmonic coupling between an overtone, or combination band, of these low frequency modes and the O mode resulting in a Fermi resonance. Let $|n_O\rangle$ be the n'th oscillator state for the vibrations of the oxygen atom whose fundamental occurs at $\nu_{\mathrm{O}} = 1129.8$ cm^{-1}, and $|m_{\mathrm{H}}\rangle$ be the m th state for a mode whose frequency, ν_{H}, is about half that of the O_i mode. The states of the coupled system are then described by linear combinations of $|n_{\mathrm{O}}, m_{\mathrm{H}}\rangle$. The effect of anharmonicity, V, is to couple together these states and second order perturbation theory gives the shift in the energy of the $|1_{\mathrm{O}}, 0_{\mathrm{H}}\rangle$ state to be dominated by

$$\frac{\langle 1_{\mathrm{O}}, 0_{\mathrm{H}}|V|0_{\mathrm{O}}, 2_{\mathrm{H}}\rangle^2}{(\nu_{\mathrm{O}} - 2\nu_{\mathrm{H}})}. \tag{5.3}$$

This follows as the energy denominator is particularly small for these modes. The perturbation is negative for the 577 cm^{-1} mode and lowers the energy of the $|1_{\mathrm{O}}, 0_{\mathrm{H}}\rangle$ state, and hence that of the fundamental transition. On the other hand, when H is replaced by D, as the frequency of this mode drops below 575 cm^{-1}, then the perturbation acts to raises the energy of the state. Thus the two cases reinforce the tendency to depress the O mode in the H case below that of D. Another way of describing the effect is an anti-crossing between the O mode and an overtone of the librational mode (as shown in fig. 5.3). If this mechanism is correct, there have to be unreported modes in the 550 cm^{-1} region.

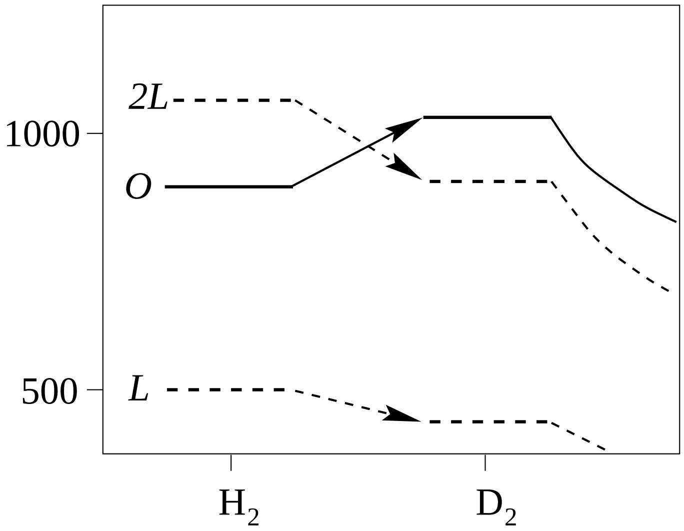


Figure 5.3: Schematic of the anti-crossing of the oxygen mode and an overtone of a lower energy librational mode, leading to the net upward shift of the oxygen related mode on increase of the hydrogen mass from ^{1}H to ^{2}H.

A second and almost degenerate configuration occurs when the molecule lies in the mirror plane containing the O_i. This has similar modes to the first and could account for the second defect which is observed. A third possibility is that the molecule lies along the Si-O-Si axis near a T_d site. This site is stable but the molecule is then close to the Si neighbour of O_i and this results in an O related mode that is strongly coupled with H in conflict with the experiment. However, the energy of this structure appears to be lower than that shown in fig. 5.2 by 0.5 eV. This may be a due to the proximity of O_i to the surface of the cluster and further investigations are needed to clarify the most stable configuration.

5.3 Calculations on H_{2i}

5.3.1 H_{2i} alignments and frequencies

To investigate the structure and vibrational properties of the isolated molecule, cluster calculations using a $Si_{84}H_{64}$ atom H-terminated cluster centred on the tetrahedral interstitial site were performed. The wave function basis consisted of independent s and p Gaussian orbitals, with eight different exponents, sited at each Si site and three at each H atom of the molecule. A fixed linear combination of two sets of s- and p-type Gaussian orbitals was sited on the terminating H atoms. In addition three Gaussian s and p orbitals were placed at each Si–Si bond centre. The charge density was fitted with eight independent Gaussian functions with different widths on each Si atom, and four (three) on the central (terminating) H atoms (with the basis on the terminating H again being a fixed combination). Three further Gaussians were placed at each bond centre. All atoms, except the terminating H, were allowed to relax. The second derivatives of the energy were found with respect to the positions of the two central H atoms.

The molecule is stable when oriented along [100], [111] or [110] at the T_d site. Displacing the molecule away from the T_d site results in a strong restoring force on the molecule back towards the site. For example if the molecule is displaced along [100], so that one of the H-atoms is positioned at the C site, the molecule relaxes back to a [100] alignment at the T_d site with an energy change of 2.6 eV.

The [100], [111] and [110] configurations at the T_d site are essentially degenerate in energy. The rotational barrier between the structures is also negligible, but the uncertainty in energy can easily exceed 0.1 eV, so the barrier is difficult to calculate accurately. This is similar to the situation in GaAs [150]. The vibrational modes are given in table 5.7. These are all close together and lie within 90 cm^{-1} of the observed ν_3 band. They are slightly lower than that found for the molecule trapped near oxygen (3855.6 cm^{-1}). The H–H length is found to be 0.785 Å for the [110] and [111] structures and 0.788 Å for the [100] alignment. These are slightly longer than the gas value (0.74 Å). It appears that the energy calculations are not sufficiently accurate to resolve the reorientation energy or the stable alignment. However, the vibrational modes do present us with a method to discriminate between the structures if the molecule is assumed to be statically aligned (see section 5.4 for the effects of molecular rotation, and an estimate for the barrier needed to prevent this happening).

Alignment	D_2	HD	D_2
[110]	3708.4	3217.1	2622.2
[111]	3713.0	3221.4, 3221.6	2625.5
[100]	3606.8	3128.6	2559.1
Expt.	3618.3	3264.8	2642.5

Table 5.7: Calculated frequencies (cm^{-1}) of H_2 molecules in Si with different alignments, and a comparison with those for the ν_3 mode observed by Pritchard *et al.*

The A_1 mode of molecules lying along [100], with D_{2d} symmetry, would be infra-red inactive; while those along [111] possesses inequivalent H atoms. The latter defect produces two distinct HD modes shown in table 5.7. As these are separated by 0.25 cm^{-1}, they would give two bands with intensities corresponding to the different populations of the defects, contrary to the observations. The ν_3 band could not arise from these defects if the HD splitting was as large as this. The [110] orientation is infra-red active with equivalent H atoms, and gives rise to only one mode in the HD case. Alternatively the molecule may possess lower symmetry, as suggested by the measurements of Zhou and Stavola, but the positions of the atoms in the molecule would have to be of greater inequivalence than for the molecule aligned along [111] to give sufficient energy to cause the reorientation necessary to depopulate one of the forms of the molecule.

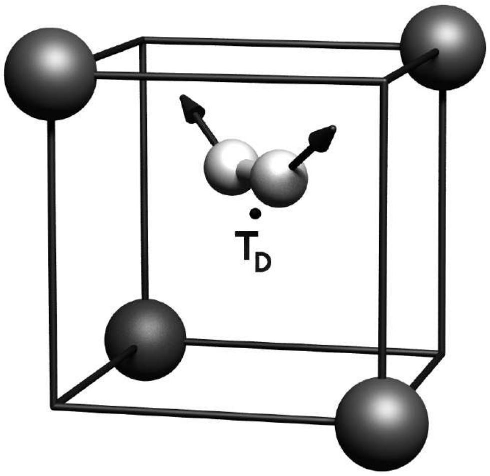


Figure 5.4: Schematic figure of the molecule at a $T_{d\,i}$ site aligned along [110]. The cube axes are aligned along $\langle 100 \rangle$. Arrows show an exaggerated movement of the atoms in the stretch mode leading to a transition dipole along [001]

The transition dipole moment of the molecule lying along [110] can be calculated directly. This is done by finding the change in the dipole moment of the cluster when the atoms are displaced according to their normal coordinates, as shown schematically in figure 5.4. The induced dipole lies along [001] and arises as there is a slight displacement of the H atoms along this direction. The effective charge of the induced dipole is 0.10 e and comparable with that of the molecule trapped by oxygen. For the [111] alignment, the effective charge is very similar at 0.08 e.

The low frequency of the molecule is due to the leaking of charge from the surrounding bonds screening the intra-molecular proton-electron attraction (see the work of Van de Walle in section 5.1.1).

5.3.2 Effects of lattice expansion on frequencies

Since the drop in the stretch frequency of the molecule from the gas–phase value is due to screening from the surrounding lattice, then if the cage size was increased, the charge density at the centre of the cage would decrease and the molecular frequency increase.

This is seen in fig 5.5 where a relaxed molecule is placed in a cage of increasing

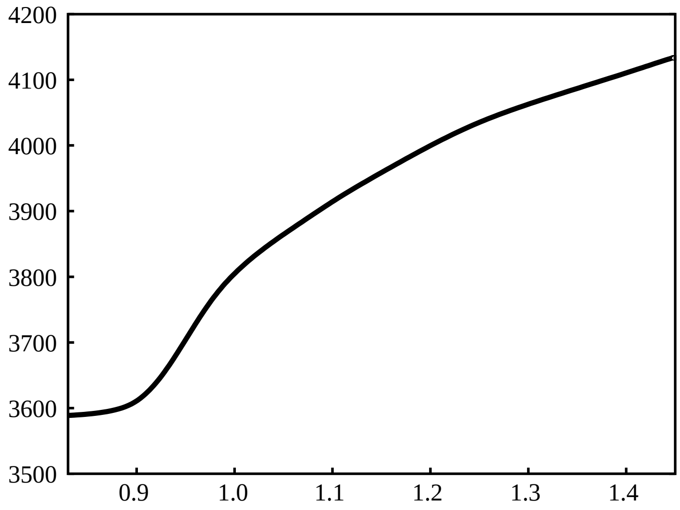


Figure 5.5: Variation of molecular stretch frequency (cm^{-1}) verses multiples of the equilibrium interstitial cage size for a [111] aligned molecule, demonstrating schematically the effect of surroundings on vibrational frequency.

size by scaling the surrounding Si-Si bonds. It is seen that the gas value is reached for voids about 1.4 times the actual cage. This shows that the large red-shift in the molecular stretch frequency is due to the size of the surrounding cage and is not an artifact of the calculational method. It is also consistent with the suggestion that the crystal field is sufficiently strong to freeze molecular rotation (see section 5.4).

These results strongly suggest that Raman active modes at 4158 cm^{-1} found in H–plasma treated Si are due to molecules in voids where the electron density arising from the lattice is small and are not due to isolated molecules at interstitial sites. A realistic model of the void must have surface dangling bonds saturated by H [151] and a correlation of the Raman signals due to these Si–H bonds and the 4158 cm^{-1} band is then to be expected, as has been found by Leitch *et al.* (see section 5.1.2). For a discussion of the possible candidates for the open defect responsible see chapter 6.

5.3.3 Molecular diffusion barriers and Fermi-level

The molecule is not stable at the centre of the puckered hexagon (Hex$_i$ site) but this site lies on a diffusion trajectory linking different $T_{d\,i}$ sites. The energy of a molecule constrained to lie at the H site is 0.72 eV greater than that at a $T_{d\,i}$ site and this then is an estimate of the migration energy of the molecule, which compares well to that measured by Markevich *et al.* and Pritchard *et al.*

The weakening of the interstitial hydrogen molecule bond, due to screening effects from the surrounding lattice, might be expected to have an effect on the barrier to the molecule passing through a hexagonal site on its diffusion path through the crystal. If the charge density within the lattice was severely affected by excess carrier generated by, for example irradiation, then the diffusion barrier would be changed, and this is indeed what we find. The diffusion barrier drops from 0.73 eV in the neutral case, to 0.33 and 0.46 eV for clusters charged to +2 or -2 e respectively, with an expansion of the bond length whilst the molecule is at the hexagonal site from 0.788 Å to 0.901 and 0.930 Å respectively (The bond length at the tetrahedral site is insensitive to charge state). Obviously such large charges are unphysical for doped material, but serve to illustrate that perturbation of the lattice charge density can effect the diffusion barrier of the molecule. Laser illumination can lead to substantial exciton concentrations, for example the intensity of 2.5 eV Ar light falls to $1/e$ within $\sim 1\ \mu$m of the surface [152], so if we were to assume all of the energy is deposited as electron-hole pairs with an energy of around twice the band gap (and a phonon of enrgy around 50 meV [152]), then for illumination of ~ 100 mW mm^2 this would lead to a carrier pair production rate of around 3×10^{23} per cm^{3}s.

Raman Spectroscopic studies performed on the plasma treated material, led to the unusual discovery [153] that the intensity of the hydrogen molecule signal within the crystal drops to essentially zero in the presence of extended laser illumination at cryogenic temperatures, this observation has recently been disputed however [154]. One possible model for this effect would be if the molecule were capable of binding an exciton; which was created by the illumination, which would then partially dissociate the molecule, leading to a lowering of the barrier to the molecule passing through a hexagonal site on its diffusion path through the crystal and allowing diffusion away from the illuminated spot. This effect in turn suggests that conduction band electrons and valence band holes could also affect such partial dissociation, in accordance with the above calculations.

5.4 The effect of rotation on hydrogen molecules

The observation of a single stretch mode of the isolated molecule is in contrast with Raman-scattering studies of H_2 molecules, introduced from a plasma, in GaAs [140]. Here, *two* H-H stretch modes, separated by 8 cm^{-1}, attributed to para- and ortho-H_2 molecules occupying the lowest energy rotational states $J = 0$ and $J = 1$, respectively, were detected in the expected intensity ratio of 1:3. This ratio comes about since protons are spin half particles and therefore Fermionic, hence the overall wavefunction for the $^1H_2$ molecule must be antisymmetric on permutation of the nuclei. Since there are two possible relative orientations for the spins, i.e., ortho (o-) ($\uparrow\uparrow$) and para (p-) ($\uparrow\downarrow$), which are symmetric and antisymmetric respectively on permutation, then to give an overall antisymmetry the rotational part of the wavefunction must be antisymmetric for the o states and symmetric for the p states. At high temperatures when the molecule form, the equilibrium ratio of o-:p- for $^1H_2$ is 3:1[8], and although the energy difference between these states corresponds to $\sim$ 170 K, they are both occupied, even at cryogenic temperatures because of the absence of strong nuclear spin-flip processes during cool-down. Similarly, for $^2H_2$, the molecule is overall Bosonic, leading to a 1:2 p-:o- ratio. It seems likely that this is also true in Si. In this case, if the molecule was freely rotating, then two H_2 (or D_2) modes would be expected, whereas only one is detected. This implies, contrary to the case of GaAs, that molecular rotation is prevented by the crystalline field.

The effects of a crystal field on molecular rotation can be understood in terms of a model where rotating the molecule by θ from its equilibrium axis is accompanied by an increase in energy equal to $V(1 - cos\ 2\theta)\left(r_0/r\right)^2$ where r_0 is the molecular bond length. The molecular Schrödinger equation is then separable and the solutions in the variables θ and ϕ have definite parity. If $VI/\hbar^2$ is small, then the lowest energy states are the rotational levels $J = 0$ and $J = 1$. Here I is the moment of inertia of the molecule. On the other hand, if $VI/\hbar^2$ is large, then they are symmetric and asymmetric combinations of harmonic oscillator states bound in the potential wells around $\theta = 0$ and $\theta = \pi$. The symmetric and asymmetric states are labeled by "+" and "−" respectively, with the ortho H_2 molecules occupying the

[8] Due to the ratio of products for the spin and rotational degeneracies of o- and p-hydrogen

"−" state, in the case of D_2 this ordering is reversed. For the molecule in Si, $VI/\hbar^2$ is presumably large and the energy difference between the "±" states, $\mu_+ - \mu_-$, is very small. This difference leads to distinct centrifugal potentials,

$$\mu_{\pm}\left(\frac{r_0}{r}\right)^2,$$

in the two cases. For free rotation, $V = 0$, and $\mu_- - \mu_+ = \frac{1}{2}\hbar^2 J(J+1)/I$. The centrifugal potentials lead to a H-H stretch mode for the $J = 1$ state lying 5.9 cm^{-1} below that of $J = 0$ [155]. However, as V increases, the separation between the two stretch modes diminishes. If we assume that the splitting in ν_3 is, in fact, less than 0.1 cm^{-1}, and has not been so far resolved, then $(\mu_- - \mu_+)$ must be less than $0.02\hbar^2/I$ and this imposes a lower limit on the rotation barrier. Using the expressions for μ_+ and μ_- given by Stern [156], the barrier must be at least 0.17 eV. This means that the molecular rotation is frozen below about 70 K.

Thus at low temperatures, the molecules are aligned along one or more axes. The observation of only one mode in the H-D case means that either the hydrogen atoms are equivalent, and the center of the molecule lies on a mirror plane or a C_2 axis, or, if they are inequivalent, then the two H-D modes must be separated by at least 20 cm^{-1} and only one is thermally populated at 10 K.

5.5 Problems unresolved by *ab initio* calculation

Whilst the calculation for the frequency of the isolated interstitial molecule, and the molecule near O_{BC} agree generally (within $\sim$ 100 cm^{-1}) with the experimentally observed spectra of both Pritchard *et al.* and Leitch *et al.* (section 5.1.2), several features seen in the spectra are not adequately explained by these calculations.

The low rotational barrier obtained in these calculations and the work of others (section 5.1.1), suggest that the behaviour of the interstitial molecule in silicon should be akin to that of H_{2i} in GaAs, i.e., with a clearly resolvable ortho-/para-structure in the spectra of the molecular stretch mode. This contrasts strongly with the experimental work of Zhou and Stavola, who find a barrier of at least 60 K to molecular rotation, combined with an asymmetric structure. The unusual splitting of *only* the ν_2 line in deuterated Cz silicon is also unexplained.

Furthermore, the question of why an interstitial hydrogen molecule should be infrared active at all remains, since freely rotating H_2 and weakly hindered H_2 in GaAs are both considered infrared inactive.

The obvious suggestion that the molecule is not in fact isolated, but instead is bound to another defect in the lattice, leads to some fairly exotic properties for the required extra centre. This X defect would have to be able to diffuse at least as rapidly as the molecule, be invisible to both infrared and Raman spectroscopy, and be present in large concentrations. If the molecule were present as part of a group of $(H_2)_n$ molecules, then this agglomerate would have to be strongly bound to compete with the entropy gained by dispersing the molecules throughout the lattice. If we were to assume that there were about 10^{15} cm^{-3} pairs of molecules and only 10^{14} cm^{-3} single molecules in the material of Pritchard *et al.*, then at room temperature the binding energy would need to be approximately 0.5 eV [157]. Provisional calculations using the same cluster and basis as for the single molecule find that pairs on molecules sharing a common cage site are strongly unbound, and molecules in neighbouring T_d sites are degenerate with the separate molecules.

Clearly further experimental and theoretical study is needed to clarify the properties of this “apparently simple” system.

Chapter 6

Hydrogen inside lattice defects

Having considered the behaviour of hydrogen dimers at interstitial sites in the lattice, we now move on to examine hydrogen pairs within open and chemically inert or active voids in the lattice.

6.1 Interaction of H_2 with inert voids

The low frequency of the interstitial hydrogen molecule can be thought of as due to the leaking of charge from the surrounding bonds screening the intra-molecular proton-electron attraction [109]. If the size of the lattice cage were to be increased, for example by a defect in the crystal, then the charge density at the centre of the cage due to the surrounding Si would fall. The molecular stretch frequency would then rise towards that of isolated hydrogen molecules.

Several types of extended defects in silicon are known to exist, with associated vibrational bands around 2100 cm^{-1}, so the behaviour of H_2 trapped at such sites becomes of interest in understanding the effects of plasma treatment on silicon (see section 5.1.2).

6.1.0.1 Nano-voids

This dependence on surroundings is confirmed by a calculation in which the inner shells of Si atoms surrounding the T_d interstitial site are removed and the resulting dangling bonds hydrogenated. Removing just one shell and inserting passivating H atoms leads to a smaller volume for the cage but it is enlarged when ten Si atoms in two shells are removed. Relaxing the cluster in this case gives a H-H bond length of 0.759 Å which is shorter than the *interstitial* molecule. The stretch frequency (table 6.1.0.1) is then increased to 4324.6 cm^{-1} (H_2) and to 3057.9 cm^{-1} (D_2). These are within 170 cm^{-1} of the modes observed in plasma irradiated material by

Molecule in void	H-H	H-D	D-D
Calculated	4324.6	3745.9, 3743.8	3057.9
Experimental	4158	3629	2990
Isolated	4373.0	3788.9	3092.2

Table 6.1: Calculated and experimental frequencies, cm^{-1}, of H_2 molecules trapped in Si at small voids.

Raman scattering (section 5.1.2) and about 100 cm^{-1} below the calculated value of the molecule in vacuum. The distance of the centre of the cage to a Si (H) atom on the void surface is 4.72 (3.36) Å. Thus molecules within voids ~50% larger than that surrounding an interstitial site possess frequencies close to the gas value.

The void discussed here has a different structure from platelets, where molecules appear to push apart neighbouring (111) planes [151], and these must be considered to be distinct defects.

6.1.0.2 Platelets

While there is no general consensus regarding the microscopic structure of $\langle 111 \rangle$ platelets, if the assumption is made that the structure is more open than the undamaged lattice (as suggested by the possible NMR observation of D_2 inside platelets [158]) then some general remarks can be made from simulation of only one of the proposed platelet models.

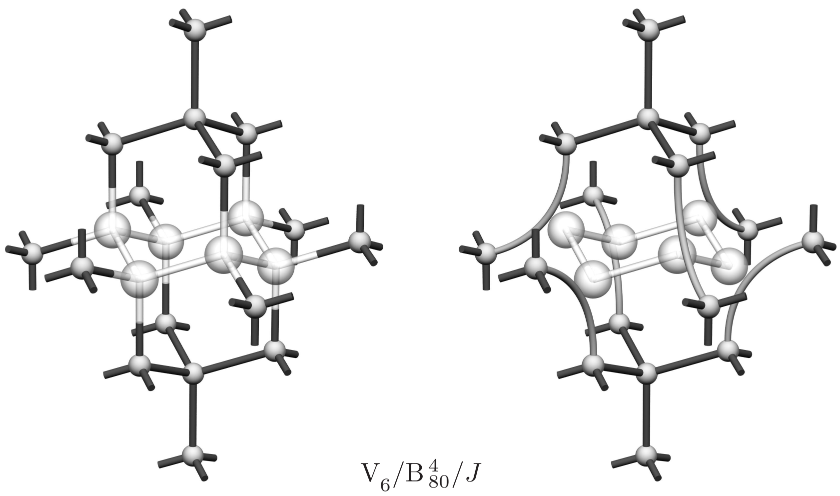
Of the models so far suggested for the $\langle 111 \rangle$ platelet, TEM studies [151] and calculations [159], favour the structure with a pair of hydrogen atoms saturating each broken bond between (111) planes displaced apart by ~ 3 Å. This structure does not include any silicon vacancies. If this defect is constructed inside a $Si_{84}H_{64}$ cluster, by cutting the cluster through the centre and hydrogenating the two exposed faces, the vibrational modes of a H_2 molecule trapped within such a platelet can be calculated. The stretch frequencies for the molecule lie at 4385 cm^{-1} for H_2, 3796 and 3804 for HD, and 3101 cm^{-1} for D_2. The modes of the hydrogen saturating the surrounding silicon atoms fall between 2117 and 2098 cm^{-1}. The experimental values for platelets [160] lie at ~ 2110 and ~ 1960 cm^{-1}. The molecule has a bond length of 0.748 Å when sited in the platelet, and is centred 2.459 (3.255) Å from the nearest passivating hydrogen (structural silicon) atoms.

Taken together, these results suggest that the 4158 cm^{-1} band might be due to molecules within voids or platelets and this is consistent with a correlation that has been found between the Raman intensities due to Si-H stretch modes arising from H atoms on the surface of a void, and the 4158 cm^{-1} band (see the work of Leitch *et al.* in section 5.1.2). The calculations also suggests that the higher frequencies found by *ab initio* calculations on molecules in small clusters (see section 5.1.1) may be due to an underestimate in the charge density at the interstitial site arising from the use of a limited basis set.

6.2 Interaction of hydrogen with V_6

Electron paramagnetic investigation of multi-vacancy centres in Si have successful identified V_1 [161], V_2 [162], V_3, V_4 and V_5 [163]. The last has been correlated with the $P1$ centre and is a non-planar defect with C_{1h} symmetry. The larger vacancy centres are formed in irradiated material when subjected to a heat treatment. Thus V_5 is formed around 170 °C and is stable until ∼450 °C.

Theoretical investigations have predicted that the ring hexavacancy (V_6) is a particularly stable defect [164, 165]. There are several metastable configurations of V_6, but they are all at least 0.8 eV higher in energy than the ring structure [165, 166]. Removing six atoms of a puckered hexagon in Si (figure 6.1) leaves 12 broken bonds. However these can be rejoined in pairs, suggesting that the reconstruction can eliminate, or reduce, the electrical activity of the centre. One would suppose that this defect would be formed after V_5 between 170 °C and 450 °C. It is tempting to argue that, among the various multi-vacancy defects, V_6 should be exceptionally numerous. Even if it is not electrically active, it should provide a trap for interstitial impurities such as transition metals [167], oxygen and hydrogen. Then evidence of a V_6 core could be found from an analysis of the spectra of impurity related defects. It is argued here that the 1.107 eV photoluminescent (PL) centre called B^4_{80} [168] or the J-line centre [72], arises from a reconstructed V_6 defect, and that a large family of related hydrogen-containing centres are due to V_6 complexed with hydrogen.



$V_6/B^4_{80}/J$

Figure 6.1: V_6 displaying the six vacancy sites, and the subsequent reconstruction caused by them.

The evidence for this assignment comes from a study of the J-line defect and a family of hydrogen related PL centres with similar optical characteristics. The J-line defect is one of the most dominant luminescence systems formed in irradiated Si after a heat treatment between 300 and 500 °C [168, 72] (see figure 6.2). However, a number of other shallow luminescence centres have been observed in Si grown in a hydrogen atmosphere after thermal neutron irradiation and subsequent heat treatment at 400-480 °C [168] which grow at the expense of the J-line defect, as shown in figure 6.2. These radiation-related defects have labels of the form B^x_{yz}, which specifies the exciton binding energy relative to the free exciton as $xy.z$ meV. So for example, B^1_{71} has an exciton binding energy of 17.1 meV, and a main PL line at 1.138 eV.

Isotopic analysis of Si samples saturated with hydrogen-deuterium mixtures have unambiguously established the direct involvement of hydrogen in the defects and have shown that they contain two or more H atoms [169, 170]. High resolution spectroscopic studies involving uniaxial stress and magnetic field measurements reveal that each of the centres has a complex structure of bound exciton states arising from the interaction between a bound electron and a hole in a defect field of particular symmetry [171, 172]. A detailed quantitative analysis of the optical transitions under external field perturbations has shown that, despite the apparent

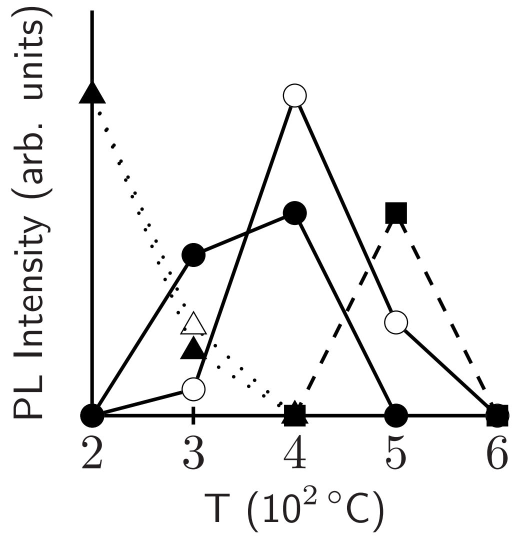


Figure 6.2: Thermal stability of the B_{80}^4 (J) centre (circles), the B_{71}^1 centres (squares), and the W centre (triangles), after Safonov [67]. The behaviour of the centres is shown in non-hydrogenated (open symbols) and hydrogenated (filled symbols) silicon. The samples used were high resistivity ($\sim$ 1000 Ωcm) fz Si with [C] $\sim 10^{17}$ cm^{-3} and [O]$< 10^{16}$ cm^{-3}.

differences between these optical systems and their multiple transition structures, all of them have a similar origin. In each case the luminescence activity arises from a single $(-/0)$ level in the vicinity of E_c [173, 170].

This suggests that these multi-hydrogen defects possess a common core. The symmetry of this core must allow the possible symmetries of the different hydrogen related centres and, in particular, B_{41} and B_{71}^1 which contain *equivalent* and *inequivalent* pairs of hydrogen atoms respectively arranged in trigonal symmetry. Other centres, such as B_{81}^1 and B_{18}^1 also contain two H atoms but have C_{1h} symmetry, while others like B_{80} and B_{19}^1 contain more H atoms without any symmetry [169, 170, 172].

The high symmetry of B_{41} and B_{71}^1 strongly restricts possible structures of the defects, as the two equivalent hydrogen atoms in B_{41} must lie along a C_3 $\langle 111\rangle$ axis and requires an additional C_2, S_2 or σ_h symmetry element. There are only two sites in the diamond lattice at which the centre of this defect can lie. These are the bond-centre and hexagonal lattice sites, which yield D_{3d} symmetry. The obvious structure, consisting of two anti-bonded hydrogen atoms attached to the pair of silicon atoms surrounding a bond centred site must be discounted, since such a defect should be stable only at low temperatures (see chapter 7).

The combined requirements of creation by irradiation and high thermal stability point towards a multi-vacancy–hydrogen complex. $V_1 - V_5$ can be excluded either because they are not centred at these sites or do not possess trigonal symmetry. V_6 has D_{3d} symmetry, and thus the simplest candidate is a complex involving V_6 and hydrogen. As stated above, V_6 is expected to form in irradiated material which is heated between ∼200 °C and 450 °C. If the material contains hydrogen molecules these will readily react with V_6 to form dimer structures.

The defects considered were each constructed in a $Si_{154}H_{108}$ cluster centred on the hexagonal site. The size of the cluster meant that there is at least two layers of Si–Si bonds separating the terminating H atoms and the H atoms in the defect.

The wave-function basis consisted of independent s and p Gaussian orbitals, with either four different exponents sited at each Si atom, or three at the H atoms of the dimers. A fixed linear combination of two Gaussian orbitals was sited on the H atoms which terminated the cluster. In addition, two Gaussian functions were placed across every Si–Si bond and the Si–H bonds the defects. The charge density was fitted with five independent Gaussian functions with different widths on each Si atom, and four (three) on the central (terminating) H atoms. Two extra Gaussian functions were placed at each bond centre. All atoms, including the terminators, were allowed to relax by a conjugate gradient method.

V_6 was first simulated by removing the six atoms nearest to the hexagonal site. The resulting structure possesses D_{3d} symmetry and strongly reconstructs on relaxation to form six new bonds of length 2.63 Å between the twelve dangling bonds, as shown in figure 6.1. The length of the two nearest Si–Si bonds along the trigonal axis is then 2.53 Å, demonstrating that tensile strain resides in this direction. The resulting electronic Kohn-Sham (KS) levels display a well defined band-gap, which is consistent with previous calculations [166]. This defect possibly possesses states in the gap very near the conduction band, but it is difficult to decide whether such near-conduction levels are localised on the defect in cluster calculations. The character of the deepest of these shallow states is a_{1g} (figure 6.3), this state is anti-bonding to both the six reconstructed bonds and the two nearby axial bonds in the defect.

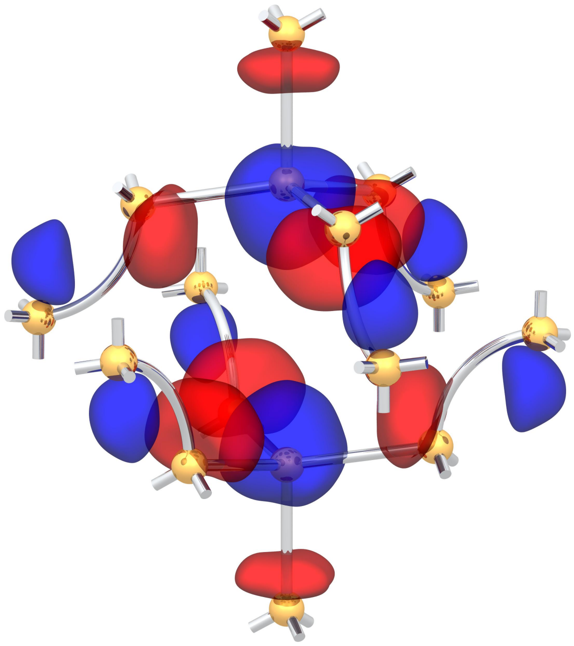


Figure 6.3: Isosurface plot of the a_{1g} pseudowavefunction corresponding to the lowest unoccupied state of V_6, as shown in figure 6.1. The red and blue lobes are at a phase difference of π. The function shows anti-bonding character between both the atoms in the six central reconstructed bonds, and also the two nearby axial Si –Si bonds.

There are several potential structures for B_{41} based on V_6 and one possible model would be a hydrogen molecule aligned along [111] and sited at the centre of inversion of the defect, or alternatively the H–H bond could be broken and the atoms placed at anti-bonding sites to the axial Si–Si bonds. Both structures are shown in figure 6.4.

The second, anti-bonded structure is unstable and spontaneously relaxes back to the [111] aligned molecule structure, however the molecule itself is only metastable at the defect's centre of inversion, and will dissociate with a barrier of less than 0.23 eV (the barrier to this reaction is shown in figure 6.5). It does this by breaking one of the reconstructed Si–Si bonds shown in figure 6.1, to form two Si–H bonds, giving rise to a structure which is 1.76 eV lower in energy and of C_{1h} symmetry, as shown in figure 6.6.

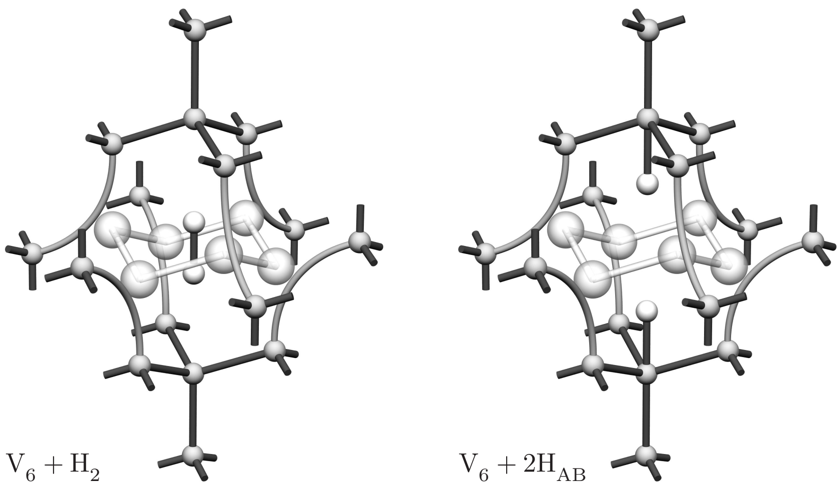


Figure 6.4: Reconstructed V_6 plus either a hydrogen molecule or two anti-bonding hydrogen atoms.

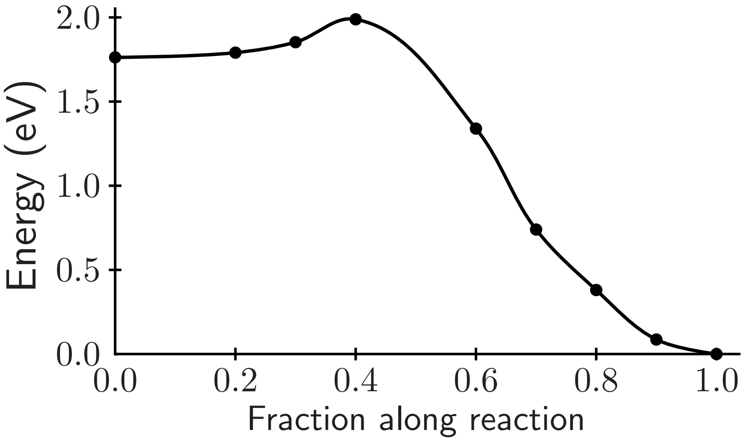


Figure 6.5: Barrier to dissociation of a H_2 molecule inside V_6, as shown in figure 6.4. The molecule dissociates by breaking one of the reconstructed bonds to form the structures shown in figure 6.6.

A third possibility is that the hydrogen atoms are near the bond-centred sites within the axial Si–Si bonds. This third, bond-centred structure, causes a further reconstruction of V_6, by breaking the two axial Si–Si bonds and forming Si–H bonds. The two dangling Si bonds thus formed pair together in the middle of the defect, leaving the structure with D_{3d} symmetry shown in figure 6.7. This new bond is then 2.42 Å, and the six reconstructed bonds relax to a length of 2.58 Å – a lowering of

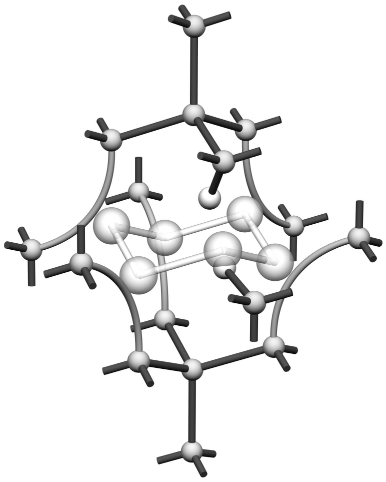


Figure 6.6: The result of dissociating a H_2 molecule inside V_6, leading to the breaking of one of the reconstructed Si–Si bonds to form two Si–H bonds.

0.02 Å. The reconstruction is energetically very favourable, being 0.80 eV lower in energy than the C_{1h} structure in figure 6.6. Thus the most stable structure has D_{3d} symmetry.

The KS levels of this defect again show evidence for a very shallow $(-/0)$ level, of symmetry a_{1g} (see figure 6.8. There are two high frequency modes related to the hydrogen in this structure of symmetries A_{1g} and A_{2u} (which is infrared active), which lie at 2033 and 2021 cm^{-1} respectively. It is clear that this structure possesses properties entirely consistent with B_{41}, although electrical field and stress alignment studies are necessary to verify this assignment.

Potential structures for B^1_{71} are now considered. Again this structure requires a C_3 axis, but due to the inequivalent hydrogen atoms the symmetry of the defect must be of lower order, i.e. C_{3v}. If it is assumed that dangling silicon bonds are energetically infeasible, there is an obvious structure which can be derived from the bond-centred model for B_{41}. This defect consists of one hydrogen atom lying at a bond centred site, and the other atom directly passivating the silicon atom which is displaced by the bond-centred H. There are two locations on either side of the silicon atom at which the hydrogen might be positioned. If the hydrogen were placed on the side nearest the other H atom, this would bring the two atoms into close proximity and thus raise the energy of the defect, this suggest that the

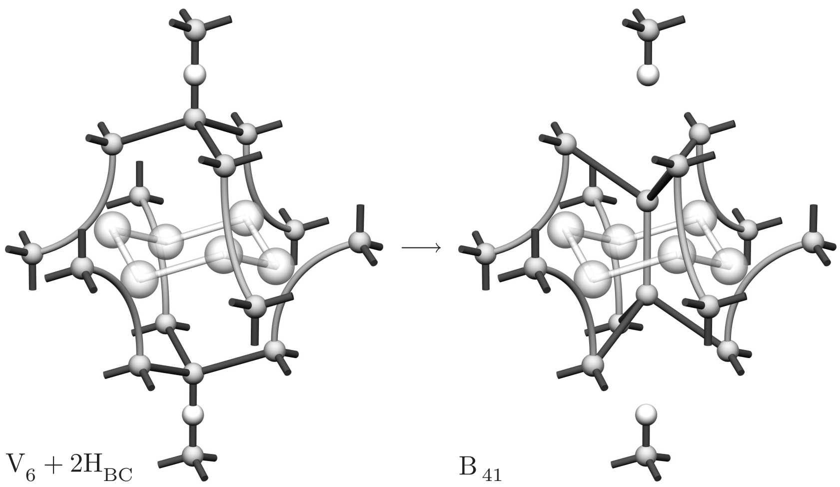


Figure 6.7: The reconstruction of V_6 when complexed with two bond-centred hydrogen atoms, giving rise to the B_{41} structure.

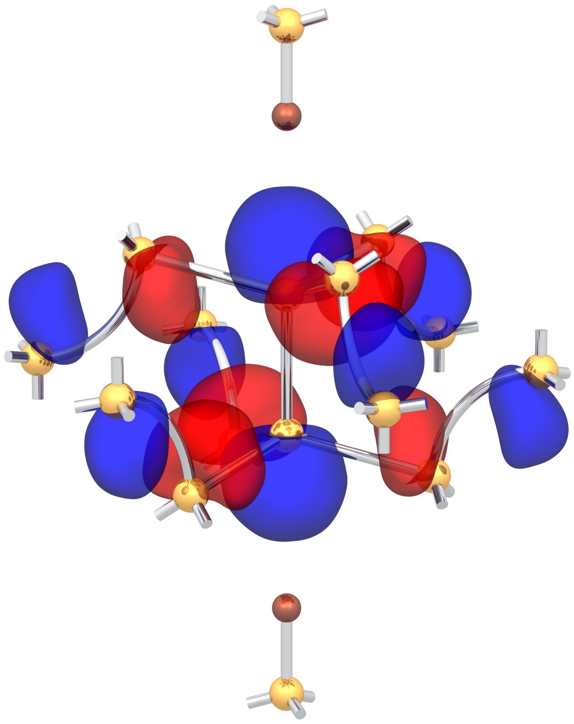


Figure 6.8: Isosurface plot of the a_{1g} pseudowavefunction corresponding to the lowest unoccupied state of the B_{41} structure, as shown in figure 6.7. As with V_6, the wavefunction shows anti-bonding character between both the atoms in the six central reconstructed bonds.

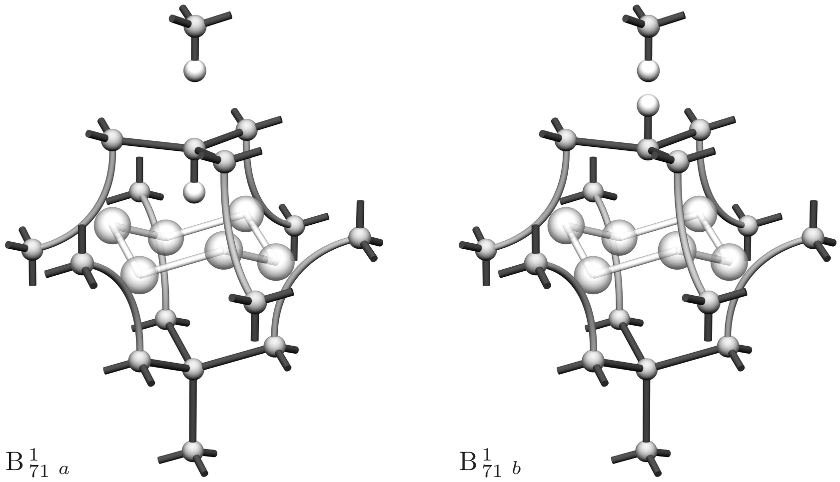


Figure 6.9: Schematic diagrams of the B^1_{71} structure, with the two inequivalent hydrogen atoms arranged as in H_2^*, or together in the same bond, similar to the dimer proposed by Aradi and Deák (see chapter 7).

hydrogen is sited on the opposite side of the silicon atom, as shown in figure 6.9a. The relative energies of these defects are 0.13 eV in favour of the structure with the hydrogen atoms on opposite sides of the silicon. The more stable structure, which is reminiscent of H_2^* (see chapter 7), is 0.09 eV higher in energy than the C_{1h} defect formed by the dissociation a hydrogen molecule inside V_6. The KS levels again show evidence of a near conduction-band a_1 state localised on the defect in both cases. Both of these models possess two high frequency A_1 vibrational modes, lying at 2149 and 2029 cm^{-1} for the structure with the two H atoms lying close together, or 2051 and 2010 cm^{-1} for the hydrogen on opposite sides of the silicon.

It is also clear that other, stable centres containing more H atoms could arise by inserting H pairs into the reconstructed bonds, as well as along the trigonal axis.

These investigations then favour V_6 to be a precursor for the family of H-related optical centres. It is clear that an unreconstructed V_6 defect has deep levels arising from the dangling bonds but these are pushed to the band edges on reconstruction. A greater degree of reconstruction results in increasingly shallow levels. As the effect of hydrogen is to reduce the strain in the defect, the expectation is that the acceptor $(-/0)$ level of V_6 will be deeper than that of V_6H_2. The KS levels also suggest that V_6 itself should be optically active.

In non-hydrogenated Si, the B_{80}^{4} centre (J-line) is observed after irradiation and subsequent heat treatment between 300 – 500 °C [72] as shown in figure 6.2. Zeeman splitting studies have also shown that B_{80}^{4} possesses trigonal symmetry [72]. However, in hydrogenated Si, the J-line intensity declines at a much lower temperature ∼450 °C, which coincides with the appearance of the hydrogenated related B-centres. figure 6.2 shows the variation in the J-line intensity and the intensity of B_{71}^{1} in non-hydrogenated and hydrogenated Si respectively, after a 30 minute isochronal anneal at different temperatures [173]. As one can see, there is an anti-correlation between B_{71}^{1} and the B_{80}^{4} centres with B_{71}^{1} created at temperatures when the B_{80}^{4} centres are disappearing. The figure also shows that B_{71}^{1} is not related to the 1.018 eV W (I_1) centre, which is also known to be trigonal [72, 75].

The identification of B_{80}^{4} with a vacancy aggregate is consistent with the strong preference for the defect to align under [111] stress during formation [174]. This is directly analogous with V_2, where stress alignment studies show that $V_2^{\pm}$ favours an orientation where the reconstructed bonds are parallel to the stress axis [162]. Uniaxial stress along the trigonal axis leads to a reduction in the exciton binding energy consistent with the idea that anti-bonding orbitals on the dilated reconstructed bonds are the origin of the gap acceptor level. Recent Zeeman and uniaxial studies demonstrate that the exciton consists of a electron in a deep $(-/0)$ level near E_c with a loosely bound hole in a similar way to B_{41} [67]. Thus this defect shares the properties expected for V_6 and can be identified with this centre. However, in the absence of experiments performed in an electric field, the question of whether B_{80}^{4} has a centre of inversion has yet to be resolved.

The dependence of the intensity of the PL due to B_{80}^{4} demonstrates that the exciton is thermally bound with an energy of 18 meV [168] and this can be taken to be the ionization energy of the hole. The exciton binding energy relative to a free electron and hole is the sum of 48.0 meV and the free exciton binding energy of 14.3 meV. From these results, the $(-/0)$ level of V_6 can be placed at $E_c - 44.3$ meV.

Chapter 7

A new hydrogen dimer in proton implanted silicon

7.1 Introduction

In addition to the hydrogen molecule already discussed (see chapter 5), a second type of hydrogen dimer is know to exist in silicon.

H_2^*, which is a pair of atoms bonded within the lattice, was first predicted to exist in diamond [175] and silicon [176] by theory, and subsequently identified experimentally [177] by infrared absorption in proton implanted silicon. Neutron and electron irradiation of silicon containing hydrogen molecules also forms H_2^* [146] at room temperature, by converting the H_2 molecule into this defect [178], with the two dimers then coexisting in the material. This defect is also directly observed in small quantities after high temperature in-diffusion of H_2 gas [179]. H_2^* consists of a nearly bond centered hydrogen atom, H_{BC}, which is displaced towards one of the two neighboring lattice atoms, with a nearby anti-bonded hydrogen, H_{AB}, further along the same $\langle 111 \rangle$ direction, and attached to the other side of the second lattice atom (see figure 7.1c for a schematic of the defect). The resulting structure is then trigonal, with C_{3v} symmetry.

It has been suggested [146] that the formation of H_2^* in irradiated silicon containing H_2 molecules proceeds by one of two mechanisms. One possibility is that a hydrogen molecule reacts with a self interstitial to form IH_2, which then captures a lattice vacancy to form H_2^*. Alternatively the molecule reacts, instead, with the vacancy to form VH_2, which then captures an interstitial to again make H_2^*. In either case, it is possible that the precursor (IH_2 or VH_2) should be produced in detectable quantities *before* H_2^*.

For proton implanted silicon, a modified form of this reaction must occur, since

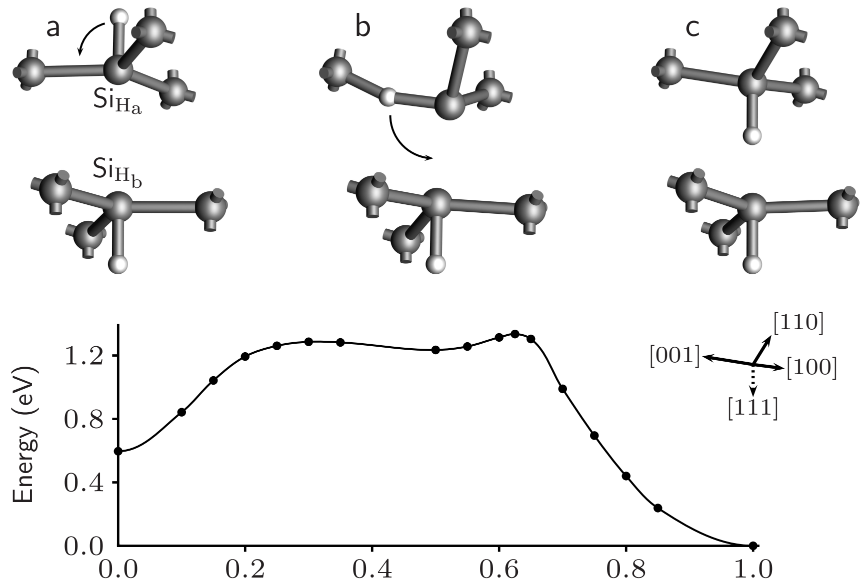


Figure 7.1: Schematic diagrams of a) H_2^{**}, b) the local minimum structure consisting of an anti-bonding hydrogen atom and a next-nearest neighboring bond-centered hydrogen atom, and c) H_2^*. The barrier to inter-conversion between H_2^{**} and H_2^*, passing through the weak local minima at the bond-center site, is shown. The barrier to the conversion from H_2^{**} to H_2^* via this path is calculated to be $\leqslant 0.7$ eV.

the majority of the hydrogen in such material at low temperatures is in the form of bond-centered H^+ [180], hence a mechanism to convert H^+ into H^0 or H^- must be present to form neutral dimers. If we assume that a large number of vacancy-interstitial pairs are created during the implantation, then these will act as trapping sites for the hydrogen in the lattice. It has been suggested that the single vacancy [161] is present in the V^{2-} state in n-type silicon [181], and can diffuse to H^+_{BC} to form immobile VH^- at about 80 K [161, 180]. As the material is heated further, H^+_{BC} becomes mobile at ~ 200 K [180], and this will then react with VH^- to form VH_2^0. Presumably VH_2^0 reacts with a mobile interstitial to form H_2^*, or other dimer structures such as the molecule. Little is known about the charge-state of the single interstitial in implanted silicon, but the absence of large quantities of IH_1 in proton-implanted silicon suggests that a similar reaction between I^{x-} and H^+ does not occur at low temperatures. IH_2 has been observed in implanted silicon annealed at room temperature [182], and may then be formed either by the reaction of a hydrogen molecule with an interstitial [183], or directly with the neutral or negative hydrogen atoms present at higher temperatures [180, 184].

Both VH_2 and IH_2 have been observed in proton implanted silicon annealed at room temperature [182, 185], and IH_2 has been observed in irradiated [186–188] silicon containing hydrogen. Both defects are of comparable thermal stability to H_2^* [182, 185], but, perplexingly, H_2^* has been observed to be present in the material *before* VH_2 has formed [189].

Estreicher *et al.* have performed molecular dynamics (MD) simulations of the interaction between interstitial H_2 and native defects [183], and find the molecule is easily dissociated by both interstitial and vacancy related radiation damage. In addition, the same authors have simulated the reaction between IH_2 and a lattice vacancy [190], and find that H_2^* is indeed formed in this process. The alternative reaction between VH_2 and an interstitial was not observed during the course of the MD simulations.

Vibrational modes lying at 1792, 812 and 1607 cm^{-1} have recently been observed, consistent with stretch, bend and overtone modes of a new defect [189]. The stretch and bend modes lie at 1307 and 581 cm^{-1} in deuteron implanted material. Mixed H/D implantation shows no new modes. These stretch and bend modes are similar to those due to the anti-bonding hydrogen atom in H_2^*, with the presence of a single bend mode suggesting that the system possesses high trigonal symmetry.

It is now suggested that these modes are consistent with the properties of a third type of hydrogen dimer present in proton-implanted silicon, consisting of two anti-bonding hydrogen atoms attached to opposite ends of a Si–Si pair, with overall D_{3d} symmetry (see figure 7.1a).

7.2 Cluster and supercell calculations

The theoretical studies in this work use local density-functional formalism as implemented in the AIMPRO code, both as clusters [21] and periodic supercells [52].

For the former calculations, the dimers were each inserted into a $Si_{98}H_{78}$ cluster centered on a Si–Si bond-center. The wave-function basis consisted of independent

s and p Gaussian orbitals, with four different exponents sited at each Si site, and three at each H atom of the dimer. A fixed linear combination of two Gaussian type orbitals was sited on the terminating H atoms. In addition, two Gaussian functions were placed at the center of each Si–H bond of the dimers and every Si–Si bond center. The charge density was fitted with five independent Gaussian functions with different widths on each Si atom, and four (three) on the central (terminating) H atoms, with two extra Gaussian functions placed at each bond center. All atoms, except the terminating H ones, were allowed to relax by a conjugate gradient method. Further details of the method can be found in Ref. [21].

The supercell calculations made use of a Si_{54} FCC cell, with the Monkhorst–Pack k-point sampling scheme [191]. The real-space basis used consisted of independent s- and p-Gaussian orbitals, with four different exponents, sited at each Si site and three at each H atom of the dimer. In addition, a single Gaussian function was placed at each Si–H bond of the dimer and every Si–Si bond center. The Hartree and Perdew–Wang [39] exchange-correlation energies were calculated in reciprocal space with an energy cut-off of 40 atomic units. All atoms were allowed to relax by a conjugate gradient method.

For both types of calculation the normal modes of the system were examined using the quasi-harmonic second derivatives of the energy, which were evaluated for the defects using the method described in Ref. [21]. The second derivatives were calculated directly for the H atoms, neighboring silicon atoms, and their neighbors, with the rest of the dynamical matrix filled using a previously evaluated Musgrave-Pople potential. The transition dipole moment of the defect was calculated directly in the cluster method by finding the change in the total dipole moment of the cluster as the atoms were displaced according to the normal coordinates of each of the modes.

The barrier to conversion between H_2^{**} and H_2^* was evaluated by performing a full structural relaxation, where the position of *one* of the hydrogen atoms in the dimer was forced to lie on a constraint surface of the form $r_a^2 - r_b^2 = c$, where r_a and r_b were the separation of the constrained H atom from two points on opposite sides of the silicon atom marked Si_{H_a} in figure 7.1a. The value of the variable c was adjusted after each constrained relaxation to force the hydrogen atom to move from

	$Si_{98}H_{78}$	$Si_{148}H_{98}^{\S}$	$Si_{98}H_{78}^{\dagger}$	54 FCC	128 FCC
$Si_H \cdots Si_H$ (Å)	3.74	3.84	3.78	3.53	3.62
Si_H–H (Å)	1.50	1.51	1.50	1.54	1.54
Si_H–Si (Å)	2.36	2.35	2.22	2.32	2.32
$H\hat{Si_H}Si$ (°)	86.4	87.3	88.2	86.2	86.6
$Si\hat{Si_H}Si$ (°)	119.6	119.8	119.9	119.6	119.7

Table 7.1: Effect of increasing cluster size (§), basis (†), or supercell size (the structure is insensitive to the number of k-points chosen between 1^3 and 4^3) on the bond lengths and angles of the H_2^{**} defect. The silicon atoms marked Si_H are bonded directly to the hydrogen in the structure shown in figure 7.1a.

its position in the H_2^{**} dimer to that in H_2^*. This constraint surface is equivalent to fixing the H atom to a plane perpendicular to the axis between the two constraint points, with the position of the plane along this axis determined by the constant c, as shown by the following analysis : Since the constraint is cylindrically symmetric the discussion can be restricted to the two dimensional case, and by applying a suitable translation, rotation and scaling the two constraining points can be transformed to lie at $[0, 0]$ and $[1, 0]$. This then leads to the constraint surface, (x_i, x_j), being of the form

$$\begin{aligned} c &= \begin{pmatrix} x_i & - & 0 \\ x_j & - & 0 \end{pmatrix}^2 - \begin{pmatrix} x_i & - & 1 \\ x_j & - & 0 \end{pmatrix}^2 \\ &= 2x_i - 1 \,, \end{aligned}$$

i.e., there is only a constraint on the x_i component of the surface.

To test the convergence of the molecular properties in the cluster calculations, the local modes of H_2^{**} were re-calculated in a larger $Si_{148}H_{98}$ (plus defect) cluster, also centered on a Si–Si bond-center. All of the atoms, except the surface terminating hydrogen were allowed to relax. Similarly, to test for effects of increasing the basis, the calculations were repeated in the $Si_{98}H_{78}$ (plus defect) atom cluster with a larger basis, using eight different exponents sited at each Si atom for both the wave-function and charge density, the hydrogen atoms in the defect were fitted using six different exponents of both types. In addition, a third Gaussian function

Mode	$Si_{98}H_{78}$	$Si_{148}H^{\S}_{98}$	$Si_{98}H^{\dagger}_{78}$
A_{1g}	2019.8	2065.0	2094.3
A_{2u}	2015.7	2061.8	2070.4
E_u	809.7	843.7	783.0
E_g	806.8	838.7	780.1

Table 7.2: Effect of increasing cluster size (§), or basis (†) on the local vibrational modes (cm^{-1}) of H_2^{**}, as described in the text.

was placed between every bonded pair of atoms in both types of basis. The effects on the molecular vibrational frequency of these changes are shown in table 7.2, the frequency is seen to be converged to within $\sim$ 100 cm^{-1} with respect to both cluster size and basis. table 7.1, shows the convergence of the structural properties for the defect, the bond lengths and angles change by $\leqslant$ 0.15 Å or 2° respectively, within the cluster calculations.

Convergence in the supercell calculations was tested by increasing the number of special k points used to sample the Brillouin zone from a single point at Γ up to the MP-4^3 sampling (i.e., 10 irreducible k points with D_{3d} symmetry), and similarly increasing the size of the supercell from 54 to 128 lattice sites, as shown in table 7.3. The vibrational and structural properties are seen to be well converged with respect to supercell size and k-point sampling, but the relative energies of H_2^* and H_2^{**} are only convergent for more than one k-point. The structure is insensitive to change in the number of k-points, and changes by $\leqslant$ 0.1 Å or 0.5° on increasing from 54 to 128 atom (plus defect) FCC supercells.

The question of the disparity between the supercell and cluster frequencies then arises. In principle, the intermediate charge-density fitting (see section 3.2) used in the supercell method should be more accurate than that of the cluster method, as for a sufficiently large cut-off the planewave fit should converge to the exact Hartree and exchange-correlation energies for the given wavefunctions. As an empirical rule of thumb, a 1.3% error the in calculation of bond lengths leads to a ~4% error in the frequency of the derived stretch-modes [192]. By comparing the bond lengths for the Si–H bonds between the cluster and supercell, the cluster results are seen

	MP-x^3 k-points		
size	1	2	4
54	0.06	0.82	0.81
128	0.43	0.75	–

MP-x^3	A_{1g}	A_{2u}	E_g	E_u
1	1870.2	1850.1	741.2	733.9
2	1880.0	1874.2	762.5	758.7

Table 7.3: Relative energy (eV) of H_2^{**} compared to H_2^* for various supercell sizes and k-point sets, and the effect of increasing the number of k-points on the vibrational modes of the 54 atom supercell containing H_2^{**}. The relative energy of the two defects is ~0.75 eV in favor of H_2^* (*c.f.* 0.60 eV from the $Si_{98}H_{78}$ cluster calculations).

Mode	cluster	supercell	Expt.
A_1	1983.0	2097.5	2061.5
A_1	1894.0	1717.0	1838.3
E	837.7	717.7	817.2

Table 7.4: Modes of H_2^* (cm^{-1}) calculated in this work using a $Si_{98}H_{78}$ cluster or a 54 atom FCC supercell with MP-2^3 k-points, compared with the experimental observations of Holbech *et al.* [177].

to be 2.7% shorter, translating to an 8.3% increase in the hydrogen-related stretch frequency. Scaling by this amount gives the A_{1g} and A_{2u} stretch modes as 1865 and 1861 cm^{-1}, i.e., comparable to the supercell results.

7.3 Results

As a preliminary to examining H_2^{**}, we first calculate the vibrational modes of H_{BC} and H_2^* using both formalisms. The derived frequencies of H_2^* are shown in table 7.4, and both types of calculation agree within $\sim$ 120 cm^{-1} with the experimental data of Holbech *et al.* [177], with the E bend mode of the cluster calculation falling within $\lesssim$ 20 cm^{-1} of experiment. The vibrational modes of H_{BC}^+ are shown in table 7.5, for the cluster calculations the A_{2u} mode of the proton falls within $\sim$ 100 cm^{-1} of the observed vibrational mode at 1998 cm^{-1} [180], and displays a shift of less than 1 cm^{-1} with variation in the mass of the neighboring silicon atoms, consistent with the lack of experimental resolution of such lines. Following the expression for the integrated intensity of the vibrational modes by Leigh and Szigeti [62], and the discussion of Clerjaud and Côte [63], if it is assumed that the mass of this A_{2u} mode is $\sim$ 1 a.m.u., the effective charge of this mode is 0.618 e, i.e., a fifth of that observed [180]. The vibrational modes of H^0_{BC} are shown in table 7.6. The proton and also the muon pseudo-isotope display an additional E_u mode above the Raman edge absent for the other isotopes. The energy of the A transitions is sufficiently large to ionize the $E_c - 0.16$ eV level of $^1H^0$ [193], effectively preventing the observation by IR, but the lower energy E_u mode of $^1H^0$ may be observable. The effective charges for the $^1H^0$ modes, again assuming a mass of $\sim$ 1 a.m.u., are 0.187 and 0.183 e for the E_u and A_{2u} modes respectively.

We now turn our attention to H_2^{**}. The optimized structure of H_2^{**} is shown in figure 7.1a, with the structural parameters given in table 7.1. It is worth noting the effect of the *two* anti-bonding hydrogen atoms on opposite ends of the same Si–Si pair, which dilates the distance between these two silicon atoms to $\sim$ 3.6 Å, effectively breaking this bond (this compares to 2.35 Å for bulk silicon), and leaving the two Si atoms almost co-planar to their silicon neighbors.

table 7.7 gives the calculated vibrational modes in comparison to experiment. Again the IR active A_{2u} stretch mode calculated in the supercell method, or the scaled cluster calculations, is within < 100 cm^{-1} of the experimental value, with the cluster calculation for the observable E_u bend mode again falling within < 20 cm^{-1} of experiment. The overtone, marked A^* in table 7.7, must be of symmetry A_{2u} and consists of a mixture of the E_u and E_g modes which lie at around 800 cm^{-1}. As shown in the mixed isotopic case, the motion of the two hydrogen atoms is practically decoupled, since the A_1 and E modes of the HD** complex are within $\sim$ 3 cm^{-1} of the A_{2u} and E_u modes of the isotopically pure H_2^{**} and D_2^{**} defects. We calculate the ratio of the intensities of the fundamental A_{2u} and E_u IR active modes to be 0.29. Assuming the mass of the oscillator is $\sim$ 1 a.m.u., the effective charges of these two modes is then 0.13 and 0.25 e respectively.

The cluster and supercell calculations both find H_2^* to be between 0.6 and 0.75 eV lower in energy than H_2^{**} (table 7.3). This energy for H_2^{**} is still $\sim$1.4 eV more stable than two isolated H^0_{BC} atoms [194]. Although this defect is less stable than H_2^*, it

a	b	μ^+		H$^+$	D$^+$	T$^+$
Cluster						
28	28	A_{2u} 5646.2	E_u 943.0	A_{2u} 1903.5	A_{2u} 1353.1	A_{2u} 1110.7
28	29	A_1 5646.2	E 943.0	A_1 1903.3	A_1 1352.8	A_1 1110.4
28	30	A_1 5646.1	E 943.0	A_1 1903.1	A_1 1352.6	A_1 1110.1
29	29	A_{2u} 5646.1	E_u 943.0	A_{2u} 1903.1	A_{2u} 1352.6	A_{2u} 1110.1
29	30	A_1 5646.1	E 943.0	A_1 1903.0	A_1 1352.4	A_1 1109.8
30	30	A_{2u} 5646.0	E_u 943.0	A_{2u} 1902.8	A_{2u} 1352.1	A_{2u} 1109.5
Supercell						
28	28	A_{2u} 6348.7	E_u 557.7	A_{2u} 2141.4	A_{2u} 1523.1	A_{2u} 1250.9
28	29	A_1 6348.6	E 557.1	A_1 2141.2	A_1 1522.8	A_1 1250.6
28	30	A_1 6348.5	E 556.7	A_1 2141.0	A_1 1522.5	A_1 1250.2
29	29	A_{2u} 6348.5	E_u 556.9	A_{2u} 2141.0	A_{2u} 1522.5	A_{2u} 1250.2
29	30	A_1 6348.4	E 556.8	A_1 2140.8	A_1 1522.2	A_1 1249.8
30	30	A_{2u} 6348.4	E_u 556.4	A_{2u} 2140.6	A_{2u} 1521.9	A_{2u} 1249.5

Table 7.5: Vibrational modes, cm^{-1}, of several different isotopic combinations of ^{a}Si–H$^+$–^{b}Si in silicon, calculated in the $Si_{98}H_{78}$ cluster or 64 atom simple-cubic supercell. The muon pseudo-isotope displays an additional E mode above the Raman edge absent for the other isotopes. The hydrogen mode calculated by the cluster code to lie at 1903.5 cm^{-1} is within $\sim$ 100 cm^{-1} of the observed vibrational mode at 1998 cm^{-1} [180], and displays a shift of less than 1 cm^{-1} with variation in the mass of the neighboring silicon atoms, consistent with the lack of experimental resolution of such lines.

a	b	mu^0		H^0		D^0	T^0
28	28	A_{2u} 4128.4	E_u 1765.9	A_{2u} 1389.7	E_u 604.2	A_{2u} 986.3	A_{2u} 808.3
28	29	A_1 4128.3	E 1765.8	A_1 1389.6	E 604.0	A_1 986.1	A_1 808.1
28	30	A_1 4128.3	E 1765.8	A_1 1389.6	E 603.9	A_1 986.0	A_1 808.0
29	29	A_{2u} 4128.3	E_u 1765.8	A_{2u} 1389.5	E_u 603.9	A_{2u} 986.0	A_{2u} 808.0
29	30	A_1 4128.3	E 1765.8	A_1 1389.5	E 603.7	A_1 985.9	A_1 807.8
30	30	A_{2u} 4128.2	E_u 1765.8	A_{2u} 1389.4	E_u 603.6	A_{2u} 985.8	A_{2u} 807.7

Table 7.6: Vibrational modes of several different isotopic combinations of ^{a}Si–H^0 –^{b}Si in silicon, calculated in the $Si_{98}H_{78}$ cluster. The proton and also the muon pseudo-isotope display an additional E mode above the Raman edge absent for the other isotopes. The energy of the A transitions is sufficiently large to ionize the $E_c - 0.16$ eV level of ^{1}H^0 [193], effectively preventing the observation by IR, but the lower energy E_u mode of ^{1}H^0 may be observable.

may exist as a metastable species if the barrier to conversion into H_2^* is large enough. The barrier between H_2^{**} and $H_2^* \sim 0.7$ eV. During the diffusion path, the hydrogen atom hops into a neighboring Si–Si bond, as shown in figure 7.1b. This structure is a weak local minimum, and can be relaxed as a separate structure. This allows an alternative method to be used to calculate the diffusion barrier. To simulate the migration of an atom through a chosen set of fully optimized local minima, the structure was relaxed at points along the vector defining the direct trajectory connecting neighboring metastable sites, while simultaneously modifiying the force experienced by the atoms such that the dot product of the new force with the constraint vector is zero. The energy of this trajectory is shown in figure 7.1.

An alternative dimer structure [195], consists of two *equivalent* hydrogen atoms sharing a common bond-centered site with C_{2h} symmetry. To compare the properties of this dimer against the observed spectra we simulate this structure in the $Si_{98}H_{78}$ cluster or MP-2^3 54 atom supercell. There are two possible forms of this defect with the hydrogen atoms either staggered or eclipsed to the neighboring Si–Si bonds (see figure 7.2). Neither of these structures possess levels in the gap. We find the staggered form to be ~ 0.3 eV higher in energy than the H_2^{**} structure.

	^{1}H^{1}H	^{2}H^{1}H	^{2}H^{2}H
	D_{3d}	C_{3v}	D_{3d}
cluster	A_{2u} 2019.8(0.0528)	A_1 2017.8(0.0256)	A_{2u} 1452.4
	A_{1g} 2015.7(0.0)	A_1 1451.2(0.0049)	A_{1g} 1450.0
	E_u 809.8(0.1838)	E 808.3(0.0834)	E_u 576.4
	E_g 806.8(0.0)	E 575.4(0.0436)	E_g 574.4
supercell	A_{1g} 1880.0	A_1 1877.2	A_{1g} 1348.6
	A_{2u} 1874.2	A_1 1347.0	A_{2u} 1345.5
	E_g 762.5	E 760.6	E_g 545.5
	E_u 758.7	E 543.8	E_u 542.2
Expt.	A 1792.	ND	1307.
	A^* 1607.6	ND	ND
	E 812.4	ND	581.

Table 7.7: Calculated and experimental modes of H_2^{**}, cm^{-1}, the A^* mode is the A overtone referred to in the text, several of the modes marked ND have not been detected. Frequencies from the $Si_{98}H_{78}$ cluster and 54 atom supercell calculations are shown with the character of the vibrational modes. $\frac{\eta^2}{M}$ is shown in brackets for modes in the cluster calculation, i.e., the square of the effective charge over the mass of the oscillator in e^2/a.m.u., only the A_{2u} and E_u modes are infrared active. The overtone mode must be of symmetry A_{2u} and consists of a mixture of the E_u and E_g modes.

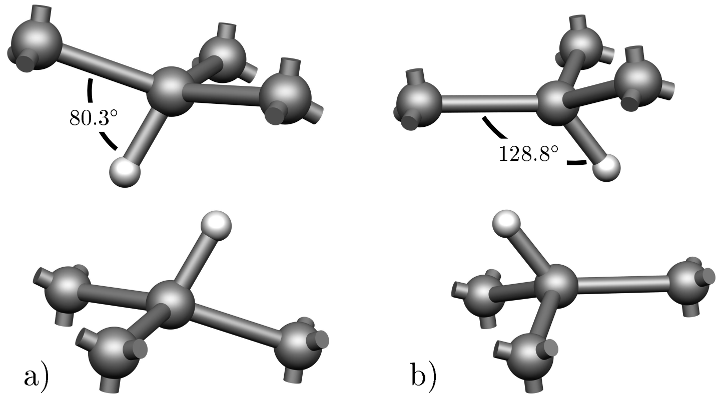


Figure 7.2: The a) staggered and b) eclipsed forms of the alternative dimer. Both structures are electronically inactive. The eclipsed structure is ~ 0.1 eV higher in energy than the staggered form, giving this as a barrier to rotation for this dimer. The vibrational modes of the staggered form are shown in table 7.8.

The eclipsed structure is found to be ~ 0.1 eV higher in energy than the staggered form, implying that the dimer can easily rotate leading to effective D_{3d} symmetry. For the fixed, staggered, form of the dimer the vibrational modes, as calculated in the same supercell as H_2^{**}, are shown in table 7.8. The IR active vibrational modes of this structure do not match those observed for H_2^{**}, particularly the low lying A_u mode which is a singlet, whereas the experimental mode is a doublet. Even if the structure rapidly rotates to give an effective symmetry of D_{3d}, this A_u mode cannot explain the observed IR active doublet unless there is another nearby mode of B_u symmetry to mix with this A_u vibration. These calculations find no such B_u mode above the Raman edge of the silicon lattice. Additionally, the large, ~ 14 cm^{-1}, shift of the high-frequency B_u mode of this defect in the isotopically mixed case is inconsistent with the absence of a distinct mixed mode of the observed dimer.

7.4 Summary

The calculations have shown that modes detected at 1792 cm^{-1} and 812 cm^{-1} are probably due to a new form of hydrogen dimer, labelled H_2^{**}. Here, the H atoms are located at AB sites to a common Si–Si bond. The high D_{3d} symmetry of the defect

^{1}H^{1}H	^{2}H^{1}H	^{2}H^{2}H
C_{2h}	C_{1h}	C_{2h}
B_u 1963.9	A' 1950.2	B_u 1408.7
A_g 1936.2	A' 1399.7	A_g 1390.7
A_g 1080.9	A' 937.6	A_g 765.3
A_u 566.2	A'' 563.9	A_u 520.2
B_g 562.3	A'' 514.3	B_g 513.3

Table 7.8: Modes of the alternative dimer [195] as shown in figure 7.2a, calculated using the 54 atom supercell with 2^3 k-points. This structure is ∼0.3 eV higher in energy than H_2^{**}, and the only infrared active modes (B_u and A_u) do not match the observed properties of H_2^{**} (as shown in table 7.7).

makes only a few of the possible modes infrared active. The defect is electrically inert and appears to be formed just before H_2^*.

Like, H_2^*, the formation mechanism of H_2^{**} is not understood. In low temperature proton implanted material, H appears to form almost exclusively H_{BC}^+ defects. This is to be contrasted with muon implantation where Mu^0 defects are prominent and may reflect the different irradiation damage in the two case. Proton irradiation creates V^- and V^{2-} defects. It is difficult to see how neutral dimers can form, without H^+ being trapped by V^- or V^{2-} first.

Subsequent reactions with interstitials must be very rapid as modes of VH_2 or VH are not detected at this stage. Nevertheless, these reactions can lead to H_2^{**} and H_2^* formation but further work is necessary to unravel the mechanisms.

Chapter 8

Calculations upon interstitial silyl and silane molecules in silicon

8.1 2222 cm^{-1} – vacancy or interstitial related?

Numerous infrared absorption lines associated with hydrogen have been observed by various experimental groups [196], with those related to complexes of native defects with hydrogen being particularly common in irradiated and proton implanted material.

A large number of infrared absorption lines have been observed in the range 1800 – 2300 cm^{-1} after proton implantation, which shift downwards in frequency by $\sim \sqrt{2}$ in deuteron implanted material [197]. This implies that the modes are due to the stretch in covalent bonds between hydrogen and silicon, since this frequency range is near to that of the Raman active A_1 mode at 2187.0 cm^{-1} and the IR and Raman active T_2 mode at 2190.6 cm^{-1} of the isolated silane molecule [198].

One particular infrared line observed at 2210 cm^{-1} in room temperature measurements, but rising to 2222 cm^{-1} on cooling to 10 K, has been a source of dispute within the literature. Different workers have assigned this line (and subsequently a class of related defects) to two distinct types of native defects: namely either to a fully hydrogenated lattice vacancy, VH_4 [199], or to a Si interstitial bonded to four interstitial hydrogen atoms as in the molecule SiH_4 which has been trapped within the Si lattice.

Early infra-red absorption studies [196] on irradiated Si containing both H and D, found that the 2222 cm^{-1} line was consistent with a defect of T_d symmetry and containing four H atoms bonded to Si. It was also suggested that, as this mode was close to those of the silane (SiH_4) molecule, the defect consisted of this molecule lying at the T_d interstitial site [200]. Confirmation that the defect had T_d symmetry came from uniaxial stress measurements [201].

Subsequent *ab initio* calculations were consistent with an assignment of the 2222 cm^{-1} mode to the fully hydrogenated vacancy, VH_4. The calculated T_2 stretch frequencies obtained by two different groups [201, 202] at 2319 and 2334 cm^{-1} respectively lie within 120 cm^{-1} of the observed line. Moreover, the calculations were able to predict that the infra-red inactive, but Raman active, A_1 mode lies *above* the T_2 mode. This was confirmed by observation of an infra-red active mode for the VH_3D defect lying above 2222 cm^{-1} [201]. Indeed, assuming this assignment, infra-red [203] and electron paramagnetic studies [204] on low temperature proton implantation experiments have indicated that VH_n $(n < 4)$ and V_mH_n $(m > 1)$ are readily formed. Naturally, the lattice damage caused by proton implantation creates both vacancies and interstitials and both defects contain unsaturated Si bonds which would be expected to complex with H. Further studies, both theoretical [205] and experimental [182], did indeed reveal that a bonded interstitial Si atom can complex with H. Moreover, the evidence that lines at 1986.5 and 1989.4 cm^{-1} and assigned to a [100] oriented split-interstitial which has trapped two H atoms is overwhelming as the effect of uniaxial stress, and the effect of isotopic substitution, leads to a detailed model entirely consistent with both the *ab initio* theory and experimental data. There is then no evidence from these studies that absorption lines arising from $(SiH_4)_i$ are present.

However, the assignment of the 2222 cm^{-1} mode to VH_4 has recently been questioned by Suezawa [206, 207]. IR studies on Si grown in a H-atmosphere reveal the 2222 cm^{-1} band whose integrated intensity increases with B, C and Au doping. These elements are assumed to lead to an increase in the production of Si_i which could then lead to increased concentrations of $(SiH_4)_i$. It is known that interstitial related A- and B-swirl defects are produced during the growth of Si doped with C and B while vacancy related D-defects are suppressed [208]. Specifically, in float-zone samples with $[C] = 5 \times 10^{16}$ cm^{-3} or $[B] = 1.5 \times 10^{16}$ cm^{-3}, the intensity of the 2222 cm^{-1} band was enhanced relative to material doped with 5×10^{18} cm^{-3} Sn. Although, this is consistent with the idea that the 2222 cm^{-1} band is interstitial related, this does not provide irrefutable evidence. It is known that both B and C react with interstitials via a kick-out mechanism, moving from substitutional sites to interstitial positions themselves [209, 210] and so removing interstitials from the lattice. Similarly, Sn can trap vacancies below 180 °C [211, 212] and this then might prevent VH_4 being formed in the Sn-doped material. It is more difficult to understand whether gold-doping should act to remove interstitials from the lattice, since Au normally sits at substitutional sites in the lattice, and so would be expected to remove vacancies from the lattice instead of interstitials.

Moreover, arguments about $(SiH_4)_i$ rather than VH_4 being produced in Si containing impurities with smaller covalent radii than Si must be treated with reservation. It is known that VH_4, unlike the vacancy, compresses the surrounding lattice since VH_4 is 67% larger than the vacancy [201]. Such an expansion can be argued to favour the production of VH_4 in material doped with elements with small atomic radii such as B or C.

Further work by Suezawa [213, 214] demonstrates that there is a growth in the intensity of the 2222 cm^{-1} band around 180 °C which is correlated with a decrease in the intensity of the 0.34 eV optical transition [215] due to V_2^-. The suggestion is made that Si interstitial clusters break up around this temperature releasing interstitials which are trapped by divacancies, or hydrogen leading to $(SiH_4)_i$ defects. Such experiments once again are capable of alternative interpretations, where divacancies are thermally dissociated releasing vacancies which are subsequently trapped by hydrogen forming VH_4.

In addition to the controversy in the assignment of the 2222 cm^{-1} line, the center responsible for the 2190.3 and 2166.1 cm^{-1} lines – previously assigned to VH_3 [201, 203] – has been reassigned by Suezawa [213] to SiH_3.

8.2 Calculations

Given these contrary viewpoints, there is a clear need to investigate theoretically the structure and vibrational properties of the interstitial silane and silyl molecules to ascertain whether assignment to the 2222 and 2166 cm^{-1} modes is credible.

The defect molecules were each inserted into a $Si_{84}H_{64}$ cluster centered on the tetrahedral interstitial site. The wave-function basis consisted of independent s and p Gaussian orbitals, with four different exponents, sited at each Si site and three at each H atom of the molecule. A fixed linear combination of two Gaussian orbitals was sited on the terminating H atoms. In addition, single Gaussian functions were placed at each Si–H bond of the molecule and every Si–Si bond center. The charge density was fitted with five independent Gaussian functions with different widths on each Si atom, and four (three) on the central (terminating) H atoms. Two extra Gaussian functions were placed at each bond center. All atoms, except the terminating H ones, were allowed to relax by a conjugate gradient method. The second derivatives of the energy were found for the atoms of the interstitial molecule.

As a preliminary to calculating the vibrational properties of *interstitial* silane, the isolated molecule in free space was simulated using the same type of basis. The calculated quasiharmonic modes are shown in table 8.1. It is worth noting that whilst the one-dimensional A_1 modes match experiment well (within ~ 5 cm^{-1}) in the isotopically pure cases, the higher dimensional (T_2 and E) modes and those of the isotopically mixed molecules are of poorer agreement (~ 60 cm^{-1}). This leads to an incorrect ordering of some of the vibrational levels, when compared with experiment.

There are two forms of interstitial $(SiH_4)_i$ with T_d symmetry, with the four Si–H bonds aligned either towards the four nearby hexagonal sites, or alternatively towards the four neighbouring Si atoms. The first configuration is calculated to be 9.3 eV lower in energy than the second. This is due to strong interaction between the molecular hydrogen atoms and the nearby silicon atoms in the surrounding cage, which are forced close together in the second configuration.

To test the convergence of the molecular properties with respect to cluster size, the lower energy structure was re-calculated in a cluster with all of the atoms, including the surface terminating hydrogen allowed to relax. In a cluster of this size, the fixed-surface relaxation could be expected to give an unphysically rigid surrounding to the molecule, and the relaxed surface cluster would give an excessively flexible model for the surrounding crystal, hence the two types of simulation will give results lying on either side of the true defect properties. Similarly, to test for effects of increasing the basis, the calculations were repeated in the 148 (plus defect) atom cluster with fixed surface, but with a larger basis using eight different exponents sited at each Si atom for both the wavefunction and charge density,

SiH_4 (T_d)	SiH_3D (C_{3v})	SiH_2D_2 (C_{2v})	$SiHD_3$ (C_{3v})	SiD_4 (T_d)
Isolated $(SiH_4)_g$				
$2190(2187.0)A_1$	$2174A_1$	$2156(2198)A_1$	$2138(2182)A_1$	$1549(1545)A_1$
$2123(2190.6)T_2$	$2122E$	$2121(2183)B_2$	$1546(1573)A_1$	$1541(1597)T_2$
	$1546A_1$	$1545(1587)A_1$	$1542(1598)E$	
		$1544(1601)B_1$		
Trapped				
Theoretical$(SiH_4)_i$				
$1812(A_1)$	$1809(A_1)$	$1806(A_1)$	$1803(A_1)$	$1312(T_2)$
$1802(T_2)$	$1801(E)$	$1801(B_2)$	$1313(E)$	$1281(A_1)$
	$1306(A_1)$	$1314(B_1)$	$1313(A_1)$	
		$1297(A_1)$	$1289(A_1)$	
Experimental(XH_4)				
$2221.9(T_2)$	$2250.4(A_1)$	$2243.5(A_1)$	$2236.1(A_1)$	$1616.6(T_2)$
NIR(A_1)	$2223.5(E)$	$2225.4(B_1)$	$1636.1(A_1)$	NIR(A_1)
	$1620.3(A_1)$	$1628.3(A_1)$	$1615.6(E)$	
		$1614.6(B_2)$		

Table 8.1: Calculated quasi-harmonic frequencies (cm^{-1}) for the fundamental vibrational transitions of isolated, non-rotating, silane molecules (experimental frequencies [198, 216] are shown in brackets), and the low energy form of trapped interstitial silane at the T_d site, as compared with observed frequencies for the tetrahedral XH_4 defect [203] (*note:* the two entries marked NIR are infrared inactive modes).

in addition, a second Gaussian was placed between every bonded pair of atoms. The effects on the molecular vibrational frequency of these changes are shown in table 8.2, the frequency is seen to be converged to within ~ 100 cm^{-1} with respect to both cluster size and basis.

8.3 Results

The calculated quasiharmonic vibrational modes of the low energy form of interstitial silane are shown in table 8.1, and compared with the experimental XH_4

mode	Fixed Surface	Relaxed Surface	Increased Basis
A_1	1812	1861	1877
T_2	1802	1755	1829

Table 8.2: Effect of allowing free-relaxation of the surface of the $Si_{84}H_{64}$ cluster, or increasing basis size on the higher frequency modes of interstitial silane, as described in the text.

defect modes [203] The calculated frequencies for the interstitial molecule differs from experimental values by > 300 cm^{-1} for all modes, which is large enough to be reasonably confident in ruling out SiH_4 as being the defect species responsible.

To further test the assignment of the 2222 cm^{-1} to interstitial SiH_4, the energy of this defect was compared against that of two experimentally known defects with an equivalent stoichiometry. The previously identified hydrogen saturated split–interstitial [182] and the interstitial hydrogen molecule [143] were chosen as obvious candidates for competing defects in the material. The calculated energy difference, using the same basis as the interstitial–hydrogen complexes, between $(IH_4)_i$ and $(IH_2)_i$ + $(H_2)_i$ (with an infinite separation between the components) is 2.74 eV in favour of $(IH_2)_i$ + $(H_2)_i$.

We consider next the assignment of the 2190.3 and 2166.1 cm^{-1} lines to $(SiH_3)_i$. There are four possible structures for trigonal interstitial silyl SiH_3 molecules with C_{3v} symmetry. The calculated relaxed forms of the neutral defect are shown in figure 8.1. Structure A has the three equivalent hydrogen atoms aligned towards three of the nearby hexagonal sites, with the molecular dangling bond pointing towards the fourth hexagonal site. Structure B has the dangling bond pointing instead towards the nearby silicon atom in the opposite direction from the hexagonal site. Structures C and D are similar to A and B respectively, but rotated about the molecular three-fold axis so that the hydrogen atoms approach nearby Si–Si bonds closely. The relative energies and quasiharmonic modes of the four structures, relaxed in the positive, neutral and negative charge states are shown in table 8.3.

The calculated vibrational modes for the interstitial silyl molecule, for both of the degenerate low–energy A and B structures and in all three charge-states,

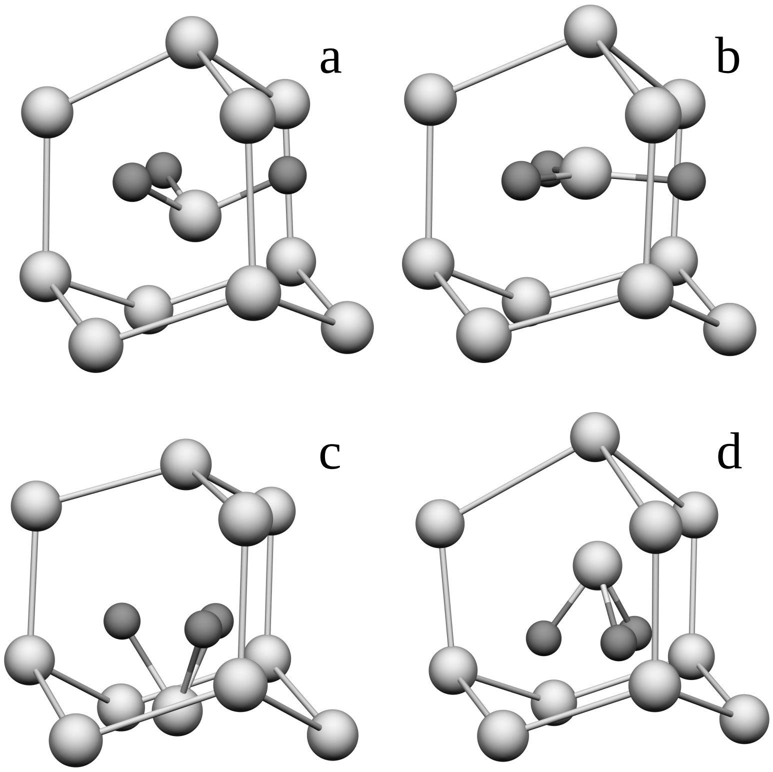


Figure 8.1: Relaxed structures for the neutral form of the four distinct C_{3v} configurations of *interstitial* SiH_3

differ from experiment by at least 300 cm^{-1}. Similarly for the intermediate energy ($\sim$ 1.6 eV above ground state for all charges) C structures, again all vibrational modes do not match the experimental values. Only the high energy ($\sim$ 4.2 eV above the lowest energy structures for the three charge states) D structure has modes that are close to the experimental values, but due to the high energy of this structure, it must be discounted as a possible candidate for the defect.

So for both the silane and silyl molecules, the calculated vibrational frequencies differ from the experimental observed spectra by more than 300 cm^{-1} for all structures, with the exception of the very high energy D structure. Taken together, these results rule out SiH_4 and SiH_3 molecules as the candidates for the defects responsible for the 2222 cm^{-1} and 2190.3/2166.1 cm^{-1} groups of lines.

The large downward shifts in the H stretch frequencies for the interstitial molecules compared with free molecules are reminiscent of the 540 cm^{-1} red-shift found for the stretch mode of the interstitial H_2 molecule (see chapter 5). This is believed

Calculated	A	B	C	D
+1				
Energy	0.21	0.00	1.84	4.37
SiH_3 modes	1853 (E)	2045 (E)	1643 (A_1)	2373 (A_1)
	1551 (A_1)	1949 (A_1)	1194 (A_1)	2239 (E)
			1067 (E)	1155 (A_1)
0				
Energy	0.00	0.04	1.61	4.29
SiH_3 modes	1668 (E)	1849 (E)	1844 (A_1)	2317 (A_1)
	1371 (A_1)	1629 (A_1)	1369 (A_1)	2162 (E)
			1228 (E)	
			1161 (E)	
-1				
Energy	0.00	0.00	1.48	4.16
SiH_3 modes	1390 (E)	1032 (E)	1971 (A_1)	1873 (A_1)
	1188 (A_1)	1006 (A_1)	1519 (A_1)	1271 (E)
			1513 (E)	
			1401 (E)	
Experimental:	2190.3 (A_1)	2166.1 (E)		

Table 8.3: Relative energies (eV) and modes (cm^{-1}) of the four C_{3v} configurations of interstitial SiH_3 in the +1, neutral and -1 charge states, and the experimental frequencies of XH_3

to arise from the screening of the molecular proton-electron attraction by the lattice charge distribution partially filling the molecular anti-bonding orbitals, leading to a weakened H–H bond. A similar mechanism may operate in the present case and suggests a universal effect.

8.4 Summary

The *ab initio* calculations offer no support to the idea that the 2222 cm^{-1} line originates from an interstitial SiH_4 molecule. The frequency of the interstitial silane molecule is found to lie around 1800 cm^{-1}. There are no examples apparent in the literature of modes in this region due to tetrahedral defects and hence we can conclude that such interstitial molecules do not exist in detectable concentrations.

Chapter 9

Hydrogen complexes with subsitutional carbon

9.1 Introduction

As-grown samples of silicon contain carbon. If the material has not been irradiated, this impurity is present at isolated substitutional sites in the lattice [217]. Such carbon is electronically inactive in silicon, but will readily complex with other atoms such as hydrogen [218], oxygen [217] or interstitial carbon [219] to form electrically active defects.

Complexes formed with single hydrogen atoms were detected by DLTS experiments by Endrös *et al.* [218, 220]. The CH pair was shown to possess a deep donor level (E_3) at $E_c - 0.16$ eV, there was an initial concern that this was very similar to the E_3' centre at $E_c - 0.17$ eV [221] (due to H_{BC}), but this was resolved by double-correlation DLTS [222].

The symmetry of the centre responsible for the E_3 donor carbon hydrogen pairs were measured to be trigonal [223] by uniaxial stress. The stability of the defect responsible for this level appears to be charge state dependent. In the dark the centre anneals out at around 275 K with an activation energy of 1.33 eV [220, 224–226]. Based on an analysis of the effects of biasing the Schottky-structure used for DLTS, it was suggested that the C–H complex dissociates by capturing an electron, changes its charge state [220, 227], and then decomposes from this neutral state with an activation energy of 0.5 eV [228, 229]. It has been suggested that the structure of this complex is a bond centred hydrogen atom between the substitutional C and a neighbouring Si atom.

A second carbon hydrogen complex, H_1, has also been observed, with a hole trap at $E_v + 0.33$ eV in p-type silicon [227, 230]. Unusually, this centre does not

stability (K)	^{12}C H	^{13}C H	^{12}C D	^{13}C D
120–220	1884.6	1884.3(-0.3)	1362.5(-522.1)	1362.3(-522.3)
120–220	595.5	578.3(-17.2)	596.0(0.5)	577.8(-17.7)
120-220	660.5	640.5(-20.0)	660.0(-0.5)	639.8(-20.7)
20–150	–	–	1363.3	1363.3(0.0)
C_s	^{12}C	^{13}C	^{14}C	
–	607	589(-18)	573(-34)	
H_{BC}	H^+	D^+		
9–200	1998	1449(-549)		

Table 9.1: Observed vibrational modes, cm^{-1}, of carbon-hydrogen related defects, including vibrational frequencies of substitutional carbon [231, 232], and bond-centred hydrogen [180].

display the Poole-Frenkel effect in fields of less than 6×10^5 Vm^{-1}, i.e., an electrical field does not cause shift in the activation energy of the level, suggesting that H_1 is not an acceptor. In the dark this complex is stable to 100 °C, and anneals out with an activation energy of about 1.7 eV. This complex is insensitive to illumination or reverse biasing, unlike the E_3 centre. Kamiura *et al.* provisionally ascribed H_1 to an anti-bonding configuration of hydrogen attached near to substitutional carbon [227].

So far, there has been only one infrared absorption study of C_sH_i complexes by Hoffman *et al.* [233]. After proton (or deuteron) implantation at ∼20 K and annealing at 180 K, two correlated modes at 596 (595) and 1884 (1363) cm^{-1} were observed. These modes are close to those of un-complexed C_s and H_{BC}, suggesting, in contrast to the DLTS data, that the structure observed does not have a direct C–H bond. The effective charge of the hydrogen related mode is found to be $\sim 2.2\,e$ which is close the effective charge of H_{BC}^+ which is 1.8 e [126]. In addition a weaker mode at 660.5 cm^{-1}, which is correlated to the 596 and 1884 cm^{-1} modes, has also recently been observed [234]. The C–H modes observed so far, and the temperature range over which each is present are summarised in table 9.1.

There are four obvious possible structures for C_sH_i that have been considered theoretically by various groups. Two structures are similar to bond-centred hydrogen, sited between the substitutional carbon and a nearest neighbouring Si atom, with a bond from the H to either carbon or alternatively the Si atom, and these are denoted as $\mathrm{CH}^{\mathrm{I}}_{BC}$ and $\mathrm{CH}^{\mathrm{I_a}}_{BC}$. The other two are hydrogen anti-bonded to the back of the $\mathrm{C_s}$–Si structure on either the carbon or the silicon end of the bond, i.e., $\mathrm{CH}^{\mathrm{I}}_{AB}$ and $\mathrm{CH}^{\mathrm{II}}_{AB}$, the H–X bond in these structures may be sufficiently long to place the H atom at the near-by tetrahedral ($\mathrm{T_d}$) site. The sites near to a substitutional carbon atom are shown schematically in figure 9.1.

Early work was performed on the neutral CH pair at the HF level of theory by Maric *et al.* [235, 236], they found that the $\mathrm{CH}^{\mathrm{I_a}}_{BC}$ and $\mathrm{CH}^{\mathrm{I}}_{AB}$ structures were the most stable, with $\mathrm{CH}^{\mathrm{I_a}}_{BC}$ being the lowest by 0.05 eV. Both structures gave rise to defect levels in the gap. The other structures were found to be considerably higher in energy.

This contrasts strongly with later LDA calculations by Kaneta *et al.* [237, 238], which found $\mathrm{CH}^{\mathrm{I}}_{BC}$ to be both the ground state structure and also the only stable bond-centred structure that the hydrogen adopted near $\mathrm{C_s}$. After a combination of further HF and post-HF cluster calculations, and also DFT supercell simulations [239] it was shown that the ground structure of the complex in the 0 and + charge states was sensitive to the effects of electron correlation (which would naturally be absent from pure HF calculations). In these calculations, the effect of electron correlation was found to be to stabilise the $\mathrm{CH}^{\mathrm{I}}_{BC}$ structure over $\mathrm{CH}^{\mathrm{I_a}}_{BC}$ in both the 0 and + charge states. The $\mathrm{CH}^{\mathrm{I}}_{AB}$ structure was found to be metastable in both charge states.

Further LDA calculations were performed by Kaneta *et al.* [237] to investigate the stability of the various structures for positive, neutral and negative charge states. $\mathrm{CH}^{\mathrm{I}}_{BC}$ was found to be the most stable structure for all three charge states (see table 9.2), with binding energy of 1.20 eV in the neutral structure. In the negative charge state, the structure with H^- sited at the tetrahedral site near a Si neighbouring the $\mathrm{C_s}$ was found to be only 0.06 eV higher in energy than the ground state, suggesting that the H^-_{BC} structure could dissociate via this intermediate. The two metastable H_{AB} structures were found to be nearly degenerate in energy (0.57 and 0.61 eV for $\mathrm{CH}^{\mathrm{II}}_{AB}$ and $\mathrm{CH}^{\mathrm{I}}_{AB}$ respectively).

Structure	Kaneta *et al.* [237].			Leary *et al.* [240].		
	+	0	-	+	0	-
CH^{I}_{BC}	0.00	0.00	0.00	0.00	0.00	0.04
CH^{I}_{AB}	0.70	0.60	0.30	0.49	0.33	0.00
$CH^{I_a}_{BC}$	Unstable in all calculations					
CH^{II}_{AB}	0.71	0.57	0.06	1.97	1.12	0.79

Table 9.2: Relative energies (eV) for the four C_sH_i structures in silicon considered by other authors.

Leary *et al.* [240] also performed cluster based LDA calculations on the CH pairs in an 86 atom system. The relative energies of the structures considered are also summarised in table 9.2. As with the work of Kaneta *et al.*, CH^{I}_{BC} was found to be the most stable structure in the neutral and +1 charge states, but $H_{AB}C_s$ was found to 0.04 eV lower in energy for the -1 charge state. These calculations found the nearly degenerate CH^{I}_{AB} and CH^{II}_{AB} structures of Kaneta *et al.* to be split in energy by 0.75 eV. The binding energy for the neutral complex was calculated to be 1.01 eV to form CH^{I}_{BC} from separated $C_s + H_{BC}$.

9.2 Calculations

The theoretical studies in this work use local density-functional formalism as implemented in the AIMPRO cluster code [21].

The defects were each inserted into a $Si_{162}H_{103}$ cluster centered on a Si atom. The wave-function basis consisted of independent s and p Gaussian orbitals, with four different exponents sited at each Si or C atom, and three at each H atom of the dimer. A fixed linear combination of two Gaussian type orbitals was sited on the terminating H atoms. In addition, a Gaussian function was placed at the center of each bond in the cluster. The charge density was fitted with five independent Gaussian functions with different widths on each Si atom, five on each carbon atom, and four (three) on the central (terminating) H atoms, with two extra Gaussian

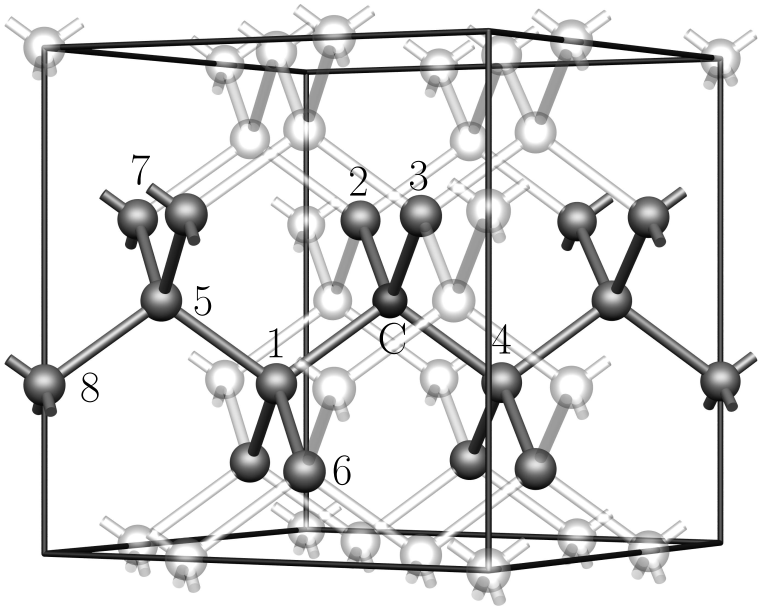


Figure 9.1: Labels of atoms around a substitutional carbon atom in the silicon lattice. A chain of atoms along [110] are marked to show the unique carbon hydrogen configurations. CH^{I}_{BC} is sited between the carbon atom and Si_1 (but closer to the C atom), CH^{I}_{AB} is anti-bonding on the back of the C atom away from Si_1. $CH^{I_a}_{BC}$ is similar to CH^{I}_{BC} but displaced towards Si_1. CH^{II}_{AB} is anti-bonding on the back of Si_1 away from the C. CH^{II}_{BC} lies between Si_1 and Si_5, $CH^{III_a}_{BC}$ lies between Si_5 and Si_7, and $CH^{III_b}_{BC}$ is between Si_5 and Si_8.

functions placed at each bond center. All atoms, except the terminating H ones, were allowed to relax by a conjugate gradient method.

The normal modes of the system were examined using the quasi-harmonic second derivatives of the energy, which were evaluated for the defects using the method described in Ref. [21]. The second derivatives were calculated directly for the C and H atoms, and also neighbouring silicon atoms with the rest of the dynamical matrix filled using a previously evaluated Musgrave-Pople potential. The effective charges of selected modes were evaluated by finding the change in the total dipole moment of the cluster as the atoms were displaced according to the normal coordinates of each of the modes.

Additionally, several many-body electronic transitions were obtained by applying the empirical correction method of Resende *et al.* [34] to the Slater electronic transition state [12] for each of the defects. The reference defects chosen for this method were the 0/+ transition of H_{BC} at $E_c - 0.17\ e$, [221] and in the absence of a $-/0$ transition between similar structures of hydrogen, the carbon interstitial at $Ev - 0.1\ e$ [241].

Further calculations have also been performed [242] using $Si_{68}H_{66}$ clusters, with a larger basis of 8 wavefunction and charge-density Gaussian functions on every carbon and silicon atom, with the same basis as used above for the H atoms, and three bond-centre functions for both basis sets.

9.3 Results

Structure	$Si_{161}H_{103}CH$ +	$Si_{161}H_{103}CH$ 0	$Si_{67}H_{66}CH$ +
$CH^{I}_{BC}(C_{3v})$	0.00	0.00	0.00
$CH^{I}_{AB}(C_{3v})$	0.44	0.40	–
$CH^{I_a}_{BC}(C_{3v})$	Unstable		
$CH^{II}_{AB}(C_{3v})$	0.43	0.39	–
$CH^{II}_{BC}(C_{1h})$	0.63	1.31	0.20
$CH^{III_a}_{BC}(C_{1})$	0.87	–	0.20
$CH^{III_b}_{BC}(C_{1h})$	0.76	–	0.20
Isolated C_s and H_i	0.83	1.56	–

Table 9.3: Relative energies (eV) for the CH defects considered in the $Si_{161}H_{103}CH$ and $Si_{67}H_{66}CH$ clusters.

The relative structural energies of the structures simulated are presented in table 9.3. The modes obtained in the large-basis cluster calculations are shown in table 9.4. The vibrational mode of isolated H^+ is found to lie within 150 cm^{-1} of experiment in these calculations.

The CH^{I}_{BC} structure is found to be the most stable in both the neutral and +1

	^{12}C, H	^{12}C, D	^{13}C, H	^{13}C, D
H^+	1852.1	1316.2(-535.9)		
$CH^{I+}_{BC}(C_{3v})$	2624.1	1928.1(-696.0)	2616.3(-7.8)	1916.8(-707.3)
	1424.7	1049.1(-375.6)	1421.0(-3.7)	1042.8(-381.9)
	1413.3	1042.0(-371.3)	1409.4(-3.9)	1035.4(-377.9)
	566.1	555.3(-10.8)	554.1(-12.0)	546.2(-19.9)
	556.6	548.9(-7.7)	548.9(-7.7)	543.2(-13.4)
	555.2	545.9(-9.3)	544.0(-11.2)	537.7(-17.5)
$CH^{II+}_{BC}(C_{1h})$	1795.0	1275.3(-519.7)	1794.9(-0.1)	1275.3(-519.7)
	642.6	642.5(-0.1)	623.7(-18.9)	623.5(-19.1)
	574.1	571.6(-2.5)	561.8(-12.3)	559.1(-15.0)
	558.6	555.4(-3.2)	546.5(-12.1)	542.8(-15.8)
$CH^{III^+_a}_{BC}(C_1)$	1802.3	1282.0(-520.3)	1802.3(0.0)	1282.0(-520.3)
	823.2	653.2(-170.0)	823.2(0.0)	634.8(-188.4)
	652.9	625.8(-27.1)	633.8(-19.1)	625.0(-27.9)
	605.0	605.0(0.0)	588.7(-16.3)	588.7(-16.3)
	597.3	597.2(-0.1)	581.0(-16.3)	580.9(-16.4)
$CH^{III^+_b}_{BC}(C_{1h})$	1706.8	1214.7(-492.1)	1706.8(0.0)	1214.7(-492.1)
	952.5	712.4(-240.1)	952.5(0.0)	712.0(-240.5)
	661.2	660.3(-0.9)	641.6(-19.6)	641.0(-20.2)
	609.5	609.3(-0.2)	592.7(-16.8)	592.6(-16.9)
	597.7	597.6(-0.1)	581.3(-16.4)	581.3(-16.4)

Table 9.4: Local vibrational modes, cm^{-1}, from the cluster calculations with the larger basis [242] for C-H complexes in Si. Isotopic shifts of the modes are shown in parenthesis.

charge states. The vibrational modes (table 9.4) of this structure do not match those observed, particularly the high frequency stretch and bend modes, which leads to an increase in the energy of this structure by $\sim$ 0.1 eV when compared to the other structures with lower frequency modes, due to the high zero point energy of the hydrogen related modes. In this cluster calculations, this leads to this structure being the lowest energy by $\sim$ 0.1 eV when compared to all of the other bond-centred structures examined. The question of why this structure is not observed is resolved by considering the intensity of these hydrogen-related mode which, as with C–H bonds generally, is weak due to the small effective charge, and will be further reduced by a factor of three due to the trigonal symmetry [63]. The large basis cluster calculations find that the square of the effective charge over the mass of the oscillating mode for the high-frequency stretch mode at 2411 cm^{-1} is 0.11 e^2/a.m.u. (for the smaller basis calculations, the effective charge of this mode is only 0.025 e^2/a.m.u. for this mode, suggesting that the basis used in these calculations is far from converged for the C atom).

Perhaps surprisingly, the 0/+ electrical level of this structure is found to lie at $E_v + 0.22$ e, in close agreement to the H_1 DLTS centre. The highest occupied Kohn-Sham level, as shown in figure 9.3, suggests why this should occur, since this level is principally localised on the dangling-bond like environment at the silicon atom. If this assignment of the H_1 centre is correct, it naturally explains the lack of a Poole-Frenkel effect at this centre, since the neutral state of the defect, before trapping a hole, would be little effected by an applied electrical field.

CH$^{\mathrm{I}}_{AB}$ is found to posses a 0/+ level at the valence band, and a $-/0$ level at $E_c - 0.33$ eV, implying it is always neutral or negative.

The 0/+ level of the CH$^{\mathrm{II}}_{AB}$ centre is found to be well below the valence band, hence again this centre is only neutral or negative. The $-/0$ transition of this defect is found to lie at $E_v + 0.27$ e (table 9.5).

CH$^{\mathrm{II}+}_{BC}$ is found to posses a set of vibrational modes at 1795.0, 642.6, and 574.1 cm^{-1} (table 9.4) which agree reasonably well with the observed modes. This structure gives the closest to experiment in terms of absolute frequencies and isotopic shifts. Unlike CH$^{\mathrm{I}+}_{BC}$, the high frequency mode at $\sim$1940 cm^{-1} is found to be intense, with the square of the effective charge over the mass of the oscillating mode

Structure	transition	type	stability
CH^{I}_{BC}	$E_v + 0.22$	0/+	–
CH^{I}_{AB}	E_v	0/+	–
CH^{II}_{AB}	$E_v - 0.33$	0/+	–
CH^{II}_{BC}	$E_c - 0.21$	0/+	–
$CH^{III_a}_{BC}$	$E_c - 0.27$	0/+	–
$CH^{III_b}_{BC}$	$E_c - 0.24$	0/+	–
CH^{I}_{AB}	$E_c - 0.33$	$-/0$	–
CH^{II}_{AB}	$E_v + 0.27$	$-/0$	–
E_3	$E_c - 0.16$	0/+	< 273 K
H_1	$E_v + 0.33$	hole-trap	<100 °C

Table 9.5: Electronic transitions of several structures compared against the observed E_3 and H_1 DLTS lines [218, 230], the thermal stability of the defects is given for the complex in the dark. The levels are calculated by applying the empirically corrected transition state method of Resende *et al.* [34] to the $Si_{162}H_{103}$ cluster, with the $E_c - 0.17$ e transition of interstitial hydrogen chosen as the 0/+ marker, and in the absence of a H $-/0$ transition between similar structures, the $Ev - 0.1$ e transition of C_i [210, 243].

being 1.02 e^2/a.m.u. The 0/+ level of this defect is found to lie at $E_c - 0.21$ with the larger basis calculation, in good agreement with that of E_3, but it is noted at this stage that E_3 is observed to be > 50 K more stable than the observed vibrational complex. The question of why E_3 has been suggested to be due to a trigonal centre (if for a moment this structure is assumed to be responsible for E_3), can perhaps be resolved by the observation that the Kohn-Sham state of the highest level in the neutral charge state is only weakly perturbed from that of the trigonal bond-centred hydrogen (see figure 9.3).

$CH^{III_a^+}_{BC}$ is found to possess a hydrogen related stretch-mode at 1802.3 cm^{-1} which

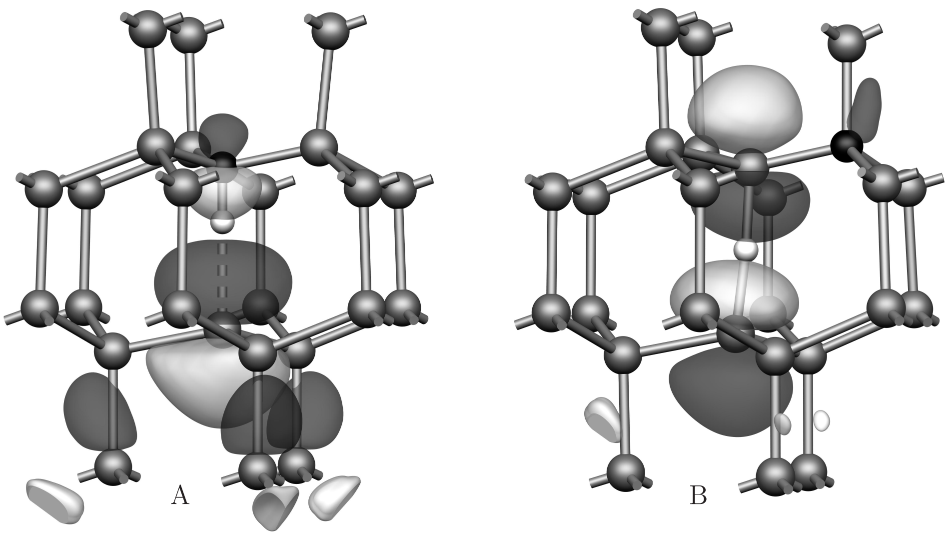


Figure 9.2: Isosurfaces of the highest occupied Kohn-Sham state associated with the neutral carbon–hydrogen defects (shown as black and white atoms respectively), for the A) C_{3v} nearest-neighbour ($\mathrm{CH}^{\mathrm{I}}_{BC}$) and B) C_{1h} next nearest-neighbour ($\mathrm{CH}^{\mathrm{II}}_{BC}$) pairs. The wavefunction shown for structure A is largely associated with a single silicon atom and the back-bonds of that atom, unlike the wavefunction of structure B which is similar to that for an isolated bond-centred hydrogen atom, but mildly perturbed by the nearby carbon atom. The dilated Si–H bond in the nearest-neighbour pair is shown as a dotted line.

shifts in the deuterated case to 1282.0 cm^{-1}. This defect also possesses three carbon-related modes at 652.9, 605.0 and 597.3 cm^{-1}, all of which have isotopic shifts similar to the observed modes at 595.5 and 660.5 cm^{-1}. The 0/+ transition of this centre is calculated to lie at $E_c - 0.27$, again similar to E_3. As with $\mathrm{CH}^{\mathrm{II}+}_{BC}$, this structure is found to be ~0.2 eV higher in energy than $\mathrm{CH}^{\mathrm{I}+}_{BC}$ in the cluster calculations.

$\mathrm{CH}^{\mathrm{III}^+_b}_{BC}$ is found to possess a hydrogen related stretch-mode at 1706.8, with three additional carbon-related modes at 661.2, 609.5, and 597.7 cm^{-1}. Again the DLTS level of this centre is similar to E_3.

9.4 Conclusions

It is suggested on the basis of these calculations that the structures responsible for the H_1 DLTS line is a hydrogen atom bound between a substitutional carbon atom and one of the neighbouring Si atoms, with the observed transition being of 0/+ character. E_3 is seen to be similar in properties to hydrogen bond-centred between two silicon atoms, which themselves are near to a substiutional carbon atom, but it has not been possible to unambiguously identify the structure responsible for E_3. The vibrational modes of the H_1 structure are shown to be unobserved due to their weak intensity, while those from the structures with hydrogen between two silicon atoms match the observed vibrational modes reasonably well in all cases, hence it is difficult to make unambiguous identification. It is suggested that all three vibrational modes can be explained by a complex of hydrogen between two silicon atoms near to a substitutional carbon, with perhaps the most likely candidate being the next-nearest neighbour site.

Chapter 10

Conclusions and further work

A general summary of the defects investigated is now presented, in addition to the specific concluding remarks at the end of each chapter. Possible directions for further work relevant to these defects is also discussed.

While some progress has been made towards understanding the behaviour of molecular hydrogen within silicon there are still a large number of unresolved questions about this system. The calculations presented in this thesis, as with the work of other groups, give a reasonable agreement between the theoretical vibrational modes and diffusion barriers and those observed for the molecule, accounting well for the downward shift in frequency of the stretch mode of the molecule due to the surrounding lattice. The binding between the molecule and interstitial oxygen is also simulated reasonably well, leading to close agreement with the observed stability of the complex. There are a number of problems with these calculations relating to the low reorientation energy of the molecule both at the isolated T_d site and bound to oxygen, leading to an inability to suggest definite candidates for the *two* complexes between O_i and the molecule and perhaps more troubling to produce agreement with the experimentally observed behaviour of the isolated molecule in the lattice.

The cause of such problems in simulating the molecule inside the lattice may lie in a variety of areas. The first possibility to consider is that the molecule is not "isolated" within the lattice, but is instead complexed with another defect, and perhaps simulations of hydrogen molecules near other impurities may shed some light on the nature of the unknown defect. Alternatively, since the rotational barrier for the molecule is estimated to be of the order of 0.2 eV, it is possible that the error in the calculations is simply too large to resolve this, but since a variety of methods have been applied by different groups all to the same effect is seems likely that this is not the problem. One common feature of all of the calculations so far is that the nuclear positions in the simulations have been treated classically as static parameters to be optimised by methods such as conjugate-gradient. It has been suggested that the dynamics of the molecule may be partly responsible for the difficulty in describing the system [244] and at least one classical molecular dynamics simulation is underway on this system to test this idea. Alternatively, the nuclear motions of the molecule within the crystal may require a quantum mechanical treatment, and while simulations methods of this type are perhaps less developed, a number of algorithms have been proposed for treating such systems, typically at large computational cost [245–247], but several cheaper methods which could be applied to this system without much additional effort beyond more conventional DFT simulations have recently appeared [248, 249].

The calculations on the molecule inside open lattice voids or platelets provides a general guide to the type of structure which would be responsible for the observed hydrogen species in a gas-like surroundings within silicon, but since the molecular environment is similar in both types of defect this does not provide a method of discrimination between the two types of structure. This coupled with the observation of a third molecular species [154] in an environment intermediate between the interstitial and open void cases also leads to the question of whether it is possible to make unambiguous assignments of structures from the molecular stretch modes, since it would appear as though almost any open lattice structure will trap hydrogen molecules, which then posses vibrational frequencies somewhere in the range between that of a molecule inside a dilated cage site near O_i at $\sim$3800 cm^{-1} to the open voids at $\sim$4150 cm^{-1}, particularly when it is realised that even the most accurate current calculations can have errors of the order of $\sim$100 cm^{-1} in the calculated stretch frequencies. This suggests that consideration of the vibrational modes of structures like Si–H which would line the inside of such voids is also needed for identifications to be made, but since such void-related stretch modes seem to fall into a range of around 200 cm^{-1} between about 2038 cm^{-1} for VH_1 [204] and 2222 cm^{-1} for VH_4, it would be difficult to perform such assignments purely from comparison with calculated modes. It is perhaps only in cases where there are large differences between the calculated and observed modes, as with the proposed IH_3 and IH_4 models for the 2166/2190 and 2222 cm^{-1} lines where one can be confident in using theory to make definite statements about observed vibrational modes.

A perhaps similar problem of identification is seen with the assignment of V_6 related complexes to some of the bound-exciton centres. While the proposed defects are certainly stable and have properties which are consistent with those observed, it is difficult to make a direct connection between theory and experiment, since the properties which can be easily calculated and those which can be readily measured do not overlap particularly. Again it is perhaps the vibrational modes of hydrogen at these centres which could be the strongest link between theory and experiment, and so it is perhaps by correlating the luminescent lines with other experimental techniques that this assignment can verified. Notwithstanding the perhaps tenuous identification of V_6, there are a large number of impurities other than hydrogen that should readily complex with this defect, and simulation of complexes between impurities such as oxygen or transition metals and V_6 appears to be a fruitful area of future work. This is perhaps of particular interest since as the hydrogen containing complexes are luminescent centres, this demonstrates that V_6 related complexes are capable of emitting light from silicon.

The suggested new hydrogen dimer discussed in this thesis perhaps raises questions about the dynamics and reactions of hydrogen at low temperatures in silicon. If this defect is readily formed at low temperatures in implanted material, before substantial quantities of VH_2 or IH_2 as is perhaps the case for H_2^*, then alternative mechanisms to bring two H_{BC}^+ together to form neutral dimers must exist. This suggests in turn that there should be precursor species to the dimers which have not yet been identified, though it is difficult to think of credible species which can form before VH_2 but contain two hydrogen atoms. The lack of IH_2 at low temperatures suggests that interstitial related reactions do not have significant impact on the behaviour of H_{BC}^+, perhaps in accordance with the recent suggestion that the AA19 EPR centre is due to I^{2+} [250] which would tend not to react with other species of the same charge.

The carbon-hydrogen complexes considered in this thesis may perhaps benefit from re-examination at a future date when more computer processing power or faster algorithms become available. Unambiguous resolution of the structure responsible for the observed vibrational modes would appear to require use of a larger basis in combination with cluster or supercell simulations with a larger number of silicon host atoms than are presently feasible. The question of the form of the defect responsible for E_3 is also somewhat vexing since its electrical properties are similar to those of the second and third shell carbon-hydrogen structures considered, but it has been measured to be trigonal in structure. Perhaps investigation of complexes containing impurity atoms in addition to carbon and hydrogen may provide a better candidate, but this will perhaps have to wait until the structure is examined by other symmetry sensitive techniques like EPR (since the neutral form of the defect is presumably a spin-half centre). It is possible to be more confident about the defect responsible for the H_1 DLTS centre, and this in turn can perhaps shed some light on the puzzle of whether there is interstitial hydrogen in diamond. If a similar situation holds for the C–H bonds in hydrogen containing defects in diamond [251], then hydrogen in this material will be difficult to observe by IR, but can perhaps instead be found by methods such as DLTS or EPR.

Bibliography

[1] B. Aspar, M. Bruel, M. Zussy, and A. M. Cartier, Electronics Lett. **32**(21), 1985 (1996).

[2] P. N. Keating, Phys. Rev. **145**, 637 (1966).

[3] J. C. Phillips, Phys. Rev. **166**, 832 (1968).

[4] A. Omeltchenko, J. Yu, R. K. Kalia, and P. Vashishta, Phys. Rev. Lett. **78**(11), 2148 (1997).

[5] S. Öberg, Private communication (1999), as a convergence test, dislocation properties properties discussed by Lehto and Öberg in *Phys. Rev. Lett.* 80, 5568 (1998), were simulated using 1200 atom LDF calculations.

[6] D. R. Bowler and M. J. Gillan, Computer Phys. Comm. **112**(2–3), 103 (1998).

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

[8] D. R. Hartree, Proc. of the Cambridge Philosophical Soc. **24**, 89 (1927).

[9] J. C. Slater, Phys. Rev. **35**, 210 (1930).

[10] W. Pauli, Z. Phys. **31**, 765 (1925).

[11] T. Koopmans, Physica **1**, 104 (1934).

[12] J. C. Slater, *The Self-Consistent Field for Molecules and Solids*, vol. IV (McGraw-Hill, New York, 1974).

[13] V. Fock, Z. Phys. **61**, 126 (1930).

[14] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

[15] J. A. Pople and R. K. Nesbet, J. Chem. Phys. **22**, 571 (1954).

[16] R. G. Parr, J. Chem. Phys. **20**, 239 (1952).

[17] J. A. Pople, J. Chem. Phys. **43**, 129 (1965).

[18] M. J. S. Dewar and W. Thiel, J. of the American Chemical Society **99**, 4907 (1977).

[19] T. A. Halgren and W. N. Lipscomb, J. Chem. Phys. **58**(4), 1569 (1973).

[20] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), international ed.

[21] R. Jones and P. R. Briddon, *The* ab initio *cluster method and the dynamics of defects in semiconductors* (Academic Press, Boston, 1998), vol. 51A of *Semiconductors and Semimetals*, chap. 6.

[22] V. A. Rassolov, J. A. Pople, and M. A. Ratner, Phys. Rev. B **59**, 15625 (1999).

[23] P. Löwdin, Advances in Chemical Physics **2**, 207 (1959).

[24] L. H. Thomas, Proc. of the Cambridge Philosophical Soc. **23**, 542 (1927).

[25] E. Fermi, Rendiconti Accademia dei Lincei **6**, 602 (1927).

[26] S. Lundqvist and N. H. March, eds., *Theory of the inhomogeneous electron gas*, Physics of solids and liquids (Plenum, New York, 1983).

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

[28] M. Levy, Proc. Nat. Acad. Sci., USA **76**, 6062 (1979).

[29] R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules* (Oxford University Press, Oxford, 1989).

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

[31] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).

[32] M. E. Casida, Phys. Rev. B **59**(7), 4694 (1999).

[33] M. K. Harbola, Phys. Rev. B **60**(7), 4545 (1999).

[34] A. Resende, R. Jones, S. Öberg, and P. R. Briddon, Phys. Rev. Lett. **82**(10), 2111 (March 1999).

[35] R. Jones, B. J. Coomer, J. P. Goss, B. Hourahine, and A. Resende, *Special defects in semiconducting materials* (Scitech Publications Ltd., Zuerich-Uetikon, Switzerland, 2000), vol. 71 of *Solid State Phenomena*, chap. The interaction of hydrogen with deep level defects in silicon, pp. 173–248.

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

[37] S. J. Vosko, L. Wilk, and M. Nusair, Chinese J. of Physics **58**, 1200 (1980).

[38] J. P. Perdew and A. Zunger, Phys. Rev. B **23**(10), 5048 (1981).

[39] J. P. Perdew and Y. Wang, Phys. Rev. B **45**(23), 13244 (1992).

[40] C. T. Lee, W. T. Yang, and R. G. Parr, Phys. Rev. B **37**(2), 785 (1988).

[41] A. D. Becke, J. Chem. Phys. **98**(2), 1372 (1993).

[42] J. M. L. Martin, Molecular Physics **86**(6), 1437 (1995).

[43] W. K. Leung, R. J. Needs, G. Rajagopal, S. Itoh, and S. Ihara, Phys. Rev. Lett. **83**(12), 2351 (1999).

[44] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[45] U. von Barth and L. Hedin, J. Phys. C Solid State **5**, 1629 (1972).

[46] J. C. Phillips and L. Kleinman, pr **116**, 287, 880 (1959).

[47] A. J. Zunger, Journal of Vacuum Sci. and Technol. **16**(5), 1337 (1979).

[48] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**(8), 4199 (1982).

[49] S. F. Boys, Proc. R. Soc. London, Ser. A **200**, 542 (1950).

[50] H. Sambe and R. H. J. Felton, J. Chem. Phys. **61**, 3862 (1975).

[51] R. Jones and A. Sayyash, J. Phys. C **19**(28), L653 (1986).

[52] P. R. Briddon, *Defects in Solids: an* Ab Initio *Study*, Ph.D. thesis, Department of Physics, University of Exeter, United Kingdom (1990).

[53] P. P. Ewald, Ann. Phys. **64**, 253 (1921).

[54] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **54**(20), 14362 (1996).

[55] H. Hellmann, *Einfunhring in die Quantenchemie* (Deuticke, Vienna, 1937).

[56] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[57] P. Pulay, *Applications of Electronic Structure Theory* (Plenum Press, New York, 1977).

[58] R. Jones, J. Goss, C. Ewels, and S. Öberg, Phys. Rev. B **50**(12), 8378 (1994).

[59] M. J. P. Musgrave and J. A. Pople, Proc. R. Soc. London, Ser. A **268**, 474 (1962).

[60] M. Stavola, ed., *Identification of Defects in Semiconductors*, vol. 51 of *Semiconductors and Semimetals* (Academic Press, San Diego, 1999).

[61] W. Kohn, *Shallow impurity states in silicon and germanium*, vol. 5 of *Solid State Physics* (Academic Press, New York, 1957).

[62] R. S. Leigh and B. Szigeti, Proc. R. Soc. London, Ser. A **301**, 211 (1967).

[63] B. Clerjaud and D. Côte, J. Phys. Cond. Matter **4**(49), 9919 (1992).

[64] J. Wagner, R. E. Pritchard, B. R. Davidson, R. C. Newman, T. J. Bullough, T. B. Joyce, C. Button, and J. S. Roberts, Semicond. Sci. Technol. **10**(5), 639 (1995).

[65] K. L. Shaklee and R. E. Nahory, Phys. Rev. Lett. **24**, 942 (1970).

[66] A. S. Kaminskii, E. V. Lavrov, V. A. Karasyuk, and M. L. W. Thewalt, Phys. Rev. B **51**(8), 4882 (1995).

[67] A. N. Safonov, *In preparation* (1999), .

[68] A. A. Kaplyanskii, V. N. Medvedev, and A. P. Skvortsov, Opt. Spectrosc. **29**, 905 (1970).

[69] A. A. Kaplyanskii, V. I. Kolyshkin, and A. P. Skvortsov, Sov. Phys. **12**, 2867 (1971).

[70] G. Davies, Semiconductors and Semimetals **51B**(PtB), 1 (1999).

[71] N. S. Minaev and A. V. Mudryi, Phys. Status Solidi A **68**, 561 (1981).

[72] R. Sauer and J. Weber, Physica B & C **116**(1–3), 195 (1983).

[73] B. J. Coomer, J. P. Goss, R. Jones, S. Öberg, and P. R. Briddon, Physica B **273–274**, 505 (1999).

[74] S. K. Estreicher, J. Weber, A. Derecskeikovacs, and D. S. Marynick, Phys. Rev. B **55**(8), 5037 (1997).

[75] G. Davies, E. C. Lightowlers, and Z. E. Ciechanowska, J. Phys. C **20**(2), 191 (1987).

[76] H. Feick and E. R. Weber, Physica B **273–274**, 497 (1999).

[77] R. N. Hall, Phys. Rev. **86**, 600 (1952).

[78] W. Shockley and W. T. Read Jr., Phys. Rev. **87**, 835 (1952).

[79] G. L. Miller, D. V. Lang, and L. C. Kimerling, Ann. Rev. Matter. Sci. **7**, 377 (1977).

[80] D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).

[81] D. V. Lang, Topics in Applied Physics **41**, 93 (1979).

[82] L. Dobaczewski, P. Kaczor, I. D. Hawkins, and A. R. Peaker, J. Appl. Phys. **76**(1), 194 (1994).

[83] K. Ikeda, H. Takaoka, and Y. Ishii, Jpn. J. Appl. Phys. Part 1 **24**(11), 1454 (1985).

[84] B. Brunwin, B. Hamilton, P. Jordan, and A. R. Peaker, Electronics Lett. **15**, 349 (1979).

[85] L. C. Allan, J. Chem. Phys. **40**, 3135 (1964).

[86] A. Abragam and M. H. L. Pryce, Proc. R. Soc. London, Ser. A **205**, 135 (1951).

[87] J. M. Spaeth, *Magneto-optical and electrical detection of paramagnetic resonance in semiconductors* (Academic Press, Boston, 1998), vol. 51A of *Semiconductors and Semimetals*, chap. 2.

[88] W. B. Mims, Proc. R. Soc. London **283**, 452 (1965).

[89] C. Gemperle and A. Schweiger, Chemical Reviews **91**(7), 1481 (1991).

[90] A. A. Kaplyanskii, Opt. Spectrosc. **16**, 329 (1964).

[91] A. E. Hughes and W. A. Runciman, Proc. R. Soc. London **90**, 827 (1967).

[92] G. Davies, E. C. Lightowlers, M. Stavola, K. Bergman, and B. Svensson, Phys. Rev. B **35**(6), 2755 (1987).

[93] A. A. Kaplyanskii, Opt. Spectrosc. **16**, 557 (1964).

[94] A. Mainwood and A. M. Stoneham, Physica B & C **116**(1–3), 101 (1983).

[95] J. W. Corbett, S. N. Sahu, T. S. Shi, and L. C. Snyder, Phys. Lett. A **93**(6), 303 (1983).

[96] A. Mainwood and A. M. Stoneham, J. Phys. C **17**(14), 2513 (1984).

[97] P. Deák, L. C. Snyder, and J. W. Corbett, Phys. Rev. B **37**(12), 6887 (1988).

[98] P. Deák, L. C. Snyder, J. L. Lindstrom, J. W. Corbett, S. J. Pearton, and A. J. Tavendale, Phys. Lett. A **126**(7), 427 (1988).

[99] K. J. Chang and D. J. Chadi, Phys. Rev. Lett. **60**(14), 1422 (1988).

[100] K. J. Chang and D. J. Chadi, Phys. Rev. B **40**(17), 11644 (1989).

[101] P. R. Briddon and R. Jones, Hyperfine Interactions **64**(1–4), 593 (1990).

[102] R. Jones, Physica B **170**(1–4), 181 (1991).

[103] D. M. Maric, M. A. Roberson, and S. K. Estreicher, Mater. Sci. Forum **143–147**, 1245 (1994).

[104] S. K. Estreicher, M. A. Roberson, and D. M. Maric, Phys. Rev. B **50**(23), 17018 (1994).

[105] C. G. Van de Walle, Physica B **170**(1–4), 21 (1991).

[106] C. G. Van de Walle, Y. Bar-Yam, and S. T. Pantelides, Phys. Rev. Lett. **60**(26), 2761 (1988).

[107] C. G. Van de Walle, Phys. Rev. B **49**(7), 4579 (1994).

[108] C. G. Van de Walle, P. J. H. Denteneer, Y. Bar-Yam, and S. T. Pantelides, Phys. Rev. B **39**(15), 10791 (1989).

[109] C. G. Van de Walle, Phys. Rev. Lett. **80**(10), 2177 (1998).

[110] C. G. Van de Walle, Phys. Rev. B **58**(3), 1689 (1998).

[111] L. D. Landau and E. M. Lifshitz, *Quantum mechanics* (Pergamon, Oxford, 1977), 3rd ed.

[112] J. K. Nørskov, Phys. Rev. B **20**(2), 446 (1979).

[113] E. Kim, K. H. Lee, H. J. Lee, and Y. H. Lee, J. Phys. Cond. Matter **4**(30), 6443 (1992).

[114] K. G. Nakamura, K. Ishioka, M. Kitajima, A. Endou, M. Kubo, and A. Miyamoto, J. Chem. Phys. **108**(8), 3222 (1998).

[115] K. G. Nakamura, K. Ishioka, M. Kitajima, and K. Murakami, Solid State Commun. **101**(10), 735 (1997).

[116] K. Ishioka, K. G. Nakamura, M. Kitajima, N. Fukata, K. Murakami, S. Fujimura, and J. Kikuchi, Mater. Sci. Forum **258–263**(Pt1–3), 235 (1997).

[117] K. G. Nakamura and M. Kitajima, Jpn. J. Appl. Phys. Part 1 **36**(4A), 2004 (1997).

[118] K. G. Nakamura, K. Ishioka, and M. Kitajima, Jpn. J. Appl. Phys. Part 2 **36**(11B), L1479 (1997).

[119] Y. Okamoto, M. Saito, and A. Oshiyama, Phys. Rev. B **56**(16), 10016 (1997).

[120] Y. Okamoto, M. Saito, and A. Oshiyama, Phys. Rev. B **58**(12), 7701 (1998).

[121] Y. S. Kim, Y. G. Jin, J. W. Jeong, and K. J. Chang, Semicond. Sci. Technol. **14**(12), 1042 (1999).

[122] Y. S. Kim, Y. G. Jin, J. W. Jeong, and K. J. Chang, Physica B **273–274**, 231 (1999).

[123] V. P. Markevich, M. Suezawa, and K. Sumino, Mater. Sci. Forum **196–201**, 915 (1995).

[124] V. P. Markevich, I. F. Medvedeva, and L. I. Murin, in R. Jones, ed., *Early stages of oxygen precipitation in silicon* (Kluwer Accademic Publishers, Dordrecht, 1996), vol. 17 of *NATO ASI – 3. High Technology*, p. 103.

[125] V. P. Markevich and M. Suezawa, J. Appl. Phys. **83**(6), 2988 (1998).

[126] B. Bech Nielsen, K. Tanderup, M. Budde, K. Bond Nielsen, J. L. Lindstrom, R. Jones, S. Öberg, B. Hourahine, and P. R. Briddon, Mater. Sci. Forum **258–263**(Pt1–3), 391 (1997).

[127] R. E. Pritchard, M. J. Ashwin, R. C. Newman, J. H. Tucker, E. C. Lightowlers, M. J. Binns, R. Falster, and S. A. McQuaid, Mater. Sci. Forum **258–263**(Pt1–3), 283 (1997).

[128] A. Van Weiringen and N. Warmholz, Physica **22**, 849 (1956).

[129] N. Fukata, S. Sasaki, K. Murakami, K. Ishioka, M. Kitajima, S. Fujimura, and J. Kikuchi, Jpn. J. Appl. Phys. Part 2 **35**(8B), L1069 (1996).

[130] K. Murakami, N. Fukata, S. Sasaki, K. Ishioka, M. Kitajima, S. Fujimura, J. Kikuchi, and H. Haneda, Phys. Rev. Lett. **77**(15), 3161 (1996).

[131] M. Kitajima, K. Ishioka, K. G. Nakamura, N. Fukata, K. Murakami, J. Kikuchi, and S. Fujimura, Mater. Sci. Forum **258–263**(Pt1–3), 203 (1997).

[132] N. Fukata, K. Murakami, K. Ishioka, K. G. Nakamura, M. Kitajima, S. Fujimura, J. Kikuchi, and H. Haneda, Mater. Sci. Forum **258–263**(Pt1–3), 211 (1997).

[133] K. Ishioka, K. G. Nakamura, M. Kitajima, N. Fukata, S. Sasaki, K. Murakami, S. Fujimura, J. Kikuchi, and H. Haneda, Appl. Surf. Sci. **117**, 37 (1997).

[134] K. G. Nakamura, K. Ishioka, M. Kitajima, N. Fukata, K. Murakami, A. Endou, M. Kubo, and A. Miyamoto, Appl. Surf. Sci. **132**, 243 (1998).

[135] N. Fukata, S. Sasaki, K. Murakami, K. Ishioka, K. G. Nakamura, M. Kitajima, S. Fujimura, J. Kikuchi, and H. Haneda, Phys. Rev. B **56**(11), 6642 (1997).

[136] M. K. Weldon, M. Collot, Y. J. Chabal, V. C. Venezia, A. Agarwal, T. E. Haynes, D. J. Eaglesham, S. B. Christman, and E. E. Chaban, Appl. Phys. Lett. **73**(25), 3721 (1998).

[137] A. W. R. Leitch, V. Alex, and J. Weber, Mater. Sci. Forum **258–263**(Pt1–3), 241 (1997).

[138] A. W. R. Leitch, V. Alex, and J. Weber, Solid State Commun. **105**(4), 215 (1998).

[139] A. W. R. Leitch, V. Alex, and J. Weber, Phys. Rev. Lett. **81**(2), 421 (1998).

[140] J. Vetterhöffer, J. Wagner, and J. Weber, Phys. Rev. Lett. **77**(27), 5409 (1996).

[141] A. W. R. Leitch and J. Weber, Phys. Rev. B **60**(19), 13265 (1999).

[142] W. Beyer, Physica B **170**(1–4), 105 (1991).

[143] R. E. Pritchard, M. J. Ashwin, J. H. Tucker, and R. C. Newman, Phys. Rev. B **57**(24), 15048 (1998).

[144] R. E. Pritchard, M. J. Ashwin, J. H. Tucker, R. C. Newman, E. C. Lightowlers, M. J. Binns, S. A. McQuaid, and R. Falster, Phys. Rev. B **56**(20), 13118 (1997).

[145] S. A. McQuaid, M. J. Binns, C. A. Londos, J. H. Tucker, A. R. Brown, and R. C. Newman, J. Appl. Phys. **77**(4), 1427 (1995).

[146] M. J. Binns, S. A. McQuaid, R. C. Newman, and E. C. Lightowlers, Semicond. Sci. Technol. **8**(10), 1908 (1993).

[147] J. A. Zhou and M. Stavola, Phys. Rev. Lett. **83**(7), 1351 (1999).

[148] J. A. Zhou, E. Chen, and M. Stavola, Physica B **273–274**, 200 (1999).

[149] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV* (Van Nostrand Reinhold, New York, 1979).

[150] S. J. Breuer, R. Jones, P. R. Briddon, and S. Öberg, Phys. Rev. B **53**(24), 16289 (1996).

[151] S. Muto, S. Takeda, and M. Hirata, Phil. Mag. A **72**(4), 1057 (1995).

[152] G. Davies, Physics Reports – Review Section of Physics Letters **176**(3–4), 83 (1989).

[153] J. Weber, private communication (1998).

[154] M. Kitajima, K. Ishioka, K. Nakanoya, S. Tateishi, T. Mori, N. Fukata, K. Murakami, and S. Hishita, Jpn. J. Appl. Phys. Part 2 **38**(7A), L691 (1999).

[155] W. Pauling and E. Wilson, *Introduction to Quantum Mechanics* (McGraw Hill, New York, 1935).

[156] T. E. Stern, Proc. R. Soc. London, Ser. A **130**, 551 (1931).

[157] R. Jones, private communication (1999).

[158] J. B. Boyce, N. M. Johnson, S. E. Ready, and J. Walker, Phys. Rev. B **46**(7), 4308 (1992).

[159] P. Deák, C. R. Ortiz, L. C. Snyder, and J. W. Corbett, Physica B **170**(1–4), 223 (1991).

[160] N. M. Johnson, F. A. Ponce, R. A. Street, and R. J. Nemanich, Phys. Rev. B **35**(8), 4166 (1987).

[161] G. D. Watkins, *Deep centers in semiconductors* (Gordon and Breach Science Publishers, Switzerland, 1992), chap. 3, 2nd ed.

[162] G. D. Watkins and J. W. Corbett, Phys. Rev. **138**(2A), A543 (1965).

[163] Y. H. Lee and J. W. Corbett, Phys. Rev. B **9**(10), 4351 (1974).

[164] D. J. Chadi and K. J. Chang, Phys. Rev. B **38**(2), 1523 (1988).

[165] S. K. Estreicher, J. L. Hastings, and P. A. Fedders, Appl. Phys. Lett. **70**(4), 432 (1997).

[166] J. L. Hastings, S. K. Estreicher, and P. A. Fedders, Phys. Rev. B **56**(16), 10215 (1997).

[167] S. K. Estreicher, Phys. Rev. B **60**(8), 5375 (1999).

[168] A. S. Kaminskii, B. M. Leiferov, and A. N. Safonov, Sov. Phys. Solid State **29**, 551 (1987).

[169] A. N. Safonov and E. C. Lightowlers, Mater. Sci. Eng. B **36**(1–3), 251 (1996).

[170] A. N. Safonov and E. C. Lightowlers, Mater. Sci. Eng. B **58**(1–2), 39 (1999).

[171] A. S. Kaminskii, E. V. Lavrov, V. A. Karasyuk, and M. L. W. Thewalt, Phys. Rev. B **50**(11), 7338 (1994).

[172] A. S. Kaminskii, E. V. Lavrov, G. Davies, E. C. Lightowlers, and A. N. Safonov, Semicond. Sci. Technol. **11**(12), 1796 (1996).

[173] A. N. Safonov, E. C. Lightowlers, and G. Davies, Phys. Rev. B **56**(24), 15517 (1997).

[174] A. S. Kaminskii and E. V. Lavrov, Solid State Commun. **106**(11), 751 (1998).

[175] P. R. Briddon, R. Jones, and G. M. S. Lister, J. Phys. C **21**(30), L1027 (1988).

[176] K. J. Chang and D. J. Chadi, Phys. Rev. Lett. **62**(8), 937 (1989).

[177] J. D. Holbech, B. Bech Nielsen, R. Jones, P. Sitch, and S. Öberg, Phys. Rev. Lett. **71**(6), 875 (1993).

[178] H. Takahashi, N. Fukata, M. Suezawa, and H. Yamadakaneta, Phys. Status Solidi B **210**(2), 581 (1998).

[179] M. Suezawa, Jpn. J. Appl. Phys. Part 2 **38**(5A), L484 (1999).

[180] M. Budde, *Hydrogen-related defects in proton implanted silicon and germanium*, Ph.D. thesis, Aarhus Center for Atomic Physics, University of Aarhus, Denmark (October 1998).

[181] G. D. Watkins, in *Lattice Defects in Semiconductors*, Inst. Phys. Conf., London (1975), Ser. 23, p. 1.

[182] M. Budde, B. Bech Nielsen, P. Leary, J. Goss, R. Jones, P. R. Briddon, S. Öberg, and S. J. Breuer, Phys. Rev. B **57**(8), 4397 (1998).

[183] S. K. Estreicher, J. L. Hastings, and P. A. Fedders, Phys. Rev. B **57**(20), 12663 (1998).

[184] K. Bond Nielsen and B. Bech Nielsen, Mater. Sci. Eng. B **58**(1–2), 163 (1999).

[185] P. Stallinga and B. Bech Nielsen, Acta Phys. Pol. A **92**(5), 989 (1997).

[186] T. S. Shi, G. R. Bai, M. W. Qi, and J. K. Zhou, Mater. Sci. Forum **10–12**, 597 (1986).

[187] L. M. Xie, M. W. Qi, and J. M. Chen, J. Phys. Cond. Matter **3**(44), 8519 (1991).

[188] Y. C. Du, Y. F. Zhang, X. T. Meng, and H. Y. Shen, Scientia Sinica A **30**(2), 176 (1987).

[189] B. Bech Nielsen, Private communication (1999).

[190] S. K. Estreicher, J. L. Hastings, and P. A. Fedders, Phys. Rev. Lett. **82**(4), 815 (1999).

[191] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**(12), 5188 (1976).

[192] P. M. Morse, Phys. Rev. **34**, 57 (1929).

[193] B. Holm, K. Bond Nielsen, and B. Bech Nielsen, Phys. Rev. Lett. **66**(18), 2360 (1991).

[194] C. P. Ewels, S. Leoni, M. I. Heggie, P. Jemmer, E. Hernández, R. Jones, and P. R. Briddon, Phys. Rev. Lett. **84**(4), 690 (2000).

[195] B. Aradi and P. Deák, *Acceptor-like hydrogen dimers in silicon*, Hydrogen '99 workshop Exeter (April 1999).

[196] G. R. Bai, M. W. Qi, L. M. Xie, and T. S. Shi, Solid State Commun. **56**(3), 277 (1985).

[197] H. J. Stein, J. Electr. Mat. **4**, 159 (1975).

[198] C. H. Tindal, J. W. Straley, and H. H. Nielsen, Phys. Rev. **62**(3–4), 151 (1942).

[199] H. J. Stein, Phys. Rev. Lett. **43**, 1030 (1979).

[200] T. S. Shi, L. M. Xie, G. R. Bai, and M. W. Qi, Phys. Status Solidi B **131**(2), 511 (1985).

[201] B. Bech Nielsen, L. Hoffmann, M. Budde, R. Jones, J. Goss, and S. Öberg, Mater. Sci. Forum **196–201**, 933 (1995).

[202] Y. K. Park, S. K. Estreicher, C. W. Myles, and P. A. Fedders, Phys. Rev. B **52**(3), 1718 (1995).

[203] B. Bech Nielsen, L. Hoffmann, and M. Budde, Mater. Sci. Eng. B **36**(1–3), 259 (1996).

[204] P. Stallinga, P. Johannesen, S. Herstrøm, K. Bond Nielsen, B. Bech Nielsen, and J. R. Byberg, Phys. Rev. B **58**(7), 3842 (1998).

[205] C. G. Van de Walle and J. Neugebauer, Phys. Rev. B **52**(20), 14320 (1995).

[206] M. Suezawa, Mater. Sci. Forum **258–263**(Pt1–3), 229 (1997).

[207] M. Suezawa, Jpn. J. Appl. Phys. Part 2 **37**(3A), L259 (1998).

[208] T. Abe, H. Harada, and J. Chikawa, Physica B & C **116**(1–3), 139 (1983).

[209] P. A. Stolk, m. Grossman, D. J. Eaglesham, D. C. Jacobson, C. S. Rafferty, G. H. Gilmer, M. Jaraíz, and J. M. Poate, J. Appl. Phys. **81**(9), 6031 (1997).

[210] P. Leary, R. Jones, S. Öberg, and V. J. B. Torres, Phys. Rev. B **55**(4), 2188 (1997).

[211] B. G. Svensson, J. Svensson, J. L. Lindstrom, G. Davies, and J. W. Corbett, Appl. Phys. Lett. **51**(26), 2257 (1987).

[212] B. G. Svensson and J. L. Lindström, J. Appl. Phys. **72**(12), 5616 (1992).

[213] M. Suezawa, Jpn. J. Appl. Phys. Part 2 **37**(7A), L806 (1998).

[214] M. Suezawa, J. Appl. Phys. **83**(4), 1958 (1998).

[215] L. J. Cheng, J. C. Corelli, J. W. Corbett, and G. Watkins, Phys. Rev. **152**, 761 (1966).

[216] J. H. Meal and M. K. Wilson, J. Chem. Phys. **24**(2), 385 (1956).

[217] R. C. Newman, Mater. Sci. Eng. B **36**(1–3), 1 (1996).

[218] A. Endrös, Phys. Rev. Lett. **63**(1), 70 (1989).

[219] A. N. Safonov, E. C. Lightowlers, G. Davies, P. Leary, R. Jones, and S. Öberg, Phys. Rev. Lett. **77**(23), 4812 (1996).

[220] A. L. Endrös, W. Krühler, and F. Koch, J. Appl. Phys. **72**(6), 2264 (1992).

[221] K. Bond Nielsen, Private communication (1999).

[222] W. Csaszar and A. L. Endrös, Phys. Rev. Lett. **73**(2), 312 (1994).

[223] Y. Kamiura, N. Ishiga, and Y. Yamashita, Jpn. J. Appl. Phys. Part 2 **36**(11A), L1419 (1997).

[224] Y. Kamiura, M. Yoneta, Y. Nishiyama, and F. Hashimoto, J. Appl. Phys. **72**(8), 3394 (1992).

[225] Y. Kamiura, M. Yoneta, and F. Hashimoto, Appl. Phys. Lett. **59**(24), 3165 (1991).

[226] M. Yoneta, Y. Kamiura, and F. Hashimoto, J. Appl. Phys. **70**(3), 1295 (1991).

[227] Y. Kamiura, M. Tsutsue, M. Hayashi, Y. Yamashita, and F. Hashimoto, Mater. Sci. Forum **196–201**, 903 (1995).

[228] Y. Kamiura, T. Okashita, and F. Hashimoto, Mater. Sci. Forum **143–147**, 921 (1994).

[229] Y. Kamiura, M. Hayashi, Y. Nishiyama, S. Ohyama, and Y. Yamashita, Jpn. J. Appl. Phys. Part 1 **36**(11), 6579 (1997).

[230] Y. Kamiura, M. Tsutsue, Y. Yamashita, F. Hashimoto, and K. Okuno, J. Appl. Phys. **78**(7), 4478 (1995).

[231] R. C. Newman and J. B. Willis, J. Phys. Chem. Solids **26**, 373 (1965).

[232] S. P. Chappell and R. C. Newman, Semicond. Sci. Technol. **2**(10), 691 (1987).

[233] L. Hoffmann, E. V. Lavrov, and B. Bech Nielsen, Mater. Sci. Eng. B **58**(1–2), 167 (1999).

[234] L. Hoffmann, *Private communication* (1999), .

[235] D. M. Maric, P. F. Meier, and S. K. Estreicher, Mater. Sci. Forum **83–87**, 119 (1992).

[236] D. M. Maric, P. F. Meier, and S. K. Estreicher, Phys. Rev. B **47**(7), 3620 (1993).

[237] C. Kaneta and H. Katayama-Yoshida, Mater. Sci. Forum **196–201**, 897 (1995).

[238] C. Kaneta and H. Katayama-Yoshida, Solid State Commun. **93**(5), 460 (1995).

[239] Y. Zhou, R. Luchsinger, P. F. Meier, H. U. Suter, D. Maric, and S. K. Estreicher, Mater. Sci. Forum **196–201**, 891 (1995).

[240] P. Leary, R. Jones, and S. Öberg, Phys. Rev. B **57**(7), 3887 (1998).

[241] L. W. Song and G. D. Watkins, Phys. Rev. B **42**(9), 5759 (1990).

[242] R. Jones, private communication (1999).

[243] L. W. Song, X. D. Zhan, B. W. Benson, and G. D. Watkins, Phys. Rev. B **42**(9), 5765 (1990).

[244] S. K. Estreicher, Private communication (1999).

[245] C. P. Herrero and R. Ramirez, Phys. Rev. B **51**(23), 16761 (1995).

[246] K. M. Forsythe and N. Makri, J. Chem. Phys. **110**(12), 6082 (1999).

[247] T. Miyake, T. Ogitsu, and S. Tsuneyuki, Phys. Rev. Lett. **81**(9), 1873 (1998).

[248] S. Reich, J. of Computational Phys. **151**(1), 49 (1999).

[249] F. A. Bornemann, P. Nettesheim, and C. Schutte, J. Chem. Phys. **105**(3), 1074 (1996).

[250] B. N. Mukashev, K. A. Abdullin, Y. V. Gorelkinskii, and S. Z. Tokmoldin, Mater. Sci. Eng. B **58**(1–2), 171 (1999).

[251] J. P. Goss, Private communication (1999).