Density Functional Modelling of Point Defects in Semiconductors

Christopher Paul Ewels

Submitted by Christopher Paul Ewels to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Theoretical Physics, in the Faculty of Physics, July 1997.

Declaration

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.

C. P. Ewels

July, 1997

Abstract

An *ab initio* local density functional cluster method, AIMPRO, is used to examine oxygen related point defects in silicon, and H-related complexes in III-V semiconductors, notably InP.

Models are proposed for three different types of shallow thermal donor in silicon, N_iO_{2i}, $(CH)_iO_{4i}$ and Al_sO_{4i}. This includes a new mechanism for converting otherwise deep level defects into shallow donor level centres, through electrostatic compression via neighbouring oxygen atoms. A reconstruction mechanism is also found whereby $(CH)_iO_{4i}$ can transform into a deep state defect.

Several thermal donor models are examined, and a centre consisting of a 'di-y-lid', O_{4i}, is shown to account for most of the observed experimental properties. Higher order thermal donor formation is discussed, as well as the role of hydrogen and silicon self-interstitials in thermal donor behaviour.

Rapid oxygen diffusion is examined in the context of the oxygen dimer, O_{2i}. A puckered dimer structure is shown to be stable with ^{16}O modes in good agreement with experiment. A low energy migration path for dimer diffusion is determined. The role of the dimer in creating other oxygen-based point defects in silicon is discussed.

Oxygen complexes with nitrogen are also modelled, and the most common N/O defect is shown to be $N_{2i}O_i$, consisting of bond centred oxygen neighbouring an interstitial nitrogen square. A N_iO_i complex is also identified and correlated with experimental data. Various vacancy–oxygen complexes are studied and their anomalous formation discussed. The structure VO_2 is unambiguously assigned to experimental infra-red absorption at 889 cm^{-1}.

Finally the interaction between hydrogen, Group-II elements and vacancies in III-V materials is examined, particularly InP. VH_4 in InP is shown to be a single shallow donor, responsible for Fe charge compensation observed in InP:Fe-H. Trends in structure with varying Group-II element and III-V material are examined.

Acknowledgements

I would like to thank both Exeter Physics Department and the EPSRC for providing resources and funding for my research assistantship throughout this PhD.

There are many people to thank for their support and encouragement, without whom this thesis would not have been possible.

Firstly my supervisor, Dr. Bob Jones, whose enthusiastic cry of 'Any Progress' will haunt me for the rest of my days – his unquenchable curiosity and love for the subject are probably the most valuable lessons I have learned from this PhD, and his continual support and encouragement have kept me going over the last three or so years.

Thanks also to Malcolm for my brief stay in Sussex, and all the other AIMPRO group members, particularly Jon Goss for loaning me 'Burn the World' and putting me off Roy Harper for life and Paul Leary for his 'Spudman' caricatures. Also Antonio, Ben, Bernd, Sven, Chris, Steve, John, Joachim, Gerd and others - I hope I'll see you all down at the New Inn some time! Also to the many collaborators I've worked with, particularly Jo Miro, Tomas Hallberg, Frank Berg Rasmussen, Brian Bech Nielsen, Bernard Pajot and Peter Deák, whose suggestions and/or guidance around their local pubs was always much appreciated.

Next there's all the friends I've made, the list is too long to mention but their friendship, drinking or dancing skills helped to make my stay in Exeter a really happy one, and hopefully I'll be able to make use of their floorspace from time to time in years to come. Thanks also to Julia for some great times and a close race to finishing our theses, one bet that I may have just won!

Finally my family, who have been great over the years and never raised an eyebrow when I claimed my thesis would be finished in the 'next two weeks' for nearly a year...

Think no more, lad; laugh, be jolly:
    Why should men make haste to die?
Empty heads and tongues a-talking
Make the rough road easy walking,
And the feather pate of folly
    Bears the falling sky.

Oh, 'tis jesting, dancing, drinking
    Spins the heavy world around.
If young hearts were not so clever,
Oh, they would be young for ever:
Think no more; 'tis only thinking
    Lays lads underground.

A. E. Housman
Excerpt from *A Shropshire Lad*, 1896

Contents

List of Published Papers

Listed in reverse chronological order:

1. 'Molecular rectification with M—DσA LB film—M junctions', A. S. Martin, A. C. Brady, B. Hodder, J. Roy Sambles, C. P. Ewels, R. Jones, P. R. Briddon, A. M. Musa, C. A. Panetta, D. L. Mattern, Science, To be submitted, 1997.
2. 'The carbon interstitial in graphite – structure and mobility', B. R. Eggen, M. I. Heggie, C. P. Ewels, G. Jungnickel, R. Jones, P. R. Briddon, Phys Rev B, To be submitted, 1997.
3. 'Shallow Thermal Donor defects in silicon' C. P. Ewels, R. Jones, S. Öberg, J. Miro, and P. Deák, Phys. Rev. Lett. **77** (5), 865-868 (1996).
4. 'Is hydrogen anti bonded in hydrogenated GaAs:Mg?' R. Bouanani-Rabi, B. Pajot, C. P. Ewels, S. Öberg, J. Goss, R. Jones, Y. Nissim, B. Theys, and C. Blaauw. To be published in the proceedings of SLCS96, 'Shallow Level Centers in Semiconductors', Amsterdam, 1996.
5. 'Shallow Thermal Donor Defects in Silicon', C. P. Ewels, R. Jones, S. Öberg, J. Miro, P. Deák. To be published in the proceedings of SLCS96, 'Shallow Level Centers in Semiconductors', Amsterdam, 1996.
6. 'Theoretical studies on nitrogen-oxygen complexes in silicon', P. Deák, J. Miro, R. Jones and C. P. Ewels, J. Phys. C : Cond. Matter **8** (41), 7711-7722 (1996).
7. 'Oxygen-Carbon, Oxygen-nitrogen and oxygen-dimer defects in Si', C. P. Ewels, R. Jones and S. Öberg, Oral Presentation and NATO ARW 'Early Stages of Oxygen Precipitation in Silicon', ed. R. Jones, Kluwer Academic Press, p. 141.
8. 'The nitrogen-pair oxygen defect in silicon', F. Berg Rasmussen, S. Öberg, R. Jones, C. Ewels, J. Goss, J. Miro, and P. Deák, NATO ARW 'Early Stages of Oxygen Precipitation in Silicon', ed. R. Jones, Kluwer Academic Press, p. 319.
9. 'The nitrogen-pair oxygen defect in Silicon', F. Berg Rasmussen, S. Öberg, R. Jones, C. Ewels, J. Goss, J. Miro, and P. Deák, Mat. Sci. Eng. B, **36**, 91-95 (1996).

10. 'Vacancy- and Acceptor- H complexes in InP', C. P. Ewels, S. Öberg, R. Jones, B. Pajot and P. R. Briddon, Semicond. Sci. Tech. **11**, 1-6 (1996).
11. 'Vielkopfige Hydra - Neue Medien verandern die wissenschaftliche Kommunikation' - 'Many-headed Hydra - The new media and scientific communication', B. R. Eggen, C. P. Ewels, Zeitschrift für Kulturaustausch, p.550, (1995), IFA, Germany.
12. 'The NNO defect in Silicon', F. Berg Rasmussen, S. Öberg, R. Jones, C. Ewels, J. Goss, J. Miro, and P. Deák, Mat. Sci. Forum Vol. **196-201**, pp 791-796 (1995), Trans Tech Publications, Switzerland.
13. 'A first principles investigation of vacancy-oxygen defects in Si', C. P. Ewels, R. Jones, S. Öberg, Mat. Sci. Forum Vol. **196-201**, pp (1995), Trans Tech Publications, Switzerland.
14. 'Vacancy- and acceptor- H complexes in InP', C. P. Ewels, S. Öberg, R. Jones, B. Pajot, P. R. Briddon, Mat. Sci. Forum Vol. **196-201**, pp (1995), Trans Tech Publications, Switzerland.
15. 'The nitrogen-pair oxygen defect in silicon', F. Berg Rasmussen, S. Öberg, R. Jones, C. Ewels, J. Goss, J. Miro, and P. Deák, E-MRS, Strasbourg, (1995).
16. 'Photoinduced changes of hydrogen-bonding in semi-insulating iron-doped InP', B. Pajot, C. Y. Song, R. Darwich, F. Gendron, C. P. Ewels, Sol. State Comm. **95** 12, 851-854 (1995).
17. 'Theoretical and Isotopic Infrared Absorption Investigations of Nitrogen-Oxygen Defects in Silicon', R. Jones, C. Ewels, J. Goss, J. Miro, P. Deák, S. Öberg, F. Berg Rasmussen, Semicond. Sci. Tech. **9**, 2145-48 (1994).
18. 'H passivated defects in InP', C. P. Ewels, S. Öberg, P. R. Briddon, J. Goss, R. Jones, S. Breuer, R. Darwich, B. Pajot, Sol. State Comm. **93** 5, pp.459-460 (1995).
19. 'The Hydrogen complexes in GaAs and InP doped with Magnesium', R. Rahbi, B. Pajot, C. P. Ewels, S. Öberg, J. Goss, R. Jones, Y. Nissim, B. Theys, C. Blaauw, Sol. State Comm. **93** 5, pp.462 (1995).
20. 'Ab Initio calculations of anharmonicity of the C-H stretch mode in HCN and GaAs', R. Jones, J. Goss, C. Ewels, S. Öberg, Phys. Rev. B, **50**, 8378-88 (1994).

List of Tables

List of Figures

Definitions and Notation

Abbreviation	Definition
AIMPRO	*Ab Initio* Modelling PROgram
CNDO	Complete Neglect of Differential Overlap
DFT	Density Functional Theory
DLTS	Deep Level Transient Spectroscopy
EMT	Effective Mass Theory
ENDOR	Electron Nuclear DOuble Resonance spectroscopy
EPR	Electron Paramagnetic Resonance
FTIR	Fourier Transform Infra-Red spectroscopy
LDA	Local Density Approximation
LVM	Local Vibrational Mode
MINDO/3	Modified Intermediate Neglect of Differential Overlap
PL	Photoluminescence
PTIS	Photo-Thermal Ionisation Spectroscopy
STD	Shallow Thermal Donor
TD	Thermal Donor (also TDD, Thermal Double Donor)

Certain standard notations have been adopted when labelling defects. Defects identified in the format NLx, where x is numeric, have been detected using EPR and ENDOR[1]. Defects detected using PL are referred to by the single letter label given to their zero phonon line, e.g. 'P-line' defects. There is no standard notation for defects detected using FTIR. Note that the same defect may be known under different names depending on the method used to detect it. Detailed correlation experiments are required to determine this. Equally, since defect signals are assigned a given notation, it may be that the same defect in different charge or spin states may have been assigned different names.

All equations are in atomic units unless stated, *i.e.* $\hbar$, e, m_e and $4\pi\varepsilon_0$ are all set to unity. The results chapters use Angstroms (Å) for length and electron volts (eV) for energy since these are generally used in the field; 1 a.u. of energy is 27.212 eV, 1 a.u. of length is 0.529 Å.

[1] 'NL' stands for the Netherlands, where one of the major ENDOR groups is based

Chapter 1

Introduction

"From the moment I picked up your book until I laid it down, I was convulsed with laughter. Some day I intend reading it."

Groucho Marx

The Silicon industry is a 10^{12} dollar industry worldwide, with Si technology lying at the core of the majority of computer systems in use today. Si is the focus of huge amounts of research, both industrially and academically. Despite this, there is still much that is not understood about the behaviour of impurities in Si, and a thorough understanding of the structure and properties of impurities is becoming increasingly important as devices become smaller and chip speeds are expected to increase.

Commercial Si is normally grown using either float zone methods or the Czochralski[1] growth technique [1]. Czochralski growth forms the vast majority due to the resultant wafers resistance to thermal stress, the speed of production and low cost. A schematic diagram of a Czochralski-Si grower is shown in Figure 1.1. The molten Si is contained in a quartz crucible which is rotated. A single crystal Si seed is placed on the surface and gradually drawn upwards and simultaneously rotated; this draws the molten Si after it which solidifies into a continuous crystal extending from the seed.

Due to the high reactivity of molten Si there are only a few possible choices for crucible, and SiO_2 is the normal choice. However the SiO_2 gradually dissolves into the melt, releasing large quantities of oxygen into the Si. Over 99% of this is lost as SiO gas from the molten surface, but the rest dissolves into the single crystal Si. Thus Cz-Si has an extremely high oxygen concentration (typically of the order of 10^{18} atoms/cm^3, or 1 ppm), and oxygen is always the highest concentration

[1] Pronounced choh-KRAL-skee

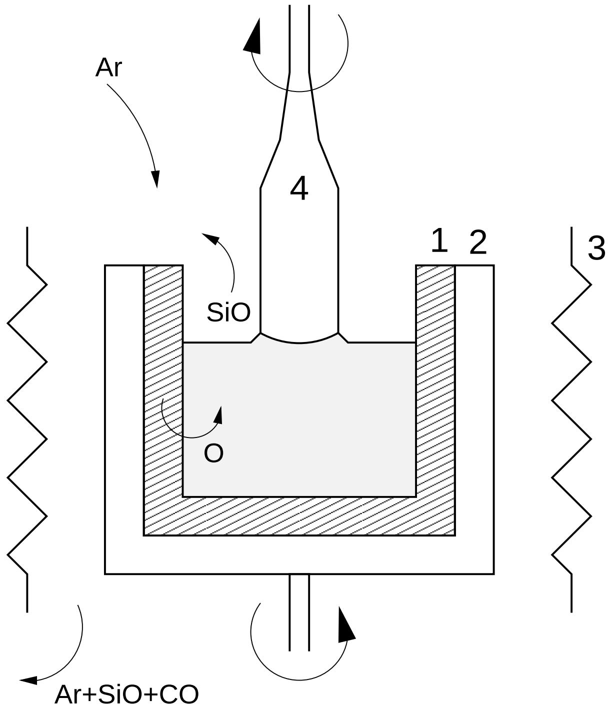


Figure 1.1: Schematic diagram showing Czochralski Si growth. (1) SiO_2 crucible, (2) Carbon susceptor, (3) Graphite Heater, (4) Single crystal Si. Note that the crystal and crucible are rotated in opposite directions. More than 99% of dissolving oxygen is lost as SiO.

impurity in Cz-Si[2]. The solid solubility of O in Si is $\sim 10^{18}$ atoms/cm^3 at melting point [3] but drops by several orders of magnitude at room temperature [4], so there is a driving force for precipitation. For extended anneals above ~450°C, small quartz precipitates form in the Si, accompanied by silicon self-interstitials (Si_i) and dislocation loops due to the volume difference of Si and SiO_2 (quartz has roughly half the Si density of pure Si). This glut of dislocations causes a large degradation in mechanical strength of the silicon (the yield point drops by a factor of almost 5 [5]). Similarly, oxygen causes warpage of the Si wafers during processing.

However, oxygen behaves differently when Cz-Si is annealed at lower temperatures. Annealing at around 450°C for several hours leads to a change in the electrical resistivity of the Si, and it is now known that this is due to the formation of *thermal double donors*, oxygen based defects with double donor levels close to the conduction band. These can form with concentrations of $\sim 10^{16}$cm^{-3}, significantly degrading the performance of the Si. Higher temperature annealing removes these ('donor kill' processing), however there is interest in lowering the annealing temperatures required during processing. High temperature anneals allow greater defect migration and increased interfacial diffusion, both of which become crucial in

[2] Large quantities of oxygen in silicon were first reported from optical absorption experiments in the 1950s [2].

extremely small devices. As well as thermal donors there are various other families of shallow donors that form, including the *shallow thermal donors*, which form over an extremely wide temperature range (some of these are stable up to 900°C), as well as the 'new donors' which form in the 500-600 °C temperature range after the thermal donors anneal out.

The obvious way to avoid these defects would be to minimise the amount of oxygen in the material. There are various ways to achieve this. One particularly successful method is to place large magnetic coils around the molten Si bath ('MCZ', or Magnetic-field-applied-Czochralski). The resultant fields control the convection fluid flow, allowing the operator to minimise the mixing between the liquid in the centre of the bath with that at the edge. This effectively creates a liquid Si crucible around the central Si bath, which can trap much of the oxygen and slow its migration into the crystal[6]. This process drops the final thermal donor concentrations by over an order of magnitude, but cannot provide a final solution to the problem, since oxygen is needed within the crystal. As well as forming unwanted electrically active defects, oxygen performs a number of very useful roles. It can pin dislocations, which greatly strengthens the crystal. Float-zone Si, which contains over two orders of magnitude less oxygen than Cz-Si, is too soft for most commercial applications and crumbles very easily. Oxygen also acts as a gettering agent for trace metal impurities in the crystal ('intrinsic gettering'). Oxygen precipitates in the wafer core suppress stacking faults [7], and oxygen makes the Si more resistant to thermal stress during processing. This is the primary reason why Cz-Si is used for integrated circuit design, where there are many thermal processing steps.

Thus to some extent we are stuck with the problem of the oxygen related donor defects in Si. Given this state of affairs, there are two available courses. Firstly to develop alternative treatments that minimise the donor defect concentration without dropping the oxygen concentration (through alternative heat treatments, addition of other impurities, and possibly even more exotic treatments such as irradiation). Secondly it is known that the formation and distribution of these defects is crucially dependent on the initial condition of the Si[3]. Therefore with improved pre-treatments and processing we may be able to tweak these to our favour, and in practise any solution will probably be a combination of both of these.

In order to pursue this approach, it is obvious that detailed knowledge is required of the behaviour of impurities in Si. Notably the defect structures and formation processes, as well as the primary impurity complexes and their diffusion mecha-

[3]Bob Falster from MEMC, at the NATO ARW 'Oxygen '96', said that the biggest question facing Si researchers was, 'What happens in the first ten minutes of annealing?'

nisms, need to be well understood. Si is the most studied semiconductor material in the world. Despite this, more than forty years of research into thermal donors[4] still leaves many fundamental questions unanswered, including their structure, and their method of formation. Some of the more important questions that need to be addressed are as follows:

- What are the structures and formation processes for thermal donors and shallow thermal donors?
- What is the role, if any, of silicon self-interstitials, both as a part of defect structures such as thermal donors, and as a catalyst for defect formation through enhanced migration or pre-cursor formation?
- What is the role, if any, for oxygen dimers? To what extent do pairs of oxygen atoms control defect formation at lower temperatures?

The last two are addressed in Chapters 6 and 9. The structure and mechanisms for the shallow thermal donors are addressed in Chapter 8. We present a model for the thermal donor core in Chapter 9. Although this does not provide a complete answer to the thermal donor problem, I have included some discussion of this and pointers towards an eventual solution, drawn from these results and the available experimental data.

As well as oxygen, there are many other light element impurities present in Si, either through accident or design. These will interact with oxygen in various ways, forming many different defect complexes. Nitrogen is not very soluble in Si, but can be incorporated in a variety of ways, either through annealing in an N_2 atmosphere, ion implantation, or through addition of Si_3N_4 to the Si melt. Nitrogen helps to suppress 'swirl defects' in the Si and stops wafer bowing during processing. It is also known to suppress thermal donor formation. Nitrogen is electrically inactive in Si, and is known to suppress vacancy and interstitial defects and increase gate oxide integrity [9]. Si_3N_4 is also used as a diffusion barrier. These effects are, in part at least, due to the interaction between nitrogen and oxygen.

Hydrogen is an extremely complex impurity in semiconductors. It is now believed to be present in most Si samples[5]. It is normally used in semiconductor materials as a passivating element, saturating electrically active dangling bonds associated with other defect centres. However, its behaviour is more complex than that. In Chapter 4 we show that in III-V materials a fully hydrogenated vacancy

[4]Thermal donors were first observed in 1954 by Fuller *et al* [8]

[5]Rob Ammerlaan: 'Its [H] traces were also observed in almost all the samples that were not intentionally diffused with hydrogen, confirming the hydrogen contamination of commercially available high-grade silicon' [10, 11]

behaves as a single shallow donor, and in InP this is responsible for charge compensation of iron dopants. In addition, in conjunction with other elements such as carbon we suggest in Chapter 8 that hydrogen may also be responsible for some of the shallow thermal donors observed in silicon. A full study of oxygen–hydrogen interactions in silicon is beyond the scope of this thesis.

Carbon is always present in concentrations of at least 5×10^{15} cm^{-3}, and can form a variety of complexes with oxygen in both its substitutional and interstitial forms. C–O complexes will be covered fully in the PhD of Paul Leary, but some discussion of their formation and behaviour is included here where relevant.

Finally, we examine the interaction between oxygen and vacancies. Vacancies and self-interstitials can be introduced into Si in a variety of ways. Irradiation produces many intrinsic defects, and although most immediately recombine, it is still the easiest way to produce large quantities of vacancies (V) and self-interstitials (Si_i). Proton irradiation hydrogenates the vacancies and prevents recombination, leading to an excess of Si_i. SiO_2 precipitation also creates Si_i, and these aggregate into rod-like defects (RLDs) that lie along $\langle 311 \rangle$ planes. Annealing Si in N_2 gas leads to Si_3N_4 precipitation on the surface, which sends a 'wind' of vacancies into the bulk. Vacancies are known to complex with oxygen and suppress thermal donor formation, and the role of Si_i in the formation of thermal donors is a hotly contested issue. Complexes involving Si_i and O_i are not believed to be stable at normal annealing temperatures and break down at low temperatures [12]. However vacancy–oxygen defects are common, and precipitate instead of thermal donors. The vacancy-oxygen centre, or 'A centre' (OV^-), is the most common impurity in electron irradiated Cz-Si at room temperature. We therefore examine various vacancy–oxygen defects, and show how vacancies act as precipitation centres for oxygen.

Our calculations are performed using AIMPRO, a density functional code working under the local density approximation (LDA). The method is *ab initio*, *i.e.* the only experimental input required is the atomic number of the elements involved. The calculations use large molecular clusters to simulate bulk crystal. These clusters contain from 35 to 98 bulk atoms, and have their surface dangling bonds terminated with hydrogen atoms (a typical 134 atom cluster is shown in Figure 4.1a)[6]. The defect structure is inserted into the centre of these clusters and some or all of the atoms allowed to relax to their ground state. Thus we are able to accurately describe the structure and symmetry of point defects. The calculated Kohn-Sham eigenvalues also provide semi-quantitative information about the localised electronic states

[6]AIMPRO can also tackle much larger problems, given sufficient computing time; it has been used to self-consistently relax an 840 atom carbon 'bucky-onion', one of the largest DFT calculations performed to date.

of the defects. Determination of the energy second derivatives for the core defect atoms also allows us to calculate the local vibrational modes of defects with different atomic isotopes, which can be directly compared to experiment. By relaxing various alternative structures it is possible to determine defect binding energies, and by applying constraints to the relaxation to hold the cluster at a saddle point, barriers to migration can be determined.

In Chapter 2 the theoretical background to AIMPRO is discussed. This includes Hartree Fock, Thomas Fermi Theory and Density Functional Theory, as well as some of the approximations used. Also included is a discussion of other theoretical methods commonly used to examine oxygen defects in Si, such as CNDO.

Chapter 3 describes the formalism used within AIMPRO. This includes discussion of the advantages and disadvantages of the method used, including the role of clusters, hydrogen termination and localised Gaussian basis sets.

Chapter 4 is concerned with hydrogenated defects in InP and GaAs. It is first shown that hydrogenated vacancies can form a range of electrically active defects, from VH_4, which is a single shallow donor, to VH, which acts as a double acceptor. VH_4 is shown to be responsible for the charge compensation of iron in InP. The role played by hydrogen in passivating group II acceptor impurities such as Be and Mg is also examined. The work is extended to passivated Group II acceptors in GaAs to examine trends with material type, and is closely tied to experimental research performed by Bernard Pajot in Paris.

The other results chapters are concerned with oxygen- related defects in silicon. Chapter 5 examines the range of vacancy–oxygen defects, V_nO_m, n = 1,2, m = 1,3. For the first time, coupled with experimental work from Århus, Denmark, we unambiguously show that VO_2 is defect responsible for the 889 cm^{-1} local vibrational mode (LVM). A discussion of the annealing behaviour of V_nO_m defects is included, as well as the role of oxygen dimers.

Chapter 6 examines the interstitial oxygen dimer, O_{2i}. This defect is believed to be responsible for the low formation energy of the thermal donors from 350-400°C, and it has been suggested that it has a lower migration barrier than O_i. In collaboration with experimental work done at Linköping by Lennart Lindström and Tomas Hallberg, we show the oxygen dimer is responsible for the 1012 cm^{-1} LVM observed in Cz-Si[7]. Our calculations show that the dimer is at least metastable in a puckered assymetric form, and does indeed have a lower migration barrier. The migration mechanism is discussed. We also extend the argument to larger oxygen complexes such as the trimer, and discuss the role of the dimer in the formation of

[7]The experimental work was also performed in collaboration with Vladimir Markevich and Leonard Murin.

other defect complexes in Si.

Nitrogen–Oxygen defects are examined in Chapter 7. We show that the primary N–O defect in silicon is the NNO defect, or $N_{2i}O_i$. This is electrically inactive and consists of O_i neighbouring an interstitial nitrogen pair. This work was done in collaboration with Frank Berg Rasmussen and Brian Bech Nielsen at Århus, Denmark. We also examine the N_iO_i defect and correlate it with a set of previously unassigned LVMs.

Chapter 8 deals with the *shallow thermal donors* and is to some extent the core of the thesis. We show that both N_iO_{2i} and $(CH)_iO_{4i}$ can form shallow thermal donors, and probably lie at the core of two of the shallow thermal donor families in Si (responsible for some of the NL10 defects). A general mechanism (the 'wonderbra mechanism') is described for turning defects with deep donor levels into shallow donor level defects, through electrostatic compression by O_i. This is shown to operate for the shallow thermal donors described here. The relevance of this mechanism for thermal donors, and possibilities for application to other materials is also discussed. This work arose from collaboration with Jo Miro and Peter Deák, theoretical researchers from Budapest. As an area of active ongoing research, we are currently involved in discussion with experimental researchers working on this problem in Japan, London, Amsterdam and Stüttgart.

Thermal donors (TDs) are examined in Chapter 9. We propose a core structure for the thermal donor based around two tri-valent oxygen atoms, and account for the various experimental data. We also discuss how the family of thermal donors can form, and possible structures for higher order TDs. A model is proposed for the NL10(Al) shallow thermal donor centre.

In Chapter 10, final conclusions are drawn and suggestions made for further work. Some general observations on the behaviour of oxygen-related species in Si are made, and the relevance of this work to the problem of thermal donors is discussed.

Chapter 2

Theoretical Background

'The man who makes no mistakes does not usually make anything.'

Edward John Phelps (1822-1900)

Electronic structure calculations attempt to determine the relationship between a system of ions and their electrons. In order to tackle problems of a useful size, given our current computational limits, it is necessary to make a number of assumptions. In the following chapter I will attempt to describe the various methods and assumptions people have used, ending with DFT/LDA that lies at the core of AIMPRO.

We assume that the ground state of a given system is determined by the time independent, non-relativistic many-body Schrödinger equation,

$$(H_{total} - E_{total})\Psi_{total}(r, R) = 0. \tag{2.0.1}$$

H_{total}, the many body Hamiltonian, is defined in terms of kinetic (T) and potential (V) terms for ions (i) of mass M_a and charge Z_a at sites $\mathbf{R}_a$, and electrons (e) at sites $\mathbf{r}_\mu$,

$$\begin{aligned} H_{total} &= T_i + T_e + V_{e-e} + V_{e-i} + V_{i-i} & (2.0.2) \\ &= \Bigg\{ -\sum_a \frac{m}{2M_a}\nabla_a^2 - \frac{1}{2}\sum_\mu \nabla_\mu^2 + \frac{1}{2}\sum_{\nu\neq\mu} \frac{1}{|\mathbf{r}_\mu - \mathbf{r}_\nu|} \\ &\quad - \sum_{\mu,a} \frac{Z_a}{|\mathbf{r}_\mu - \mathbf{R}_a|} + \frac{1}{2}\sum_{a\neq b} \frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|} \Bigg\}. & (2.0.3) \end{aligned}$$

However this equation is insoluble for all except the most simple systems. We therefore invoke the *Born-Oppenheimer* approximation. This assumes that on the time scale at which the ions move, the electrons are able to respond instantaneously.

Therefore the ionic wavefunction is independent of the electrons, and we can split the wavefunction of the system into the product of ionic and electronic terms

$$\Psi_{total}(r, R) = \chi(R)\Psi(r, R). \tag{2.0.4}$$

We can then eliminate the co-dynamic terms from the equation for Ψ and thus solve the Schrödinger equation for the electronic wavefunction only, given a set of fixed atomic coordinates.

In practise the ions and electrons are not decoupled and there is also a similar equation for $\chi(R, r)$ that should include ionic/electronic coupling. However this term is only relevant for near degenerate ground states where additional effects such as Jahn-Teller distortions become important. These effects are not important in this study.

2.1 The Hartree Equations

The Hartree equations were derived by considering the electronic wavefunction as a product of many single particle wavefunctions, ψ_λ, each subject to its own external potential. Thus

$$\begin{aligned}\Psi(\mathbf{r_1}s_1, \mathbf{r_2}s_2, \cdots, \mathbf{r_N}s_N) &= \psi_1(\mathbf{r_1}s_1)\psi_2(\mathbf{r_2}s_2)\cdots\psi_N(\mathbf{r_N}s_N), \\ &= \Pi_\lambda \psi_\lambda(\mathbf{r}_\lambda s_\lambda)\end{aligned} \tag{2.1.5}$$

The Schrödinger equation for each single particle state, λ, is given by

$$\left(-\frac{1}{2}\nabla^2 + V(r) - E_\lambda\right)\Psi_\lambda(r) = 0 \tag{2.1.6}$$

Hartree suggested that in a many electron system it was possible to replace $V(r)$ with $V_{eff}(r)$, which included the mean electrostatic potential from all of the electrons [13]. This leads to the Hartree self-consistent equations

$$\left(-\frac{1}{2}\nabla^2 + V_{eff}(r) - E_\lambda\right)\Psi_\lambda(r) = 0, \tag{2.1.7}$$

$$V_{eff}(r) = V(r) + \int \frac{n(r')}{|r - r'|}dr', \tag{2.1.8}$$

$$n(r) = \sum_{\lambda=1}^{N} |\Psi_\lambda(r)|^2. \tag{2.1.9}$$

The total many body wavefunction is then formed as the product of these one-electron spin-orbitals. $V(r)$ is the potential due to the ions. This simple approximation leads to surprisingly accurate results, but suffers from the lack of electronic exchange, *i.e.* the many body wavefunction, $\Psi(r, R)$ is not antisymmetric with respect to a switching of two of the electrons. This is in breach of the Pauli principle, which in general imposes a change of sign of the total wavefunction when two electrons are switched, and in the independent electron approximation described above, does not allow more than one electron to occupy a given state.

2.2 Hartree-Fock Theory

The lack of exchange in Hartree theory is overcome using Hartree-Fock theory, where the many body electronic wavefunction is rewritten [14]. Instead of using a simple product wavefunction, (Equation 2.1.5 above), a Slater determinant is used,

$$\Psi(\mathbf{r}_1 s_1, \mathbf{r}_2 s_2, \ldots, \mathbf{r}_N s_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1 s_1) & \psi_1(\mathbf{r}_2 s_2) & \ldots & \psi_1(\mathbf{r}_N s_N) \\ \psi_2(\mathbf{r}_1 s_1) & \psi_2(\mathbf{r}_2 s_2) & \ldots & \psi_2(\mathbf{r}_N s_N) \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ \psi_N(\mathbf{r}_1 s_1) & \psi_N(\mathbf{r}_2 s_2) & \ldots & \psi_N(\mathbf{r}_N s_N) \end{vmatrix}.$$

Incorporating electron spin through a spin function, $\chi_\alpha(s)$, this can be rewritten as

$$\Psi(r) = \frac{1}{\sqrt{N!}} \det|\psi_\lambda(r_\mu)|$$
$$\psi_\lambda(r) = \psi_i(\mathbf{r})\chi_\alpha(s)$$
$$\sum_s \chi_\alpha^*(s)\chi_\beta(s) = \delta_{\alpha\beta}.$$

We can define the electron density in terms of these orbitals,

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

In order to find the ground state it is necessary to minimise the energy while keeping the orbitals, ψ_λ, orthonormal. This is done by applying the variational principle, adding Lagrange multipliers, $E_{\lambda\mu}$ to Equation 2.1.7,

$$\left(-\frac{1}{2}\nabla^2 + V_{eff}(r) - E_\lambda\right)\psi_\lambda(r) - \sum_{\mu\neq\lambda} E_{\lambda\mu}\psi_\mu(r) = 0.$$

These set of equations are known as the *Hartree-Fock equations*. Writing them more completely, for a given orbital, λ,

$$\left\{-\frac{1}{2}\nabla^2 + V_{e-i}(\mathbf{r}) + V_H(\mathbf{r}) + V_\lambda^x(\mathbf{r}) - E_\lambda\right\}\psi_\lambda(r) = \sum_{\mu\neq\lambda} E_{\lambda\mu}\psi_\mu(r), \tag{2.2.10}$$

$$V_H(\mathbf{r})\psi_\lambda(r) = \frac{\delta E_H}{\delta\psi_\lambda^*} = \int \frac{n(\mathbf{r}_1)\psi_\lambda(r)}{|\mathbf{r}-\mathbf{r}_1|}d\mathbf{r}_1,$$

$$V_\lambda^x(\mathbf{r})\psi_\lambda(r) = \frac{\delta E_x}{\delta\psi_\lambda^*} = -\sum_{\mu s_1}\int \psi_\mu^*(r_1)\psi_\lambda(r_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}|}\psi_\mu(r)d\mathbf{r}_1,$$

with the orthonormality constraint

$$\sum_s \int \psi_\mu^*\psi_\lambda d\mathbf{r} = \delta_{\lambda\mu}.$$

Diagonalising $E_{\lambda\mu}$ removes the Lagrange multiplier term on the right hand side, and the equation can then be solved self-consistently. To find the total energy, E_{TOT}, Equation 2.2.10 is multiplied by $\psi_\lambda^*(r)$, integrated over $\mathbf{r}$ and summed over s and λ to give

$$E_{TOT} = \sum_\lambda E_\lambda - E_H - E_x + E_{i-i}.$$

E_λ is the ionisation energy for the λ electron excluding any redistribution in the remaining orbitals once an electron is removed.

The orbitals are normally expanded in terms of plane waves, and for a homogeneous electron gas the exact plane wave solution is given by

$$\psi_\lambda(\mathbf{r}, s) = \frac{1}{\sqrt{\Omega}}e^{i\mathbf{k}.\mathbf{r}}\chi_\alpha(s).$$

In the spin polarised case these are split into the spin 'up' and spin 'down' populations, each with its own Fermi wavevector; the total energy is a sum of these two populations (n_s is the electron density of spin state s). For the homogeneous electron gas this is given by

$$E = \Omega\sum_s\left\{\frac{3n_s}{10}(6\pi^2 n_s)^{\frac{2}{3}} - \frac{3}{2}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}} n_s^{\frac{4}{3}}\right\}.$$

2.2.1 Discussion

Hartree-Fock theory is a useful tool for many problems. Results can be systematically improved through an increasingly rigorous choice of basis set (AIMPRO is limited in this respect by overcompleteness problems; if the basis set becomes too large then it becomes difficult for the system to achieve self-consistency). It gives good structures and vibrational modes for small molecules, although tends to underpredict bond lengths with a corresponding overprediction of frequencies.

A major flaw with the method is the lack of electron correlation (leading to zero density of states at the Fermi level for Jellium). This precludes the use of Hartree-Fock theory in problems such as, for example, metallic systems. The difference in correlation between ground and excited states in real systems means that Hartree-Fock greatly overpredicts the width of the band gap.

Hartree-Fock provides the basis for methods covering a wide spectrum of accuracies. At the lower level there are semi-empirical methods such as MINDO/3, MNDO, etc., which are described in more detail below. These attempt to compensate for the lack of correlation through their parameterisation methods, additionally fitting to experimental ionisation energies and/or heats of formation of molecules. At the other end of the scale, many computationally expensive improvements can be utilised to overcome the correlation problem, such as the configuration interaction (CI) method [15], where the wavefunction is constructed from a combination of determinants. These are only really viable for small molecular systems at present.

In addition, the calculation size scales as $O(N^4)$, *i.e.* doubling the system size leads to a 16-fold increase in calculation time. This poor scaling with system size, coupled with the heavy computational cost of Hartree-Fock theory, currently limits it to smaller systems. Nonetheless, Hartree-Fock based calculations (when incorporating the high level corrections such as CI) often provide the benchmarks that other theoretical methods aspire to.

2.3 Thomas-Fermi Theory

Thomas-Fermi theory marks a change in approach from Hartree and Hartree-Fock theory, as it was the first method to propose using the electronic charge density as its fundamental variable instead of the wavefunction. It is thus the earliest form of density functional theory.

The model can be understood with reference to Figure 2.1. Although the charge density is that of a non-uniform electron gas, the number of electrons in a given element, dr, can be expressed as $n(r)dr$, where $n(r)$ is the charge density for a *uniform* electron gas at that point. It is then possible to express the energy as a

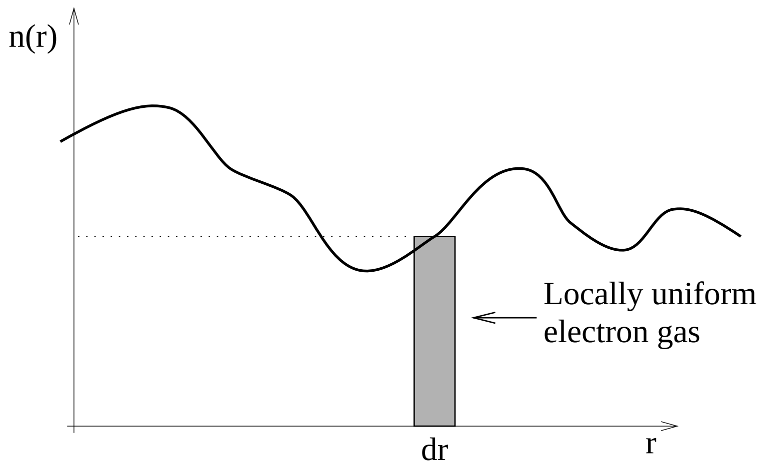


Figure 2.1: Schematic diagram showing the principle of the local density approximation and also Thomas Fermi theory, namely that for a given radial slab, dr, the local charge density can be considered to be $n(r)$, the density of an equivalent uniform homogeneous electron gas.

function of $n(r)$,

$$E_{TF}[n(r)] = \int v(r)n(r)dr + \int \frac{3}{10}(3\pi^2)^{\frac{2}{3}} n^{\frac{5}{3}}(r)dr + \frac{1}{2}\int \frac{n(r)n(r')}{|r-r'|}drdr'. \quad (2.3.11)$$

Again the variational principle is applied and the energy, E_{TF} is minimised subject to the condition

$$\int n(r)dr = N.$$

A Lagrange parameter of μ incorporates this and leads to the Thomas-Fermi self-consistent equations,

$$n(r) = \frac{1}{3\pi^2}\left\{2\left[\mu - v_{eff}(r)\right]\right\}^{\frac{3}{2}}, \quad (2.3.12)$$

$$v_{eff}(r) = v(r) + \int \frac{n(r')}{|r-r'|}dr'. \quad (2.3.13)$$

Note that μ is equivalent to the chemical potential, $\frac{\partial E}{\partial N}$. This approach is extremely simple, and is qualitatively correct for atoms. However it does not give a binding energy for molecules, and the proof is not rigorous. To overcome these problems it is necessary to move on to the work of Hohenberg, Kohn and Sham.

2.4 Density Functional Theory

Density Functional Theory (DFT) is a way of simplifying the many body problem by working with the electronic charge density as fundamental variable rather than the wavefunction. It is a ground state theory that incorporates both exchange and correlation effects, and is more computationally efficient than Hartree-Fock methods. However it brings with it various problems that will be outlined below. These include having eigenvalues that do not necessarily correspond with the physical eigenvalues of the system (with the exception of the highest occupied state), as well as poor excited state modelling and hence band gap prediction.

DFT was originally developed by Hohenberg and Kohn [16] and later developed by Kohn and Sham [17]. The crux of this work is the proof that it is valid to use the charge density as the fundamental system variable, *i.e.*, that for a given non-degenerate non-polarised ground state wavefunction, there is a unique electron density. Thus Ψ is a unique functional of the charge density, n, and hence the energy E is uniquely defined by n.

We first define for a given wavefunction, Ψ_n, a Hamiltonian, H_n given by

$$H_n = T + U + V_n.$$

The proposal is that the ground state electron density is in unique correspondence with the external potential, V_n (T and U are the common kinetic and interaction energy operators respectively). If this was not the case there would be two or more such potentials, say, V_1 and V_2 having the same n. By the Rayleigh-Ritz variational principle for the ground state energy, this would imply that

$$\begin{aligned} E_1 = \langle\Psi_1|H_1|\Psi_1\rangle &< \langle\Psi_2|H_1|\Psi_2\rangle, \\ &= \langle\Psi_2|H_2 + V_1 - V_2|\Psi_2\rangle, \\ &= E_2 + \langle\Psi_2|V_1 - V_2|\Psi_2\rangle, \\ &= E_2 + \textstyle\int (V_1 - V_2) n(\mathbf{r}) d\mathbf{r}. \end{aligned}$$

However, the expression is equally valid with the terms reversed,

$$E_2 < E_1 + \int (V_2 - V_1) n(\mathbf{r}) d\mathbf{r}.$$

Adding these two inequalities gives

$$E_1 + E_2 < E_1 + E_2,$$

which is clearly a contradiction unless the energies are degenerate.

This remarkable proof shows that since the charge density can unambiguously specify the external potential, then contained within the charge density is the total information about the stationary system; eigenvalues, wavefunctions, Hamiltonians, etc. Thus what was a $4M$ variable problem (where M is the number of electrons, each one having three Cartesian variables and electron spin) is now reduced to the four variables needed to define the charge density at a point.

It is now theoretically possible to write an *energy functional* in terms of the charge density,

$$E[n] = E_{i-i} + E_{e-i} + F[n(r)], \tag{2.4.14}$$

where $F[n(r)]$ is a universal functional of density; *i.e.*, applying the variational principle, when $n(r)$ corresponds exactly to that induced by the external potential, the total energy functional will be at a minimum, the ground state energy. In principle, at this stage this formalism is still exact.

$F[n(r)]$ can be broken down into its components,

$$F[n(r)] = \frac{1}{2}\int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + \text{Kinetic} \tag{2.4.15}$$

where the first term is the *Hartree energy.* The simplest approximation to the kinetic term comes from Thomas Fermi theory [18],

$$\text{Kinetic} = \gamma \int n(\mathbf{r})^{\frac{5}{3}} d\mathbf{r}$$

However, for an exact form we can expand the kinetic energy to give

$$\text{Kinetic} = T[n(r)] + E_{xc}[n(r)]$$

T is the kinetic energy of a set of *non-interacting* particles with a given charge density. However there is also a term, E_{xc}, the *exchange-correlation energy,* that includes all the kinetic terms due to interactions within the gas. This includes exchange energies discussed previously, as well as correlation energy. The electrons will tend to dynamically repel other electrons from their region of space through Coulombic repulsion, creating a 'correlation hole', and this is incorporated in the E_{xc}.

Both the kinetic energy, T, and the exchange correlation energy, E_{xc} are harder to express in terms of the charge density; to determine T we have to invoke Kohn-Sham theory, but to express E_{xc}, we commonly use the *local density approximation.*

2.5 The local density approximation

The local density approximation (LDA) states that, for regions of a material where the charge density is slowly varying, the exchange correlation energy at that point can be considered the same as that for a locally uniform electron gas of the same charge density (see Figure 2.1). In this case we can write E_{xc} as

$$E_{xc} = \int n(\mathbf{r})\varepsilon_{xc}(n)d\mathbf{r}.$$

The spin polarised variation (local spin density approximation, or LSDA) replaces the spin averaged energy density in the above equation with the energy density for a polarised homogeneous electron gas.

Although this approximation is extremely simple, it is surprisingly accurate, and forms the core of most modern DFT codes. It even works reasonably well in systems where the charge density is rapidly varying. However it tends to underpredict atomic ground state energies and ionisation energies, while overpredicting binding energies. It is also known to overly favour high spin state structures. For these reasons there have been attempts to move beyond the LDA, notably through the addition of *gradient corrections* to incorporate longer range gradient effects [19]. However in practise, although these improvements seem to give better total energies the resultant structure is often worse, and at a greatly increased computational cost. In general, the LDA is worse for small molecules and improves with system size.

The only remaining problem is to find an approximate solution for the homogeneous electron gas exchange-correlation term, $\varepsilon_{xc}(n)$. There are several parameterised prescriptions for this, the one used in AIMPRO is based on the work by Ceperley and Alder [20].

2.5.1 Parameterising the correlation energy

The correlation energy, E_c, for a homogeneous electron gas comes from a set of quantum Monte Carlo calculations performed by Ceperley and Alder [20]. These were later parameterised by Perdew and Zunger [21]. Slight improvements to these parameterisations have been made since, for example Perdew and Wang [22], but AIMPRO uses the Perdew and Zunger parameterisations in order to remain consistent with the pseudopotential parameters of Bachelet, Hamman and Schlüter [23] (see Section 2.7 below).

The results of Ceperley and Alder apply to low density electron gases, and can be combined with results from perturbation theory for high density gases to cover a wide density range. Defining the correlation energy per electron, ε_c, polarisation ξ and Wigner-Seitz radius of each electron, r_s as:

$$E_c = \Omega n \varepsilon_c(n,\xi), \quad \xi = \frac{(n_\uparrow - n_\downarrow)}{n}, \quad r_s = \left(\frac{4\pi n}{3}\right)^{\frac{1}{3}},$$

where n is the electron density, for the non-polarised ($\xi = 0$) and fully-polarised ($\xi = 1$) cases, ε_c is given by[21]:

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

Table 2.1: Parameters for the correlation energy per electron given in Equation 2.5.16. Taken from Reference [21].

$$\varepsilon_c = \begin{cases} \gamma \left\{1 + \beta_1\sqrt{r_s} + \beta_2 r_s\right\}^{-1} & r_s \geq 1 \\ B + (A + Cr_s)\ln(r_s) + Dr_s & r_s < 1 \end{cases} \tag{2.5.16}$$

The value of the coefficients γ, β_1 and β_2 are given in Table 2.1. When the gas is only partially polarised ($0 < \xi < 1$), an average of the polarised and non-polarised cases is taken[24],

$$\varepsilon_c(n, \xi) = \varepsilon_c^{np}(n) + f(\xi)(\varepsilon_c^{p} - \varepsilon_c^{np}), \tag{2.5.17}$$

$$f(\xi) = \frac{(1+\xi)^{4/3} + (1-\xi)^{4/3} - 2}{2^{4/3} - 2}. \tag{2.5.18}$$

It is possible to simplify these expressions however; to a good approximation the non-polarised and fully polarised energies vary as αn^{s_1} and βn^{s_2} respectively. Therefore for the exchange-correlation energy used in AIMPRO, we use a simplified expression:

$$\begin{array}{ll} \mathrm{Non-polarised} & E_{xc} = \Omega A n^{1.30917} \\ \mathrm{Polarised} & E_{xc} = \Omega \sum_{i,s} A_i n_s^{p_i+1} n_{1-s}^{q_i} \end{array} \tag{2.5.19}$$

The coefficients, A_i, p_i and q_i are given in Table 2.2. For larger density values ($n > 1$), an alternative set of parameters are used, A'_i, p'_i and q'_i (also listed in Table 2.2), which are more accurate in the high density regime. The error when using these approximations is normally less than 3% for densities up to $n = 15$. For calculations such as these with no heavier metal impurities, the density is normally in the range $0 < n < 1$.

2.6 Kohn-Sham Theory

Obtaining an expression for the kinetic energy of the electrons in terms of the charge density is a harder problem. The first approximation, the local density type

i	A_i	p_i	q_i
1	-0.9305	0.3333	0.0000
2	-0.0361	0.0000	0.0000
3	0.2327	0.4830	1.0000
4	-0.2324	0.0000	1.0000
i	A'_i	p'_i	q'_i
1	-0.9305	0.3333	0.0000
2	-0.0375	0.1286	0.0000
3	-0.0796	0.0000	0.1286

Table 2.2: Parameters for the exchange-correlation energy used in AIMPRO, see Equation 2.5.19.

of Thomas Fermi approach described above is not sufficiently accurate [18]. It is therefore necessary to break up the charge density into a set of orthonormal orbital functions, as first proposed by Kohn and Sham [17]. These are single particle wave-functions in a non-interacting system (since interaction terms have been included through E_{xc}). For simplicity here we only consider the spin averaged theory,

$$n(\mathbf{r}) = \sum_{\lambda=1,s}^{N} \delta(s_\lambda, s)|\psi_\lambda(\mathbf{r}, s)|^2. \tag{2.6.20}$$

This means we can now write an expression for T, the kinetic energy, as

$$T = -\frac{1}{2}\sum_{\lambda,s}\int \psi_\lambda^* \nabla^2 \psi_\lambda d\mathbf{r}. \tag{2.6.21}$$

Once the total number of electrons and spin of the system are fixed, these orbitals can be determined using two constraints. Firstly we minimise the energy with respect to the charge density. Secondly different orbitals are kept orthogonal, and normalised through the introduction of a set of Lagrange multipliers, E_λ. Therefore we minimise

$$E[n] - \sum_\lambda E_\lambda \left\{ \sum_s \int |\psi_\lambda(r)|^2 \, d\mathbf{r} - 1 \right\}. \tag{2.6.22}$$

Differentiating this with respect to $\psi_\lambda^*(r)$ gives

$$-\frac{1}{2}\nabla^2\psi_\lambda(\mathbf{r}) + \left(v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}) \right) \psi_\lambda(\mathbf{r}) = E_\lambda \psi_\lambda(\mathbf{r}) \tag{2.6.23}$$

$$v(\mathbf{r}) = -\sum_a \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|}$$

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}$$

It can be seen that Equation 2.6.23 is a single particle Schrödinger equation, as we originally proposed in this section. If we therefore rewrite Equation 2.6.23 in terms of an effective potential, $v_{eff}(\mathbf{r})$, we obtain

$$\left(-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right)\psi_\lambda(\mathbf{r}) = E_\lambda \psi_\lambda(\mathbf{r}) \qquad (2.6.24)$$

$$v_{eff}(\mathbf{r}) = -\sum_a \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|} + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \qquad (2.6.25)$$

$$n(\mathbf{r}) = \sum_{\lambda=1,s}^{N} \delta(s_\lambda, s)|\psi_\lambda(\mathbf{r}, s)|^2. \qquad (2.6.26)$$

These three equations together constitute the *Kohn-Sham equations*, and the self-consistent solution of these leads to the ground state charge density of the system. Note that this does not necessarily correspond to the ground state of the total system since the ionic component has been removed; the energy with respect to ionic position has to be minimised in addition to this. In practise the forces on the ions are iteratively minimised, and for each iteration the Kohn-Sham equations are solved to find the charge density ground state for that ionic configuration.

2.6.1 Discussion

Expanding the charge density into a set of orthonormal wavefunctions is purely used as an artificial way of separating the charge density, and no checks are included to map these onto a set of physical single particle states. Nevertheless there often appears to be a good correlation between these and single particle states [25], and they are routinely used as semi-quantitative indicators of the behaviour of real states. The exception to this is the ionisation energy, *i.e.* the energy of the highest occupied orbital, which is given exactly in this theory.

All wavefunction plots included in this thesis show Kohn-Sham wavefunctions rather than true single particle plots, and are labelled accordingly. Equally any interpretation of the associated energies, E_λ as quasi-particle energies is not correct although there is normally a qualitative agreement between the two. Note that this issue is further complicated through the use of clusters, which artificially widens the band gap.

Density functional theory has a number of advantages, discussed above. The exchange potential does not depend on the orbitals, λ, unlike Hartree-Fock theory. Working with the charge density minimises the number of system variables, and in conjunction with approximations such as the LDA large, accurate calculations are possible. However DFT also has a number of shortcomings. As a ground state theory any physical interpretation of calculated excited states is not theoretically justified. It consistently underpredicts semiconductor band gaps, although in AIMPRO calculations this is compensated for through the use of finite clusters, which tend to increase the band gap (see Section 3.7.3). In some cases the band gap predictions are qualitatively wrong, e.g. for NiO, DFT predicts metallic behaviour whereas experimentally it is a good insulator.

2.7 Pseudopotentials

DFT as described above would still prove computationally too difficult for system sizes useful for point defect modelling. It is therefore necessary to simplify the system further, and one important way of achieving this is through the use of pseudopotentials. These rely on the fact that the core electrons sit tightly bound to their host nuclei, and only the valence electrons are involved in chemical bonding. Therefore it is possible to incorporate the core states into a bulk nuclear potential, or pseudo-potential, and only deal with the valence electrons separately.

Core wavefunctions are rapidly varying with many nodes, and require a large number of fitting functions to model accurately; pseudopotentials overcome this as the pseudo-wavefunction is smoothly varying and nodeless within a certain core radius. In addition the Coulombic interaction of core states would tend to swamp any total energy calculations trying to compare small differences in valence bond energies. In order to remain orthogonal to the core states, the valence states are also forced to become rapidly oscillating near the atomic core, requiring large numbers of fitting functions to model accurately. Pseudopotentials overcome all of these problems. They become particularly important for the heavier elements when the number of core electrons vastly outweighs the valence electrons. Finally for heavier atoms in which relativistic effects are important, and so the Dirac equation is required, the valence electrons can be treated non-relativistically. Therefore removal of the core electrons allows a non-relativistic approach to be maintained (some corrections are required in the core electron pseudopotentials to account for relativistic effects, see below).

Pseudopotentials assume that the localised core states do not take part in bonding, and so can be incorporated as a change in the charge state of the nucleus. Be-

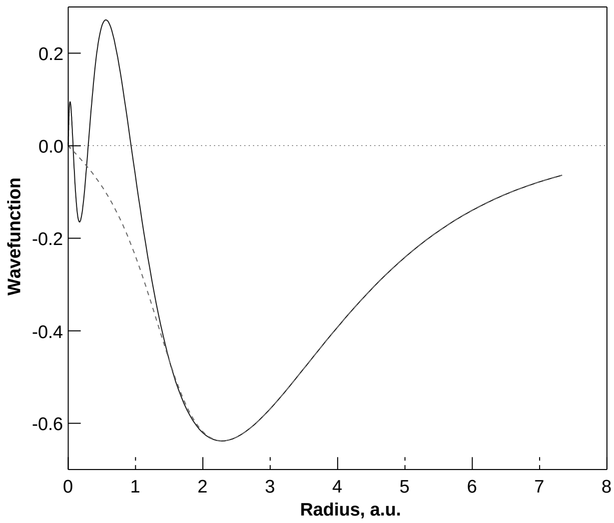


Figure 2.2: The 4s (full) and pseudo- (dashed) radial wavefunction (in atomic units) for the Ni atom.

tween the nucleus and some cut-off radius r_c, the all-electron valence wavefunctions can be replaced with a smoothed form; beyond this radius the pseudo-wavefunction must be identical to the full all-electron wavefunction. In addition there will still be exchange interactions between core and valence electrons which must be included in the potentials. Finally, if self-consistency is to be easily achieved in the calculations then the potentials must have the correct atomic charge density outside the core (*'norm-conserving'* pseudo-potentials).

Figure 2.2 shows the radial 4s full- and pseudo- wavefunction for Ni. It can be seen that the rapidly varying core region has been smoothed in the pseudo-wavefunction but beyond the cut-off radius, r_c, the pseudo- and all-electron- wavefunctions are identical.

The pseudopotentials used in AIMPRO come from work done by Bachelet, Hamann and Schlüter [23, 26] who produced pseudopotentials for all elements up to Pu. There are four fundamental properties which these pseudo-potentials display:

1. Real and pseudo- valence eigenvalues agree for a chosen "prototype" atomic configuration.

2. Real and pseudo- atomic wave functions agree beyond r_c, a chosen "core radius".

3. The integrals from 0 to r of the real and pseudo charge densities agree for $r > r_c$ for each valence state (norm conservation).

4. The logarithmic derivatives of the real and pseudo wave function and their first energy derivatives agree for $r > r_c$.

The last two points are essential if the pseudopotential is to have optimum transferability.

2.7.1 Generating pseudopotentials

Firstly the Kohn-Sham (K-S) equations are solved in an all-electron calculation for a single atom. The configuration of the K-S levels is important, since this is used as a reference atom for the pseudopotential. Therefore two special criteria are applied. Firstly configurations are chosen so the charge density is spherically symmetric, and so can be labelled with angular momentum numbers (or $j = l \pm 1/2$ for heavy elements requiring Dirac formalism). Secondly for the higher angular momentum states a small fraction of the valence electrons are shifted into apparently unoccupied states (e.g. In Si, the d-pseudopotential comes from the configuration $1s^2 2s^1 p^{0.75} d^{0.25}$). This is so that the d-type states are included correctly when constructing the wavefunctions in solids. The electronic fraction chosen for this excitation is picked to eliminate 'bumps' in the potential and is different for different atomic species[1].

The solution of the K-S equations gives the all-electron potential, $V^{\mu}(r)$ which has a Coulomb singularity at $r = 0$. An initial guess pseudopotential is generated for each configuration and angular momentum which eliminates this singularity,

$$V_l(r) = V^{\mu}(r)\left(1 - f(\frac{r}{r_{c,l}})\right) + c_l^{\mu} f\left(\frac{r}{r_{c,l}}\right). \tag{2.7.27}$$

$f(x) = 1$ at the origin and vanishes rapidly as x increases (in their work, $f(x) = e^{-x^{3.5}}$). Thus at the origin, $V_l(r) = c_l^{\mu}$, and this is set so that the lowest energy level for each l matches the equivalent all-electron solution. By shifting the value of c_l^{μ} downwards it is possible to decrease the curvature of the pseudopotential, making it easier to fit; such pseudopotentials are known as *ultra-soft* pseudopotentials, and are essential for plane wave calculations of elements such as oxygen. However these are not always as accurate or transferable as the 'bhs' pseudopotentials described here. $r_{c,l}$ is the core radius and defines a minimum radius beyond which the pseudo wavefunction matches the all-electron wavefunction. As such it is an indication of the quality of the pseudopotential (the closer this lies to the core, the more realistic the potential). However this is a trade off between the minimum preferred radius before the potential becomes too rapidly varying and the benefit of using pseudopotentials is lost; $r_{c,l}$ normally lies halfway to the outermost node of the all-electron wavefunction.

[1]There seems no obvious pattern to the choice of occupancies and for this reason pseudopotential generation is often seen as something of a 'black art'.

The next step is to normalise the wavefunction (since the pseudowavefunction is currently only proportional to the all-electron one). If $V_l(r)$ has a corresponding wavefunction w^{ν}_{1l}, then a new wavefunction is created,

$$w^{\nu}_{2l} = \gamma^{\nu}_{l}\left\{w^{\nu}_{1l}(r) + \delta^{\nu}_{l} r^{l+1} f\left(\frac{r}{r_{cl}}\right)\right\}. \tag{2.7.28}$$

γ^{ν}_{l} and δ^{μ}_{l} are varied until w^{ν}_{2l} exactly equals the all-electron wavefunction outside the core. The Schrödinger equation is then inverted, using the eigenvalues corresponding to the all-electron eigenvalues, which produces the potential that gives rise to w^{ν}_{2l}. This potential nearly conforms to all of the requirements listed above. Finally it is necessary to subtract the Hartree and exchange-correlation contributions to this potential due to the pseudo-wavefunctions themselves (since these are later included explicitly in the calculation). This leaves a bare ion potential. This is exact for the Hartree potential but is only approximate for the exchange-correlation potential as it is non-linear. However the approximation can be improved by subtracting the exchange-correlation potential derived from an all-electron calculation, with the added benefit of improving the transferability of the potential [27].

For relativistic solutions, the pseudopotential ($V_{ps}(r)$) is constructed from an average of the states $j = l \pm \frac{1}{2}$ ($V_l(r)$, the scalar relativistic potential) and a spin-orbit potential, $V^{so}_{l}(r)$,

$$V_{ps}(\mathbf{r}) = \sum_{l} |l\rangle\{V_l(r) + V^{so}_{l}(r)\mathbf{L}.\mathbf{S}\}\langle l|, \tag{2.7.29}$$

$$V_l(r) = \frac{1}{2l+1}\{lV_{l-\frac{1}{2}}(r) + (l+1)V_{l+\frac{1}{2}}(r)\}, \tag{2.7.30}$$

$$V^{so}_{l}(r) = \frac{2}{2l+1}\{V_{l+\frac{1}{2}}(r) - V_{l-\frac{1}{2}}(r)\}. \tag{2.7.31}$$

In practise each pseudopotential is split into two groups, the *local* and *non-local* pseudopotential, depending on whether the terms are independent of l or not, respectively. These two sets of terms are then parameterised to a set of error functions, Gaussian, and $r^2 \times$ Gaussian based functions. It is the coefficients of these which are given in Bachelet, Hamann and Schlüter's paper [23].

2.8 Other theoretical approaches

The problem of oxygen complexes in silicon has been addressed using many other theoretical approaches. Although *ab initio* methods such as ours provide a very complete description of the systems examined, we are limited in the system sizes that we can study. In addition we are unable to tackle molecular dynamics type problems which would require too much computing power. There are therefore other computational methods used, the most notable of these being semi-empirical methods and parameterised potential methods. Deák and Snyder have had much success using CNDO and MINDO/3 semi-empirical methods for pioneering work on large oxygen complexes and y-lid based structures [28], and potentials such as the Stillinger-Weber (SW) and Jiang-Brown (JB) have been used to dynamically model self-interstitial clustering [29] and oxygen diffusion processes [30, 31]. However parameterised methods often suffer from transferability problems and do not give many oxygen systems correctly (e.g. Si-Ô-Si of 180° in O_i and O_s under CNDO [28] and JB [31] respectively). With *ab initio* calculations used to obtain many of the parameters anyway, first-principles studies should be preferred over empirical or semi-empirical calculations where the size of the problem permits.

2.8.1 Semi-Empirical Hartree Fock

Hartree-Fock theory as it stands is too time consuming for use in large systems such as point defect complexes in Si.

However it can be used in a parameterised form, and this is the basis of many of the semi-empirical codes such as *CNDO* and *MINDO/3*. CNDO and related schemes have been used to great effect by Snyder, Corbett and Deák [28] in their work on the behaviour of interstitial oxygen and thermal donors in silicon. By parameterising the overlap matrix to varying degrees, the calculation speed is vastly increased, allowing 130 atom clusters to be routinely optimised on desktop PCs. I will describe in the following section the basis for early parameterisations of CNDO [32]. A lot of information is removed from the calculation when the overlap between orbitals of different atoms is neglected, which in theory could destroy any chance of realistic simulations. However much of the information removed in the simplifications is brought back into the calculations through the parameterisation of experimental data.

The first step in such methods is to expand the one-electron wavefunctions as a set of atomic orbitals; this is known as the *tight binding* method, or sometimes *LCAO* (Linear Combination of Atomic Orbitals),

$$\psi_i = \sum_\mu c_{\mu i}\phi_\mu.$$

This reduces the problem to the determination of the linear expansion coefficients, $c_{\mu i}$. Different ranges of atomic orbitals can be used depending on the accuracy required; commonly these are either Minimal (all atomic orbitals up to and including the valence shell), Extended (Minimal with a few more), or Valence (only the orbitals in the valence shell of the atom). Despite being the least accurate, the valence basis set is commonly employed for very large systems due to its speed and is used in CNDO.

Since ψ_i are orthogonal, it follows that

$$\sum_{\mu\nu} c^*_{\mu i} c_{\nu j} S_{\mu\nu} = \delta_{ij}$$

where δ_{ij} is the Kronecker delta and $S_{\mu\nu}$ is the overlap between atomic orbitals,

$$S_{\mu\nu} = \int \phi_\mu(r)\phi_\nu(r)dr.$$

This means that the Hartree Fock equations can be rewritten as the *Roothaan equations*, a set of algebraic expressions

$$\sum_{\nu}(F_{\mu\nu} - \varepsilon_i S_{\mu\nu})c_{\nu i} = 0. \tag{2.8.32}$$

ε_i is the one-electron energy. $F_{\mu\nu}$ is the 'Fock matrix' and consists of two sets of matrix elements, the core Hamiltonian elements, $H_{\mu\nu}$ and those due to interactions between the different orbitals. These are expressed in terms of a 'density matrix', $P_{\lambda\sigma}$ which is a matrix of the orbital coefficients $c_{\mu i}c_{\nu j}$, coupled with various two-electron integrals over atomic orbitals, $\langle\mu\nu|\lambda\sigma\rangle$,

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma}\left(\langle\mu\nu|\lambda\sigma\rangle - \frac{1}{2}\langle\mu\lambda|\nu\sigma\rangle\right), \tag{2.8.33}$$

$$H_{\mu\nu} = \int \phi_\mu(r)H^{\text{core}}\phi_\nu(r)dr,$$

$$P_{\lambda\sigma} = 2\sum_{i}^{\text{occ}} c^*_{\lambda i}c_{\sigma i},$$

$$\langle\mu\nu|\lambda\sigma\rangle = \int\int \phi_\mu(r_1)\phi_\nu(r_1)\frac{1}{|r_1 - r_2|}\phi_\lambda(r_2)\phi_\sigma(r_2)dr_1dr_2.$$

Now it is possible to make several selective approximations. The first of these is the *zero differential overlap* approximation (ZDOA). This says that since the atomic orbitals on a given atom are roughly orthogonal, the overlap between these can be considered zero; in terms of the two-electron integrals, most of these are therefore set to zero,

$$\langle\mu\nu|\lambda\sigma\rangle = \langle\mu\mu|\lambda\lambda\rangle\delta_{\mu\nu}\delta_{\lambda\sigma}.$$

The corresponding overlap integrals, $S_{\mu\nu}, \mu \neq \nu$ are also ignored for the same reason. This approximation removes all the three and four centre integrals, vastly simplifying the calculation. The chemical bonding is however reliant on the overlap character, and this is maintained through $H_{\mu\nu}$, the core integrals. These are parameterised to experimental results, compensating for the ZDOA to some extent. For CNDO,

$$H_{\mu\nu} = \int \phi_\mu(1) \left(-\frac{1}{2}\nabla^2 - \sum_B V_B \right) \phi_\nu(1) dr_1, \quad \mu \neq \nu,$$

where $-V_B$ is the potential due to the nucleus and core electrons of atom B.

Applying the ZDOA simplifies the Roothaan equations to

$$\sum_\nu F_{\mu\nu} c_{\nu i} = \varepsilon_i c_{\mu i},$$

with

$$F_{\mu\mu} = H_{\mu\mu} - \frac{1}{2} P_{\mu\mu} \langle \mu\mu | \mu\mu \rangle + \sum_\lambda P_{\lambda\lambda} \langle \mu\mu | \lambda\lambda \rangle,$$

and for $\mu \neq \nu$,

$$F_{\mu\nu} = H_{\mu\nu} - \frac{1}{2} P_{\mu\nu} \langle \mu\mu | \nu\nu \rangle.$$

However in the form described here, the ZDOA leads to inconsistencies. The orbitals on a given atom can be transformed into a different orbital set which would breach the zero differential overlap approximation, *i.e.*

$$\phi_\mu \phi_\nu (\mu \neq \nu) \rightarrow \sum_{\mu\nu}^{\text{all } \mu\nu} \phi'_\mu \phi'_\nu A_{\mu\nu}, \quad A_{\mu\nu} \text{ constant}$$

This means that the results would be basis dependent, and to avoid this, further constraints need to be applied. The simplest scheme to achieve this is CNDO.

2.8.1.1 CNDO

Under CNDO, only the valence electrons are considered explicitly. Interactions between electrons on the same atom are handled simply. Since the atomic orbital basis is built from orthonormal atomic functions, then the interaction between different orbitals μ and ν, $U_{\mu\nu} = 0, \mu \neq \nu$, due to their symmetry. The core Hamiltonian elements, $U_{\mu\mu}$, are fitted to experiment. In CNDO/1 these were fitted to atomic ionisation energies, but in CNDO/2 this was modified by fitting to the average of the ionisation potential and electron affinity, making the system more applicable to both electron gain and loss.

In order to overcome the basis dependency of the ZDOA, the remaining two-centre integrals only depend on their host atoms and not on the orbital type. This is achieved by using an average electrostatic repulsion on any atom A caused by atom B, γ_{AB} (the 'electron repulsion integral'), so that

$$\langle \mu\mu|\lambda\lambda\rangle = \gamma_{AB}.$$

In CNDO/1, γ_{AB} is approximated by s-orbital integrals on the atoms,

$$\gamma_{AB} = \int\int s_A^2(r_1)\frac{1}{|r_1 - r_2|}s_B^2(r_2)dr_1 dr_2.$$

Similarly if the core potential (consisting of the nucleus and core electrons) on an atom B is $\mathbf{V}_B$ then to maintain rotational invariance, the interaction between an electron on atom A and this core on atom B must be a constant, V_{AB} (the 'potential integrals').

Similarly to γ_{AB}, in CNDO/1 V_{AB} is calculated using an $s-$orbital on atom A:

$$V_{AB} = Z_B \int \frac{s_A^2(r)}{|r - R_B|}dr.$$

In CNDO/2, an improved parameterisation, this was replaced with a simple approximation, $V_{AB} = Z_B\gamma_{AB}$.

The lack of monatomic differential overlap removes all terms of the form

$$\langle \mu|\mathbf{V}_B|\nu\rangle = 0, \quad \mu \neq \nu.$$

Next the interaction between an electron and the field of two atoms, A and B is simplified (the 'resonance integral', $\beta_{\mu\nu}$)[2], by setting it proportional to the overlap integral, $S_{\mu\nu}$,

$$H_{\mu\nu} = \beta_{\mu\nu} = \beta_{AB}^0 S_{\mu\nu},$$

where β_{AB}^0 are known as the bonding parameters and are dependent on the atom types A and B. Under CNDO/1 these are approximated to $\beta_{AB}^0 = \frac{1}{2}(\beta_A^0 + \beta_B^0)$, and β_A^0 is fitted to *ab initio* calculations for each species.

The precise parameterisation technique used to determine $S_{\mu\nu}$, $U_{\mu\nu}$, V_{AB}, γ_{AB} and β_{AB}^0 varies. Although I have described the method used for CNDO/1 and CNDO/2 (the earliest parameterisations), there have been many refinements to these since. CNDO is extremely fast, and this has meant that when computing

[2]These determine the chemical bonding in the system.

resources were limited, CNDO calculations have been able to perform pioneering calculations on system sizes inaccessible to their *ab initio* cousins. Much of the accuracy lost during the above simplifications is compensated for by the use of parameterisation to experimental and *ab initio* results. Although CNDO is not as accurate as *ab initio* methods such as AIMPRO it remains a useful investigative tool in semiconductor calculations, and a valuable intermediate between empirical potentials and full *ab initio* calculations.

2.8.2 Interatomic Potentials

Interatomic potentials consist of sets of two- and three-body interaction terms of varying complexity, which are parameterised using a variety of experimental and *ab initio* data to fit to. Due to their high speed compared to *ab initio* calculations they can be routinely used with large supercells, and many molecular dynamics timesteps can be calculated. For example, recent studies of VO diffusion in Si used a 512 atom supercell under the Jiang-Brown (JB) potential, with multiple time steps and diffusion constraints [31].

We use a Musgrave-Pople (MP) potential fitted to results from AIMPRO calculations when determining LVMs. The potential can be used to calculate bulk phonon dispersion curves (see Section 4.3). However it is primarily used by our group to determine the energy double derivative terms in the dynamical matrix for bulk Si atoms, since to calculate all of these from first principles would be too time consuming. For this reason I describe here the MP potential, although 'stand-alone' O in Si calculations are normally performed using either the Jiang-Brown[30] or Stillinger-Weber[33] potentials.

The Musgrave-Pople potential [34] is a three- and four- body potential made up of bond stretch and bond bending terms. It is given for atom i by the formula

$$\begin{aligned} V_i = \quad & \frac{1}{4}\sum_j k_r^{(i)}(\Delta r_{ij})^2 + \frac{r_0^2}{2}\sum_{j>k} k_\theta^{(i)}(\Delta\theta_{jik})^2 + r_0\sum_{k>j} k_{r\theta}^{(i)}(\Delta r_{ij}+\Delta r_{ik})\Delta\theta_{jik} \\ & + \sum_{k>j} k_{rr}^{(i)}\Delta r_{ij}\Delta r_{ik} + r_0^2\sum_{l>k>j} k_{\theta\theta}^{(i)}\Delta\theta_{jik}\Delta\theta_{kil}. \end{aligned} \tag{2.8.34}$$

Here Δr_{ij} and $\Delta\theta_{jik}$ are the changes in the length of the $i-j$ bond and angle between the $i-j$ and $i-k$ bond respectively (see Figure 2.3), and the summation is over the nearest neighbours j, k, and l. r_0 is the equilibrium bond length and is often set to be the experimental value. The coefficients k_r, k_θ, $k_{r\theta}$, k_{rr}, and $k_{\theta\theta}$ are defined in Table 2.8.2.

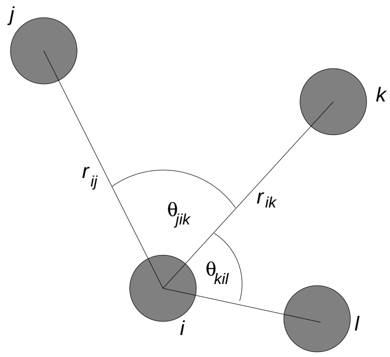


Figure 2.3: Schematic diagram showing terms included in the Musgrave Pople potential

Coefficient	Description
k_r	bond stretch
k_θ	bond bending
$k_{r\theta}$	stretch-bend interaction
k_{rr}	stretch-stretch interaction
$k_{\theta\theta}$	bend-bend interaction

Table 2.3: Physical interpretation of the coefficients in Equation 2.8.34, see schematic diagram, Figure 2.3.

2.8.3 Use of the Musgrave-Pople potential in this work

Local vibrational modes can be considered to be independant of any lattice coupling if they are sufficiently far above the Raman frequency, as is the case with H- stretch modes. However for lower frequency vibrational modes we can expect the mode to couple with the surrounding lattice, particularly once below the Raman frequency. It is therefore essential to take into account the vibrational properties of the surrounding lattice when calculating such modes for defects. Calculating the double derivatives of energy for each atom in the cluster has a prohibitive computational cost, and so it is therefore necessary to make some assumptions.

The way we tackle this problem is using the Musgrave-Pople interatomic potential to model the vibrational properties of those bulk atoms for which we do not calculate the energetic double derivatives directly. This can be used in conjunction with the AIMPRO double derivative data for the core defect atoms.

In order to parameterise the potential it is necessary to calculate the double derivatives for a small group of core atoms in a cluster of pure bulk material, and the parameters are then fitted to these derivatives. The Musgrave-Pople potential parameters are calculated from a least-squares fit to the *ab initio* energy double derivatives, calculated for the core of a fully relaxed bulk cluster. For example, the

InP potential parameters given in Section 4.3 were determined using an 86 atom C_{3v} cluster $In_{22}P_{22}H_{42}$. This was relaxed using AIMPRO, and then the energy double derivatives were calculated for the inner eight atoms, four of each chemical species. This leads to 156 double derivative terms which produces a very accurate fit.

Examples of the application of the MP potential to defect problems is given in Chapter 4, the parameters used for Si are given in Reference [35].

Chapter 3

AIMPRO methodology

"Frequently the approximations that enter solid state calculations are so intricate that they are never specified in the publication in which the calculations are described and would not be of interest if they were. Nevertheless, without knowledge of the approximations, it is sometimes difficult to assess a result that is characterised only as being the best the author could do."

W. A. Harrison, 1980
'Electronic Structure and the properties of solids'

In this chapter I will describe how LSDF theory has been implemented in AIMPRO. I will also discuss some of the practical problems encountered when using AIMPRO and how these can be minimised. A full review of the methodology behind AIMPRO is given in Reference [36] and vibrational mode calculations in Reference [37].

3.1 Expanding the Wavefunction

The wavefunctions are expanded using a basis of localised orbitals, $\phi_i(\mathbf{r}-\mathbf{R}_i)$, where

$$\psi_\lambda(\mathbf{r}, s) = \chi_\alpha(s) \sum_i c_i^\lambda \phi_i(\mathbf{r} - \mathbf{R}_i), \tag{3.1.1}$$

which converts the Kohn-Sham equations into matrix equations for c_i^λ. A set of Gaussian functions are used, multiplied by spherical functions to set the orbital quantum number:

$$\phi_i(\mathbf{r} - \mathbf{R}_i) = (x - R_{ix})^{n_1}(y - R_{iy})^{n_2}(z - R_{iz})^{n_3} e^{-a_i(\mathbf{r}-\mathbf{R}_i)^2}. \tag{3.1.2}$$

The choice of n_1, n_2 and n_3 in Eq. 3.1.2 sets the orbital type; $n_1 = n_2 = n_3 = 0$ gives spherically symmetric s-orbitals, setting one of n_1, n_2 or $n_3 = 1$ gives p-orbitals in the x, y or z direction respectively, whereas setting $\sum_i n_i = 2$ produces a combination of five d- and one $s-$ type orbital.

The charge density for a given spin state can then be described in terms of the *density matrix*, $b_{ij,s}$ (the total charge density is simply the sum of the spin dependent charge densities):

$$\begin{aligned} n_s(\mathbf{r}) &= \sum_{ij} b_{ij,s}\phi_i(\mathbf{r}-\mathbf{R}_i)\phi_j(\mathbf{r}-\mathbf{R}_j) && (3.1.3) \\ b_{ij,s} &= \sum_{\lambda} \delta(s, s_\lambda) c_i^\lambda c_j^\lambda && (3.1.4) \end{aligned}$$

Here, s gives the spin state and λ is an occupied orbital. Substituting this expression for $n(r)$ back into Equation 2.4.14 is only simple for the kinetic and pseudopotential energy terms:

$$\begin{aligned} T_{ij} &= -\frac{1}{2}\int \phi_i(\mathbf{r}-\mathbf{R}_i)\nabla^2\phi_j(\mathbf{r}-\mathbf{R}_j)d\mathbf{r} && (3.1.5) \\ V_{ij}^{ps} &= \int \phi_i(\mathbf{r}-\mathbf{R}_i)\sum_a V_a^{ps}(\mathbf{r}-\mathbf{R}_a)\phi_j(\mathbf{r}-\mathbf{R}_j)d\mathbf{r} && (3.1.6) \end{aligned}$$

However the Hartree energy requires $O(N^4)$ integral terms, and when N, the number of basis functions is large, this rapidly becomes unfeasible. Equally the exchange-correlation energy needs some sort of simplification. We therefore introduce an *approximate* charge density for each spin, $\tilde{n}_s$, which allows analytic solution of these energies and the forces acting on each atom.

3.2 Approximation to the charge density

The approximate charge density is expanded in a set of basis functions [38, 39],

$$\tilde{n}(\mathbf{r}) = \sum_k c_k g_k(\mathbf{r}). \qquad (3.2.7)$$

We can now write an approximation to the Hartree energy in Equation 2.4.15. This is always lower than E_H, but tends to E_H as $\tilde{n}(\mathbf{r})$ tends to $n(\mathbf{r})$ and so provides a way of estimating the quality of the basis function fit,

$$\tilde{E}_H = \int \frac{n(\mathbf{r}_1)\tilde{n}(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 - \frac{1}{2}\int \frac{\tilde{n}(\mathbf{r}_1)\tilde{n}(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2. \qquad (3.2.8)$$

Then we can select c_k in Eq. 3.2.7 to minimise the error

$$E_H - \tilde{E}_H = \frac{1}{2}\int \frac{\{n(\mathbf{r}_1) - \tilde{n}(\mathbf{r}_1)\}\{n(\mathbf{r}_2) - \tilde{n}(\mathbf{r}_2)\}}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2. \tag{3.2.9}$$

If this is now differentiated with respect to c_k in order to determine the minimum we get an important expression,

$$\begin{aligned}
\sum_l G_{kl} c_l &= \sum_{ij} t_{ijk} b_{ij} && (3.2.10)\\
t_{ijk} &= \int \phi_i(\mathbf{r}_1 - \mathbf{R}_i)\phi_j(\mathbf{r}_1 - \mathbf{R}_j) g_k(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 && (3.2.11)\\
G_{kl} &= \int g_k(\mathbf{r}_1) g_l(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 && (3.2.12)
\end{aligned}$$

Evaluation of these 't matrices' is often the slowest step in these cluster calculations, so obviously a careful choice of fitting function, g_k, is required. Simple Gaussian functions have the advantage of being analytically soluble, but if there are roughly as many g_k as basis functions ϕ_i (commonly the case), this means the t matrices are still $O(N^3)$ integrals. Nonetheless, this is the way all the O in Si work was performed. However it is possible to split the g_k into two types of function,

$$\begin{aligned}
g_k &= A e^{-b_k(\mathbf{r} - \mathbf{R}_k)^2} && (3.2.13)\\
A &= \begin{cases} 1 - \frac{2b_k}{3}(\mathbf{r} - \mathbf{R}_k)^2 \\ 1 \end{cases} && (3.2.14)
\end{aligned}$$

The first of these gives a Gaussian potential,

$$\int \frac{g_k(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1 = \frac{3b_k}{2\pi} e^{-b_k(\mathbf{r} - \mathbf{R}_k)^2}, \tag{3.2.15}$$

making t_{ijk} a simple product of three Gaussians which is quick to calculate. However the form of these expressions means that their integral vanishes, and hence they do not contribute to the total charge density. Therefore the second set of expressions for g_k are introduced (simple Gaussians) to ensure the integrated charge density gives the correct number of electrons.

This formalism was adopted for the III-V work in Chapter 4, however when testing different fitting functions for small Si–O based molecules such as disiloxane, $(SiH_3)_2O$, we found that a much better fit was obtained using simple Gaussian functions. Therefore these have been used for all the Si work despite the loss in computing speed.

The exchange-correlation is similarly approximated as

$$\tilde{E}_{xc} = \int \varepsilon_{xc}(\tilde{n}_\uparrow, \tilde{n}_\downarrow)\tilde{n}d\mathbf{r}. \tag{3.2.16}$$

However the expression used to minimise the error in $\tilde{n}$ for E_H (Equation 3.2.9) was integrated over all $\mathbf{r}$ and does not necessarily minimise the error at any specific value of $\mathbf{r}$. Therefore it is necessary to choose a new approximation for $\tilde{n}_s(\mathbf{r})$ and a good choice is a sum of Gaussians,

$$\tilde{n}_s(\mathbf{r}) = \sum_k d_{k,s}h_k(\mathbf{r}), \tag{3.2.17}$$

with $d_{k,s}$ obtained by minimising

$$\int \{n_s(\mathbf{r}) - \tilde{n}_s(\mathbf{r})\}^2 d\mathbf{r}. \tag{3.2.18}$$

If this is differentiated with respect to $d_{k,s}$ we obtain

$$\sum_l H_{kl}d_{l,s} = \sum_{ij} u_{ijk}b_{ij,s} \tag{3.2.19}$$

$$H_{kl} = \int h_k(\mathbf{r})h_l(\mathbf{r})d\mathbf{r} \tag{3.2.20}$$

$$u_{ijk} = \int \phi_i(\mathbf{r}-\mathbf{R}_i)\phi_j(\mathbf{r}-\mathbf{R}_j)h_k(\mathbf{r})d\mathbf{r}. \tag{3.2.21}$$

The integrals are the same for both s, and if g_k is set to the first form in Equation 3.2.13 then u_{ijk} are just proportional to t_{ijk} which makes calculation extremely quick.

3.2.1 Spin averaged $\tilde{E}_{xc}$ and improvements to the approximations

In the spin averaged case,

$$\tilde{E}_{xc} = \sum_k d_k \int h_k(\mathbf{r})\varepsilon_{xc}(\tilde{n})d\mathbf{r}. \tag{3.2.22}$$

If h_k is a Gaussian then the integrals are proportional to $\langle\varepsilon_{xc}(\tilde{n})\rangle_k$. To a first approximation, since $\varepsilon_{xc}(n)$ only varies slowly with n we can approximate this to $\varepsilon_{xc}(\langle\tilde{n}\rangle_k)$, where

$$\langle \tilde{n} \rangle_k = \frac{\sum_l d_l \int h_k h_l d\mathbf{r}}{I_k} \tag{3.2.23}$$

$$I_k = \int h_k d\mathbf{r}. \tag{3.2.24}$$

This approximation can be improved through some mathematical tricks, since we know that a highly accurate form for $\varepsilon_{xc}(n)$ is given by $\varepsilon_{xc}(n) = An^s$ where $s = 0.30917$ (see Eq. 2.5.19).

If we define

$$f(s) = \ln\left(\frac{\langle \tilde{n}^s \rangle_k}{\langle \tilde{n} \rangle_k^s}\right), \tag{3.2.25}$$

we can now do a limited expansion of this to give an approximation for $f(s)$,

$$f(s) = \frac{1}{2}s(s-1)f(2). \tag{3.2.26}$$

This step is one of the largest approximations within our method, leading to errors of up to 10%; however in practise its contribution to the total energy is minute and so the approximation is not unreasonable.

To find $f(2)$ we need to know the second moment of $\tilde{n}$, which can be determined analytically as

$$\langle \tilde{n}^2 \rangle_k = \frac{\sum_{lm} d_l d_m \int h_k h_l h_m d\mathbf{r}}{I_k}. \tag{3.2.27}$$

Finally substituting all of these new terms back in, we get an expression for $\tilde{E}_{xc}$:

$$\tilde{E}_{xc} = \sum_k d_k \varepsilon_k \tag{3.2.28}$$

$$\varepsilon_k = I_k \varepsilon_{xc}\left(\langle \tilde{n} \rangle_k\right) e^{f_k} \tag{3.2.29}$$

$$f_k = \frac{1}{2}s(s-1)\ln\left(\frac{\langle \tilde{n}^2 \rangle_k}{\langle \tilde{n} \rangle_k^2}\right). \tag{3.2.30}$$

3.2.2 Spin polarised exchange correlation

An equivalent set of exchange-correlation terms have been derived for the spin polarised case. Since all of the calculations performed in this thesis are spin averaged I will not describe the methodology here but list the relevant equations for completeness.

$$\tilde{E}_{xc} = \sum_{ks} d_{k,s}\varepsilon_{k,s}, \tag{3.2.31}$$

where

$$\begin{aligned}
\varepsilon_{k,s} &= \sum_i A_i I_k \langle \tilde{n}_s^{p_i} \tilde{n}_{1-s}^{q_i} \rangle_k, && (3.2.32)\\
\langle \tilde{n}_s^p \tilde{n}_{1-s}^q \rangle_k &= \langle \tilde{n}_s \rangle_k^p \langle \tilde{n}_{1-s} \rangle_k^q e^{f(p,q)}, && (3.2.33)\\
f(p,q) &= \ln\left(\frac{\langle \tilde{n}_s^p \tilde{n}_{1-s}^q \rangle_k}{\langle \tilde{n}_s \rangle_k^p \langle \tilde{n}_{1-s} \rangle_k^q}\right), && (3.2.34)\\
&\sim \frac{1}{2}p(p-1)f(2,0) + \frac{1}{2}q(q-1)f(0,2) + pqf(1,1).
\end{aligned}$$

3.2.3 Summary so far

I will now attempt to summarise the method shown so far, and draw out the salient calculational points. Using the approximate functions for E_H and E_{xc} derived in the last few sections we can rewrite the total energy as

$$E = \sum_{ij} \{T_{ij} + V_{ij}^{ps}\} b_{ij} + \tilde{E}_H + \tilde{E}_{xc} + E_{i-i} \tag{3.2.35}$$

where

$$\begin{aligned}
b_{ij} &= \sum_\lambda c_i^\lambda c_j^\lambda,\\
\tilde{E}_H &= \frac{1}{2}\sum_{kl} c_k c_l G_{kl},\\
\tilde{E}_{xc} &= \sum_{ks} d_{k,s}\varepsilon_{k,s},\\
E_{i-i} &= \frac{1}{2}\sum_{a\neq b} \frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|}.
\end{aligned}$$

E is now minimised, however an additional constraint has to be applied in order to keep the wavefunctions orthonormal; this can be imposed in terms of the overlap matrix, S,

$$\begin{aligned}
\sum c_i^\lambda c_j^\mu S_{ij} &= \delta_{\lambda\mu}, && (3.2.36)\\
S_{ij} &= \int \phi_i(\mathbf{r} - \mathbf{R}_i)\phi_j(\mathbf{r} - \mathbf{R}_j) d\mathbf{r}. && (3.2.37)
\end{aligned}$$

Now all we need is Lagrange undetermined multipliers, E_λ to allow unconstrained minimisation with respect to c_i^λ, *i.e.*:

$$\frac{\partial E}{\partial c_i^\lambda} = \frac{\partial\{\sum_{ij\lambda} c_i^\lambda\{T_{ij} + V_{ij}^{ps} - E_\lambda S_{ij}\}c_j^\lambda + \tilde{E}_H + \tilde{E}_{xc} + E_{i-i}\}}{\partial c_i^\lambda} = 0 \tag{3.2.38}$$

$$\sum_j \{T_{ij} + V_{ij}^{ps} + V_{ij}^H + V_{ij,s\lambda}^{xc} - E_\lambda S_{ij}\}c_j^\lambda = 0 \tag{3.2.39}$$

$$\sum_j (H_{ij} - ES_{ij})c_j = 0 \tag{3.2.40}$$

Equation 3.2.39 is the Kohn-Sham equations once more, and is written in terms of two generalised eigenvalue problems (one for each spin).

3.2.4 Matrix Formalism

Re-writing Equation 3.2.40 in matrix notation gives

$$(H - ES)c = 0. \tag{3.2.41}$$

S is now written in terms of an upper triangular matrix using Choleski decomposition,

$$S = U^t U. \tag{3.2.42}$$

If we then define a vector d through $Uc = d$, we can rewrite Equation 3.2.41 as a standard eigenvalue problem

$$\left\{(U^{-1})^t H U^{-1} - E\right\} d = 0. \tag{3.2.43}$$

Evaluation of U and U^{-1} is an $O(N^3)$ problem, and the eigenvalues are then found using a Householder scheme. Since the occupied states only represent a small fraction of the Hamiltonian eigenstates, the required eigenvectors are most efficiently found by inverse iteration since only the required states are calculated. Nonetheless this is normally the time dominant step, scaling as $O(N^3)$.

3.3 Self-consistency

Self-consistency involves redistribution of charge throughout the cluster until a minimum energy is reached, thereby hopefully producing an accurate simulation of the charge distribution in an equivalent real system. This equilibrium distribution of charge for a given set of atomic coordinates is achieved by iteratively solving the Kohn-Sham equations until the charge density distribution produced by the Kohn-Sham orbitals gives the same potential that was used to generate it.

In order to achieve this, a first guess set of charge density coefficients, c_k and $d_{k,s}$ are required; initially we take these from the neutral atom cases but during structural optimisation they are taken from the result of the previous iteration. The Kohn-Sham equations are then solved to determine the density matrix, $b_{ij,s}$, which is then used to produce a set of output charge density coefficients, c_k^0, $d_{k,s}^0$ (Equations 3.2.10 and 3.2.19).

The next choice of charge density coefficients are formed from a weighted combination of the previous two, *i.e.*

$$c_k' = c_k + w(c_k^0 - c_k). \tag{3.3.44}$$

The choice of weighting factor, w, is important. If we solve once for a particular value of w ('w_1') giving a specific charge density c_{1k}^0, then the deviation from self-consistency of c_k is given by

$$e_k = \frac{(1-w)(c_k - c_k^0)}{w} - \frac{(c_{1k}^0 - c_k)}{w_1}. \tag{3.3.45}$$

This e_k can be used to form a pseudo-charge density, $\sum_k e_k g_k(\mathbf{r})$ which has a corresponding electrostatic energy,

$$\frac{1}{2}\sum_{kl} e_k G_{kl} e_l. \tag{3.3.46}$$

w is then chosen by minimising this energy. In practise this can be generalised to include the coefficients from all of the previous iterations, and the cycle converges exponentially, normally in only a few iterations (six or so).

3.3.1 Fermi statistics and 'level smearing'

Convergence is often either slow or impossible when there is only a tiny band gap, since 'charge sloshing' can occur. In one iteration electrons may transfer from a filled level to a lower energy empty level, but in doing so switch the ordering of the levels, so that in the next iteration they switch back. In such situations, or when the structural symmetry leads to degenerate levels that are only partially filled, the self-consistency fails.

In these cases it is either necessary to choose a different structure, or else 'smear out' the occupation of the levels using Fermi statistics, simulating a finite temperature distribution. In practise this can often be used during optimisation and then later switched off once the structure has relaxed past the point where the levels cross.

If we say that a given level, E_λ contains f_λ electrons, then the energy term E that is to be minimised must now be expanded into a free energy, F containing both an entropy term and a constraint to keep the total number of electrons fixed to M,

$$F = E + k_b T \sum_\lambda \{f_\lambda \ln f_\lambda + (1 - f_\lambda)\ln(1 - f_\lambda)\} - \mu\{\sum_\lambda f_\lambda - M\},$$

$$\begin{aligned} f_\lambda &= \frac{1}{e^{(E_\lambda - \mu)/k_b T} + 1}, \\ \sum_\lambda f_\lambda &= M. \end{aligned}$$

It is also necessary to generalise the density matrix to

$$b_{ij,s} = \sum_\lambda \delta(s, s_\lambda) f_\lambda c_i^\lambda c_j^\lambda. \tag{3.3.47}$$

In this work, $k_b T$ was always chosen to be 0.04 eV where used, and as long as the final level splitting was greater than that there is not normally any partial occupancy. Where Fermi statistics have been used it is stated in the text. Use of Fermi statistics does not always lead to the correct structures since, for example, the driving force for a Jahn-Teller distortion will be removed in this way.

3.4 Evaluation of Forces

Now the electronic system is self-consistent it is possible to evaluate the forces on the atoms in order to optimise the structure.

The forces are determined from the gradient of the free energy, F determined above, with respect to displacement,

$$f_{la} = -\frac{\partial F}{\partial R_{la}} \tag{3.4.48}$$

These forces can be determined analytically by considering the change in each term in the total energy for a displacement of ΔR_{la}. Thus

$$\Delta E = \sum_{ij} b_{i,j}\Delta\left\{T_{ij} + V_{ij}^{ps}\right\} + \sum_{ij}\left\{T_{ij} + V_{ij}^{ps}\right\}\Delta b_{ij} + \Delta\tilde{E}_H + \Delta\tilde{E}_{xc} + \Delta E_{i-i},$$

where

$$\Delta\tilde{E}_H = \sum_{kl} c_k G_{kl}\Delta c_l + \frac{1}{2}\sum_{kl} c_k c_l \Delta G_{kl},$$

$$\Delta\tilde{E}_{xc} = \sum_{k,s}\varepsilon_{k,s}\Delta d_{k,s} + \sum_{k,s} d_{k,s}\Delta\varepsilon_{k,s}.$$

These are rearranged and substitutions made to eliminate Δc_l and $\Delta d_{l,s}$, leaving terms in only $\Delta S_{ij}, \Delta G_{kl}, \Delta H_{kl}$ and $\Delta\varepsilon_{k,s}$. $\Delta\varepsilon_{k,s}$ contains terms in $\Delta\langle\tilde{n}_s\rangle_k$ and $\Delta\langle\tilde{n}_s^2\rangle_k$ which can be determined from Equations 3.2.23 and 3.2.27. Although T_{ij} and S_{ij} only depend on R_{la} through the basis functions, $\phi_i(\mathbf{r} - \mathbf{R}_a)$, the pseudopotential term is also dependent on R_{la} due to $V_a^{ps}(\mathbf{r} - \mathbf{R}_a)$; this is determined by integrating by parts to replace with terms in $\Delta\phi$:

$$\int \phi_i(\mathbf{r} - \mathbf{R}_i)\Delta V_a^{ps}(\mathbf{r} - \mathbf{R}_a)\phi_j(\mathbf{r} - \mathbf{R}_a)d\mathbf{r} =$$
$$-\int \left\{\phi_j(\mathbf{r} - \mathbf{R}_j)\Delta\phi_i(\mathbf{r} - \mathbf{R}_i) + \phi_i(\mathbf{r} - \mathbf{R}_i)\Delta\phi_j(\mathbf{r} - \mathbf{R_j})\right\} V_a^{ps}(\mathbf{r} - \mathbf{R}_a)d\mathbf{r}$$

The forces are relatively quick to evaluate compared to determination of the self-consistent energy.

3.5 Structural Optimisation

Once the forces are known it is possible to move the atoms until the structure lies in a system energy minimum. AIMPRO uses a *conjugate gradient* method [40], usually considerably more efficient than alternatives such as 'steepest descent'.

If an atom a has forces f'_{la} acting on it in direction l, the atoms are then moved along a conjugate direction d'_{la} so that they are now located at

$$R'_{la} = R_{la} + wd'_{la},$$

where w is chosen using quadratic or cubic interpolation in order to minimise the free energy. The directions d'_{la} are chosen using

$$d'_{la} = f'_{la} - x d_{la},$$

where d_{la} was the previous search direction with a force of f_{la}, and x, initially zero, is set to

$$x = \frac{\sum_{la} f'_{la}(f'_{la} - f_{la})}{\sum_{la} f^2_{la}}.$$

This method is an efficient way of finding a system minimum, however it cannot guarantee finding the global system minimum. Schemes such as genetic algorithms are better for this, but are computationally too expensive for general use. The best way to ensure the result is really a global minimum appears to be attempting several relaxations from different starting structures.

It is also possible to perform constrained relaxations, for example to find saddle points for defect migration. This is discussed further in Section 6.2.1.

3.6 Calculation of Vibrational Modes

An important piece of information that can be extracted from these calculations and compared directly to experiment is the local vibrational modes of the defect. These are vibrations characteristic of the defect core that do not resonate with bulk modes and so remain localised at the defect site. Since these are directly dependent on the masses of the atoms involved, the influence of different isotope substitutions on these modes is an invaluable probe of the defect structure and content. These are primarily detected using Fourier Transform Infrared Spectroscopy (FTIR) through observation of IR absorption at different wavelengths, however they can also be measured as phonon sidebands to PL absorption and Raman Spectroscopy for the Raman active modes.

The vibrational modes of a cluster can be determined from the energetic double derivatives with respect to displacement of the atoms. Initially the cluster is fully relaxed until all atomic forces are zero. A given atom a is displaced along axis l by a distance ε (usually chosen to be 0.025 a.u.), and the self-consistent charge density recalculated. The forces on the atoms are then no longer zero, so for example, atom b along axis direction m will feel a force of $f^+_{mb}(l,a)$. If this is repeated with a displacement of $-\varepsilon$ then a new force $f^-_{mb}(l,a)$ is determined. Given these three points it is possible to now calculate the energy second derivatives up to second order in ε:

$$D_{la,mb} = \frac{f^{+}_{mb}(l,a) - f^{-}_{mb}(l,a)}{2\varepsilon}.$$

Once the full dynamical matrix is determined it can be diagonalised, scaled by the root of the relevant atomic masses, to give as eigenvalues the vibrational modes of the system. Note that this method includes some contribution from all even powers of ε and not just the quadratic term since the step size is not infinitesimal, and thus there will also be some anharmonic contribution included. These frequencies are thus sometimes referred to as quasi-harmonic, and this is discussed in more detail in Reference [37]. Surface H atoms are given a large enough mass (^{1000}H) that they are effectively fixed.

To calculate the double derivatives for every atom in the cluster would be too time consuming. Since most of the cluster atoms approximate to bulk and only act to perturb the localised vibrational modes, the derivatives are only calculated for the core defect atoms. The rest of the bulk material is included using a Musgrave-Pople valence force potential, discussed further in Section 2.8.2.

3.6.1 Induced Dipole Moments

The induced dipole moment is a useful quantity since it allows determination of η, the effective charge associated with a given vibrational mode of a defect. This is directly proportional to the absorption intensity of the mode, something that can be directly determined by experiment. If I is the integrated intensity of absorption then

$$I = \frac{2\pi^2\rho}{ncM}\eta^2$$

where c is the velocity of light, n the refractive index of the material, ρ the concentration of the impurity, and M its mass [41]. Previous calculations using our method have shown accurate determination of dipole moments for small molecules.

After determining a dynamical matrix from the energy double derivatives as described in the previous section, this is diagonalised to give a set of eigenvectors associated with each vibrational mode. These are scaled by the square root of the atomic masses as described above.

The atoms in the cluster are then shifted along their normal mode directions and the energy and adiabatic cluster dipole determined. This is repeated in a reverse direction, and the difference between these two gives the induced dipole moment of the cluster for that vibrational mode. η is simply the square of this value.

3.7 Discussion of the methods used in AIMPRO

AIMPRO is a real space Gaussian hydrogen-terminated cluster method, all of which bring advantages and disadvantages to the calculational procedure.

3.7.1 Density functional theory / Local Density Approximation

One of the greatest advantage of density functional schemes under the local density approximation is their reliance on the charge density as the fundamental variable, which greatly simplifies the amount of calculation required. However they consistently underpredict bandgaps (this is to some extent compensated by the use of clusters, see Section 3.7.3 below), and as a ground state theory they are restricted in their application to excited state problems. Various methods have been used to try and overcome this, notably constrained relaxation techniques to determine saddle points (see Section 6). In addition since the wavefunction eigenvalues are a computational tool and do not directly link to experimental eigenvalues, physical interpretation of these states is somewhat dangerous (although there is normally qualitative agreement between the two). Current work suggests Slater transition methods may provide a way of quantitatively relating the two [42].

Density functional theory under the local density approximation is well known for overbinding of molecules such as O_2, but can be greatly improved through the use of gradient corrections. AIMPRO actually does better than the standard DFT result, shown in Table 3.1. These results support the conclusion that we are obtaining good structural and vibrational data, but need to improve the accuracy of our total energies. However the magnitude of such errors normally decreases with system size. The local density approximation is good for smoothly varying systems, so given that other DFT/LDA calculations on oxygen/silicon structures have also had trouble with total energies (see, for example, Chapter 5), it suggests that part of the problem of innacurate total energies could be due to the local density approximation.

3.7.2 Real space

Calculating everything in realspace means that complex conjugate integrals can be greatly simplified (removing all imaginary terms), and it is much easier to directly interpret wavefunction data, etc. in terms of bonding and anti-bonding orbitals (given the caveat that they are Kohn-Sham wavefunctions, Section 2.4). It removes the need to fill the whole calculational real space with fitting functions, thus large vacuum regions (e.g. seperating molecular species from bulk surfaces) do not in-

Source	Binding Energy (eV)	O stretch mode (cm^{-1})	Bond Length (a.u.)
AIMPRO, spin			
averaged	8.79	1565	2.28
polarised	6.41	1581	2.28
Experiment	5.23	1580	2.28
Quoted DFT	7.54	1610	2.31
Hartree Fock	1.43	2000	2.18

Table 3.1: Properties of the O_2 molecule. AIMPRO values are obtained using the standard basis set described elsewhere in the thesis, *i.e.* six Gaussian fitting functions on each O atom for both charge density and wavefunction, with three more Gaussian functions on the bond centre for charge density and two for the wavefunction. All non-AIMPRO values from [43].

crease the size of a calculation; unlike k-space plane wave calculations. In addition it makes it easy to increase the local basis size around key atoms such as defect cores or O atoms where a large basis set is required to accurately model the charge.

In order to accurately model oxygen in silicon a large basis set is required. For real space methods such as this it is easy to locally increase the basis size. However for plane wave calculations the basis is determined for the whole supercell by the cut-off energy of the plane waves. Using bhs pseudopotentials [23] this is typically of the order of 15 Ry for bulk[1], and 5-12 Ry for surface calculations where the size of the supercell is necessarily larger in order to include a vacuum region [44]. However when including light elements such as carbon, nitrogen or oxygen this figure increases to something over 50 Ry, for example, recent calculations on bulk GaN used a cut-off of 150 Ry [45]. In order to increase the basis locally it is necessary to increase the energy cut-off for the whole calculation, leading to a rapid increase in calculation size. The only way to overcome the computational restriction of this is by changing the pseudo-potential, and using *ultra-soft* pseudopotentials. These are more slowly varying nearer the core and so are easier to fit. However they have a higher core energy and so are no longer norm-conserving, which leads to errors in self-consistency. In addition, ultra-soft pseudopotentials are less transferable and less able to cope with unusual structures. In the GaN calculation described above the use of ultra-soft pseudopotentials drops the required cut-off energy to 65 Ry. In general, structure converges quicker with cut-off than band structure [44].

[1] 1 Ry = 13.6058eV

3.7.3 Clusters

Clusters have many advantages over periodic systems, one of the strongest of these being the variety of symmetries that can be accurately reproduced by producing a cluster of the correct proportions. This makes symmetry constraining point defects much easier, and allows easy examination of extended structure defects such as line defects (the non-zero Burgers vector means that it is impossible to handle a single line defect in a supercell and they must be constructed in groups with cancelling Burgers vectors). Since a cluster is not periodically repeated it is easy to calculate induced dipole moments.

Charged species are easier to handle with clusters, since the number of electrons can be arbitrarily fixed. In a supercell an overall non-neutral charge would be infinitely repeated leading to an infinite system charge. This problem is avoided by adding an equivilent opposite uniform background charge to the supercell to maintain its charge neutrality. Such problems become worse when, for example, applying an external electric field; whereas with clusters a uniform field can be used, for a supercell it has to be a sawtooth field in order to keep the field continuous from one cell to its neighbour[46].

However, clusters also have several disadvantages. The main one of these is the potential for interaction between the cluster surface and the defect contained within. This problem is avoided when supercells are used, but instead is replaced with potential defect–defect interaction between neighbouring supercells. Finite cluster size can also constrain electron delocalisation. Hopefully defect–surface interaction should be minimised with the use of increasingly large clusters as we have access to more powerful computing facilities. Such surface effects can be seen in small polyaromatic carbon sheets used to simulate graphite, where the edge termination can impose a Kekulé structure across the sheet if it is too small. It is sometimes useful to repeat calculations in different cluster sizes to check the dependancy of the result on surface proximity. Alternatively repeating a calculation with surface hydrogen atoms held fixed and then allowed full relaxation gives an indication of surface effect depending on the degree of relaxation.

Using finite clusters increases the calculated bandgap of semiconductor material due to the surface confinement; this works against density functional theory which tends to decrease band gaps[2].

Zinc-Blende type clusters present an additional problem over their diamond lattice equivilents, due to the different valency of the two atomic species present. If such a cluster is atom centred and radially symmetric, it will terminate with an

[2] In the case of diamond these two effects almost cancel, giving an AIMPRO calculated bandgap of 5.0 eV in a 71 atom cluster, $C_{35}H_{36}$, compared to the experimental figure of 5.5 eV[47].

outer shell of only one atomic species, so the cluster will be non-stoichiometric. In order to have each covalent bond fully populated it is then necessary to charge the cluster (e.g. a 'neutral' 131 atom $In_{31}P_{40}H_{60}$ cluster has a charge of +9 to remove the excess P electrons). The alternative is a bond centred stoichiometric cluster. Although this does not have to be artificially charged, there will be a net cluster dipole due to the non-symmetric arrangement of atoms, and this leads to anisotropy in the bonds. The effect of these two types of cluster on the core defect structure and vibrational modes can be seen in Chapter 4.

The use of clusters with a fixed electronic charge means that it is sometimes possible to obtain unphysical solutions. For example with highly delocalised states such as effective-mass type states, these are normally very delocalised in physical systems. However in a cluster approach they are forced to remain within the confines of the cluster and hence can artifically occupy anti-bonding states, etc. For problems such as the thermal donors this can be surmounted by running the defects in their ionised charge state, and the effect of localisation can be tested by then checking the structural change when the donor electrons are added once more.

3.7.4 H termination

If a realspace calculation is performed without using either periodic boundary conditions[48] or an embedding technique then the only other alternative is to use clusters. A realspace cluster of atoms has a surface array of dangling bonds that need to be saturated in some way in order to remove them from the gap and correctly simulate bulk material. This is achieved by terminating the dangling bonds with H atoms, concentrating surface charge in correctly orientated chemical bonds.

Justification for the use of a H terminated cluster approach comes from the apparent scalability in chemical bonding from molecular species into semiconductor bulk. For example, disiloxane, $(SiH_3)_2O$ has a Si-O-Si vibrational stretch mode of 1107 cm^{-1}, only 2.5% less than that of O_i in Si[49]. It has a Si-O bond length of 1.634 Å[50], less than 1.5% longer than the bondlengths we calculate here for O_i in Si (see Chapter 6).

If the surface H atoms are allowed to move, the number of iterations required to reach a minimum energy structure is vastly increased. Surface H sits in a shallow potential and as the defect core relaxes there is a 'relaxation ripple' which runs out to the edge of the cluster. Often over half of the conjugate gradient iterations are spent moving the surface hydrogens with little overall change in the cluster energy or defect core structure. If the surface H atoms are not allowed to move then it is questionable to what extent the resultant cluster is in a fully relaxed ground state structure. Ideally there should be no motion of the cluster surface indicating minimal interaction with the defect at the cluster core, but in practise such clusters are prohibitively large. Fixing surface H atoms can lead to incorrect structures where there is a large volume change, e.g. substitutional carbon and VO_n centres.

A compromise between these two solutions that has been used for much of this work is to confine the hydrogen atoms in an additional quadratic 'spring' potential. This is a crude first approximation to the additional rigidity that would be provided if the cluster was embedded in bulk. Using a standard Si-Si bond spring constant gave a potential that was too restricting, so where used in these calculations, the spring strength is 5.0 eV/Å. Various other simple methods have been tried with less success.

3.7.5 Gaussian orbitals

Gaussian orbitals have the advantage that most integrals involving them can be evaluated analytically. They are a relatively localised basis set that are ideally suited to modelling rapidly varying functions such as wavefunction tail-off away from atomic cores such as oxygen. In this case the basis can easily be increased in the regions of a cluster where it is needed, whereas plane wave methods have to increase the plane waves throughout the supercell to achieve the same effect. However they do not individually approximate solutions to the Kohn-Sham equations in the same way as Slater orbitals. In addition, increasing the number of Gaussian orbitals used to fit the orbitals can quickly lead to *'over-completeness'*. In this case the system has functions which are very close with high overlaps, and can generate unphysical eigenvalues as a result.

States that are highly delocalised in space, for example weak bonding between distant molecular species, or fragments far from a surface, can suffer from the use of Gaussian orbitals, since they are by their nature localised. However they work well in large molecular clusters where often only a surprisingly small basis is required to accurately model the structure. Using fixed electronic charge means that it is impossible for a system to spontaneously lose or gain electrons. For example, O_2^- has a single very weakly bound electron that is distributed over a large space; this delocalisation is not correctly modelled by AIMPRO and the preference of the system to lose the electron is manifested in positive occupied eigenvalues.

3.7.6 Bond centred Gaussian fitting functions

The use of bond centred fitting functions has a dual purpose. Firstly it allows better modelling of the charge distribution in a bond centre, which is normally where the valence charge is localised in a cluster. Secondly it removes the need to include $d-$

orbital fitting functions on the bulk Si atoms. This is because the bond centred Gaussians act on their nearby atomic sites with s, p, d, etc. character, the strength of this effect tailing off with distance between the Gaussian and the atomic site and increasing quantum number. In order to maximise d-like character it is possible to arrange these fitting functions symmetrically around an atomic site so that their p-character exactly cancels; the minimum number of functions required for this is an icosohedral arrangement. We tried this for some of the oxygen diffusion calculations, where bonds are breaking and reforming and so it is not always obvious where bond centred functions should be located (see Chapter 6), however in practise they did not increase the accuracy of the calculation and significantly slowed calculation time.

However bond centred Gaussian functions bring associated disadvantages. Such additional fitting functions will locally improve the wavefunction and charge density modelling, thus lowering the cluster energy. This makes it harder to compare different structures, where the energy may change simply because the bond centred functions are no longer in such optimal locations. This has proved to be a particular problem in the case of oxygen in silicon where the cluster energy is critically dependant on the number and location of bond centred fitting functions. They often have little effect on final structure but the improved local charge description can improve the local vibrational modes.

3.8 Application of AIMPRO to the oxygen in silicon problem

Oxygen in silicon is a problem well suited to modelling using AIMPRO. Oxygen has a rapidly varying potential close to its atomic core which is much more suited to fitting with real space Gaussian functions than sets of plane waves. Oxygen related point defects in silicon exhibit a variety of different defect symmetries which AIMPRO is able to constrain accurately. Our use of additional fitting functions placed at bond centred sites means that d orbital fitting functions are not required on the Si atoms, greatly simplifying the computational problem. Finally, FTIR has proved to be one of the most important experimental tools used to study these problems, and we are able to theoretically calculate both the vibrational modes and induced dipole moments of these modes. In combination with isotopic shift data these are a good way of comparing experimental data with the theoretical models.

Chapter 4

Hydrogen in III-V materials

"You've been in my life so long... I can't remember anything else."

Ripley, Alien 3 (1992)

Hydrogen plays many different roles in III-V materials. It is traditionally thought of as a passivating element, able to diffuse rapidly through the lattice and terminate dangling bonds associated with other point defects, resulting in electrically inactive complexes. However hydrogen plays many more roles than this, and is actually able to form electrically active donors when fully hydrogenating Group-III vacancy sites (see below). In order to have a better understanding of the role played by hydrogen in these systems we examine two sets of problems; hydrogenated vacancies in InP, and passivated Group II elements in InP and GaAs.

This chapter initially examines how hydrogen can complex with indium vacancies in InP, producing a variety of donor and acceptor-like defects depending on the number of hydrogen atoms present. It has been suggested that iron in InP is compensated by a donor, related to the 2316 cm^{-1} local vibrational mode and previously assigned to the fully hydrogenated indium vacancy, $V_{In}H_4$. We find that $V_{In}H_4$ acts as a single shallow donor. It has a triplet vibrational mode close to this value, consistent with this assignment. We also analyse the other hydrogenated vacancies $V_{In}H_n, n = 1, 3$ and determine their structure, vibrational modes, and charge states.

Substitutional Group II impurities act as acceptors in InP, but can be passivated by hydrogen. We investigate the passivation of beryllium by hydrogen and find that the hydrogen sits in a bond centred site and is bonded to its phosphorus neighbour. Its calculated vibrational modes are in good agreement with experiment. We then look at H passivation of Be and Mg in InP and GaAs, to examine trends with variation in host material.

In order to accurately calculate the vibrational modes a set of parameters for the Musgrave Pople potential is required for the host material (see Section 2.8.2).

As our group did not have this for InP, the first task was to calculate them. This also allows calculation of the bulk phonon dispersion curves, which provides a way of evaluating the quality of the parameterisation.

4.1 Background

InP has the potential to be an extremely important III-V semiconductor material for high speed electronic devices. It has both a higher mobility and thermal conductivity than GaAs [51], making it suitable for high-speed switching circuits, and its direct band gap makes it highly suitable for optoelectronic devices such as field-effect transistors. It also has great potential for use in single-junction photovoltaic cells, as its energy gap is close to that required for optimum solar radiation to electricity conversion [52]. It is also relatively radiation resistant making it suitable for solar cell use in space.

Trace impurities can lead to a deterioration in the electrical properties of InP when such impurities are electrically active. For this reason there has been recent interest in the possibility of passivating these residual impurities using hydrogen [53, 54]. Hydrogen is also routinely used during technological processing of InP. A hydrogen plasma is used during the fabrication of high-speed InP-based devices and methane is used during reactive ion etching [55]. For all of these reasons, a detailed understanding of hydrogen related defects in InP is required.

4.1.1 Electrical effects and the role of Fe

There is great interest in the electrical effects associated with heavy iron doping of InP. Iron doping is used to produce mid-gap levels, allowing the use of InP as a semi-insulating substrate [56, 57]. However Fe has many unusual properties in InP which are not yet fully understood. In particular, the charge compensation mechanism of iron has not yet been fully explained [58].

It is known that in InP:Fe, the Fe substitutes for In and primarily adopts the neutral Fe_{In}^{3+} state (labelled 3+ to denote its oxidation state). However, some of the iron is compensated, forming Fe_{In}^{2+}, and metallic donor defects are not present in sufficiently high concentrations to account for this. Therefore it has been proposed that another donor must be present, however its identity is not established. Recent work has shown a correlation between the intensity of the 2316 cm^{-1} LVM and the Fe_{In}^{2+} concentration [58], suggesting that this defect could be the unknown donor.

When InP is grown by the LEC method using wet boric oxide as encapsulant, the water vapour dissociates and hydrogen is incorporated into the material. The most intense H-related mode in as-grown crystals occurs at 2316 cm^{-1} (found in InP crystals both with and without Fe), and has been associated with $V_{In}H_4$ [59].

This assignment has been made for several reasons. Uniaxial stress measurements show the centre to have T_d symmetry [59], consistent with $V_{In}H_4$. The LVM lies close to that of other P-H related modes, and this is consistent with results obtained for VH_4 in Si [60]. The defect can be created by proton irradiation [61], which also leads to the creation of the required In vacancies.

H/D mixed defects have not been observed by mixed H+/D+ implantation because the signal to noise ratio of published spectra [61] is not high enough, and the FWHPs of the proton- and deuteron-induced LVMs are too broad to resolve the individual components. This last point is probably the origin of the apparent decrease in intensity of the 2316 cm^{-1} line reported with mixed implantation, since there are several components near 2316 cm^{-1}. This is not the case for the LVM at 2202 cm^{-1} tentatively assigned to $V_{In}H$, and this contrasting result also supports the assignment.

Other LVMs in SI InP include H-acceptor modes (mainly (Zn,H)) due to inadvertent contamination, as well as other LVMs whose origin is not clear. In slightly n-type InP:Fe, the (Zn,H) LVM is not observed, but another LVM at 2285 cm^{-1} is always observed. Annealing of LEC InP produces LVMs attributed to partially hydrogenated vacancies, but these are much lower in intensity than the line at 2316 cm^{-1}.

There has been no theoretical examination of these defects performed to date. We therefore investigate the structure, local vibrational modes and charge states of a range of hydrogenated vacancies, $V_{In}H_n, n = 0, 4$.

4.1.2 Group II impurities in InP

Group II residual impurities are normally present in InP at low concentrations. They are known to act as acceptors, producing shallow levels in the gap [62]. It has been observed that the presence of hydrogen is able to passivate these impurities, and it is only recently that there have been attempts to fully understand this process [59, 53].

Experimental investigations of H- and D-plasma passivated Be, Mg, Cd and Zn have been performed [59]. Infra-red absorption studies show hydrogen local vibrational modes (LVMs) which lie in the frequency range where P-H stretch modes are expected ($\approx$2200 - 2300 cm^{-1}). This suggests that hydrogen bonds to phosphorus atoms in the lattice instead of the acceptor impurities.

Uniaxial stress measurements show that these acceptor – hydrogen defects have C_{3v} symmetry [59]. Therefore, since these Group II acceptors are known to substitute for indium, hydrogen must lie somewhere along the ⟨111⟩ acceptor-P direction. Passivation with a mixed H,D plasma produces no peak splitting, suggesting that there is only one hydrogen atom present in each defect. It was suggested that passivation occurred through a chemical recombination, allowing a lower coordination for the Group II acceptor atoms and tying up the resultant dangling bond on the neighbouring phosphorus atom [59].

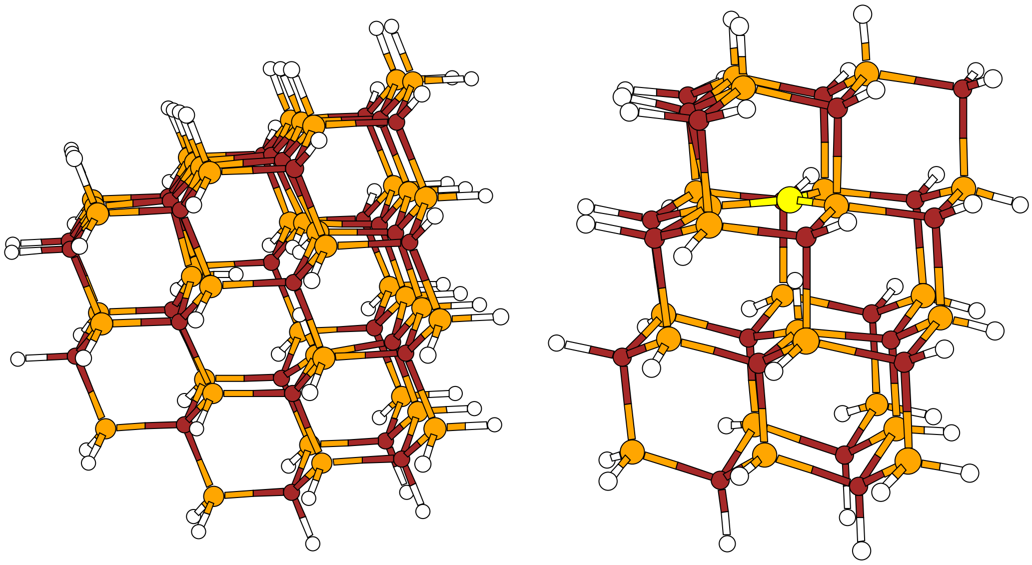
We examine several possible Be-H complexes in InP in order to determine the lowest energy structure and the passivation mechanism of the hydrogen. We then examine how H passivates Be and Mg in InP and GaAs, to examine trends with variation in impurity or host material.

4.2 Method

Initial calculations on pure InP used an 86 atom bond centred cluster, $In_{22}P_{22}H_{42}$ with C_{3v} symmetry. We fitted a Musgrave Pople interatomic potential to the pure InP dynamical matrix and used this to determine the bulk phonon modes (see Section 2.8.2). We used a combination of charge neutral Gaussian orbitals with highly localised charge varying Gaussian orbitals near the atom cores (see Section 3.2). Eight Gaussian orbitals were used to fit each In and P atom for charge density, with a similar number for the wavefunction. Two Gaussian orbitals were used to model the wavefunction of the surface H atoms, and three for their charge density. Bond centred fitting functions were placed on all bonds between the inner eight atoms. Symmetry was maintained at C_{3v} throughout. Hereafter for brevity the number of fitting functions used will be listed after the element name, so that the basis sets described above would be referred to as In(8,8), P(8,8) and H(2,3).

In order to examine the hydrogenated vacancies, we generated a 131 atom hydrogen terminated InP cluster, $In_{31}P_{40}H_{60}$. The central In atom was replaced by up to four hydrogen atoms (see Figure 4.1a). Symmetry was only enforced for $V_{In}H_4^+$ (T_d). Two electrons were added for each bond so that the resultant cluster charge was +9. In order to keep the cluster charge constant, the structures optimised were $V_{In}H_4^+$, $V_{In}H_3$, $V_{In}H_2^-$, $V_{In}H^{2-}$ and V_{In}^{3-} (Charge state is discussed further in Section 4.4). The electronic spin component was averaged. The basis sets used were In(8,8), P(4,5) and H(2,3). Bond centred Gaussian orbitals were placed on the twelve In-P bonds neighbouring the vacancy.

For examination of the Be-H complexes we generated an 86 atom hydrogen terminated InP cluster, $In_{22}P_{22}H_{42}$. This is In-P bond centred. The centre In atom was replaced with Be, and a H atom inserted in one of three different sites; anti-bonded behind the Be, anti-bonded behind the P, and in the bond centre between



a. b.

Figure 4.1: Clusters used in the InP work. (a) 134 atom $In_{30}P_{40}H_{64}$ containing $V_{In}H_4^+$, (b) 88 atom cluster $BeIn_{21}P_{22}H_{43}$ containing H passivated Be

Be and P. This resulted in the cluster $BeIn_{21}P_{22}H_{43}$ (see Figure 4.1b). Symmetry was constrained to C_{3v}. The basis sets used were In(8,8) and P(8,8) for the seven core InP atoms, In(4,5) and P(4,5) for the rest, Be(6,6) and H(3,4). Bond centred Gaussian orbitals were placed between the inner nine atoms.

Cluster choice is discussed further in Section 3.7.3.

4.3 Pure InP

Initial investigations centred on pure InP, in order to determine vibrational frequencies and a phonon dispersion curve, and to obtain a set of Musgrave Pople interatomic potentials for frequency calculations in the defect clusters.

After relaxation, we obtained central InP bond lengths of 2.480Å along the $\langle 111 \rangle$ axis parallel to the C_{3v} axis, and 2.421Å otherwise (an error of 2.4% and 4.7% respectively, as the experimental bond length is 2.54 Å[49]). The bond elongation in the $\langle 111 \rangle$ direction is due to the intrinsic dipole moment of a bond centred cluster. The nearest neighbour P-In-P bond angle was 111.5°, and 107.3° when using the In-P bond parallel to the C_{3v} axis; these are within 2° of the tetrahedral angle.

We next fitted the Musgrave Pople potential to the derivatives for the inner eight atoms (this is described further and the potential given in Section 2.8.2). The coefficients are shown in Table 4.1.

A supercell calculation of the full phonon dispersion curve was performed using

Atom	k_r	k_θ	$k_{r\theta}$	k_{rr}	$k_{\theta\theta}$
In	7.187	0.088	-0.125	0.283	0.013
P	7.187	0.154	-0.156	0.917	0.065

Table 4.1: Parameters for Musgrave-Pople Potential for InP in eV/Å^2, $r_0 = 2.421$

Location	Mode	Neutron Diffraction (300K)[63]	Raman Scattering (300K)[64]	Far I.R. (20K)[65]	Theoretical
Γ(0,0,0)	LO		345.4	351	354.80
	TO	307 $\pm$ 7	303.3		354.80
X(1,0,0)	LO	332 $\pm$ 3		328.5	324.59
	TO	324 $\pm$ 7		326.5	317.46
	LA	194 $\pm$ 10		190.5	186.86
	TA	68 $\pm$ 3		67.5	69.59
$L(\frac{1}{2},\frac{1}{2},\frac{1}{2})$	LO	340 $\pm$ 10		340.5	(5,6) 335.27
	TO	317 $\pm$ 5		315.5	(4) 314.03
	LA	167 $\pm$ 3		167.5	171.19
	TA	55.0 $\pm$ 0.7		53.5	49.49

Table 4.2: Phonon frequencies for pure InP (cm^{-1})

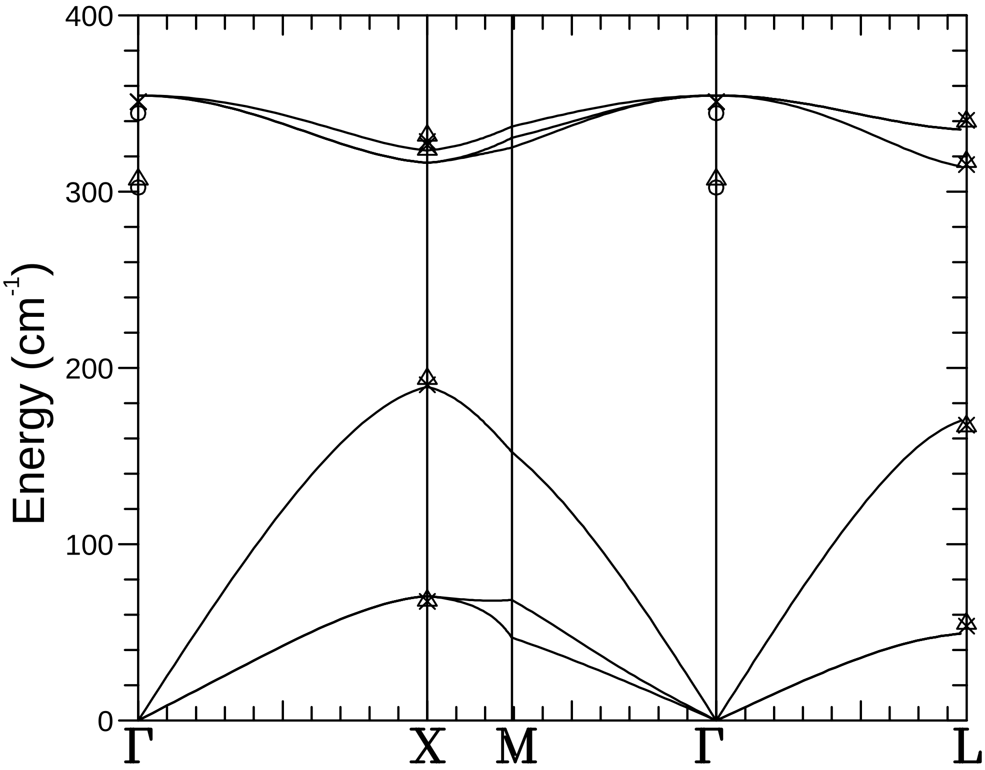


Figure 4.2: Calculated Phonon Dispersion Curve for InP. ×=Koteles *et al* [65], ○=Hilsum *et al* [64], △=Borcherds *et al* [63]

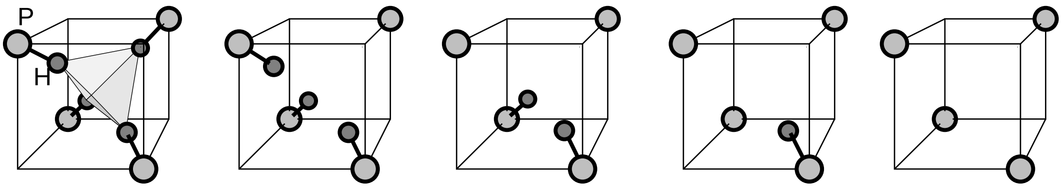


Figure 4.3: Schematic diagrams showing $V_{In}H_n, n = 4, 0$. The tetrahedral symmetry is shaded in $V_{In}H_4$.

the Musgrave-Pople interatomic potential. This is given in Figure 4.2 along with various experimental values. Selected values from this curve are given in Table 4.2. We do not obtain any LO-TO splitting because of an absence of long range electric field effects in the potential. The highest bulk phonon modes are 354.80 (351), 324.59 (328.5), and 335.27 cm^{-1} (340.5) at Γ, X, and L respectively (the experimental modes are in brackets[65], at 20K), thus our frequency errors are ≈ 4 cm^{-1}. The calculated band gap for InP was 0.84 eV, compared to an experimental value of 1.34 eV deduced from optical data[66], and is in good agreement with a previous LDF value of 0.8 eV[67] (see Section 4.2 above). Our calculated band gap is underestimated, which is unusual for a cluster calculation.

We conclude that our method gives bond lengths and local vibrational modes within an acceptable error range.

4.4 Hydrogenated Vacancy Centres

Schematic diagrams of the hydrogenated vacancy centres are shown in Figure 4.3. The top Kohn-Sham eigenvalues of the hydrogenated vacancies are shown in Figure 4.4. These show that $V_{In}H_4$ is a single shallow donor. Each of the defects considered has a single shallow level close to the conduction band. We chose cluster charge states to empty this level, since in practise it would thermally depopulate. $V_{In}H_4$ contains one electron in this level when neutral and hence is a single shallow donor. All of the defects also possess a t_2-like state above the top of the valence band which splits as the number of H atoms decreases.

As each hydrogen atom is removed from the vacancy, an electron is removed from the highest occupied state. The resulting defects can act as acceptors that fill the t_2 -level when ionised. This level is already filled in the case of $V_{In}H_3$ so this defect is electrically neutral. However $V_{In}H_2$, $V_{In}H$ and V_{In} will behave as single, double and triple acceptors respectively.

The t_2-level is split due to the lower symmetry of the partially hydrogenated vacancies. $V_{In}H_4^+$ has T_d symmetry and therefore the level does not split. $V_{In}H_3$ has C_{3v} symmetry, so the t_2 splits into an $e-$ and a_1- level. $V_{In}H_2^-$ has C_{2v} symmetry

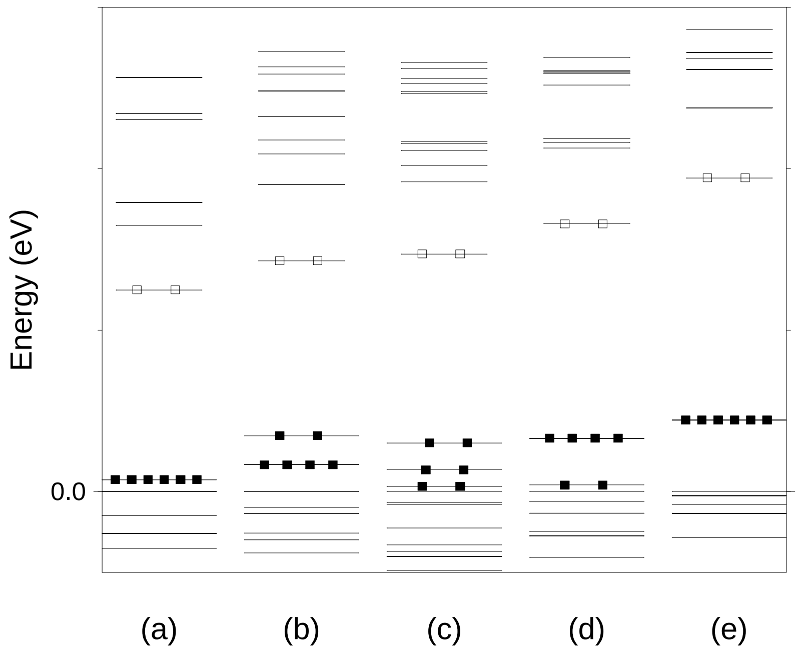


Figure 4.4: Top Kohn-Sham eigenvalues of the hydrogenated vacancies, (a) $V_{In}H_4^+$, (b) $V_{In}H_3$, (c) $V_{In}H_2^-$, (d) $V_{In}H^{2-}$, (e) V_{In}^{3-}. Filled boxes indicate electrons and empty boxes indicate holes. The eigenvalues have been arbitrarily shifted to align the highest filled level below the t_2-like state with zero.

which leads to three separated singlets. $V_{In}H^{2-}$ and V_{In}^{3-} are directly comparable with $V_{In}H_3$ and $V_{In}H_4^+$, with the dangling and hydrogen terminated bonds reversed, in addition to the $a - e$ level ordering. This t_2-like state gradually moves upwards away from the valence band as hydrogen atoms are removed. In V_{In}^{3-} the t_2-level is quite deep in the gap.

The indium vacancy, V_{In}^{3-}, fits the trend of increasing acceptor character with decreasing number of hydrogen atoms present. Previous calculations [68] also show V_{In} to be a triple acceptor.

The local vibrational modes for $V_{In}H_4$ with pure and mixed isotope composition are shown in Table 4.3. The highest IR visible mode for $V_{In}H_4^+$ at 2356.4 cm^{-1} is in excellent agreement with experiment (2315.6 cm^{-1}, a 1.8% error), as is the 1690.8 cm^{-1} mode for $V_{In}D_4^+$ (experimental value of 1683.4 cm^{-1}, an error of 0.4%). This small drop in error with deuteration suggests only limited anharmonic character in the bonding [37].

The LVMs for the range of hydrogenated vacancies, $V_{In}H_n, n = 0, 4$ are shown in Table 4.4. The results show that as the number of H atoms in the vacancy increases, the P-H bonds are shorten and the vibrational modes increase. This is due to the compressive effect of the other hydrogen atoms on each P-H group, coupled with the removal of dangling bonds from 'unsaturated' P atoms which would have acted

H_4	H_3D	H_2D_2	HD_3	D_4
* 2387.85	2380.18	2372.34	2364.31	* 1713.01
T 2356.40	D 2356.32	2356.24	1707.26	T 1690.82
D 618.57	1696.43	1701.74	D 1690.95	D 445.34
T 565.87	D 603.12	1691.08	D 549.65	T 411.94
T 408.69	565.71	595.02	D 428.92	
	D 508.72	556.04	412.03	
	408.69	541.80		
		481.28		
		425.83		

Table 4.3: Calculated LVMs for the fully hydrogenated vacancy in InP, $V_{In}H_nD_m^+, n + m = 4$. Note that **: IR Inactive, T: Triplet, D: Doublet*

to attract the hydrogen away from its phosphorus neighbour. This is consistent with previous results obtained for Si [60, 69].

The shift in P-H length from $V_{In}H_4$ to $V_{In}H$ is only 2.1% (from 1.419 Å to 1.450 Å), but it leads to a 9.9% shift in vibrational mode (2387.8 cm^{-1} to 2150.7 cm^{-1}). The calculated H- stretch modes of the partially hydrogenated vacancies could account for a group of experimentally observed vibrational modes lying between 2200 and 2290 cm^{-1} [59, 70].

High temperature annealing of InP:Fe reduces the concentration of $V_{In}H_4^+$. This is due to the partial dissociation of this centre and, up until now, only $V_{In}H$ has been identified with certainty. It is shown here that $V_{In}H_2$ and $V_{In}H$ are acceptors. In annealed high-resistivity material they should cause a drop in the Fe^{2+} concentration. This will be in addition to the decrease in concentration of Fe^{2+} due to loss of the $V_{In}H_4$ donor.

In contrast, Bardeleben *et al* [71] found that thermal annealing of Fe doped InP in the range 660-820 °C led to an increase in Fe^{2+} concentration. They suggested this was due to the formation of some unidentified deep thermal donors. This could be associated with other hydrogen complexes. However, from this work it seems unlikely that hydrogenated vacancies are responsible.

In summary, the fully hydrogenated vacancy, $V_{In}H_4$, acts as a single donor due to a partially filled singlet near the top of the gap; thus $V_{In}H_4$ will compensate Fe_{In}^{3+} in InP. Removal of hydrogen atoms from the vacancy leads to increased acceptor character as the triplet state starts to empty.

Defect	Local Vibrational Modes					Symmetry	Bond
	Exp [61]	[59] ([70])	Calculation				Length
$V_{In}H_4^+$	2315.2	2315.6	2387.8*	2356.4T	618.6D	T_d	1.419
$V_{In}D_4^+$	1683.4		1713.0*	1690.8T	445.3D		
$V_{In}H_3$			2324.1	2286.3D	695.3D	C_{3v}	1.429
$V_{In}D_3$			1667.6	1640.8D	498.2D		
$V_{In}H_2^-$			2256.2	2216.8	730.0	C_{2v}	1.439
			617.7	611.8	602.0		
$V_{In}D_2^-$			1619.4	1591.6	523.0		
			446.4	444.4	438.8		
$V_{In}H^{2-}$	2201.7	(2202.4)	2150.7	644.9	644.7	C_{3v}	1.450
$V_{In}D^{2-}$	1603.8		1544.7	467.5	467.4		

Table 4.4: Local vibrational modes (cm^{-1}), symmetry, and calculated bond lengths (Å) of hydrogenated vacancies in InP **: IR Inactive, T: Triplet, D: Doublet*

Experimental		Bond Centred		Anti-Bonded behind P	
H	D	H	D	H	D
2236.5	1630.9	2287.7	1639.9	2469.9	1770.9

Table 4.5: Calculated and Experimental LVMs for H passivated Be in InP (cm $^{-1}$)

4.5 Passivation of Group II Impurities

Be was used to determine the preferential site for hydrogen in the passivated Group II acceptor-hydrogen complex. In order to maintain C_{3v} symmetry along the (111) axis, hydrogen can sit in one of only three sites: at the P-Be bond centre, or one of the two anti-bonded sites behind P or Be (see Figure 4.5).

The structure with H lying in an anti-bonded site behind Be was 1.77 eV higher in energy than the bond centred case, and so was energetically unfavourable. The structure with H anti-bonded behind P was only 0.13 eV higher in energy than the bond centred case, close to the expected errors in our calculation. In all cases the gap was cleared of any acceptor levels by the presence of the hydrogen atom, showing that it has passivated the acceptor activity. In order to determine whether the bond centred or anti-bonded behind P model is correct, it is necessary to examine the LVMs, shown in Table 4.5.

As can be seen, the calculated LVMs for the bond centred case are in good agreement with experiment (errors of 2.3% and 0.6% for H and D respectively)

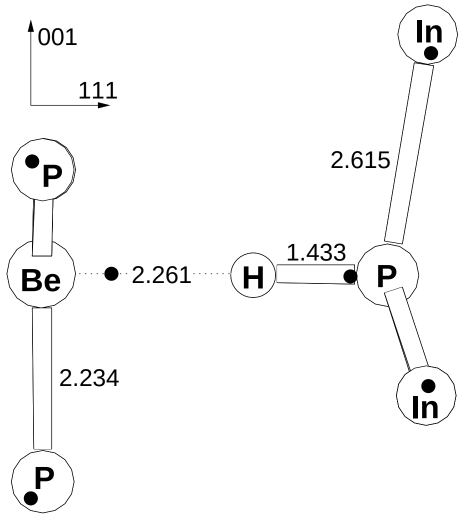


Figure 4.5: Calculated structure for H passivated Be in InP (bond lengths in Å). Black dots mark ideal lattice positions.

whereas the anti-bonded results are poor (10.4% and 8.6% for H and D respectively). From this we conclude that H sits in the bond centred location.

The defect is shown in Figure 4.5. The large H-Be separation shows a lack of bonding between them; the H atom is strongly bonded to P. The P-H bond length is comparable to the P-H bond lengths in the hydrogenated vacancy, and compares to an experimental bond length for the P-H_3 molecule of 1.43 Å[49]. The Be-P bond lengths are similar to crystalline Be-P lengths in Be_3P_2 (2.16 Å). The impurity is displaced along the $\langle 111 \rangle$ direction away from the hydrogen atom to achieve this, until it is co-planar with the phosphorus atoms.

Therefore we conclude that substitutional Be acceptors are passivated by a single H sitting at a bond centred position between the Be atom and a nearest neighbour P atom. We find a LVM at 2287.71 cm^{-1}, 2.3% larger than the experimental frequency. There is also an undetected wag mode at 644 cm^{-1}, similar to the case of the C-H defect in GaAs and AlAs [37, 72].

4.5.1 Varying acceptor and host – InP, GaAs, Be and Mg

We next examined H passivation of Be and Mg in InP and GaAs to examine the effect of changing acceptor and host type. We used Ga(4,5), As(4,5), In(4,5), P(4,5) and H(2,3) for surface H atoms. The core H was H(3,4), with Mg(10,10) and Be(6,6). Bond centred Gaussian orbitals(2,3) were placed in every bond centre of the cluster, and symmetry was constrained to C_{3v}. Clusters were 87 atoms as in the previous section unless specified otherwise.

In all cases the energy difference between the bond-centred site and the anti-bonded site behind the Group V element (either P or As) was normally tiny and much smaller than the expected error in our calculations. Again, in every case the Group II element dropped back to a planar position in order to form three strong bonds with its Group V neighbours, while the H atom formed a single strong bond with its Group V neighbour, which also lay co-planar with its Group III neighbours.

Vibrational mode agreement was not quite as good as that of InP:Be-H, however in all cases the bonds centred LVMs were in reasonable agreement with experiment, and were more accurate than the anti-bonding modes. Previous work on GaAs:Be-H by Briddon *et al* [73] produced better vibrational mode agreement, and this was probably due to their choice of cluster. 86-atom clusters have the advantage of being stoichiometric, but then have an inherent cluster dipole moment which will compress the core bond artificially. However atom centred clusters such as the 131-atom cluster have large charges to maintain the correct electron bond populations, and this charging will also effect the local bond lengths. Although much of Patrick's work used smaller 56 atom clusters, these have the advantage of being atom centred and also only having a small cluster charge, leading to more accurate core bonds. For comparison I have included results for both an 87 atom and 132 atom cluster of InP:Mg, and it can be seen that both the BC and AB modes converge on the experimental figure with the increase in cluster size.

In the only case where lower frequency modes are known, GaAs:Be, there is excellent isotopic agreement with the BC site, unlike the AB site. This is the only case theoretically where the Group-II metal A_1 mode lies *above* the H wag mode; in all other cases the lower quoted frequency is that for H wag. This is reflected in the tiny H/D isotopic shift of the 587 cm^{-1} mode of only 1 cm^{-1} for GaAs:Be.

These results therefore suggest that both Be and Mg are passivated in InP and GaAs through the addition of a single H atom to the bond centred site. This bonds tightly to the Group-V atom, which moves co-planar with its three Group-III neighbours, while the acceptor atom also moves away to lie co-planar with its Group-V neighbours. As an additional check on the method, test calculations on the same defects in GaN gave good vibrational modes for Mg with H in the AB location, as well as predicting this to be more stable than the BC site, consistent with current theories [74].

4.6 Conclusions

We have shown that $V_{In}H_4$ is a single donor forming $V_{In}H_4^+$, with a donor level in the upper part of the gap. It is tetrahedrally symmetric and the calculated vibrational modes are in excellent agreement with experimental data for the 2316 cm^{-1} mode.

Material	Experimental		Theoretical			
			BC		AB	
InP:^{24}Mg	2366.4	(646.1)	2308	(649)	2418	(683)
			644	(181)	581	(167)
InP:^{24}Mg	2366.4	(646.1)	2357	(664)	2381	(670)
132 atom			519	(143)	869	(246)
GaAs:^{9}Be	2037.1	(565.9)	2128	(614)	2257	(651)
	555.7	(2.1)	587	(1)	699	(105)
			470	(135)	594	(100)
GaAs:^{24}Mg	2144.0	(596.7)	2175	(628)	2265	(654)
			526	(149)	680	(200)
GaAs:^{9}Be (Pat)						
56 atom	2037		2083		2069	
			346,383		852,853	
86 atom			2163			
			495			
102 atom			2018			
			301			

Table 4.6: Local vibrational modes of H-passivated Mg,Be in InP and GaAs, with H in the BC or AB site neighbouring the P/As (all modes in cm^{-1}) - figures in brackets show drop with D isotope). The lower wag-type modes are doublets, the higher stretch mode a singlet. Clusters are 87 atoms unless specified otherwise.

$V_{In}H_3$ is shown to be electrically inactive, but the partially hydrogenated vacancies, $V_{In}H_2$ and $V_{In}H$ act as acceptors, forming $V_{In}H_2^-$ and $V_{In}H^{2-}$ with acceptor levels in the lower gap region. $V_{In}H_4$ compensates Fe_{In}^{3+} leading to the formation of Fe_{In}^{2+}. The donor electron could also be trapped by the partially hydrogenated vacancy centers, $V_{In}H_2^-$, and $V_{In}H^{2-}$, or even V_{In}^{3-}, however these are present in much lower concentrations than $V_{In}H_4^+$.

Darwich et al.[59] have observed a number of vibrational modes in their InP:Fe samples between 2202.39 cm^{-1} and 2282.8 cm^{-1} which have been attributed to H-related complexes[53], and we suggest that the partially hydrogenated vacancy centers could account for at least some of these modes.

In agreement with experiment, the calculations show that H is able to passivate Group II acceptors such as Be. It lies bond centred between Be and P, and forms a strong bond with P rather than Be. Both the Be atom and its P neighbour drop away from each other along the C_3 axis so as to lie co-planar with their three other neighbouring atoms. This result seems to hold true for both Be and Mg, in InP as well as GaAs. Since it appears that H passivates Mg in the AB site for GaN it would be interesting to extend this current study to further acceptor impurities such as Zn and Cd, as well as other III-V compounds, to determine any trend in passivation site with host material and acceptor type. In addition, fully hydrogenated vacancies should act as donors in other III-V materials and this study could be extended to examine that.

Chapter 5

Vacancy - Oxygen complexes in Silicon

5.1 Introduction

In this chapter we examine the interaction between oxygen and vacancies in Si. Vacancies are introduced into Si in large quantities through irradiation treatments such as electron irradiation or ion implantation. Although most of the resultant vacancies and self-interstitials immediately recombine, large quantities remain. Annealing causes a mixture of recombination and complexing with other impurities in the crystal. In Cz-Si grown wafers there is a radial band of high vacancy concentration with radially varying vacancy and self-interstitial concentration. In all of these cases it is important to know how vacancies behave in the presence of oxygen. An attraction between V and O_i would be expected due to lattice strain compensation arguments.

The most common oxygen–vacancy complex is the VO centre (or 'A' centre), consisting of a single off-site substitutional oxygen. Annealing at 300°C leads to the gradual loss of VO centres, as they are replaced, initially with a defect with vibrational modes at 914 and 1000 cm^{-1}, and then with a defect with a mode at 889 cm^{-1}. One of the primary objectives of this section is to positively identify this defect, and we show it to be VO_2, a prediction which has since been verified by experiment. Annealing at 450°C and above leads to higher order V_m-O_n defects that form from the 889 cm^{-1} defect.

We show here that the VO_2 defect has D_{2d} symmetry and only one calculated O-related high frequency IR active mode at 807 cm^{-1}. The VO_3 defect has three high frequency IR active modes. The V_2O defect has one such LVM, at lower frequency. These results provide strong support for the assignment of the 889 cm^{-1} (300 K) local vibrational mode to VO_2.

We also discuss the annealing behaviour of the VO_n centres including intermediate species that form between VO and VO_2, the role of oxygen dimers, and the possibilities for higher order VO_n species.

5.1.1 The VO Centre

Irradiation of Si at room temperature produces mobile vacancies and interstitials. The former trap O_i yielding A-centres, or VO complexes [75]. These are the primary oxygen-related defect in radiation damaged Cz-Si, and form irrespective of the irradiation method (electrons, protons, neutrons, α-particles, γ-rays, etc.) The structure of the negatively charged centre has been established by EPR [76], whereas uniaxial stress studies on the local vibrational mode (LVM) at 836 cm^{-1} (4.2K) has shown that the neutral defect has a similar structure [77]. This consists of an O atom bridging a pair of Si neighbours of the vacancy (Figure 5.1). The defect has associated vibrational modes at 836 and 885 cm^{-1} for the neutral and negatively charged species respectively. The defect has an acceptor level at $E_c - 0.17$ eV, possibly due to the strained Si-Si bond between the other pair of Si atoms bordering the vacancy. The lines are, however, broad and no ^{29}Si isotope shifts are seen [78]. The reorientation energy of the centre is surprisingly low, at only 0.38 eV [75]. Their barrier to migration is 1.46$\pm$0.29 eV [79] (also given as 1.86 eV [80, 31]), allowing them to migrate at lower temperatures than O_i.

5.1.2 The formation of VO_2

The OV defects anneal out at around 300°C and a new line appears at 894 cm^{-1} (4 K) which then grows in intensity [141]. This band shifts to 889 cm^{-1} at room temperature. The original proposal [141] was that the absorption is due a VO_2 defect, *i.e.* two O atoms sharing a vacancy, each bonded between two of the Si vacancy neighbours (Figure 5.2). The symmetry is then D_{2d}. There are three processes which contribute to the demise of VO.

1. Initially a fast process in which the $[O_i]$ rapidly increases. This may be related to the break up of aggregates of Si_i introduced during the irradiation. This then leads to the reaction: VO + $Si_i \rightarrow O_i$ [81, 82, 79].

2. A second slower process. This is assumed to be: VO + $O_i \rightarrow VO_2$ and caused by VO diffusion to the immobile (at 350°C) O_i. This assignment is based on the 1.86 eV activation energy [80], the migration barrier for VO diffusion.

3. Not all VO is lost by this second process and a third involves VO being captured by some (unknown) defect and producing O_i. Roughly half of VO can meet their fate in this third process.

Support for the VO_2 assignment of the 889 cm^{-1} defect comes from its IR-absorption intensity which is approximately proportional to $[O_i]^2$ when VO has disappeared. However, there is no loss of O_i, as measured by IR-absorption, during the slow stage when the 889 cm^{-1} defect is being formed [81]. This initially provided the necessity for the third formation mechanism given above, however recent studies by Londos et al [79] show two distinct activation energies in the slow process, with a switch in dominance at about 360°C. In addition they showed that above 380°C VO_2 absorption continues to increase, even though VO levels have almost stabilised.

Further, early uniaxial stress studies on the 889 cm^{-1} defect [77] concluded that its symmetry was lower than D_{2d}; unless the symmetry was D_{2d}, two O-related LVMs would be expected. However, prompted by our theoretical investigations, Bech Neilsen *et al* [83] re-examined the defect with uniaxial stress and showed it to possess either C_{2v} or D_{2d} symmetry. Although it is difficult to distinguish between the two, they have recently performed a range of stress-induced dichroism experiments that were able to unambiguously show D_{2d} symmetry [83] (see below). Finally, for mixed ^{16}O-^{18}O samples, only two LVMs were detected and not three [84, 85], with isotopic shifts indicative of a single oxygen atom. This shows that the two O-modes in this model for the 889 cm^{-1} centre are decoupled even though their separation can only be a few Å. This objection is not fatal as earlier modelling studies [86] found that the O atoms were decoupled in the LVMs.

The difficulties [87] described here have led to an alternative assignment of the 889 cm^{-1} LVM to V_3O [78]. This defect would possess only one LVM and possess C_{1h} symmetry. Presumably, in this model vacancies are thermally released from VO or V_n complexes and form stable divacancies which subsequently trap VO. EPR studies have suggested [88] that complexes such as V_2O and V_3O, with $S = 1$, form and anneal out around 300°C although Davies et al [89] suggest that V_2O may be responsible for an LVM at 1005 cm^{-1} (10 K). This defect anneals out around 450°C.

5.1.3 Higher order VO_n defects

The 889 cm^{-1} defect also anneals out around 450°C, whereupon LVMs around 910, 976 and 1005 cm^{-1} (10K) are found. The intensities of these LVMs appear to be correlated. Another band at 986 cm^{-1} also arises but its intensity is not correlated with the other three. Corbett et al [141] assign the three LVMs to VO_3 where O_i is trapped near the VO_2 centre, since at this temperature O_i is now mobile. Thus VO_2 does not dissociate but traps O_i. Support for this assignment is that the rate of growth of 976 cm^{-1} line is proportional to the product of the intensities of the 889 cm^{-1} line and that due to O_i [90]. However, the interpretation here is dependent

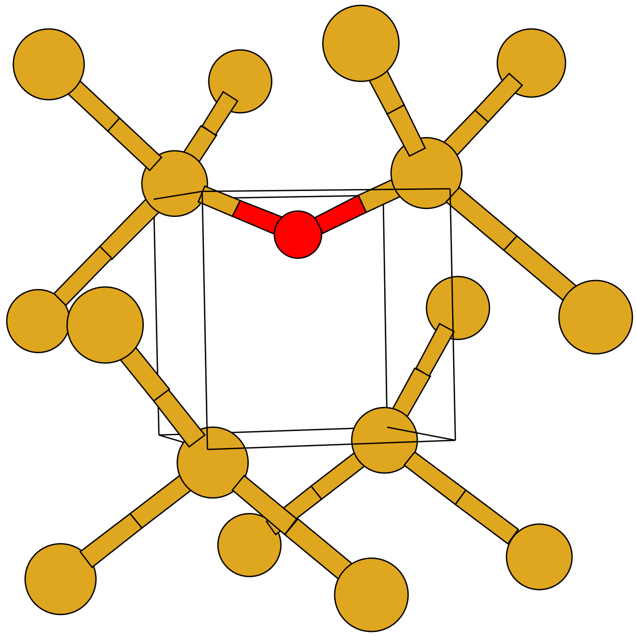


Figure 5.1: The OV centre in Si. The box indicates the [100] directions.

on the assignment of the 889 cm^{-1} defect.

Further lines have also been seen and tentatively assigned to VO_4 complexes, and a study of higher order VO_n defects is currently underway by Lindström and Hallberg.

5.2 Method and Results

We used a 130 atom cluster surrounding a vacancy to which O atoms were successively added, $Si_{70}H_{60}O_n$. All the atoms other than the surface H atoms were relaxed, and symmetry was constrained where appropriate. Unless stated otherwise the defects were all charge neutral and spin averaged. For many of the later runs we experimented with many of the cluster parameters, e.g. the status of the surface H atoms (fixed, free, or constrained in quadratic 'spring' potentials), and the arrangement of additional bond centred fitting functions. These are discussed below in the relevant sections.

5.2.1 The VO Defect

The O atom lay 0.934 Å away from the centre of the vacancy and formed Si-O bonds of length 1.754 Å. The Si-O-Si angle is 138°. There are two short back Si-Si bonds (2.341 Å) and one long one (2.51 Å). This is similar to the C_i defect in Si [91], in that the $\langle 110 \rangle$ bonds in the Si-O-Si plane are dilated, while the bonds above are compressed. The other pair of Si atoms bordering the vacancy have a shared

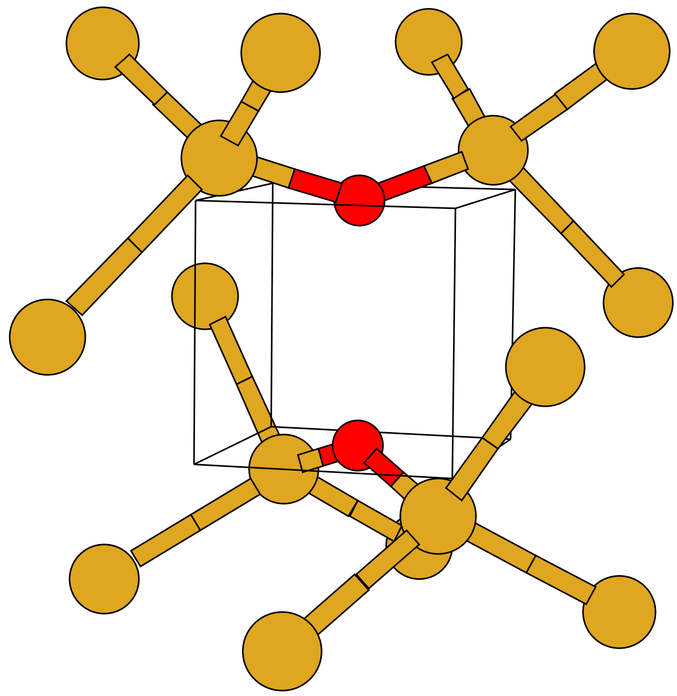


Figure 5.2: The O_2V centre in Si. The box indicates the [100] directions.

Si-Si bond length of 3.77 Å. The O atom is known to lie in a shallow $\langle 100 \rangle$ oriented potential valley, and so the precise bond angle is difficult to calculate; however this is in sensible agreement with earlier calculations which gave 149°[92] and 152°[93].

There is an empty mid-gap level, consistent with the acceptor level seen experimentally at $E_c - 0.17$ eV. The LVMs are given in Table 5.1 and Fig. 5.1 shows the structure of the defect. The defect is very tensile and the Si atoms are pulled in from their ideal lattice sites to form either Si-O bonds or the reconstructed Si-Si bond. The highest mode at 787 cm^{-1} is an asymmetric stretch of the two Si-O bonds and is within 44 cm^{-1} of the observed mode. This is much lower than the stretch of bond centred O_i in agreement with the dilated bond lengths. It shifts 38 cm^{-1} with ^{18}O in good agreement with the observed shift of 33 cm^{-1} [85]. This shows that the method gives isotopic shifts more accurately than the absolute frequencies.

In the last column of the table, one of the ^{28}Si neighbours has been replaced by ^{29}Si. This gives a shift of 1 cm^{-1}: less than the 3 cm^{-1} width of the line [78]. Presumably the mode is broad because of anharmonic effects: the reorientation energy of the defect is only 0.38 eV.

5.2.2 The VO_2 Defect

This structure (Fig. 1b) has D_{2d} symmetry with four Si-O bonds of length 1.710 Å. These are shorter than those in VO and lead to higher frequency vibrations. They are, however, longer than those of O_i (1.596 Å [94]), and the defect is still tensile. The Si-O-Si angles are 147° and the O-O separation is 2.61 Å. There are six LVMs (see Table 5.1). The highest at 807 cm^{-1} are E modes and represent independent motion of the two O atoms. There are no additional modes in the mixed O-isotope case. The mode at 656 cm^{-1} represents stretch of the O-O bond and is IR inactive. The E mode at 574 cm^{-1} does display coupled motion of the O atoms. This mode then gives additional O-O bands in the mixed isotopic case. There are no gap levels associated with the defect. The LVM at 807 cm^{-1} is about 90 cm^{-1} below that possibly assigned to the defect, but no other modes have been reported. It has good agreement with the experimental isotope shifts, and shows the correct upwards shift from VO $\rightarrow$ VO_2.

Since performing this work, new symmetry investigations by Lisby and Bech Nielsen [83] have determined the symmetry of the defect giving rise to the 889 cm^{-1} defect as being D_{2d}. Initial uniaxial stress experiments determined that it was either C_{2v} or D_{2d}, but the resolution of the fits was not able to distinguish between the two. They have since performed stress induced reorientation experiments, where stress is applied at 200°C, maintained while the sample cools, and then released [83].

If a defect is either tensile or compressive it will tend to reorient itself at the higher temperature with respect to the applied stress field. If, for example, this stress field is applied along $\langle 100 \rangle$, and the defect has inequivalent $\langle 100 \rangle$ directions, this means the defects will preferentially align. If the defect is tensile along $\langle 100 \rangle$ it will line up with the stress field, and if compressive along $\langle 100 \rangle$ it will line up along one of the two remaining unstressed $\langle 100 \rangle$ directions. If the defect symmetry means it has no preferential $\langle 100 \rangle$ direction, they will remain randomly distributed with respect to the stress direction.

The sample is then cooled to quench this orientation in, and the stress removed. If the sample is now probed using FTIR along different $\langle 100 \rangle$ directions, if the defects are preferentially aligned along one of these the signal will exhibit *dichroism*, *i.e.* the signal will appear stronger in one of the $\langle 100 \rangle$ directions than the others. In practise, as the chosen $\langle 100 \rangle$ direction of stress is known, applying FTIR along one of the other $\langle 100 \rangle$ directions and comparing to the sample before stress is sufficient to show such dichroism. If the defect possesses no preferential $\langle 100 \rangle$ direction, there will be no alignment effect and the signal before and after stress will be the same.

Stress applied in this way for VO_2 along $\langle 100 \rangle$ and $\langle 110 \rangle$ produced dichroism whereas ⟨111⟩ stress did not. This is consistent with D_{2d} symmetry but not C_{2v} (which has inequivalent ⟨111⟩ directions). This result is sufficient to exclude both the V_2O model and the C_{2v} structure described below, and coupled with our theoretical investigation show that the defect responsible for the 889 cm^{-1} line is indeed VO_2.

While investigating this structure we examined various other possible configurations of V + O_2, including a C_{2v} structure where the oxygen atoms were inserted in O_i positions in bonds neighbouring the vacancy, one either side along the same ⟨110⟩ orientation. Unfortunately we found that this was a lower energy structure than the VO_2 model presented here. We investigated varying the structure to see if we were inadvertently trapped in some local minimum, and also tried varying the basis set, however in every case it remained lower in energy. We also varied the cluster surface treatment, holding the surface H atoms fixed, allowing them to move freely, and constraining them to move in a quadratic spring potential, but again none of these reversed the energies.

This model cannot be the correct one for several reasons; it does not have the correct (D_{2d}) symmetry and the vibrational modes are wrong. In addition there is coupling between the oxygen atoms, so multiple modes are observed in the mixed isotope case in contradiction with experiment. We are unable at present to explain why we obtained the energies incorrectly, but note that J. Chadi also found that addition of O_i to a VO centre was an endothermic process [95]. One possibility is that there is some charge redistribution to the cluster surface which helps to stabilise pure V-based structures, however this alone is insufficient to explain the difference. Alternatively our charge density modelling may be inaccurate for vacancy-based structures, and the addition of gradient corrections to the exchange-correlation energy or the use of a plane wave basis set may overcome this problem. Recent tests with a new supercell version of AIMPRO that avoids intermediate fits to the charge density correctly gives the D_{2d} structure as 1.918 eV more stable than the C_{2v} structure; further investigation is required to determine whether this is an effect of the basis or the supercell.

5.2.3 Alternatives to VO_2 – the V_2O Defect

This possesses Si-O lengths of 1.73 and 1.72 Å with a Si-O-Si angle of 143°. The Si-O bonds are longer than in VO_2 and close to those in the VO centre. There is very little reconstruction of the Si-Si bonds in the divacancy and consequently there is an empty state very close to a filled mid-gap state. The LVMs are given in Table 5.1, and the structure is shown in Figure 5.3. There is a single mode at 791 cm^{-1}: almost the same frequency as in the VO defect. Thus the calculations give

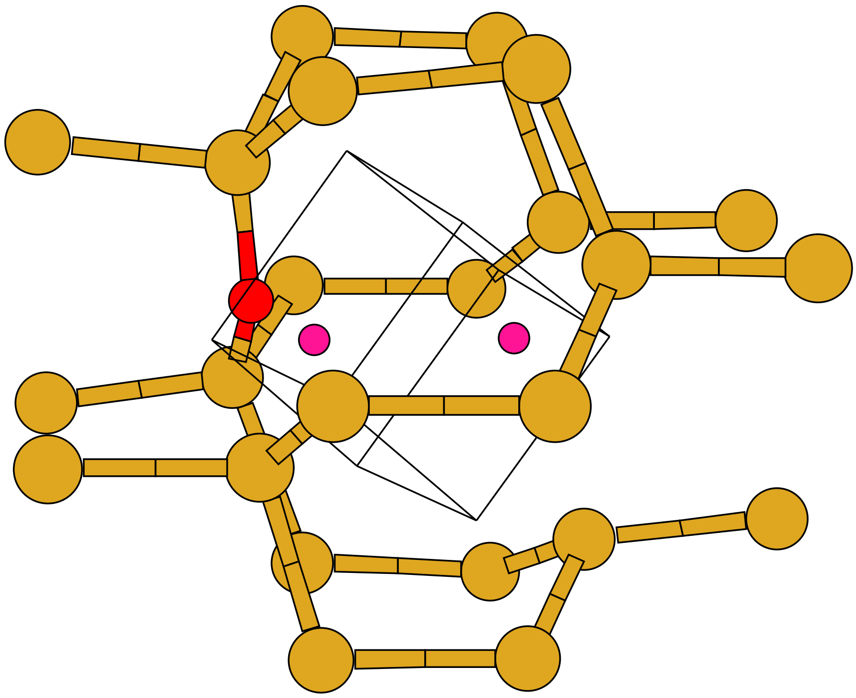


Figure 5.3: The OV_2 centre in Si. The box indicates the [100] directions. The two central dots mark the vacant lattice Si sites.

no support for an assignment of the 889 cm^{-1} defect to a V_2O complex [89]. It is then unlikely that V_3O complexes have high frequency modes and this casts doubt on the assignment [78] of the 889 cm^{-1} line to this defect. In addition the V_2O and V_3O defects do not have the required D_{2d} symmetry [83]. The table also shows the effect of replacing a ^{28}Si neighbour of O by ^{29}Si.

Divacancy complexes are a common vacancy impurity, which combine with O_i to form V_2O. These centres are more common in heavily damaged samples. V_2O and V_3O can, in addition to our studies, be ruled out as candidates for the 889 cm^{-1} line on the basis of various annealing studies. Lee *et al* showed V_2O anneals out at the same temperatures as VO [88]. Other studies show it to anneal out at around 450 °C to form V_3O and V_2O_2 [89], with V_2O_2 going on to form either V_3O_2 or V_3O_3 [96]. Although these results are not consistent with each other, neither fits the annealing properties of the 889 cm^{-1} mode. For these reasons V_3O can also be ruled out as a candidate for the 889 cm^{-1} line.

5.2.4 The VO_3 Defect

VO_3 forms at 450°C when VO_2 traps an additional O_i. This defect (shown in Figure 5.4) has an additional O atom, say O_c, next to VO_2. The bonds of the O_c atom are 1.632 Å to the Si atom bordering the vacancy and 1.686 Å to the further

Symm.	$^{28}Si^{16}O$	$^{28}Si^{18}O$	$^{28}Si^{16}O^{18}O$	$^{29}Si^{16}O$
VO				
B_1	787.4	38.1	-	0.9
Expt. [75, 85]	831	33		
VO_2				
E	806.5	38.8	0.0, 38.8	0.0, 1.0
A_1	656.4	18.5	8.8	1.3
E	574.2	3.5	0.1, 3.5	0.0, 2.7
B_2	563.8	8.6	4.7	1.4
Expt. [75, 84]	894 (4 K)	49 (300 K)		
V_2O				
A	790.7	38.2	-	1.1, 0.9
VO_3	^{16}O	$^{18}O_c$	$^{18}O_b$	$^{18}O_a$
A	997.8	22.7	0.1	12.8
A	902.2	23.5	0.1	30.5
B	813.5	0.3	39.2	0.1
A	669.1	4.1	5.6	6.6
A	634.3	6.5	8.2	2.1
B	611.6	4.5	0.0	1.9
A	597.5	1.1	2.4	1.8
Expt. [141]	1005, 976, 910			

Table 5.1: Calculated and Experimental Frequencies, cm^{-1}, of LVMs of VO_n defects - isotope values show downwards shift of the modes with shifting isotope. a, b, c for VO_3 refer to Figure 5.4.

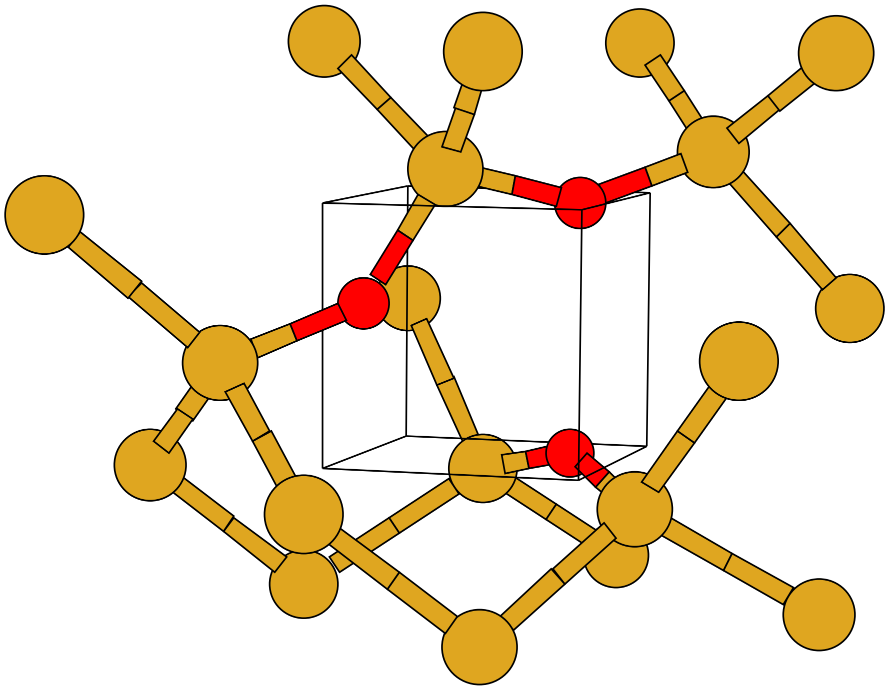


Figure 5.4: The O_3V centre in Si. The box indicates the [100] directions.

Si. These are much shorter bonds and naturally lead to higher O-related local modes. O_c is bonded to a Si atom which is also bonded to another O-atom, say O_a, with length 1.643 Å. O_a is also 1.696 Å away from the other Si atom bordering the vacancy. The third O atom, O_b, also lies within the vacancy and possesses two bonds of length 1.696 Å. Thus the presence of the additional O atom has only slightly perturbed the O atoms within the vacancy. Their separation is 2.646 Å.

The LVMs are given in the table. The 998 and 902 cm^{-1} modes are localised on O_c and O_a. The 813 cm^{-1} is localised on O_b. The 665 cm^{-1} and other modes are extended over all the O atoms. Thus the effect of an additional O is to displace upwards the mode due to VO_2 by about 100 cm^{-1} and introduce an additional mode $\sim$1000 cm^{-1}. This top mode is in good agreement with a mode found at 1005 cm^{-1}, although the other modes are about 70-100 cm^{-1} too low. A similar error was found for VO_2, and this error lies within the uncertainties of the method, and again, the upward mode shift from $VO_2 \rightarrow VO_3$ is correctly predicted. The prediction and observation of three O-related modes supports the assignment of the modes at 1005, 976, 910 cm^{-1} to VO_3 and therefore indirectly the 889 cm^{-1} mode to VO_2.

Just as we had trouble finding the lowest energy structure for VO_2, we also found a low energy structure for VO_3 consisting of an OV centre with two O_i atoms in the dilated bonds neighbouring the Si-O-Si of the OV. This is similar to the alternative VO_2 model described above but with an additional O_s sitting in the vacancy. This

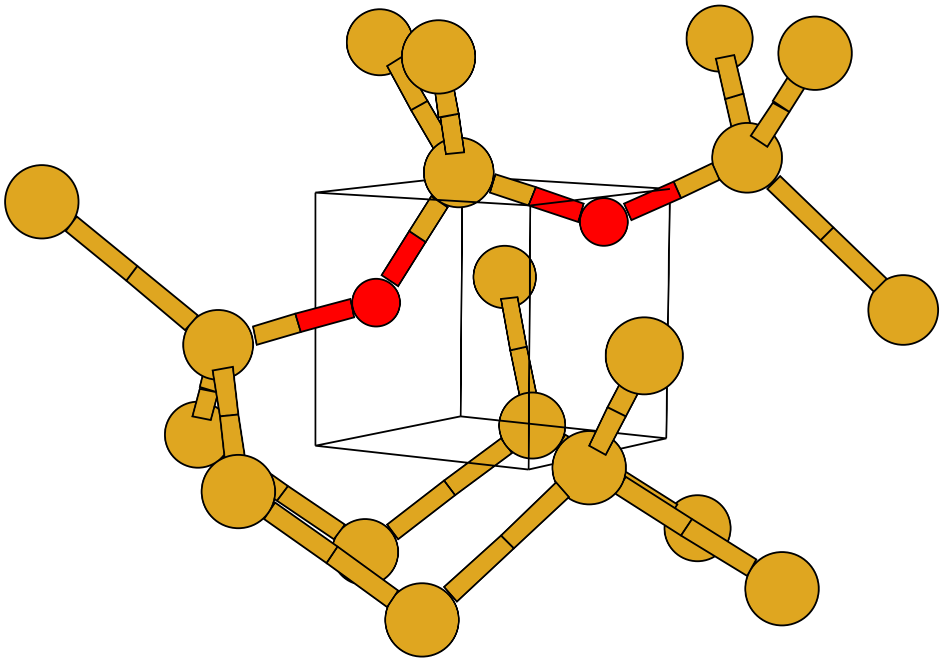


Figure 5.5: The OV-O_i centre in Si, pre-cursor to the O_2V. The box indicates the [100] directions.

proved to be lower energy than the VO_3 model described above, and also had good agreement with experimental LVMs. However since the energies of VO_2 have to be corrected in order to find the actual lowest energy structure we have assumed this also applies to VO_3, and so exclude this model. Experimental stress induced alignment experiments should be able to differentiate between these two centres, since the VO_3 presented above has C_{1h} symmetry, whereas this low energy VO_3 has C_{2v} symmetry. Work is currently underway on the higher order VO_n complexes by Lindström and Hallberg, and hopefully they should be able to check this.

5.2.5 The OV-O_i pre-cursor to the VO_2 centre

Annealing OV centres at 300°C or above leads to their breakdown and the formation of VO_2. However, a set of careful annealing experiments by Londos *et al* after neutron bombardment [97, 98] showed that at 272-340°C in long time anneals (4-5 hours), a pair of LVMs at 914 and 1000 cm^{-1} appear. As the 889 cm^{-1} VO_2 line grows, these two peaks anneal out. These peaks are only seen when the material is initially neutron bombarded; electron irradiation does not lead to these pre-cursors.

They speculated that these modes were due to a metastable intermediate structure consisting of a VO centre with a neighbouring O_i, which forms when VO diffuses through the lattice to a O_i atom. This then has to overcome an energy barrier to form VO_2. We modelled the proposed structure with O_i in a Si-Si bond neighbouring the Si-O-Si of the VO centre, in the same $\langle 110 \rangle$ plane. This is the

$^{16}O_i$–$^{16}O_s$	$^{16}O_i$–$^{18}O_s$	$^{18}O_i$–$^{16}O_s$	$^{18}O_i$–$^{18}O_s$
Relaxed			
926	17	9	41
886	26	33	45
656	25	3	30
Shifted Shared Si			
994	3	35	42
905	39	10	46
656	24	4	31
Experimental			
1000			
914			

Table 5.2: LVMs associated with O_i-O_s, the pre-cursor to $O_{2i}V$. All modes are in cm^{-1}, later columns show downwards shift as isotopes are changed as listed.

natural bond site for the O_i to adopt if it is not yet in the vacancy, since the bond is dilated and also highly polarised due to the neighouring O_s. The shared Si atom is able to move into the vacancy towards O_s, so that the Si-O bond lengths of both O atoms are very similar (this behaviour also occurs in VO_3). The O_i has an outer Si-O bond of 1.73 Å and a shared Si-O bond of 1.64 Å, O_s also has an outer bond of 1.73 Å but its shared bond is slightly longer at 1.68 Å. Both have Si-O-Si bond angles of 138°. The resultant vibrational modes are given in Table 5.2.

As can be seen, the modes are in reasonable agreement with experiment. However, the LVMs are highly sensitive to the position of the shared central Si atom sitting between the two O atoms. Moving this Si atom along ⟨-0.053,0.011,-0.011⟩ Å, less than 0.08 Å towards the O_i, causes the modes to shift into excellent alignment with experiment, as can be seen from the table. There is currently no isotopic shift data available on these peaks.

5.3 Discussion

The results give strong support to the idea [141, 90] that VO_n complexes preferentially form during annealing of irradiated O-rich Si. The VO_2 defect gives three IR active modes but only one has been detected. VO_3 gives three high frequency LVMs which have all been detected. V_nO complexes on the other hand give modes close to VO, *i.e.* around 830 cm^{-1}. VO_2 appears to have no connection with TDs [90] but may be a nucleus for oxygen precipitation.

One of the remaining questions is why $[O_i]$ does not decrease when VO anneals out at 350° and VO_2 is formed. We originally proposed this was due to oxygen dimers [99] complexing with vacancies, and recent work by Londos *et al* [79] seems to confirm this. The third, slow, formation process for VO_2 occurs while VO levels remain relatively static, suggesting that this is not a VO transformation process. It is known that vacancy aggregates exist in the material and frequently release and re-trap isolated vacancies. It therefore seems likely that these isolated vacancies in the structure are soaking up oxygen dimers. Alternatively there could be a breakdown of VO centres as they trap dimers, and the release of O_i through this mechanism roughly balances the loss of O_i through trapping of O_i by VO centres.

Different treatments will place different emphasis on each reaction step, but it now seems that the following processes contribute to the breakdown of VO and the formation of VO_2 (the numbering corresponding to the three regimes discussed in Section 5.1.2):

$$
\begin{array}{rrcl}
1a: & \mathrm{VO} & \rightleftharpoons & \mathrm{V+O_i} \\
b: & \mathrm{VO+Si_i} & \rightarrow & \mathrm{O_i} \\
2: & \mathrm{VO+O_i} & \rightarrow & \mathrm{VO\cdots O_i} \\
 & \mathrm{VO\cdots O_i} & \rightarrow & \mathrm{VO_2} \\
3: & \mathrm{V+O_{2i}} & \rightarrow & \mathrm{VO\cdots O_i(?)} \\
 & \mathrm{VO\cdots O_i} & \rightarrow & \mathrm{VO_2}
\end{array}
$$

Processes such as

$$
\begin{array}{rcl}
\mathrm{VO+VO} & \rightarrow & \mathrm{V_2O_2} \\
\mathrm{V_2O_2} & \rightarrow & \mathrm{VO_2+V}
\end{array}
$$

are unlikely due to the low concentration of VO centres as compared to other defects such as O_i. Note that unlike normal Cz-Si there will be a large number of Si_i present in this material due to the irradiation. This complex behaviour of VO destruction and VO_2 formation provides strong evidence for the presence of interstitial oxygen dimers in Cz-Si, since without them it is difficult to explain the observed reaction processes.

There is also some uncertainty about the formation kinetics of the VO_3 defects. Corbett et al [141] pointed out that the growth of the VO_3 bands was anomalously fast and required either a very non-uniform distribution of O_i or a diffusivity of VO_2 higher than that of O_i. However more recent work by Lindström *et al* [90] states that the activation energy for the growth of VO_3 is close to that of the diffusion of O_i, suggesting that O_i diffuses to immobile VO_2. If Corbett is correct there could naturally be O_i present quite close to many of the VO_2 centres. This could occur by

$$\mathrm{VO + O_{2i} \rightleftharpoons V + O_i + O_{2i} \rightarrow VO_2 + O_i}$$

which could leave the O_i quite close to the VO_2 centre. An alternative is that VO_2 is breaking down to VO and then trapping a dimer before the VO can lose its second oxygen. The activation energy observed by Lindström is then actually the activation energy for the formation of the dimers in the first place, since once formed they can rapidly migrate to these VO centres. It seems a new study of the growth of VO_3 is called for which should also include the effects of mobile dimers.

Finally we note that when O_i is complexed with a defect imposing tensile strain on the lattice, such as VO, N_2 [100] or C_s [101], the O_i-related LVM is reduced from 1136 cm^{-1} to around 1000 cm^{-1}. TDs also have modes in this region [102] and this suggests that they might consist of O atoms surrounding a core which gives rise to tensile strain. This need not be a vacancy, as many defects such as interstitial N_2 pairs [103], C_i [104] and Si_i [105] can also give dilated Si-Si bonds. This will be discussed further in Chapter 9.

Chapter 6

Oxygen–Oxygen defects

Although oxygen is able to complex with a wide variety of other point defects such as vacancies, carbon and nitrogen atoms, probably the most important reactions of O in Cz-Si are those involving aggregation of more than one O_i atom. Oxygen aggregates are believed to be responsible for the family of thermal donors (with the possible addition of Si_i), and it has been speculated that fast diffusing pairs of oxygen atoms may dominate many of the defect formation processes.

In this chapter, we examine the way oxygen behaves in silicon without any other point defects present. We initially examine isolated O_i, and then an oxygen dimer, O_{2i}. This includes structural and vibrational mode analysis, as well as determination of the migration barrier and path. We then go on to look briefly at higher order oxygen complexes such as the trimer, O_{3i}, and discuss their implications in the development of other defect species such as thermal donors. Since this is still very much work in progress, many of the conclusions are necessarily open ended.

6.1 Background

6.1.1 Interstitial Oxygen

The primary configuration for oxygen in silicon is single interstitial oxygen, O_i (see Figure 6.1). O_i sits in a bond centred (BC) site between two lattice Si atoms, dilating the Si-Si bond by around 36% [106, 107, 108]. It has a Si-O-Si bond angle of 160-180°, and is able to precess around its Si-Si bond with a low energy barrier (less than 0.09 eV). It is electrically neutral. Table 6.1 gives various properties of O_i, both experimentally determined and as calculated by various theoretical groups.

Thus general consensus puts the Si-O bond length at ∼1.6 Å, similar to that in α-quartz at 1.60-1.61 Å[119]. There is also remarkable consistency with the properties of smaller Si–O molecules. The smallest of these is disiloxane, (SiH_3)-O-(SiH_3), which has a Si-O bond length of 1.634 Å, Si-O-Si bond angle of 144.1° and

Source	Si-O length (Å)	Si-O-Si angle	LVMs (cm^{-1})
Experimental			
Bosomworth and Newman[77, 109]		164	1136, 518
Theoretical			
Previous AIMPRO[110, 111, 94]	1.59	172	1104, 554
Snyder *et al*[112]	1.61	180	1275, 699
Estreicher[113]	1.57	164	
Oshiyama *et al*[93]	1.68	152	1187
Kelly[114]	1.77	140	
Needels *et al*[115]	1.64	140	
Artacho *et al*[116]	1.56	180	1150, 517
Disiloxane [50, 117, 118]	1.634	144.1	1107, 606

Table 6.1: Theoretical and Experimental Properties of O_i. The Si-O length is in Å and the Si-O-Si angle in degrees.

vibrational modes at 1107 and 606 cm^{-1} [50, 117, 118]. This 'scalability' from small molecular systems to the Si bulk provides indirect support to the cluster method which similarly assumes that a large molecular cluster is a reasonable physical approximation to the bulk.

The variation in bond angle is due to the potential surface perpendicular to the Si-Si bond which appears to be fairly flat providing the bond lengths are kept approximately constant. Bosomworth *et al.* estimated this angle by fitting overtones of the 29.3 cm^{-1} bend mode to a simple potential for which they deduced a Si-O-Si angle to be 164°, supposing a Si-O length of 1.6 Å[77].

The symmetric and asymmetric stretch modes are well reproduced by most calculations (see Table 6.1). Previous AIMPRO calculations gave isotopic shifts in good agreement with experiment [94]. The 1136 cm−1 mode is an assymetric O-Si stretch mode. The 518 cm^{-1} mode is a symmetric O-Si stretch, where both Si atoms move symmetrically along the bond while the O atom remains fixed (hence the mode does not shift when the O isotope is changed). There is also a low frequency 29 cm^{-1} bend mode. These modes, along with the AIMPRO calculated values are given in Table 6.3. The primary 1136 cm^{-1} absorption peak has been callibrated to the O_i concentration (ASTM procedure F121-83) [120]

$$
\begin{aligned}
[\mathrm{O_i}] &= 0.94 \times 10^{16} \times \mathrm{Icm}^{-3} \\
[\mathrm{O_i}] &= 3.14 \times \alpha_{\mathrm{MAX}} \times 10^{17}\mathrm{cm}^{-3}
\end{aligned}
$$

Where I is the integrated absorption coefficient of the band and α_{MAX} is the peak absorption coefficient. The value of 3.14 replaced an earlier calibration coefficient of 2.45 in 1989, and so any work based on these calibrations before that time should be treated with caution.

O_i diffuses by hopping between bond centred sites. In order to achieve this it has to break one of its current Si-O bonds, move to some intermediate state often referred to as a 'y-lid' [121] (since in principal the O atom has three bonds), and then push its way into the next Si-Si bond. The diffusion constant for this motion has been determined by several groups, and Mikkelsen [4] fitted data from six independent experiments, leading to a diffusion constant for this motion of

$$D = 0.13e^{-2.53\mathrm{eV}/k_BT}\ \mathrm{cm}^2\mathrm{s}^{-1}$$

This is valid over a wide temperature range, from 300-1200°C. Stress induced dichroism experiments show this barrier at temperatures around 380°C [106, 122, 123]. Secondary ion mass spectrometry (SIMS) measurements [4, 124] also lead to the same activation energy for the oxygen diffusivity at temperatures above 700°C, while small angle scattering measurements on the growth of small oxygen precipitates show that the precipitation is controlled by oxygen diffusion at temperatures above 650°C [125].

Below this temperature, the process is sufficiently slow to become unusable. An important question is whether long range oxygen diffusion and the growth of oxygen precipitates at temperatures below 400°C is controlled by the same process. It is known that some impurities (such as hydrogen) can increase the low temperature hop rate [126, 127] and reduce its activation energy to about 1.6 eV [113, 128]. This then increases the rate of oxygen precipitation as the thermal donor production rate also increases. It has also been proposed that pairs of oxygen atoms may be able to diffuse faster than O_i, and this is discussed further below.

Various theoretical groups have attempted to model O_i diffusion, and the results are summarised in Table 6.2. Apart from Jiang and Brown[129] they all obtained a saddle point midway between the two bond centre sites, either as an implicit assumption in the calculations, or determined by mapping out the potential energy surface [110, 130]. There are large differences between the calculations, which may be because the potential energy near the saddle point is rather flat.

Needels *et al.*, [115] suggested the adiabatic saddle point is 0.7 eV lower than the experimental one. They argued that although an O atom provided with about 1.8 eV of kinetic energy moves close to the saddle point, it fails to open up the attacked Si-Si bond and returns to its starting point. They suggested that the O atom must develop kinetic energy of above 2.3 eV to avoid this return. More recent work by Ramamoorthy and Pantelides [130] developed this idea and suggested that the total barrier is a mixture of two independent barriers, namely that required

Author	Calculation / Measurement	Barrier (eV)
Experimental		
Mikkelsen [4]	Best-Fit to various data	2.53
Mikkelsen *et al* [131, 106, 132]		2.56
Stavola *et al* [133, 122]		2.54
Bosomworth and Newman [77, 109]		2.5
Theoretical		
Jones *et al* [110]	Previous AIMPRO	2.8
Jiang and Brown [129]	Empirical fit potential	2.43
Needels *et al.* [115]	local DFT supercell	1.8
Oshiyama and Saito	local DFT supercell [134]	2.0
	Cluster calculation [93]	1.2
Estreicher [113]	PRDDO	4.1
Kelly [114]	s-c Greens Function	2.5
Snyder and Corbett [121], Kelly [135]	s-c CNDO / MINDO3	2.3 - 2.5
Ramamoorthy and Pantelides [130, 136]	DFT supercell	~2.5

Table 6.2: Experimental and calculated diffusion barriers for O_i (eV)

to break and reform the relevant Si-O bonds, and a second barrier to break and reform the other two Si-Si bonds. These two barriers together give a potential surface crossed by a 'saddle ridge', the average height of which is around ~2.5 eV.

6.1.2 The oxygen dimer

The first stage of O precipitation is considered to be the formation of a dimer, O_{2i}. It was originally proposed that rapidly diffusing O_2 molecules could lead to enhanced oxygen diffusion [137]. However theoretical calculations [138] showed that an O_2 molecule in Si dissociates into two O_i with a large associated gain in energy. Later calculations showed oxygen molecules to be unstable compared to nearest neighbour pairs of bonded interstitial oxygen defects [135, 114, 121, 115].

The evidence for the dimer started with results in 1981 showing thermal donor formation rates decreased towards the sample surface. Interpreting this as oxygen out-diffusion gave oxygen diffusivities enhanced by more than four orders of magnitude [139].

There are 17 thermal donors which are likely to involve increasing numbers of oxygen atoms. The growth rates of these donors have been monitored and it is found that they grow with an activation energy around 1.8 eV – much lower than the energy for a single O_i hop [140]. There have been suggestions that VO_2 complexes could diffuse rapidly [141] but the evidence is not conclusive. The early suggestion [137] that oxygen pairs could diffuse rapidly has been supported by SIMS measurements of the out-diffusion of oxygen [142] between 500 and 700°C, and recent experiments [143, 144] studying carefully the loss of O_i from solution at temperature as low as 350°C. These show that the process is second order below 400°C but of higher orders, up to 8, as T increases to 500°C. The low temperature process must be controlled by the formation of oxygen dimers. The change in order of the process is then explained by an instability of the dimer above about 450°C and thus larger numbers of oxygen atoms are required to diffuse together to form a stable precipitate.

A plane wave supercell calculation [115] gave a binding energy for the oxygen dimer of 1.0 eV. However other calculations suggest that dimers will not bind [135]. Semi-empirical CNDO/S techniques showed that the dimer has a binding energy of 0.1 eV and the saddle point for its migration is 1.36 eV [121]. The saddle point for dimer diffusion was found to lie close to an over-coordinated oxygen square defect – rather similar to the known structure of the nitrogen pair [145] (see Chapter 7). However this was obtained from total energy calculations of manually selected intermediate structures, and re-examination of this problem using *ab initio* methods with constrained optimisation would be desirable.

6.1.3 The 1012 cm^{-1} LVM

Until recently no LVMs were assigned to the oxygen dimer. However it has been suggested that a vibrational mode at 1012 cm^{-1} (first observed in 1972 [146]) could be due to the oxygen dimer [147].

There are several reasons for this. Firstly it is present after heat treatments at a wide range of temperatures: 650°C [102, 148], 1100°C [149] and 1350°C [148], as well as in as-grown material at concentrations of about 1.5×10^{15}cm^{-3} [270]. The line is definitely oxygen related given its 43 cm^{-1} isotopic shift when going from ^{16}O to ^{18}O [239]. Its annealing properties (discussed in Section 6.5.1) are also consistent with that of a rapidly diffusing dimer. Given that it is present over such a range of conditions it seems sensible that it should be the simplest oxygen complex possible, namely a dimer.

Assuming that the 1012 cm^{-1} mode is responsible for the dimer, simple rate equation analysis led to a binding energy in the range $0.1 \leq E_b \leq 0.2$ eV [147]. This is consistent with the idea that the dimer should not be important in the kinetics of TD formation at higher temperatures [272]. Fits to plots of Log[1012] vs. Log(1/T) gave a more accurate dimer binding energy of 0.16 eV [147, 239]. Recent detailed analysis has suggested a figure of $E_b = 0.3 \pm 0.05$ eV [150].

In addition to the 1012 cm^{-1} mode, there is a secondary weaker mode at

1006 cm^{-1} which is normally observed under the same conditions. It is also present in as-grown material, but at concentrations of about 1.5×10^{14} cm^{-3} [270], and shows similar annealing properties. If the 1012 cm^{-1} mode is produced by the dimer, given that the most common defect in silicon is O_i, it is likely that the 1006 cm^{-1} mode is caused by the oxygen trimer, O_{3i}. Little is known about this defect, but crucially, from fitting their models to experimental data, Markevich and Murin showed the fits were greatly improved for the 1006 cm^{-1} mode if they assumed that this defect was highly mobile [147].

There is a strong correlation between the 1012 cm^{-1} band and the 1107 band due to O_i [102]. The band shows very little temperature dependance (unlike O_i, suggesting any bonding in the defect is very different to O_i); it shifts from 1013 to 1012.2 cm^{-1} on going from 300K to 10K. With ^{18}O it drops to 969 cm^{-1}. It is always present in as-grown material.

After we suggested a dimer should give rise to more modes, further examination has determined modes at 1105 and 1062 cm^{-1} (1060 cm^{-1} at 10K) which are correlated with the 1012 cm^{-1} [151]. The 1062 cm^{-1} mode is nearly twice as intense as the other two, and is very broad (FWHM $\sim$7 cm^{-1} at 10K). More recently the 1105 cm^{-1} was shown to be unconnected. There are also two lower modes at 685 and 552 cm^{-1}, the 552 cm^{-1} exhibiting only a small shift with oxygen isotope.

Further evidence has come from irradiation experiments by Lindström and Hallberg [150]. Electron irradiated material was annealed, giving a high concentration of VO_2 centres. This was then proton irradiated (which leads to Si_i production), and using FTIR a decrease in the VO_2 889 cm^{-1} signal was observed concurrently with the appearance of lines at 1012 and 1062 cm^{-1}. This is consistent with the reaction $VO_2 + Si_i \rightarrow O_{2i}$.

The 1012 cm^{-1} band is not seen in C-rich material [239], where presumably most dimers are being trapped by C_sO_n complexes.

Note that there is another, electrically active, defect that forms after long time annealing, which also possesses a LVM at 1012 cm^{-1}. The FWHP of this new 1012 cm^{-1} is much broader than the early 1012 cm^{-1} line and rapidly swamps it. This is also correlated with a lower mode at 744 cm^{-1}. These modes come from a fundamentally different defect and should not be confused with the dimer modes. This mode overlap originally led to incorrect assignment of the 1012 cm^{-1} line to some sort of shallow thermal donor centre[239].

6.2 Method

The calculations were performed using a variety of clusters. The structural and vibrational data presented here comes from large 150 atom clusters, $Si_{79}H_{68}O_2$, the same size as used to study the STDs and TDs. The diffusion data was obtained from 113-115 atom clusters, $Si_{57}H_{56}O_n, n = 1, 2$, similar to those used to explore the structure and vibrational modes of O_i, VO, VO_{2i} and VO_{3i} defects [99].

In order to investigate diffusion barriers and determine the saddle point configuration of O_i and O_{2i} it is necessary to introduce constraints on the positions of atoms during the relaxation. Unless this is done, the cluster would relax away from the saddle point to the nearest stable configuration.

There are two methods that we have used to constrain the system, one based on an orthogonality constraint and the other on constraining certain bond lengths within the system. The first is a method common amongst the literature but has a tendency to overconstrain the system.

6.2.1 Orthogonality Constraints

In order to constrain the system in this way it is necessary to firstly obtain the ground state structures of the two end points of the diffusion path, A and B. A vector is then constructed in order to translate A into B, where N is the number of atoms in the system.

$$\mathbf{R_i} = \mathbf{B_i} - \mathbf{A_i}, i = 1, 3N \tag{6.2.1}$$

The atoms are then moved a fraction α along this vector, where α=0 corresponds to structure A, α=1 corresponds to structure B. The atoms are allowed to relax using a conjugate gradient algorithm at each point, subject to the constraint that on each iteration the total system relaxation vector must be orthogonal to $\mathbf{R_i}$. If only one atom is moving then this is equivalent to allowing the atom to relax in a plane perpendicular to its direction of movement.

This method is simple to implement but can sometimes result in a system that is overconstrained; it tends to be best for systems where the diffusion motion primarily involves a single atom [152]. In particular when examining compound motion such as the oxygen dimer this method will only sample diffusion paths where both atoms are moving simultaneously, which need not be the lowest energy diffusion mechanism.

Therefore although initial investigations were performed with this method, we later switched to a second, bond-length constraint method detailed in the following

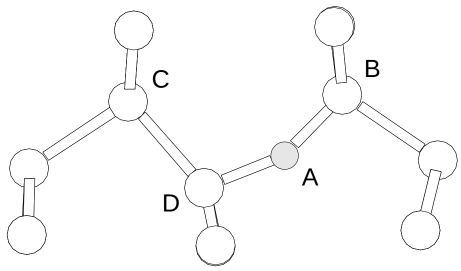


Figure 6.1: Schematic of interstitial oxygen in silicon. For explanation of the labelling, see text.

section.

6.2.2 Constraining Bond Lengths

Suppose that an atom A is hopping from one site to another, during which one bond $A-B$ is broken and the bond $A-C$ is created (see Figure 6.1). Then a constraint is introduced on the relative bond lengths, r_{A-B} and r_{A-C}, and the cluster relaxed maintaining this constraint. The actual constraint used is

$$c_1 = r_{A-B}^2 - r_{A-C}^2 \qquad (6.2.2)$$

Now c_1 is clearly negative in the configuration where the bond $A-B$ is short and positive in the configuration where the bond $A-C$ has been formed.

The saddle point usually, but not always, corresponds to a value of c_1 around zero where the $A-B$ and $A-C$ bond lengths are equal. For interstitial O motion, a constraint is selected so that A, B and C correspond to the O atom and the two Si atoms which swap bonds with O. However, the imposition of one constraint is insufficient [128]. During the hop, the O atom initially bonded to the Si atoms D and B, becomes bonded to D and C. However at the same time the Si atom D also breaks a bond with C and makes a bond with B. To deal with this a second constraint is required. Here,

$$c_2 = r_{D-B}^2 - r_{D-C}^2. \qquad (6.2.3)$$

In the initial configuration c_1, c_2 are both negative but become positive after the hop. The cluster is relaxed using a conjugate gradient algorithm subject to these constraints for a range of values of c_i. The saddle point can then be directly found by interpolation on the resultant 2D energy surface.

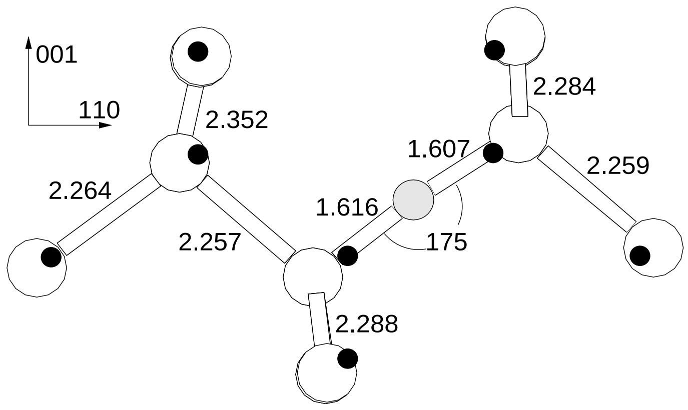


Figure 6.2: Structure of O_i, bond lengths in Å. Dots mark the ideal lattice sites.

6.3 Interstitial oxygen

A single oxygen atom was placed at a central BC site of the 148 atom cluster and all atoms relaxed. The Si-O bond lengths are 1.607 and 1.616 Å, with a 175° bond angle. The Si-Si bond is dilated to 3.219 Å, an increase of 37%, by the presence of the oxygen atom. These results are in good agreement with experiment (36% bond dilation, 160-180° angle), as well as the quartz bond length of 1.60-1.61 Å[119]. The next Si-Si bonds along the $(0\bar{1}1)$ plane are compressed by $\sim$3.9%. The structure is shown in Figure 6.2. Local vibrational modes are given in Table 6.3, with the corresponding atomic motion shown in Figure 6.3.

The isotope shifts are in excellent agreement, the absolute values are reasonable although the asymmetric stretch is a bit low. The symmetric stretch is also in excellent agreement, since it is experimentally believed to lie at 612.2 cm^{-1}, but due to its lack of dipole is only observed via a combination mode at 1749 cm^{-1} [108]; this is very close to our value of 603.3 cm^{-1}. The v_2 wag modes around 650 cm^{-1} are too high in our calculation, and suggest that the curvature of our potential perpendicular to the Si-Si bond for O is too steep. The splitting in the wag modes is due to slight deviations from the symmetric structure, as symmetry was not constrained during relaxation.

The saddle point for hopping to an adjacent bond centre was found using the method described above, and corresponds [128] to $c_1 = c_2 = 0$. The saddle configuration then possesses C_{2v} symmetry with O equidistant from the Si atoms B and C, and the Si atom D is equidistant from B and C. The configuration in which an O atom is equidistant from two second neighbour Si atoms, such as B and C is called the ‘Y-lid’, as there is some bonding with the third Si atom D. The calculated energy barrier is 2.6 eV, in excellent agreement with the experimental value of 2.53 eV. This is an improvement on the value previously obtained using AIMPRO

	^{16}O	^{17}O	^{18}O	^{28}Si^{29}Si	^{28}Si^{30}Si	^{30}Si
Experiment						
Asymmetric v_3	1136.36	26.9	51.4	1.9	3.7	7.2
Symmetric v_1	[612.2]	[1.3]	[1.2]	[4.5]	[9.5]	
Si Asymmetric-wag v_2	517.8		0			
This work						
Asymmetric v_3	1058.9	26.4	50.3	1.1	2.1	5.6
Symmetric v_1	603.3	0.2	0.5	4.5	8.7	18.6
Si Asymmetric-wag v_2	530.2	0.1	0.1	3.7	7.3	11.3
	544.6	0.0	0.1	3.2	6.3	11.0
Coupled v_2 wag	664.0	7.0	13.0	2.2	4.1	11.3
	655.5	6.6	12.2	2.4	4.4	11.5

Table 6.3: Local vibrational modes of O_i (cm^{-1}). Experimental results from References [153, 154, 77, 155] (LHeT). Figures in square brackets are determined indirectly [155]. All except first row show the downward shift with change in isotope. Theoretical results are new to this work. The split in theoretical results are slight numerical variation due to cluster assymetry. Atomic motion associated with each mode is shown in Figure 6.3.

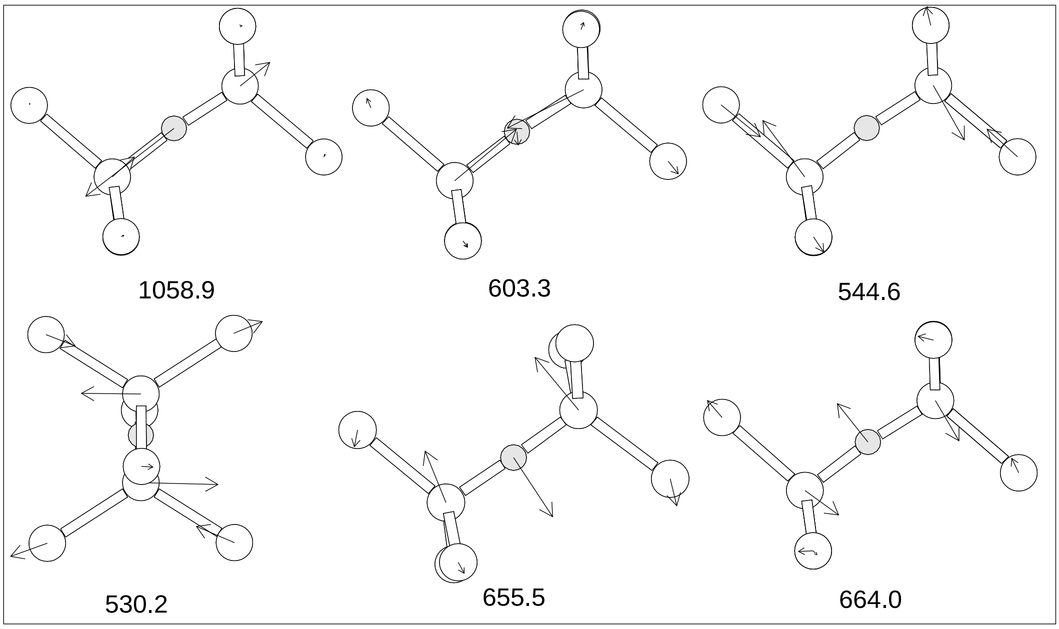


Figure 6.3: Atomic motion associated with each vibrational mode for the oxygen interstitial (modes in cm^{-1}) (see Table 6.3).

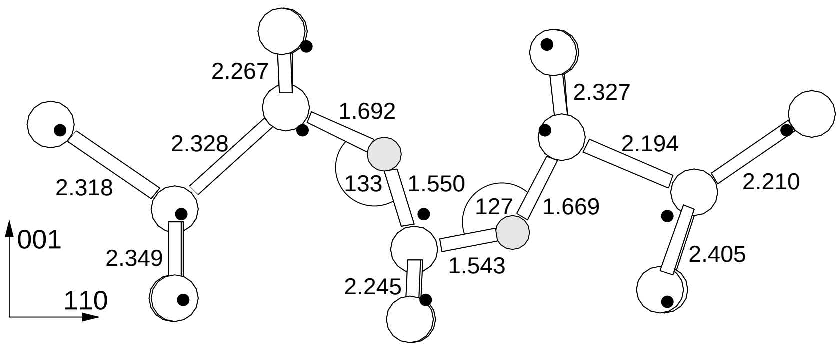


Figure 6.4: The stable assymmetric dimer in silicon (lengths in Å).

[110] of 2.8 eV. The saddle point structure consists of B-O and D-O bonds of length 2.5 and 1.55 Å respectively. These long bonds reflect the inability of the O atom at the Y-lid position to strain the lattice and pull in the Si atoms along [011]. There is then a net tensile strain at the saddle point pulling the Si atoms B and C inwards along [011]. This contrasts with the compressive strain at the stable BC site where the backbonds to D and B are compressed.

These results show that we are able to accurately obtain the ground state structure for O_i, and obtain reasonable agreement on the vibrational modes and migration saddle point energy.

6.4 The oxygen dimer

In the absence of impurities, especially those with tensile strain fields which promote oxygen aggregation, the first oxygen complex to be formed is likely to be a dimer. In this section we explore its stability and kinetics of motion.

An oxygen dimer consisting of two neighbouring BC sited O_i was placed at the centre of the cluster and subsequently relaxed. The lowest energy structure was a 'puckered dimer' structure, where both oxygen atoms sit bond centred in a kinked Si-O-Si bond, with the kinks pointing in the same direction along $\langle 110 \rangle$. Rather than having equal Si-O bond lengths as O_i, each atom has one short and one long bond, the short bond being formed with the shared central Si atom (see Figure 6.4). This strong bonding is due to the quadropole interaction between the two oxygen atoms; both draw charge from their Si neighbours making them $\delta+$, and thus their shared Si becomes extremely positively polarised. The stability of the dimer is thus attributed to the creation of a [011] chain of aligned polar bonds as in $\mathrm{Si}^{\delta+}$–$\mathrm{O}^{\delta-}$–$\mathrm{Si}^{\delta+}$–$\mathrm{O}^{\delta-}$–$\mathrm{Si}^{\delta+}$. The defect is electrically neutral with no gap levels. It is notable from the shift in position of the Si atoms compared to the ideal lattice that the defect has a strong strain field in the $\langle 001 \rangle$ direction. However along $\langle 110 \rangle$ its strain field is directional; 'behind' the dimer the Si atoms show very little displacement from their ideal sites, but there is a large displacement in front of it. This suggests that as the dimer moves through the lattice it produces in front of it a large 'bow wave' of lattice compression.

During our investigation we also found a stable symmetric dimer structure, shown in Figure 6.7. However this is 0.259 eV higher in energy than the assymetric case (and has vibrational modes that disagree with experiment). This is discussed further below with reference to the 1105 cm^{-1} mode defect. It exhibits the same strong binding to the shared Si atom. The influence of cluster size on the energy of the dimer was investigated by displacing it so that it is centred on a nearby Si atom, *i.e.* it is displaced by $a/4$ (011). This energy changed by less than 0.08 eV, showing the cluster surface has negligible effect.

This puckered structure is consistent with work by Needels *et al* [115], but in contrast to Greens Function calculations by Kelly [114] as well as previous AIMPRO calculations on the dimer [156], who both found a symmetric C_{2v} structure. This is discussed further in Section 6.4.1 below. Seperating the dimer by one Si-Si bond along $\langle 110 \rangle$ raised the cluster energy by 1.7 eV, which is a crude first estimate at a binding energy. However such energy differences are very dependant on the number of bond centred fitting functions used, and the error on such a value may be large. This compares to a dimer binding energy of 2.25 eV for earlier AIMPRO calculations [156] (the energy difference between the dimer and a Si-O-Si-Si-O-Si structure) and 1.0 eV by Needels *et al* [115] (see Section 6.1.2).

The vibrational modes of the asymmetric dimer are given in Table 6.4. The associated eigenvectors are shown in Figure 6.5. As can be seen, the top two modes roughly correspond to assymetric stretch modes for the inner and outer O atoms respectively. The modes are in good agreement with the experimental values for the 1012 cm^{-1} defect, including the shifts with $^{16}O \rightarrow ^{18}O$. There are some calculated modes which are not observed, however these have a lower (dipole moment)^2 and thus absorption intensity. Thus these modes serve to confirm that the 1012 cm^{-1} defect is indeed the puckered oxygen dimer.

There is a problem however with the mixed dimer modes; experimental FTIR on mixed ^{16}O / ^{18}O samples show no mixed mode absorption. This suggests the atoms are decoupled in some way, e.g. through greater seperation than that of the puckered dimer. Current models of the O-H defect suggest that O and H sit in bonds on opposite sides of a hexagonal interstitial ring site [157]. A possible alternative dimer model would be two O atoms in these positions, which is shown schematically in Figure 6.6. Their strain fields could couple in this way, and if the one buckled into the hexagonal site through Coulombic interaction with the Si neighbours of

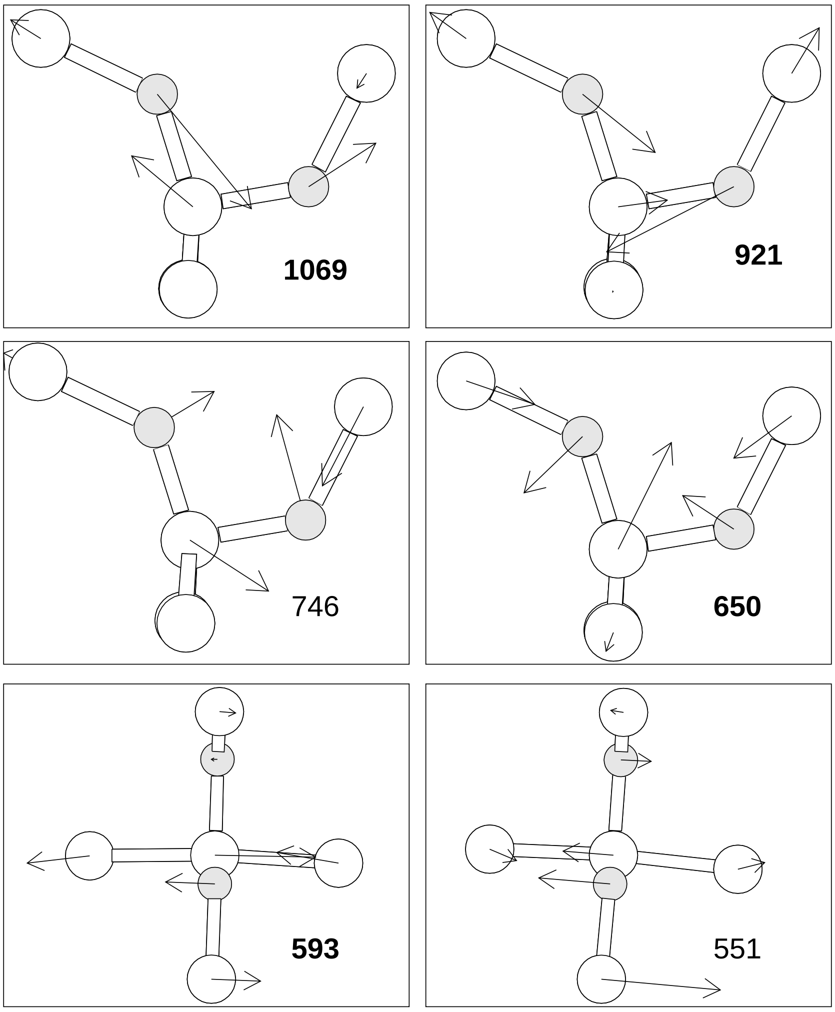


Figure 6.5: Atomic motion associated with each vibrational mode for the assymetric dimer (modes in cm^{-1}). The top two modes roughly correspond to the assymetric stretch for the inner and outer O atoms respectively, the lower modes are wag modes. The 593 cm^{-1} mode is largely localised on the core Si atoms and hence shows little shift with oxygen isotope, in agreement with experiment.

the other O atom, this would push the other atom to buckle outwards. Thus the two would be inequivilent, leading to the observed experimental mode splitting. In this case rapid migration could still occur with the puckered dimer structure, which would split to give this hexagonal ring as the stable structure. We are currently investigating such a model.

6.4.1 The symmetric dimer

The symmetric dimer structure is a local minimum on the potential surface. It seems to cover quite a large region of phase space, so that the relaxed structure is very dependant on the choice of starting structure, and most starting structures we chose produced a symmetric dimer structure. This was originally published as

	^{16}O	^{17}O	^{18}O	$^{16}O^{18}O$	$^{18}O^{16}O$	Intensity
Experimental [151]						
1	1059.6		48.1			0.9
2	1012.5		43			1.0
3	685		10			0.08
4	552		0.5			0.15
Assymetric Dimer						
1	1069.3	24.8	47.2	8.9	33.2	1.16
2	921.3	21.8	41.6	30.9	15.7	1.00
	746.4	10.3	19.3	13.5	5.3	0.26
3	649.6	5.4	10.2	4.1	6.7	0.33
4	592.6	1.6	2.8	0.1	2.7	0.50
	551.0	3.9	7.5	1.0	6.5	0.07
Symmetric Dimer						
	1169.6	27.4	52.3	18.6	20.0	1.00
	1077.9	27.6	52.6	33.8	32.5	0.76
	643.9	0.8	1.5	0.8	0.6	0.42
	637.7	3.0	5.3	2.7	2.9	
	581.3	5.3	9.9	4.7	4.7	
	536.0	7.5	14.2	7.0	6.8	
Split Dimer						
	1201.5	26.7	50.9	23.1	8.2	
	1152.0	26.0	49.4	28.7	40.3	
	647.4	6.5	11.7	7.0	4.5	
	632.1	5.7	10.5	8.1	1.5	
	622.2	3.2	6.9	2.2	5.2	
	569.6	2.2	4.1	2.7	1.7	
	539.7	4.4	8.3	4.1	4.5	

Table 6.4: LVMs for the dimer (cm^{-1}). Isotopic values show downwards shift for change in isotope. The assymetric dimer is more stable than the symmetric one by 0.259 eV. Where mixed isotope results are listed, the first isotope refers to the 'inner' atom of the assymetric dimer. Calculated intensity is the dipole moment squared for the ^{16}O case for the given mode, divided by that of the 921.3cm^{-1} mode. The 'symmetric dimer' was not symmetry constrained and slight variations in position account for the difference between the $^{16}O^{18}O$ and the $^{18}O^{16}O$ values.

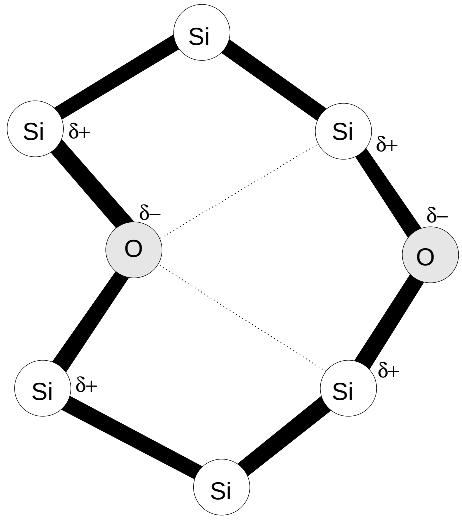


Figure 6.6: Schematic diagram of a possible alternative dimer model that should show no mixed isotope vibrational mode coupling. The atoms sit bond centred, on opposite sides of the hexagonal interstitial site.

the dimer structure using AIMPRO [156]. However according to our calculations this is 0.259 eV less stable than the puckered dimer structure. The defect is shown in Figure 6.7. This structure is also electrically neutral with no states in the gap. The vibrational modes are shown in Table 6.4, and as can be seen, they are very different from the experimentally observed dimer modes. Thus these results exclude the symmetric dimer as a model for the 1012 cm^{-1} LVM defect.

However, there is an additional mode seen at 1105 cm^{-1} in the presence of the dimer modes [151], and it was speculated that this may be two O_i atoms seperated by a single Si-Si bond along $\langle 110 \rangle$. An alternative explanation is that this is due to the symmetric dimer. Although the calculated symmetric dimer does not have a mode exactly at this value, there is some evidence in favour of this model for the 1105 cm^{-1} mode. The defect appears to have an experimental binding energy of 0.2 eV as compared to 0.3 eV for the dimer [151], so an energy difference of 0.1 eV between the two structures is quite close to our calculated value of 0.26 eV. The mode is only observed at 10K, and disappears by 300K. This would be consistent with symmetric dimers getting 'frozen in' at lower temperatures with a finite reorientation barrier. At 300K they are able to overcome this and no symmetric dimers are observed. In addition the 1105 cm^{-1} mode drops by 50/52 cm^{-1} when ^{16}O is switched to ^{18}O, very close to our calculated value of 52.6 cm^{-1}. To fully check this model it would be necessary to calculate the energy barrier between the symmetric and assymmetric structures, as well as locate further experimental modes to improve the assignment.

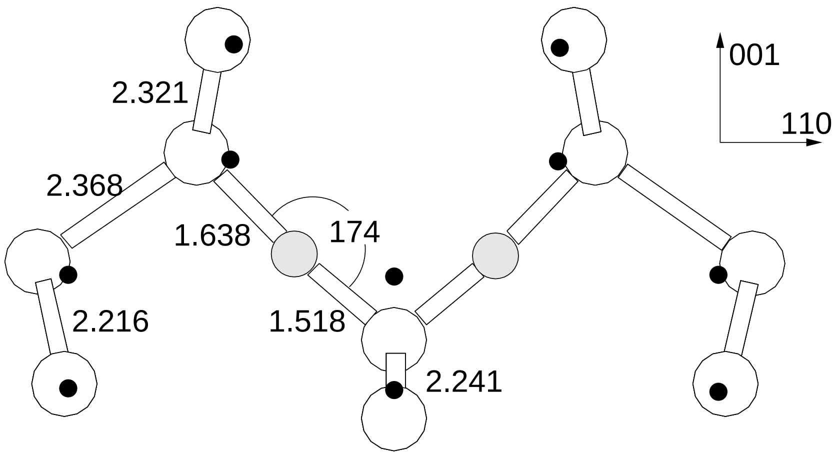


Figure 6.7: The symmetric dimer in silicon (lengths in Å). This is 0.259 eV higher in energy than the assymetric dimer. Black dots mark the ideal lattice sites.

6.4.2 The split dimer

We tried seperating the dimer to give a single Si-Si bond along ⟨110⟩ between the O_i atoms. It was suggested that this might be the source of the 1105 cm^{-1} vibrational mode often seen at the same time as the dimer modes. The resultant structure is shown in Figure 6.8, with vibrational modes in Table 6.4. As can be seen, the modes are not in good agreement with the 1105 cm^{-1} LVM, excluding this as a model. Instead of dropping, the assymetric O_i stretch modes increase in frequency due to the proximity of the atoms. In addition the calculations gave this structure as over 1.5 eV less stable than the puckered dimer, which although probably an overestimate, suggests the defect is not stable in this configuration. The defect is not positioned symmetrically in the cluster, and this manifests itself in the assymetric atomic shifts from their ideal lattice positions, and the split top mode; in the fully symmetric structure this would be degenerate with D_{2h} symmetry.

Earlier calculations showed that the energy of this structure could be lowered by placing the O_i atoms in perpendicular ⟨110⟩ planes, thus lowering the conflict between their stress fields. This roughly corresponds to cis- and trans- isomers. The separation of the O atoms is now $a/8$ (233). However the energy differences between these structures is quite dependant on the basis set used and number of bond centres, and so although this earlier work predicted a binding energy of 0.7 eV for the dimer, we might expect an error on this of 50% or more.

6.4.3 Dimer Migration

We turn now to investigating the movement of the dimer. There are several ways in which a diffusive jump could be executed. We have investigated two mechanisms.

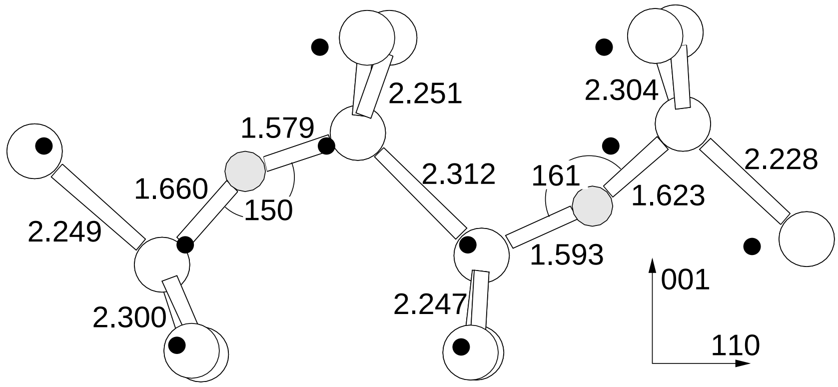


Figure 6.8: The split dimer configuration, *i.e.* two O_i atoms with an empty Si-Si bond between them. All bond lengths in Å. This structure is less stable than the puckered dimer structure.

The first involves cooperative dimer diffusion where one of the O atoms drags the other into a Y-lid configuration, *i.e.* both atoms are in a y-lid configuration at the same time (a Si-O-Si-O square is formed). So far, we have failed to find a low energy route by this mechanism. If the constraints are chosen so that each O atom lies in a Y-lid configuration, this means that the O atoms bond to the Si atoms that are second neighbours to each other, and these two Si-O bonds have equal length. The energy of this configuration is 2.4 eV above that of the stable dimer. One reason for the high energy is that there is a large tensile stress acting on the dimer although this is offset by the absence of any broken Si bonds. It may be possible that there are low energy migration paths by this mechanism but they must involve other configurations.

An alternative diffusive mechanism occurs when the dimer partially dissociates in the $(0\bar{1}1)$ plane. This involves a diffusive jump where one O atom of the dimer moves to its Y-lid configuration and the other O atom is free to remain at its BC site. This configuration was found by imposing a pair of constraints so that the two arms of the Y-lid are equal and the Si atom at the base of the Y-lid is equidistant from the arms (see Figure 6.9). Remarkably, the energy of this structure is only 1.1–1.5 eV above that of the symmetric dimer, *i.e.* 1.36–1.76 eV above the stable puckered dimer. Again the uncertainty in energy is due to the different numbers of bond centres employed. This is in good agreement with the experimental formation energy for thermal donors of ∼1.7 eV, believed to be the dimer migration barrier. The crucial result is that the dimer energy is now lowered by the presence of the other O atom.

The structure of the 'y-lid' saddle point is very revealing (Figure 6.9). The Si-O bonds across the arms of the Y-lid are now strong (bond length of 1.69 Å) while the third Si-O bond is 2.62 Å and essentially broken. This structure of two strong Si-O

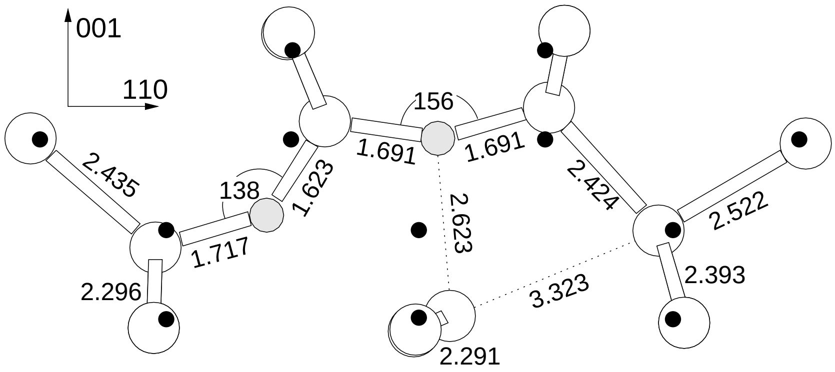


Figure 6.9: A saddle point for oxygen dimer diffusion. This structure is 1.36–1.76 eV higher in energy than the stable puckered dimer structure. Black dots mark ideal lattice sites

bonds has to be contrasted with that of the Y-lid obtained for single O diffusion where only one strong Si-O bond was found. The extra strong Si-O bond has been made because of the compressive force acting on the arms of the Y-lid from the other O atom. It is interesting that a shallow double donor level also occurs in this structure, reinforcing the claim that overcoordinated oxygen atoms are the source of donor activity.

This configuration where the two Si-O bonds have equal lengths is not necessarily the lowest saddle point because of the asymmetry imposed by the second O atom. However this does represent a saddle point, since if the constraints are removed and the structure is moved slightly off the saddle point and allowed to relax, it returns to either one of the equilibrium end points. Thus this represents a saddle point to a low energy migration path for the oxygen dimer. Work is currently underway to explore nearby points on the potential surface and hence determine the potential surface around the barrier.

These results therefore show that an oxygen pair, O_{2i}, can migrate with a significantly lower barrier than a single oxygen atom, O_i. This lowering of the barrier occurs due to compression of the 'y-lid' saddle point structure for one of the O atoms by the presence of the neighbouring O_i, shortening the Si-O bond lengths at the saddle point and hence lowering the energy barrier.

6.5 Dimer Formation

Given a uniform distribution of O_i throughout the sample, statistically there is a chance that some of these atoms will be close enough that they are able to form dimers through standard O_i diffusion. When the O_i concentration is very high, as in Cz-Si, the number of dimers able to form in this way may be appreciable. Taking the dimer concentration equation from [143], and assuming no breakup of

1hr anneal		Equilibrium	450°C	
Temp (°C)	$[O_{2i}]$ (cm^{-3})	$[O_{2i}]$ (cm^{-3})	Time (mins)	$[O_{2i}]$ (cm^{-3})
300	0.53×10^{10}	0.30×10^{16}	5	0.98×10^{13}
320	0.27×10^{11}	0.25×10^{16}	10	0.20×10^{14}
340	0.12×10^{12}	0.21×10^{16}	15	0.29×10^{14}
360	0.51×10^{12}	0.18×10^{16}	20	0.39×10^{14}
380	0.20×10^{13}	0.15×10^{16}	25	0.49×10^{14}
400	0.69×10^{13}	0.13×10^{16}	30	0.59×10^{14}
420	0.23×10^{14}	0.11×10^{16}	35	0.69×10^{14}
440	0.69×10^{14}	0.99×10^{15}	40	0.79×10^{14}
450	0.12×10^{15}	0.93×10^{15}	45	0.88×10^{14}
460	0.20×10^{15}	0.87×10^{15}	50	0.98×10^{14}
480	0.54×10^{15}	0.78×10^{15}	55	0.11×10^{15}
500	0.14×10^{16}	0.69×10^{15}	60	0.12×10^{15}

Table 6.5: Natural dimer concentration as a function of temperature, T, or anneal time, t, assuming no dimer dissociation. $[O_i] = 10^{18}cm^{-3}$, $r_c = 4$ Å, $D_{O_i} = 0.13e^{-2.53/kT}$, $[O_{2i}] = 8\pi r_c D_{O_i}[O_i]^2 t$. For variable temperature data anneal time is set to 1 hour, for variable anneal time the temperature is set to 450°C. Also included is the equilibrium dimer concentration at various temperatures assuming a binding energy of 0.3 eV.

the dimers, then initially:

$$[O_{2i}] = 8\pi r_c D_{O_i}[O_i]^2 t$$
$$D_{O_i} = 0.13e^{-2.53/kT}$$

Taking $[O_i] = 10^{18}cm^{-3}$, and the cut-off radius, r_c, to be 4 Å, we can now vary either the temperature or the anneal time, as shown in Table 6.5 (capture radius, r_c, is normally taken to be the separation at which the interaction energy is equal to kT [158]).

Note that this result does not allow for dimer dissociation. It is useful to compare these figures to the equilibrium dimer concentration, using

$$[O_{2i}]_{equilibrium} = \frac{[O_i]^2}{N} e^{B/kT}$$

where B is the binding energy, and N is the number of BC sites. Taking B=0.3 eV (see above), N= $10^{23}cm^{-3}$, and $[O_i] = 10^{18}cm^{-3}$, we get the results shown in the third column of Table 6.5. These show the approximation of no dimer dissociation breaks down at higher temperatures; at 500°C after a 1 hour anneal a zero dissociation approximation predicts a dimer concentration higher than the equilibrium value.

These results suggest that for a typical 450°C anneal, after an hour we should expect over 10^{14} cm^{-3} dimers. This result assumes a uniform oxygen distribution

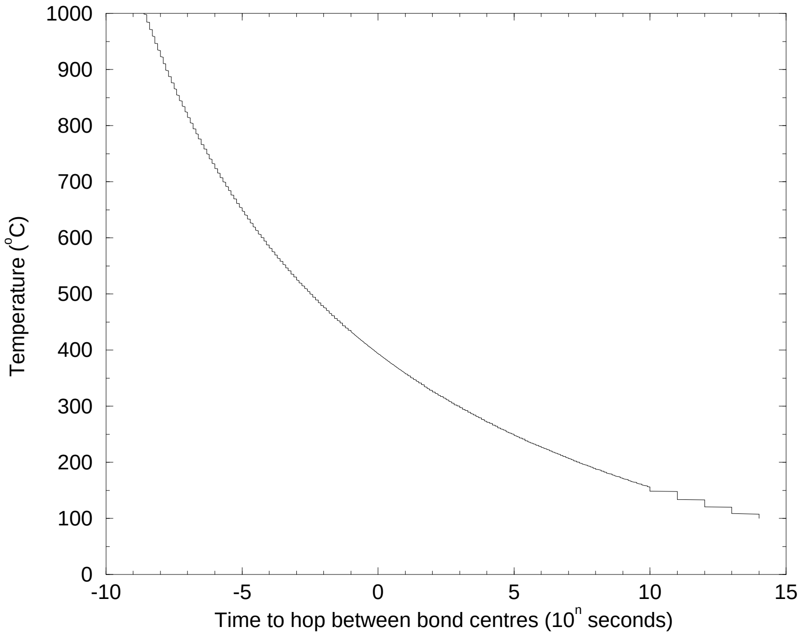


Figure 6.10: Time taken for a single O_i to hop from one bond centred site to another at varying temperature. Calculated using $x = \sqrt{\mathrm{Dt}}$ where the distance between sites, x, comes from AIMPRO cluster calculations and is set to 3.639Å. $\mathrm{D} = 0.13\mathrm{e}^{-2.53/\mathrm{kT}}$.

throughout the material, but in practise this will probably not be the case. Even in quenched anneal experiments, at higher temperatures the oxygen is able to diffuse rapidly from one BC site to another. Treating this diffusion step using simple kinetics, $x = \sqrt{\mathrm{Dt}}$ where x is the distance between BC sites, we obtain hop rates as shown in Figure 6.10.

This suggests that the oxygen concentration need not be uniform. Any localised clustering will increase the dimer concentration figure calculated above, and thus experimental as-grown concentrations of $1.5 \times 10^{15}\mathrm{cm}^{-3}$ are not inconsistent with this analysis [239]. This is also particularly important in heavily carbon doped material, where C_s exerts a long range strain field in the lattice and can attract O_i in this way. It has been observed that the oxygen concentration can vary by a factor of 1.6 over a given sample, independant of cutting direction[159].

This result gives us an initial pool of dimers to work with, which should be present in as-grown material in quantities similar to those observed for the defect responsible for the 1012 cm^{-1} mode. The rate of loss of $[O_i]$ during annealing agrees with that expected for O_i-O_i interaction with the normal O_i diffusion constant [82]. It therefore seems to be the case that the dominant O_i loss will occur through the formation of dimers, and no enhanced diffusion process for dimer formation needs to be invoked (such as trimer migration to O_i followed by decomposition into two dimers).

6.5.1 Experimental Evidence

Lindström and Hallberg annealed Cz-Si samples at 370°C and 450° for 25 hours, and the variation in absorption of the various observed LVMs provide strong support for assignment of the 1012 cm^{-1} mode to the oxygen dimer. Their results are shown in Figure 6.11 and are reproduced from Reference [239]. Note that the bands at 975, 988 and 999 cm^{-1} are associated with the early thermal donors, TD1, TD2 and TD3 respectively.

At 370°C where O_i is not mobile, there is a gradual decrease in the 1012 cm^{-1} dimer band from its residual levels, accompanied with an associated increase in the 975-1006 cm^{-1} bands. These changes have been directly correlated [160], and are consistent with a thermal donor model constructed from dimers (see Chapter 9). There is no notable change in the O_i levels, suggesting the available dimers are becoming depleted as they form higher order complexes, but the lack of O_i migration means that new dimers are not being formed and hence the dimer concentration is dropping.

At 450°C when oxygen is mobile (Figure 6.11b), the story is very different. There is a rapid initial drop in $[O_i]$, accompanied with a rapid rise in dimer concentration. Meanwhile there is also an increase in the other modes. The dimer concentration then tails off, before reaching a steady state growth along with a similar growth in the 975-1006 modes, and an associated drop in $[O_i]$. This can be understood in terms of an initial glut of dimer formation due to O_i atoms which are close to the capture radius required to form a dimer. However once these are exhausted, there is then a drop to steady state O_i diffusion from a more homogenous background level; the behaviour of the 975-1006 modes is then simply a knock-on from the fluctuation in dimer levels (note that this is different from the explanation given in Reference [239]).

6.6 Dimer interaction with other defects

Rapidly diffusing dimer pairs allows us to explain many of the apparently anomalous defect formation behaviour of other complexes in silicon. These have been discussed in the relevant chapters but are summarised briefly here:

1. There are several different formation process for VO_2, and one of these does not result in any loss of $[O_i]$; indeed, O_i is produced when $VO \rightarrow VO_2$ by

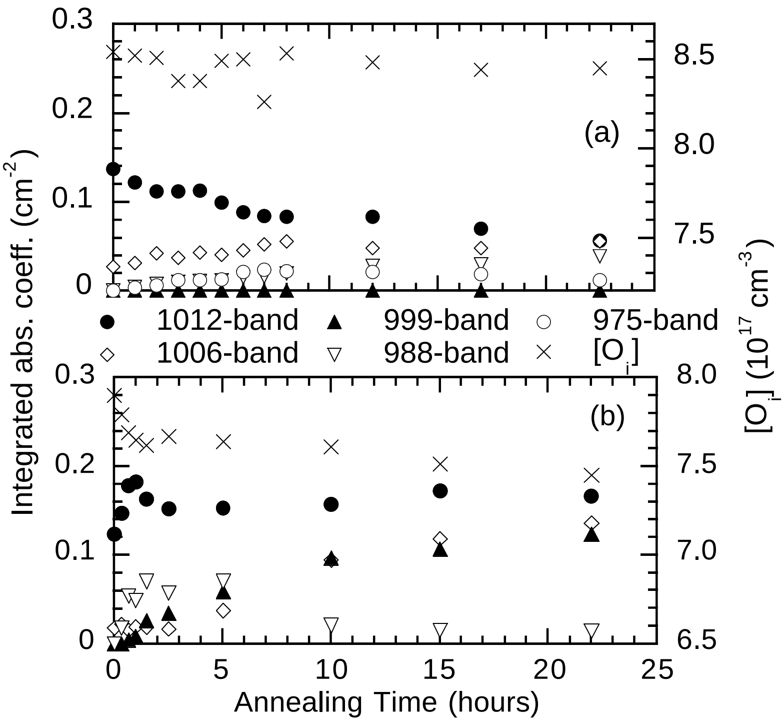


Figure 6.11: The initial changes of the bands 975-1012 cm^{-1} and [O_i] for annealing at (a) 370°C (top) and (b) 450°C (bottom). Taken from Reference [239]. In the 450°C run, the 975 cm^{-1} signal was too weak for inclusion, and has a maximum that appears and disappears too quickly to appear on this scale.

this unknown process. It seems likely this is the breakup of VO followed by trapping of V by a dimer (Chapter 5).

2. The anomalous formation of thermal donors has been discussed briefly here, and in more depth in Chapter 9. It seems that thermal donors are formed through the aggregation of rapidly diffusing dimers.

3. In the early stages of annealing C-rich Cz-Si there is strong C_s and C_sO_i signals; after anneals at 450°C these are replaced with a weaker C_sO_2 signal and stronger C_sO_3 signal [239]. These can form through dimer aggregation to the earlier C_s and C_sO_i defects, and this is discussed further in Chapter 10.

The work performed on NNO defects involved anneals in temperature ranges at which the dimer is no longer stable.

6.7 Trimer formation

It is next important to consider the 1006 cm^{-1} mode. It was suggested earlier that this could be assigned to the trimer, O_{3i}, since it too is normally present with the

1012 cm^{-1} mode, and since O_i is the most common impurity in Cz-Si it seems likely that O_{2i} will complex with this. Of the modes discussed here, only the 1012 and 1006 modes are present in as-grown Cz-Si [239]. This also suggests that it must be one of the earlier complexes to form, with a trimer being the next simplest structure after the dimer. This mode has not been correlated with any of the early TDs (TD1-4).

The 1012 cm^{-1} dimer mode anneals out under first order reaction kinetics [149]. This is similar to the

$$\mathrm{VO} + \mathrm{O_i} \rightarrow \mathrm{VO_2}$$

reaction [79], which also obeys first order reaction kinetics since $[O_i]>>[VO]$ and so effectively remains constant throughout the reaction. Therefore this observation is consistent with a reaction scheme $\mathrm{O_{2i}} + \mathrm{O_i} \rightarrow \mathrm{O_{3i}}$.

Both the 1012 and 1006 behave similarly when exposed to oxygen dispersion treatments (see Figure 6.12). This behaviour is consistent with a dimer / trimer model, since in a 2 hour dispersion treatment at 1100°C, the integrated intensity level of the 1012 cm^{-1} level halves and the 1006 cm^{-1} practically disappears. In the following anneal at 390°C the 1012 level then drops further as the dimers are not being replaced by O_i diffusion. However the 1006 rapidly starts to reappear. This seems inconsistent if it is assigned to a large defect centre but follows naturally if it is assigned to trimers that are forming from the remaining supply of dimers in the material. The graph is also useful in later discussions of TD1, since the dispersion treatment effectively removes the as-grown 'pool' of trimers which has a knock-on effect on the TD1 formation rate. We therefore conclude that the 1006 cm^{-1} can be assigned to the oxygen trimer, O_{3i}.

6.7.1 Theoretical Results

This section covers our work to date, however further work is required before we can make any definitive statements on the nature of the trimer.

There are many possible configurations for three oxygen atoms. Electrically active configurations will be examined in Chapter 9 so we restrict ourselves to neutral complexes here. In this case there are two primary possible structures, a linear chain (an extension of the dimer along $\langle 110 \rangle$, see Figure 6.14), or placing all three O_i atoms around the same core Si as shown in Figure 6.13 (in this structure the oxygens all bow 'clockwise' around the defects C_3 axis, earning it the name the 'Manx' defect!). We relaxed both of these structures in a 134 atom cluster, $Si_{71}H_{60}O_3$, and found the Manx structure to be 0.248 eV more stable. There is in addition a third possible structure consisting of a dimer with an appended oxygen

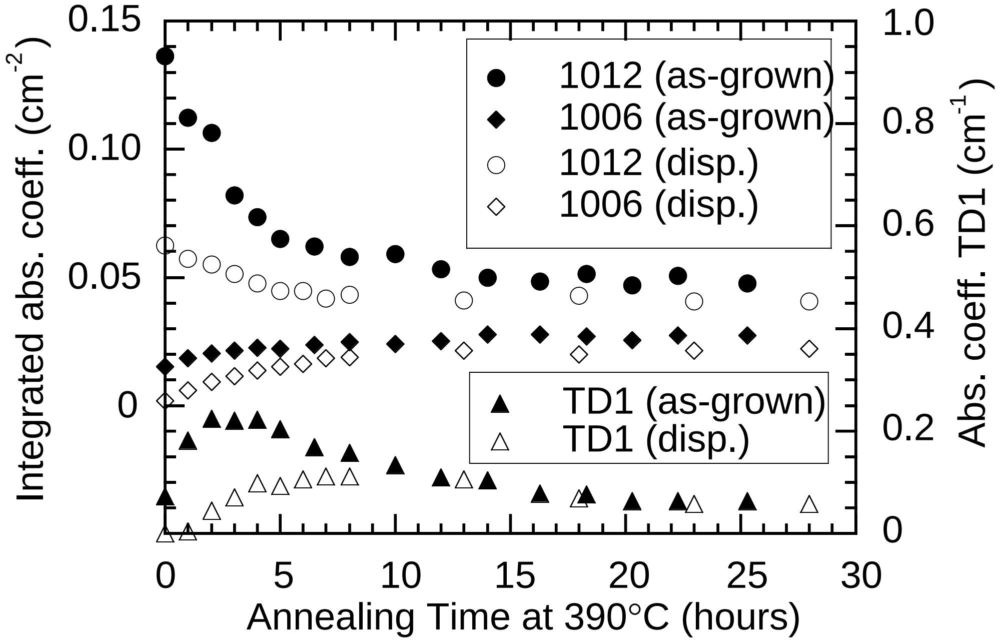


Figure 6.12: The initial changes of the 1006 and 1012 bands during the adjustments to their respective equilibrium levels and the growth of TD+1 (1s-3p±) for an as-grown sample and a dispersed sample (2h at 1100°C). Note that the TD data applies to the r.h.axis, and the 1006/1012 data to the left, hence the initial dispersed 1006 level is almost zero. Taken from Reference [239].

atom that sits in a perpendicular $\langle 110 \rangle$ plane, thus oxygen atoms sit in three of the six bonds around a hexagonal ring site. This may be more stable as the stress will be spread over two planes. Nothing is currently known experimentally concerning the symmetry of the trimer.

The vibrational modes of the linear and Manx structures are given in Table 6.7.1. The linear structure gives modes which are close to those of the dimer, although these calculations would suggest slightly higher. The Manx structure has modes that are very different. This suggests that the trimer may be the linear O_{3i} structure, since experimentally its vibrational mode at 1006 cm^{-1} is only 6 cm^{-1} away from the primary dimer vibration at 1012 cm^{-1}. Although these calculations predict a slight increase in this mode as opposed to a drop, that could be a cluster size effect, as the end of the trimer chain is quite close to the cluster edge. Alternatively, the 'half hexagon' structure described above might be expected to have perturbed dimer-like modes.

6.8 Conclusions

In this chapter we have shown that there is a binding energy between two O_i atoms to form an assymetric oxygen dimer, O_{2i}. Our calculations predict a binding energy in the range 0.7 $\rightarrow$ 1.7 eV. This assymetric dimer is 0.26 eV more stable than a

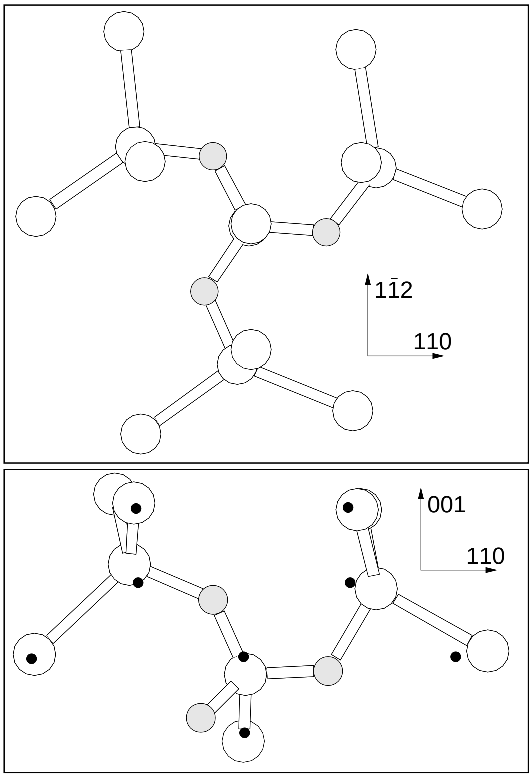


Figure 6.13: The 'manx' oxygen trimer in silicon. The top diagram shows the view along the $\langle\bar{1}11\rangle$, C_3 symmetry axis.

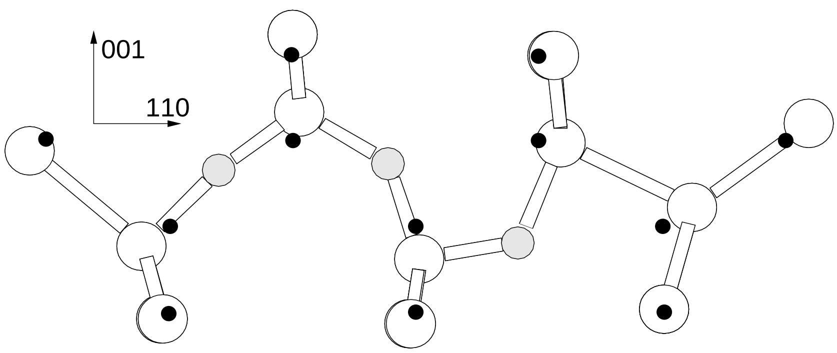


Figure 6.14: The linear chain trimer in silicon; dots mark ideal Si lattice sites.

'Linear Chain' structure			'Manx' structure		
^{16}O	^{17}O	^{18}O	^{16}O	^{17}O	^{18}O
1080.5	24.4	46.5	1176.8	23.6	44.7
948.9	21.6	41.1	1154.3	23.6	44.9
849.9	20.3	38.7	978.2	27.0	51.6
733.3	11.3	21.2	737.9	12.5	23.6
655.7	13.2	24.8	713.5	10.4	19.8
630.1	5.4	10.5	598.4	8.0	14.5
570.5	8.3	15.4	555.1	7.5	14.7
552.2	1.9	4.2	511.4	6.1	11.8
525.3	4.4	8.2			

Table 6.6: LVMs and isotope shifts (cm^{-1}) for the trimer in the $\langle 110 \rangle$ linear chain structure, and shared central Si 'Manx' structure. The 'Manx' structure is 0.248 eV more stable.

symmetric dimer pair, previously predicted using AIMPRO and other methods. We have shown that the dimer can migrate with a barrier substantially lower than that of isolated O_i.

In addition we have shown that it has vibrational modes in excellent agreement with the experimentally observed 1012 cm^{-1} and associated modes. An additional mode at 1105 cm^{-1} has been tentatively correlated with the symmetric dimer. We have attempted to show how the dimer is crucial in the formation processes of many oxygen related defects in Cz-Si, that form in the temperature range ~300-450°C in which the dimer is stable.

Fast diffusing dimers will rapidly encounter relatively slow moving single O_i to produce an oxygen trimer, O_{3i}. We have performed some preliminary investigations of possible trimer structures, which show three bond centred O_i atoms around a shared Si atom to have the lowest energy by 0.248 eV, but a linear structure to have vibrational modes closer to experiment. The role of dimers and trimers in thermal donor formation are discussed further in Chapter 9.

Chapter 7

Nitrogen–Oxygen defects in Silicon

Nitrogen can be incorporated into Si in a variety of ways, although it only has a low solubility. Anneals in an ambient environment usually use N_2 gas. Si_3N_4 is used as a surface diffusion barrier or as a dielectric layer in metal insulator semiconductor systems. In addition, N doping in Cz-Si suppresses vacancy and interstitial defects and increases gate oxide integrity. It suppresses 'swirl defects' and stops wafer bowing during processing, as well as suppressing thermal donor formation. Many of these properties are dependent on nitrogen–oxygen interactions.

In this chapter we study the interaction between nitrogen and oxygen in silicon. Previous calculations using AIMPRO have shown that N in Si exists primarily in the form of interstitial $(N_i)_2$ pairs [145]. Isolated N_i has been observed experimentally but only in small quantities, and it rapidly diffuses to form $(N_i)_2$. It is shown here that the primary N—O defect is $(N_i)_2O$, or NNO, consisting of O_i neighbouring a $(N_i)_2$ pair [103]. This is electrically inactive and can reversibly break down to give $(N_i)_2$ and O_i. We also examine a N_iO_i complex and show that small quantities of this are also present over similar temperature ranges. Electrically active N_iO_{2i} complexes are considered in Chapter 8.

7.1 Method

All the defects considered here have been symmetry constrained - C_{1h} for N_iO_i, $N_{2i}O$, C_{2v} for N_i, etc. unless stated otherwise. However we also breached the symmetry of the defects and allowed them to relax; in every case they returned to the constrained symmetry structure. A variety of different cluster sizes have been used. Initial results were obtained using 113 atom clusters, and the final results shown here come from 134 atom clusters for the NNO work, and 148-152 atom clusters for the rest. The results are qualitatively identical, showing them to be

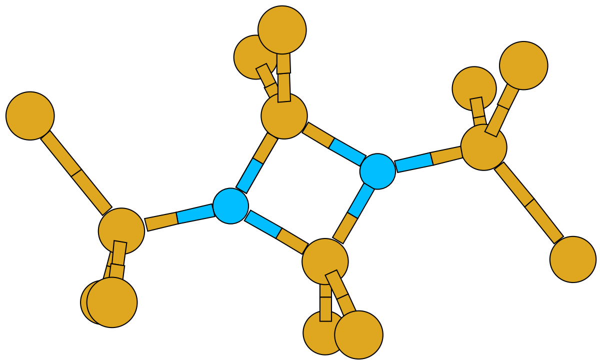


Figure 7.1: The N-pair defect in Si. Vertical direction is $\langle 001\rangle$, horizontal is $\langle 110\rangle$

independent of cluster size.

7.2 Background

Nitrogen is commonly used as an inert atmosphere for processing Si, but it has important effects on the material. For example, N impurities can pin dislocations [161, 162] and form electrically active defects such as the substitutional deep donor [163, 164]. There have been suggestions that nitrogen-oxygen complexes can form shallow donors [164, 165, 166], although whether N forms an essential component of these defects is controversial [167, 168], and this is discussed further in Chapter 8.

Nitrogen can be introduced into Si by adding Si_3N_4 to the melt, annealing in an atmosphere of N_2, or by implantation. N is known to enhance O precipitation and inhibit C enhanced O precipitation at 750°C [169]. Photoluminescence (PL) experiments showed a decrease in oxygen precipitate strain-induced defects in samples containing a higher N content.

The principal N defect in O-free silicon is a nitrogen pair consisting of two adjacent [100] oriented nitrogen interstitials, N_i, with C_{2h} symmetry (see Figure 7.1) [145, 170, 145]. It is stable until about 800°C [170]. Isolated N_i defects have also been observed (see Section 7.3), and although substitutional N_s is a rare defect it has been detected by EPR [171] and local vibrational mode spectroscopy [172].

In N and O implanted FZ-Si [173], or in Cz-Si grown in a N_2 atmosphere [174, 175], N–O complexes are formed by annealing between 400 and 700°C. Under these conditions, the intensity of the localised vibrational modes (LVMs) due to the pair are progressively reduced and three other IR-absorption lines at 1026, 996 and 801 cm^{-1} (room temperature) become increasingly prominent. The intensities of these lines are correlated with each other [175] and they are believed to be due to NNO complexes. This is because O_i atoms become mobile around 400°C and become trapped by N-pair defects which are stable at this temperature. Beyond $\approx$ 700°C, these complexes break up with the emission of O_i leaving the N-pair. Cooling allows the NNO complexes to reform. This dissociation and reformation is reversible.

7.3 The N_i defect

Isolated N_i defects also exist in implanted material, but lie below the detection limit in doped as-grown material. They possess a local vibrational mode (LVM) around 691 cm^{-1} [170], which is in close agreement with the calculated N stretch mode of the [100] oriented nitrogen-silicon split interstitial at 700 cm^{-1} [145] (see Table 7.2). In this defect N and Si share a lattice site. Berg Rasmussen *et. al.* [176] observed N_i in N-implanted Si singly annealed at 600°C. They concluded that N_i was mobile during annealing, and rapid cooling was necessary if N was to be trapped as N_i rather than forming other defect complexes.

N_i is a [100] split interstitial, structurally similar to interstitial carbon, C_i. N is much more electronegative than Si, and hence draws charge from its neighbouring atoms. However, unlike C_i, its lone pair is already filled, and so it cannot remove the excess electron from the neighbouring undercoordinated silicon radical. The wavefunction of this state is small in the vicinity of the N atom and lies deep in the gap (see Figure 8.2). Like C, the small size of the N atom suggests that the Si-Si bonds along [011] are dilated, and are therefore sites where O atoms would preferentially precipitate.

By pairing up as $(N_i)_2$ this leaves all the Si and N atoms fully coordinated and removes any associated electrical activity. However the localised lattice polarisation due to the strong electronegativity of N remains, as does the tensile lattice strain.

A single interstitial N atom was inserted into a 148 atom cluster, $Si_{79}H_{68}N$ and all atoms were allowed to relax. The Kohn-Sham eigenvalues are given in Figure 8.2, and it can be seen that the single donor state lies quite low in the gap. The top N–Si bonds oriented along [011] are 1.83 Å and the [100] N-Si bond is 1.75 Å. These Si atoms are pulled towards the core leading to a dilation of the Si-Si backbonds (the two $\langle 110\rangle$ back-bonds are dilated to 2.397 Å, and the $\langle 1\bar{1}0\rangle$ pair above the defect are compressed to 2.293 Å; these compare to ideal Si-Si bond lengths of 2.349 Å). This makes these bonds attractive sites for oxygen precipitation. Vibrational modes of the defect have been given earlier [145] and are in good agreement with an observed 691 cm^{-1} line (see Table 7.2).

7.4 $N_{2i}O$ – The dominant N–O defect in silicon

In N and O implanted FZ-Si [173], or in Cz-Si grown in a N_2 atmosphere [174, 175], NNO complexes are formed by annealing between 400 and 700°C. Under these conditions, the intensity of the localised vibrational modes (LVMs) due to the N-pair are progressively reduced and three other IR-absorption lines at 1026, 996 and 801 cm^{-1} (room temperature) become increasingly prominent. The intensities of these lines are correlated with each other [175] and they are believed to be due to NNO complexes. This is because O_i atoms become mobile around 400°C and become trapped by N-pair defects which are stable at this temperature. Beyond ≈ 700°C these complexes break up with the emission of O_i leaving the N-pair. Cooling allows the NNO complexes to reform. This dissociation and reformation is reversible.

7.4.1 Experimental Work

There were no reports of isotopic effects on the LVMs of the NNO defect even though this information is of the greatest importance in deducing its structure. Therefore concurrently with our theoretical investigation, experimental IR studies were performed by Berg Rasmussen *et al* [177]. Float-zone n-type Si samples were implanted at room temperature with various combinations of ^{16}O, ^{17}O, ^{14}N and ^{15}N, with a concentration of 6.5×10^{19} cm^{-3} for both O and N. All samples were annealed under flowing N_2 ambient for an hour in the temperature range from 500 to 650°C. Infrared absorption measurements (shown in Figure 7.2) were carried out at 77 K. The LVMs due to the NNO defect are given in Table 7.1.

The ^{14}N frequencies agree with the earlier studies [174, 175, 178], allowing for their shift (≈ 4 cm^{-1}) between low and room temperature. All the modes shift with ^{15}N showing that at least one N atom is involved in the defect. Unexpectedly, the peak in the 1030 cm^{-1} LVM is shifted by around 3 cm^{-1} with ^{15}N and this suggests that this mode is almost entirely due to Si-O stretch. This lies between the LVMs for O_i (1136 cm^{-1}[179]) and O_s (836 cm^{-1}[77]), suggesting that O_i bridges a dilated Si bond near the N-pair. The modes at 999 and 805 cm^{-1} are slightly above those of an isolated N-pair (967 and 770 cm^{-1}) and show a similar shift with ^{15}N (25 and 19 cm^{-1} compared to 26 and 17 cm^{-1} for the N-pair), suggesting that the pair remains intact but with compressed bonds due to the nearby O_i. No additional LVMs are observed in the mixed implantation, implying that the N atoms are no longer dynamically coupled but vibrate independently of each other. They must also be inequivalent atoms as they possess distinct LVMs. This implies that O lies closer to one N atom than the other.

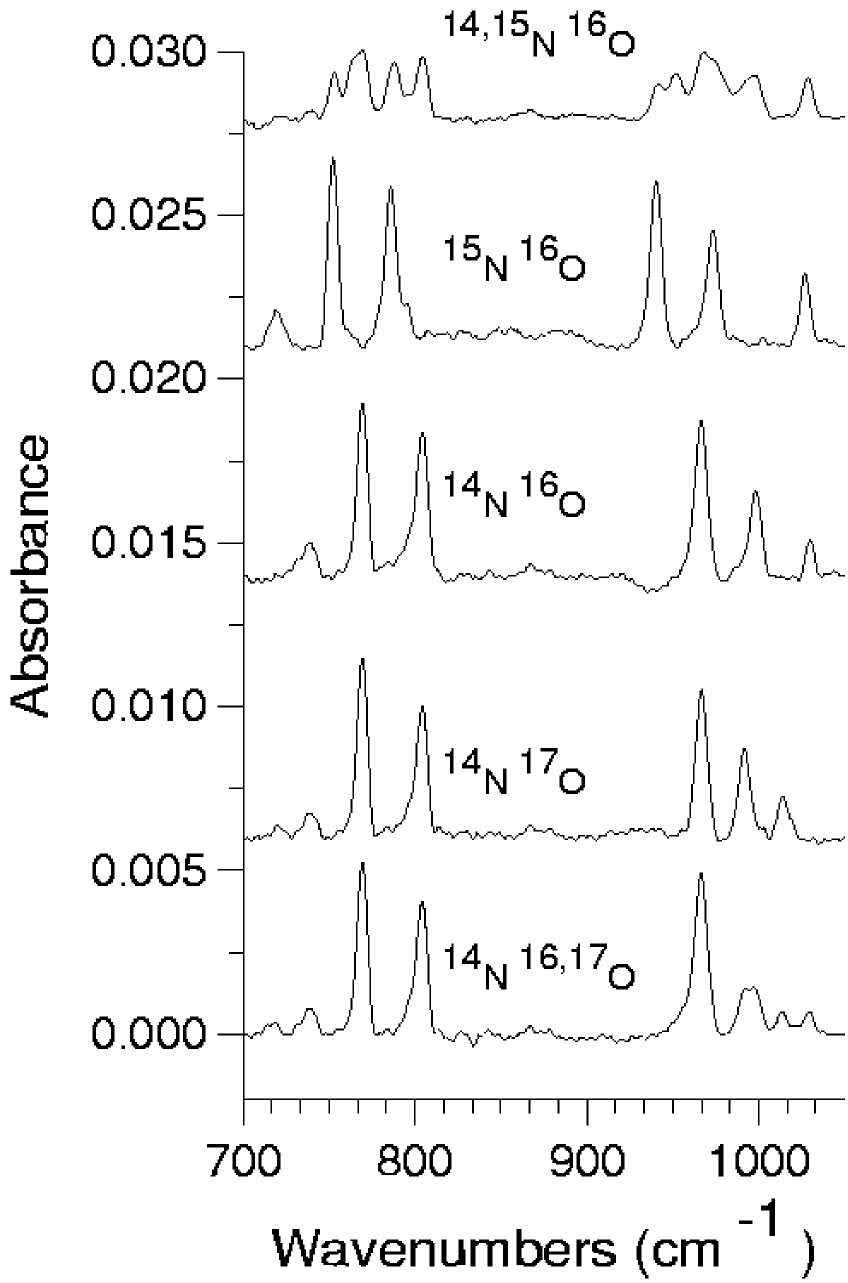


Figure 7.2: Absorbance spectrum observed after annealing at 650°C for 1 hour, of samples with various combinations of ^{14}N, ^{15}N, ^{16}O and ^{17}O. In all cases the total N and O doses are equal. Data obtained by Berg Rasmussen *et al* [103]

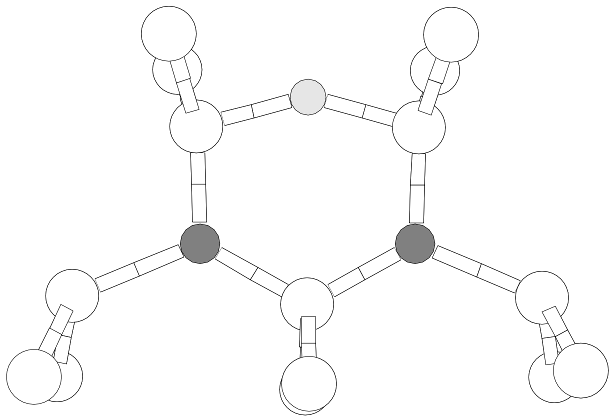


Figure 7.3: The alternative 'Humble Ring' model of NNO. This was discarded since it did not match experimental data.

7.4.2 Theoretical Work

Calculations on the N pair defect [145] showed that N_i dilated the Si(1)-Si(5) bond (Fig. 7.4) by 4%, making it a prime site for O_i. The other Si bonds bordering the pair along [100] are compressed making these unlikely sites for the O_i. Previous investigations on the trapping of O_i at dislocation cores showed a marked preference for O to bridge dilated Si bonds [156].

We investigated several possible configurations for the NNO defect, using a 134 atom cluster, $Si_{71}H_{60}N_2O$, and found that the structure shown in Figure 7.4 was the only one that fitted the experimental data. Moving the O atom in this model to bridge the second nearest Si-Si bond along [011] between Si(5) and Si(6) was 1.2eV less stable. Neither moving the O_i to lie between N(1) and Si(3) in Figure 7.4 or using O to bridge the dilated bond in the Humble N-pair model [145] (Figure 7.3) gave modes in the right region or produced the correct isotope shifts. Our conclusion is that the structure in Figure 7.4 represents the NNO defect.

The vibrational modes of the structure are given in Table 7.1. The two modes at 1070 and 808 cm^{-1} are clearly N-related, as they barely shift with the O-isotope, whereas the 861 cm^{-1} mode is O-related. The mixed isotopic results show very little coupling between the N-related modes. This explains why no new modes are introduced in this case, since the nearby O_i makes the two N atoms inequivalent.

Thus although the modes are in the correct region and display the correct isotopic behaviour, their ordering is incorrect: the O mode should lie above the two highest N modes. This ordering and their positions are, however, very sensitive to the position of the common Si atom, Si(1) (Figure 7.4). If this is displaced by 0.07 Å along $[0\bar{1}\bar{1}]$, decreasing the Si(1)-O and increasing the N(1)-Si(1) lengths by 0.05 and 0.07 Å respectively, then the O and N modes are within 70 cm^{-1} of

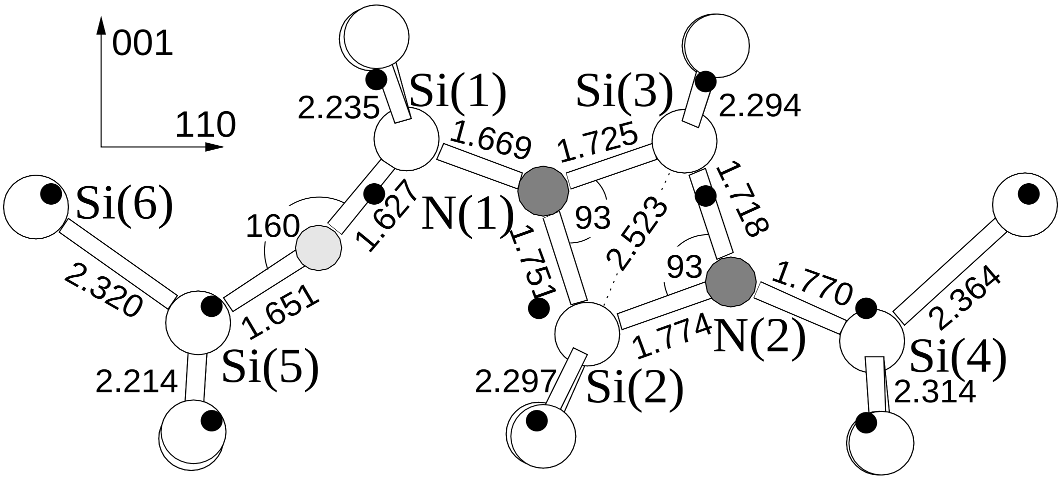


Figure 7.4: Structure of the NNO defect. All bond lengths in Å.

the observed frequencies (Table 7.1). Once again the N atoms are decoupled. The changes in bond lengths caused by the shift of Si(1) and the discrepancies between the calculated and observed modes lie within the errors of the calculation.

The highest O-mode is simply an asymmetric stretch in the Si-O bonds, with a symmetric stretch mode at lower frequency. It shifts downwards by 23 cm^{-1} when ^{16}O is substituted by ^{17}O, which is somewhat larger than the observed shift of 17 cm^{-1}. The calculation found no O-shift of the 949 cm^{-1} mode whereas a downwards shift of 8 cm^{-1} is observed (Table 7.1). Note however, that with the displaced Si(1)-atom the calculated O-mode and highest N-mode cross when ^{16}O is substituted by ^{17}O. Such a crossing is not observed experimentally. The extreme sensitivity of the position of the common Si-atom (Si(1)) and the possibility of crossing of the two highest modes makes it rather difficult to calculate the isotopic shifts for these modes. In the calculations almost no interaction between the two highest modes were found. However, just a small interaction would automatically cause a smaller O-shift of the 963 cm^{-1} mode and a (larger) O-shift of the 949 cm^{-1} mode. Contrary to this, the mode at 813 cm^{-1} shows no shift with ^{17}O in close agreement with the observed shift of 1 cm^{-1}.

Consider now the shifts with N isotopes. Substituting ^{14}N by ^{15}N, the calculated shifts of the 963, 949 and 813 cm^{-1} modes are 0, 29 and 21 respectively in good agreement with the observed shifts of 2, 25 and 19 cm^{-1}. Finally, a mode involving motion of both N-atoms were found at 727 cm^{-1} in the calculation. The 727 cm^{-1} shifts 19 cm^{-1} with ^{15}N but does not show any shift with O. The position and shift of the 727 cm^{-1} calculated mode is extremely close to the observed mode at 739 cm^{-1} that displays a 18 cm^{-1} shift with ^{15}N and no resolvable shift with O. Furthermore, the 727 cm^{-1} mode is expected to split into four components when

	^{14}N ^{16}O	^{15}N ^{16}O	^{14}N ^{15}N ^{16}O	^{15}N ^{14}N ^{16}O	^{14}N ^{17}O	^{14}N ^{18}O
Calculated						
N(1)	1070.4	30.8	0.6	30.1	0.1	0.1
O	861.1	1.0	0.4	0.7	20.6	37.2
N(2)	808.2	20.1	20.0	0.1	1.3	3.7
	723.9	18.4	15.6	1.7	0.4	0.0
	671.5	15.7	2.0	14.8	2.2	0.9
Displaced Si(1) atom						
O	962.9	0.2	0.1	0.1	23.4	44.0
N(1)	948.9	28.8	0.9	27.6	-0.5	-0.3
N(2)	813.3	20.7	20.6	0.3	0.0	0.1
	727.0	18.8	13.0	3.1	0.0	0.0
	688.1	15.8	4.6	13.9	0.4	0.7
Observed						
O	1030	3	2		17	
N(1)	999	25	0, 25		8	
N(2)	805	19	0, 19		1	
	739	28			0	

Table 7.1: Calculated and Observed LVMs, cm^{-1}, due to the NNO defect in Si. Isotopic values show drop in modes when different atomic isotopes are used. For modes where primarily one atom is moving, this atom is given in the first column of the table (numbering refers to Figure 7.4).

both N-isotopes are present (Table 7.1). Considering the small intensity of the observed mode at 739 cm^{-1}, this explains why this mode is not observed in samples implanted with both ^{14}N and ^{15}N.

7.5 N_iO_i – pre-cursor to STDs and the NNO

Wagner et al [174] have reported two weak LVMs at 810 and 1018 cm^{-1} (813 and 1020 cm^{-1} at room temperature [100]) which were seen in the same materials as the NNO defects. These were also reproduced in experiments by Berg Rasmussen *et al* [103] in implanted Si after annealing at 600°C. These modes are correlated, and disappear after annealing at 650°C. Long time annealing at 600°C increases their concentration at the expense of N_2 modes and the O defect responsible for the 1012 cm^{-1} mode, suggesting that oxygen and nitrogen complexes are breaking up while this defect is forming. They are always of a much weaker intensity than the NNO vibrational modes. These modes have not to our knowledge ever been properly characterised, and their origin is unknown.

After high temperature pre-annealing even the N_2 pairs are broken up, and it seems reasonable that in material with a high O_i concentration, some of the N_i could diffuse to O_i before reaching a second N_i and reforming a pair. This N_iO_i

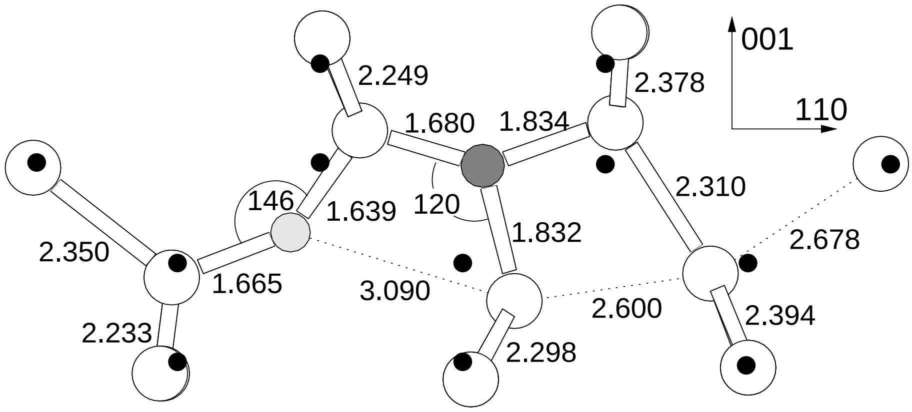


Figure 7.5: The N_iO_i defect in the neutral charge state. Lengths in Å, angles in degrees.

would form as an intermediate structure between N_i and either N_iO_{2i} or $N_{2i}O_i$, and is a possible source of the 810 and 1018 cm^{-1} vibrational modes (there are other alternatives, see below). If this is the source of the modes it would be important since it shows clear evidence of a N–O precursor to the N_iO_{2i} defect, and adds much support to the claim that the N_iO_{2i} shallow thermal donor can form in silicon (see Chapter 8).

We examined N_i–O_i using the 149 atom cluster $Si_{79}H_{68}NO$. In the neutral charge state the defect has roughly planar C_{1h} symmetry with a deep donor level close to that of isolated N_i (see Figure 8.2). This eliminates it as a model for the shallow thermal donor. The N atom moves very slightly out of plane by 0.156 Å which allows its lone pair to become more assymetric, similar to the nitrogen coordination in ammonia, NH_3. This is discussed further below. The O atom sits in a stable bond-centred location with Si-O lengths of 1.639 and 1.665 Å with the shorter bond on the Si shared with the N, and a Si-O-Si bond angle of 146° (see Figure 7.5). The N-Si bonds are 1.680, 1.834 and 1.832 Å, with the shorter bond once more shared with the Si bonded to the oxygen.

The deep donor level is concentrated on a $\langle 110 \rangle$ $p-$ type orbital located on the central Si atom. This atom moves away from the O. The asymmetry of the defect allows the donor wavefunction to concentrate on the lobe of the $p-$type orbital furthest from the oxygen, thus minimising the Coulombic repulsion with O. This explains why the donor level remains deep in the gap. A plot of the Kohn-Sham wavefunction for this level is given in Figure 7.7. This electronic behaviour is discussed further in Section 8.10.

Since the defect possesses such a deep level it raises the possibility that it could act as an acceptor, trapping an additional electron to fully populate this level. The implantation used to add isotopic N/O creates radiation defects, pinning the

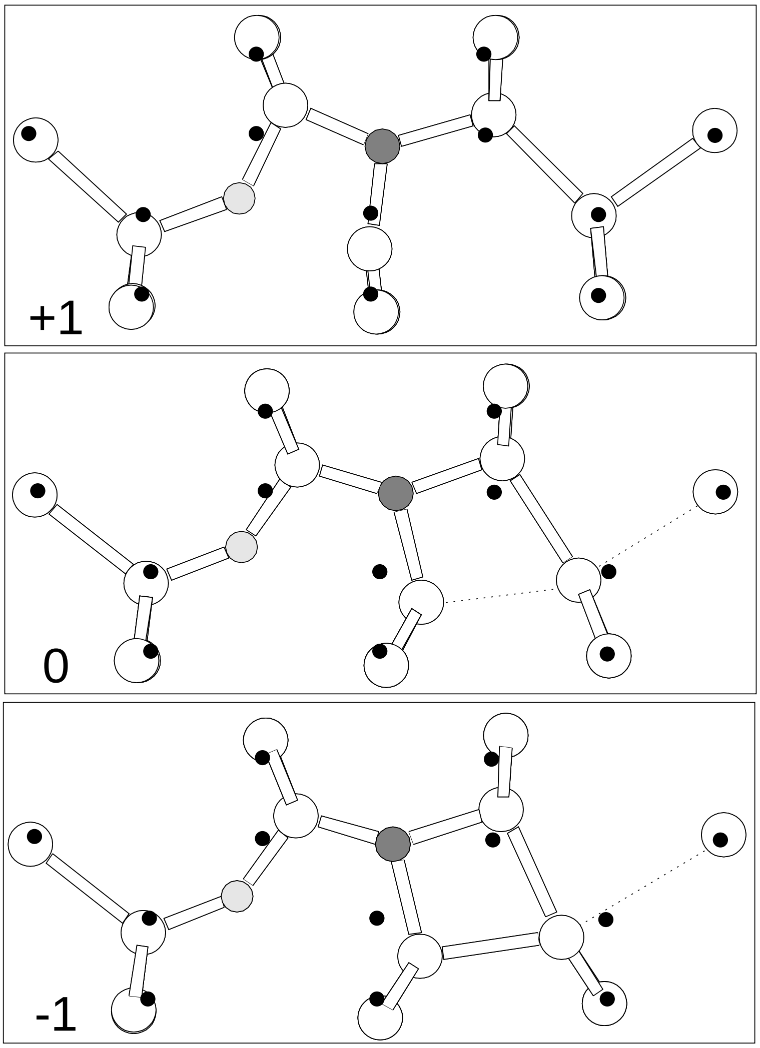


Figure 7.6: The N_iO_i defect in the +1, neutral, and -1 charge state. The core Si atom has a $p-$type orbital in the plane of the defect that is empty in the +1 charge state. This is partially and fully occupied in the neutral and -1 charge states respectively, which is reflected in the change in bonding character.

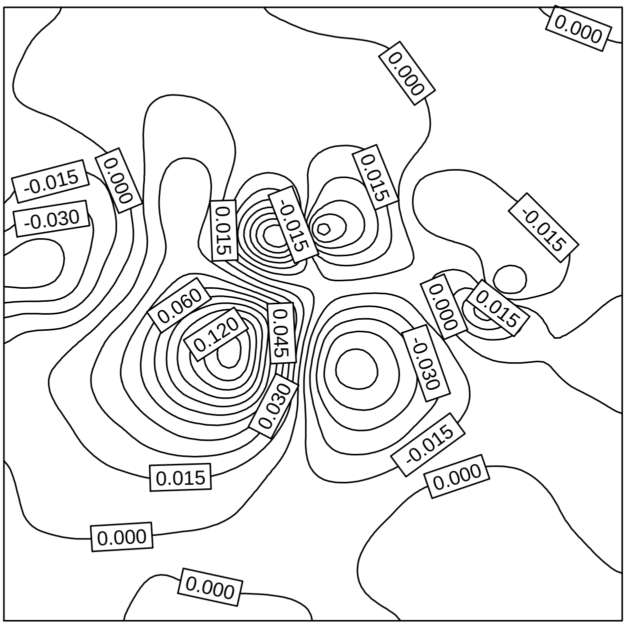


Figure 7.7: Kohn-Sham wavefunction for the partially occupied level of N_iO_i in the neutral charge state. Note that the diagram is reversed from the ball and stick diagrams, *i.e.* the O atom is on the right of the plot. The wavefunction lies primarily on the core Si atom, and is localised on the side away from the O_i atom.

fermi-level near to midgap. Both the isotope implanted and as-grown material were *n*-type[100] with no difference in modes, which suggests the defects possessed the same charge state in both. This suggests that the defects would be negatively charged. We therefore calculated the structure and vibrational modes of the defect in both the neutral and negative charge states.

In the -1 charge state the core Si atom starts to form a bond with its Si next neighbour on the opposite side to the oxygen; this bond shortens to 2.35 Å, dilating the Si-Si back bond to 3.16 Å (see Figure 7.5). Thus the lone pair on the core Si appears to have migrated out of the defect core and onto a next neighbour Si atom of the defect. There is little change in the Si-O-Si lengths (1.63 and 1.65 Å, 153° bond angle), however the N-Si lengths dilate as the core Si is pulled away from the N, the new lengths being 1.66, 1.86 and 1.93 Å.

The change in bonding with charge state is shown in Figure 7.6, and shows an interesting effect. In the +1 charge state the core $p-$type orbital is empty and thus there is an attraction between it and the $O^{\delta-}$ atom. It therefore shifts off site towards the oxygen. However in the neutral charge state this orbital is occupied and so there is now a driving force to form a covalent bond with its Si next-neighbour on the other side. Since this Si atom is fully coordinated already then it has to choose between bonding to the core Si and bonding to its standard neighbour along $\langle 110 \rangle$.

In practise it appears to form an extended bond between the two with roughly equal bond lengths (see Figure 7.5). Finally in the -1 charge state there are now sufficient electrons to completely fill the dangling $p-$type bond. However it is energetically more favourable to form instead a Si-Si bond with this next neighbour along $\langle 110 \rangle$, thus creating a dangling bond on the Si atom further from the defect core. This is probably since the dangling bond, if it were to stay on the core Si, would have to sit quite close to the oxygen atom, so the defect can lower the overall Coulombic repulsion by moving it out of the defect centre in this way. Whether it would be possible to grow this defect in p-type material and observe the shifted +1 behaviour is an open question.

The vibrational modes are given in Table 7.2. As can be seen, there is extremely good agreement between the calculated and experimental values. The change in charge state has only a small effect on the modes. The modes for the -1 charge state are slightly better than the neutral result and add weight to the suggestion that the defect is negatively charged, however there is not sufficient difference to allow us to distinguish the defect charge state purely on this basis. The experimental resolution was 6 cm^{-1} and most of the isotopic shifts lie well within this accuracy. The 805 cm^{-1} mode appears experimentally as a shoulder to the stronger 813 cm^{-1} mode due to NNO, and as this barely shifts with O isotope it could mask any movement in the 805 cm^{-1} mode.

Both the N and O atoms have one very short Si bond with their mutually shared Si atom. This is extremely similar to the oxygen pair (see Chapter 6), where the two oxygen atoms bind tightly to their shared Si atom. This is due to the large electronegativities of both the oxygen and nitrogen atom, which draw charge from their Si neighbours, making the shared Si atom strongly $\delta+$. However the binding for the N_iO_i defect will be much stronger than that of the oxygen dimer, since as well as this electrostatic bonding there is the additional binding due to strain compensation with is absent in the dimer. This is in agreement with their dissociation temperatures; oxygen dimers are not important in TD formation above about 450°C whereas the N_iO_i modes anneal out at 650°C.

We originally relaxed this defect with the N atom starting exactly in-plane, but this is a metastable minimum. Allowing the N extremely small out-of-plane relaxation allows the wavefunction of its lone pair to distribute correctly. Without this the N does not attract as much surrounding charge, decreasing the polarisation of the surrounding lattice and so weakening the strong shared bonds with the neighbouring Si in the N-Si-O chain. This then shifts their vibrational modes, leading to an incorrect ordering of the N and O modes. These earlier modes are included in Table 7.2 for reference.

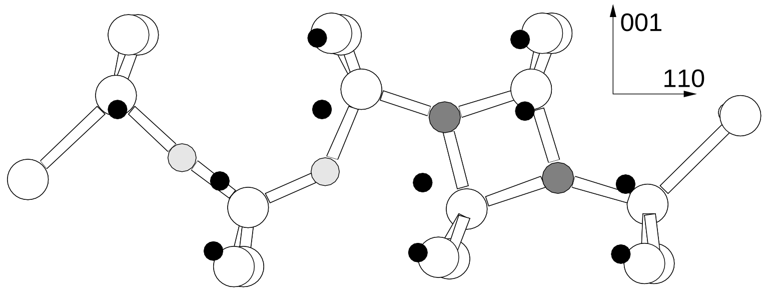


Figure 7.8: The NNOO defect.

Note that these results for N_iO_i differ to those given in earlier work performed using semi-empirical PM3 clusters[180], where it was suggested that N_iO_i represented a single, rather shallow, EMT donor.

7.5.1 Alternative Models - NNOO

There are other alternative explanations for this defect that need to be investigated. Oxygen dimers are present in the material, and are probably responsible for the 1012 cm^{-1} absorption line (see Chapter 6); one explanation could be that these are simply diffusing to N pairs and forming $N_{2i}O_{2i}$ defects. It is known that N suppresses thermal donor formation, and the normal explanation for this is that it simply 'soaks up ' the O_i by forming NNO defects. However the $N_{2i}O_{2i}$ defect provides an intriguing alternative, whereby if dimer migration is the crucial step in thermal donor formation (see Chapter 9), then thermal donors can be suppressed by trapping these to form $N_{2i}O_{2i}$.

To investigate this, we relaxed the 151 atom cluster, $Si_{79}H_{68}N_2O_2$, which contains the most likely structure for NNOO, i.e. the standard [001] split interstitial N_2 pair neighbouring an O_{2i} dimer; the result is a defect lying along $\langle 110\rangle$, with C_s planar symmetry. The structure is extremely similar to NNO with the shared Si atoms between the N_i moved a long way off site and the O_i atoms only slightly distorted (see Figure 7.8). As expected, it possesses no gap levels. The furthest N_i has bond lengths of 1.70 Å and 1.76 Å within the square (bond angle 92°) and 1.73 Å outside, with the inner N_i having lengths of 1.68 and 1.75 Å within the square and 1.62 Å to the Si neighbouring O. The oxygen nearest the square bows a long way from its bond centred site (136° Si-Ô-Si bond angle), whereas the outer O_i has standard bond angles with assymetric bond lengths characteristic of the dimer (see Chapter 6).

This similarity to both the dimer and the N-square suggests that the NNOO defect will not have modes in the correct region, as shown by the calculated values

Defect	Source				
		^{14}N	^{15}N		
N_i	Experiment [170, 181]	690	17		
	Theory [145]	700	17		
		582	15		
		$^{14}N^{16}O$	$^{15}N^{16}O$	$^{14}N^{17}O$	$^{14}N^{18}O$
N_iO_i	Experiment [100, 174]	1020	8	9	
		813	18	0	
This work					
	-1	1014	3	20	37
		875	20	5	11
		745	0	14	28
	Neutral	992	2	21	39
		850	21	3	8
		694	0	12	23
	Local Minimum	954	25	0	0
	(see Section 7.5)	828	0	20	38
		625	15	0	0
$N_{2i}O_{2i}$	$^{14}N^{14}N$	$^{16}O^{16}O$	$^{16}O^{18}O$	$^{18}O^{16}O$	$^{18}O^{18}O$
	This work	1142	0	0	0
		1100	26	13	42
		963	0	0	0
		868	17	23	37
		767	0	1	2
		741	0	0	0
		629	2	0	3
		591	12	3	17
	$^{16}O^{16}O$	$^{14}N^{14}N$	$^{14}N^{15}N$	$^{15}N^{14}N$	$^{15}N^{15}N$
		1142	1	31	32
		1100	0	1	1
		963	26	1	28
		868	0	1	1
		767	0	17	18
		741	17	0	17

Table 7.2: Local vibrational modes (cm^{-1}) for various $(N_i)_n$–$(O_i)_m$ defects in Silicon. Later columns show the drop in frequency with the change of isotope. For NNOO the second N and O atoms are the ones in the defect centre.

given in Table 7.2. Firstly only two modes are experimentally observed, whereas this NNOO defect has four higher frequency modes. Secondly the experimental oxygen mode is lower than that due to the NN-O defect, but in NNOO the O atom is more compressed and so gives rise to a higher mode. Therefore we can discard the NNOO defect as a model for the two LVMs at 813 and 1020 cm^{-1}.

An alternative structure for the NNOO defect would be one O$_i$ atom at either side of the N$_2$ core, but we have not yet investigated that. This could not be a candidate for the 813 and 1020 cm^{-1} modes either since it would lead to more than one nitrogen mode.

In addition, the experimental anneals were performed at 600°C, and the dimer is not expected to be stable at these temperatures (see Chapter 6). However, if the annealing was repeated at 300-350°C, then it is possible that this defect would indeed be observed. This would be an interesting experiment to perform to confirm the presence of dimers in N doped Si. Such a low temperature anneal would also provide more information about the mobility of the N$_2$ pair.

7.6 Conclusions and Discussion

Our conclusion is that the structure shown in Figure 7.4 represents the NNO defect, the primary N–O complex in silicon. The NNO defect has no deep or shallow donor levels, and so is not responsible for shallow thermal donors (see Section 8).

The three atoms of the defect are only weakly dynamically coupled because of their inequivalence. The O atom is bound to the defect because the N-pair strains the lattice. The dissociation barrier of the NNO complex is then approximately the energy necessary for O to migrate to the barrier top starting from O lying between Si(6) and Si(5) and ending between Si(5) and Si(1). Taking the first figure to be that for O$_i$ diffusion, 2.5 eV [4] and the second one as calculated here, 1.2 eV, we find this dissociation barrier to be 3.7 eV. The equilibrium concentration of NNO defects would equal,

$$\frac{[N_2][O_i]}{N_s} e^{(1.2eV/kT)},$$

where N_s is the density of bond-center sites for O$_i$. The formation energy of NNO from N$_2$ and O defects is estimated to be around 1.2 eV which gives a concentration of NNO equal to that of N$_2$ below 800°C if [N$_2$]= [O$_i$] = 10^{-6}N$_s$ = 10^{17} cm^{-3}.

We remark that as the trapping of O by N pairs has been attributed to the strain field surrounding the pair, then this mechanism would be also be expected to operate for other impurities. H and F would behave like O. However, C would be expected to substitute for Si atoms compressed by the pairs whereas Ge and substitutional transition metals would replace Si atoms subject to tensile stress. This would lead to ⟨011⟩ oriented aggregates.

In addition we conclude that the 813 and 1020 cm^{-1} vibrational modes are produced by the N_iO_i defect. This is present in much smaller concentrations than the standard NNO defect, and acts as a deep donor defect similar to N_i. It will form as a pre-cursor to either NNO or the N_iO_{2i} shallow thermal donor, depending on whether it traps N_i or O_i, and this adds weight to our claim that N_iO_{2i} can form in N/O codoped material (see Chapter 8). The two models examined above differ in that the NNOO model has no gap levels at all, whereas the N_iO_i model has a single deep donor state. Therefore it should be possible to experimentally verify that the defect is N_iO_i rather than NNOO, for example through detection of an associated deep level using DLTS (although this may prove difficult to distinguish from other centres in the material) or else electronic transition IR.

An important question that needs to be addressed when discussing the kinetics and formation of N/O defects is the migration process for N_2 pairs. It is currently unknown whether the N atoms are able to diffuse as a pair, or have to break up into individual N_i before diffusing and later recombining. An experiment which may be able to distinguish between these two is as follows [182]. A sample should be ion implantated using ^{14}N and briefly annealed, so the N forms pairs of ^{14}N. The sample is then similarly ion implanted, but using ^{15}N; another quick anneal should pair these atoms off, leaving a collection of ^{14}N and ^{15}N nitrogen pairs, but very few mixed isotope pairs. This can be verified by FTIR, before performing a longer time anneal. If, after the annealing, the sample shows a spread of pure and mixed isotope pairs, then this suggests the pairs are breaking up and reforming. However if the pure isotopic pairs are maintained then it suggests that the N atoms diffuse as a pair.

Chapter 8

Shallow Thermal Donors in Silicon

Many experimental groups have observed a family of *shallow thermal donors* (STDs) in Si doped with both O and N. Our calculations predict that a defect consisting of a single N_i surrounded by two O_i acts as a STD [183]. Isolated N_i possesses a deep donor state localised on a neighbouring Si. However in $N_i(O_i)_2$ this donor level is electrostatically 'squeezed' by the O_i, which pushes it towards the conduction band to form the observed shallow level. The mechanism is not restricted to N_i based defects and also occurs in the defect $(C\text{-}H)_iO_{4i}$ with a similar defect core. We show that these structures may be responsible for two of the the 'NL10' defects observed in Si, and discuss some of the ramifications of such a deep-to-shallow conversion mechanism.

In addition we show that $(C\text{-}H)_iO_{4i}$ is able to structurally re-arrange when switching from the neutral to the -1 charge state, by rotating one of the oxygen atoms away from the defect core. This suggests that the defect could act as a deep acceptor to become -1 charged, or a shallow donor to become +1 charged. The structure of the neutral charge state defect is almost identical to that of the +1 defect.

8.1 Background

The shallow thermal donors (STDs) are a family of defects which can be formed by annealing N rich Cz-Si around 650°C [184, 185, 186, 187, 165]. As single donors they have ionisation energies from 0.0347 to 0.0374 eV [188]; this is comparable with P, and is much shallower than the thermal double donor (TD) (0.0530 to 0.0693 eV). N is known to suppress TD formation [189].

EPR experiments of Hara *et. al.* [190, 191] found STD defects to have $S = 1/2$ with a g-tensor extremely close to that of the NL10 defect. This led them to suggest that they were the same defect. The g-tensor shows that the STD possesses an apparent C_{2v} symmetry. However EPR and ENDOR studies have failed to resolve any hyperfine splitting in the defect core [190, 192] and so the presence of N in the defect is only inferred from its presence in the material and remains controversial. Although several studies claim to have shown that both O and N must be present in the Si for STDs to form [189, 190, 165], other workers [168] find STDs in nominally N free Cz-Si. In addition, large N concentrations can actually *suppress* STD formation [193]. We believe that the explanation of these conflicting observations lies in a) that other defects, not containing N, can be STDs, and b) there are other electrically inactive nitrogen-oxygen defects which form in competition with the N-related STDs (see Chapter 7.4).

The IR electronic absorption lines observed by Suezawa *et. al.* [165] originated from several different families of STDs. In addition, calculations using effective mass theory by Griffin *et. al.* [188] suggest that the different families of STDs are due to different numbers of oxygen atoms surrounding a common core, with larger oxygen complexes leading to increasingly shallow donor behaviour.

Recently, Yang *et. al.* annealed N-doped Si containing STDs [189]. As their concentration decreased, the concentrations of O_i and N-pairs, as monitored by IR spectroscopy, increased. Conversely when STDs formed, the concentration of both O_i and N-pairs decreased. This unfortunately cannot be taken as direct evidence that these elements are part of the STD, because it is known that electrically inactive NNO defects can be formed around the same temperature as STDs [103] (Section 7.4), and this confuses the issue. It seems to us that while N pairs would complex with O_i creating inactive NNO defects [177], N_i defects could form active centres.

Liesert *et. al.* [168] observed PL from STDs in samples that were nominally N-free. It may be that there was an undetected level of N in these materials but it is worth considering whether STDs can be formed with different core centres. One such possibility to be considered below is an oxygen complex with C_i-H. The latter defect is to be distinguished from the marginally stable substitutional carbon – hydrogen defect [194, 195]. A C_iH defect complexed with a substitutional carbon atom forms the T-centre [196] which is observed when Cz-Si containing carbon is annealed between 450°C and 600°C [197]. There is no deliberate H doping of these materials showing that H is a common and unintentional impurity in Cz-Si [198, 199]. The chemistry of C_iH seems very similar to that of N_i and hence it is plausible that $(CH)_iO_n$ defects could form in C, O and H rich Si.

Recent work has shown that there does indeed exist a family of STDs that contain hydrogen, since there is a shift in the electronic infra-red resonances associated with the $1s \rightarrow 2p_{\pm}$ transitions of the STD when hydrogen is replaced with deuterium [192, 200]. This shift is not reproduced in the thermal donor resonances. This work correlates IR absorption strengths of these STDs to EPR spin concentration of the NL10 spectrum, suggesting they are the same defect. However there are at least two different sources of NL10 containing either H or Al [192], and these results do not explain the nitrogen containing samples. This therefore suggests there are at least three types of STD, one N-based, one containing H, and one containing Al. It is suggested that the H-based STDs could be H-passivated TDs [192, 200]. However the H-related STDs are stable up to 520°C [201], whereas TDs have broken down by then. In addition, H passivated TDs have been identified and are known to lose hydrogen and reactivate at less than 200°C [202, 203].

The NL10 centres were originally observed by Müller *et al* [204]. These are electron spin resonance (ESR) active centres with C_{2v} symmetry after annealing from 450 to 600 °C. The g tensor of Si-NL10 is less anisotropic than that of NL8 (TD^+), suggesting its paramagnetic electron is only weakly bound, consistent with a shallow donor state, and it has spin $S = \frac{1}{2}$ [11]. NL10 defects have a series-like character [205, 206, 207, 208, 209], and ENDOR work shows they include oxygen within one $\{011\}$ lattice plane forming a planar structure[233, 205]. In addition, ENDOR suggests NL10 is structurally similar to NL8 (associated with the thermal donors), with an identical oxygen structure [192]. Therefore identification of NL10 could give important clues as to the structure of the thermal donors.

8.2 Method

All defects discussed here were symmetry constrained to C_{1h} or C_{2v} as appropriate. In every case symmetry was later breached and in most cases the defects relaxed back to the symmetrical structures; exceptions are discussed in the text. 113 atom clusters were used for initial results, but the final results discussed here were obtained using 148-152 atom clusters. The results were qualitatively identical, showing that they are independent of cluster size. In some cases, e.g. the $(CH)_iO_{4i}$, the final structure shows slight deviation from its overall C_{2v} symmetry; where relevant this has been discussed in the text.

The Kohn-Sham energy levels arising from the calculations give only an approximate estimate of donor and acceptor levels. This being so, it is difficult to conclude that an occupied level lying close to the conduction band of unoccupied levels is a shallow donor as opposed to a donor whose ionisation energy is a few tenths of an eV. It certainly appears in this work that certain levels are very close to the conduction band and will be described as shallow to distinguish them from deep mid-gap states possessed by other defects.

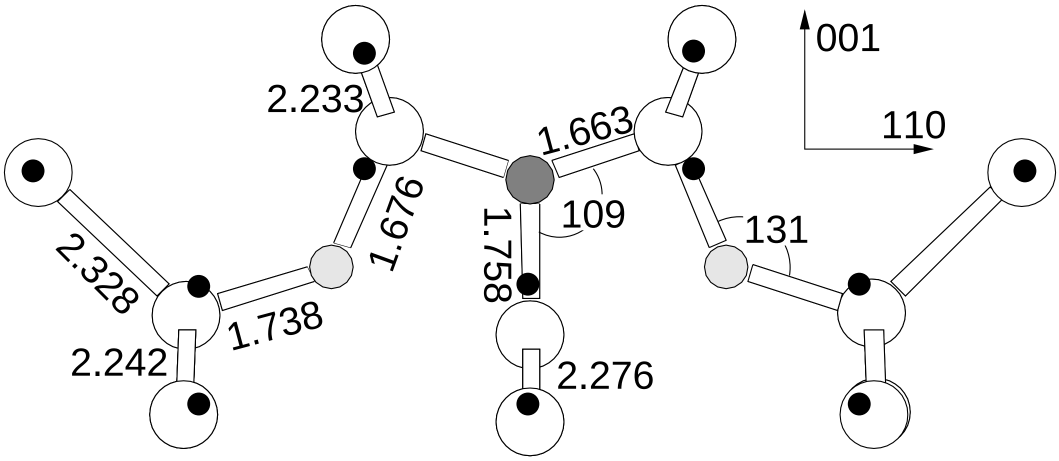


Figure 8.1: The core structure of the +1 N_i-O_{2i} defect. All bond lengths are in Å. N atoms are dark grey, O atoms light grey, black dots mark the ideal lattice sites.

8.3 N_iO_{2i} – The N–O Shallow Thermal Donor

There are two likely structures for the N_iO_{2i} defect which have C_{2v} and C_{1h} symmetries respectively. These correspond to placing one oxygen atom either side of N_i, or both on the same side, respectively. Unlike the case of C_iO_{2i}, we found the first defect to have the lower energy by 1.52 eV, and hence the N_iO_{2i} has the same symmetry as that found for nitrogen related STDs [190]. Crucially the defect has a very shallow donor level, which is more properly described by effective mass theory, as shown in Figure 8.2. The calculated local vibrational modes of N_i-O_{2i} are given in Table 8.1.

The 150 atom cluster, $Si_{79}H_{68}NO_2$, was used for investigating this defect, centred on the N-Si bond centre lying on the C_2 axis; this ensured that all atoms up to second shell neighbours of the defect atoms were bulk Si atoms. Its relaxed C_{2v} defect structure is shown in Figure 8.1. The oxygen atoms have an Si–O–Si bond angle of 140°, considerably lower than the standard interstitial angle of ≈ 170°, and closer to that of substitutional oxygen (≈ 120°). In addition the Si–O bond lengths are 1.69 Å and 1.64 Å for the outer and bonds respectively, again lying somewhere between normal interstitial and substitutional bond lengths (1.61 Å and 1.75 Å respectively [99]). The distance between the O atoms and the central Si atom is 2.54 Å: considerably longer than would be expected if there was a Si–O bond forming.

The N atom has two short 1.65 Å top Si–N bonds and a longer 1.72 Å [100] Si–N bond, so the presence of the O_i causes a 9.5% compression of the top bonds, slightly compensated by a 4.1% extension of the lower N–Si bond.

We also optimised the defect structure in the positive charge state, since at the

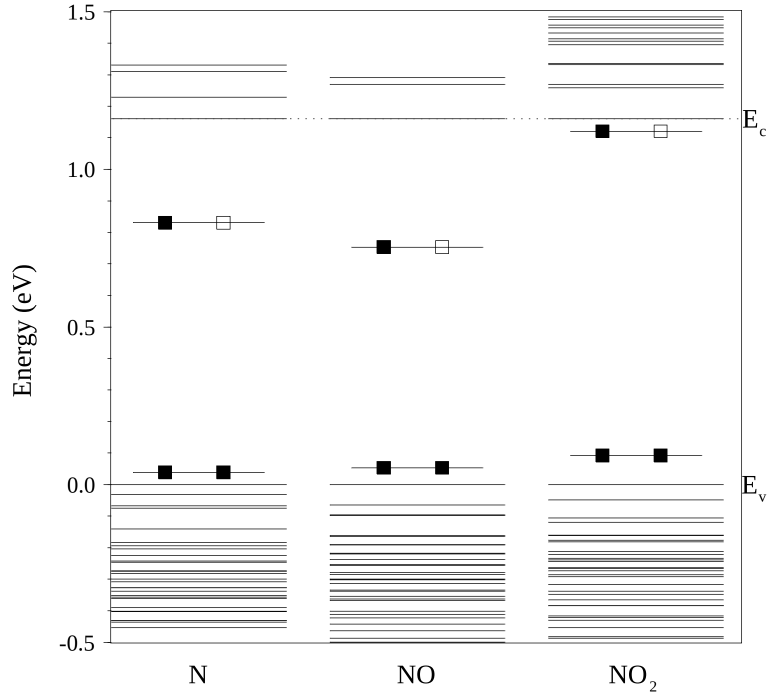


Figure 8.2: Top Kohn-Sham eigenvalues for neutral N_i, N_iO_i, and N_iO_{2i}. Filled (empty) boxes denote electrons (holes). These have been scaled to the experimental 1.16 eV band gap.

annealing temperatures described the shallow level will depopulate. This revealed a very similar structure, where the O atoms move only 0.15 Å towards the central Si. Thus there is a slightly stronger third Si–O bond, although not three equivalent bonds as are shown in related 'y-lid' models [121, 92]. This small change suggests that cluster size is not playing a significant role.

A plot of the wavefunction of the partially filled shallow level is given in Figure 8.3. The level is diffusively localised on the $p-$ type dangling bond of the central Si radical. There is negligible localisation on the interstitial nitrogen, consistent with the lack of any hyperfine interaction with N observed by EPR/ENDOR experiments. The nodal surface between O and the Si radical demonstrates an anti-bonding character to the state. There must then be a bonding orbital lying below this state and this shows that the O atoms are to some extent overcoordinated. A larger cluster would probably lead to a greater delocalisation of the donor state, which would then be properly described by effective mass theory.

The local vibrational modes for the defect are given in Table 8.1. As would be expected, the longer Si–O bonds lead to oxygen related LVMs that are lower than those of the standard interstitial, and closer to those of substitutional O_s. Interestingly the oxygen modes are strongly coupled due to their shared interaction with the core Si, yet are largely decoupled from the N_i. Although the STDs are typically present in concentrations too low to be observed by infra-red spectroscopy,

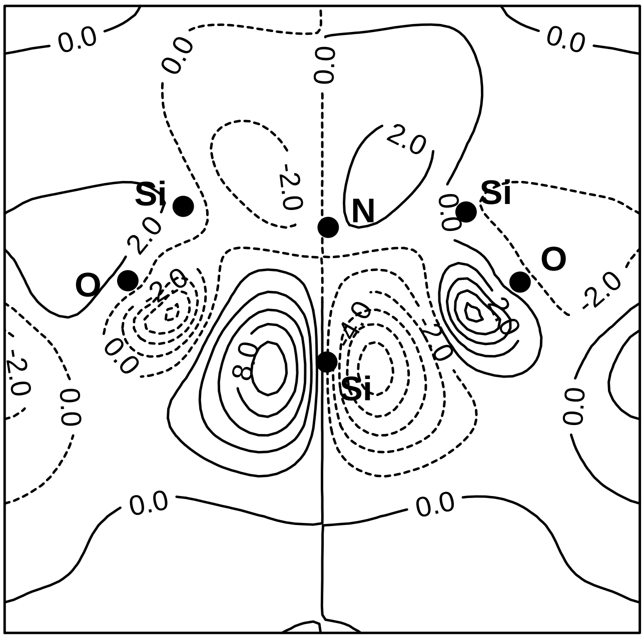


Figure 8.3: The pseudo-wavefunction (× 100 a. u.) of the highest occupied orbital of neutral N_i-O_{2i}. Note that it has little amplitude on N and is localised on the Si radical. There are nodal surfaces lying between this atom and the O atoms demonstrating anti-bonding behaviour.

$^{14}N^{16}O^{16}O$	$^{14}N^{18}O^{16}O$	$^{14}N^{18}O^{18}O$	$^{15}N^{16}O^{16}O$
1051.4	0.4	0.8	28.6
893.6	10.3	36.4	1.5
857.9	28.3	40.9	0.8
720.3	1.1	2.2	12.3
673.4	2.0	3.8	12.1

Table 8.1: Local vibrational modes of N_iO_{2i} (cm^{-1}) - later columns give downwards shift in modes with change of isotope.

there should be enough of them for detection by photoluminescence methods.

Further calculations show that the defect with two more interstitial oxygen atoms in the ⟨110⟩ plane is stable, N_i-O_{4i}. These atoms sit in standard interstitial sites but are able to relieve some of their compressive strain through a compression of the weaker central Si-O-Si bonds. These O atoms move closer to the Si radical. This should result in pushing upwards the shallow donor level but the limitations of the cluster method prevent a quantitative estimate of the change.

This mechanism whereby larger oxygen aggregates push the nearest oxygen atoms towards the Si radical, increasing the repulsive central cell potential, demonstrates how the family of shallow donor states develops. It also explains the observation of Suezawa *et. al.* that the absorption lines of different shallow donor defects at first sight appear to be single absorption line splitting due to small strains at the defect core [165]. A more detailed discussion is given in Section 8.10.

Since performing these calculations we have learned that independant research by Chadi also shows N_i-O_{2i} to possess a shallow donor level in this configuration [211].

8.4 Other potential STD models

We also considered a variety of other defects as candidates for the nitrogen related shallow thermal donor.

8.4.1 $(N_{2i})_n$–O_i – Suezawa's model

A kinetic study by Suezawa *et. al.* [165] linking the rate of loss of N-pairs to the increase in STDs (using IR electronic absorption intensities), suggested that STDs consisted of an unknown O-containing core, to which N-pairs attached themselves. However this modelling neglected dissociation of N_n-O_m complexes, which will be crucial in the kinetics of these defects. In addition it is known that the NNO defect consisting of an N-pair neighbouring a single O_i is electrically inactive [103, 177].

It was proposed that the STD could consist of a core containing O_i, surrounded by N_{2i} pairs in the ⟨110⟩ direction [165]. However in order to maintain C_{2v} symmetry with only one O_i, the O_i must lie along the central C_2 axis. This structure will therefore be of a [100] split interstitial type, similar to N_i. Since it contains an even number of N atoms it is EPR inactive in the neutral state.

We investigated the cluster $Si_{57}H_{56}N_4O$, containing this defect and constrained it to possess C_{2v} symmetry. No shallow donor level was found and we therefore conclude that this cannot be a valid model for the STD. However, N_{2i} pairs could still co-exist with the N_iO_{2i} shallow thermal donor core proposed above. Although

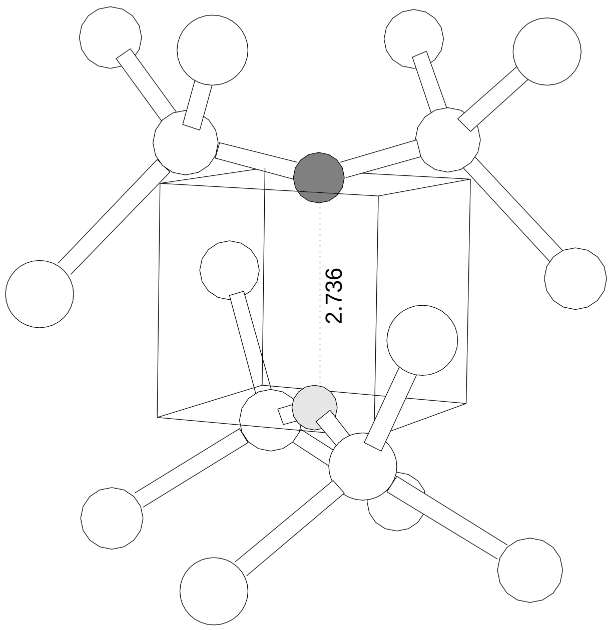


Figure 8.4: Structure of the proposed VO_sN_s defect. This does not have the required bonding or gap states to behave as a shallow donor defect. The cube marks the ⟨100⟩ directions, corners on ideal lattice sites. The N–O distance in Å.

O_i atoms will continue to agglomerate at the defect core due to the polar bond formation, the outer O_i atoms will not benefit from the tensile strain field of the defect core. They will themselves still exert tensile stress on the surrounding lattice, and therefore there will be a driving force for N_{2i} pairs to also collect along the same ⟨110⟩ plane. This formation mechanism is identical to that of the NNO defect [103, 177] but requires N_{2i} to be mobile.

8.4.2 Substitutional Nitrogen

It was previously proposed by members of our group that a *substitutional* N_sO_i pair could form the core of a STD [212]. The problem here is that the substitutional defect seems to be a rare defect. However, in analogy with the VO_2 centre, we examined a substitutional split N_s-O_s pair. This, however, was found to possess a deep level, with no bonding between the N_s and O_s atoms. The addition of extra O_i in neighbouring bond centres did not provide sufficient strain to push the gap state up to give a shallow donor level. A diagram of the structure is shown in Figure 8.4.

8.4.3 Bistable NNO structures

Another proposal for the STD structure was based on the possibility that the electrically inactive NNO defect was actually *bistable*, and had a singly charged alternative structure with shallow donor properties [213] (see Figure 8.5). Initial investigations

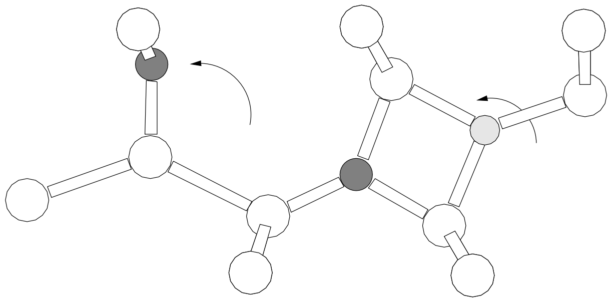


Figure 8.5: Structure of the proposed bistable NON(+) defect core. Arrows show approximate movement of N_i and O_i atoms from the electrically inactive NNO defect structure.

Charge State	PM3-CCM	AIMPRO	AIMPRO using PM3-CCM relaxed structure
0	+1.27	*	+6.09
+1	-1.18	+3.97	+1.96

Table 8.2: Calculated energy difference (eV) between the NNO and NON structures, using the PM3 cyclic cluster method and AIMPRO in the neutral(0) and positive(+1) charge state. Positive numbers show NNO to be more stable. The $*$ indicates NON(0) spontaneously restructures into NNO(0) with no barrier.

using PM3-CCM semi-empirical methods proved promising [180]. We therefore investigated this using the same clusters as used for the NNO work in both the neutral and +1 charge states.

Both the standard NNO structure (Section 7.4) and the new proposed alternative were relaxed in the neutral and +1 charge states. The original NNO structure was the most stable in both charge states; the neutral N-ON spontaneously restructured to form NN-O with no barrier. In the positive charge state NO-N^+, although metastable, was 3.97 eV higher in energy than NN-O^+. This suggests that contrary to the previous results from PM3-CCM calculations[214], we find that NON(+) is not a stable form of the NNO system, and is not predicted to occur in real systems in either the neutral or +1 charge state. In addition it does not display the correct C_{2v} defect symmetry, and in light of these results we rule it out as a possibility for the STD.

8.5 Formation processes for N–O defects

A complication with the unravelling of the processes by which oxygen atoms aggregate in the presence of these impurities is the existence of more than one type of oxygen complex. In N-doped material the dominant nitrogen defect is the pair which itself forms electrically inactive NNO defects [103, 177]. These have to be distinguished from the shallow donor defects originating from N_i, and the question then arises as to which defect is actually formed.

High temperature annealing is normally used to break up any defect complexes present in the material, and this will include N_{2i} pairs. On cooling, the diffusing N_i defects are trapped either by O_i or N_i. A high concentration of N_i implies that N_{2i} pairs will be preferably formed, whereas a high oxygen concentration may result in N_iO_i defects. Both defects subsequently complex with oxygen to create either electrically inactive NNO or STD defects in the form of NO_{2i}. Sun *et al* [169] showed that for N to have an effect on O precipitation it is essential to pre-anneal at 1100°C; this step will be required to break up the N_{2i} pairs and any larger oxygen precipitates, and homogenously distribute the N_i and O_i.

It is likely that the N-pair is more stable than N_iO_i and hence slow cooling probably results in higher concentrations of NNO. Hara *et. al.* [190] examined the effect of cooling rate on N-rich Cz-Si, and found that in quenched Si they obtained high concentrations of STDs, whereas in slow cooled Cz-Si the STD concentration was much lower. This is consistent with the above explanation. Slow cooling will also lead to out-gassing of nitrogen, leaving less N_{2i} pairs to mop up O_i forming NNO. Similarly, evidence that high N concentrations do not necessarily result in high concentrations of STDs has been shown by Griffin *et. al.* [193] using PTIS on N doped samples annealed at 450°C. They found that N suppressed TD growth — behaviour similar to that of carbon doped Si where C–O complexes are formed. In addition, for low doping levels of N (1.4×10^{15} cm^{-3}), they observed a large number of STDs, whereas for a higher N doping level (9×10^{15} cm^{-3}), the STD concentration was strongly reduced. This result is consistent with N_{2i} pairs forming preferentially at higher concentrations, since N_i atoms are able to combine before encountering any O_i.

The formation kinetics of the NO_{2i} defects will be dependent on the quantities of each reaction component, and the rate constants for many different reactions:

$$\begin{aligned}
\mathrm{N_i + N_i} &\rightleftharpoons \mathrm{N_{2i}} \\
\mathrm{N_i + O_i} &\rightleftharpoons \mathrm{N_iO_i} \\
\mathrm{O_i + O_i} &\rightleftharpoons \mathrm{O_{2i}} \\
\mathrm{N_{2i} + O_i} &\rightleftharpoons \mathrm{N_{2i}O} \\
\mathrm{N_iO_i + N_i} &\rightleftharpoons \mathrm{N_{2i}O} \\
\mathrm{N_iO_i + O_i} &\rightleftharpoons \mathrm{N_iO_{2i}} \\
\mathrm{N_i + O_{2i}} &\rightleftharpoons \mathrm{N_iO_{2i}} \\
\mathrm{N_{2i} + O_{2i}} &\rightleftharpoons \mathrm{N_{2i}O_{2i}}
\end{aligned}$$

Here N_{2i} is the nitrogen pair defect, $N_{2i}O$ is the electrically inactive NN–O defect, and N_iO_{2i} is the defect proposed here as a STD. In theory there could be other reaction components such as N_{3i}, as well as reaction paths involving other point defects such as C.

8.6 (C-H) based shallow thermal donors

The observation of STDs in material that is nominally N free [168] suggests that other families of STDs may exist. Annealing Cz-Si between 450 and 600°C produces a number of PL bands [197] some of which, e.g. the T-line, are due to defects containing interstitial C-H complexes [196]. The hydrogen is present as an unintentional impurity and the concentrations of these centres are comparable with those of thermal donors. These interstitial defects are to be contrasted with the marginally stable substitutional carbon-hydrogen centres [194, 195].

$(C\text{-}H)_i$ is believed to be a fast diffusing species in silicon [215] which is highly stable in either a bond centred or three-fold coordinated site depending on whether the carbon is sp^2 or sp^3 hybridised. In the three-fold coordinated site it sits as a [100] split-interstitial similar to C_i. The C-H bond is parallel to $[0\bar{1}1]$ and lies perpendicular to the plane containing the three C-Si bonds. The C-H bond length is exceptionally small, 1.1 Å, and as this unit is isovalent with N, it suggests that an STD might arise if N in Fig. 8.1 was replaced by C-H. We therefore tried replacing N_i with $(C\text{-}H)_i$ in the STD model.

There are three primary structures to be considered. Firstly the $(CH)_iO_{2i}$ defects. In direct analogy with the N_iO_{2i} defects these have two possible structures, with either one oxygen atom on each side of the $(CH)_i$, or with both oxygen atoms forming a dimer pair on one side of the $(CH)_i$. We therefore investigated both of these. Then there is also the $(CH)_iO_{4i}$ defects, i.e. $(CH)_i$ flanked by two oxygen pairs, one on each side.

8.6.1 $(C\text{-}H)_iO_{2i}$

This was modelled using a 151 atom cluster in the +1 charge state, $CSi_{79}H_{69}O_2$ (Figure 8.6). The presence of the (C-H) bond gives the defect C_{1h} symmetry rather than C_{2v}. However, more recent ENDOR symmetry analysis [10] has shown that NL10(H), the EPR/ENDOR signal associated with H-related shallow thermal donors (see Section 8.7) was misassigned using EPR to C_{2v} symmetry. Using ENDOR they have shown it to be triclinic, with H lying along $[1\bar{1}0]$ perpendicular to the plane containing O_i, in agreement with this model.

The O forms assymetric bonds of 1.66 and 1.62 Å, the shorter bond with the more positive Si atom shared with the $(CH)_i$. The Si-O-Si bond angle is 158°. C

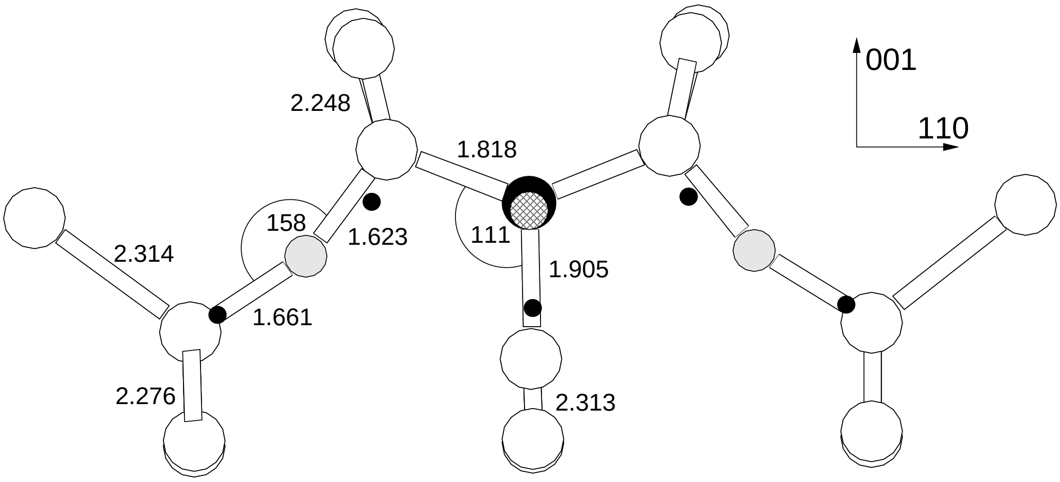


Figure 8.6: The (C-H)$_i$-O$_{2i}$ 'C_{2v}' defect core in the +1 charge state. Numbered atoms are referred to in the text. Symmetry with H included is C_{1h} (see text). Black dots mark the ideal lattice sites. This is less stable than the assymetric structure in Figure 8.7. The cross-hatched H atom lies perpendicular to the defect plane.

forms three strong Si bonds, the top two of 1.818 Å and the bond with the core Si of 1.905 Å. The C-H unit has a short C-H bond of 1.115 Å, and C moves out of plane by 0.159 Å. However, unlike N$_i$O$_{2i}$, this defect was not itself a shallow donor, the level instead lying closer to mid-gap. Changing the charge state of the defect to neutral or -1 through the addition of 1 or 2 electrons respectively, had little effect on the structure, and the level remained deep in the gap. In this case the empty dangling bond on the Si was either singly (neutral) or doubly (-1 charged) occupied. The C-H bond dilated (1.138 Å when neutral, 1.145 Å when -1), and the distance between the C and the core Si increased (1.965 Å when neutral, 2.001 Å when -1). The O atoms were also pushed back into their bond centres (Si-O-Si bond angles of 173°), and all of these shifts are consistent with increased charge in the defect core. These results show that, by itself, the (CH)$_i$O$_{2i}$ defect in this configuration is not a single shallow donor. We deliberately breached the defect symmetry after relaxation but in all cases it returned to the symmetric structures described above with the same energies.

We next examined the stability of this defect and found it to be less stable than the alternative structure with both oxygen atoms on the same side of the C (similar to its unhydrogenated cousin, C$_i$-O$_{2i}$, responsible for the 'P-line' PL absorption). This lop-sided defect structure is 1.36 eV more stable than the symmetric one in the neutral charge state (1.33 and 1.76 eV in the +1 and -1 charge states respectively). This also has a deep donor level rather than a shallow one. The structure is shown in Figure 8.7. We again examined the three possible charge states of this defect, +1, 0, or -1, corresponding to different populations of the Si $p-$type dangling bond in the

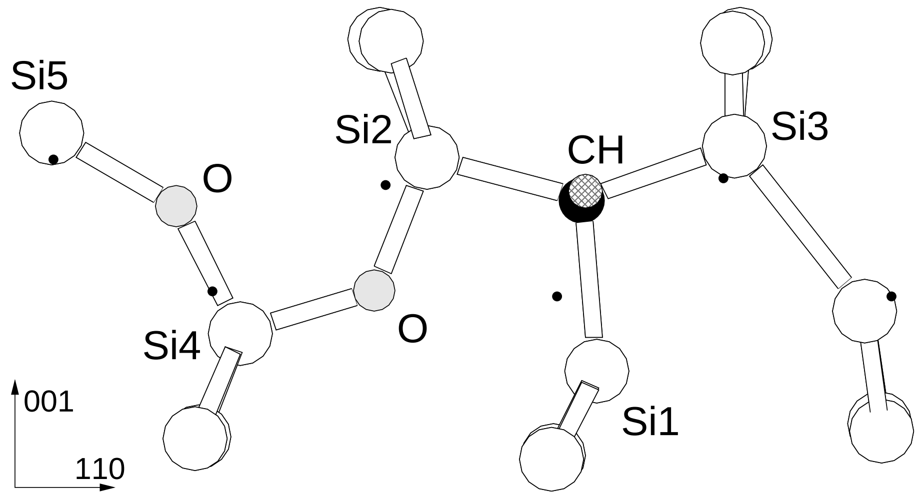


Figure 8.7: The (C-H)$_i$-O$_{2i}$ defect with O$_{2i}$ neighbouring (CH)$_i$, in the neutral charge state. Numbered atoms are referred to in the text. This structure is 1.36 eV more stable than its equivalent with one O atom on either side of the (CH)$_i$ unit. Black dots mark the ideal lattice sites.

core of the defect. In the +1 charge state the core Si stayed close to the $\langle 001 \rangle$ axis, but in the neutral and -1 charge state it bows away from the neighbouring oxygen atom, allowing it to localise the filled $p-$type orbital as far from the oxygen lone-pairs as possible. In this respect its behaviour mimics that of N$_i$O$_i$, and presumably the exact analogy, (CH)$_i$O$_i$ would also behave in a similar way.

The lack of shallow donor properties for (CH)$_i$O$_{2i}$ in the C$_{2v}$ configuration is initially surprising, given that it is iso-electronic with the N$_i$O$_{2i}$ shallow thermal donor. However the (CH) unit is less electronegative than N$_i$ and so the core Si atom is less polarised. This in turn provides less driving force to draw the oxygen atoms inwards and indirectly compress the gap level. In addition the back-bonds of (CH)$_i$ are not as dilated as those of N$_i$, and so there is less strain energy to be gained from placing the oxygen atoms next to the (CH)$_i$; this is why the oxygen atoms prefer to sit on one side where they gain energy through quadrapole interactions through their shared lattice Si.

However the C_{1h} defect could still act as a *STD core*, and would become stable with the addition of further oxygen atoms, similar to N$_i$-O$_{4i}$. This would also help to compress the defect core. We therefore relaxed (C-H)$_i$-O$_{4i}$, and found that it possessed a single shallow donor level. The Kohn-Sham eigenvalues are shown in Figure 8.10, showing the donor level. (CH)$_i$O$_{4i}$ is examined further, below in Section 8.8.

In conclusion, it appears that C$_i$HO$_{4i}$ is a stable shallow donor defect, possesses C_{1h} symmetry and contains a C-H bond lying along $[0\bar{1}1]$. There is very little spin-polarised density residing on C or H. The defect core, C$_i$HO$_{2i}$ is more stable in a

C_1 configuration similar to C_iO_{2i}. In n-type material the $(CH)_iO_{2i}$ defect would be either neutral or -1 charged.

8.7 Identification of NL10

The recent work of Newman and Ammerlaan [200, 192] suggests that there are actually at least three different types of STD defect. The first, not yet been identified by them, is the N-based defect of Hara, Suezawa, etc. discussed above. This defect is stable to ~900°C [165]. The second contains hydrogen and corresponds to at least one of the NL10 EPR signals. This defect is stable to 520°C, and has an electronic IR signal distinct from that of STDs found in N-rich material [201]. Finally there is a defect that contains Al on the core C_2 axis but no hydrogen [192], which again has a distinct electronic IR signal. None of these can be attributed to thermal donors, since TDs break down in anneals at 500°C, and can only be H-passivated below 200°C[203].

We suggest that the first two STD families correspond to those with an N_iO_{2i} core, and those with a $(C\text{-}H)_iO_{4i}$ core respectively. This would be consistent with the experimental evidence, as well as the thermal stabilities. N_i-O_{2i} is a tightly bound defect core and should be highly thermally stable (NNO is known to be stable to ~700°C). However $(C\text{-}H)_iO_{4i}$ does not have such tightly bound outer O_i atoms, and once these have diffused away to leave $(C\text{-}H)_iO_{2i}$, this is no longer a shallow donor. Therefore we would expect this to be stable to lower temperatures. In addition, other (C-H) based defects have similar thermal stabilities; the defect responsible for the 'T-line' observed in PL experiments, $(C\text{-}H)_iC_s$, is stable to 600 °C [196], consistent with the idea that the C-H bond is a strong one.

Recent work by Newman has shown the H-based and N-based STDs are indeed two different types of defect[201], and after our suggestions they are now re-examining their ENDOR results for carbon related signals. In their work correlating NL10 to STDs, they assigned the H-containing NL10 signals to H-passivated TDs [200]. However in light of the work of Weber *et al* (suggesting H-passivated TDs are only stable at low temperatures) [203], and the discussion above, we believe this to be an incorrect assignment. They have also shown that high carbon concentrations suppress STD formation [201]. This is similar to the result for nitrogen, and is probably due to several mechanisms: C_s forming complexes with dimers, C_i and C_iH.

Strong confirmation of the assignment of $(CH)_iO_{4i}$ to NL10(H) has come from recent work by Markevich on the D1, D2 and D3 EPR centres, where D2 and D3 correspond to NL10(H) centres that form in hydrogenated irradiated material at the same time that the absorption signal from C_iO_i disappears. This is discussed below in Section 8.8.

Spaeth examined P-doped samples that were subjected to gamma irradiation [216]. This created OV^- centres, which lowered the fermi level and switched the P from a neutral to a P^+ state. However the concentration of neutral STD centres remained unaffected (and actually appeared to increase a little). They explained this by suggesting the STD levels must lie relatively deep in the gap, over 65meV below E_c. However an alternative possibility is that the gamma irradiation was creating Si_i atoms, which in turn liberated more C_i (through $C_s+Si_i \rightarrow C_s$) and thus created further STDs, similar to the irradiation production method of Markevich discussed above. Their assignment was complicated since the STD absorption only appears as a shoulder on the larger OV^- absorption.

A previous study by Jones *et al* [111] proposed an Al-based STD consisting of substitutional Al surrounded by two over-coordinated oxygen atoms, which remains a possible candidate for the third type of NL10 defect. However an alternative can be based around thermal donor models with Al substituting for Si in the core. This is examined further in Section 9.5.

8.8 $(CH)_i$-O_{4i} and the 'pop-out' mechanism

8.8.1 Experimental Background

As well as the experiments described above, there is some work specifically relevant to the $(CH)_iO_{4i}$ defect.

Recent experimental work by Markevich and Suezawa has identified an EPR active centre that forms in irradiated material containing both H and O [217]. This EPR 'D1 centre' is stable in the singly negative charge state, but has an alternative structure which is only marginally stable in the neutral charge state before spontaneously forming a metastable +1 charge state structure. The -1 charge state is a deep level defect, however both the neutral and +1 charge state defects are shallow (identified using DLTS). It is present in concentrations of up to 10^{15} cm^{-3}, and is a negative-U centre with U $\sim$ -0.065 eV [218]. It has a g-tensor remarkably similar to that of NL10(H), and an activation energy for generation of 1.68 eV [219], suggesting it grows by dimer addition (since the dimer migration barrier is approximately 1.7 eV, see Chapter 6). They are produced in hydrogenated Cz-Si crystals after electron irradiation and annealing at 300-400°C [220, 217]. Above around 400°C they transform into further D2 and D3 centres, all of which anneal out at 550-600°C. Their formation corresponds with the loss of C_iO_i absorption signal. D1 displays well-resolved hyperfine interactions of the unpaired electron with a ^{29}Si atom which shows isotropic s-type character; this is not consistent with either our calculations or published EPR results on the NL10 centres. The electronic IR signals for the D centres are in some cases very close to those of NL10(H), in other cases identical (for 1s $\rightarrow$ 2p$_0$ transitions, the D centres absorb at 249.7, 204.4 and 198.8 cm^{-1}, compared to 208.7, 204.2 and 198.2 cm^{-1} for NL10(H)). In addition the D series donor levels lie at 42.6, 37.0 and 36.3 meV [217], whereas the nearest NL10(H) members lie at 37.8, 37.0 and 36.3 meV [200].

Thus it appears that the D1 centre is a precursor to NL10(H), whereas D2 and D3 correspond to two of the NL10(H) centres. D1 dies away at the same temperature as NL10(H) increases, which adds weight to this suggestion. The isotropic EPR signal of D1 suggests it is quite different from any of the centres modelled here.

8.8.2 Results

We modelled the $(CH)_iO_{4i}$ defect in the +1, neutral and -1 charge states using the cluster described in Section 8.2 above. The neutral structure is given in Figure 8.8, along with bond lengths and angles. The +1 structure is almost identical (similar to the result for N_iO_{2i}). The +1 and -1 structures are shown in Figure 8.9, and the eigenvalues in Figure 8.10. Both the neutral and +1 defects act as a shallow donor, but the gap level becomes very deep for the -1 charge state. The vibrational modes for the +1 structure are shown in Table 8.3. These show high frequency C-stretch modes at 1256.3 and 1229.3 cm^{-1}, much higher than C-stretch modes in either C_i, C_iH or C_sC_iH; this is due to compression of the C-Si bonds by the neighbouring O.

In the +1 and neutral charge state, the oxygen atoms force inwards towards the core Si, behaving in exactly the same way as N_iO_{2i}. However in the -1 charge state, one pair of O_i atoms switch so that their Si-O-Si bond angle points away from the core instead of towards it, the oxygen atom is then further from the core Si atom. This allows the core to distort, similar to the N_iO_i defect, with the majority of the single donor state localising on a lobe of the $p-$orbital of the Si furthest from the remaining compressive O_i atom. This removes the primary method of electrostatic compression, so the level can once more drop to lie deep in the gap. If this is in $n-$type material (e.g. P doped material), then it will spontaneously trap a second electron to become -1 charged.

The +1 charge state structure shows some assymetry, with the core Si atom sitting slightly closer to one of the O atoms than the other. This would appear to breach the experimentally observed symmetry of C_{2v}. However, the energy saved relaxing from the symmetric to the off-centre structure is only 0.0152 eV. This represents the energy barrier to the core Si atom switching from one side to the

$^{12}CH^{16}O$	^{17}O	^{18}O	^{13}C	D
2866.0	0.0	0.0	5.5	783.4
1256.3	2.2	3.9	22.0	17.1
1229.3	0.8	1.6	23.6	135.6
1132.8	0.5	1.0	6.3	142.6
1022.8	5.9	9.5	15.5	51.2
975.2	13.2	23.4	5.0	27.6
971.4	22.1	42.4	2.8	75.0
947.9	15.6	31.9	10.7	67.2
880.1	20.9	38.3	1.1	30.6
855.5	13.9	29.1	7.3	34.8
848.8	20.1	38.4	1.0	37.7
740.6	12.6	22.9	0.1	1.6
737.9	12.9	24.4	0.8	8.0
682.0	9.9	18.8	5.9	12.0
634.0	3.7	7.6	1.1	3.1
616.4	4.7	8.7	0.1	0.0
602.8	6.9	13.4	0.3	1.5
554.0	11.9	13.7	0.1	1.0

Table 8.3: Vibrational modes of the $(CH)_iO_{4i}$ defect (cm^{-1}). Later columns show downward shift in mode with isotope change.

other, and since it is so low the core will in practise be switching rapidly between the two sites. Thus the defect will have a time averaged dynamic symmetry of C_{1h}, or C_{2v} if the H atom is not observed.

These results provide an explanation of why the D1→D3 centres are metastable +1 defects that can switch structure to become deep level -1 charged defects. Isothermal capacitance transient (ICTS) measurements under DLTS gave a barrier to electron capture of 0.05 eV. The barrier for major structural reconstruction should be much higher than this. However the energy barrier for the 'pop-out' mechanism might be expected to be small. The energy for precession of an isolated O_i atom around its Si-Si bond is $<$0.09 eV [221], and this would be the mechanism by which the defect could restructure itself between the deep and shallow defect state.

Sample treatment would also suggest $(CH)_iO_{4i}$. The D1 defect is only observed in hydrogenated irradiated Cz-Si samples. Irradiation produces C_i which goes on to form C_iO_i; this disappears as the D1 defect is formed, consistent with a C_i based defect core. The defect forms through dimer addition, which is also consistent with $(CH)_iO_{4i}$.

The results of Markevich suggest that this is a negative-U centre, although the

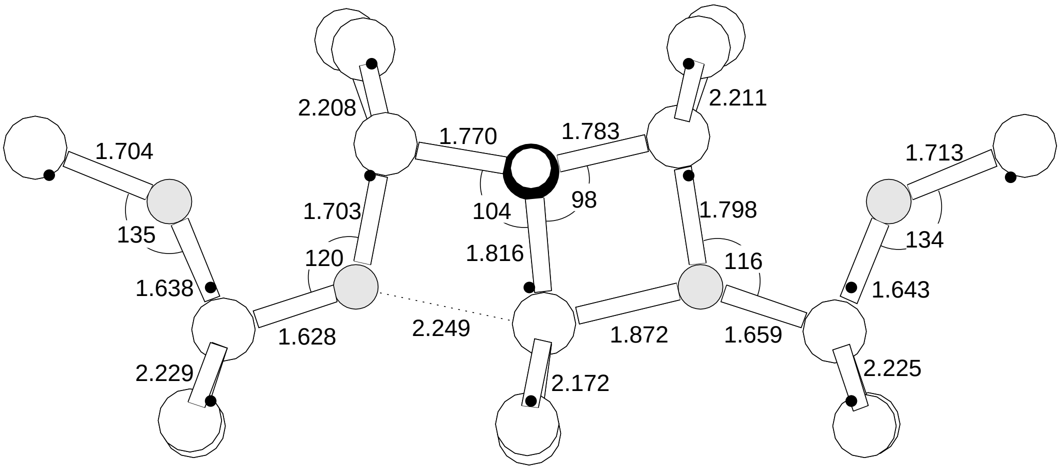


Figure 8.8: The neutral $(CH)_i$-O_{4i} defect (bond lengths in Å). Black dots mark ideal lattice sites, black is carbon, light grey oxygen, with the C-H bond perpendicular to the plane of the diagram.

negative-U energy is extremely small and outside the accuracy of our code.

8.9 Summary of (CH)-O_{ni} defects and further work

In summary, it appears that the $(CH)_i$-O_{2i} defect forms with the oxygens in a pair on one side of the $(CH)_i$. This defect is not a shallow donor defect but has a partially filled level deep in the gap. It is unclear whether it would be naturally neutral or -1 charged since this depends on the respective locations of the defect level and the fermi level, and our Kohn-Sham eigenvalues are not quantitatively accurate enough for such a prediction.

When this defect attracts further oxygen atoms to form $(CH)_i$-O_{4i} this acts as a single shallow donor. This may happen through the trapping of a fast diffusing oxygen dimer. The donor level is effective mass like, AIMPRO predicts it to sit very close to the conduction band. However, on trapping of another electron, filling the core Si $p-$type orbital and making the defect -1 charged, it is able to undergo a structural change by rotating two of the oxygen atoms on one side of the defect core so that they bow out of their bond centred sites *away* from the defect core (in effect, one side of the wonderbra defect is 'popping out'). This allows the $p-$ lobe of the Si orbital to redistribute the majority of its charge away from the remaining O atom in the core, and the defect level drops to mid-gap. This is indicative of negative-U behaviour, but within the accuracy of our code we cannot determine whether the defect is negative-U based on our cluster energies.

Photo-luminescence work has observed a number of lines such as the I- and M-lines in Si, which are believed to be C-H-O defects of some kind [222]. It is possible that these are associated with $(CH)_iO_n$ defects such as these, possibly

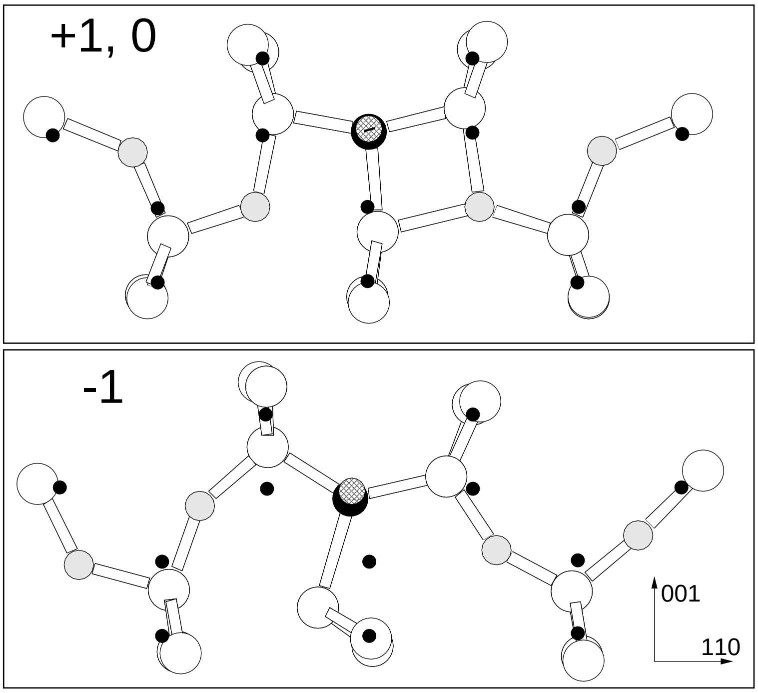


Figure 8.9: Structure of $(CH)_i$-O_{4i} in the +1, 0 and -1 charge states. Black dots mark ideal lattice sites. The neutral and +1 charge state structures are almost identical and so the neutral diagram has not been included seperately. In the -1 charge state the oxygen pair on one side bow out of the defect core, removing the elecrostatic compression on the core Si atom; this shifts off site to localise the deep gap state on its p-type orbital.

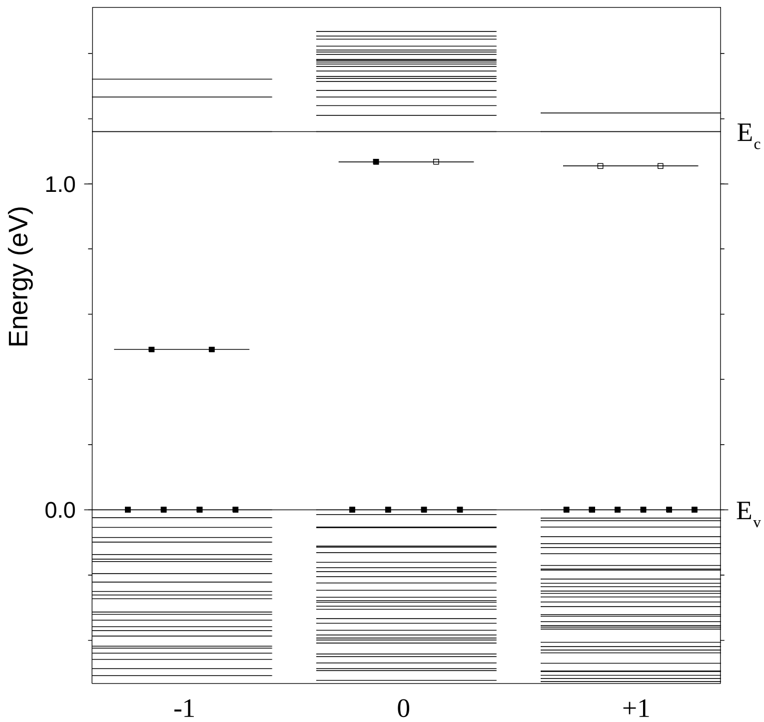


Figure 8.10: Kohn-Sham eigenvalues for $(CH)_iO_{4i}$ in the +1, neutral, and -1 charge states, showing its shift from a shallow to a deep level defect. The eigenvalues have been aligned and scaled to the experimental Si band gap of 1.16 eV.

either $(CH)_iO_i$ or $(CH)_iO_{2i}$ (or $(CH)_iCO_n$). Further work would be required to correctly assign these, however. The I-line forms after 30 minutes annealing at 600 °C, and at this temperature dimers are unstable. Its symmetry was found with uniaxial stress to be monoclinic-I [222]. This would exclude $(CH)_iO_{2i}$ but could still be $(CH)_iO_i$ since $(CH)_i$ can also sit in a bond centre with an sp^2 coordinated C atom, in which case it could form a tight pair with the O_i similar to the dimer, and all the defect atoms would lie in the same $\langle 110 \rangle$ plane. If these PL-lines could be assigned to a $(CH)_iO_{ni}$ defect this would be important, as it would show proof of pre-cursors to the $(CH)_iO_{4i}$ STD complex.

8.10 The 'Wonderbra Mechanism'— Deep to Shallow Donor Conversion

The N_iO_{2i} defect exhibits the 'Wonderbra Mechanism'[223], a method of converting a deep level to a shallow one through electrostatic compression. In isolated N_i, there are two lone pair orbitals, a filled $p-$ orbital located on the N, and a single occupied $p-$ orbital on its Si neighbour. This Si orbital normally leads to a deep state. In the presence of a single O_i, forming N_iO_i, the level remains deep, since the single electron is able to distribute most of its wavefunction to the lobe furthest from the O_i. However in N_i–O_{2i} this orbital is squeezed between the two highly negative oxygen atoms. This pushes the energy of the level up to become a shallow donor level.

There are many attractive forces holding the defect together, including the dilated bonds which allow tensile strain relief, and large local polarisation of the lattice due to the electronegative N_i and O_i atoms. The N_i draws charge away from the central Si core, leaving it overall positively charged, despite the electron in the $p-$ orbital. This attracts the oxygen atoms and forms a weak dative bond, inadvertently causing local electrostatic repulsion between the single electron on the Si $p-$ orbital and the oxygen atoms. The drawing of charge to N_i from its Si neighbours also acts as an extra initial attraction for O_i attack, in addition to the dilated bonds.

Figure 8.3 shows the pseudo-wavefunction of the partially filled shallow level of the neutral N_i-O_{2i} defect. The occupation of this orbital is anti-bonding in nature, suggesting that there is repulsion from the oxygen lone-pairs, which will also contribute to the elevation of the Si dangling bond state to a shallow level. There must then be a bonding orbital lying below this state and this shows that the O atoms are to some extent overcoordinated. Therefore the transformation of the deep to shallow level occurs through a combination of electrostatic repulsion together with some dative bonding of the Si radical with the oxygen atoms.

8.10.1 Discussion and Future Applications

One of the fascinating aspects of this model is that this is a new mechanism for the formation of shallow donor levels, and only initially requires a defect with a deep level state. In effect the compressive neighbouring O_i atoms act to form a localised quantum well that contains the electronic gap state. This quantum well is compressed due to the Coulombic repulsion and dative bonding of the confining atoms, and this increases the energy of the locally confined electronic state, which is trapped in this well.

There are many implications of this work beyond the direct attempt to determine the structure of STDs. This mechanism could provide a mechanism for creating shallow donor levels in a wide range of materials, including commercially important shallow donors in diamond. Ultimately it may even be possible to attempt 'defect level engineering', treating a material in such a way as to produce defects with energetically 'squeezed' levels lying at precisely the energies required for device design.

In order that the wonderbra mechanism can operate, a core defect has to have only two fundamental properties:

1. It must possess a deep lying gap state. Although the precise depth is not important, a relatively localised level is useful since the compression mechanism is also relatively localised, so it helps if the electronic state is initially confined.

2. It has to be able to attract further defect atoms in order to confine and compress the gap state; the spatial arrangement of this gap state will be important in this respect.

It is possible to imagine various ways in which this could occur, although the $\langle 110\rangle$ type N-O_2 structure is possibly one of the simplest. For example, a similar mechanism could operate through a $\langle 111\rangle$ dangling bond state, which could be compressed by a non-bonded substitutional atom sitting on the original neighbour site. A third caveat is that where the electrostatic attraction towards the defect core is only weak, such as in the (CH)- shallow thermal donors, some form of external compression may also have to be applied, such as by the addition of further O_i atoms.

The mechanism provides an alternative method for producing shallow donor states in materials that are traditionally hard to dope, for example, $n-$ type diamond[224]. Instead of finding ways of introducing shallow dopants, the challenge now becomes one of finding ways to compress deep donor states already available in the material in order to shift them upwards in the gap. It is possible to dope diamond with substitutional nitrogen, which in the neutral charge state has a donor level 1.7 eV below the conduction band [225]. This sits in an anti-bonding orbital between N and one of its C neighbours, and hence N does not remain on the T_d lattice site but instead drifts away to form a C_{3v} defect. It could be imagined that the anti-bonded C could be replaced with some other defect such as Ge, Si, O or F in order to compress the N donor state through this mechanism. However in practise incorporating such impurities is extremely difficult. In addition any charge redistribution effects will be less marked due to the higher electronegativity of C than Si, and so any application of the wonderbra mechanism to diamond doping will probably be more complicated than this. Another alternative source of this mechanism could come through sp^2 hybridised carbon atoms where the filled non-bonded $p-$ orbital could be compressed.

Some preliminary test calculations by Latham and Jones have shown that a N_sO_{2i} structure similar to that described above is potentially a single shallow donor in diamond [226]. This could prove useful, as N_s is a common impurity in diamond. The challenge is now thrown open to experimental groups to determine ways in which these defects can be constructed in real materials.

Chapter 9

Thermal Donors in Si

" If the brain were so simple we could understand it, we would be so simple we couldn't."

Lyall Watson

Thermal donors remain one of the great unsolved problems in defect physics. First discovered over 45 years ago, their microscopic structure still remains a mystery. They are a family of up to 17 double donor defects [227] that appear in oxygen-rich Si (TD1 $\rightarrow$ TD17) when annealed in the temperature range 300-550 °C [227, 228, 229]. Higher temperature anneals lead to their destruction but also the growth of other, 'new donor' centres. In addition, higher temperature anneals lead to quartz precipitation, and associated build-up of Si_i and prismatic dislocation loops. Therefore Si is often subjected to rapid high temperature anneals to break up these oxygen clusters followed by quenching before other defect centres have time to form.

There is incentive to understand thermal donors from two camps. Firstly decreased tolerance chip design means that future processing will demand lower temperature anneals and less processing steps, which means that a thorough understanding of the thermal donors will become increasingly important. Secondly, despite being probably the most studied point defects in semiconductor materials to date they have remained intransient, and thus understanding them provides one of the greatest challenges to the point defect research community[1].

Experimental examination has revealed no other impurity atoms in these centres apart from O_i, and so it appears that they can only consist of O_i, possibly Si_i or V, and lattice Si. Any complete model appears doomed to conflict with some of the experimental evidence, given the amount of material collected over the years.

[1]Rob Ammerlaan, in the NATO Oxygen'96 ARW proceedings says, 'By inspection of the [...] unsolved problems related to TDs one concludes that whereas the general properties of these centres are well studied and, for the major part, understood, the microscopic picture is nearly completely missing'.

Hence we attempt to postulate here a model which agrees with the majority of results currently available; there are no doubt exceptions in the literature[2].

In this chapter we consider a number of different models for the thermal donor based around only O_i and lattice Si, and show that a model consisting of four O_i atoms, based around a 'di-y-lid' core, can account for many of the experimental observations. We discuss possible formation methods for these defects.

9.1 Experimental data on the thermal donors

To cover in detail the available experimental data on the family of thermal donors could easily fill a book in itself; the reader is referred to various articles in the NATO ARW proceedings from "Oxygen '96" [230]. Therefore I will only summarise the results here that are of key importance for determining the microscopic structure of the thermal donors.

9.1.1 EPR and ENDOR

The results of EPR and ENDOR concerning thermal donors are well summarised in [10, 231]. Relatively high TD concentration is required which corresponds to quite long sample anneal times, and hence EPR and ENDOR has only ever been applied to the thermal donors TD$n, n \geq 3$, since TD1 and TD2 anneal out before the experiments are performed [232].

EPR detects two different types of heat treatment centre in Si, NL8 and NL10 [233, 204]. These have since been identified as the singly positive charge states of the thermal donors and shallow thermal donors respectively. The symmetry of these was shown to be orthorhombic-I, or C_{2v}, with possible small deviations [207]. Thus they have two perpendicular (110) mirror planes and a twofold axis $\langle 100 \rangle$ [234, 204]. Longer anneals showed a shift in the $g-$tensor, indicating the gradual change from earlier TDs to later species in the family. The anisotropy of the $g-$tensor suggested the NL8 defects were shallow, but not as shallow as the NL10 centres. No hyperfine splitting of the spectrum was observed with ^{17}O, since the hyperfine interaction is below the detection limit of EPR [210, 235].

ENDOR showed $s-$type character to the donor wavefunction, indicative of a shallow EMT donor [10]. Only super-hyperfine interactions (shf) with Si and O were observed, suggesting that no other impurities are involved in the centres. It failed to detect a single Si atom with prominent spin localisation, suggesting that the donor activity does not originate from a single Si atom [10]. Examination of

[2] Seamus McQuaid, private communication, 'Some of the experimental (not to mention theoretical) data must be misleading'

the ^{17}O shf interaction showed the O lies in a (110) plane [235, 232]. No evidence of an electrically active core was found, including no signal corresponding to oxygen lying on the C_2 axis. Notably ENDOR did not detect a change in the number of oxygen atoms in the different TD centres, in contradiction with other techniques such as FTIR [10]. However the number of ^{17}O signals increases with long anneal time, suggesting oxygen addition [231], and analysis of the growth in ^{29}Si and ^{17}O signals with time shows an over-proportional increase in ^{17}O signal, *i.e.* the amount of oxygen per thermal donor seems to be increasing with anneal time. The TDs which anneal out earlier have deeper donor states as can be seen by their ^{29}Si shf interaction [231], *i.e.* the later TDs in the series have increasingly shallow levels. There were two shells of oxygen nuclei observed [232]. Thus, since the defect possesses C_{2v} symmetry it was assumed this translated into four oxygen atoms [232]. ENDOR also shows that each TD has only one Si atom on the C_2 axis[231, 232].

So to summarise, EPR and ENDOR provide a great deal of information about the core structure of the thermal donor. It must have C_{2v} symmetry, contain only O and Si, with all O lying in the $\langle 110 \rangle$ plane. It can only have one Si atom on the C_2 axis (which cannot be the sole source of defect activity) and no oxygen atoms. The core O atoms then lie in two distinct shells, and coupled with the planar constraint on oxygen that implies at least four oxygen atoms in the core. The addition of further oxygen atoms must not change the core signal, however ENDOR predicts the addition of oxygen to TDs with annealing, leading to increasingly shallow levels.

9.1.2 Stress-induced alignment experiments

Stress-induced alignment experiments [236, 237], primarily on TD3, show that the defects can reorient with a barrier of 2.7±0.2eV, likely to be through interstitial oxygen motion. The time constant is however 5 times longer than that of isolated interstitial oxygen, suggesting that reorientation time is limited by ∼5 O_i hops. It was shown with IR that this time constant increases with higher order TDs [238].

Analysis of the NL8 alignment gives the average effective C_{2v} strain coupling tensors shown in Table 9.1. This means the defect must be compressive along $\langle 001 \rangle$, tensile along $\langle 110 \rangle$, and weakly tensile along $\langle 1\bar{1}0 \rangle$. IR excitation spectra were used to show that the B_3 coupling coefficient dropped off with increasing TD^+ species [237].

9.1.3 FTIR

Many of the FTIR results are discussed in detail in the text so will not be covered here. There has been a recent major breakthrough in FTIR analysis of thermal

Average effective strain coupling tensor			Anneal Time	Primary Thermal Donor
B_1	B_2	B_3	(minutes)	
$\langle 1\bar{1}0\rangle$	$\langle 110\rangle$	$\langle 001\rangle$		
1.9	10.3	-12.2	90	TD3
1.1	8.5	-9.6	360	TD3, TD4
1.1	6.6	-7.7	600	TD5

Table 9.1: Average effective C_{2v} strain coupling tensor for the NL8 signal, heat treated under 600MPa stress at 460°C [236]

donors, as a series of vibrational absorptions in the region 975-1015 cm^{-1} have been identified as corresponding to the thermal donors [239]. In addition each of these centres also has a corresponding lower frequency mode at around 700-800 cm^{-1}. These lie very close to the dimer vibration at 1012 cm^{-1} (see Chapter 6), and the earlier TD1/2 modes are also observed in as-grown material. The annealing characteristics of these lines has been observed at 370, 420 and 450°C over long time scales, as well as their behaviour when subjected to high temperature dispersion treatments.

In addition, IR has been used to identify the electronic transitions associated with the thermal donors [240, 241], where the (0/+) and (+/++) occupancy levels were shown to lie at around $E_c - 0.07$ and 0.15 eV respectively, with excited states in agreement with effective mass theory predictions.

9.1.4 Other information

There is various other important kinetic, electronic and impurity-complex related information that is relevant to this work. The initial total TD formation rate at 450°C is proportional to $[O_i]^n$, $n = 4$ [179] initially suggesting early TDs contained 4 O atoms; later work has shown this to be a function of both $[O_i]$[242] (for $[O_i]$ $\leq 8 \times 10^{17}$ cm^{-3} n was close to 1), and temperature[243]. At lower temperatures (T $\leq$ 400°C) $n = 2$, and this increases at 450°C to $n = 4$ as seen earlier. This suggests that at lower temperatures, oxygen dimerisation is the rate dominating step, but at higher temperatures this is replaced with simple O_i diffusion. At 500°C, n increases further.

Uniaxial stress measurements on the electronic IR absorption associated with the TDs has also allowed symmetry assignments. This also showed C_{2v} symmetry [244, 245], with the effective-mass like donor state constructed from a single pair of conduction band valleys [246] along [001] [247]. The C_2 axis is compressive and lies along [001] [248].

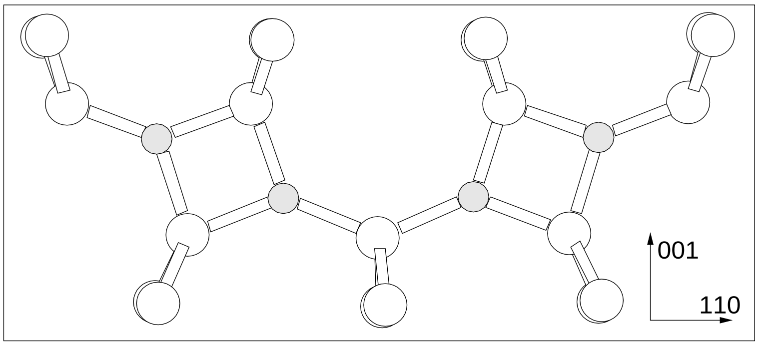


Figure 9.1: Schematic diagram of the 'di-oxygen square' model for the thermal donor, as proposed in [252]. This structure was found to be lower energy than the di-y-lid structure.

9.1.5 Previous theoretical models

There have been several thermal donor models proposed in the past, notably the O_2Si_i model [105], the O_{3i} model [249, 211], and the OBS model [250].

As the initial concentration of TDs was proportional to $[O_i]^4$, initial models concentrated on defects containing four O_i atoms. Kaiser *et al* proposed a SiO_4 tetrahedral structure [179]. This was believed to become inactive on addition of a fifth O_i, and the model was abandoned when it was discovered that there was a whole family of TDs.

Later work concentrated on trivalent oxygen in different complexes. The strongest contender here is O_{3i} with the central O atom in a 'y-lid' trivalent structure (see Figure 9.13). This was originally proposed in 1983 [249], and has been modelled by several groups since [93, 251, 211]. Since this model features O on the C_2 axis it cannot account for TD3 or higher (ENDOR shows no O on the C_2 axis), however it could be TD1 or TD2. This was shown using first principle supercell methods to be 0.7 eV lower in energy than three neighbouring BC O_i atoms [211]. This structure was also proposed as a core with two further flanking O_i atoms, forming a 5O species [28] (see Figure 9.11). The central O atom becomes trivalent in response to the compression from the other oxygen atoms, and is the source of the electrical activity. However ENDOR excludes this as a model for TD3+ for the same reasons as the 3O 'y-lid' species.

To avoid the problem of O on the C_2 axis, a di-y-lid model was proposed by two groups [111, 252] (see Figure 9.4), and later work suggested that the 'y-lid' trivalent structure discussed above could transform into this with the addition of further O_i atoms [211]. CNDO/S calculations found this to be 1.5 eV unstable with respect to another alternative 4O species, the di-oxy-square [105] (see Figure 9.1), consisting of four trivalent oxygen atoms in two neighbouring squares. This model was also discarded since the donor wavefunction is concentrated on the square structures

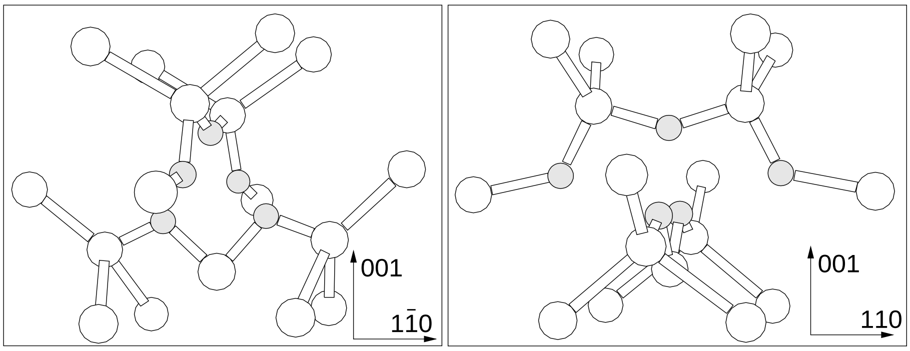


Figure 9.2: Schematic diagram of the OSB model for the thermal donor [250]. The diagrams show the same defect along perpendicular directions. The bottom Si atom is a lattice atom displaced a long way from its original site.

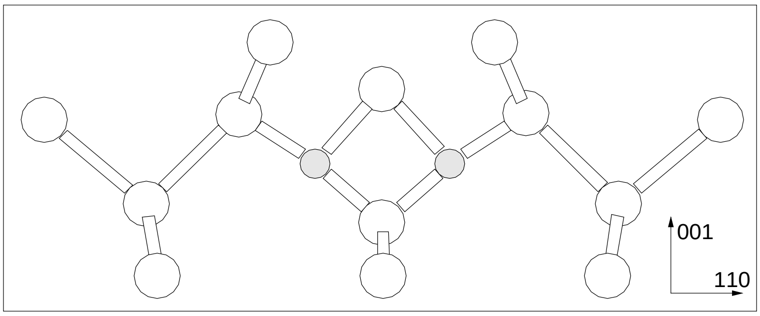


Figure 9.3: Schematic diagram of the Deák / Snyder / Corbett thermal donor model containing two trivalent O_i atoms and a Si_i [253].

[252] in disagreement with that deduced by ENDOR. The di-ylid consisting of two oxygen atoms in y-lid configurations without the flanking O_i atoms was modelled, but this is extremely tensile and unstable with respect to other configurations such as the BC dimer [105].

Prior to the ENDOR result that all O lies on the same C_{2v} plane with no O on the C_2 axis, one of the more popular models was the OSB model, shown schematically in Figure 9.2 [250]. This assumes that oxygen aggregation and compression pushes a core lattice Si off its site to become a divalent bridging Si atom, which is the source of the donor activity. The donor can then become inactive through the ejection of this atom as an interstitial. The model is compressive along $\langle 1\bar{1}0 \rangle$ and was abandoned in light of the new ENDOR data.

The model that has received strong support in the last few years is the Deák model, shown in Figure 9.3. This consists of two tri-valent oxygen atoms pinning a Si_i between them which is divalent [253, 105]. The electrical activity comes from the trivalent O atoms, which donate electrons to the empty states on the Si_i.

There are several reasons why the Deák model cannot be correct. One problem comes through the bistability of the early TDs, TD1 and TD2, which exhibit isomeric electrically inactive structures [254]. There seems no easy way of reconstructing the Deák model such that the Si_i core atom can become fully coordinated. It is possible for several Si_i to aggregate in such a way that all of the dangling bonds combine, but not with a single atom. Alternative structures with trivalent Si_i were proposed as the inactive forms, since the CNDO calculations gave its dangling bond state in the valence band [253]. However AIMPRO calculations of trivalent Si give this state deep in the band gap [255]. In addition, the vibrational mode agreement of the model with experiment is poor. The corrected calculated CNDO modes lie at 984 and 901 cm^{-1}, when calculated with AIMPRO we get modes at 823.5 and 805.4 cm^{-1}. Experimental modes for TD3 are observed at 999 and 728 cm^{-1} so the agreement is poor. There are other subtleties, for example ENDOR shows only a single Si atom on the C_2 axis whereas this model has two; this has been explained in terms of low electron density on the Si_i[256].

In conjunction with this, Si_iO_i was proposed as the structure of TD1 which would rapidly migrate and trap an additional O_i. However Newman points out that Si_iO_i has been observed and is only stable to 50 °C [257], with a binding energy of somewhere near 0.8 eV. Thus this is ruled out as a model for TD1.

Various other models have been proposed over the years. It was suggested that an O_2 molecule in a vacancy might have double donor character if its antibonding $p\pi^*$ state crossed the gap[258], but this was disproved with self-consistent local density Greens function calculations of Kelly[259].

9.2 Results

We have modelled several of the complexes described above, the results of which are given here.

9.3 The 4 O di-y-lid thermal donor

The structure of the di-y-lid is somewhat different to that of the STDs examined in the previous chapter. The defect consists of a pair of oxygen dimers; all the other atoms are lattice Si. These dimers lie in the same $\langle 110 \rangle$ plane but are separated by two empty Si-Si bonds; the electrical activity occurs when the inner pair of O atoms become trivalent, both bonding to the central lattice Si (see Figure 9.4). The defect has no oxygen atoms on the C_2 axis along $\langle 100 \rangle$, consistent with the EPR and ENDOR data available on the thermal donors. Crucially it has a double donor level very close to the conduction band. Since at anneal temperatures this

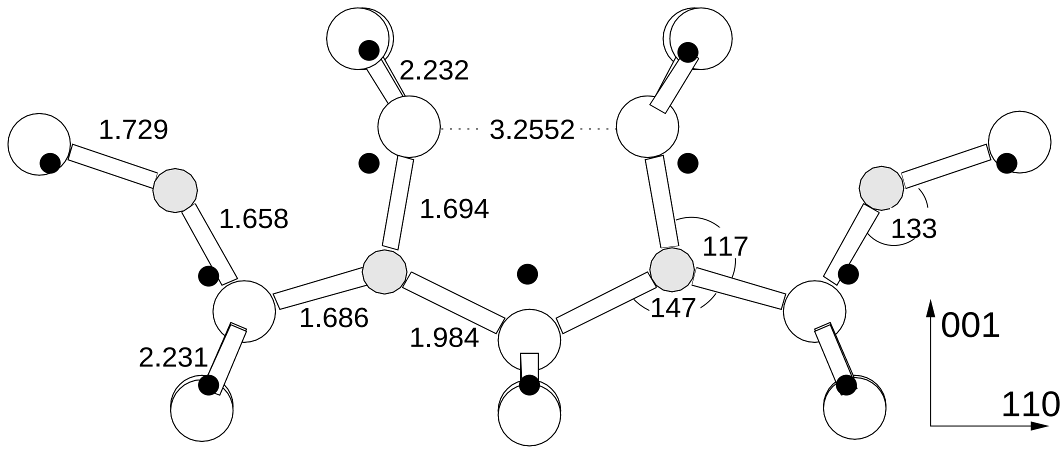


Figure 9.4: The core of the di-y-lid thermal donor. Bond lengths are in Å, bond angles are in degrees. Si atoms are white, O atoms are grey. Black dots mark the ideal Si lattice sites.

would normally be depopulated, in all of the simulations discussed below we have modelled the defect in the +2 charge state, *i.e.* this shallow level is depopulated. Exceptions to this are discussed in the text.

The cluster used was $Si_{79}H_{68}O_4$, charged +2. Weak additional quadratic spring potentials ($k_r = 5.0$ eV/Å^2) were used to constrain the surface H atoms, in order to simulate the effect of the rest of bulk. The structure is shown in Figure 9.4, along with bond lengths (Å) and angles (degrees). The black dots mark out the ideal Si lattice sites and show the displacements and resultant stress caused by the defect. The whole cluster, with surface hydrogen atoms removed for clarity, is shown in Figures 9.7 and 9.8. These diagrams clearly show the tensile/compressive effect the defect has on the surrounding lattice. It is highly compressive along $\langle 100 \rangle$, and tensile along $\langle 110 \rangle$ and $\langle 1\bar{1}0 \rangle$, at least closer to the core. The $\langle 1\bar{1}0 \rangle$ tensile strain is only extremely weak, and it may be that the shape of our cluster, with very little bulk Si along $\langle 1\bar{1}0 \rangle$ from the defect, provided too great a constraint in this direction.

The defect essentially consists of a pair of dimers separated by what would normally be two $\langle 110 \rangle$ Si-Si bonds. However the dimers have tilted in so that the inner two O atoms are somewhat overcoordinated (whether they are truly trivalent is a debatable point, since the inner Si-O bond is extremely long for a covalent bond). Meanwhile the top two Si atoms in the defect core form a shared reconstructed bond. Thus it is possible to either view the structure as a defect core with fully coordinated Si atoms and two trivalent oxygen atoms, or alternatively with two roughly divalent oxygen atoms and an undercoordinated core Si atom, whose remaining non-bonded electron pair are being electrostatically compressed by the oxygen atoms. The reconstructed Si-Si bond is extremely long, and it is possible

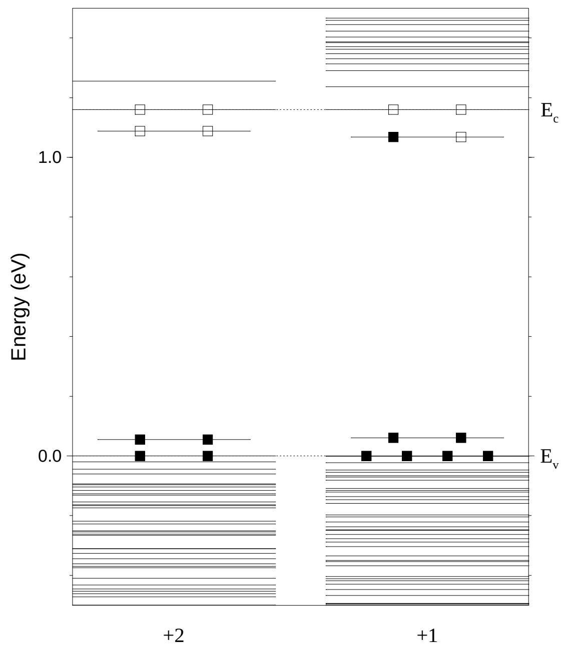
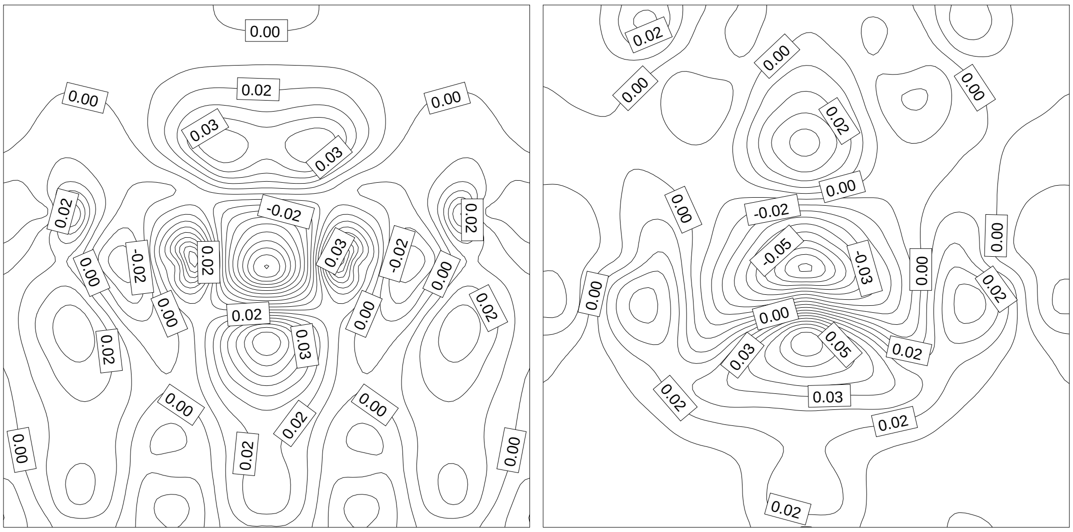


Figure 9.5: The top Kohn-Sham eigenvalues for the $4O_i$ di-y-lid model for the thermal donor. Charge state of the defect is given beneath each column. Black boxes denote filled state, white boxes for empty states. The eigenvalues have been arbitrarily scaled to the experimental band gap of 1.16eV.

that this is also contributing an anti-bonding state to the defect, similar to the 'A-centre' (OV^-) which has a next neighbour dilated Si-Si bond, whose antibonding equivilent can act as an electron trap to make the defect negatively charged. This reconstructed bond is also important when considering strain effects of further dimer addition (see below).

Several of the core atoms move slightly away from the ideal $(1\bar{1}0)$ lattice plane, and without this shift the defect is not stable. All the displacements described are along the same direction. The top two Si atoms and the two core O atoms are displaced by 0.062 Å and 0.058 Å from the $(1\bar{1}0)$ plane respectively (these figures are the same within the numerical limits of AIMPRO). There is some assymetry in the back Si-Si bonds of the core Si atom (2.218 and 2.230 Å). The outer two O atoms are displaced out-of-plane by 0.074 Å. The other three 'core' Si atoms are displaced by less than 0.025 Å. These displacements may be sufficient to explain the 'slight deviations' from C_{2v} symmetry observed by EPR[207].

The Kohn-Sham eigenvalues are given in Figure 9.5, scaled to the experimental band gap of 1.16 eV. As can be seen, the defect possesses a shallow double donor level which is depopulated in the +2 charge state. Adding a single electron to the system and relaxing once more does not produce any significant change in structure or vibrational modes. The donor level appears to drop very slightly away from the



(a) (b)

Figure 9.6: Cross-sections through the Kohn-Sham wavefunction for the donor state of the di-y-lid thermal donor model in the +2 charge state. Vertical direction is $\langle 001 \rangle$, horizontal direction is (a) $\langle 110 \rangle$, (b) $\langle 1\bar{1}0 \rangle$.

conduction band in this case. Slices through the donor wavefunction are shown in Figure 9.6. Note that in practise the donor states will be effective-mass like and will be considerably more delocalised, however they are confined here in a finite cluster.

We next examined the vibrational modes of the defect core, along with the square of the dipole moment for each mode, which is linearly proportional to its absorption strength in FTIR. The results are given in Table 9.2. As can be seen, there are two modes lying at 999 cm^{-1} and 734 cm^{-1} which are more intense than the others. We also examined the shift in these modes with differing oxygen isotope, in Table 9.2. The agreement with the experimentally observed LVMs for the thermal donors is remarkable. Unfortunately the accuracy of our method does not allow us to distinguish between the different thermal donors, whose vibrational modes only differ by a few wavenumbers. There is however agreement with the observed TD3 LVMs to less than 0.5% in the higher mode, with $\sim$3% error in the lower mode. The core atom displacements associated with each mode are shown in Figure 9.9. It can be seen that the two most intense modes are primarily a high frequency asymmetric oxygen stretch mode associated with the core O_i atoms (stretch in the $\langle 001 \rangle$ direction), and a lower frequency out-of-plane wag mode with the two core O_i atoms once more, this time vibrating in phase. Thus they crudely approximate to a coupled stretch, and a wag mode.

LVM (cm^{-1})	(Dipole moment)^2 (Intensity)	Source	^{16}O	^{17}O	^{18}O
1050	0.126	Experiment			
1000	**0.249**	TD2	988		945
903	0.206		724		
891	0.151				
808	0.190	TD3	999		955
762	0.019		728		
751	**0.323**				
745	0.001	Di-ylid	1000	976	955
633	0.113		751	738	726

Table 9.2: Vibrational modes (cm^{-1}) and associated absorption intensities for the di-y-lid thermal donor model. Stronger modes are picked out in bold. Second section shows shift with isotope for the di-y-lid thermal donor model, and the experimentally observed values for TD2 and TD3 (room temperature).

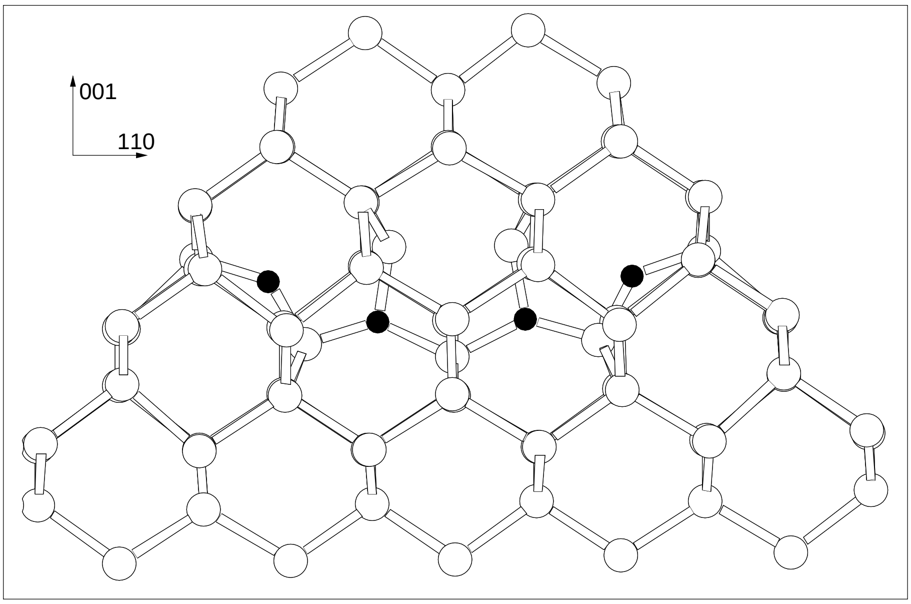


Figure 9.7: Side view of the $4O_i$ atom di-y-lid TD. This shows the complete cluster except for the surface hydrogen atoms. The oxygen atoms are shown in black.

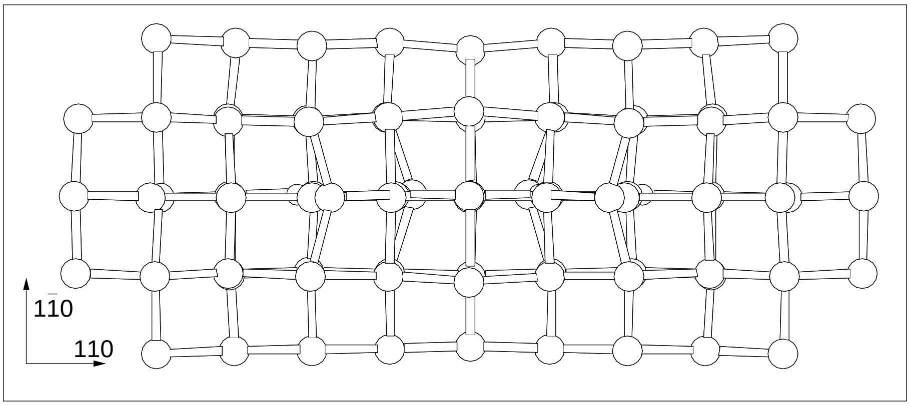


Figure 9.8: Overhead view of the 4O$_i$ atom di-y-lid TD. This shows the complete cluster without surface H atoms. The tensile stress in the $\langle 110\rangle$ and $\langle 1\bar{1}0\rangle$ is visible in the distortion to the lattice.

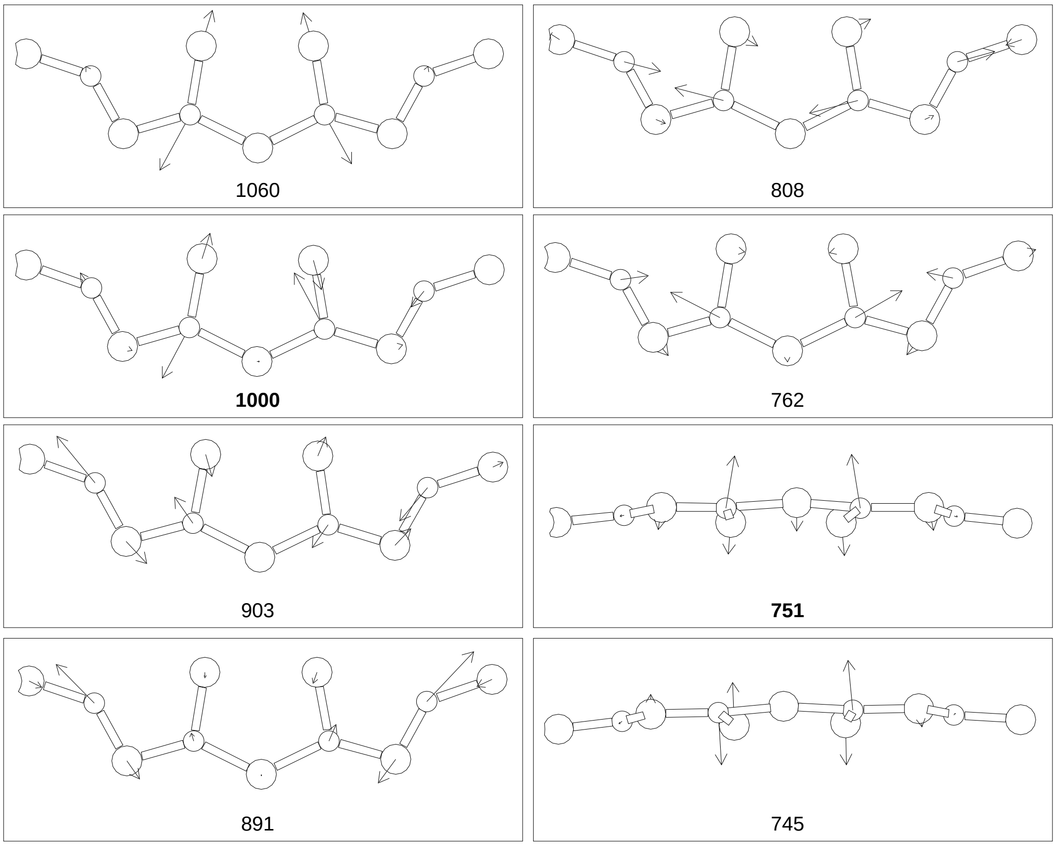


Figure 9.9: Displacement of the core atoms associated with each calculated LVM (given in cm^{-1}). Vector length has been exaggerated for clarity. Overhead views of the 751 and 745 cm^{-1} modes have been displaced slightly off axis so that it is possible to separate the vectors due to the O and Si atoms in the core. Stronger modes are given in bold.

9.3.1 Discussion of the di-y-lid model

The di-y-lid model is consistent with a wide range of experimental thermal donor observations, many of which have been discussed above.

Firstly it contains four oxygen atoms in agreement with early oxygen loss experiments [179]. It consists of a pair of dimers, consistent with dimer formation being the low temperature rate limiting step[143]. It has the correct C_{2v} symmetry with all O atoms in the same $\langle 110 \rangle$ plane, and none of them on the C_2 axis. In addition it contains only one Si atom on the C_2 core which is displaced significantly from its ideal lattice site [231, 232]. The donor activity arises from O instead of a single Si atom in agreement with ENDOR[10]. The defect has two shells of O nuclei [232] corresponding to the trivalent and then divalent O atoms. It exhibits the correct compressive and tensile properties along the various $\langle 001 \rangle$, $\langle 110 \rangle$ and $\langle 1\bar{1}0 \rangle$ directions[236].

The defect gives rise to the correct number of vibrational modes, with both absolute values and isotope shifts extremely close to those experimentally observed[239]. Since it consists of a pair of dimers, its entire formation path can consist of dimer migration and addition, hence the formation barrier is always 1.7 eV and never needs to involve isolated O_i.

It should have an associated inactive form where the two trivalent O atoms become divalent once again; this could either relax back into two dimers separated by a pair of Si-Si bonds, or else into a central dimer flanked by two O_i atoms separated by a single empty Si-Si bond (with further structural rearrangement it is possible to imagine other inactive forms). A final alternative is if the two trivalent O atoms form next-neighbour bonds along $\langle 110 \rangle$, while the top two Si atoms form a stronger bond which when H passivated could be inactive (see below). The modes of this inactive form might be expected to be slightly higher than isolated dimers given their proximity; experimentally whereas the dimer absorbs at 1012 cm^{-1} the inactive TD form absorbs at 1020 cm^{-1} [239], consistent with this interpretation.

In agreement with stress induced alignment[236] the defect should be able to reorient within about five 'standard' O_i hops, if some of the motion is coupled and slightly faster. For example, the two core O atoms initially perform a 'coupled hop' downwards into the Si-Si bonds lying along $\langle 1\bar{1}0 \rangle$. Next the outer two O_i atoms perform standard 'hops' to replace the original core atoms. The two O atoms now lying along $\langle 1\bar{1}0 \rangle$ next hop outwards, and finally the two atoms remaining on $\langle 110 \rangle$ perform a 'coupled hop' downwards to adopt the new core sites along $\langle 1\bar{1}0 \rangle$. If these 'coupled hops' in total take the same time as a standard O_i hop this would match experiment.

Earlier experiments showed that TD concentration decreased near the sample surface, with rates consistent with oxygen diffusion that was enhanced by around four orders of magnitude[139]. From this it is expected that the out-diffusion should increase with $[O_i]$, but it was found that the depth of depleted material decreases with increasing O concentration[260, 261]. This is consistent with increasing oxygen concentration leading to more dimers, which have thus a higher chance of being trapped by a second dimer to form a thermal donor before out-gassing. Thus this result supports a thermal donor model consisting of the binding of two or more oxygen dimers.

The reconstructed Si-Si bond is interesting. Since it is quite dilated it could potentially absorb a lot of the surrounding lattice strain. In this case oxygen dimers might sit in bond sites near to the top of the defect so that their lattice strain is absorbed in this way. This is discussed further below with relation to higher order TDs. Such dimers would be able to rapidly respond to applied stress if the defect core reoriented.

Results of Hallberg showed an initial 1.2 eV formation rate for the TDs that rapidly reverted to the standard 1.7 eV[239]. If the 1006 cm^{-1} mode does correspond to the trimer, then as-grown material contains an initial pool of trimers, and this 1.2 eV could correspond to the migration barrier for the trimer to travel to O_i in order to form the di-y-lid thermal donor. Once this pool is exhausted, the dimer migration barrier then becomes the rate limiting step. This is discussed further below.

Carbon passivation of thermal donors may occur by C substitution into the core of this instead of the central Si atom, forming C_sO_4 which would be inactive. This would either form by C_s trapping dimers initially, or else through C_i migrating to a TD and kicking out the core Si atom to replace it. This could go on to kick out a further C_s atom elsewhere and so the process is catalysed and only requires a small number of Si_i atoms to start it off.

Note that this discussion does not address the question of whether this structure represents TD2 or TD3, which is discussed further below.

9.3.2 Hydrogen passivation

It is known that hydrogen can passivate thermal donors, in a fully reversible manner that is only stable to $T \sim 200°C$ [203]. There are a number of unusual features of this hydrogen passivation:

1. The attempt frequencies for dissociation are at least three orders of magnitude larger than expected for an atomic jump process. This could be explained by an associated large lattice relaxation.

2. Dissociation of TD-H occurs in a single stage. Earlier work suggested this occurred with the same dissociation enthalpy of 1.1 eV for both charge states and different TDn-H species [262]. This suggests that a single hydrogen atom can passivate a double donor, which would not be expected in a simple 'dangling bond' termination model.

3. More recent work [203] showed the dissociation enthalpy for TD1-H is 1.90 eV, whereas for the other TDs is $\sim$1.64 eV, suggesting that the binding in TD1-H is different to the other TDs. This is consistent with the suggestion that the core of TD1 is different to the other TDs.

4. Earlier TDs are more effectively passivated than the later ones [263].

There are two potential ways in which hydrogen can passivate the di-y-lid structure. The simplest is for the H atom to form a Si-H bond with one of the Si atoms connected by the reconstructed bond at the top of the defect. If both of these Si atoms form Si-H bonds, then they become fully coordinated in combination with their two top Si-Si bonds and their shared bond. This allows them to relax away from the defect core in order to shorten the reconstructed bond. At the same time the oxygen atoms can drop back to become divalent and sit in next-neighbour positions between the core Si atom and the Si shared with their outer O atom. This structure is then fully passivated with no gap states, and exhibits fairly standard Si-H bonds which should have dissociation enthalpies in the region of the experimentally observed values. However, this process requires two H atoms for full defect passivation, in contradiction with experiment.

When we attempted to model this by placing H between the core O and the reconstructed Si atoms, this process began to occur. However the structure continued to relax with the H atoms attempting to migrate around the Si atoms to anti-bonding sites, resulting in extensive rebonding in the top of the cluster and eventually the breakdown of the calculation. This suggests that these Si- bonded sites were not stable locations for the H, and casts doubt on this passivation mechanism.

An alternative way to 'passivate' the defect is to place a single H atom in the centre of the reconstructed Si-Si bond. This disrupts the core bonding, and the defect is able to relax backwards into two standard dimers, possibly with them both buckling in the same direction instead of towards each other. This would be a neutral, electrically inactive defect. Since H has a partially filled deep level and the material contains many TDs, etc., this level will fill, forming H^-. In this case the H is not covalently bonded to any of the Si atoms, and any bonding would be Coulombic. This model disagrees with experiment, since the H^- still has a filled level deep in the gap, possibly quite close to the valence band. However it is possible that experiments using DLTS were only measuring signals near to the conduction band (or at least, in the top half of the gap), and hence registered only the loss of the thermal donor levels and not the emergence of such a deep level. We are awaiting further information on this subject from J. Weber, who has performed the DLTS experiments.

This bonding has some similarity with the H_2^* defect. H^- normally prefers to sit anti-bonded behind a lattice Si atom. In the H_2^* defect, a H^+ atom sits at a bond centre and a H^- atom sits at the neighbouring anti-bonding site. It is stable to $\sim$150°C, although two Si-H bonds are formed in this process.

In this passivated di-y-lid model, the Coulomb attraction at the central site is much stronger, although there are no actual Si-H covalent bonds forming. The H^- ion is positioned between two very positive Si atoms, since they both have O neighbours and so are $\delta+$. In addition, each dimer is pushing these Si atoms towards each other, so the H^- ion is sat closer to the atoms than it would normally be in a next neighbour site. For this reason it is feasible that the binding energy might be higher than that of H_2^*, at $\sim$1.6 eV.

The fact that earlier TDs are more effectively passivated could be because in the larger TDs, the di-y-lid core is relieving the strain of many surrounding dimers. There is therefore less energetic incentive for it to switch back to the double dimer inactive form. This is similar to the argument as to why earlier TDs exhibit bistability whereas later ones do not.

We tried relaxing this model with H in the defect core, overall charged -1. As expected the H formed two reasonably strong bonds (1.540 Å) with its neighbouring Si atoms, however this was at the expense of Si-Si bonds to the core Si atom, which dropped down along $\langle 001 \rangle$ from the defect core. Thus the resultant structure had levels in the gap due to this roughly divalent Si, and the H did not passivate the di-y-lid. We are currently examining variations on this structure to see if it is possible to passivate the di-y-lid in this way with small structural rearrangments.

9.4 Other models

It is possible to construct various other defect structures that act as double donors, involving different numbers of oxygen atoms (O_{ni}, $n \geq 3$). We have also modelled many of these for comparison with the di-y-lid model, and to determine whether they could be responsible for other members of the thermal donor family.

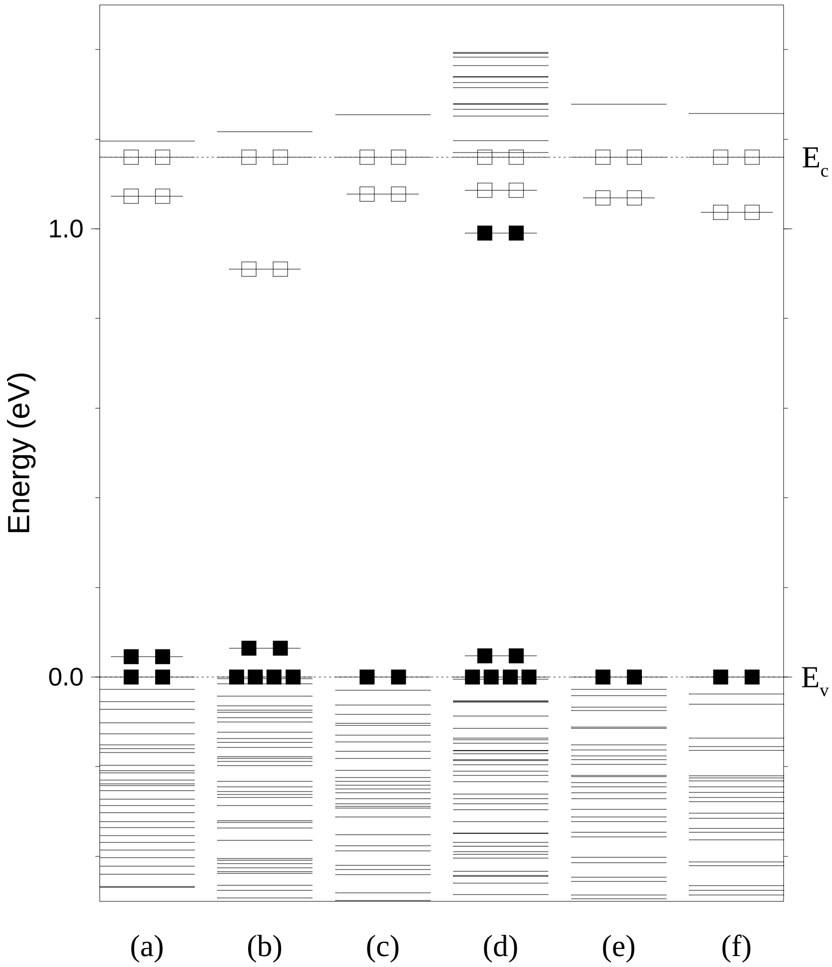


Figure 9.10: The top Kohn-Sham eigenvalues for various alternative thermal donor models, all in the +2 charge state. Black boxes denote filled state, white boxes for empty states. The eigenvalues have been arbitrarily scaled to the experimental band gap of 1.16eV. Structures are $3O_i$ species (a) Snyder/Stavola, (b) Partially dissociated trimer, $4O_i$ species (c) 'Flanked square' model, (d) Di-square model, (e) $5O_i$ STD analogue, and (f) $6O_i$ di-y-lid with dimer in central Si-Si backbonds.

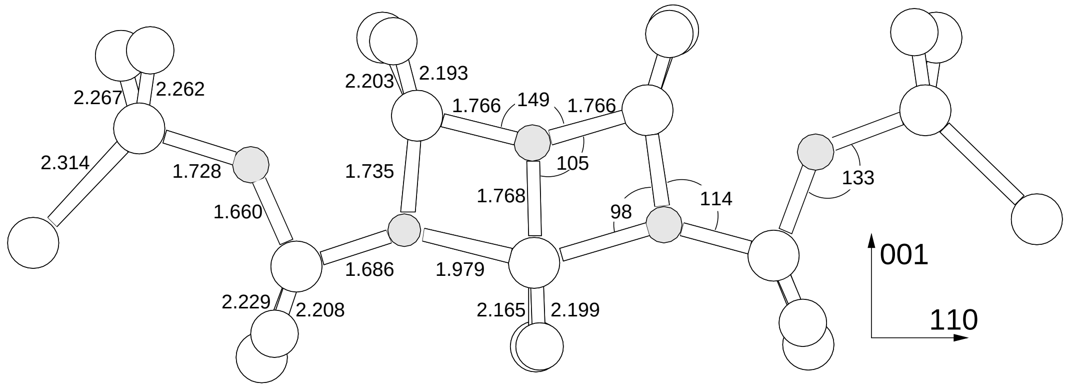


Figure 9.11: The core of the 5O thermal donor model. Bond lengths in Å, angles in degrees. Symmetrically equivilent bonds have not been marked with lengths or angles. Si atoms are white, O atoms are grey. The difference between the central y-lid and its two neighbours can be seen in the bond lengths.

9.4.1 The $5O_i$ Thermal Donor

This structure is a direct analogue of the $(CH)_i$-O_{4i} shallow thermal donor defect, with the $(CH)_i$ unit replaced by a single trivalent O_i atom, forming O_{5i}. The two outer atoms provide compression, which stabilises an inner core of three O atoms, all of which are to varying degrees tri-valent. A central lattice Si atom sits at the defect core and is between 3- and 5- valent depending on the definition of a bond in this core. Since this has O_i on the C_2 axis it can only be a potential model for TD1 or TD2 since ENDOR shows TD3 and above have no O_i on the C_2 axis.

The resultant structure is largely C_{2v} and very similar to that of the shallow thermal donor structure, $(CH)_iO_{4i}$ (see Figure 9.11). There is slight deviation from C_{2v} in that the two roughly tri-valent oxygen atoms either side of the core (B and C) move 0.01 Å out of the defect plane in the same direction. The core O atom remains, to within 0.005 Å, on the C_2 axis. It is compressive along $\langle 001\rangle$ and tensile along $\langle 110\rangle$.

The structure breaches ENDOR observation [231, 10] that the higher TDs ($TD_n, n \geq 3$) have no oxygen on the C_2 axis, and in addition have two shells of oxygen atoms around the core (this defect displays three: the core O atom, the next roughly tri-valent O atoms, and then the outer O atoms). However this does not necessarily exclude the model since EPR has previously been unable to detect light element impurities lying on the C_2 axis of a defect (e.g. C_i^- [264]), and so it may be that the core O atom is invisible to this method.

The Kohn-Sham eigenvalues for this structure are given in Figure 9.10e, scaled to the experimental Si band gap of 1.16eV. Since the defect is +2 charged the shallow donor level is empty.

Local Vibrational Modes (cm^{-1}) ^{16}O	^{17}O	^{18}O	Dipole moment squared for ^{16} O
1644.1	**1604.8**	**1569.2**	0.386
1092.3	1067.0	1044.1	0.024
1081.3	**1057.2**	**1035.2**	0.407
1002.9	975.3	950.2	0.006
972.2	946.4	923.1	0.004
921.2	**898.2**	**877.5**	0.315
892.5	870.6	850.8	0.002
860.7	839.5	820.4	0.039
850.5	**832.3**	**815.9**	0.417
826.3	805.8	787.3	0.012
825.9	804.5	784.9	0.018
739.5	739.0	738.3	0.001
653.9	650.3	647.2	0.099

Table 9.3: Vibrational modes of the $5O_i$ TD model (cm^{-1}), IR-active modes are given in bold. Last column gives the associated absorption intensity of each mode (dipole moment squared).

This thermal donor model produces a number of vibrational modes, and these are shown in Table 9.3. Altogether there are four strong modes, one much higher than any experimentally observed values. The two central ones straddle the experimental figure, and both give reasonable isotope shifts. Experimentally the isotope shift from ^{16}O to ^{18}O is 44 cm^{-1}, for the 1081 cm^{-1} mode it is 46 cm^{-1} and for the 921 cm^{-1} mode it is 44 cm^{-1}. The existance of four modes with one of those extremely high does not provide good agreement with experiment, and the lower 850 cm^{-1} mode is a long way from the experimental lower frequencies of the TDs at 716-748 cm^{-1}. This therefore probably rules it out as a candidate for the thermal donors, in conjunction with its breach of ENDOR observation. This could be confirmed by examination of experimental FTIR spectra in the higher 1300-1700 cm^{-1} frequency region, where this model predicts an absorption. The two highest modes of this model are caused by out-of-plane wag motion of the core O, and its two O neighbours respectively; it is unusual for wag modes to have such high frequencies, and further investigations are required to ensure this is not simply a basis effect. By analogy to the dimer, the 917 cm^{-1} mode is an extremely coupled relation to the asymmetric stretch mode, and the 849.2 cm^{-1} mode is similar to the symmetric stretch.

In light of the poor LVMs and disagreement with ENDOR we therefore exclude this structure as a model for any TDs above TD2, and probably also as a model for TD1 or TD2.

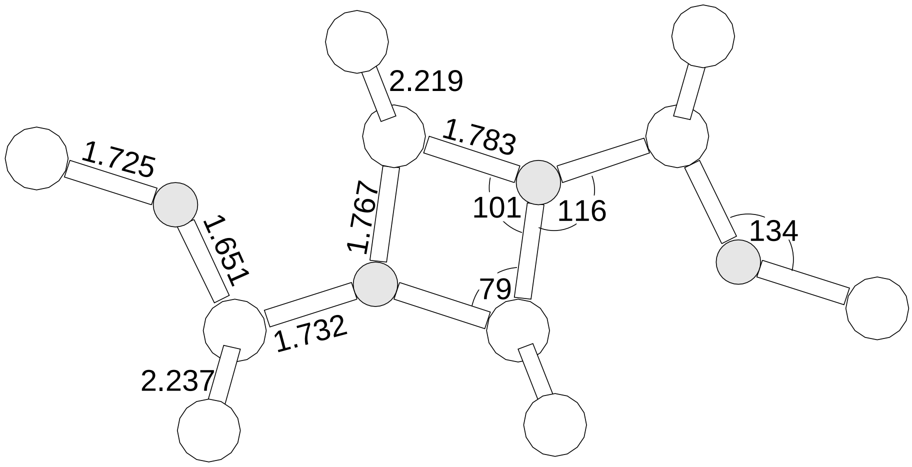


Figure 9.12: The core of the alternative 4O 'flanked square' thermal donor model. Bond lengths are in Å, bond angles are in degrees. Si atoms are white, O atoms are grey. Symmetrically equivilent bond lengths and angles have only been marked once.

9.4.2 Alternative 4O_i structures

9.4.2.1 The 'flanked square' structure

It is possible construct a thermal donor consisting of 4 O_i atoms in a different way to the di-y-lid. This can be done by removing O_i from the 5 O_i thermal donor described above. Removing the oxygen atom on the C_2 axis leads to the di-y-lid model, however removing one of the outer oxygen atoms, simply shortens the chain and removes some compression from the defect core.

The defect relaxed by rotating the core around to give an oxygen 'square' flanked by two standard divalent O_i atoms. This has C_{2h} rotational symmetry, the same as that of the N_{2i} pair. The eigenvalues are given in Figure 9.10c and its double donor character can be seen. Distortion in the structure was caused by cluster assymetry, so the calculations were repeated in a C_{2h} symmetry constrained 180 atom cluster, $Si_{98}H_{78}O_4$.

The structure is shown in Figure 9.12. The bonding in this defect is quite different to that of the di-y-lid, having no core reconstructed Si-Si bond. Whereas the di-y-lid can be thought of as two dimers held apart by two empty Si-Si bonds, this square-based structure could be formed from two dimers with only a single empty Si-Si bond between them. Such a structure does not have the required C_{2v} symmetry, and thus this could only be a candidate for TD1.

The vibrational modes of this structure are given in Table 9.4, along with their isotope shifts and their absorption intensities. This model has two strong absorptions at 904 and 742 cm^{-1}. The experimental TD1 modes lie at 975 and 716 cm^{-1}, unfortunately their ^{18}O isotope shifts have not yet been reported[239]. Thus the

Local Vibrational Modes (cm^{-1})			Dipole moment squared
^{16}O	^{17}O	^{18}O	for ^{16} O
946	921	898	0.000
904	**882**	**862**	0.179
854	831	811	0.000
811	791	773	0.100
742	**725**	**711**	0.206
741	723	706	0.000

Table 9.4: Vibrational modes of the alternative $4O_i$ 'flanked square' TD model (cm^{-1}), IR-active modes are given in bold. Last column gives the associated absorption intensity of each mode (dipole moment squared).

calculated modes are in roughly the right range, and without isotopic shift data these LVMs are not sufficient to exclude this as a model for TD1. The primary absorptions for TD1 are still quite weak, so it is unlikely that the mode at 811 cm^{-1} will be observable. In the assymetric cluster this structure was nearly 1 eV more stable than the di-y-lid and was the lowest energy 4O donor structure, although the error on such a figure will be large, and also influenced by cluster surface effects.

Symmetry analysis of TD1 would be a useful tool to distinguish between these models since this structure has well defined C_{2h} symmetry. Note that if this is TD1 it would add weight to the suggestion that the di-y-lid is TD2. This is because TD1 and TD2 are believed to be isomeric, since their respective electrically inactive forms show vibrational absorption at the same frequency[239]. In conclusion, symmetry excludes this 'flanked square' structure as anything except TD1, but the calculations would be consistent with what is currently known of TD1. Further experimental data is required to either confirm the assignment or eliminate it from enquiries.

9.4.2.2 The di-square structure

We also modelled the di-square structure shown schematically in Figure 9.1. Earlier CNDO/S calculations showed this to be more stable than the di-y-lid structure [105]. However there are several things wrong with this model, the primary fault being the associated defect states. The defect actually has four electrons in the gap near to the conduction band, since each square acts as a double donor. Hence running it in the +2 charge state still leaves two electrons close to C_b (see Figure 9.10d). This is sufficient to exclude it as a thermal donor model. In addition it was less energetically stable than the di-y-lid model.

For these reasons we exclude the $4O_i$ di-square model as a potential thermal donor model.

Calculating the modes for one of the two squares gave strong modes at 725, 544 and 533 cm^{-1}, and hence this suggests that O-square structures on their own cannot account for the observed experimental TD vibrational modes. However the mode at 725 cm^{-1} is close to the lower modes observed, and hence it is possible that a square structure in combination with other O_i could be responsible for TD1. This mode corresponded to an in-plane stretch of one of the trivalent oxygen atoms.

9.4.3 Three O atom TD species

We also modelled various three-oxygen species; a y-lid flanked by two BC O_i atoms, previously proposed as a thermal donor core, and hereafter referred to as the Snyder-Stavola model[249, 211]. We also examined a partially dissociated structure consisting of a pair separated by two Si-Si bonds from isolated O_i. In this second structure the 'inner' O of the pair can become trivalent, allowing a reconstructed Si-Si bond to form similar to the di-y-lid structure (see Figure 9.14). A third possibility is an oxygen square structure with one flanking O_i, and in practise the Snyder-Stavola model relaxes towards this with a tiny saving in energy.

9.4.3.1 Snyder-Stavola model

The relaxed structure is shown in Figure 9.13 with eigenvalues in Figure 9.10a. When unconstrained, the central Si atom moves ~0.05 Å off the C_2 axis to form a structure somewhere between the Snyder-Stavola and an oxygen square with a flanking O_i, saving 0.063 eV in the process. This is consistent with calculations by Chadi who found an energy saving of 0.07 eV for the same process[211]. This is presumably the diffusion barrier for reorientation of the defect between the two off-site configurations, and hence it would appear time averaged C_{2v}.

The vibrational modes of the Snyder-Stavola structure are given in Table 9.5. The high frequency mode is due to out-of-plane wag of the core trivalent oxygen atom, and seems remarkably high for such a motion. The modes are not in good agreement with either TD1/TD2, which exhibit two modes each at 975/988 and 716/724 cm^{-1}. The only possibility for correlation is if the calculated modes at 978.9 and 771.6 cm^{-1} are those that have been observed, and there are additional experimental modes in the 850-900 cm^{-1} range which have not yet been observed. The TD2 988 cm^{-1} mode drops by 43 cm^{-1} when ^{16}O is replaced with ^{18}O, which is not very good agreement with the 978.9 cm^{-1} mode shift of 46.9 cm^{-1}.

We therefore conclude on the basis of the vibrational modes that the Snyder-Stavola model is unlikely to be TD2, but may be TD1 if there are further vibrational modes which have not yet been detected by experiment.

Local Vibrational Modes (cm^{-1}) ^{16}O	^{17}O	^{18}O	Dipole moment squared for ^{16} O
1230.2	26.2	49.8	0.421
1038.5	25.2	48.2	0.031
978.9	24.6	46.9	0.140
946.2	26.0	49.6	0.005
899.2	20.6	39.3	0.120
884.6	22.3	42.7	0.143
771.6	17.4	32.8	0.190
755.6	17.8	33.8	0.117
697.5	13.4	24.5	0.034
630.8	1.8	3.8	0.005
601.2	2.9	6.4	0.014
560.7	0.9	1.9	0.001
538.7	0.3	0.6	0.020
527.6	0.2	0.5	0.001

Table 9.5: Vibrational modes of the Snyder-Stavola 3O TD model (cm^{-1}). Last column gives the associated absorption intensity of each mode (dipole moment squared).

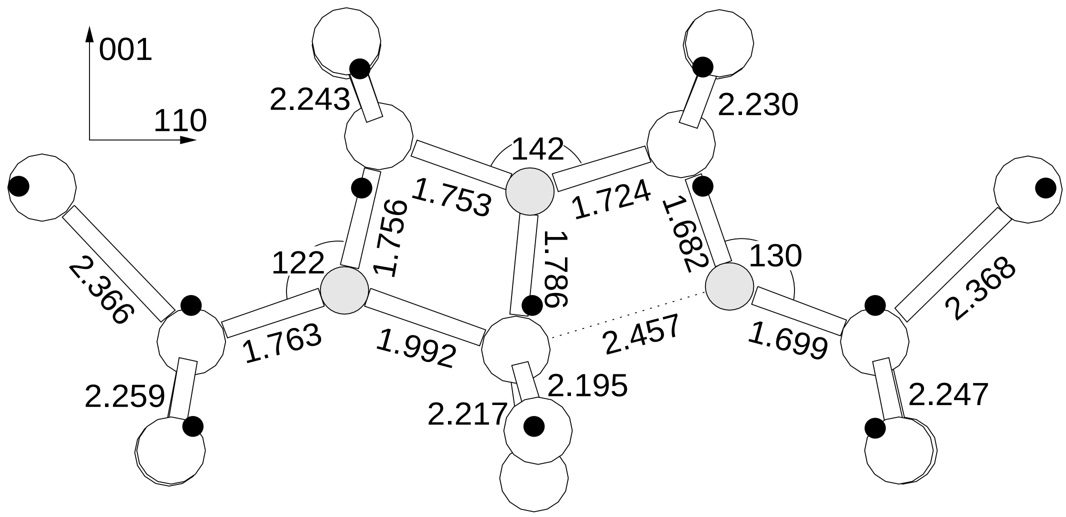


Figure 9.13: The core of the Snyder/Stavola 3O thermal donor model (lengths in Å). Black dots mark ideal Si lattice sites, showing the [001] compression and [110] lattice tension. The deviation from C_{2v} symmetry saves 0.063 eV over the relaxed C_{2v} symmetric structure.

9.4.3.2 Partially dissociated 3O model

An alternative 3O model is the structure of the di-y-lid, with one of the outer O_i atoms removed. According to our calculations this is 1.789 eV less stable than the Snyder-Stavola model. The structure is shown in Figure 9.14, with eigenvalues in Figure 9.10b.

As can be seen, one of the core y-lid atoms drops back into its bond centre, leading to a separated defect core. However, the reconstructed Si-Si bond remains. Symmetry excludes this structure from being any of the higher order TDs. The local vibrational modes are given in Table 9.6. At first site these do not appear to bear any relation to the experimentally observed modes, and hence this would exclude it as a model for any TD. However the top 1123.2 cm^{-1} mode is largely O_i stretch of the isolated atom. 1058.8 cm^{-1} is primarily bond stretch of the trivalent O, while the 912.4 cm^{-1} is bond stretch of its neighbouring O_i. 790.4 cm^{-1} is the out-of-plane wag of the trivalent oxygen.

Thus the top mode should lie roughly over the O_i absorption, which leaves stronger modes at 1058.8 cm^{-1} and 790.4 cm^{-1}. There are various reasons for considering this as a model for TD1. Apart from the isolated O_i mode, the nature of these modes are very similar to those of the di-y-lid, except that the out-of-plane wag only occurs with a single O_i instead of a pair; the low frequency absorption for TD1 is exceptionally weak and took a long time to find[239], and this would explain the lower signal strength. These calculated modes are too high for TD1 (they should be lower than those of TD2 or TD3). The calculated donor state lies deeper in the gap than any of the other TD models such as the di-y-lid, and the donor level for TD1 is known to lie deeper than the others[203].

TD1 forms quickly, and then gradually transforms to TD2; this would be consistent with a scheme where, for example, this defect trapped O_i. Alternatively, if it is not very stable (the dimer is quite separated from the isolated O_i) it could be a transient species, and eventually instead of dimers meeting O_i to form this structure they pair up to form a di-y-lid TD2. Such binding could be a Fermi level effect. As a dimer migrates towards the O_i, one of its oxygen atoms passes through a y-lid as a saddle point. If at this point it loses its two electrons it becomes stable in this structure, forming this partially dissociated TD1 structure. At a later point when the higher order TDs are forming (TD2 onwards), these have shallower defect levels than TD1, and so could depopulate into the TD1 level. This would instantly destabilise the defect and allow it to break up into a neutral separated dimer and O_i. Hence it would be a rapidly forming defect that would quickly be lost as soon as defect complexes were formed which possessed a shallower level.

Local Vibrational Modes (cm^{-1})			Dipole moment squared
^{16}O	^{17}O	^{18}O	for ^{16} O
1123.2	25.5	48.6	0.178
1058.8	25.4	48.4	0.185
912.4	20.9	39.7	0.163
790.4	15.6	29.6	0.366
722.3	10.9	20.5	0.145
698.4	11.6	21.8	0.198
680.4	9.9	18.6	0.129
617.4	2.4	4.6	0.004

Table 9.6: Vibrational modes of a 'partially dissociated' 3O TD model (cm^{-1}). Last column gives the associated absorption intensity of each mode (dipole moment squared).

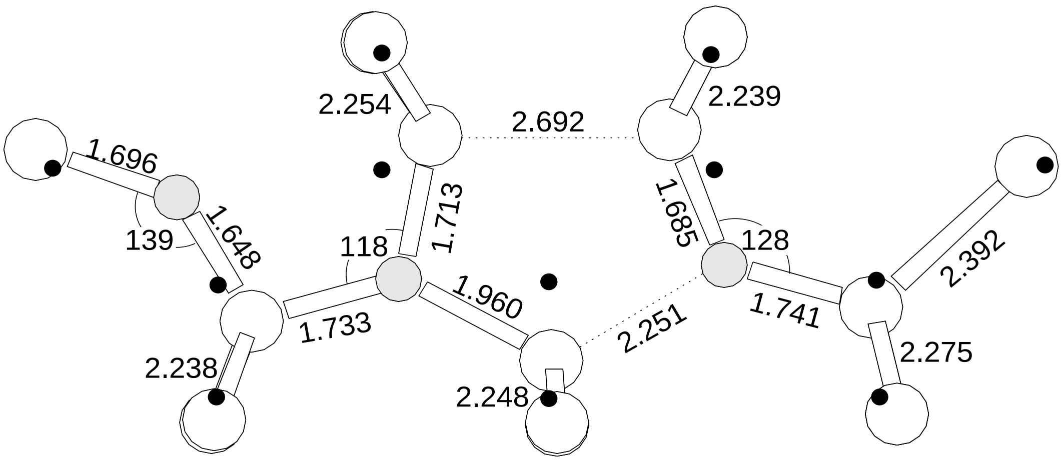


Figure 9.14: An alternative 3O thermal donor model similar to the di-y-lid but missing an end O_i; the trivalent O in the pair is stabilised by the presence of the neighbouring O_i.

Local Vibrational Modes (cm^{-1}) ^{16}O	^{17}O	^{18}O	Dipole moment squared for ^{16} O
1127.0	23.9	45.8	0.093
1062.1	18.4	34.9	**0.308**
973.4	18.9	34.2	0.022
945.5	17.6	35.2	**0.219**
933.8	18.7	35.7	**0.271**
917.3	18.3	35.2	**0.143**
849.2	17.6	33.6	0.035
827.7	14.4	27.7	0.019
652.1	2.5	4.9	
643.1	9.1	17.4	
630.3	0.2	0.4	
591.0	1.3	2.8	

Table 9.7: Vibrational modes (cm^{-1}) of the NL10(Al) proposed structure, a 'di-y-lid' with centrally substituted Al. Also included is the dipole moment squared for the ^{16}O higher modes, which is proportional to absorption intensity.

9.4.4 Six oxygen atom TD species

It is possible to insert an additional dimer into the di-y-lid structure, with one oxygen atom in each of the Si-Si bonds below the core Si atom. This pushes the core Si atom upwards into the defect until the trivalent oxygen atoms are closer to this core atom than their outer Si neighbours. We relaxed such a defect in the 148 atom cluster used above, but the defect is then really too large for such a cluster and more bulk Si material is needed to accurately simulate the surrounding lattice relaxation, which is extensive. However the defect core maintains its shallow double donor nature (see Figure 9.10f) and the structure does not spontaneously relax into any other; thus it remains a possible candidate for the next in the TD series.

9.5 Al in the di-y-lid core - NL10(Al)?

Several of these thermal donor structures feature a core Si atom surrounded by oxygen atoms, notably the di-y-lid model. In this structure the core Si would be a prime candidate for substitution by Al. For full background description of the NL10(Al) defect, see Chapter 8.

When we relaxed this structure we found it to be stable in a single positive charge state, where it possessed a (now depopulated) level very close to the conduction band, in agreement with experiment. It also has a filled singlet quite close to the valence band (see Figure 9.15).

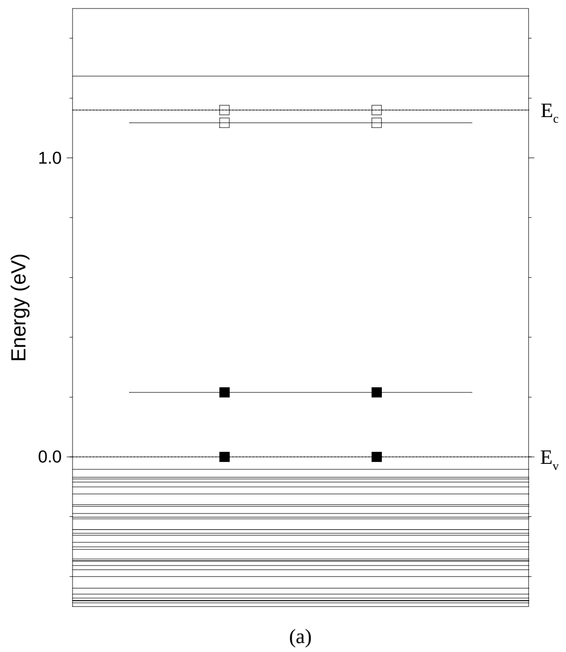


Figure 9.15: The top Kohn-Sham eigenvalues for the NL10(Al) Al/di-y-lid model, Al_sO_{4i}. Defect is in the +1 charge state. Black boxes denote filled state, white boxes for empty states. The eigenvalues have been arbitrarily scaled to the experimental band gap of 1.16eV.

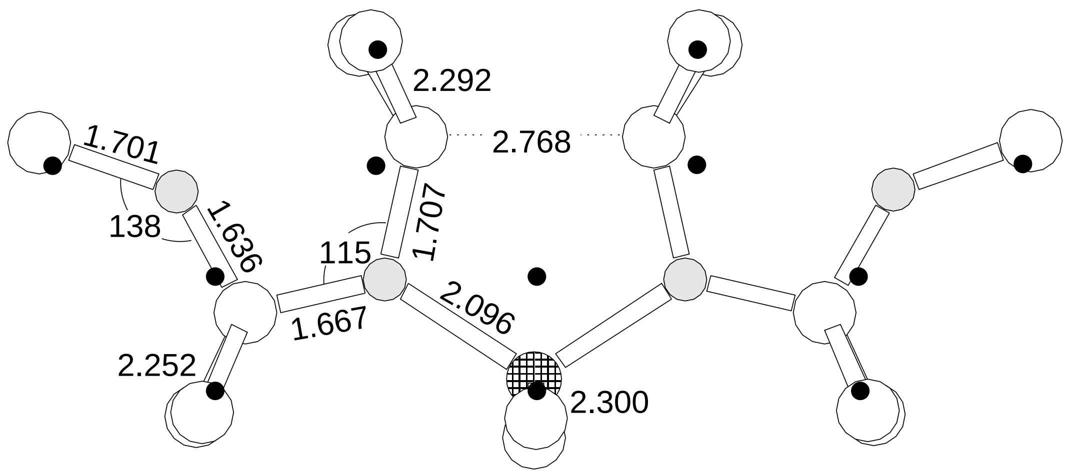


Figure 9.16: The proposed NL10(Al) structure, consisting of the di-y-lid structure with the core Si atom replaced by Al (cross-hatched atom). All bond lengths in Å, black dots mark ideal Si lattice sites.

It possesses the correct symmetry of approximately C_{2v} (there are small out of plane deviations similar to the di-y-lid discussed above). In addition it has a single Al atom on the C_2 axis and no oxygen on this axis, in agreement with ENDOR studies. The relaxed structure is shown in Figure 9.16. There are several variants on the NL10(Al) structure observed [10] which are explainable in terms of Al subsituting for different lattice Si atoms; in the di-y-lid structure there are other Si atoms neighbouring more than one oxygen atom, all of which would be sites suitable for Al attack. Isotopic ENDOR studies led to the conclusion that the local structures around oxygen must be identical for both NL8 and NL10 [10], which is the case in these di-y-lid based models.

Unfortunately there is no unambiguous vibrational data to compare to, so we cannot use this to help strengthen the assignment[265]. However an absorption at 992 cm^{-1} has been observed in Al-doped Cz-Si which has an integrated absorption intensity linearly proportional to the Al content [239]. This lies mid-way between the absorptions of TD2 and TD3 (988, 999 cm^{-1} respectively), suggesting that this could be a perturbed form of these early thermal donors. If the Al_sO_{4i} complex described here forms from dimer aggregation at Al_s, then we might expect to first observe the formation of Al_sO_{2i}, which could be the source of this absorption. The calculated vibrational modes of Al_sO_{4i} are given in Table 9.7 along with the square of their dipole moments.

Substitution for the core Si in the 5O model is also stable, although the structure exhibits slightly more distortion. However this cannot be a model for NL10(Al) since it has O on the C_2 axis, in contravention of ENDOR observation[10].

Thus the di-y-lid with Al substituted for the core Si seems a sensible candidate for the NL10(Al), which fits all of the available ENDOR structural restrictions, and possesses a single shallow gap state at around the right value.

9.6 Y-lids – Tri-valent oxygen

It seems that for oxygen to move to a trivalent position in Si there are two constraints:

1. Firstly there has to be some compression from at least one side, increasing the strain on a 'BC' O_i atom and bringing it closer to a potential third neighbour.

2. Secondly there has to be a strong quadrupole attraction that binds it tightly to one of its Si neighbours. In this work that has been induced by the next neighbour presence of another highly electronegative element, in these cases always O. This pulls the O atom close to this shared neighbour, but crucially

weakens the second O-Si bond in turn. This allows the O sufficient freedom to move out of its divalent site and form a weak third bond.

Thus these results do not show trivalent oxygen to form three uniform bonds; instead it seems to form one strong bond, one intermediate, and a third much weaker bond. This is contrary to previous CNDO/S calculations which show three equivilent bonds [105]. The exception to this is tri-valent oxygen atoms on the C_2 axis of defects, such as the core of the 5O thermal donor and to a lesser extent the Snyder-Stavola structure. These have bond lengths that are more uniform, particularly the 5O species. This is probably because these atoms have two next neighbour O atoms which roughly equalise the Coulombic attraction along both top bonds.

9.7 The role of silicon self-interstitials

In later oxygen precipitation, quartz precipitates are formed with an associated ejection of Si_i. This has led to a general assumption that there are many Si_i present in Cz-Si, and from this to the idea that they may lie in the core of the thermal donor. This has been supported by the Deák model which is based around Si_i. This is further supported by the fact that carbon suppresses thermal donors, which has led to another common assumption that this is occurring through Si_i absorption by the reaction

$$\mathrm{C_s} + \mathrm{Si_i} \rightarrow \mathrm{C_i}$$

This has led to many unusual postulates to explain experimental data, for example that there are two types of dimer complex, 'normal' and 'O_2^*' which can eject Si_i atoms[266]. However these assumptions are flawed. It is known that C_iO_n defects (the 'P-line' complexes) form in the first 15 minutes of annealing[197] at 450°C, showing the presence of either C_i or Si_i to produce C_i via the mechanism given above. However these are at extremely low concentrations, and indeed, act as evidence that Si_i concentrations are low as a consequence.

In addition, as described above it is not possible as far as we are aware to construct a single Si_i structure that does not possess gap states, and so thermal donor models based around a core Si_i are unable to explain the electrically inactive forms of the TDs. Carbon suppression of thermal donors can occur through other mechanisms, primarily C_s absorption of O_i and O_{2i}[267], and possibly also a self-catalysed C_i suppression mechanism (described briefly in Section 9.3.1).

Long time annealing FTIR studies[239] show absorption at 999 cm^{-1} associated with the thermal donors. However after long times this disappears and is replaced with broad absorptions at 1012, 1006, and finally 1015 cm^{-1}, believed to be associated with the shallow thermal donors. This would be consistent with a picture whereby over long times the thermal donors transform into quartz, releasing large quantities of Si_i. These are then able to initially transform all available C_s into C_i, leading to a glut of C_i-based STD formation (CHO_4 complexes). Thereafter the remaining Si_i combine to form $\langle 311 \rangle$ rod-like precipitates.

This therefore suggests that Si_i are not present in large quantities in the early stages of Cz-Si annealing but are formed later when quartz precipitation occurs. In addition they are not involved in the structure of thermal donors[268].

9.8 Earlier Thermal Donors - O_i to TD3

Unfortunately these results are not sufficient alone to determine whether the 4 O_i di-y-lid is TD2 or TD3. There are then at least two possible schemes that could describe the early thermal donor family consistent with these results. The following section details these, and attempts to select one using further experimental data [3]. We first present a scheme where the di-y-lid is TD2 and show that this contradicts experimental evidence. We then propose a scheme where the di-y-lid is TD3 and show that this may agree with the experimental data.

In this discussion a few things are assumed. Firstly that the trimer, O_{3i} is responsible for the 1006 cm^{-1} absorption line, for the reasons discussed in Chapter 6. Secondly that the experimental data concerning the inactive form of TD1 and TD2 is correct [239], and these defects are isomers of one another. Finally we make use of the fact that TD3 is the primary thermal donor species observed in annealing studies. As discussed in Chapter 6, kinetic analysis by Markevich *et al* suggests the trimer is mobile [269].

9.8.1 Di-y-lid is TD2 - a serial process

The first of the alternatives is that the di-y-lid is TD2. The formation energy of TD1 is initially 1.2 eV[270] and later reverts to 1.7 eV. In this case the 1.2 eV would be the trimer migration barrier, slightly lower than that of the dimer. Thus the 1.2 eV represents an initial pool of trimers present in the as-grown material that migrate to form TD1; once these are exhausted the rate determining step becomes trimer formation and the barrier for this is 1.7 eV, *i.e.* the dimer migration barrier.

The trimer then migrates until it reaches O_i, at which point the strain energy is sufficient to force it to restructure, giving either TD1 or TD2. The structure for

[3] Be warned that these arguments necessarily become quite involved, and a stiff drink may be required!

TD1 is unclear, but may involve either a single y-lid or an oxygen 'square' structure such as the 4O flanked square structure discussed above. TD2 is the di-y-lid. Now the di-y-lid and TD1 can restructure into their isomeric electrically inactive form, which is probably either two dimers separated by a bond centre, or a linear chain of four O_i atoms.

We then need to explain why TD1/2 gradually change into TD3, which becomes dominant. There are two possibilities here. The first is that TD1/2 are mobile (possibly only one of the two species) and encounter O_i, implying TD3 is a 5O species. However it seems impossible to design a 5O structure that has the observed C_{2v} symmetry without putting O on the C_2 axis, which is in breach of ENDOR results.

The second possibility is that TD1/2 are immobile, and so TD3 forms through dimer addition to TD1/2. We next need to explain the predominance of TD3. If TD1/2 are formed through rapidly migrating trimers encountering O_i, this should happen very quickly, since the O_i concentration is so high. However if TD3 relies on TD1/2 trapping a dimer, this requires the meeting of two much lower concentration species. The dimer is probably also not migrating as fast as the trimer. Hence we would expect TD1/2 to form much more rapidly than TD3, and there to be no difference in the formation rate of TD3 and higher order TDs. We would still expect TD3 to be more predominant than any other higher order TD species since it is the first of these (there should be a tailed distribution of these depending on the binding energy of dimers to each species). This is a flaw in this model since it appears to predict more TD1/2 than TD3.

A possible way around this would be that TD1/2 were unstable and regularly broke up into two dimers. However the kinetic analysis of Newman *et al* [143] suggests that dimers form through O_i migration. If TD1/2 broke up into dimers, under this scheme, they would encounter O_i to form trimers, which would encounter O_i to form TD1/2 once again, and so dimer formation would be a self-catalysed reaction that did not require O_i diffusion, in contradiction with their results.

An alternative is that TD1/2 are mobile, but for some reason on meeting O_i do not form a thermal donor species. This seems unlikely, since no other pure O species are observed apart from quartz precipitates after long time anneals. It seems unlikely that O_i would not bind to TD1/2 but a dimer would. It cannot be that these 5O species break up into a trimer and a dimer, for the kinetic arguments given above. Another more unusual model that appears consistent is to postulate that as well as TD1/2 trapping dimers to form TD3, there is a parallel formation stage as they migrate (possibly very slowly) to O_i, and form $5O_i$ species which are the first step towards quartz precipitates.

Annealing studies at 450°C (see Figure 6.11b) show that TD2 forms in the first few hours, but then tails off with longer anneal time until it practically disappears after 10-15 hours. Thereafter there is only trace TD2 signal, however the TD3 signal continues to increase in line with the trimers; *i.e.* there is no observed correlation between the TD1/2 trace and TD3. This loss of TD1/TD2 signal with annealing time presents a problem to any scheme which requires a serial transformation from dimer $\rightarrow$ trimer $\rightarrow$ TD1/2 $\rightarrow$ TD3, since there should always be some sort of correlation. It eliminates the model proposed here in all the forms discussed above, as well as any scheme that requires a serial progression through TD1/2 to higher order TDs.

The only possible way to salvage this model is if TD1/2 are becoming unstable for some reason (e.g. Fermi level movement, see below), and they are by-passed in the formation process thereafter, e.g. through dimer/trimer aggregation. However this seems unlikely given the low concentration of dimer and trimer centres.

9.8.2 Di-y-lid is TD3 - a parallel process

A second possibility is that the di-y-lid is actually TD3, and both TD1 and TD2 are pre-cursors to the di-y-lid. Therefore, being bistable isomers, TD1 and TD2 have to be different forms of electrically active trimer.

In this case, the initial 1.2 eV formation energy of TD1 could be explained as a barrier to internal restructure of the trimer. If the dimer is able to migrate to O_i and form a weakly bound trimer, in the initial material there is already a small pool of trimers (there are also some early thermal donors present). These trimers can restructure with a barrier of 1.2 eV. Once these are exhausted, the limiting barrier becomes the trimer formation barrier, *i.e.* the dimer migration barrier of 1.7 eV. In addition both TD1 and TD2 have an electrically inactive form, with LVMs different to those of the trimer. Thus there has to be a third form for the trimer, although this form may be much higher in energy than the others and therefore only frozen in at lower temperatures.

In such a scheme, presumably TD1 and TD2 could remain in dynamic equilibrium with the inactive trimer form. There are two ways of forming the $4O_i$ di-y-lid. Either the trimers could migrate, or else they could be only weakly stable and so break into dimers and O_i, so that dimers would eventually meet each other and bind into the di-y-lid. Markevich found his concentration models for the dimer and trimer fitted experiment much better if he assumed a mobile trimer, and such a mechanism would lead to faster di-y-lid formation given the relatively high concentration of O_i compared to O_{2i}.

This also helps to account for the relative TD concentrations. TD3 is the primary

TD. If both TD1 and TD2 are mobile, they can form through dimer addition to O_i. TD3 can then form by trimer addition to O_i. However, if TD3 is immobile, it cannot migrate to O_i. O_i is only moving very slowly on the timescales considered, and thus in order to form higher order TDs, TD3 has to rely on dimers migrating to it; therefore the formation process will slow considerably. Hence there will tend to be a build up of TD3 species.

Against this argument is the necessity that there has to be dimers present in the material to bond with the higher order TDs. However if dimers are reacting with O_i to form trimers, that should rapidly take most dimers out of circulation. This could still be consistent with the above model if the trimer binding energy is so small that the dimers only bond briefly, and most trimers dissociate before forming any of the thermal donors.

The crux of this model then seems to be that TD1 and TD2 form in *parallel* with the main formation scheme, which is dimers to trimers, and on to the di-y-lid as TD3. However in this form it is not consistent with annealing studies (see Figure 6.11). In the 450°C plot, it can be seen that the TD3 absorption exactly maps that of the trimer. This is consistent with a scheme where the trimer migrates to O_i to form a di-y-lid TD3. However, the TD2 signal rises rapidly in line with the dimer concentration to begin with, and then tails off after two hours, until it has almost disappeared after 10 hours or so. If TD1/2 were isomers of the trimer, we might expect the change in their absorption to roughly match that of the trimer, which is not the case.

However this might be understood if the conditions in the material change to make the switch from trimer to TD1/2 less favourable. As the anneal progresses increasing numbers of TDs are produced, which should lead to a shift in the Fermi level. This could shift the relative stability of the trimer and TD1/2, making TD1/2 less favourable. In the first 3 hours of annealing at 450°C the trimer peak appears to remain roughly constant, whereas the TD2 peak roughly traces that of the dimer. Thereafter the trimer peak starts to increase while TD2 drops away (see Figure 6.11). We would expect the combined signal from the trimer, TD1 and TD2 to roughly correlate with the dimer; in Figure 6.11b, this appears qualitatively to be the case.

This idea is supported by results from samples subjected to dispersal treatment [271], where during subsequent annealing the decrease in the 1012 cm^{-1} band correlates with the growth of the 1006 and 975 cm^{-1} bands. From these results it appears that TD1 forms simultaneously with the trimers, but as soon as the donor concentration starts to creep up, the TD1 transform into TD2. From this we might tentatively suggest that TD1 is closer in structure to the trimer, and is slightly less stable than TD2.

9.8.3 General discussion of early stages of TD formation

There are also other possibilities beyond those considered above. One of these is that the dimer can structurally transform into TD1/2. This is unlikely since it would require four distinct structures for an oxygen pair; the dimer, TD1 and TD2, and the electrically inactive form of TD1/2. In addition, such a model also suffers from the lack of direct correlation between [TD1/2] and the 1012 cm^{-1} dimer absorption. We have not been able to determine any electrically active double donor structures for the dimer (although structures such as a di-y-lid core could theoretically be produced by only two O_i atoms, the resultant lattice strain would be huge).

It has also been suggested that trimers do not form [143], but the models discussed here all assume they are at least marginally stable. If trimers do not form, it is difficult to explain the 1006 cm^{-1} LVM. However rather than being a stable defect complex, it is possible the trimer is simple the result of a dimer pushing O_i along for a few steps before moving on separately again.

There are many other ways these arguments may break down, not least that some of the experimental evidence may be misleading; for example the isomeric nature of TD1 and TD2 could actually just be due to very similar inactive forms that vibrate at the same frequencies.

It has been stated at above that it is impossible to generate a C_{2v} structure with an odd number of O_i atoms without putting an oxygen on the C_2 axis. To further complicate matters it may be possible to have a structure such as the Snyder-Stavola model where a single oxygen atom lies off the C_2 axis, but has a tiny barrier to the symmetrically mirrored structure with the oxygen on the other side of the axis. In this case the structure would in practise flip rapidly between the two structures and would obtain time averaged C_{2v} symmetry, have an odd number of O_i, and yet might not show O_i on the actual C_2 axis.

Unfortunately without the ability to accurately compare energies between these different structures there is little more that theory can do for the thermal donor problem at this point, beyond proposing a variety of models, and eliminating those that it can on the basis of the available experimental evidence. The variability in oxygen energies does not appear to be just a problem with AIMPRO but of many DFT codes; Chadi finds that addition of O_i to OV is endothermic in contradiction with experiment [95], and work by Pantelides *et al* on the O_i diffusion barrier has varied between 1.8 eV and 2.5 eV in recent papers[115, 130].

This work proposes many models that seem to be suitable as either TD1, TD2 or TD3. The di-y-lid structure fits TD3 very well in terms of structure, electronic properties, and vibrational modes. Other structures we cannot eliminate include the Snyder-Stavola, a new 4O 'flanked square' structure, and even the partially dissociated trimer structure given above. However all of these three would appear to be models for only TD1 and possibly TD2 for the Snyder-Stavola, on symmetry grounds. It is definitely the case that there are a variety of possible models with shallow double donor character that only require trivalent O_i, BC O_i and lattice Si, and do not need to invoke Si self-interstitials.

9.8.4 Early Stages - Summary

Given that the di-y-lid is either TD2 or TD3, there only appears to be one mechanism consistent with the experimental and theoretical data. O_i atoms diffuse together normally, to form dimers. These are faster diffusing, with a barrier to diffusion of 1.7 eV. They migrate, and normally encounter a third O_i. In this case they may form a trimer, but if they do so it is only briefly stable. The trimer itself is also capable of diffusion, and so can move to another O_i. At this stage the 4O_i atoms experience sufficient lattice strain to switch them into the di-y-lid structure, and are responsible for TD3. This is not mobile, and so can only grow through further dimer addition (O_i migration is assumed to be too slow for growth purposes). Hence TD formation slows down at this stage, and TD3, the di-y-lid, becomes the primary thermal donor.

In *parallel* with this process, the trimer is able to restructure into either TD1 or TD2 with a barrier of 1.2 eV. These two defects are also able to restructure into their electrically inactive shared isomer, which is different from the standard trimer structure, and probably not stable at room temperature. The low stabilities of TD1/2 are Fermi level dependant, so as the Fermi level rises with the addition of more thermal donors, first the TD1 species tend to transform to TD2, and eventually these are also unstable and transform back to trimers to rejoin the main thermal donor formation path[4].

We now need electrically active trimer structures for TD1 and TD2. The Snyder-Stavola 3O model is a candidate, assuming there are a few modes which are weaker and not yet observed by experiment (this was the case with the oxygen dimer, where theory predicted modes which were not initially observed by experiment).

The results of initial dispersion treatments and isothermal anneals [239] suggest that TD1 forms in the initial anneal stages along with the trimers, but rapidly converts to either trimers or TD2. TD2 is stable for a longer time. This suggests that there is only a small transformation barrier between the trimer and TD1, with

[4]Using an analogy of Bobs, TD1 and TD2 are 'neanderthal species' that form in parallel to the evolution of trimers to TD3, and die out during the annealing.

a larger barrier to TD2; however once formed, TD2 is more stable. There are presumably several schemes that could fit this information. One possibility is if the trimer is stable in the 'Manx' form, as our preliminary results suggest. This can then become a +2 TD1 structure by pushing one of the oxygens into a y-lid; in this case TD1 is a y-lid with O_i atoms in both back bonds of the core Si. This requires little structural rearrangement. Later the O_i atoms are able to migrate around to lie in the same plane as the y-lid, forming the Snyder-Stavola structure as TD2.

Watkins has shown that the reorientation time for TD1 is roughly half that of TD2 [236]. This is consistent with this model, since the TD1 model proposed here can reorient by the y-lid O atom dropping back into a bond centre, and one of the other bond centred O_i atoms moving into a y-lid configuration. However, for the Snyder-Stavola TD2 structure to reorient requires several hops for the O_i atoms at the ends of the defect, and hence would be slower. The TD1 model here would only be a double donor if both the trivalent oxygen and the core trivalent Si atom both lost an electron. There is no obvious electrostatic compression mechanism exerting on the empty p-type orbital of the core Si, so it is not obvious that such a +2 donor state would be shallow. Experimentally however TD1 has a much deeper gap state than TD2. As TD1 and TD2 are isomeric, it is notable that during the reorientation experiments, transformation into the thermodynamically more stable phase is not observed, *i.e.* the TD1 species do not all transform into TD2. Presumably this must be because the reorientation does not bring the atoms into a position where they can hop directly into TD2 sites; the model described here would fulfill that criterion.

9.9 Later Thermal Donors - TD3 and beyond

The problem of later thermal donors is a very different one. Here, experiment suggests a common core that undergoes small perturbation with increasing oxygen content. The standard model is one of oxygen aggregation along the $\langle 110 \rangle$ plane, resulting in a defect with a fixed core and two tails of O_i stretching out along $\langle 110 \rangle$.

There have been a total of 16 different thermal donors observed[241]. In this discussion I will assume TD3 is the di-y-lid and hence has four oxygen atoms in it. Additionally I will assume that this lies at the core of all following thermal donors; these assumptions do not drastically alter any of the models should they later be revised. Analysis of oxygen loss from solution by Newman *et al* [143] suggests an average loss of ten oxygen atoms from solution per thermal donor, at 500°C.

There are a number of possible formation models, but these can be roughly divided depending on whether they assume a serial development of thermal donors or allow isomers. The argument also varies depending on whether the TDs grow through O_i addition or dimer addition, which is currently not clear [272]. However standard O_i migration alone cannot account for the quantity of higher order TDs formed [272], and TD formation from TD2 to TD6 has been shown to have a formation barrier of 1.7 eV [273], suggesting that dimers are the dominant active species rather than O_i.

9.9.1 1D thermal donors

The first model is the linear tail model described above, the currently accepted 'standard model'. In this case, O_i atoms aggregate in the same $\langle 110 \rangle$ plane as the defect core, in neighbouring BC sites. To form TD16 we require 14 new thermal donors after TD3. If each of these grows through a dimer adding to the previous defect that implies $13 \times 2 + 4 = 30$ oxygen atoms in TD16. If the scheme is slightly modified and the TDs grow through O_i addition then this number can be reduced to 17. Thus there is a chain of either 7 or 13 O_i atoms on each side of the defect core.

This model seems very unphysical. O_i strains its surrounding lattice, causing 30% Si-Si bond dilation. When paired to form O_{2i}, there is once again a large strain field produced both above (001) and in front (110) of the defect. Two dimer structures together, the structure of the di-y-lid, seem sufficiently strained to force the core oxygen atoms into tri-valent positions. A model where O_i atoms collect along $\langle 110 \rangle$ next to the defect core is reasonable for the first few O_i atoms, since TD3 is highly tensile along this direction. But it would be unable to equalise the strain field of 7 to 13 O_i atoms on each side. Although the strain field in such a chain might be high enough to force more of these chain O_i atoms into y-lid sites, these would be electrically active, and the Coulombic repulsion with the TD core would be too high.

One way in which a chain model can achieve these 13 different structures is if the chains on either side of the core can be of different lengths, and thus isomers are allowed. In this case, structures such as O_i-TD3-O_{3i} would be allowed. Table 9.8 shows that a minimum of 6 total 'tail' O_i atoms are required for this to explain all the TDs, *i.e.* a maximum tail length of $5O_i$ on each side.

The higher order annealing studies do not allow differentiation between an isomer and a serial development model. TDs 12 to 16 were seen simultaneously in a sample annealed for 3 hours at 470°C [241], and so there is no information on their relative development; they are named in sequence of decreasing ground state binding energy.

Finally, extremely long tail models may not be consistent with reorientation

Number of 'tail' O_i	Number of TD isomers	Cumulative number of TDs
1	1	1
2	2	3
3	2	5
4	3	8
5	3	11
6	4	15

Table 9.8: Number of thermal donor structures possible with a fixed common core and O_i adding in two linear tails, one either side of the core. The defect is assumed symmetric. The table lists number of isomeric combinations possible for a given tail length, and cumulative total. There are 13 experimentally observed TDs after TD3.

data[236]. For TD3 ~5 O_i hops are required for reorientation; this would be possible with the di-y-lid model. For higher order TDs there is a small increase in the reorientation time prefactor, corresponding to more hops. However for a long tail thermal donor to reorient through each atom hopping individually, the number of reorientation steps would rapidly increase, as the line has to sweep out a circle to realign itself. One possibly way around this is if the atoms hop towards the core, through the core, and back out along the new defect line; in this case the number of hops required would equal the number of O_i atoms in each 'arm' of the defect plus some constant for the core to reorient (presumably, from the TD3 result, five hops for the core).

9.9.2 2/3D thermal donors

An alternative model for higher order thermal donors is a more general aggregation of oxygen to a common TD core. In this case dimers are attracted to the core but bond into a variety of bond centred sites around the core including, but not limited to, those along the $\langle 110 \rangle$ defect plane. Although the defect core causes compression along $\langle 100 \rangle$, there are Si-Si bond sites at angles to this which are dilated due to at least one of the Si atoms being pushed off-site by the strain field of the TD core, and these sites would provide locations for dimer attack. When discussing the mechanism in this general way it is impossible to predict whether the process would be serial, or whether several sites would have equivilent energies, leading to isomeric thermal donors. In a similar way to the calculations shown in Table 9.8, if the TD has six sites surrounding it, four of which lie off the C_{2v} mirror planes and are symmetrically equivilent, and the other two lie on a mirror plane and are a symmetric pair, there are 18 different permutations obtainable by filling between 0 and 4 of these sites. Hence only four dimers are required to aggregate to the TD core to produce the complete set of observed TDs, if TD isomers are allowed.

However, it appears that any perturbations of the thermal donor core are only extremely small, and directly bonding dimers to the sides of the core might be expected to perturb it to a larger degree. In addition, ENDOR results suggest all oxygen lies along the $\langle 110 \rangle$ plane, and O_i aggregates bonded to the defect core would be expected to provide an ENDOR signal.

An alternative model, and one which seems consistent with experimental data, is one whereby such sites exist much further from the defect core. In this case they are regions of the lattice arranged around the core which exhibit some Si-Si bond dilation, and hence provide weak binding sites for dimers. These adopt these sites, forming a 'cloud' of dimers around the core like flies. Thus they would only be expected to weakly perturb the core structure and states. Such sites can be seen in our cluster results; although the bonds directly along $\langle 001 \rangle$ are compressed, those to one side of these show some dilation as the Si atoms directly above the defect core are pushed upwards and away from their neighbours further off this plane.

The thermal donors would be expected to show extremely similar behaviour under this picture with only slight variations from one another. For example, in the reorientation experiments discussed above the core would still rotate as before, with the dimers just shifting position to the new slightly tensile sites. This would be consistent with the ENDOR result since the O atoms would only be weakly bound and not very close to the core, and hence would not show up as part of the core structure. Hence although the total defect could display low symmetry, ENDOR and EPR would only measure the C_{2v} symmetry of the defect core. Oxygen atoms in neighbouring dilated sites may be responsible for the broad 1060 cm^{-1} peak observed after anneals, the broadness of the peak explained by the slight variety in local environment of each of the atoms.

If trimers can occupy these sites it provides an additional variable when determining the number of dimers/trimers needed to give the required number of structural permutations to explain all of the observed TDs, *i.e.* less dimer sites would be required to explain all of the TDs. For an average of 10 O atoms per thermal donor, this requires (with a core of 4 atoms) an average of either two trimers or three dimers per thermal donor, which is a very modest figure.

In this picture then, there is a dynamic equilibrium concentration of dimers, and eventually trimers, adopting sites around the TDs. The position of the donor level in the gap should gradually increase as the surrounding lattice strain goes up with dimer addition. There is essentially no perturbation of the core and instead an increasing strain of the surrounding lattice, which compresses the $\langle 001 \rangle$ conduction band valleys used to construct the EMT-like donor state, making it increasingly shallow. If this was not the case we would expect a spectrum of vibrational absorption peaks from TDs with slightly perturbed cores, whereas in long time anneals there are only four peaks in the 999-1020 cm^{-1} range associated with electrical activity [239]: the 999-TD3 band, the 1006 and 1012 cm^{-1} electrically active bands, and a later 1015 cm^{-1} band, with the latter three thought to be associated with shallow thermal donors.

A further key piece of evidence for such a picture is that the figure of on average 10 atoms lost from solution per TD appears to be independant of annealing temperature or time [274]. This suggests that TDs form rapidly, and then higher order TDs become either more unstable or unable to trap further atoms. This would be consistent with a result where around 10-14 O_i atoms could bind to the TD with cumulative loss of energy, but thereafter there was an energy increase[274]. This is consistent with this 'cloud' picture, where further dimers have to attach further from the defect core and are increasingly weakly bound, until the higher order TDs are in a thermal equilibrium with dimers sticking and releasing at regular intervals.

Such a picture would also be consistent with a revised dimer model where O_i atoms sat on opposite sides of the hexagonal ring site.

9.9.3 Later TDs - Summary and Conclusion

We propose that a scheme consistent with the available experimental evidence for later thermal donors is one whereby all the later TDs contain a common thermal donor core which is not modified in the later species. Instead they could develop through nearby weak dimer binding to the core, forming a 'cloud' of O_i around the donor core. This binding will be increasingly weak with increasing defect size, and these atoms will not significantly perturb the core. Thus there will be a common set of vibrational modes as observed, and any symmetry idenfication will only measure the C_{2v} symmetry of the core. The defect levels will become increasingly shallow as these aggregate atoms gradually increase the stress exerted on the defect core. This model is also consistent with the rapid observed reorientation under stress.

Such a picture does not attempt to tackle the final stages of TD formation. After ~900 hours annealing at 420°C the 999 cm^{-1} thermal donor vibrational absorption drops away[239], while three new bands appear at 1006, 1012 and 1015 cm^{-1}. It seems likely that this stage represents the thermal donors finally transforming into quartz precipitates, with these new bands being associated with other electrically active defect complexes that form at this point, possibly shallow thermal donors. In addition once quartz precipitates form there are other 'new donors' formed, possibly associated with the quartz boundaries[275].

There are other explanations for these higher order modes, apart from shallow thermal donors vibrations. They could be caused by dimer / trimer trapping near to the quartz. Alternatively if quartz precipitates are being produced in these regimes, the associated lattice strain punches out prismatic dislocation loops in the Si. These dislocations can be pinned by O_i, and so it seems reasonable they may also be pinned by dimers and trimers. Hence these broadened modes could be associated with dimers and trimers pinned at dislocations, the electrical activity coming from the dislocations themselves.

The best that theory can do at this stage is to propose various models, as has been done above, and therefore suggest discriminatory experiments to select between them. We are involved in on-going discussions with several experimental groups on this subject (Lindström and Hallberg, McQuaid, Newman and Markevich amongst others), and hope to be able to reach some firmer conclusions soon.

Chapter 10

Conclusions and Further Work

> Normally screws are so cheap and small and simple you think of them as unimportant. But now, as your Quality awareness becomes stronger, you realise that this one, individual, particular screw is neither cheap nor small nor unimportant. Right now this screw is worth exactly the selling price of the whole motorcycle, because the motorcycle is actually valueless until you get the screw out. With this re-evaluation of the screw comes a willingness to expand your knowledge of it.
>
> R. Pirsig
> Zen and the Art of Motorcycle Maintenance

Although conclusions have been drawn at the end of each individual chapter, it is useful to discuss some general conclusions here, as well as further work that could stem from that presented in this thesis.

The results for H in III-V materials has potential for expansion to other members of the III-V family, notably GaN which shows great potential as a wide bandgap semiconductor. Such an investigation for the Group II passivation is currently underway by Torres and Öberg [276], and preliminary results suggest a transition from BC to AB hydrogen for various sized Group II impurity in GaN. The advantage of such work is the variety of different III-Vs and Group II impurities which makes it possible to draw general trends depending on host and impurity type.

There are many conclusions to be drawn from the O in Si work. Our successful collaboration with a variety of experimental groups has shown the advantage of productive experimental/theoretical discussion. Recent developments in computing power and parallelisation of the code mean we can now tackle reasonably large defect systems such as the $(CH)_iO_{4i}$ shallow thermal donor which would have been impossible only a few years ago. Theory is now equipped to take a place as a serious investigative tool alongside conventional methods such as FTIR and DLTS, and theoretical investigations can now help to point experimental groups in directions for their research[1].

The problem of oxygen in silicon is a tricky one, since oxygen seems able to complex with most other impurities, both extrinsic and intrinsic. The difficulty is compounded for theory since many of the defect structures formed are only metastable (for example, thermal donors), and so absolute ground state energy comparison cannot necessarily be used as a tool for determining the dominant defect structures.

Formation energies with AIMPRO for O-related complexes have proved difficult (such as VO_n complexes), possibly due to inadequate modelling of the charge density via our intermediate fits. In future we may be able to improve on this through the use of the new supercell AIMPRO code under development in Newcastle, possibly in conjunction with gradient corrections to the exchange-correlation energy. Although many kinetic models for TD formation and destruction have been proposed[147, 143, 277], for these to achieve the desired accuracy, formation energies of the individual species need to be determined. This has not been done to date. Hopefully with the improvements to AIMPRO described above, coupled with structural models for the various component species, we may be able to achieve this.

Significant advances have been made in our understanding of O in Si as a result of the work in this thesis. The behaviour of O in the presence of N has been well characterised, both through the primary electrically inactive NNO defect, and the more elusive NO_2 shallow thermal donor, as well as the intermediate N_iO_i species which for the first time shows direct evidence of interaction between N_i and oxygen rather than N_2. In future there will hopefully be direct evidence for the incorporation of N in the shallow thermal donor centres, and they may be produced in sufficient numbers for clear identification through their vibrational modes. Various N/O experiments have been suggested here which would fill in the remaining gaps in our understanding of how N_2 pairs diffuse, and complex with either O_i or O_{2i} dimers.

Crucially we have demonstrated a general mechanism whereby defects that normally possess a deep gap level (such as N_i) can have this level electrostatically compressed, pushing it up to become a shallow level lying close to the conduction band edge. It may be possible to invoke a similar mechanism in other materials, possibly those where it has traditionally been difficult to produce shallow donor doping. We have shown that this mechanism can occur in defects containing both N_i and $(CH)_i$, and although there is no direct experimental evidence yet that N_i or $(CH)_i$ lie in the core of the shallow thermal donor defects, we have presented a

[1]Dimer modelling described here led successful experimental research to find further absorption modes.

strong case in favour of these models.

The results on rapidly diffusing O_{2i} dimers appear to shed light on many otherwise confusing results, such as the anomalous formation rates of VO_2 and the thermal donors. Their role in thermal donor formation is now clearer, although there is still a long way to go on this problem.

Finally the results on the thermal donors show a variety of different O-based complexes which display shallow double donor character. The di-y-lid structure, formed from a combination of two dimers, seems to fit most of the available symmetry and structural evidence, and possesses vibrational modes very close to those of experiment. A variant of this with Al substituted for the core Si acts as a single shallow donor and is also consistent with experimental data for the third type of NL10 defect, NL10(Al).

We thus have a remarkably consistent set of structural predictions for shallow donor defects in Cz-Si, all based around the combination of two dimers. The thermal donors form around a di-y-lid, which occurs when two dimers combine. When these two dimers aggregate at Al_s, N_i or $(CH)_i$, they form instead a shallow thermal donor defect. Such a picture does not yet explain the whole family of thermal donors, including the various anomalous properties of TD1, although the work presented in this thesis goes a long way to selecting between various models.

Long time annealing studies show that the shallow thermal donors form after the thermal donors [10], which is consistent with FTIR studies; these show the TD absorption at 999 cm^{-1} drop off at 450° after 150 hours or so, at the same time as strong absorption at 1012, 1006 and 1015 cm^{-1} begins. These lines are believed to be associated with the shallow thermal donors [239]. Discussion in Section 9.7 suggests that early Si_i concentrations are low. This leads to a consistent picture whereby small initial Si_i concentrations lead to low C_i and hence STD $(CH)_iO_{4i}$ concentrations. After long annealing times thermal donors form, and then transform into quartz precipitates, releasing Si_i. These convert the remaining C_s into C_i, providing a sudden glut of $(CH)_iO_{4i}$ defects. Thereafter, any more Si_i forming has to aggregate into $\langle 311 \rangle$ rod-like aggregates, as seen with HREM[278].

For reasons of limited space, results on C/O defects such as C_iO_{2i} and C_sO_{ni} where $1 \leq n \leq 3$ have been excluded from this thesis, but this work is important and a thorough study of C/O defects is essential in order to understand O in Si. C primarily exists as a substitutional defect (C_s), but can switch sites with a silicon self-interstitial (Si_i) through a 'kick-out' mechanism to produce interstitial carbon (C_i). Thus $[C_i]$ can be taken as an indirect marker of the Si_i concentration. P-line defects are observed using PL (believed to be C_iO_{2i}) within the first 15 minutes of annealing Cz-Si at 450°C[197], suggesting that there is either an initial pool of C_i, or some initial Si_i. However the concentration of these defects is extremely small, probably $\sim 10^{13}$ cm^{-3}[279]. Annealing C rich CZ-Si gives rise to a set of vibrational lines associated with C/O defect complexes instead of thermal donors. Due to the high frequency of the C-modes it has been suggested that these are C_iO_{ni} defects [239], and if this was the case it would show that Si_i must be produced during early Cz-Si annealing. However preliminary modelling results with AIMPRO show that C_sO_{ni} defects produce vibrational modes close to these frequencies, and the modes due to the C are high because the C_s is compressed by neighbouring oxygen. In combination with the low initial concentration of P-line defects, these results suggest that Si_i is *not* present in significant quantities in the earlier stages of annealing, and thus provide more evidence that Si_i does not lie in the core of thermal donors. In addition the role of C in suppressing thermal donors is not fully understood (since C appears to increase the rate of thermal donor dissociation rather than decreasing the rate of formation). Further studies of C/O defects are currently underway[280].

The work in this thesis has opened up a number of new questions (and failed to answer others!) and many of these deserve further study. As well as modelling and/or explaining current experimental data we have also attempted to make predictions, and there are some experiments it would be useful to perform to examine these. In addition there are more general questions such as the mobility of N_2 pairs which could be checked by experiment. Where such experiments suggest themselves, these have been described in the text. This following section is therefore restricted to further theoretical modelling work that could extend the results in this thesis.

It would be useful to produce a full potential surface map for dimer diffusion to prove that we have located the lowest energy path between the end structures. If the assymetric dimer structure is freezing in at low temperatures we could calculate a diffusion barrier between the assymetric and symmetric forms and hence a temperature at which they could no longer switch between the two.

The trimer requires more investigation of various alternative structures in order to determine both the lowest energy, and a structure that could give rise to the absorption at 1006 cm^{-1}. The results presented here were performed in quite small clusters given the size of the defects, and for reliability the convergence with cluster size should be checked by repeating these in larger clusters. In addition, determination of binding energies and potential diffusion paths for the trimer would be important to calculate.

The thermal donor problem still has many unanswered questions that we can address. There are several other possible TD structures not investigated here (notably the +2 variant of the 'Manx' trimer), and it would be useful to repeat many of these calculations using the supercell to provide reliable energy differences between alternative structures. We are still unsure of the nature of TD1 and TD2, and are far from a definitive explanation of higher order TDs. Such analysis could be helped through a detailed correlation of electronic IR with vibrational IR absorption data showing the formation stages of the TDs, and it is clear that this problem will not be solved without extensive collaboration with experimental groups.

Notable related problems are the mechanism for H-passivation of thermal donors, and carbon suppression of thermal donors, both of which are currently under investigation using AIMPRO. Neither of these are obvious, for example, H-passivation seems to remove both donor states from the gap with the addition of only a single H [203], suggesting the H is stabilising the TD in an inactive form rather than saturating dangling bonds.

Clues to the aggregation processes in TDs could come from studying the aggregation of O in higher order VO_n complexes, $n \geq 4$, and increased computational power means we now have the ability to do this. If we can increase the ease with which diffusion barriers are obtained, it would be nice to model the migration barrier for other species, notably N_i and N_{2i}.

Finally, there are many other atomic species that can complex with oxygen which have not been addressed here. A variety of hydrogen–oxygen complexes are known to form [157], and oxygen can also complex with many metallic species, from Group II elements [281] to rare earth impurities such as Er [282].

Work of this kind could be accused of being 'defect stamp collecting', *i.e.* a defect is examined, a structure produced and stuck in the album, and the process then moves on to the next. However such accusations are somewhat unfair. The real strength of these results comes from considering them as a whole, in the context of the additionally available experimental work. In this way they serve to show a whole cycle of defect formation processes and how they relate to the treatment and properties of the sample.

Applied theoretical modelling such as this fulfils three roles. Its first role is to *confirm*; either to produce results that can be compared to other calculations performed using different methods, or to confirm ideas built up from experimental results (see, for example, results on O_i in this thesis). More usefully, theory can *explain*, *i.e.* take unexplained, and maybe even apparently contradictory experimental results, and produce a theory that ties them all together. This is true of much of the work in this thesis, notably the VO_n work and N/O complexes. However, the third and most important job of theory is to *predict*. In this case, theory should identify general mechanisms and trends, suggest new experiments and techniques, and open up new fields of work. Ideally we would like it to be able to provide new ways of examining things, and find underlying patterns, trends and mechanisms. It is at this stage that theory moves far beyond the ‘stamp collecting’ metaphor, and hopefully this has been achieved with some of the results in this thesis such as the ‘wonderbra mechanism’.

Bibliography

[1] J. Czochralski, Z. Phys. Chem. **92**, 219 (1918).

[2] K. Waiser, P. H. Keck, and C. F. Lange, Phys. Rev. **101**, 264 (1956).

[3] T. Carlberg, J. Electrochem. Soc., **133**, 1940 (1986).

[4] J. C. Mikkelsen, Jr., Mat. Res. Soc. Symp. Proc. **59**, 19 (1986).

[5] F. Shimura, Chapter 1, 'Semiconductors and Semimetals, Vol.42 - Oxygen in Silicon', p.5, Academic Press Ltd, Ed. F. Shimura (1994).

[6] T. M. Tkacheva, G. N. Petrov, K. L. Enisherlova, N. A. Iasamanov, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 517-526, ed. R. Jones, Kluwers Academic Press, (1996).

[7] G. A. Rozgonyi, R. P. Deysher and C. W. Pearce, J. Electrochem. Soc. **123** 1910 (1976).

[8] C. S. Fuller, N. B. Ditzenberger, N. B. Hannay and E. Buehler, Phys. Rev. **96**, 833 (1954).

[9] W. Von Ammon, P. Dreier, W. Hensel, U. Lambert, L. Koster, Mat. Sci. Eng. B, **36**, 1-3, p.33-41 (1996).

[10] C. A. J. Ammerlaan, I. S. Zevenbergen, Yu. V. Martynov and T. Gregorkiewicz, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 61-82, ed. R. Jones, Kluwers Academic Press, (1996).

[11] Yu V. Martynov, T. Gregorkiewicz, C. A. J. Ammerlaan, Phys. Rev. Lett. **74** 11, 2030-2033 (1995).

[12] Kh. A. Abdullin, B. N. Mukashev, Yu. V. Gorelkinskii, Mat. Sci. Forum, Vol. 196-201, p.1007-1012 (1995).

[13] D. R. Hartree, Proc. Camb. Phil. Soc. **24**, 89, 111, 426 (1928).

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

[15] V. Sahni, Int. Journ. Quantum Chem. **56**, 4, p.265-283 (1995).

[16] P. Hohenberg, W. Kohn, Phys. Rev. **136** 3864-71 (1964).

[17] W. Kohn, L. J. Sham, Phys. Rev. **140** A1133-8 (1965).

[18] W. Kohn, P. Vashishta, 'Theory of the inhomogenous electron gas', ed. S. Lundquist, N. H. Marck (New York, Plenum) (1983).

[19] F. W. Kutzler, G. S. Painter, Phys. Rev. B **45**, 3236 (1992).

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

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

[22] J. P. Perdew, Y. Wang, Phys. Rev. B **45**, 13244 (1992).

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

[24] U. von Barth, L. Hedin, J. Phys. C **5**, 1629 (1972).

[25] See, for example, R. O. Jones, O. Gunnarsson, Rev. Mod. Phys. **61**, No.3, p. 689-740 (1989).

[26] D. R. Hamann, M. Schlüter, C. Chiang, Phys. Rev. Lett **43** 20, 1494 (1979).

[27] S. G. Louie, S. Froyen, M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).

[28] L. C. Snyder, J. W. Corbett, Proc. ICDS, Ed. L. C. Kimerling, J. M. Parsey Jr., p. 693-699 (1984); L. C. Snyder, J. W. Corbett, Mat. Res. Soc. Symp. Proc. **59**, p. 207-220, Ed. J. C. Mikkelsen Jr., S. J. Pearton, J. W. Corbett, S. J. Pennycook, (1986).

[29] M. Kohyama, S. Takeda, Phys. Rev. B **46**, 19, p.12305-12315 (1992).

[30] Z. Jiang, R. A. Brown, Chem. Eng. Sci. **49**, 2991 (1994).

[31] P. J. Grönberg, J. von Boehm, R. M. Nieminen, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 441-446, ed. R. Jones, Kluwers Academic Press, (1996).

[32] A good description of CNDO and related semi-empirical methods can be found in 'Approximate Molecular Orbital Theories', by J. A. Pople and D. L. Beveridge, McGraw-Hill, 1970. Reference for this thesis was also made to the thesis of J. P. Hagon, University of Newcastle Upon Tyne, 1986.

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

[34] M. J. P. Musgrave, J. A. Pople, Proc. Roy. Soc. **A268**, 474 (1962).

[35] R. Jones, J. Phys. C **20**, 271 (1987).

[36] R. Jones, P. R. Briddon, To appear in *'Identification of Defects in Semiconductors'*, Ed. M. Stavola, in *Semiconductors and Semimetals*, Ed. R. K. Willardson, A. C. Beer, E. R. Weber, Academic Press (1997).

[37] R. Jones, J. Goss, C. P. Ewels, S. Öberg, Phys. Rev. B **50** 12, p.8378-8388 (1994).

[38] B. I. Dunlap, W. J. Connolly, J. R. Sabin, J., Chem. Phys. **71**, 4993 (1979)

[39] R. Jones, A. Sayyash, J. Phys. C: Solid State Phys. **19**, L653 (1986)

[40] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge (1987).

[41] R. S. Leigh, B. Szigeti, Proc. Roy. Soc. A **301**, 211 (1967).

[42] A. Resende, J. Goss, P. R. Briddon, S. Öberg, R. Jones, ICDS '97, To be published (1997).

[43] W. Kohn, Oral Presentation, Brazil International Defects Conference (1991).

[44] S. J. Jenkins, G. P. Srivastava, J. C. Inkson, Surf. Rev. and Lett., **1**, 4, p.473-476 (1994).

[45] S. J. Jenkins, G. P. Srivastava, J. C. Inkson, J. Phys. C : Cond. Matter **6**, 8781-8794 (1994).

[46] D. Lamoen, P. Ballone, M. Parrinello, Phys. Rev. B **54**, 7, p.5097-5105 (1996).

[47] J. Goss, Thesis, Exeter University, (1997).

[48] J. Miro, P. Deák, C. P. Ewels, R. Jones, J. Phys. C : Cond. Matter, to be published (1997).

[49] J. Slater, *"Quantum Theory of Molecules and solids"*, Volume 2, Mc. Graw-Hill (1965).

[50] A. Alminningen, O. Bastiansen, V. Ewing, K. Hedberg, M. Traetteberg, Acta. Chem. Scand. **17**, 2544 (1963).

[51] H. H. Weider, *Thin Solid Films 100, Some Defects of the Technology of III-V Compound Semiconducting Layers* (Elsevier Sequoia, Lausanne), p. 391 (1983).

[52] J. J. Loferski, J. Appl. Phys. **27**, 777 (1956).

[53] For a recent review see, for example, J. Chevallier, B. Clerjaud, and B. Pajot, in *Hydrogen in Semiconductors*, Semiconductors and Semimetals, Vol 34, edited by J. I. Pankove and N. M. Johnson (Academic, London), p. 447 (1991).

[54] S. J. Pearton, W. S. Hobson and C. R. Abernathy, Appl. Phys. Lett. **61** 1558 (1992).

[55] T. Sugino, H. Ninomiya, K. Matsuda and J. Shirafuji, Jpn. J. Appl. Phys. **33** Part 2, 3A (1994).

[56] O. Mizuno and H. Watanabe, Electron. Lett. **11**, 118 (1975).

[57] A. Carnera, A. Gasparotto, M. Tromby, M. Caldironi, S. Pellegrino, F. Vidimari, C. Bocchi and C. Frigeri, J. Appl. Phys. **76**, 9 (1994).

[58] F. X. Zach, J. Appl. Phys. **75**, 12 (1994).

[59] R. Darwich, B. Pajot, B. Rose, D. Robein, B. Theys, R. Rahbi, C. Porte and F. Gendron, Phys. Rev. B **48**, 24 (1993).

[60] B. Bech Nielsen, F. Berg Rasmussen, L. Hoffmann, M. Budde, R. Jones, J. Goss and S. Öberg, Proceedings of ICDS Conference, Sendai, Japan (1995).

[61] D. W. Fischer, M. O. Manasreh and G. Matous, J. Appl. Phys. **71**, 4805 (1992).

[62] S. T. Pantelides, in *Festkörperprobleme (Advances in Solid State Physics)*, edited by P. Grosse (Pergamon, New York), Vol. XV, p. 149 (1975).

[63] P. H. Borcherds, G. F. Alfrey, D. H. Saunderson and A. D. B. Woods, J. Phys C **8**, 2022 (1975).

[64] C. Hilsum, S. J. Fray and C. Smith, Solid State Commun. **7**, 1057 (1969).

[65] E. S. Koteles and W. R. Datars, Solid State Commun. **19**, 221 (1976).

[66] O. G. Folberth and H. Weiss, Z Naturforsch. Teil A **10**, 615 (1955).

[67] P. K. Khowash, Phys. Rev. B **49**, 16 (1994).

[68] A. P. Seitsonen, R. Virkkunen, M. J. Puska and R. M. Nieminen, Phys. Rev. B **49** 8, 5253 (1994).

[69] M. A. Roberson and S. K. Estreicher, Phys. Rev. B **49** 24, 17040 (1994).

[70] B. Clerjaud, D. Côte and C. Naud, Phys. Rev. Lett. **58**, 1755; J. Cryst. Growth **83**, 190 (1987).

[71] H. J. von Bardeleben, D. Stievenard, K. Kainosho and O. Oda, J. Appl. Phys. **70** (12), 7392 (1991).

[72] R. Pritchard, B. R. Davidson, R. C. Newman, T. J. Bullough, T. B. Joyce, R. Jones, S. Öberg, Semicond. Sci. Tech. **9** 140 (1994).

[73] P. R. Briddon, R. Jones, Phys. Rev. Lett. **64** 21, p.2535-2538 (1990).

[74] J. Neugebauer, C. G. Van der Walle, Phys. Rev. Lett. **75**, 24, p.4452-4455 (1995).

[75] J. W. Corbett, G. D. Watkins, R. M. Chrenko, and R. S. McDonald, Phys. Rev., **121**, 1015 (1961).

[76] G. D. Watkins, and J. W. Corbett, Phys. Rev.,**121**, 1001 (1961).

[77] D. R. Bosomworth, W. Hayes, A. R. L. Spray, and G. D. Watkins, Proc. Roy. Soc., **A 317**, 133 (1970).

[78] B. Pajot, Chapter V in *Oxygen in Silicon*, ed. by F. Shimura, in the series: *Semiconductors and Semimetals*, ed. by R. K. Willardson and A. C. Beer, **42**, (Academic Press 1994), p. 191.

[79] C. A. Londos, N. V. Sarlis, L. G. Fytros, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 477-484, ed. R. Jones, Kluwers Academic Press, (1996).

[80] B. G. Svensson, J. L. Lindström, J. W. Corbett, Appl. Phys. Lett. **47** (8), p.841-843 (1985).

[81] B. G. Svensson, and J. L. Lindström, Phys. Rev. B **34**, 8709 (1986).

[82] R. C. Newman, A. K. Tipping, and J. H. Tucker, J. Phys. C: Solid St. Phys, **18**, L861 (1985).

[83] T. Lisby, B. Bech Nielsen, to be published (1997).

[84] H. J. Stein, Mater. Sci Forum, **10-12**, 935 (1986).

[85] F. A. Abou-el-Fatouh, and R. C. Newman, Solid State. Commun. **15**, 1409 (1974).

[86] G. D. DeLeo, W. B. Fowler, and G. D. Watkins, Phys. Rev. B **29**, 3193, (1984).

[87] R. C. Newman, R. Jones, Chapter VI in *Oxygen in Silicon*, ed. by F. Shimura, in the series: *Semiconductors and Semimetals*, ed. by R. K. Willardson and A. C. Beer, **42**, (Academic Press 1994), p. 289.

[88] Y. H. Lee, and J. W. Corbett, Phys. Rev. B, **13**, 2653 (1979).

[89] G. Davies, E. C. Lightowlers, R. C. Newman, and A. S. Oates, Semicond. Sci. and Technol., **2**, 524 (1987).

[90] J. L. Lindström, and B. G. Svensson, Mat. Res. Soc. Symp., **59**, 45 (1986).

[91] R. Jones, P. Leary, S. Öberg, V. J. T. Torres, Mat. Sci. Forum **196-201**, 785 (1995).

[92] L. C. Snyder and J. W. Corbett, in *Proceedings of the 13th International Conference on Defects in Semiconductors, Coronado, California, 1984*, ed. L. C. Kimerling and J. M. Parsey, Jr. (The Metallurgical Society of the AIME, Warrendale), p.693 (1985).

[93] M. Saito and A. Oshiyama, Phys. Rev. B **38** 15, 10711 (1988).

[94] R. Jones, A. Umerski, and S. Öberg, Phys. Rev. B, **45** 19, 11321, (1992).

[95] J. Chadi, Oral Presentation, NATO ARW, Oxygen '96 (1996).

[96] A. Uedono, Y. K. Cho, S. Tanigawa, A. Ikari, Jap. J. Appl. Phys. 1, **33**, 1A, p.1-5 (1994).

[97] C. A. Londos, G. I. Georgiou, L. G. Fytros and K. Papastergiou, Phys. Rev. B **50** 16, 11531-11534 (1994).

[98] C. A. Londos, N. Sarlis, L. G. Fytros and K. Papastergiou, Phys. Rev. B **53** 11, 6900-6903 (1996).

[99] C. P. Ewels, R. Jones, S. Öberg, Mat. Sci. Forum Vol.196-, p.1297-1301, (1995).

[100] F. Berg Rasmussen, S. Öberg, R. Jones, C. Ewels, J. Goss, J. Miro and P. Deák, Mat. Sci. Forum Vol. 196-201 pp. 791-796, (1995).

[101] C. Kaneta, T. Sasaki, and H. Katayama-Yoshida, Phys. Rev., B **46**, 13179 (1992).

[102] J. L. Linström, and T. Hallberg, Phys. Rev. Lett., **72**, 2729-2732 (1994).

[103] R. Jones, C. Ewels, J. Goss, J. Miro, P. Deák, S. Öberg, F. Berg Rasmussen, Semicon. Sci. and Technol., **9**, 2145-2148 (1994).

[104] R. Jones, and S. Öberg, Phys. Rev. Lett., **68**, 86, (1992).

[105] P. Deák, L. C. Snyder, and J. W. Corbett, Phys. Rev. B **45** 20, 11612, (1992).

[106] J. W. Corbett, R. S. McDonald and G. D. Watkins, J. Phys. Chem. Solids, **25**, p.873 (1964)

[107] W. L. Bond and W. Kaiser, J. Phys. Chem. Solids, **16**, 44 (1960).

[108] B. Pajot, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 283-302, ed. R. Jones, Kluwers Academic Press, (1996).

[109] R. C. Newman, '*Infra-Red Studies of Crystal Defects*', Taylor and Francis, London (1973).

[110] R. Jones, S. Öberg, and A. Umerski, Mat. Sci. Forum, **83-7**, 551 (1991).

[111] R. Jones, Semicond. Sci. Tech. **5**, 225-260 (1990).

[112] L. C. Snyder, R. Wu, and P. Deák, Radiation Effects and Defects in Solids, **111** & **112**, 393 (1989).

[113] S. K. Estreicher, Phys Rev. B **41**, 9886 (1990).

[114] P. J. Kelly, R. Car, Phys. Rev. B, **45** 12, 6543-6563 (1992).

[115] M. Needels, J. D. Joannopoulos, Y. Bar-Yam and S. T. Pantelides, Phys. Rev. B **43**, 4208 (1991); Mat. Res. Soc. Symp. Proc. **209**, 103 (1991).

[116] E. Artacho, A. Lizón-Nordström, and F. Ynduráin, Phys. Rev. B **51**, 7862-7865 (1995).

[117] R. C. Lord, D. W. Robinson, W. C. Schumb, J. Am. Chem. Soc. **78**, 1327 (1956).

[118] J. R. Aronson, R. C. Lord, D. W. Robinsin, J. Chem. Phys. **33**, 1004 (1960).

[119] L. Levien, C. T. Prewitt, and D. J. Weidner, American Mineralogist, **65**, 920 (1980).

[120] A. Baghdadi, W. M. Bulles, M. C. Croarkin, Yue-Zhen Li, R. I. Scace, R. W. Series, P. Stallhofer and M. Watanabe, J. Electrochem. Soc. **136**, 2015-2024 (1989).

[121] L. C. Snyder and J. W. Corbett, Mater. Res. Soc. Symp. Proc. **59**, 207 (1986); **104**, 179 (1988).

[122] M. Stavola, J. R. Patel, L. C. Kimerling, P. E. Freeland, Appl. Phys. Lett. **42**, 73 (1983).

[123] R. C. Newman, J. H. Tucker, F. M. Livingstone, J. Phys. C **16**, L151 (1983).

[124] S. -T. Lee, D. Nichols, Appl. Phys. Lett. **47**, 1001 (1985).

[125] Newman R. C., Jones R., in *Oxygen in Silicon*, edited by Shimura F., *Semiconductors and Semimetals*, edited by Willardson R. K. and Beer A. C., **42**, p. 289, Academic Press (1994).

[126] S. A. McQuaid, R. C. Newman, J. H. Tucker, E. C. Lightowlers, A. Kubiak and M. Goulding, Appl. Phys. Lett. **58**, 2933 (1991).

[127] H. J. Stein, S. K. Hahn, Appl. Phys. Lett. **56**, 63 (1990).

[128] R. Jones, S. Öberg and A. Umerski, Mater. Sci. Forum **241** 551 (1992).

[129] Z. Jiang, R. A. Brown, Phys. Rev. Lett. **74** (11), p.2046 (1995).

[130] M. Ramamoorthy, and S. T. Pantelides, Phys. Rev. Lett. **76**, 267-270, (1996).

[131] J. C. Mikkelsen, Appl. Phys. Lett. **40**, 336 (1982).

[132] G. D. Watkins, J. W. Corbett, and R. S. McDonald, J. Appl. Phys. **53**, 7097 (1982).

[133] S. -Tong Lee, D. N. Nichols, Appl. Phys. Lett. **47**, 1001 (1985).

[134] A. Oshiyama and M. Saito, in *Defect Control in Semiconductors*, edited by K. Sumino (Elsevier Science Publishers, North Holland, 1990).

[135] P. J. Kelly, Mater. Sci. Forum **38-41**, 269 (1989).

[136] M. Ramamoorthy and S. T. Pantelides, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 197-205, ed. R. Jones, Kluwers Academic Press, (1996).

[137] U. Gösele, T. Y. Tan, Appl. Phys. A, **28**, 79 (1982).

[138] M. A. Roberson, S. K. Estreicher, and C. H. Chu, J. Phys. C: Cond. Matter, **5**, 8943 (1993).

[139] P. Gaworzewski and G. Ritter, Phys. Stat. Sol. (a) **67**, 511-516 (1981).

[140] P. Wagner, J. Hage, J. M. Trombetta and G. D. Watkins, Mat. Sci. Forum **83-87**, 401 (1992).

[141] J. W. Corbett, G. D. Watkins, R. S. McDonald, Phys. Rev. B, **135**, A1381 (1964).

[142] S. -T. Lee, P. Fellinger, S. Chen, J. Appl. Phys. **63** 1924 (1988).

[143] S. A. McQuaid, M. J. Binns, C. A. Londos, J. H. Tucker, A. R. Brown and R. C. Newman, J. Appl. Phys. **77**, 1427 (1995).

[144] H. Yamanaka, Jpn. J. Appl. Phys. **33**, 3319 (1994).

[145] R. Jones, S. Öberg, F. Berg Rasmussen, B. Bech Neilsen, Phys. Rev. Lett. **72**, 1882 (1994).

[146] A. R. Bean and R. C. Newman, J. Phys. Chem. Solids **33**, 255 (1972).

[147] V. P. Markevich, I. F. Medvedeva, L. I. Murin, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 103-122, ed. R. Jones, Kluwers Academic Press, (1996).

[148] B. Pajot, H. J. Stein, B. Cales, and D. Naud, J. Electrochem. Soc. **132** 3034-3037 (1985).

[149] T. Hallberg, Thesis, Linköping University, Sweden (1995).

[150] T. Hallberg, J. L. Lindström, L. I. Murin, V. P. Markevich, Proceedings ICDS'97, Portugal, to be published (1997).

[151] T. Hallberg, Private Communication (1997).

[152] S. J. Breuer, R. Jones, P. R. Briddon, S. Öberg, Phys. Rev. B **53**, 24, p.16289-16296, (1996).

[153] H. J. Hrostowski, R. H. Kaiser, Phys. Rev. **107**, 966 (1957).

[154] B. Pajot, B. Cales, Mat. Res. Soc. Symp. Proc. **59**, 39 (1986).

[155] B. Pajot, E. Artacho, C. A. J. Ammerlaan and J. -M. Spaeth, J. Phys. C : Cond. Matt. **7**, 7077-7085 (1995).

[156] A. Umerski and R. Jones, Phil. Mag. A **67**, 4, 905-915 (1993).

[157] B. Hourahine, R. Jones, S. Öberg, unpublished (1997).

[158] R. Newman, MRS Symp. **104**, Boston, (1987).

[159] P. Rava, H. C. Gatos, and J. Lagowski, *Semiconductor Si*, edited by H. Huff, R. Kriegler and Y. Takeishi (Electrochem. Soc., Princeton, NJ), p.232 (1981).

[160] J. L. Lindström, T. Hallberg, J. Appl. Phys. **77**, 6, p.2684-2690 (1995).

[161] K. Sumino, I. Yonenaga, M. Imai, and T. Abe, J. Appl. Phys., **54**, 5016 (1983).

[162] M. I. H. Heggie, R. Jones, and A. Umerski, Phys. Stat. Sol., (a) **138**, 383 (1993).

[163] K. L. Brower, Phys. Rev., B **26**, 6040 (1973).

[164] H. J. Stein, in *Oxygen, Carbon, Hydrogen and Nitrogen in Crystalline Silicon*, Mat. Res. Soc. Symp., Pittsburgh PA, ed J. C. Mikkelsen Jr., S. J. Pearton, J. W. Corbett, and S. J. Pennycook, **59**, 523 (1985).

[165] M. Suezawa, K. Sumino, H. Harada, and T. Abe, Jpn. J. Appl. Phys., **27**, 62 (1988).

[166] A. Hara, M. Aoki, M. Koizuka, and T. Fukuda, J. Appl. Phys. **75**, 2929 (1994).

[167] J. A. Griffin, H. Navarro, J. Weber, and L. Genzel, Mater. Sci. Forum, **10-12**, 997 (1986).

[168] B. J. Heijmink Liesert, T. Gregorkiewicz, and C. A. J. Ammerlaan, Phys. Rev. B **47**, 7005 (1993).

[169] Q. Sun, K. H. Yao, H. C. Gatos, J. Lagowski, J. Appl. Phys. **71** 8, p.3760-3765 (1992).

[170] H. J. Stein, Appl. Phys. Lett. **47**, 1339, (1985).

[171] K. L. Brower, Phys. Rev. B **26**, 6040 (1982).

[172] H. J. Stein, Appl. Phys. Lett. **52**, 153 (1988).

[173] H. J. Stein, MRS Proceedings, **46**, edited by N. M. Johnson, S. G. Bishop, and G. D. Watkins, MRS Pittsburgh, PA (1985), p. 287.

[174] P. Wagner, R. Oeder, and W. Zulehner, Appl. Phys. A, **46**, 73 (1988).

[175] M. W. Qi, S. S. Tan, B. Zhu, P. X. Cai, W. F. Gu, X. M. Xu, T. S. Shi, D. L. Que, and L. B. Li, J. Appl. Phys., **69**, 3775 (1991).

[176] F. Berg Rasmussen, R. Jones, S. Öberg, Phys. Rev. B **50** 7, 4378 (1994).

[177] F. Berg Rasmussen, S. Öberg, R. Jones, C. Ewels, J. Goss, J. Miro, P. Deák, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 319-328, ed. R. Jones, Kluwers Academic Press, (1996).

[178] M. W. Qi, T. S. Shi, S. S. Tan, B. Zhu, P. X. Cai, L. Q. Liu, D. L. Que, and L. B. Li, Material Science Forum, **83-7**, 263 (1992).

[179] W. Kaiser, H. L. Frisch and H. Reiss, Phys. Rev. **112**, 1546 (1958).

[180] A. Gali, J. Miro, P. Deák, C. P. Ewels, R. Jones, J. Phys. C : Cond. Matter, **8**, 7711-7722 (1996).

[181] F. Berg Rasmussen, Thesis, Aarhus University, Aarhus, Denmark (1994).

[182] The original idea has been greatly improved thanks to discussion with Brian Bech Neilsen (1996).

[183] C. P. Ewels, R. Jones, S. Öberg, J. Miro, P. Deák, Phys. Rev. Lett. **77** 5, p.865-868 (1996).

[184] T. Abe, T. Masui, H. Harada, J. Chikawa, in *VLSI Science and Technology*, edited by W. M. Bullis and S. Broydo (Electrochemical Society, Pennington, NJ), p.543 (1985).

[185] H. J. Stein in *Oxygen, Carbon, Hydrogen and Nitrogen in Crystalline Silicon*, Mat. Res. Soc. Symp., Pittsburgh PA, ed Mikkelsen Jr. J. C., Pearton S. J., Corbett J. W. and Pennycook S. J., **67**, 523 (1986).

[186] H. Navarro, J. Griffin, J. Weber, L. Gentzelm, Solid State Comm. **58**, 151 (1986).

[187] M. Suezawa, K. Sumino, H. Harada, T. Abe, Jpn. J. Appl. Phys. **25**, L859 (1986).

[188] J. A. Griffin, H. Navarro, J. Weber, L. Genzel, J. T. Borenstein, J. W. Corbett, L. C. Snyder, J. Phys. C : Solid State Phys. **19**, L579-L584, (1986).

[189] D. Yang, D. Que, K. Sumino, J. Appl. Phys. **77**, 2, p.943 (1995).

[190] A. Hara, T. Fukuda, T. Miyabo, I. Hirai, Jpn. J. Appl. Phys. **28** 1, p.142-143 (1989).

[191] A. Hara, L. Hirai, A. Ohsawa, J. Appl. Phys. **67**, 2462 (1990).

[192] Yu V. Martynov, T. Gregorkiewicz, and C. A. J. Ammerlaan, Phys. Rev. Lett. **74** 11, 2030-2033 (1995).

[193] J. A. Griffin, J. Hartung, J. Weber, H. Navarro, L. Genzel, Appl. Phys. A **48** 41-47 (1989).

[194] A. Endrös, Phys. Rev. Lett. **63**, 70 (1989).

[195] A. L. Endrös, W. Krüher, J. Grabmaier, Physica B **170**, 365 (1991).

[196] A. N. Safonov, E. C. Lightowlers, G. Davies, P. Leary, R. Jones, S. Öberg, Phys. Rev. Lett. **77**, 23, p.4812-4815 (1996).

[197] N. S. Minaev, A. V. Mudryi, Phys. Stat. Solidi A, **68**, 561 (1981).

[198] A. N. Safonov, E. C. Lightowlers, Mat. Sci. Forum **143-147**, 903 (1994).

[199] E. C. Lightowlers, R. C. Newman, and J. H. Tucker, Semicond. Sci. and Tech. **9**, 1370 (1994).

[200] R. C. Newman, J. H. Tucker, N. G. Semaltianos, E. C. Lightowlers, T. Gregorkiewicz, I. S. Zevenbergen and C. A. J. Ammerlaan, Phys. Rev. B **54** 10, R6803-6806 (1996).

[201] H-related STDs break down at 520°C, R. C. Newman, Private Communication (1996).

[202] D. I. Bohne and J. Weber, Mat. Sci. Forum **143-147** pp.879-884, Trans Tech Publications, Switzerland (1994).

[203] J. Weber and D. I. Bohne, Proceedings of NATO ARW, "Early Stages of Oxygen Precipitation in Silicon", NATO ASI Series 3, Vol. 17, p. 123-140, ed. R. Jones, Kluwer Academic Publishers, (1996).

[204] S. H. Muller, M. Sprenger, E. G. Sieverts, and C. A. J. Ammerlaan, Solid State Commun. **25**, 987 (1978).

[205] T. Gregorkiewicz, H. H. P. Th. Bekman, and C. A. J. Ammerlaan, Phys. Rev. B **38**, 3998 (1988).

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

[207] T. Gregorkiewicz, H. H. P. Th. Bekman, and C. A. J. Ammerlaan, Phys. Rev. B **41**, 12628 (1990).

[208] T. Gregorkiewicz, H. H. P. Th. Bekman, and C. A. J. Ammerlaan, Phys. Rev. B **45**, 5873 (1992).

[209] T. Gregorkiewicz, H. H. P. Th. Bekman, and C. A. J. Ammerlaan, Phys. Rev. B **46**, 4582 (1992).

[210] T. Gregorkiewicz, D. A. Van. Wezep, H. H. P. Th. Bekman, and C. A. J. Ammerlaan, Phys. Rev. Lett. **59**, 1702 (1987).

[211] D. J. Chadi, Phys. Rev. Lett. **77** 5, p.861-864 (1996).

[212] R. Jones, S. Öberg, A. Umerski, Mat. Sci. Forum, **72** 287 (1991).

[213] P. Deák, Private Communication (1995).

[214] A. Gali, J. Miro, P. Deák, Proceedings of NATO ARW, "Early Stages of Oxygen Precipitation in Silicon", NATO ASI Series 3, Vol. 17, p. 419-426, ed. R. Jones, Kluwer Academic Publishers, (1996).

[215] P. Leary, R. Jones, S. Öberg, ICDS, Aveiro, Portugal, Mat. Sci. Forum, to be published (1997).

[216] N. Meilwes, J. -M. Spaeth, V. V. Emtsev, G. A. Oganesyan, W. Götz and G. Pensl, Mat. Sci. Forum **143-147**, 141 (1994).

[217] V. P. Markevich, T. Mchedlidze, M. Suezawa, submitted to Phys. Rev. Lett. (1997).

[218] V. P. Markevich, I. F. Medvedeva, L. I. Murin, T. Sekiguchi, M. Suezawa and K. Sumino, Mat. Sci. Forum Vols. 196-201, p.945-950, Trans Tech Public., Switzerland (1995).

[219] H. Hatakeyama, M. Suezawa, V. P. Markevich, K. Sumino, Mat. Sci. Forum Vols. 196-201, p.939-944, Trans Tech Public., Switzerland (1995).

[220] V. P. Markevich, M. Suezawa and K. Sumino, J. Appl. Phys. **76**, 7347 (1994).

[221] M. Stavola, Appl. Phys. Lett. **44**, 514 (1984).

[222] J. E. Gower, E. C. Lightowlers, G. Davies and A. N. Safonov, Proceedings of NATO ARW, "Early Stages of Oxygen Precipitation in Silicon", NATO ASI Series 3, Vol. 17, p. 463-468, ed. R. Jones, Kluwer Academic Publishers, (1996).

[223] Just citing the rightful author of the title 'wonderbra'! Name courtesy of R. Jones, Private Communication (1996).

[224] J. F. Prinns, Diamond and related materials, **4**, 5-6, p.580-585 (1995).

[225] Vermeulen LA, and Farrer, RG, Diamond Research (Suppl. Ind. Diam. Rev.) pp. 18-23 (1975).

[226] C. D. Latham and R. Jones, private communication (1996).

[227] D. Wruck and P. Gaworzewski, Phys. Status Solidi A **56**, 557 (1979).

[228] M. Suezawa and K. Sumino, Mater. Lett. **2**, 85 (1983).

[229] B. Pajot, H. Compain, J. Lerouille and B. Clerjaud, Physica B **177-118**, 110 (1983).

[230] *'Early Stages of Oxygen Precipitation in Silicon'*, NATO ASI Series, Sub Series 3, Vol. 17, Ch.43, ed. R. Jones, Kluwers Academic Press, ISBN 0-79-234296-8 (1996).

[231] J. -M. Spaeth, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 83-102, ed. R. Jones, Kluwers Academic Press, (1996).

[232] J. Michel, J. R. Niklas, J. -M. Spaeth, Phys. Rev. B **40**, 1732 (1989).

[233] T. Gregorkiewicz, D. A. van Wezep, H. H. P. Th. Bekman, C. A. J. Ammerlaan, Phys. Rev. B **35**, 3810-3817 (1987).

[234] E. G. Sieverts, Phys. Stat. Sol. (b) **120**, 11-29 (1983).

[235] J. Michel, J. R. Niklas, J. -M. Spaeth, Mat. Res. Soc. Symp. Proc. Vol. 104, p.185-188, Pittsburgh, (1988).

[236] G. D. Watkins, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 1-18, ed. R. Jones, Kluwers Academic Press, (1996).

[237] J. M. Trombetta, G. D. Watkins, J. Hage, P. Wagner, J. Appl. Phys. **81**, 3, p.1109-1115 (1997).

[238] P. Wagner, J. Hage, J. M. Trombetta and G. D. Watkins, Mat. Sci. Forum **83-87**, 401 (1992).

[239] J. L. Lindström, T. Hallberg, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 41-60, ed. R. Jones, Kluwers Academic Press, (1996).

[240] P. Wagner and J. Hage, Appl. Phys. A **49**, 123 (1989).

[241] W. Götz, G. Pensl and W. Zulehner, Phys. Rev. B **46**, 4312-4315 (1992).

[242] V. V. Emtsev, T. V. Mashovets, G A. Oganesyan, K. Schmalz, Semiconductors, **27** (9), p.854 (1993).

[243] C. A. Londos, M. J. Binns, A. R. Brown, S. A. McQuaid, R. C. Newman, Appl. Phys. Lett. **62** (13), p.1525 (1993).

[244] K. M. Lee, J. M. Trombetta and G. D. Watkins, in *Microscopic Identification of Electronic Defects in Semiconductors*, edited by N. M. Johnson, S. G. Bishop, and G. D. Watkins, MRS Symposia Proceedings No. 46 (Mat. Res. Soc., Pittsburgh), p.263 (1985).

[245] M. Stavola and K. M. Lee in *Oxygen, Carbon, Hydrogen and Nitrogen in Silicon*, edited by J. C. Mikkelsen Jr., S. J. Pearton, J. W. Corbett and S. J. Pennycook, MRS Symp. Proc. No. 59 (Mat. Res. Soc. Pittsburg, 1986), p.95.

[246] M. Stavola, K. M. Lee, J. C. Nabity, P. E. Freeland, L. C. Kimerling, Proc. Spring Mat. Res. Soc., San Fransisco, ed. N. Johns (1985).

[247] M. Stavola., K. M. Lee, J. C. Nabity, P. E. Freeland, L. C. Kimerling, Phys. Rev. Lett **54**, 2630 (1985).

[248] P. Wagner, H. Gottschalk, J. M. Trombetta, G. D. Watkins, J. Appl. Phys. **61**, 346 (1987).

[249] M. Stavola and L. C. Snyder, in *Defects in Silicon*, edited by L. C. Kimerling and M. W. Bullis (Electrochem. Soc., Pennington, NJ), p.61 (1983).

[250] A. Ourmazd, W. Schröter, A. Bourret, J. Appl. Phys. **56**, 1670 (1984).

[251] M. Saito, A. Oshiyama, Mat. Sci. Forum Vol. 38-41, p.655-660, Trans Tech Publications, Switzerland, (1989).

[252] L. C. Snyder, P. Deák, R. Z. Wu and J. W. Corbett, Mat. Sci. Forum Vol. 38-41, p.329-334, Trans Tech Pub., Switzerland, (1989).

[253] P. Deák, L. C. Snyder, J. W. Corbett, Phys. Rev. Lett. **66** (6), p. 747-749 (1991).

[254] Y. A. I. Latusho, L. F. Markevich, and L. I. Murin, Phys. Status Solidi A **93**, K181 (1986).

[255] P. Leary, S. Öberg, R. Jones, unpublished (1996).

[256] P. Deák, private communication (1997).

[257] R. C. Newman, R. Jones, *'Oxygen in Silicon'*, Semiconductors and Semimetals, Vol. 42, Chapter 8, p. 290-347, Ed. Fumio Shimura, Academic Press (1994).

[258] G. G. DeLeo, C. S. Milsted Jr., J. C. Kralik, Phys. Rev. B **31**, 3588 (1985).

[259] P. J. Kelly, Mat. Sci. Forum, Vols. 38-41, p.269-274, ed G. Ferenczi, Trans Tech Publications, Switzerland, (1989).

[260] S. Hahn, in *Oxygen, Carbon, Hydrogen and Nitrogen in Silicon*, Mat. Res. Soc. Pittsburgh, PA, p.181-186, ed S. J. Pearton, J. W. Corbett, J. C. Mikkelsen Jr., and S. J. Pennycook (1986).

[261] Y. Tokuda, M. Katayama and T. Hattori, Semicond. Sci. Technol. **8**, 163-166 (1993).

[262] S. J. Pearton, in *Hydrogen in Semiconductors, Semiconductors and Semimetals*, Vol. 34, 84, Academic Press Inc., London (1991).

[263] N. M. Johnson, S. K. Hahn, Appl. Phys. Lett. **48**, 709-711 (1986).

[264] G. Davies, R. C. Newman, *Carbon in Monocrystalline Silicon*, Chapter 21 *Handbook on Semiconductors*, Ed. T. S. Moss, Vol. 3. ed. S. Mahajar, Pub. Elsevier Science BC (1994).

[265] This work is currently underway by T. Hallberg on samples with various Al doping levels.

[266] U. Gösele, K. -Y. Ahn, B. P. R. Marioton, T. Y. Tan and S. -T. Lee, Appl. Phys. A **48**, 219 (1989).

[267] H. Yamada-Kaneta, Y. Shirakawa, C. Kaneta, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 389-396, ed. R. Jones, Kluwers Academic Press, (1996).

[268] A longer discussion of this is in progress, with reference to the behaviour of thermal donors in the presence of carbon.

[269] L. I. Murin, V. P. Markevich, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 329-336, ed. R. Jones, Kluwers Academic Press, (1996).

[270] T. Hallberg, J. L. Lindström, Mat. Sci. Eng B **36**, 1-3, p.13-15 (1996).

[271] J. L. Lindström, T. Hallberg, Phys. Rev. Lett. **72**, 17, p.2729-2732, (1994).

[272] R. C. Newman, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 19-39, ed. R. Jones, Kluwers Academic Press, (1996).

[273] M. Claybourn and R. C. Newman, Appl. Phys. Lett. **51**, 26, 2197 (1987).

[274] U. Gösele, E. Schroer, P. Werner, T. Y. Tan, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 243-262, ed. R. Jones, Kluwers Academic Press, (1996).

[275] G. Pensl, M. Schulz, K. Hölzlein, W. Bergholz, and J. L. Hutchison, Appl. Phys. A **48**, 49 (1989).

[276] S. Öberg, V. J. Torres, unpublished (1997).

[277] M. Suezawa, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 207-222, ed. R. Jones, Kluwers Academic Press, (1996).

[278] See for example, M. Kohyama, S. Takeda, Phys. Rev. B **46** 19, p.123405-12315, (1992).

[279] E. C. Lightowlers, Private Communication (1996).

[280] P. Leary, Thesis, Exeter Universiy, Exeter (1997).

[281] E. McGlynn, M. O. Henry, S. E. Daly, K. G. McGuigan, NATO ARW, 'Oxygen'96', ASI Series Vol. 17, p. 355-362, ed. R. Jones, Kluwers Academic Press, (1996).

[282] Er/O complexes in Silicon, C. P. Ewels, A. Resende, B. Hourahine, R. Jones, S. Öberg, P. R. Briddon, unpublished (1997).