An *ab initio* Study of Deep-Level Defects in Silicon

Submitted by **António Luís Santos Ferreira-Resende** to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Physics, November 1999.

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

"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,

. . . ."

November 3, 1999.

Dedicated to my parents, Maria and Adelino.

Abstract

Theoretical investigations into deep-level defects in crystalline silicon are presented in this thesis. The calculations are carried out using the AIMPRO code, an *ab initio* pseudopotential local spin density method applied to large hydrogen terminated clusters containing up to 346 atoms.

By definition, deep-level defects are those with localised states, *i.e.*, states with decaying wavefunctions in real space. As a result, these defects can, and often do, give rise to a number of levels lying within the silicon bandgap. Due to the fact that the presence of these levels within the Si forbidden band can dramatically change the optical and electrical properties of Si integrated devices, the understanding of their microscopic properties is of paramount importance.

Two distinct types of deep-level defects are investigated: (*i*) radiation-induced defects and (*ii*) transition-metal (TM) related defects. Most of these defects are unstable against Jahn-Teller distortions. Since their electrical properties are highly sensitive to their atomic arrangement, it is therefore necessary to correctly describe their ground state configurations.

The back-bone of this thesis is a novel theoretical approach to the calculation of the electrical level of deep-level defects. This method has allowed the successful characterisation a number of common defects in radiation damaged Si material, like VO (*A*-centre), VOH, C_iP, C_iO_i, C_iC_s–H (*T*-centre), etc.

The method is then applied to the study of the structural and electrical properties of the lattice di-vacancy. As a result, we confirm the predictions of Watkins and Corbett for the structural properties of paramagnetic $V_2^{\pm}$ defects. Accordingly, these defects undergo a strong Jahn-Teller distortion by *pairing* of four of the six atoms constituting the defect.

This is followed by a study on the structural, vibrational and electrical properties of vacancy–hydrogen-related defects produced by low-temperature proton implantation.

Finally, the method is applied to the analysis of the structural and electrical properties of substitutional transition-metal centres—gold, silver, platinum and palladium. The main goal of these calculations is the study of the influence on the electrical properties of these defects of bringing atomic hydrogen close to the defect's core. Concerning the structure of the TM–H defects, we find no evidence for a direct interaction between the hydrogen impurity and the TM ion. In accordance with the *vacancy model* of Watkins, the TM ions do seem to interact weakly with the surroundings, in an attempt to reproduce their environment as isolated species. Despite an enlargement of the vacancy cage, the hydrogen atoms sit at 'anti-bonding' lattice positions, being back-bonded to the Si vacancy atoms. No electrically inactive TM–H_n, with $n = 1 \ldots 4$, were found and an alternative model for the neutralisation of the electrical properties of these centres is proposed.

Acknowledgements

There are many people to thank for their support and encouragement, without whom for one reason or another, my PhD in Exeter would have not been possible.

First of all, I would like to thank my supervisor Dr. Bob Jones for his guidance, patience and contagious enthusiasm.

Thanks to all the members of the AIMPRO group—in particular Ben, Jamie, Zé Pedro, Jon, Chris, Paul, Sven, Malc and Patrick—as well as to all the other members of the semiconductor physics group.

Much of the work presented here involved the collaboration with experimental groups. I am indebted to the following groups for their support and release of experimental results, in many cases, before publication: IFA at the University of Aahrus (Denmark), CEM at the UMIST – Manchester, Max-Planck-Institut für Festkörperforschung – Stuttgart (Germany) and Lehigh University – Bethlehem, PA (USA), in particular Dr. Brian Bech Nielsen, Prof. Tony Peaker, Dr. Joerg Weber and Dr. Michael Stavola.

Contents

Abstract iii

Acknowledgments iv

List of Publications x

List of Figures xiv

List of Tables xvii

Glossary of Abbreviations and Acronyms xviii

1 The Quantum Many Body Problem 1

1.1 Introduction . . . 1

1.2 The Schrödinger equation . . . 1

1.3 The Born-Oppenheimer approximation . . . 2

1.3.1 Degenerate systems . . . 5

1.4 The one-electron potential . . . 6

1.5 The wavefunction variational principle . . . 7

1.5.1 The Hellmann-Feynman theorem . . . 8

1.5.2 The virial theorem . . . 9

1.6 The Hartree self-consistent field method . . . 10

1.7 Hartree-Fock theory . . . 12

1.7.1 Koopmans' theorem . . . 15

1.8 The X_α method . . . 15

1.9 Shortcomings of the one-electron approximation . . . 17

1.10 The classical density functional models . . . 17

1.10.1 Introduction . . . 17

1.10.2 The model of Thomas and Fermi . . . 18

1.10.3 Exchange effects: Dirac's correction . . . 20

1.10.4 A conventional gradient correction: The Weizsäcker correction . . . 21

2 Modern Density Functional Theory 22

2.1 Introduction . . . 22

2.2 Basic formalism . . . 23

2.2.1 The Hohenberg and Kohn theorems 23
2.2.2 The v-representability of an electron density 26
2.2.3 The constrained search process 27
2.3 The Kohn-Sham scheme 28
2.3.1 Introduction 28
2.3.2 The Kohn-Sham equations 29
2.3.3 The chemical potential 31
2.3.4 Exchange and correlation 32
2.3.5 Interpretation of the Kohn-Sham equations 33
2.4 Local density approximation 33
2.4.1 Discussion of the local density approximation 35
2.5 Spin polarised systems 37
2.5.1 The local spin density approximation 38
2.6 Correlation effects 38
2.7 Beyond the local spin density approximation 40
2.8 Fractional occupancy 42
2.8.1 The Janak theorem 43
2.8.2 Slater's 'transition-state' argument 44
2.8.3 Ionisation potential and electron affinity 46
2.8.4 The band gap problem 48

3 Pseudopotential Theory 50
3.1 Introduction 50
3.2 Basic idea 50
3.3 *Ab initio* pseudopotentials 52
3.3.1 Bachelet, Hamann and Schlüter pseudopotentials 53

4 The Real Space Cluster Method 56
4.1 Introduction 56
4.2 Methodology 56
4.3 The Hartree energy 58
4.4 The exchange-correlation energy 60
4.4.1 Spin polarisation 63
4.5 Full matrix formulation 64
4.6 Approach to self-consistency 66
4.6.1 Fermi statistics 67
4.7 Evaluation of atomic forces 68
4.8 Structure optimisation 69
4.9 Determination of vibrational modes 70
4.9.1 Equations of motion in the harmonic approximation 70
4.9.2 Evaluation of energy double derivatives 71
4.10 Practical considerations 72

4.10.1 Choosing a basis set . . . 72
4.10.2 Basis size . . . 74
4.10.3 Termination effects . . . 74

5 Deep Level Characterisation 76
5.1 Introduction . . . 76
5.2 Shallow *versus* deep levels . . . 77
5.3 The effects of lattice relaxation . . . 77
5.4 Carrier emission and recombination . . . 79
5.5 Detection and characterisation of deep levels . . . 81
5.5.1 Deep level transient spectroscopy . . . 84
5.5.2 Laplace deep level transient spectroscopy . . . 88
5.5.3 Electron paramagnetic resonance . . . 89
5.5.4 Local vibrational mode spectroscopy: infra-red absorption . . . 95
5.6 Photoluminescence . . . 98
5.7 Modelling deep levels . . . 100
5.7.1 Electrical level calculations with supercells . . . 101
5.7.2 Novel approach to the calculation of deep levels . . . 102

6 Radiation Defects 105
6.1 Introduction . . . 105
6.2 Electrical properties characterisation . . . 108
6.2.1 The V–O pair: the A-centre . . . 108
6.2.2 Interaction of hydrogen with the A-centre . . . 110
6.2.3 The di-vacancy (V_2) . . . 112
6.2.4 The V_2O complex . . . 112
6.2.5 The VP pair: the E-centre . . . 112
6.2.6 The isolated C_i impurity . . . 113
6.2.7 The C_iP complex . . . 113
6.2.8 The C_iO_i pair: the K-centre . . . 114
6.2.9 Interaction of hydrogen with the K-centre . . . 114
6.2.10 The C_sC_i pair . . . 115
6.2.11 The C_sC_i–H defect: the T-centre . . . 116
6.3 Summary . . . 117

7 The Lattice Divacancy 118
7.1 Introduction . . . 118
7.2 Experimental background . . . 119
7.2.1 Structural properties . . . 119
7.2.2 Electrical and optical properties . . . 120
7.3 Previous theoretical work . . . 122
7.4 Results . . . 125

7.4.1 Cluster and basis . . . 125
7.4.2 The undistorted di-vacancy . . . 125
7.4.3 The V_2^+ and V_2^- paramagnetic states . . . 127
7.4.4 Optical properties . . . 131
7.4.5 Electrical levels . . . 132
7.5 Summary . . . 134

8 Vacancy–Hydrogen Defects 136
8.1 Introduction . . . 136
8.2 Background . . . 137
8.2.1 Electrical properties . . . 141
8.3 Previous theoretical studies . . . 143
8.4 Results . . . 145
8.4.1 Cluster and basis . . . 145
8.4.2 VH and V_2H . . . 146
8.4.3 VH_2 . . . 149
8.4.4 VH_3 *versus* V_2H_6 . . . 151
8.4.5 VH_4 . . . 152
8.4.6 Electrical properties of multi-hydrogen–vacancy defects . . . 154
8.5 Summary . . . 155

9 Platinum and Palladium Defects 157
9.1 Introduction . . . 157
9.2 Experimental work . . . 158
9.2.1 Microscopic models for the electronic structure . . . 159
9.3 Results . . . 169
9.3.1 Cluster and basis . . . 169
9.3.2 Microscopic structure as isolated species . . . 169
9.3.3 Transition-metal–hydrogen defects . . . 170
9.3.4 Vibrational properties . . . 174
9.3.5 Electrical properties . . . 176
9.3.6 Electrical passivation of Pt and Pd centres . . . 180
9.4 Final remarks . . . 181

10 Gold and Silver Defects 182
10.1 Introduction . . . 182
10.2 Experimental background . . . 183
10.2.1 Structural properties as isolated centres . . . 183
10.2.2 Interaction with atomic hydrogen . . . 184
10.2.3 FTIR measurements . . . 184
10.2.4 Electrical properties . . . 184
10.3 Results . . . 186

10.3.1 Cluster and basis 186
10.3.2 Isolated gold and silver centres: structural properties 186
10.3.3 Structural and vibrational properties 187
10.3.4 Electrical levels 189
10.3.5 On the passivation of ioslated transition metal centres 191
10.4 Concluding remarks 192

11 Conclusions and Future Work 194

Bibliography 196

List of Publications

Listed in reverse chronological order:

1. *Microscopic structure and electrical levels of platinum and platinum–hydrogen defects in silicon: A first principles study*, A. Resende. R. Jones, S. Öberg and P. R. Briddon, To be submitted to Phys. Rev. B

2. *An ab initio study of the structural and electrical properties of divacancy-tin complexes in silicon*, A. Resende. M. Kaukonen, R. Jones, S. Öberg and P. R. Briddon, To be submitted to Phys. Rev. B

3. *The Yarykin conjecture – defect profiles and the chemical composition of defects*, R. Jones, B. J. Coomer, A. Resende and S. Öberg and P. R. Briddon, To be submitted to Phys. Rev. B

4. *Antithetical distortions of the divacancy in silicon and diamond*, B. J. Coomer, J. P. Goss, A. Resende and R. Jones, S. Öberg and P. R. Briddon, To be submitted to Phys. Rev. Lett.

5. *Why do divacancies distort in semiconductors?* B. J. Coomer, A. Resende, J. P. Goss, R. Jones, S. Öberg, and P. R. Briddon, in the Proceedings of the 20th ICDS, 26–30 July 1999, Berkeley, USA (To appear in Physica B).

6. *The Interaction of Hydrogen with Deep Level Defects in Silicon*, R. Jones, B. J. Coomer, J. P. Goss, B. Hourahine, and A. Resende. To be published in Special Defects in Semiconducting Materials, edited by R. P. Agarwala.

7. *Interstitial clusters in silicon*, B. J. Coomer, J. P. Goss, A. Resende and R. Jones, S. Öberg and P. R. Briddon, To appear in "Materials Science in Semiconductors"

8. *Theoretical investigation of electrically active defects in silicon*, R. Jones, A. Resende, S. Öberg and P. R. Briddon, To appear in "Materials Science in Semiconductors"

9. *Calculations of electrical levels of deep centers: Application to Au-H and Ag-H defects in silicon*, A. Resende, R. Jones, S. Öberg and P. R. Briddon, Phys. Rev. Lett. **82**, 2111 (1999).

10. *The structural properties of transition metal hydrogen complexes in silicon*, A. Resende, R. Jones, S. Öberg and P. R. Briddon, Mater. Sci. & Eng. B **28**, 146 (1999).

11. *The electronic properties of transition metal hydrogen complexes in silicon*, R. Jones, A. Resende, S. Öberg and P. R. Briddon, Mater. Sci. & Eng. B, **28**, 113 (1999).

12. *Theory of gold-hydrogen complexes in silicon*, A. Resende, J. Goss, P. R. Briddon and R. Jones, Mat. Sci. Forum **258–263**, 295 (1997).

List of Figures

2.1 Optimised spin-polarised one-electron eigenvalue spectra for H_2O and its positive and negative ions, and the corresponding intermediate transition states . . . 46

3.1 The real and pseudo-radial $4s$ wavefunction for the Ni atom . . . 50

4.1 Variation of polarised and non-polarised exchange-correlation energy per unit volume with density . . . 62

5.1 The electronic energy of a twofold degenerate localised state and the elastic energy . . . 78
5.2 Thermal emission and capture processes from electron and hole traps, generation and recombination centres . . . 80
5.3 Distribution of filled and empty traps at steady state in a p^+-n step junction under zero bias and reverse bias . . . 82
5.4 Isothermal capacitance regime evolution after the application of a reverse bias when the p^+-n junction is initially reversed biased. . . . 83
5.5 Implementation of a DLTS rate window by means of a boxcar integrator with gates set at times t_1 and t_2 . . . 84
5.6 Diagrams of some rate-window operations: box-car, lock-in amplifier and two-phase lock-in amplifier . . . 85
5.7 Schematic representation of a defect potential in an electric field . . . 86
5.8 p^+-n junction diagram showing the dominant photon-semiconductor interactions during the fill pulse of a MCTS experiment . . . 87
5.9 Comparison between a standard DLTS and a Laplace spectrum of Pt-doped Si. The spectra shown originate from the single acceptor level of substitutional Pt^-. . . . 88
5.10 The phonon dispersion spectrum of a diatomic linear chain plotted in the first Brillouin zone . . . 96
5.11 Schematic representation of low temperature photoluminescence . . . 98
5.12 Schematic configuration co-ordinate diagram of the luminescence transitions 99
5.13 A schematic of the localised wavefunction of the ionised defect and that of the marker . . . 103

6.1 Schematic of the implantation and damage processes using energetic ion beams 106
6.2 Lowest-energy structures of the VO and VOH defects in the transition-state 108
6.3 Optimised structure of the isolated carbon interstitial 112
6.4 Lowest-energy configuration of the C_iP defect 113
6.5 Low energy structures of the C_iO_i–H complex 115
6.6 Optimised structures of the C_sC_i defect (A-form) and the C_iC_s–H defect (T-centre) with a transition-state occupancy of $-\frac{1}{2}\,e$ 116

7.1 The two normal modes of the Jahn-Teller distortion responsible for the lowering of symmetry of V_2 from D_{3d} to C_{2h} 123
7.2 The effect of symmetry lowering for the two possible Jahn-Teller distortions on the one-electron levels of the undistorted di-vacancy 126
7.3 3D-isosurface plot of the wavefunction corresponding to the two highest occupied and the lowest unoccupied spin-up Kohn-Sham levels of V_2^- 130
7.4 3D-isosurface plot of the wavefunction correspoding to the two highest occupied and the lowest unocupied spin-up Kohn-Sham levels of V_2^- 131
7.5 Scaled Kohn-Sham eigenvalue spectra for the lowest-energy structures of V_2 (C_{2h}) in three different charge states, in the vicinity of the band gap 132

8.1 The H_2^* defect 138
8.2 Relaxed structures for the neutral form of the four distinct C_{3v} configurations of *interstitial* SiH_3 139
8.3 Scaled spin-polarised Kohn-Sham eigenvalue spectra of neutral VH_n defects in the vicinity of the band gap 147
8.4 3D-isosurface plot of the wavefunction corresponding to the highest occupied and the lowest unoccupied spin-up Kohn-Sham levels of C_{1h} VH^0 148
8.5 Two views of the lowest-energy structure of monoclinic VH^0 149
8.6 3D-isosurface plot of the wavefunction corresponding to the highest occupied spin-up Kohn-Sham level of VH_3 152

9.1 Representation of the one-electron levels of the negative silicon vacancy in full tetrahedral symmetry, the d^{10} states of the transition-element impurity, and the substitutional impurity in silicon 160
9.2 Energy-level diagram depicting the level of Pt and PtH as observed by DLTS and EPR techniques 166
9.3 Optimised structure of C_{1h} PtH_1^- complexes with the hydrogen atom at anti-bonding (AB) sites, outside and inside the vacancy 171
9.4 Scaled spin-polarised Kohn-Sham eigenvalue spectra of singly negatively charged Pt and PtH defects in the vicinity of the band gap 172
9.5 Optimised structure of C_{2v} PtH_2^- defect with the hydrogen atom at anti-bonding (AB) sites, outside the vacancy cage 177

9.6 Structure and corresponding electrical levels of platinum and palladium and their complexes with hydrogen 180

10.1 One-electron molecular orbital picture of the formation of the gold impurity states in silicon 183
10.2 Scaled spin-polarised Kohn-Sham eigenvalue spectra of neutral Au and AuH defects in the vicinity of the band gap 188
10.3 A gold mono-hydride molecule in a hydrogen-induced microvoid 190
10.4 Structure and corresponding electrical levels of gold and silver and their complexes with hydrogen in Si 192

List of Tables

2.1 Parameters of the fit to the correlation energy per particle and the high- and low-density expansions . . . 40

2.2 Typical errors for atoms, molecules and solids from self-consistent Kohn-Sham calculations within the LSD and GGA approximations . . . 41

2.3 Calculated and observed first ionisation energy potentials and electron affinities the H_2O molecule. . . . 48

4.1 Parametrisation of the spin-polarised exchange-correlation energy . . . 63

4.2 Lowest Kohn-Sham level, highest occupied level, calculated and experimental band gap for various Si clusters . . . 75

6.1 Oxygen and carbon content in different silicon materials . . . 107

6.2 Calculated $(-/0)$ level (eV) versus Si–Si separation for the VO pair . . . 109

6.3 Structural parameters for the optimised structure of neutral VOH_1 and VOH_2 defects . . . 111

6.4 Summary of calculated and observed level positions of common deep-level point defects in irradiated and/or implanted crystalline silicon . . . 117

7.1 Observed capture cross-sections for the levels of V_2 from DLTS measurements 121

7.2 Character table for the C_{2h} point-symmetry group . . . 122

7.3 Atomic basis and basis expansion type used for the calculations on V_2 . . . 125

7.4 Total cluster energy differences for different configurations and charge states of V_2 . . . 128

7.5 Calculated geometries and distortion types for several charge states of V_2 . 129

7.6 One-electron occupancies for the transition-state used to calculate the ionisation and affinity of the defect in two different configurations . . . 133

7.7 Comparison between calculated and observed electrical levels of the lattice di-vacancy . . . 135

8.1 Relative thermal stability of the several neutral vacancy-hydrogen defects . 141

8.2 Comparison between calculated and observed local vibrational modes of multi-hydrogen–vacancy defects . . . 144

8.3 Atomic basis and basis expansion type used for the calculations on vacancy-hydrogen defects . . . 146

8.4 Structural parameters for the optimised neutral vacancy–hydrogen defects containing up to four hydrogen atoms 147
8.5 Comparison between the calculated and observed H-stretch frequencies of neutral VH_1 and V_2H 150
8.6 Variation of the calculated hydrogen and deuterium local vibrational modes of VH_3 with charge state 150
8.7 Comparison between the calculated and observed H-stretch frequencies of neutral VH_3 and V_2H_6 defects 152
8.8 A summary of the calculated hydrogen and deuterium stretch frequencies for multi-hydrogen–vacany defects in silicon 153
8.9 Calculated electrical levels of VH_n defects 155

9.1 Activation energy ΔE, majority capture cross sections, activation energy of the capture cross section, the enthalpy and the entropy factor for Pt and Pd energy levels 163
9.2 Activation energy ΔE, majority capture cross sections, activation energy of the capture cross section, the enthalpy and the entropy factor for PtH levels 164
9.3 Activation energy ΔE, majority capture cross sections, activation energy of the capture cross section, the enthalpy and the entropy factor for PdH levels 165
9.4 Penetration lengths L_n and the corresponding experimental assignments for PtH and PdH complexes 168
9.5 Atomic basis and basis expansion type used for the calculations on Pt and Pd–H defects 169
9.6 Structural parameters for the Pt_s^- and singly negatively charged PtH_n centres. PtH_1 values are given for the two competing configurations 173
9.7 Calculated and observed hydrogen and deuterium stretch frequencies for C_{2v} PtH_2 complexes in three charge states 174
9.8 Calculated and observed hydrogen and deuterium stretch frequencies for two configurations of PtH_1 (C_{1h}) in three charge states 175
9.9 Calculated and observed hydrogen and deuterium stretch frequencies (cm^{-1}) of the symmetric and anti-symmetric modes for two configurations (see text) of C_{2v} $(PtH_2)^q$ complexes, the corresponding values of M_N and ω_H/ω_D ratios. 176
9.10 Comparison between calculated and experimental electrical levels of Pt and Pd–H defects 178

10.1 Activation energy E_a, majority capture cross sections, activation energy of the capture cross section, the enthalpy and the entropy factor for Au and Ag energy levels 185
10.2 Activation energy E_a, majority capture cross sections, activation energy of the capture cross section, the enthalpy and the entropy factor for AuH and AgH levels 185

10.3 Atomic basis and basis expansion type used for the calculations on Au and Ag–H defects 187
10.4 Structural parameters for the Au_s and neutral AuH_n centres. AuH_1 values are given for the two competing configurations 188
10.5 Calculated and observed hydrogen and deuterium stretch modes for three configurations of AuH_1 in three different charge states 189
10.6 Calculated and observed hydrogen and deuterium stretch frequencies for three configurations of AuH_2 (C_{2v}) complexes 189
10.7 Comparison between calculated and experimental electrical levels of Au– and Ag–H defects 191

Glossary of Abbreviations and Acronyms

AIMPRO	*Ab initio* modelling program
BHS	Bachelet-Hamann-Schlüter (pseudopotentials)
CB	Conduction band
Cz	Czochralski (silicon material – oxygen-rich)
DFT	Density functional theory
DLTS	Deep level transient spectroscopy
ENDOR	Electron nuclear double resonance spectroscopy
EPR	Electron paramagnetic resonance spectroscopy
FTIR	Fourier transform infra-red spectroscopy
FZ	Float zone (silicon material – carbon-rich)
GGA	General gradient approximation (exchange-correlation)
HF	Hartree-Fock (theory)
JT	Jahn-Teller (effect/distortion)
KS	Kohn-Sham (theory/eigenvalues)
LCAO	Linear combination of atomic orbitals
LDA	Local density approximation (exchange-correlation)
LDLTS	Laplace deep level transient spectroscopy
LSD	Local spin density (theory)
LVM	Local vibrational mode
MCTS	Minority carrier transient spectroscopy
ODMR	Optical detection of magnetic resonance
PL	Photoluminescence
RBA	Reverse bias annealing
RT	Room temperature
TM	Transition metal (impurity)
TRIM	Transport of ions in materials (Monte Carlo program)
VB	Valence band (maximum)
ZPL	Zero phonon line

1

The Quantum Many Body Problem

1.1 Introduction

The properties of a material are dependent on its composition and its structure: this is the general law of nature. A *crystal* consists of a periodic array of atoms. An atom is made of a nucleus of charge Z, surrounded by a number of neutralising electrons.

Core electrons occupy the filled inner shells of the atom and their main effect is to screen the nuclear charges from the valence electrons. It is, therefore, tempting to eliminate them in a practical application by replacing them with an effective potential acting on valence electrons. This is the main idea behind the concept of *pseudopotentials*, which will be introduced and discussed in a forthcoming chapter of this thesis.

The chemically important valence-electrons, as well as all those in lighter elements, move at speeds much less than the velocity of light, and can be considered to be non-relativistic particles. All the important characteristics of the material are due to interactions between valence electrons and nuclei, and with themselves: the basic interaction between electrons and nuclei of charge Z is electrostatic or Coulombic. Other interactions arising from the quantum-mechanical character of the electron, are also important. This is the case of exchange and correlation effects, which act like a *glue* without which atoms would bond weakly if at all, to form the solid.

In order to tackle the many-body problem, two initial approximations have to be made. First, since the electrons of the inner shells do not take part in physical processes in the solid, only electrons from outer shells will be considered. The second approximation is due to Born and Oppenheimer. It allows the decoupling of the electronic and nuclear motions, treating the nuclei as classical particles. This is the subject of the next section of this thesis.

1.2 The Schrödinger equation

The stationary state of a quantum-mechanical system is described by the non-relativistic time-independent Schrödinger equation,

$$\widehat{\mathcal{H}}\,\Psi = E\,\Psi\ , \tag{1.2.1}$$

where $\widehat{\mathcal{H}}$ is the crystal Hamiltonian operator, Ψ its wavefunction,

$$\Psi \equiv \Psi(\vec{r}_1, \sigma_1; \vec{r}_2, \sigma_2; \ldots, \vec{r}_N, \sigma_N; \vec{R}_1, \vec{R}_2, \ldots, \vec{R}_N) \equiv \Psi(\vec{r}_i, \vec{R}_\alpha) , \tag{1.2.2}$$

and E its eigenvalue or crystal energy. Note that the total wavefunction Ψ is a function of the whole set of electron spatial and spin co-ordinates, $\{\vec{r}_i, \sigma_i\}$, and nuclear co-ordinates $\vec{R}$ of all the particles existing in the system.

For a system of N electrons moving in a field of N ions of charge Z_α at sites $\vec{R}_\alpha$, the full Hamiltonian has the following form:[1]

$$\begin{aligned}\widehat{\mathcal{H}}(\vec{r}, \vec{R}) = &-\frac{1}{2}\sum_i \nabla_i^2 - \sum_\alpha \frac{1}{2M_\alpha}\nabla_\alpha^2 + \frac{1}{2}\sum_i \sum_{\substack{j \\ j\neq i}} \frac{1}{|\vec{r}_i - \vec{r}_j|} \\ &+ \frac{1}{2}\sum_\alpha \sum_{\substack{\beta \\ \beta\neq\alpha}} \frac{Z_\alpha Z_\beta}{|\vec{R}_\alpha - \vec{R}_\beta|} - \sum_i \sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} .\end{aligned} \tag{1.2.3}$$

The terms with electron $i = j$ and nuclear indices $\alpha = \beta$ are excluded to prevent self-interactions. In an obvious notation, (1.2.3) can be re-stated as

$$\widehat{\mathcal{H}} \equiv \widehat{T}_e + \widehat{T}_i + \widehat{V}_{e-e} + \widehat{V}_{i-i} + \widehat{V}_{e-i} . \tag{1.2.4}$$

The presence of the kinetic term associated with the nuclear motion, $\widehat{T}_i$, turns the study of the hypothetical many-body system into a very complex problem.

The Schrödinger equation contains $3(Z+1)N$ variables, where N is the number of atoms in the crystal. Since 1 cm^3 of a crystalline material contains about 5×10^{22} atoms, for $Z = 14$ (silicon), the number of variables involved reaches more than 2×10^{24} variables per cm^3!

Obviously, this equation cannot be solved in a general form. This is not only because of difficulties of a practical nature, for modern quantum mechanics lacks the means of solving many-particle problems.

1.3 The Born-Oppenheimer approximation

The *adiabatic* approximation, or the Born-Oppenheimer approximation (1925), takes account of the difference in motion of light (electrons) and heavy (nuclei) particles. Important for fast moving electrons is the instantaneous position of the nuclei. Obviously, the nuclei, on account of their large masses, do not follow the motion of every electron but move only in the averaged field of all the electrons. At the same time the ion moving relatively slowly is accompanied by the electrons.

If we neglect the kinetic energy of the nuclei, the Hamiltonian is simply

$$\widehat{\mathcal{H}}_e = \widehat{T}_e + \widehat{V}_{e-e} + \widehat{V}_{i-e} + \widehat{V}_{i-i} , \tag{1.3.1}$$

[1]All quantities are expressed in atomic units (unless otherwise specified). In terms of these units, $\hbar$, e, m and $4\pi\varepsilon_0$ are taken to be unity. The unit of length, 1 a.u. = 0.529 Å and the unity of energy 1 a.u. = 27.211 eV.

with the associated eigenfunction ψ satisfying

$$\widehat{H}_e\,\psi(\vec{r},\vec{R}) = \mathcal{E}_e(\vec{R})\,\psi(\vec{r},\vec{R})\ , \tag{1.3.2}$$

dropping the subscripts i and α, as well as the electron spin s_i, for sake of simplicity. In this equation, $\widehat{H}_e$ is the total Hamiltonian in the limit $M \to \infty$, or electronic Hamiltonian. Note that the nuclear co-ordinates enter in (1.3.2) as parameters: $\mathcal{E}_e(\vec{R})$ is the energy of the electron in a field of stationary ions. Analytically

$$\mathcal{E}_e(\vec{R}) = \int \psi^*(\vec{r},\vec{R})\,\widehat{H}_e(\vec{r})\,\psi(\vec{r},\vec{R})\,d^3r\ . \tag{1.3.3}$$

As long as we consider the nuclei as static particles, we can consider the term $\widehat{V}_{i-i}$ as a simple additive constant in the Hamiltonian. This will have no effect on the wavefunction ψ but will add directly to the eigenvalue $\mathcal{E}_e$. We may now write the total Hamiltonian as

$$\widehat{\mathcal{H}} = \widehat{H}_e - \sum_{\alpha}^{N} \frac{1}{2M_\alpha}\,\nabla^2_\alpha\ . \tag{1.3.4}$$

What Born and Oppenheimer (1925) did was to show that, because the ratio (m/M_α) is so small, we can treat $\widehat{H}_n$ as a perturbation. The relevant expansion is $\kappa = (m/M)^{1/4}$, where M is the average nuclear mass ($\kappa \approx 0.1$). Born and Oppenheimer carried out an expansion of the eigenfunctions and eigenvalues in powers of κ. We shall proceed to show the validity of this approximation.

In the zero-order Born-Oppenheimer approximation (which is all that one normally uses) to the wavefunction, one writes an approximate total eigenfunction of $\widehat{\mathcal{H}}$ as a simple product

$$\Psi(\vec{r},\vec{R}) = \phi(\vec{R})\,\psi(\vec{r},\vec{R})\ . \tag{1.3.5}$$

Here ϕ depends only on the nuclear co-ordinates $\vec{R}$. Replacing the total wavefunction $\Psi(\vec{r},\vec{R})$ into (1.2.1), and re-calling (1.3.5) results in

$$\begin{aligned}
\psi(\vec{r},\vec{R}) &\sum_{\alpha} \frac{1}{2M_\alpha}\nabla^2_\alpha\,\phi(\vec{R}) + \mathcal{E}_e(\vec{R})\,\phi(\vec{R})\,\psi(\vec{r},\vec{R}) \\
&+ \frac{1}{2}\sum_{\substack{\alpha,\beta\\ \beta\neq\alpha}} \frac{Z_\alpha\,Z_\beta}{|\vec{R}_\alpha - \vec{R}_\beta|}\,\phi(\vec{R})\,\psi(\vec{r},\vec{R}) - E\,\phi(\vec{R})\,\psi(\vec{r},\vec{R}) \\
&= \sum_{\alpha} \frac{1}{2M_\alpha}\left(\phi(\vec{R})\,\nabla^2_\alpha\,\psi(\vec{r},\vec{R}) + 2\vec{\nabla}_\alpha\cdot\vec{\nabla}_\alpha\psi(\vec{r},\vec{R})\right)\ .
\end{aligned} \tag{1.3.6}$$

Multiplying all terms by ψ^* and integrating over all electronic co-ordinates, we have[2]

$$\begin{aligned}
&\left[\sum_{\alpha} \frac{1}{2M_\alpha}\nabla^2_\alpha\ +\ \mathcal{E}_e(\vec{R})\ + \frac{1}{2}\sum_{\alpha,\beta\neq\alpha} \frac{Z_\alpha\,Z_\beta}{|\vec{R}_\alpha - \vec{R}_\beta|} - E\right]\phi(\vec{R}) \\
&\qquad = \sum_{\alpha} \frac{1}{2M_\alpha}\left(\phi(\vec{R}) \int \psi^*(\vec{r},\vec{R})\,\nabla^2_\alpha\,\psi(\vec{r},\vec{R})\,d^3r\right. \\
&\qquad\quad \left. +\ 2\,\vec{\nabla}_\alpha\,\phi(\vec{R})\cdot\int \psi^*(\vec{r},\vec{R})\,\vec{\nabla}_\alpha\,\psi(\vec{r},\vec{R})\,d^3r\right)\ .
\end{aligned} \tag{1.3.7}$$

[2]Since $\{\psi\}$ is a complete orthonormal set.

At this stage, we might define, explicitly, the nuclear Hamiltonian, $\widehat{H}_n$:

$$\widehat{H}_n\,\phi \equiv \left[\widehat{T}_i + \widehat{V}_{i-i} + \mathcal{E}_e(\vec{R})\right]\phi = \mathcal{E}_n\,\phi\ . \tag{1.3.8}$$

Pre-multiplying the above equation by ϕ^* and integrating over ionic co-ordinates, we obtain the expression for the energy of the crystal,

$$\int \phi^*(\vec{R})\,\widehat{H}_n\,\phi(\vec{R})\,d^3R = E + \delta E\ , \tag{1.3.9}$$

with

$$\begin{aligned}\delta E \equiv \sum_\alpha \frac{1}{2M_\alpha}\Bigg(&\int \psi^*(\vec{r},\vec{R})\,\nabla_\alpha^2\,\psi(\vec{r},\vec{R})\,d^3r \\ &+ 2\int \phi^*(\vec{R})\,\vec{\nabla}_\alpha\phi(\vec{R})\,d^3R\cdot\int \psi^*(\vec{r},\vec{R})\,\vec{\nabla}_\alpha\,\psi(\vec{r},\vec{R})\,d^3r\Bigg)\ .\end{aligned} \tag{1.3.10}$$

In order to asses the quantity δE (1.3.10), it is necessary to make certain assumptions concerning the nature of the dependence of ψ on the co-ordinates $\vec{r}$ and $\vec{R}$.

If the electron-electron interaction is neglected in $\widehat{H}_e$ (*i.e*, $\widehat{V}_{e-e} = 0$), it turns into a Hamiltonian of a system of non-interacting particles. The wavefunctions ψ will be simple products of one-electron wavefunctions, each of which only depends on the difference $(\vec{R}_\alpha - \vec{r}_i)$. Under these conditions,

$$\vec{\nabla}_\alpha^k\,\psi(\vec{R}-\vec{r}) = (-1)^k\,\vec{\nabla}_i^k\,\psi(\vec{r},\vec{R}) \qquad \text{with} \quad k = 1 \text{ and } 2\ . \tag{1.3.11}$$

The first term in (1.3.10) may be re-written in the form,

$$\begin{aligned}\sum_\alpha \frac{1}{2M_\alpha}\int \psi^*(\vec{r},\vec{R})\,\nabla_\alpha^2\,\psi(\vec{r},\vec{R})\,d^3r &= \sum_{i,\alpha}\frac{1}{2M_\alpha}\int \psi^*(\vec{r},\vec{R})\,\nabla_i^2\,\psi(\vec{r},\vec{R})\,d^3r \\ = -\sum_{i,\alpha}\frac{m}{2M_\alpha}\int \psi^*(\vec{r},\vec{R})\left(-\frac{\nabla_i^2}{2m}\right)\psi(\vec{r},\vec{R})\,d^3r &= -\sum_{i,\alpha}\frac{m}{M_\alpha}\,\langle\widehat{T}_e\rangle\ ,\end{aligned} \tag{1.3.12}$$

where $\langle\widehat{T}_e\rangle$ is the mean kinetic energy per electron. Hence,

$$\sum_\alpha \frac{1}{2M_\alpha}\int \psi^*(\vec{r},\vec{R})\,\nabla_\alpha^2\,\psi(\vec{r},\vec{R})\,d^3r = -N\,\frac{m}{M_\alpha}\,\langle\widehat{T}_e\rangle \tag{1.3.13}$$

where N is the number of electrons in the system. This term may be neglected. In comparison with E, the error introduced thereby into the expression for the crystal energy is quite small—of the order of the ratio of the electron mass to the ionic mass, which, for example, for ^{28}Si is $\sim 10^{-5}$.

The second term in (1.3.10) can be analysed in the same way:

$$\frac{1}{M_\alpha}\left(\int \phi^*(\vec{R})\,\vec{\nabla}_\alpha\,\phi(\vec{R})\,d^3R\cdot\int \psi^*(\vec{r},\vec{R})\,\vec{\nabla}_\alpha\,\psi(\vec{r},\vec{R})\,d^3r\right) = -\frac{1}{M_\alpha}\left(\langle\vec{p}_i\rangle\cdot\langle\vec{p}_e\rangle\right) \tag{1.3.14}$$

In classical statistics, for a state at thermodynamical equilibrium, we have

$$\langle p_e\rangle^2 = \frac{8}{3\pi}\,\langle p_e^2\rangle \quad \text{and} \quad \left\langle\frac{\vec{p}_e^{\,2}}{2m}\right\rangle = \left\langle\frac{\vec{p}_i^{\,2}}{2m}\right\rangle\ . \tag{1.3.15}$$

It follows that

$$\langle p_e \rangle \approxeq \sqrt{\frac{m}{M_\alpha}}\, \langle p_i \rangle \; . \tag{1.3.16}$$

To analyse the corrections in (1.3.9), above we have considered the most disadvantageous situation to illustrate the validity of the adiabatic approximation. This is when the electron wavefunction is written as a combination of atomic wavefunctions (the so-called strongly-bound electron approximation). If the electron wavefunction is independent of the ionic co-ordinates (the free electron approximation), *both corrections* in (1.3.10) will be *zero*.

We arrive at the conclusion that by neglecting both corrections in (1.3.10), we introduce an error that in the value of E that is not greater than the square root of the m/M ratio. It is now clear that the adiabatic approximation gives a sufficiently accurate energy value if the assumption is made that the total wavefunction may be written in the form

$$\Psi(\vec{r}_i, \vec{R}_\alpha) = \phi(\vec{R}_\alpha)\, \psi(\vec{r}_i) \; , \tag{1.3.17}$$

ϕ and ψ being determined from the equations,

$$\widehat{H}_e\, \psi = \mathcal{E}_e\, \psi \qquad \text{and} \qquad \widehat{H}_n\, \phi = \mathcal{E}_n\, \phi \equiv E\, \phi \; . \tag{1.3.18}$$

Thus, in the Born-Oppenheimer approximation, the electron wavefunction is determined by the instantaneous position of the ions (the term $\widehat{V}_{e-i}$ in $\widehat{H}_e$), the ionic wavefunction, on the other hand, being determined by the average electron field (term $\mathcal{E}_e$ in $\widehat{H}_n$).

It important to note that the Born-Oppenheimer approximation breaks down for Jahn-Teller systems. This is the case when an interaction between electron and nuclear orbitals occurs. Such a state is usually known as a *vibronic* state, resulting from the electron-lattice coupling.

1.3.1 Degenerate systems

For a ground-state degenerate system, the total wavefunction can be expanded, for a fixed $\vec{R}$, as follows (Gross *et al.*, 1991):

$$\Psi_n(\vec{r}, \vec{R}) = \sum_m \phi_{nm}\, \psi_m(\vec{r}, \vec{R}) \; . \tag{1.3.19}$$

Inserting this representation of Ψ into the full Schrödinger equation and pre-multiplying all terms by ψ_k^* and integrating over the electron co-ordinates, we have

$$\left[\widehat{T}_i + \widehat{T}_{i-i} + E_k(\vec{R})\right]\, \phi_{nk}(\vec{R}) + \sum_m \phi_{nm}(\vec{R}) <\psi_k \,|\, \widehat{T}_i \,|\, \psi_m> \, = \mathcal{E}_n\, \phi_{nk}(\vec{R}) \; , \tag{1.3.20}$$

using Dirac's 'ket-bra' notation. The matrix elements $<\psi_k \,|\, \widehat{T}_i \,|\, \psi_m>$ are functions of $\vec{R}$, since we are integrating over electron co-ordinates. As a result, we have an infinite system of coupled equations.

The Born-Oppenheimer approximation consists of neglecting all the *off-diagonal* matrix elements $<\psi_k \,|\, \widehat{T}_i \,|\, \psi_m>$. In this approximation, (1.3.20) may be written as,

$$\left[\widehat{T}_i + \widehat{T}_{i-i} + E_k(\vec{R}) \; + <\psi_k \,|\, \widehat{T}_i \,|\, \psi_k>\right] \phi_{nk}(\vec{R}) = \mathcal{E}_n \, \phi_{nk}(\vec{R}) \; , \tag{1.3.21}$$

where the electron eigenvalues E_k and the matrix elements $<\psi_k|\, \widehat{T}_i \,|\psi_k>$ act as an effective potential for the ionic motion.

1.4 The one-electron potential

As a result of the adiabatic approximation the electron wavefunction should satisfy the equation

$$\left\{\widehat{T}_e + \widehat{V}_{e-e} + \widehat{V}_{e-i}\right\} \psi(\vec{r}, \vec{R}) = \mathcal{E}_e(\vec{R}) \, \psi(\vec{r}, \vec{R}) \; . \tag{1.4.1}$$

This equation, too, needs to be simplified. It should be first transformed into an equation for one particle, or a system of one-particle equations. It follows from (1.4.1) that this equation for an electronic system turns into a system of equations if it is assumed that the electrons do not interact ($\widehat{V}_{e-e} = 0$). Therefore the problem arises how to take account of the electron-electron interaction so that ultimately we would be able to deal with a system of non-interacting particles instead of interacting electrons. This is achieved through the introduction of the so-called *self-consistent electron field* (Hartree, 1928).

In order to solve the Schrödinger equation for the atom, Slater, following Hartree's intuitive approach for the atom, introduced an approximate electronic wavefunction to the many-body wavefunction (Slater, 1928)

$$\psi(\vec{r}_1, \sigma_1; \vec{r}_2, \sigma_2; \ldots \; ; \vec{r}_N, \sigma_N) = \prod_i^N \psi_i(\vec{r}_i, \sigma_i) \; . \tag{1.4.2}$$

Here the ψ_i are *one-electron* wavefunctions and σ_i the electronic spin.

Such a simple product function is the simplest conceivable function expressing the idea that the system state is characterised by a configuration specified by the one-electron ψ_i wavefunctions. Clearly, such a function is only approximate because it completely omits the *correlation* between the other electrons in the system, *i.e.*, the probability distribution of an electron at $\vec{r}_1$ is independent of wherever $\vec{r}_2, \ldots \, , \vec{r}_N$ might be at the time. In fact, the Coulomb repulsion will tend to keep the electrons away from each other in the true wavefunction. Since no allowance is made for that here, the energy of the Hartree wavefunction (1.4.2) must exceed the correct (and minimum) energy. However, in the space of all functions that can be written as simple products, there should be a 'best' one characterised by a set of functions $\psi_2, \ldots \, , \psi_N$ which give together the lowest energy. These functions are the so-called *Hartree functions*, which have been calculated numerically for many atoms.

In the spirit of Hartree, the electron-electron potential $\widehat{V}_{e-e}$ is replaced by an effective one-electron potential, V_i, seen by each electron and given as the average of the potential

produced by the *other* electrons ($i \neq j$) and the nucleus. Based on an intuitive argument—the action of all electrons on a particular electron cloud can be approximated by the electrostatic action of their averaged charge density—Hartree defined the one-electron potential, assumed spherical, as

$$V_i(\vec{r}_i) = \left\langle \sum_{j \neq i}^{N} \frac{1}{|\vec{r}_i - \vec{r}_j|} \right\rangle = \int \frac{n(\vec{r}_j)}{|\vec{r}_i - \vec{r}_j|} \, d^3 r_j \, . \tag{1.4.3}$$

Hartree introduced the concept of atomic electron density, $n(\vec{r}_j)$, defined as,

$$n(\vec{r}_j) = \sum_{j \neq i}^{\text{occ}} |\psi_j(\vec{r}_j)|^2 \, . \tag{1.4.4}$$

Equation (1.4.3) represents the average electrostatic interaction between the ith-electron and the 'ensemble' of all the electrons in the system. According to Hartree, the quantity $\psi_j^* \psi_j$ represents the contribution of the j-th orbital to the charge density of the atom, so that the total charge density arising from all the electrons in the system is given by (1.4.4).

The one-particle Hartree Hamiltonian, $\widehat{h}_{\mathrm{H}}$, can now be written as

$$\widehat{h}_{\mathrm{H}} \equiv -\frac{1}{2} \nabla_i^2 + \sum_{\alpha} \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} + V_i(\vec{r}) \, . \tag{1.4.5}$$

The total Hamiltonian is now a superposition of the one-electron Hamiltonians, $\widehat{h}_{\mathrm{H}}$. This implies that the total energy for the electronic system, $\mathcal{E}_e$, can be defined as the sum of one-electron energies of separate electrons. Analytically,

$$\mathcal{E}_e = \sum_{i}^{N} \varepsilon_i \, , \tag{1.4.6}$$

with ε_i satisfying,

$$\widehat{h}^{\mathrm{H}} \, \psi_i(\vec{r}_i, \sigma_i) = \varepsilon_i \, \psi_i(\vec{r}_i, \sigma_i) \, . \tag{1.4.7}$$

In the following section, we discuss the Hartree approximation in more detail.

1.5 The wavefunction variational principle

So far, we have applied simple concepts from classical electrostatics to analyse the many-body problem, following a simple physical reasoning. A well known principle of quantum mechanics provides us with a more accurate tool for the determination of the average value of an observable, like the the ground-state energy E_0.

The Schrödinger equation (1.2.1) is equivalent to a wavefunction variational principle: Minimise $<\Psi \,|\, \widehat{\mathcal{H}} \,|\, \Psi>$ subjected to the constraint $<\Psi \,|\, \Psi> = 1$, *i.e.*,

$$\delta \left\{ \frac{<\Psi \,|\, \widehat{\mathcal{H}} \,|\, \Psi>}{<\Psi \,|\, \Psi>} \right\} = 0 \tag{1.5.1}$$

The ground-state energy and wavefunction are found by minimising the expression in curly brackets, *i.e.*, the total-energy functional, $E[\Psi]$. Analytically,

$$E_0 = \inf_{\Psi} E[\Psi] \,. \tag{1.5.2}$$

This self-consistent procedure is quite simple but very effective. Starting from a trial wave function, written in terms of one-electron wavefunctions, *i.e.*, the trial parameters, the energy is minimised. The minimisation of the total energy with respect to these one-electron wavefunctions leads to a set of one-electron-like equations allowing the determination of these wavefunctions. Such a procedure is without doubt valid for the ground-state and sometimes for excited states.[3]

The unrestricted solution of $\delta E[\Psi] = 0$ is equivalent by the method of Lagrange multipliers to the solution of

$$\delta \left\{ <\Psi \,|\, \widehat{\mathcal{H}} \,|\, \Psi> - \; E <\Psi \,|\, \Psi> \right\} = 0 \,, \tag{1.5.3}$$

or

$$<\delta\Psi \,|\, (\widehat{\mathcal{H}} - E) \,|\, \Psi> \; = 0 \,. \tag{1.5.4}$$

Since $\delta\Psi$ is an arbitrary variation, we recover the Schrödinger equation. Every eigenstate of $\widehat{\mathcal{H}}$ is an extremum of $E[\Psi]$ and *vice-versa*.

The above principle implies the Hellmann-Feynman and virial theorems.

1.5.1 The Hellmann-Feynman theorem

Let λ be a parameter in the Hamiltonian $\widehat{\mathcal{H}}$ and Ψ_λ a corresponding normalised eigenstate. Then

$$\frac{\mathrm{d}E_\lambda}{\mathrm{d}\lambda} = 2 <\Psi_{\lambda'} \,|\, \widehat{\mathcal{H}} \,|\, \frac{\partial \Psi_\lambda}{\partial\, \lambda'}> + <\Psi_\lambda \,|\, \frac{\partial\, \widehat{\mathcal{H}}}{\partial \lambda} \,|\, \Psi_\lambda> \,, \tag{1.5.5}$$

where E_λ is defined as $<\Psi_\lambda \,|\, \widehat{\mathcal{H}} \,|\, \Psi_\lambda>$. The first term in (1.5.5) vanishes by the variational principle and we find the differential Hellmann-Feynman theorem

$$\frac{\mathrm{d}\, E_\lambda}{\mathrm{d}\lambda} = \; <\Psi_\lambda \,|\, \frac{\partial\, \widehat{\mathcal{H}}}{\partial \lambda} \,|\, \Psi_\lambda> \,. \tag{1.5.6}$$

Integrating from λ_1 to λ_2 gives

$$E_{\lambda_1} - E_{\lambda_2} = \int_{\lambda_1}^{\lambda_2} \frac{<\Psi_\lambda \,|\, (\partial\, \widehat{\mathcal{H}}/\partial\lambda) \,|\, \Psi_\lambda>}{<\Psi_\lambda \,|\, \Psi_\lambda>} \, d\lambda \,, \tag{1.5.7}$$

which is the integrated Hellmann-Feynman formula.

Applying the Hellmann-Feynman theorem, the force acting on nucleus α is

[3] This is the case where orthogonality with the ground-state is ensured, in most cases, by symmetry.

$$
\begin{aligned}
-\frac{\partial E}{\partial \vec{R}_\alpha} &= -<\Psi \,|\, \frac{\partial \widehat{\mathcal{H}}}{\partial \vec{R}_\alpha} \,|\, \Psi> \\
&= \int n(\vec{r})\, \frac{Z_\alpha\,(\vec{r}-\vec{R}_\alpha)}{\left|\vec{r}_i-\vec{R}_\alpha\right|^3}\, d^3r + \sum_{\beta\neq\alpha} \frac{Z_\alpha\, Z_\beta\,(\vec{R}_\alpha-\vec{R}_\beta)}{\left|\vec{R}_\alpha-\vec{R}_\beta\right|^3}\ .
\end{aligned}
\tag{1.5.8}
$$

This is the famous electrostatic theorem of Feynman (1939).

This result, which can be derived by applying pure classical electrostatics, can be used to find the equilibrium geometries of a solid by varying all the $\vec{R}_\alpha$ until the energy minimum is reached, *i.e.*, $-\partial E/\partial \vec{R}_\alpha = 0$.

It is important to note that although the Helmann-Feynman theorem (1.5.6) holds for any eigenstate of $\widehat{\mathcal{H}}$, the electrostatic formula (1.5.8) is only true for the ground state.

1.5.2 The virial theorem

The virial theorem relates the kinetic energy and potential energy components of the total energy under certain circumstances.

Let $\psi(\vec{r}_1,\dots,\vec{r}_N)$ by any extremum of $<\Psi\,|\,\widehat{\mathcal{H}}\,|\,\Psi>$ over normalised wavefunctions, *i.e.*, any eigenstate or optimised trial wavefunction. For any scale parameter $\gamma>0$, define the uniformly-scaled wavefunction

$$
\psi_\gamma(\vec{r}_1,\dots,\vec{r}_N) = \gamma^{3N/2}\,\psi_\gamma(\gamma\,\vec{r}_1,\dots,\gamma\,\vec{r}_N)\ , \tag{1.5.9}
$$

and

$$
<\psi_\gamma\,|\,\psi_\gamma> \;=\; <\psi\,|\,\psi> \;=1\ . \tag{1.5.10}
$$

The density corresponding to the scaled wavefunction is the scaled density

$$
n_\gamma(\vec{r}) = \gamma^3\, n(\vec{r})\ , \tag{1.5.11}
$$

which clearly conserves the electron number

$$
\int n_\gamma(\vec{r})\, d^3r = \int n(\vec{r})\, d^3r = N\ . \tag{1.5.12}
$$

The condition $\gamma > 1$ leads to densities n_γ that are higher (on average) and more 'contracted' than n, while for $\gamma < 1$ produces densities that are lower and more 'expanded'. The main effect of the scaling is a contraction, or expansion, of the electron cloud without changing the normalisation for $\gamma > 1$ and $\gamma < 1$ respectively.

Now consider what happens to $<\widehat{\mathcal{H}}> = <\widehat{T}+\widehat{V}>$ under scaling. By definition,

$$
\frac{\mathrm{d}}{\mathrm{d}\gamma}\, <\psi_\gamma\,|\,(\widehat{T}+\widehat{V})\,|\,\psi_\gamma>\Big|_{\gamma=1} = 0 \tag{1.5.13}
$$

The kinetic energy component $\widehat{T}$ is homogeneous of degree -2 in electron coordinates $\vec{r}$, so

$$
<\psi_\gamma\,|\,\widehat{T}\,|\,\psi_\gamma> = \gamma^2\,<\psi\,|\,\widehat{T}\,|\,\psi>\ , \tag{1.5.14}
$$

and (1.5.13) becomes

$$2 < \psi \,|\, \widehat{T} \,|\, \psi > + \frac{\mathrm{d}}{\mathrm{d}\gamma} < \psi_\gamma \,|\, \widehat{V} \,|\, \psi_\gamma > \Big|_{\gamma=1} = 0 \;, \tag{1.5.15}$$

or

$$2 < \widehat{T} > - < \sum_{i=1}^{N} \vec{r}_i \cdot \frac{\partial \widehat{V}}{\partial \vec{r}_i} > = 0 \;. \tag{1.5.16}$$

Since the potential is homogeneous of degree 1, (1.5.15) can now be written as

$$2 < \psi \,|\, \widehat{T} \,|\, \psi > - < \psi \,|\, \widehat{V} \,|\, \psi > = 0 \;. \tag{1.5.17}$$

1.6 The Hartree self-consistent field method

Due to the importance of the Hartree method, it is convenient to deduce and analyse the Hartree equations in some detail. Slater (1930) was the first to apply the wavefunction variational principle to minimise the total energy, going beyond Hartree's intuitive formulation.

If we multiply (1.4.1) by ψ_i^*, using (1.4.2), and integrate over electron co-ordinates, we have for the i-th electron, neglecting spin for sake of simplicity,

$$\begin{aligned}
\mathcal{E}_e = \sum_i^N \Bigg\{ & \left(\int \psi_i^*(\vec{r}_i) \left[-\frac{1}{2} \nabla_i^2 + \sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} \right] \psi_i(\vec{r}_i)\, d^3 r_i \right) \\
\times & \left(\int \psi_1^*(\vec{r}_1)\, \psi_1(\vec{r}_1)\, d^3 r_1 \right) \cdots \left(\int \psi_N^*(\vec{r}_N)\, \psi_N(\vec{r}_N)\, d^3 r_N \right) \\
+ & \frac{1}{2} \sum_{\substack{j \\ j \neq i}} \left(\iint \frac{|\psi_i(\vec{r}_i)|^2\, |\psi_j(\vec{r}_j)|^2}{|\vec{r}_i - \vec{r}_j|}\, d^3 r_i\, d^3 r_j \right) \\
\times & \left(\int \psi_1^*(\vec{r}_1)\, \psi_1(\vec{r}_1)\, d^3 r_1 \right) \cdots \left(\int \psi_N^*(\vec{r}_N)\, \psi_N(\vec{r}_N)\, d^3 r_N \right) \Bigg\} \;.
\end{aligned} \tag{1.6.1}$$

Here, the terms involving only the co-ordinates of the i-th atom were separated from those involving both, i- and j-th atom co-ordinates; the two first terms in (1.6.1) only involve the co-ordinates of the i-th atom.

By means of the condition of orthogonality, $\int \psi_i^* \, \psi_i \, d^3 r = 1$, (1.6.1) and (1.4.4), may be simplified to,

$$\begin{aligned}
\mathcal{E}_e = \sum_i \Bigg\{ & \int \psi_i^*(\vec{r}_i) \left[-\frac{1}{2} \nabla_i^2 + \sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} \right] \psi_i(\vec{r}_i)\, d^3 r_i \\
& + \frac{1}{2} \iint \frac{n(\vec{r}_i)\, n(\vec{r}_j)}{|\vec{r}_i - \vec{r}_j|}\, d^3 r_i\, d^3 r_j \Bigg\} \;,
\end{aligned} \tag{1.6.2}$$

or

$$\mathcal{E}_e = \sum_i \left\{ \int \psi_i^*(\vec{r}_i) \left(-\frac{1}{2} \nabla_i^2 \right) \psi_i(\vec{r}_i)\, d^3 r_i + \int |\psi_i(\vec{r}_i)|^2 \sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|}\, d^3 r_i + U^{\mathrm{H}} \right\} \;, \tag{1.6.3}$$

with

$$U_{\mathrm{H}} \equiv \frac{1}{2} \iint \frac{n(\vec{r}_i)\, n(\vec{r}_j)}{|\vec{r}_i - \vec{r}_j|}\, d^3r_i\, d^3r_j \tag{1.6.4}$$

defining the Hartree energy, U_{H}.

The simplest way to minimise the total energy following the variational principle is to build a new functional Φ,

$$\Phi = \mathcal{E}_e - \sum_i \sum_j \varepsilon_{ij} \int \psi_i^*\, \psi_j\, d^3r = \mathcal{E}_e - \sum_i \varepsilon_i \int |\psi_i|^2\, d^3r\ . \tag{1.6.5}$$

Here, we have used the fact that the parameters ε_{ij} are Hermitian. Then, the variational problem can be posed as,

$$\frac{\delta\Phi}{\delta\psi_i^*} = \frac{\delta}{\delta\psi_i^*}\, \mathcal{E}_e - \frac{\delta}{\delta\psi_i^*} \left(\sum_i \varepsilon_i \int |\psi_i(\vec{r}_i)|^2\, d^3r \right) = 0\ . \tag{1.6.6}$$

The solution of the previous equation yields the famous set of equations called Hartree's equations. These can be written as,

$$\left\{ -\frac{1}{2}\nabla_i^2 + \sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} + \sum_{j\neq i} \int \frac{|\psi_j(\vec{r}_j)|^2}{|\vec{r}_i - \vec{r}_j|}\, d^3r_j \right\} \psi_i(\vec{r}_i) = \varepsilon_i\, \psi_i(\vec{r}_i)\ , \tag{1.6.7}$$

with i running from 1 to N.

This famous set of integro-differential equations are called *Hartree's equations*. One immediate difficulty is caused by the fact that the self-consistent field $\widehat{V}_i$—probably one of the most important concepts in solid-state computational physics—can only be evaluated if the ψ_j are known, and *vice-versa*. The one-electron potential $\widehat{V}_i$, now renamed $\widehat{V}^{\mathrm{H}}$ for historical reasons, can now be expressed as

$$\widehat{V}_i(\vec{r}_i) \equiv \widehat{V}^{\mathrm{H}}(\vec{r}_i) = \sum_{j\neq i} \int \frac{\psi_j^*(\vec{r}_j)\, \psi_j(\vec{r}_j)}{|\vec{r}_i - \vec{r}_j|}\, d^3r_j = \int \frac{n(\vec{r}_j)}{|\vec{r}_i - \vec{r}_j|}\, d^3r_j\ . \tag{1.6.8}$$

The above expressions show clearly why the solutions of (1.6.7) have to proceed by successive approximations. This self-consistent process is as follows:

1. Take some $\psi_i^{(0)}(\vec{r}_j)$—the trial function—as a zero approximation, and evaluate the one-electron potential $\widehat{V}^{\mathrm{H}}_{(0)}(\vec{r}_i)$;

2. Substitute $\widehat{V}^{\mathrm{H}}_{(0)}(\vec{r}_i)$ into (1.6.7) to obtain the first approximation for the one-electron wave-function, $\psi_i^{(1)}(\vec{r}_j)$. Use this to re-calculate the potential, $\widehat{V}^{\mathrm{H}}_{(1)}(\vec{r}_i)$;

3. Repeat the process until the $(n+1)$th approximation coincides with the nth approximation within the limits of a predetermined error.

Finding the solutions of Hartree's equation is nothing else than a mere problem of electrostatics. One difficulty is the fact that each effective one-electron Schrödinger equation contains the unknown wavefunctions in the charge density term.

1.7 Hartree-Fock theory

The shortcoming of the Hartree scheme is that it takes no account of the Pauli exclusion principle. The introduction of this principle, *i.e.*, the inclusion of exchange effects, turns it into the Hartree-Fock scheme (Fock, 1930).

The Pauli principle requires the electron wavefunction to be antisymmetric under the interchange of any two fermions. At the same time $\prod \psi_i(\vec{r}_i, \sigma_i)$ (1.4.2) does not satisfy this requirement. The total wavefunction is written in the form of an antisymmetrised product of one-electron wavefunctions,

$$\psi_e(\vec{r}, \sigma) = \frac{1}{\sqrt{N!}} \widehat{A}(N) \left[\psi_1(\vec{r}_1, \sigma_1)\, \psi_2(\vec{r}_2, \sigma_2) \cdots \psi_N(\vec{r}_N, \sigma_N)\right] \tag{1.7.1}$$

The antisymmetrising operator $\widehat{A}(N)$ is usually defined as (Fulde, 1995),

$$\widehat{A}(N) = \sum_\nu \xi_\nu\, \tau_\nu\,, \tag{1.7.2}$$

where the sum is over all permutations τ_ν of the electrons and ξ_ν is the parity of τ_ν. Consequently, ξ_ν is 1 if the permutation τ_ν involves an even number of electrons and -1 for an odd number of electrons.

Thereafter, we will only consider *closed-shell* systems, which are characterised by doubly occupied orbitals ψ_i only, *i.e.*, when $\psi_\lambda(\vec{r}_i, \alpha)$ is occupied so is $\psi_\lambda(\vec{r}_i, \beta)$. Thus, there is a complete symmetry with respect to spin-up and spin-down electrons.

The right combination of N one-electron wavefunctions takes the form of a *Slater determinant*,

$$\psi(\vec{r}, \sigma) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\vec{r}_1, \sigma_1) & \cdots & \psi_N(\vec{r}_1, \sigma_1) \\ \vdots & \ddots & \vdots \\ \psi_1(\vec{r}_N, \sigma_N) & \cdots & \psi_N(\vec{r}_N, \sigma_N) \end{vmatrix} \equiv \frac{1}{\sqrt{N!}} \det |\psi_\lambda(\vec{r}_i, \sigma_i)| \tag{1.7.3}$$

with,

$$\psi_\lambda(\vec{r}_i, \sigma_i) = \phi_\lambda(\vec{r}_i)\, \chi_\sigma \tag{1.7.4}$$

The subscript λ represents the quantum state. The dimensionless quantity χ_σ is a spin-function, with components $\alpha = \binom{1}{0}$ for spin-up and $\beta = \binom{0}{1}$ for spin-down electrons.

The spin-independent orbitals ϕ_λ satisfy

$$\int \phi_\lambda^*(\vec{r})\, \phi_\mu(\vec{r})\, d^3r = \delta_{\lambda\mu}\,. \tag{1.7.5}$$

The single-particle wavefunctions (or *spin orbitals*, after Slater (1974)), are formally obtained from the equation,

$$\widehat{h}_{\mathrm{HF}}\, \psi_\lambda(\vec{r}_i, \sigma_i) = \varepsilon_i\, \psi_\lambda(\vec{r}_i, \sigma_i) \tag{1.7.6}$$

However, so long as we do not know the exact form of the operator $\widehat{h}_{\mathrm{HF}}$, we cannot, of course, calculate single particle orbitals from the equation above.

Applying the Ritz variational principle,[4] it is possible to overcome this difficulty by optimising the *ansatz* of single-particle orbitals. The variational problem can be posed as follows,

$$\frac{\delta}{\delta\psi_\lambda^*}\left[<\psi_\lambda\,|\,\widehat{H}\,|\,\psi_\lambda> - \sum_{\lambda>\mu}^{N}\varepsilon_{\lambda\mu}\int\psi_\lambda^*(\vec{r},\sigma)\,\psi_\mu(\vec{r},\sigma)\,d^3r\right] = 0 \tag{1.7.7}$$

The Lagrange matrix elements $\varepsilon_{\lambda\mu}$ arise from the constraint that the spin orbitals must be normalised.

The normalisation integral $<\psi_\lambda\,|\,\psi_\lambda>$ is equal to unity and the energy expectation value is found to be given by

$$E = \sum_\lambda <\lambda\,|\,(\widehat{T}_e + \widehat{V}_{e-i})\,|\,\lambda> + \frac{1}{2}\sum_\lambda\sum_\mu\Big(<\lambda\mu\,|\,\widehat{V}_{e-e}\,|\,\lambda\mu> - <\lambda\mu\,|\,\widehat{V}_{e-e}\,|\,\mu\lambda>\Big) \tag{1.7.8}$$

The second and third terms are sometimes called *direct* and *exchange* integrals (Gross *et al.*, 1991). It is now advantageous to introduce a notation customary in quantum chemistry, where these integrals are referred to as $J_{\lambda\mu}$ and $K_{\lambda\mu}$, respectively.[5]

The summation $\frac{1}{2}\sum_{\lambda,\mu}<\lambda\mu\,|\,\widehat{V}_{e-e}\,|\,\mu\lambda> \equiv \sum_{\lambda,\mu}K_{\lambda\mu}$ defines the *exchange* energy

$$E_x = \frac{1}{2}\sum_\lambda\sum_\mu \delta_{\sigma_\lambda\sigma_\mu}\,\delta_{\sigma_\tau\sigma_\rho}\iint \phi_\lambda^*(\vec{r})\,\phi_\tau^*(\vec{r}\,')\,\frac{1}{|\vec{r}-\vec{r}\,'|}\,\phi_\mu(\vec{r})\,\phi_\rho(\vec{r}\,')\,d^3r\,d^3r' \tag{1.7.9}$$

The Kronecker delta pre-factors, $\delta_{\sigma_\lambda\sigma_\mu}$ ensure that the matrix elements for which ϕ_λ and ϕ_μ have different spins are zero.

What is the physical meaning of the exchange energy E_x? If one is to take the Pauli principle into account, two electrons with parallel spins cannot be found at the same point in space. As a consequence, the average distance, separating the two electrons, will be greater and the electrostatic repulsion energy lower by a quantity corresponding to the exchange energy. The exchange integral, $K_{\lambda\mu}$, has no classical counterpart and should be regarded as a quantum mechanical correction to the Coulomb integral $J_{\lambda\mu}$.

It is now possible to solve the variational problem posed by (1.7.7). The result of the minimisation may be expressed as,

$$\widehat{F}_\lambda\,\psi_\lambda(\vec{r},\sigma) = \sum_{\mu\neq\lambda}^{N}\varepsilon_{\lambda\mu}\,\psi_\mu(\vec{r},\sigma) \tag{1.7.10}$$

$\widehat{F}_\lambda$ is the Fock operator assuming the following form,

$$\widehat{F}_\lambda \equiv \widehat{T}_e + \widehat{V}_{e-i} + \widehat{g}\,. \tag{1.7.11}$$

The Hartree-exchange operator $\widehat{g}(\vec{r},\vec{r}\,')$ is defined as

$$\widehat{g} = \widehat{V}^{\mathrm{H}}(\vec{r}) + \widehat{V}_\lambda^x(\vec{r},\vec{r}\,')\,. \tag{1.7.12}$$

[4] The trial wavefunction is defined as a linear combination of one-particle wavefunctions, and the standard variational principle (1.5.1) is applied.

[5] An important equality applies here: $K_{\lambda\lambda} = J_{\lambda\lambda}$, with $K_{\lambda\mu}, J_{\lambda\mu} \geqslant 0$.

$\widehat{V}^{\rm H}$ and $\widehat{V}^x_\lambda$ are the Hartree- and exchange-potential operators, defined as functional derivatives of the Hartree and exchange energies, respectively:

$$\widehat{V}^{\rm H}(\vec{r})\,\psi_\lambda(\vec{r},\sigma) = \frac{\delta U^{\rm H}}{\delta \phi^*_\lambda} = \int \frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,\psi_\lambda(\vec{r},\sigma)\,d^3r' \tag{1.7.13}$$

and,

$$\widehat{V}^x_\lambda(\vec{r},\vec{r}\,')\,\psi_\lambda(\vec{r},\sigma) = \frac{\delta E_x}{\delta \phi^*_\lambda} = -\sum_{\mu\neq\lambda}\delta_{\sigma\sigma_\lambda}\int \frac{\phi_\mu(\vec{r}\,')\,\phi^*_\lambda(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,\psi_\mu(\vec{r},\sigma)\,d^3r' \tag{1.7.14}$$

with,

$$n(\vec{r}\,') = \sum_{\mu}^{N/2} |\phi_\mu(\vec{r}\,')|^2\,. \tag{1.7.15}$$

Again, the factor $\delta_{\sigma\sigma_\lambda}$ implies that only those spin orbitals ψ_μ which have spin σ contribute. In (1.7.15), the summation is performed over occupied states only, *i.e.*, $N/2$ for the the present case of a closed-shell system.

It important to note that the exchange potential is a *non-local* operator. In order to show this unambiguously, let us write the HF equations (1.7.10) explicitly as,

$$\left(\widehat{T}_e + \widehat{V}_{e-i} + \int \frac{n(\vec{r}\,') - n^x_\lambda(\vec{r},\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,d^3r'\right)\,\phi_\lambda(\vec{r}) = \varepsilon_\lambda\,\phi_\lambda(\vec{r}) \tag{1.7.16}$$

The non-local exchange density term, $n^x_\lambda(\vec{r},\vec{r}\,')$, is

$$n^x_\lambda(\vec{r},\vec{r}\,') = \sum_{\mu\neq\lambda}\frac{\phi^*_\lambda(\vec{r})\,\phi_\mu(\vec{r})}{|\phi_\lambda(\vec{r})|^2}\,\phi^*_\mu(\vec{r}\,')\,\phi_\lambda(\vec{r}\,') \tag{1.7.17}$$

The non-local character of the exchange density complicates considerably finding a self-consistent solution of the HF equations. From a computational viewpoint, this is confirmed by the fact that the evaluation of the exchange energy term involves a four-centre integral.

Integration of n^x_λ over r', upon using orthogonality of ϕ_λ and ϕ_μ, with $\lambda \neq \mu$, shows that this exchange charge density contains exactly one electron,

$$\int n^x_\lambda(\vec{r},\vec{r}\,')\,d^3r' = -1 \tag{1.7.18}$$

Further, at $\vec{r} = \vec{r}\,'$,

$$n^x_\lambda(\vec{r},\vec{r}) = \sum_{\mu\neq\lambda}|\phi_\mu(\vec{r})|^2 \tag{1.7.19}$$

That is, it exactly subtracts out the charge density of all electrons with spin parallel to that at a λ-th state and located at $\vec{r} = \vec{r}\,'$. At a crude approximation, it will have the effect of removing all the parallel spin-charge from a sphere centred at $\vec{r}$ and of radius R such that

$$\int n^x_\lambda(\vec{r},\vec{r})\,d^3r \simeq \frac{4\,\pi}{3}\sum_{\mu\neq\lambda}|\phi_\mu(\vec{r})|^2\,R^3 = 1 \tag{1.7.20}$$

In other words, each electron moves in the field of the nuclei, the electrons of opposite spin, and those of parallel spin outside an *exchange-hole*, or 'Fermi-Coulomb hole' (of radius R), which follows it around wherever it goes.

Multiplying (1.7.10) by $\psi^*_\lambda(\vec{r})$, and integrating, one obtains the formula for *orbital energies*,

$$\varepsilon_\lambda \equiv \varepsilon_{\lambda\mu} = <\psi_\lambda \,|\, \widehat{F} \,|\, \psi_\lambda> \;=\; <\lambda| \left(\widehat{T}_e + \widehat{V}_{e-i}\right) |\lambda> + \frac{1}{2}\sum_{\mu}^{N}(J_{\lambda\mu} - K_{\lambda\mu}) \tag{1.7.21}$$

Summing over λ and comparing with (1.7.8), we get, including ion-ion repulsive interaction

$$E_{\mathrm{HF}} = \sum_{\lambda} \varepsilon_\lambda - E_{e-e} + E_{i-i} \tag{1.7.22}$$

Note that E_{HF} is not equal to the sum of the orbital energies. $\widetilde{V}_{e-e}$ is now the total electron-electron potential (including exchange):

$$E_{e-e} = \frac{1}{2}\sum_{\lambda>\mu}(J_{\lambda\mu} - K_{\lambda\mu}) = U^{\mathrm{H}} - E_x\,. \tag{1.7.23}$$

1.7.1 Koopmans' theorem

The non-linear Hartree-Fock equations have the form of one-particle Schrödinger equations, with an effective non-local potential. The correspondent one-electron eigenvalues ε_λ were formally introduced as Lagrange parameters, which do not have *per se* any physical meaning. These were simply introduced to account for the normalisation conditions of the wavefunction with respect to which the stationary point of the energy is found.

Koopmans' theorem (1933) attaches a physical meaning to those eigenvalues, stating that they are the ionisation potentials of the system for which the HF equations have been set up. Analytically,

$$-I_\lambda = E_{\mathrm{HF}}(N - 1_\lambda) - E_{\mathrm{HF}}(N) = \varepsilon_\lambda\,. \tag{1.7.24}$$

I_λ is the ionisation potential of an electron at a level λ and 1_λ simply means that the electron is extracted from the one-electron level λ.

The above expression, in which the ionisation potential is identified as a finite difference, implies a severe approximation. It is assumed that the removal of the electron in state λ from the system does not influence the wavefunctions of the remaining electrons.

Effectively, we calculate the energy of the perturbed systems, with $(N - 1)$ particles, using the unperturbed wavefunctions of the initial system containing N particles. In other words, the $E_{\mathrm{HF}}(N - 1_\lambda)$ and $E_{\mathrm{HF}}(N)$ are calculated from the *same* set of spin orbitals, obviously ignoring the consequent relaxation process to compensate the lack of *screening* due to removal of an electron.

1.8 The local exchange potential: The X_α method

It has been pointed out in this thesis that one of the major difficulties on applying the HF scheme arises from the non-local character of the exchange potential. Slater (1972) has

suggested replacing this non-local potential by a local one, a step carried out by means of the X_α method.

This method assumes that the exchange potential at a certain point in space $\vec{r}_0$ of a given system, characterised by a density distribution $n(\vec{r})$ is that of a homogeneous electron gas with density $n \equiv n(\vec{r}_0)$.

This is a reasonable approximation as long as $n(\vec{r})$, as a function of $\vec{r}$, varies sufficiently slowly, *i.e.*, as long as

$$\left| \frac{\vec{\nabla}\, n(\vec{r})}{n(\vec{r})} \right| \ll k_{\mathrm{F}}(n(\vec{r}))\ . \tag{1.8.1}$$

The local Fermi momentum is given by $k_{\mathrm{F}} = \sqrt[3]{(3\pi)^2\, n}$.

For the case of a homogeneous electron gas, the solutions of the HF equations assume the form of plane waves,

$$\psi_\lambda(\vec{r}, \sigma) = \frac{1}{\sqrt{\Omega}}\ \mathrm{e}^{i\,\vec{k}\cdot\vec{r}}\ \ \sigma\ . \tag{1.8.2}$$

Here Ω is the volume of the system and λ refers to the wavevector $\vec{k}$ and spin state σ. The energy levels are:

$$E_k = \frac{1}{2}k^2 + V^x_{\vec{k}} \equiv E_{\vec{k},\sigma}\ , \tag{1.8.3}$$

where the exchange potential $V^x_{\vec{k}}$ is given by,

$$V^x_{\vec{k}}(\vec{r}) = -\sum_{|\vec{k}'|<k_{\mathrm{F}}}^{\mathrm{occ}} \int \frac{\mathrm{e}^{i\,(\vec{k}-\vec{k}')\cdot(\vec{r}'-\vec{r})}}{\Omega\,|\vec{r}-\vec{r}'|}\, d^3r'\ . \tag{1.8.4}$$

This is in fact independent of $\vec{r}$, depending only on the magnitude k of $\vec{k}$ alone. The sum is over all occupied states with spin σ. Analytically,

$$V^x_{\vec{k}} = -\sum_{|\vec{k}'|<k_{\mathrm{F}}}^{\mathrm{occ}} \frac{4\pi}{\Omega\,|\vec{k}-\vec{k}'|} = -\frac{1}{8\,\pi^3} \int_{|\vec{k}'|<k_{\mathrm{F}}} \frac{4\pi}{|\vec{k}-\vec{k}'|}\, d^3k\,. \tag{1.8.5}$$

The solution of the above integral is, replacing k_{F} by its explicit form

$$V^x_{\vec{k}} = -4\,F(\eta)\,\left(\frac{3\,n(\vec{r})}{8\,\pi}\right)^{\frac{1}{3}}\ , \tag{1.8.6}$$

where η defined as k/k_{F}, equals

$$F(\eta) = \frac{1}{2} + \frac{1-\eta^2}{4\,\eta}\,\ln\left(\frac{1+\eta}{1-\eta}\right)\ . \tag{1.8.7}$$

The function $F(\eta)$, known as the Linhard function, tends to 1 when $\eta \to 0$, and to $\frac{1}{2}$ as $\eta \to 1$. Its derivative has a weak singularity as η tends to 1. This fact has catastrophic implications for the applicability of this theory to simple metals.

We obtain Slater's form of a local exchange potential by averaging $F(\eta)$ over all occupied states, *i.e.*, the Fermi sphere. This results in

$$F_{\mathrm{av}} \equiv \frac{3}{4\,\pi\,k_{\mathrm{F}}^3} \int\limits_{|\vec{k}|<k_{\mathrm{F}}} F(\eta)\, d^3k = 3 \tag{1.8.8}$$

Replacing k_{F} by $k_{\mathrm{F}}(n(\vec{r}))$, we derive from (1.8.6) the averaged quantity,

$$V_{\mathrm{av}}^{x}(n(\vec{r})) = -\frac{3}{2\,\pi}\, k_{\mathrm{F}}(n(\vec{r})) = -\frac{3}{2\,\pi}\,[3\,\pi^2\, n(\vec{r})]^{\frac{1}{3}} \,. \tag{1.8.9}$$

This is the required local approximation for the exchange potential.

Adopting a more phenomenological point of view, we multiply F_{av} by an adjustable parameter α which is fitted to observable quantities. The X_α potential which replaces the non-local exchange potential can then be expressed as

$$V_{X_\alpha}(n(\vec{r})) = -3\,\alpha\,\left(\frac{3}{8\,\pi}\right)^{\frac{1}{3}}\, n(\vec{r})^{\frac{1}{3}} \,. \tag{1.8.10}$$

1.9 Shortcomings of the one-electron approximation

The one-electron approximation implies that an electron interacts with other electrons in the system according to their *average* locations. In reality, however, the electronic motion occurs according to the other electrons' *actual* positions. The Coulomb repulsion becomes sufficiently reduced only if a correlated motion of the electronic system takes place.

What is missing is the correlation-hole energy that the electron carries along to prevent other electrons from coming too close, which would provoke a considerable potential energy loss. The difference between the exact N-electron wavefunction and its HF counterpart is therefore related to the correlation aspect of the electron motion. The correlation contribution, E_c^{HF}, is then defined as,

$$E_c^{\mathrm{HF}} = E - E_{\mathrm{HF}} \,. \tag{1.9.1}$$

1.10 The classical density functional models

1.10.1 Introduction

In previous sections of this thesis, we have described some approximations that allow us to to solve the full Schrödinger equation by decoupling the electronic and nuclear motions and reducing the many-body problem to a single-particle one. Despite this, analytical solutions are only possible for a few very simple systems, and numerically exact solutions can only be found for a small number of atoms or molecules. Large systems, like a solid, are still out of reach.

The appealing picture resulting from the Hartree-Fock scheme is so familiar that it is sometimes overlooked that the approximation behind it is not a particularly good one. The lowest-lying configuration is generally only one of very many which are almost degenerate

in energy, which suggests that a better one would result from taking a linear combination of those configurations. Such an approach, where effects beyond the HF approximation (*correlation* effects) are included by improving the many-particle wavefunction, is known as *configuration interaction* (CI).

The CI correction leads, in principle, to the exact wavefunction from which most properties of interest can be calculated. Unfortunately, due to the fact that the number of configurations *explodes* with the number of electrons considered, only systems with a very small number of electrons can be studied with high accuracy. Furthermore, the complexity of the resulting solutions means that a simple interpretation of the results is often difficult.

Thomas (1927) and Fermi (1928) proposed a new scheme based on the density of electrons in the system, $n(\vec{r})$. This original idea is remarkable in the sense that it allows us to replace the complicated N-electron wavefunction and the associated Schrödinger equation with the simpler electron density.

This model assumes that the electronic properties of a real system are similar to those of a *gas* of fermions, *i.e.*, a classical ensemble of non-correlated particles. Although this approach has had a limited success in reproducing the properties of real systems, it is a prototype for later density functional theories in its focus on the density, *i.e.*, the integrated quantity,

$$n(\vec{r}') = N \int d^3r_2 \cdots \int \mathrm{d}^3r_N \, \Psi^*(\vec{r}_1, \ldots, \vec{r}_N) \times \Psi(\vec{r}_1, \ldots, \vec{r}_N) \,, \tag{1.10.1}$$

which is easier to describe than the precise details of the total wavefunction Ψ. This model also provides a well-defined model whose mathematical properties have received much attention (Lieb, 1981).

1.10.2 The model of Thomas and Fermi

We are interested in finding the electronic properties of a system described by the following Hamiltonian:

$$\widehat{\mathcal{H}} \equiv -\frac{1}{2}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{j\neq i} \frac{1}{|\vec{r}_i - \vec{r}_j|} + v(\vec{r}_i) \,, \tag{1.10.2}$$

with

$$v(\vec{r}_i) = -\sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} \,, \tag{1.10.3}$$

an *external* potential defining the interaction between the electrons and the nuclear frame.

In the approximation of Thomas and Fermi (TF), the total electron-electron interaction energy, including non-classical terms, arises solely from the Coulombic interactions between particles within a classical charge distribution. The quantity

$$U_{\mathrm{H}}[n] = \frac{1}{2} \iint \frac{n(\vec{r})\, n(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d^3r \, d^3r' \,, \tag{1.10.4}$$

a Hartree electrostatic self-repulsion of the electron density, replaces the total electron-electron interaction (the second term in right side of (1.10.2)), which includes non-classical

terms. Re-calling the kinetic term in (1.10.2), the kinetic energy is given by

$$T[n] = \int t[n(\vec{r})]\, d^3r \,, \tag{1.10.5}$$

where $t[n]$ is the kinetic energy density for a system of *non-interacting* electrons with density n. It is assumed that $n(\vec{r})$ varies so slowly in space that an electron at a point $\vec{r}$ sees an essentially homogeneous medium with density $n(\vec{r})$.[6]

The kinetic energy density is derived from the corresponding functional for the homogeneous electron gas. We have,

$$t[n] = \frac{2}{(2\pi)^3} \int\limits_{|\vec{k}|<k_{\mathrm{F}}} \frac{1}{2}\, k^2\, d^3k \,, \tag{1.10.6}$$

where k_{F} is the Fermi momentum and

$$n = \frac{1}{(3\pi)^2}\, k_{\mathrm{F}}^3 = \frac{3}{4\,\pi\, r_s^3} \,, \tag{1.10.7}$$

where r_s is the Wigner-Seitz radius defined as the radius of a sphere whose volume is the effective volume of an electron. This gives

$$T_0[n] = C_{\mathrm{F}} \int n(\vec{r})^{\frac{5}{3}}\, d^3r \,, \tag{1.10.8}$$

where C_{F} is $\frac{3}{10}(3\pi^2)^{\frac{2}{3}} = 2.871$. This equation shows that in this theory the kinetic energy is determined explicitly by the density $n(\vec{r})$.

Equation (1.10.8) is the famous TF kinetic energy functional which Thomas and Fermi dared to apply to electrons in atoms and molecules. Here, we first encounter one of the most important ideas in modern DFT: the *local density approximation* (LDA). In this approximation, electronic properties are determined as functionals of the electron density by applying, *locally*, relations appropriate for a homogeneous electronic system.

The total electronic energy, $E_{\mathrm{TF}}[n;v]$, can now be written as

$$E_{\mathrm{TF}}[n;v] = \int v(\vec{r})\, n(\vec{r})\, d^3r + U_{\mathrm{H}}[n(\vec{r})] + C_{\mathrm{F}} \int n(\vec{r})^{\frac{5}{3}}\, d^3r \,. \tag{1.10.9}$$

It is important to note that E_{TF} is a functional of v, the potential energy of the nuclear framework—the external potential—as well as of the electronic density n. Note that the expression of E_{TF} does not contain any exchange or correlation terms.

Imposing the constraint that the total number of electrons remains constant, the ground-state density $n(\vec{r})$ is determined by minimising (1.10.9). Analytically, the Euler-Lagrange equation is

$$\delta \left\{ E_{\mathrm{TF}}[n;v] - \mu \int n(\vec{r})\, d^3r \right\} = 0 \,, \tag{1.10.10}$$

where μ is a Lagrange parameter playing the role of a chemical potential. This is evident from the form of the variational equation,

$$\delta(E_{\mathrm{TF}} - \mu\, N) = 0 \quad \Rightarrow \quad \mu = \frac{\delta E_{\mathrm{TF}}}{\delta N} \,, \tag{1.10.11}$$

[6] It is important to note that a molecular electron cloud, the foundation of a crystal, is *not* obviously a uniform gas.

assuming the constancy of μ.

If we bring (1.10.9) in, the solution of the variational problem can by expressed as,

$$\mu = \frac{5}{3}\,C_{\mathrm{F}}\,n(\vec{r})^{\frac{2}{3}} + v(\vec{r}) + \int \frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,d^3r'\ , \tag{1.10.12}$$

which is the basic relation between electron density $n(\vec{r})$ and potential energy $v(\vec{r})$ of the Thomas-Fermi theory.

Replacing the value of C_{F} in the above equation, and denoting v_i as the electrostatic, or *internal*, potential produced by the electronic density,

$$v_i(\vec{r}) \equiv \int \frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,d^3r'\ , \tag{1.10.13}$$

the electron density can be written as

$$n(\vec{r}) = \frac{1}{3\,\pi^2}\left\{2\left[\mu - v(\vec{r}) - v_i(\vec{r})\right]\right\}^{\frac{3}{2}}\ , \tag{1.10.14}$$

for $\mu \geqslant v(\vec{r}) + v_i(\vec{r})$. One way of determining the form of $v_i(\vec{r})$ is to relate this quantity to the electron density $n(\vec{r})$ (1.10.14), through the Poisson's equation

$$\nabla^2\, v_i(\vec{r}) = -4\,\pi\, n(\vec{r})\ . \tag{1.10.15}$$

From (1.10.14) and (1.10.15) we can eliminate $n(\vec{r})$ and arrive at the Thomas-Fermi equation

$$\nabla^2\, v_i(\vec{r}) = -\frac{2^{\frac{7}{2}}}{3\pi}\left[\mu - v(\vec{r}) - v_i(\vec{r})\right]^{\frac{3}{2}}\ . \tag{1.10.16}$$

The form of the above equation reveals the attractiveness of the Thomas-Fermi model. It allows the determination of $n(\vec{r})$ directly from $v(\vec{r})$, by-passing the wavefunctions. This is the main achievement of this theory. Of course, the choice of $v(\vec{r})$ remains problematic.

Although the Thomas-Fermi method suffers from a crude treatment of the kinetic energy and the neglect of exchange, it contains all the important ingredients of a density functional theory.

The Thomas-Fermi scheme has been widely used in the past. March (1983) provides a good background review of this method. This method has been found to give a rough estimative of the charge density and the electrostatic potential. Additionally, the charge density is infinite at the nucleus, and it does not decay exponentially far from the nucleus, but as r^{-6} (Gross & Dreizler, 1979).

Another weakness of the TF is the lack of shell structure, which means that the observed periodic variation when changing the atomic number cannot be reproduced. In fact, the atoms shrink with increasing atomic number Z ($Z^{-1/3}$) (Jones & Gunnarsson, 1989).

1.10.3 Exchange effects: Dirac's correction

The addition of a correction to the TF model due to exchange effects is due to Dirac (1930). Within this model—the Thomas-Fermi-Dirac model—the energy functional is written as,

$$E_{\mathrm{TFD}}[n;v] = E_{\mathrm{TF}}[n(\vec{r}), v(\vec{r})] + E_x[n(\vec{r})]\ . \tag{1.10.17}$$

The additional term is defined as,

$$E_x[n] = -C_x\, n(\vec{r})^{\frac{4}{3}} \quad \text{with} \quad C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} = 0.7386\ . \tag{1.10.18}$$

The Thomas-Fermi-Dirac energy functional, labelled TFD accordingly, can then be written, explicitly, as

$$E_{\text{TFD}}[n;v] = T_0[n(\vec{r})] + \int v(\vec{r})\, n(\vec{r})\, d^3r + U_{\text{H}}[n(\vec{r})] - C_x \int n(\vec{r})^{\frac{4}{3}}\, d^3r\ . \tag{1.10.19}$$

The expression for the exchange term can be derived from the Hartree-Fock theory in the density matrix formalism, using first order reduced matrices (Gross & Dreizler, 1979).

1.10.4 A conventional gradient correction: The Weizsäcker correction

Attempting a better description of the properties of an inhomogeneous electron gas, von Weizsäcker introduced a correction to the TF kinetic energy (von Weizsäcker, 1935):

$$T_{\text{W}}[n] = \frac{\hbar^2}{8m}\int \frac{\vec{\nabla}\, n(\vec{r}) \cdot \vec{\nabla}\, n(\vec{r})}{n(\vec{r})}\, d^3r\ . \tag{1.10.20}$$

In (1.10.20), in order to display the dependence on $\hbar$, atomic units are not used.

The total kinetic energy thus becomes

$$T_{\text{TF}\lambda\text{W}}[n] = T_{\text{TF}}[n(\vec{r})] + \lambda\, T_{\text{W}}[n(\vec{r})]\ , \tag{1.10.21}$$

where the parameter λ was taken to be unity in the original work of von Weizsäcker, but was later shown to lie in the interval $\frac{1}{9} \leqslant \lambda \leqslant 1$ (Meyer *et al.*, 1976).

To second order, we can write the Thomas-Fermi-Dirac-Weizsäcker (TFDW) model for the total energy functional as

$$\begin{aligned} E_{\text{TFDW}}[n] = {} & C_{\text{F}} \int n(\vec{r})^{\frac{5}{3}}\, d^3r + \int v(\vec{r})\, n(\vec{r})\, d^3r \\ & + U_{\text{H}}[n(\vec{r})] - C_x \int n(\vec{r})^{\frac{4}{3}}\, d^3r + \frac{1}{8}\,\lambda \int \frac{\left|\vec{\nabla}\, n(\vec{r})\right|^2}{n(\vec{r})}\, d^3r\ . \end{aligned} \tag{1.10.22}$$

The corresponding Euler-Lagrange equation is as follows:

$$\mu = \frac{5}{3}\, C_{\text{F}} \int n(\vec{r})^{\frac{2}{3}}\, d^3r + v(\vec{r}) - \frac{4}{3}\, C_x \int n(\vec{r})^{\frac{1}{3}}\, d^3r + \lambda\, \frac{1}{8}\left[\frac{\left|\vec{\nabla}\, n(\vec{r})\right|^2}{n^2(\vec{r})} - 2\,\frac{\nabla^2\, n(\vec{r})}{n(\vec{r})}\right]\ . \tag{1.10.23}$$

The TFDW model provides considerable improvement over the TF or TFD models. The electron density is finite at the nucleus instead of diverging as it does in the TF or TFD models:

$$n(\text{small}\ \vec{r}) \propto r^{-\frac{3}{2}} \to \infty \quad \text{when} \quad r \to 0\ . \tag{1.10.24}$$

Far from the nucleus, the electron density is given by (Gross & Dreizler, 1979)

$$n \propto \left(\frac{1}{r^2}\right) \exp\left\{-\left(\frac{-8\,\mu}{\lambda}\right) r\right\}\ , \tag{1.10.25}$$

in contrast with the improper power decay of r^{-6}.

2

Modern Density Functional Theory

2.1 Introduction

Density functional theory (DFT) has had a major impact on structure-electronic calculations. The density functional approach expresses ground-state properties—such as total energies, equilibrium positions or magnetic moments—in terms of the electronic density $n(\vec{r})$ and provides an effective scheme for calculating the referred properties. The method avoids the problem of calculating the ground-state many-body wavefunction, as it reduces the number of degrees of freedom to a minimum.

As described previously (§1.10.2), Thomas and Fermi had attempted to formulate such an approach much earlier; however their model suffered from inaccuracies in the description of the kinetic energy and a complete neglect of exchange-correlation effects. Despite this fact, the TF method served as a starting point for the development of a more advanced one by Hohenberg, Kohn and Sham (Hohenberg & Kohn, 1964; Kohn & Sham, 1965).

The DFT formalism is easily translated into computer language. It can handle infinite periodic systems (no N-dependence),[1] and non-periodic systems of very many atoms, currently on the order of $\sim 10^3$. These properties, together with many extensions and improvements, have made DFT a popular method in several branches of physics and chemistry. It is now possible to study distinct fields of physics like time-dependent phenomena, superconductivity or excited states, etc, using DFT.

The aim of the present chapter is to provide a brief account for the original idea behind the modern DFT and present the latest extensions and improvements to it. Special attention will be paid to generalisations of DFT that are relevant to the work presented in this thesis.

[1] In contrast to the HF method.

2.2 Basic formalism

For an electronic system described by the Hamiltonian (1.10.2)

$$\widehat{\mathcal{H}} \equiv -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\sum_{j\neq i}^{N}\frac{1}{|\vec{r}_i - \vec{r}_j|} + v(\vec{r}_i)\ , \tag{2.2.1}$$

where $v(\vec{r})$ is some external potential, both the ground-state energy and the ground state wavefunction are determined by the minimisation of the energy functional (1.5.1). From the form of the total Hamiltonian operator it is obvious that two quantities completely fix the referred Hamiltonian. These are the number, N, of electrons in the system and the external potential v. In other words, N and v determine all the properties for the ground state. This is not surprising since v defines the whole nuclear frame, which together with the number of electrons determines the electronic properties of the system under study.

Hohenberg and Kohn (1964) proposed the use of the electron density n as a basic variable, instead of N and v.

2.2.1 The Hohenberg and Kohn theorems

Theorem 1. *The density as a basic variable:*
The external potential v is uniquely determined within a physically irrelevant additive constant by the electron density n:

$$n(\vec{r}) \rightarrow v(\vec{r})\ . \tag{2.2.2}$$

This statement has far-reaching consequences. Since $n(\vec{r})$ determines v, and trivially N, it determines the Hamiltonian $\widehat{\mathcal{H}}$; hence, implicitly, also all the properties derivable from $\widehat{\mathcal{H}}$.

Despite its importance, the proof of the first theorem of Hohenberg and Kohn is disarmingly simple. It follows by *reductio ad absurdum*, having as premise the fact that two different external potentials, v and v', lead to the same charge density n.

Proof. Let v be the external potential of the system with N particles, defined as $\int n(\vec{r})\, d^3r$, an associated ground state density $n(\vec{r})$, Hamiltonian $\widehat{\mathcal{H}}$ and ground-state wavefunction and energy, Ψ and E, respectively. Schematically,

$$v:\ \widehat{\mathcal{H}},\ N,\ n,\ \Psi,\ E\ . \tag{2.2.3}$$

A second system,with N particles, is characterised as:

$$v':\ \widehat{\mathcal{H}}',\ N,\ n',\ \Psi',\ E'\ , \tag{2.2.4}$$

where $v' \neq v + C$, and hence $\Psi' \neq \Psi$.

By the Rayleigh-Ritz variational principle,

$$\begin{aligned} E &= <\Psi|\,\widehat{\mathcal{H}}\,|\Psi> < \ <\Psi'|\,\widehat{\mathcal{H}}'\,|\Psi'> = <\Psi'|\,\widehat{\mathcal{H}'}\,|\Psi'> + <\Psi'|\,(\widehat{\mathcal{H}} - \widehat{\mathcal{H}}')\,|\Psi'> \\ &= <\Psi'|\,\widehat{\mathcal{H}}'\,|\Psi'> + \int [v(\vec{r}) - v'(\vec{r})]\, n'(\vec{r})\, d^3r\ , \end{aligned} \tag{2.2.5}$$

or

$$E < E' + \int [v(\vec{r}) - v'(\vec{r})]\, n(\vec{r})\, d^3r \;. \tag{2.2.6}$$

The above inequality follows from the fact that $\Psi' \neq \Psi$. Similarly,

$$E' < E + \int [v'(\vec{r}) - v(\vec{r})]\, n'(\vec{r})\, d^3r \;. \tag{2.2.7}$$

Adding up (2.2.6) and (2.2.7), one obtains

$$(E + E') < (E + E') + \int [v'(\vec{r}) - v(\vec{r})]\, [n'(\vec{r}) - n(\vec{r})]\, d^3r \tag{2.2.8}$$

The possibility $n' \equiv n$ has to be excluded since it would lead to $(E + E') < (E + E')$. Thus any potential v' except $v + C$, leads to $n' \not\equiv n$.

Quod Erat Demonstradum.

The total ground-state energy of the system can be written as

$$E = <\Psi|\, \widehat{V}_{\rm ext}\, |\Psi> + <\Psi|\, (\widehat{T} + \widehat{V}_{ee})\, |\Psi> \;, \tag{2.2.9}$$

where the terms on the right hand side are the expectation values of the external potential, kinetic energy and interaction energy (including classical and non-classical terms), operators. Clearly,

$$\begin{aligned} <\Psi|\, \widehat{V}_{\rm ext}\, |\Psi> &= <\Psi| \sum_{i=1}^{N} v(\vec{r}_i)\, |\Psi> \\ &= \int v(\vec{r})\, n(\vec{r})\, d^3r \end{aligned} \tag{2.2.10}$$

are

$$F_{\rm HK}[n] \equiv <\Psi|\, (\widehat{T} + \widehat{V}_{ee})\, |\Psi> \;. \tag{2.2.11}$$

functionals of n. $F_{\rm HK}$ is *universal* in the sense that does not depend on v. It depends on n, via Ψ, defined for any physical n corresponding to *some* v.[2] Explicitly,

$$\begin{aligned} F_{\rm HK}[n] &= T[n(\vec{r})] + V_{ee}[n(\vec{r})] \\ &= T[n(\vec{r})] + (U[n(\vec{r})] + \text{non-classical term}) \;, \end{aligned} \tag{2.2.12}$$

where $V_{ee}[n]$ represents all the electron-electron effects, including $U_{\rm H}$, the classical repulsion of (1.10.4), and a non-classical term. $T[n]$ is the kinetic energy contribution.

Using (2.2.10) and (4.6.6), we can explicitly write the total energy functional as

$$E \equiv E_v[n(\vec{r})] = \int v(\vec{r})\, n(\vec{r})\, d^3r + F_{\rm HK}[n(\vec{r})] \;. \tag{2.2.13}$$

The second Hohenberg-Kohn (HK) theorem provides the density variational principle:

[2] Such densities are called v-representables.

Theorem 2. *The density variational principle:*
For a trial density $n'(\vec{r})$, such that $n'(\vec{r}) \geqslant 0$ and $\int n'(\vec{r})\, d^3r = N$,

$$E_v[n'(\vec{r})] \geqslant E_0 \tag{2.2.14}$$

where $E_v[n'(\vec{r})]$ is the energy functional defined by (2.2.13).

Using the Rayleigh-Ritz (RR) variational principle for the ground-state energy leads rather simply to the conclusion that, for a given v, the expression (2.2.13) is a *minimum* for the correct ground-state $n(\vec{r})$ (Hohenberg & Kohn, 1964):

Proof. Let Ψ' be a trial state. Note that the previous theorem ensures that $n'(\vec{r})$ determines its own v'. Then, by the conventional RR principle,

$$E[\Psi'] \equiv <\Psi'|\,(\widehat{T} + \widehat{V}_{\text{ext}} + \widehat{V}_{ee})\,|\Psi'> \geqslant E_0 \;, \tag{2.2.15}$$

or

$$E_v[n'(\vec{r})] = \int v(\vec{r})\, n'(\vec{r})\, d^3r + F_{\text{HK}}[n'(\vec{r})] \geqslant E_0 \;. \tag{2.2.16}$$

The equality sign holds only if $\Psi' = \Psi$.[3]

Quod Erat Demonstradum.

This is the HK energy variational principle. In other words, the variational principle for the ground-state energy (§1.5) can now be re-written: the charge replaces the total wavefunction as the direct variational variable.

Note that, so far, $F_{\text{HK}}[n]$ was defined only abstractly through (4.6.6), but not constructively. The HK theorems offer no practical guide to built this particular functional.[4] The knowledge of the explicit form of $F_{\text{HK}}[n]$ could allows us to apply this formalism to any system, independently of the number, N, of electrons in the system. This is easily seen if one applies the stationary principle explicitly:

$$\delta \left\{ E_v[n(\vec{r})] - \mu \int n(\vec{r})\, d^3r \right\} = 0 \;, \tag{2.2.17}$$

which is equivalent to the Euler equation

$$\mu = \frac{\delta}{\delta n} E_v[n(\vec{r})] = v[n(\vec{r})] + \frac{\delta F_{\text{HK}}[n(\vec{r})]}{\delta n(\vec{r})} \;. \tag{2.2.18}$$

In conclusion, Hohenberg and Kohn have showed that all the terms in the expression for the total energy may be evaluated as a sum of functionals of the charge density:

$$E[n] = T[n] + E_{e-i}[n] + U_{\text{H}} + E_{xc}[n] + E_{i-i}. \tag{2.2.19}$$

Clearly $E_{e-i}[n]$ and $U_{\text{H}}[n]$, the Hartree term, are manifestly functionals of the charge density:

[3] This proof can be extended to degenerate ground states, leading again to expression (2.2.15).
[4] The crudest *ansatz* for $F_{\text{HK}}[n]$ gives the familiar Thomas-Fermi approximation.

$$E_{e-i}[n] = -\int n(\vec{r}) \sum_{\alpha} \frac{Z_\alpha}{\left|\vec{r} - \vec{R}_\alpha\right|} \, d^3r, \tag{2.2.20}$$

$$U_{\rm H}[n] = \frac{1}{2} \iint \frac{n(\vec{r})\, n(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d^3r \, d^3r', \tag{2.2.21}$$

an the ion-ion term is

$$E_{i-i} = \frac{1}{2} \sum_{\alpha \neq \beta} \frac{Z_\alpha \, Z_\beta}{\left|\vec{R}_\alpha - \vec{R}_\beta\right|} \; . \tag{2.2.22}$$

The kinetic and the non-classical exchange-correlation term are more difficult to express as a function of $n(\vec{r})$. This issue will be discuss in forthcoming sections of this thesis.

2.2.2 The v-representability of an electron density

A function $n(\vec{r})$ is a *pure-state* v-representable if it is the density of a (degenerate or not) ground state of the total Hamiltonian associated with *some* suitable chosen local external potential v.[5]

Having defined a v-representable (VR) charge density, one question arises immediately: Given a well-defined and non-negative, charge density, normalised to N, can one find always an external potential, so that $n(\vec{r})$ is a ground-state density of the total Hamiltonian? Obviously, a given charge density, $n(\vec{r})$, may or may not be v-representable.

This is important since the inequality (2.2.16),

$$E_v[n(\vec{r})] = \int v(\vec{r})\, n(\vec{r})\, d^3r + F_{\rm HK}[n(\vec{r})] \geqslant E_0 \;, \tag{2.2.23}$$

ceases to apply in a practical application for a non-VR density.

The original hope of Hohenberg and Kohn (1964) that all well behaved functions are VRs, turned out to be too optimistic. Levy (1982) and Lieb (1983) have shown that there are densities, which are *à priori* reasonable candidates, are not indeed VR. Furthermore, Englisch and Englisch (1983) showed that even for a single-particle system there are densities that do not come from a ground-state wavefunction of any v. In this context, we can restate the first HK theorem as the fact that there is a one-to-one mapping between the ground-state wavefunctions and the VR densities. It is through this unique mapping that a VR density determines the properties of its ground state. In their original paper, HK (1964) proved v-representability for the case of a density that is nearly uniform.

Kohn (1983) has proved, for the case of a lattice (grid) where the kinetic energy operator is represented by some finite difference, that $n(\vec{r})$ will be a VR, provided it is close enough to a known VR density. Unfortunately, there is no such theorem for the case of a continuum.

[5] The number of particles, N, and their mutual interaction being specified.

2.2.3 The constrained search process

Of particular importance is the functional $F_{\rm HK}[n]$ of (4.6.6) defined as

$$F_{\rm HK}[n] \equiv <\Psi\,|\,(\widehat{T}+\widehat{V}_{ee})\,|\,\Psi> .$$

where Ψ is the ground-state wavefunction associated with $n(\vec{r})$, which has to be VR, *i.e.*, comes from an anti-symmetric N-electron wavefunction, in order to determine all the relevant properties of the system.

We have already established the one-to-one correspondence between the ground-state density, n_0, and the ground-state wavefunction Ψ_0. It is obvious that Ψ_0 provides n_0 (by quadrature). But there is an infinite number of antisymmetric wavefunctions (not necessarily from ground states) that give the same charge density. Given one of these functions that integrates to n_0, say Ψ', how do we distinguish them from the ground-state Ψ_0? Although having proved that such a charge density exists, Hohenberg and Kohn (1964) have not provided the tools to construct it.

The answer to this question was provided by Levy (1982) and further investigated by Lieb (1983) and Levy and Perdew (1985). This formalism is usually classified under the heading of a *constrained search.*

Re-applying the variational principle for the functions Ψ', we can write

$$<\Psi'\,|\,(\widehat{T}+\widehat{V}_{ee})\,|\,\Psi'> \geqslant <\Psi_0\,|\,(\widehat{T}+\widehat{V}_{ee})\,|\,\Psi_0> \qquad (2.2.24)$$

since the potential energy due to the external potential $v_{\rm ext}(\vec{r})$ is a simple function of N.

We conclude that among all the wavefunctions giving the same charge density n_0, the ground state Ψ_0 minimises the expectation value of $<\widehat{T}+\widehat{U}>$.

Analytically, given *any* n' of one of many ground-state wavefunctions Ψ' (Dreizler & Gross, 1990),

$$\begin{aligned} F_{\rm LL}[n'] \equiv & \inf_{\substack{\Psi \\ (n'[\Psi']=n'(\vec{r}))}} <\Psi'\,|\,(\widehat{T}+\widehat{V}_{ee})\,|\,\Psi'> \\ = & <\Psi_{n'}\,|\,(\widehat{T}+\widehat{V}_{ee})\,|\,\Psi_{n'}> , \end{aligned} \qquad (2.2.25)$$

where $\Psi_{n'}$ is the wavefunction that gives n' and minimizes $F_{\rm LL}[n']$. This is a constrained search definition for $F_{\rm LL}[n']$. The search of the infinimum proceeds over all the trial antisymmetric wavefunctions Ψ', subject to the constraint that $n'[\Psi'] = n'(\vec{r})$, *i.e*, whose density equals the prescribed charge density, n'. The variation is *constrained* because the space of trial functions comprises only those that give the density n', in opposition to the search (1.5.1). The requirement of v-representability is then no longer necessary.

An immediate consequence of the Rayleigh-Ritz principle is

$$F_{\rm LL}[n'(\vec{r})] = F_{\rm HK}[n'(\vec{r})] \qquad (2.2.26)$$

for all pure-state VR functions n'.

This formalism is easily incorporated into the HF scheme by rewriting the RR variational principle as

$$E_0 = \inf_{\Psi'} <\Psi' \,|\, \widehat{H} \,|\, \Psi'> . \tag{2.2.27}$$

The above minimisation proceeds in two stages (Dreizler & Gross, 1990):

$$\begin{aligned} E_0 &= \inf_{n'} \Big[\inf_{\substack{\Psi \\ (n'[\Psi']=n'(\vec{r}))}} <\Psi' \,|\, (\widehat{T} + \widehat{V}_{ee} + \widehat{V}_{\text{ext}}) \,|\, \Psi'> \Big] \\ &= \inf_{n'} \Big[\inf_{\Psi'} <\Psi' \,|\, (\widehat{T} + \widehat{V}_{ee}) \,|\, \Psi'> + \int v(\vec{r})\, n'(\vec{r})\, d^3r \Big] \\ &= \inf_{n'} \Big[F_{\text{LL}}[n'(\vec{r})] + \int v(\vec{r})\, n'(\vec{r})\, d^3r \Big] . \end{aligned} \tag{2.2.28}$$

In the second line, the inner minimisation is constrained to all wavefunctions that give n', while the outer minimisation releases this constrained by searching all n'.

Note that the $F_{\text{HK}}[n']$ functional only applies to ground-state densities, while the definition (2.2.28) belongs to a broader class, covering the case of degenerate states.

2.3 The Kohn-Sham scheme

2.3.1 Introduction

As already stated in previous sections, the Hohenberg-Kohn provides no practical guide for the explicit construction of the functional $F_{\text{HK}}[n]$.

An important question still remains to be answered. Can the properties of the homogeneous electron gas be used in theoretical studies of inhomogeneous systems, like solids? So far, none of the available functions led to satisfactory results. This was mainly due to a deficient description of the kinetic contribution $T[n]$.

This picture changed dramatically with the scheme introduced by Kohn and Sham (1965). The main achievement of the later was the introduction of an orbital single-particle picture that can be established rigorously. It leads quite naturally to a more appropriate representation of the kinetic energy contribution.

The Kohn-Sham (KS) scheme has its roots in the Hartree and X_α methods, which are, like the KS, self-consistent methods featuring an effective single-particle Schrödinger equation with a local, particle dependent, one-particle potential. A crucial difference from the Hartree and Slater's X_α and the KS method is that the later includes correlation effects exactly.

In following sections, we will present the derivation of the KS equations based on a variational principle. This will involve an energy functional depending on single-particle orbital and their (fractional) occupancy numbers. This will allow us to discuss the possibility of calculating excitation, or transition, energies with the aid of the so-called *transition-state* argument. This concept was introduced by Slater (1974) to calculate approximate excitation energies within the Hartree-Fock-Slater method.

2.3.2 The Kohn-Sham equations

Using a non-interacting reference system of N particles, Kohn and Sham (1965) introduced the following Hamiltonian $\widehat{H}_s$,

$$\widehat{H}_s = \widehat{T} + \widehat{V}_s = \sum_i^N \left(-\frac{1}{2}\nabla_i^2 \right) + \sum_i v_s(\vec{r}_i) \; . \tag{2.3.1}$$

Obviously, this Hamiltonian does not contain any electron-electron interaction terms. According to the first HK theorem, there exists a unique energy functional,

$$E_s[n] = T_s[n(\vec{r})] + \int v_s(\vec{r})\, n(\vec{r})\, d^3r \tag{2.3.2}$$

for which the variational principle $\delta E_s[n] = 0$ yields the exact ground-state density $n_s(\vec{r})$ corresponding to $\widehat{H}_s$. $T_s[n]$ denotes the universal kinetic energy functional of a system of *non-interacting* particles.

The main assertion in the KS scheme is: for any system of interacting particles, there exists a *local* single-particle potential $v_s(\vec{r})$ such that the exact ground-state density $n(\vec{r})$ of the interacting system satisfies

$$n(\vec{r}) \equiv n_s(\vec{r}) \; . \tag{2.3.3}$$

Here $n(\vec{r})$ is the charge density of the real (inhomogeneous) electron gas.

The determinantal ground-state wavefunction for this system of non-interacting particles is

$$\Psi_s \equiv \frac{1}{\sqrt{N!}} \det\left[\psi_1\, \psi_2 \, \cdots \, \psi_N\right] . \tag{2.3.4}$$

The single-particle wavefunctions ψ_i, assumed non-degenerate, define n_s through

$$n_s(\vec{r}) = \sum_i^N |\psi_i(\vec{r})|^2 \; , \tag{2.3.5}$$

with the sum over all occupied states, ie, N, are obtained from the one-particle Schrödinger equation

$$\widehat{h}_s\, \psi_i(\vec{r}, s) \equiv \left\{ -\frac{1}{2}\nabla_i^2 + v_s(\vec{r}) \right\} \psi_i = \varepsilon_i\, \psi_i, \quad i = 1, 2, \ldots, N \; . \tag{2.3.6}$$

Once the existence of a potential v_s, generating $n(\vec{r})$, via (2.3.5) and (2.3.6), is assumed, the uniqueness of v_s follows from the first HK theorem.

The kinetic energy term $T_s[n]$ is, according to Kohn and Sham (1965)

$$\begin{aligned} T_s[n] &= <\Psi_s \,|\, \sum_i^N \big(-\frac{1}{2}\nabla_i^2 \big) \,|\, \Psi_s> \\ &= \sum_i^N <\psi_i \,|\, \big(-\frac{1}{2}\nabla^2 \big) \,|\, \psi_i> \; , \end{aligned} \tag{2.3.7}$$

It is now convenient to separate the several component of $F[n]$ as,

$$F[n] = T_s[n(\vec{r})] + U_{\mathrm{H}}[n(\vec{r})] + E_{xc}[n(\vec{r})] \tag{2.3.8}$$

where

$$\begin{aligned} E_{xc}[n] &= F_{\mathrm{HK}} - U_{\mathrm{H}}[n(\vec{r})] - T_s[n(\vec{r})] \\ &= T[n(\vec{r})] - T_s[n(\vec{r})] + V_{ee}[n(\vec{r})] - U_{\mathrm{H}}[n(\vec{r})] \;. \end{aligned} \tag{2.3.9}$$

The quantity $E_{xc}[n]$ is the exchange-correlation energy; it contains the difference between T and T_s, presumably fairly small as well as the non-classical part of $V_{ee}[n]$.

Although we can perform a search over all anti-symmetric N-electron wavefunctions directly on T_s,

$$T_s[n] = \inf_{n} < \Psi_s \,|\, \widehat{T}_s \,|\, \Psi_s > = < \Phi_n \,|\, \widehat{T} \,|\, \Phi_n > \;, \tag{2.3.10}$$

the minimising wavefunction Φ_n for a given density will be a non-interacting wavefunction for some external potential v_s such as

$$\frac{\delta\, T_s}{\delta n(\vec{r})} + v_s(\vec{r}) = \mu \;, \tag{2.3.11}$$

as in (2.2.18). Lieb (1983) has conjectured that the Thomas-Fermi kinetic energy functional $T_0[n]$ (1.10.8) is the lower limit of $T_s[n]$, *i.e.*,

$$T_0[n] \leqslant T_s[n] \;. \tag{2.3.12}$$

The potential v_s assumes the form of an *effective* potential,

$$\begin{aligned} v_s(\vec{r}) &= v(\vec{r}) + \frac{\delta\, U_{\mathrm{H}}[n(\vec{r})]}{\delta n(\vec{r})} + \frac{\delta E_{xc}[n(\vec{r})]}{\delta n(\vec{r})} \\ &= v(\vec{r}) + V^{\mathrm{H}}([n];\vec{r}) + V_{xc}([n];\vec{r}) \;, \end{aligned} \tag{2.3.13}$$

with the exchange-correlation potential as

$$V_{xc}([n];\vec{r}) \equiv \frac{\delta E_{xc}(\vec{r})}{\delta n(\vec{r})} = \frac{\delta}{\delta n}\big(F_{\mathrm{HK}}[n(\vec{r})] - U^{\mathrm{H}}[n(\vec{r})] - T_s[n(\vec{r})]\big) \tag{2.3.14}$$

and

$$V^{\mathrm{H}}([n];\vec{r}) \equiv \frac{\delta\, U_{\mathrm{H}}[n]}{\delta n(\vec{r})} = \int \frac{n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d^3r' \;. \tag{2.3.15}$$

According to Kohn and Sham recipe (1965), we do not try to solve (2.3.11) directly, since it is just a re-arrangement of (2.2.18) and the explicit form of $T_s[n]$ is still unknown.

Instead, their alternative treatment is as follows: For a given v_s, one obtains the n that satisfies (2.3.11) simply by solving the following N one-electron equations

$$\widehat{h}_s\, \psi_i([n];\vec{r}) \equiv \left\{ -\frac{1}{2}\nabla_i^2 + v_s(\vec{r}) \right\} \psi_i(\vec{r}) = \varepsilon_i\, \psi_i(\vec{r}) \quad \text{with} \quad \varepsilon_1 \leqslant \varepsilon_2 \leqslant \cdots \leqslant \varepsilon_N \;, \tag{2.3.16}$$

and

$$v_s(\vec{r}) = v(\vec{r}) + V^{\rm H}([n];\vec{r}) + V_{xc}([n];\vec{r}) \ , \tag{2.3.17}$$

defining the density n like in (2.3.5), *i.e.*,

$$n(\vec{r}) = \sum_i^N |\psi_i([n];\vec{r})|^2 \ . \tag{2.3.18}$$

Obviously, this set of equations, known as the Kohn-Sham equations, can only be solved self-consistently. This is was already expressed in (2.3.18) by representing the orbital ψ_i as a functional of n, $\psi_i \equiv \psi_i[n]$.

Once a self-consistent solution of the KS equation has been obtained, the single particle energies

$$\varepsilon_i = \int \psi_i^*(\vec{r})\Big(-\frac{1}{2}\nabla_i^2\Big)\,\psi_i(\vec{r})\,d^3r + \int v_s(\vec{r})\,|\psi_i(\vec{r})|^2\,d^3r \ , \tag{2.3.19}$$

can be used to re-write the kinetic energy functional as

$$T[n] = \sum_{i=1}^N \varepsilon_i - \int v_s(\vec{r})\,n(\vec{r})\,d^3r \tag{2.3.20}$$

The total energy can now be expressed as a function of the KS eigenvalues as follows,

$$E = \sum_i^N \varepsilon_i - U_{\rm H}[n(\vec{r})] + E_{xc}[n(\vec{r})] - \int V_{xc}([n];\vec{r})\,n(\vec{r})\,d^3r \ , \tag{2.3.21}$$

where

$$\begin{aligned} \sum_i^N \varepsilon_i &= \sum_i^N < \psi_i|\,-\frac{1}{2}\nabla^2 + v_s(\vec{r})\,|\psi_i> \\ &= T_s[n(\vec{r})] + \int v_s(\vec{r})\,n(\vec{r})\,d^3r \end{aligned} \tag{2.3.22}$$

2.3.3 The chemical potential

Let us assume that a charge density $n_N(\vec{r})$ is a solution of (2.3.11), that is

$$\frac{\delta E_s[n]}{\delta n(\vec{r})} = \mu \ , \tag{2.3.23}$$

for a system with N particles and ground-state energy E_N, *i.e.*, $E_N = E_s[n_N(\vec{r})]$.

For each solution $n_N(\vec{r})$, (2.3.23) is satisfied for a certain value μ with $\mu = \mu(N)$. It is then obvious that the Lagrange multipliers, incorporating the subsidiary condition of a specific number of particles in a HK variational problem described by (2.2.17), *i.e.*,

$$\frac{\delta}{\delta n}\Big(E_s[n] - \mu\int n(\vec{r})\,d^3r\Big) = 0 \ , \tag{2.3.24}$$

correspond to the exact chemical potential of the system, *i.e.*,

$$\mu(N) = \left.\frac{\partial E_N}{\partial N}\right|_{v(\vec{r})} \ , \tag{2.3.25}$$

i.e., the chemical is the slope of the E versus N curve at constant $v(\vec{r})$. Now, consider the difference in ground state energies,

$$E_{N+\eta} - E_N = E_s[n_{N+\eta}(\vec{r})] - E_s[n_N(\vec{r})] \ . \tag{2.3.26}$$

From the definition of functional derivative, assuming a infinitesimal change, η, in the total number of particles, N, of a particular system, the above expression is equivalent to

$$\begin{aligned} \delta E_s &= \int \left.\frac{\delta E_s[n(\vec{r})]}{\delta n(\vec{r})}\right|_{n_N} \left[n_{N+\eta}(\vec{r}) - n_N(\vec{r})\right] d^3r \\ &= \int \mu(N) \left[n_{N+\eta}(\vec{r}) - n_N(\vec{r})\right] d^3r = \mu(N)\,\eta \ , \end{aligned} \tag{2.3.27}$$

using (2.3.25). In other words, the Lagrange parameter μ is exactly the system's chemical potential when $\eta \to 0$. Additionally, (2.3.25) is formally similar to the chemical potential of an open system (Parr *et al.*, 1978), which provides another argument towards the identification of the parameter μ as a chemical potential.

2.3.4 Exchange and correlation

The exchange-correlation energy, $E_{xc}[n]$ can be decomposed into the exchange and correlation components as

$$E_{xc}[n] = E_x[n] + E_c[n] \ . \tag{2.3.28}$$

The exchange term is (Levy, 1996)

$$E_x[n] = <\Phi_n \,|\, \widehat{V}_{e-e} \,|\, \Phi_n> - U_{\mathrm{H}}[n] \ . \tag{2.3.29}$$

Using (2.3.28) and (2.3.8), the correlation energy can be written as

$$\begin{aligned} E_c[n] &= F[n] - \{T_s[n] + U_{\mathrm{H}}[n] + E_x\} \\ &= <\Psi_n \,|\, (\widehat{T} + \widehat{V}_{e-e}) \,|\, \Psi_n> - <\Phi_n \,|\, (\widehat{T} + \widehat{V}_{e-e}) \,|\, \Phi_n> \ , \end{aligned} \tag{2.3.30}$$

where Ψ_n and Φ_n are the wavefunctions yielding n (2.2.26 and 2.3.10) and minimising the expectation values of $(\widehat{T} + \widehat{V}_{e-e})$ and $\widehat{T}$ respectively. The above expression shows that

$$E_c[n] \leqslant 0 \ . \tag{2.3.31}$$

Noting that

$$<\Phi_n \,|\, \widehat{V}_{e-e} \,|\, \Phi_n> = T_s[n] + U_{\mathrm{H}}[n] + E_x[n] \ , \tag{2.3.32}$$

the exchange and correlation contributions are, in the one-electron limit, *i.e.* non-interacting system ($\widehat{V}_{e-e} = 0$), as follows (Perdew & Zunger, 1981):

$$E_x[n] = -U_{\mathrm{H}}[n] \quad \text{and} \quad E_c[n] = 0 \quad \text{for} \quad N = 1 \ . \tag{2.3.33}$$

This expression shows that the exchange-correlation energy of a one-electron ($N = 1$) system simply cancels the spurious self-interaction $U_{\rm H}[n]$. In a similar way for the KS potential:

$$\frac{\delta E_x[n]}{\delta n(\vec{r})} = -V^{\rm H}([n];\vec{r}) \quad \text{and} \quad \frac{\delta E_c[n]}{\delta n(\vec{r})} = 0 \quad \text{for} \quad N = 1 \,. \tag{2.3.34}$$

Thus

$$\lim_{r\to\infty} \frac{\delta E_{xc}[n]}{\delta n(\vec{r})} = -\frac{1}{r} \,, \tag{2.3.35}$$

for $N = 1$.

2.3.5 Interpretation of the Kohn-Sham equations

Comparing the KS equations with the single Euler-Lagrange equation (2.2.18), we can see a major advantage: through the introduction of N orbitals, the KS equations deal with $T_s[n]$, the dominant part of the true kinetic energy $T[n]$, indirectly but exactly. The price for this gain in accuracy is that there are now N one-electron to solve as opposed to just one equation for the total density derived from a direct approximation to $T_s[n]$ of a Thomas-Fermi type. It is very important to note that at this stage, the other unknown part E_{xc} remains intact, entering in the KS in the same way.

The KS equations have the same form as the Hartree equations (1.6.7), except that they contain a more general effective potential. Despite this, the computational effort needed to solve the KS equations is not much more than that to solve the equations of Hartree but less than those of Hartree and Fock.

The Hartree-Fock equations contain a non-local potential operator in the one-electron Hamiltonian and hence are not a particular case of the KS equations. Nevertheless, all three theories—Hartree, Hartree-Fock and Kohn-Sham—provide one-electron equations for describing many-electron systems.

The KS scheme, exact in principle, is distinguished form the HF theory in its capability to account for correlation-exchange effects of electrons. In HF theory, approximate by definition, electron correlation effects are lacking (§1.9) and their incorporation is by no means easy, involving wavefunction techniques like configuration interaction. The KS equations are open for improvements to the approximation to E_{xc} and would give exact n, and total energy, E if E_{xc} was known precisely.

2.4 Local density approximation

Despite the fact that the KS equations (2.3.16)–(2.3.18) incorporate the kinetic energy $T_s[n]$, they still leave the correlation-exchange term undetermined. It is therefore necessary to find a way of circumventing this difficulty.

In this section, we will described the simplest approximation to $T_s[n]$, offered by Kohn and Sham (1965). This is the *local-density approximation* (LDA) for the exchange and correlation energy.

Before presenting the approximation, it is convenient to recall that the uniform-electron gas was used to obtain the Thomas-Fermi functional $T_{\rm TF}$ (1.10.8) and the Dirac functional for the exchange energy (1.10.18). Now, since the kinetic energy is rigorously treated in the KS scheme, we can use the uniform-electron gas to treat solely the unknown part of the exchange-correlation functional. We write

$$E_{xc}^{\rm LDA}[n] = \int \epsilon_{xc}(n(\vec{r}))\, n(\vec{r})\, d^3r \ , \tag{2.4.1}$$

where ϵ_{xc} indicates the exchange and correlation energy per particle of a non-polarised uniform electron gas of density n.

The local density model for the exchange-correlation energy is obtained from the homogeneous electron gas via the prescription

$$\epsilon_{xc}^{\rm LDA}(n(\vec{r})) \longrightarrow \epsilon_{xc}^{\rm hom}(n_0)\Big|_{n_0\rightarrow n(\vec{r})} \ , \tag{2.4.2}$$

where n_0 is the constant density of the homogeneous gas and $n(\vec{r})$ is the local density of the actual inhomogeneous gas. The one-particle exchange-correlation energy is known to great accuracy ($\sim$ 0.1 %) from studies of the uniform electron gas (Pines, 1963).

The corresponding potential then becomes

$$V_{xc}^{\rm LDA}(n) \equiv \frac{\delta E_{xc}^{\rm LDA}[n]}{\delta n(\vec{r})} = \epsilon_{xc}(n(\vec{r})) + n(\vec{r})\, \frac{\delta \epsilon_{xc}(n)}{\delta n(\vec{r})} \ , \tag{2.4.3}$$

and the self-consistent KS orbital equations read

$$\left\{ -\frac{1}{2}\nabla^2 + v(\vec{r}) + U_{\rm H}([n];\vec{r}) + V_{xc}^{\rm LDA}(n;\vec{r}) \right\} \psi_i(\vec{r}) = \varepsilon_i\, \psi_i(\vec{r}) \ . \tag{2.4.4}$$

The function $\epsilon_{xc}(n)$ can be divided into exchange and correlation contributions, *i.e.*,

$$\epsilon_{xc}(n) = \epsilon_x(n(\vec{r})) + \epsilon_c(n(\vec{r})) \ . \tag{2.4.5}$$

As described previously, the exchange part is already known thanks to Dirac (§1.10.3). Analytically, for the spin-average case

$$\epsilon_x(n) = -C_x\, n(\vec{r})^{\frac{1}{3}} \ ,$$

with C_x given by (1.10.18).

Imposing a condition of neutrality, the total ground-state energy of a uniform electron gas, including the positive background, may be written as

$$\begin{aligned} E[n] &= T_s[n(\vec{r})] + E_{xc}[n(\vec{r})] \\ &= T_s[n(\vec{r})] + E_x[n(\vec{r})] + E_c[n(\vec{r})] \\ &= \int \epsilon_t(n(\vec{r}))\, n(\vec{r})\, d^3r + \int \epsilon_x(n(\vec{r}))\, n(\vec{r})\, d^3r + \int \epsilon_c(n(\vec{r}))\, n(\vec{r})\, d^3r \ . \end{aligned} \tag{2.4.6}$$

The total energy for the spin-compensated case can now be written as a function of the Wigner-Seitz radius r_s (1.10.7), as

$$E[n] = \int \epsilon(n(\vec{r}))\, n(\vec{r})\, d^3r \,, \tag{2.4.7}$$

with

$$\epsilon(n) = \frac{1.1049}{r_s^2} - \frac{0.4582}{r_s} + \epsilon_c(r_s) \,. \tag{2.4.8}$$

where the first and second terms on the right side are numerical values for $\epsilon_t(n)$ and $\epsilon_x(n)$, respectively. Note that large r_s means low density and small r_s high density: r_s at the nucleus of a hydrogen-like atom of nuclear charge Z is $0.91/Z$; at the first Bohr radius is $1.77/Z$.

Exact analytic expressions for $\epsilon_c(r_s)$ are known only in extreme limits. We will present and discuss the available expressions for the correlation energy per particle in the next section, where we consider the LSD approximation, *i.e*, the generalisation of LDA approximation to fully polarised systems.

2.4.1 Discussion of the local density approximation

As pointed out in connection with the $X\alpha$ method (§1.8), the LDA is expected to hold when the density varies sufficiently slowly in space, *i.e.*, $|\vec{\nabla}\, n(\vec{r})| \ll k_{\mathrm{F}}(n(\vec{r}))\, n(\vec{r})$. However, it has been found that this approximation provides surprisingly good results even for systems beyond this limit. This condition is not always met as exemplified by the case of metallic Cu (Gunnarsson *et al.*, 1979).

In order to understand this success, a variation of the coupling constant integration technique has been used. In particular, the analysis of the exchange-correlation or Fermi-Coulomb hole has proven to be rather fruitful (Harris & Jones, 1974; Langreth & Perdew, 1975; Gunnarsson & Lundqvist, 1976; Langreth & Perdew, 1977; Harris, 1984).

Let us define Ψ_n^λ as the normalised and antisymmetric wavefunction which provides the density $n(\vec{r})$ and minimises the expectation value of $\widehat{T} + \lambda\, \widehat{V}_{e-e}$. λ is a non-negative coupling constant.

When $\lambda = 1$, Ψ_n^λ is Ψ_n, is the ground state wavefunction of an interacting particle system of density $n(\vec{r})$. When $\lambda = 0$, Ψ_n^λ is Φ_n, the non-interacting or Kohn-Sham wavefunction for density $n(\vec{r})$. Consequently, varying λ between the two limits, for a fixed density $n(\vec{r})$ is equivalent to varying the external potential $v_\lambda(\vec{r})$: at $\lambda = 1$, $v_\lambda(\vec{r})$ is the true external potential, while at $\lambda = 0$, it assumes the form of a Kohn-Sham effective potential $v_s(\vec{r})$. It is customary to assume a smooth, 'adiabatic', relationship between interacting and non-interacting ground states as λ varies from 1 to 0. It is important to note that only $\lambda = 1$ is real or has any physical meaning, while the Kohn-Sham system at $\lambda = 0$, or any other value in between are just a convenient mathematical fiction.

Within the framework introduced by the coupling constant formalism, the exchange-correlation energy $E_{xc}[n]$ (formally defined previously by equations 2.3.28–2.3.30) can now be re-written as

$$
\begin{aligned}
E_{xc}[n] &= <\Psi_n^\lambda \,|\, \widehat{T} + \lambda\, \widehat{V}_{e-e} \,|\, \Psi_n^\lambda> \Big|_{\lambda=1} - <\Psi_n^\lambda \,|\, \widehat{T} + \lambda\, \widehat{V}_{e-e} \,|\, \Psi_n^\lambda> \Big|_{\lambda=0} - U^{\mathrm{H}}[n] \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\lambda} <\Psi_n^\lambda \,|\, \widehat{T} + \lambda\, \widehat{V}_{e-e} \,|\, \Psi_n^\lambda> d\lambda - U^{\mathrm{H}}[n] \; . \qquad (2.4.9)
\end{aligned}
$$

Applying the Hellmann-Feynman theorem of section (§1.5.1), E_{xc} can be expressed as

$$
E_{xc}[n] = \int_0^1 <\Psi_n^\lambda \,|\widehat{V}_{e-e} \,|\, \Psi_n^\lambda> d\lambda - U^{\mathrm{H}}[n] \; . \qquad (2.4.10)
$$

Using the concept of reduced density matrices (Parr & Yang, 1989), it is possible to evaluate the N-electron expectation value of a sum of two-body operator like $\widehat{V}_{e-e}$. According to Gunnarsson and Lundqvist (1976) and Gunnarsson *et al.* (1979), the exchange-correlation energy (2.4.10) can be written as

$$
E_{xc}[n] = \frac{1}{2} \iint \frac{n(\vec{r}\,)}{|\vec{r} - \vec{r}\,'|} \, \widetilde{n}_{xc}(\vec{r}, \vec{r}\,') \, d^3r \, d^3r' \; , \qquad (2.4.11)
$$

where

$$
\widetilde{n}_{xc}(\vec{r}, \vec{r}\,') = \int_0^1 n_{xc}^\lambda(\vec{r}, \vec{r}\,') \, d\lambda = n(\vec{r}\,') \int_0^1 \left[g_\lambda(\vec{r}, \vec{r}\,'; [n]) - 1 \right] d\lambda \; . \qquad (2.4.12)
$$

Here, $g_\lambda(\vec{r}, \vec{r}\,'; [n])$ is a pair-distribution function defined by the probability of finding an electron in a homogeneous gas, of density $n(\vec{r}\,)$ and electron-electron interactions described by $\lambda \widehat{V}_{e-e}$, at point $\vec{r}$ provided there is another electron at $\vec{r}\,'$. The second integral in (2.4.12) defines an average pair-distribution function.

Obviously, the exchange-correlation can be written in terms of the pair-distribution function, *i.e.*,

$$
E_{xc}[n] = \frac{1}{2} \iint n(\vec{r}\,) \, \frac{\left[\widetilde{g}(\vec{r}, \vec{r}\,'; [n]) - 1 \right]}{|\vec{r} - \vec{r}\,'|} \, n(\vec{r}\,') \, d^3r \, d^3r' \; . \qquad (2.4.13)
$$

The quantity $n(\vec{r}\,') \left[\widetilde{g}(\vec{r}, \vec{r}\,') - 1 \right]$ describes the depletion region in density due to the exchange-correlation hole around an electron at $\vec{r}\,'$, when an average with respect to λ is taken. The exchange-correlation energy may be viewed as the energy resulting from the interaction between an electron and its exchange-correlation hole.

The exact relation (2.4.13) is very appealing because it enables one to formulate all approximation for E_{xc} as ones for $\widetilde{g}(\vec{r}, \vec{r}\,')$. In particular, for the use of the local-density approximation the exact density depletion around an electron at $\vec{r}$ is replaced by that of a homogeneous electron gas of density $n(\vec{r}\,)$, *i.e.*,

$$
n(\vec{r}\,') \left[\widetilde{g}(\vec{r}, \vec{r}\,') - 1 \right] \longrightarrow n(\vec{r}\,) \left[\widetilde{g}_{\mathrm{hom}}(|\vec{r} - \vec{r}\,'|; n(\vec{r}\,') - 1 \right] \; . \qquad (2.4.14)
$$

One important feature of the above expression can be understood right away. The density depletion due to the exchange-correlation hole around an electron corresponds

exactly to one particle, independently of the size of the electron-electron coupling strength. Consequently, the sum rule

$$n(\vec{r})\left[\widetilde{g}_{\mathrm{hom}}(|\vec{r}-\vec{r}\,'|;n(\vec{r}\,')-1\right] = \int \widetilde{n}_{xc}^{\mathrm{LDA}}(\vec{r},|\vec{r}\,'-\vec{r}|)\,d^3r' = -1\ , \tag{2.4.15}$$

must hold as a consequence of electron-number conservation.

A variable substitution $\vec{R}\equiv\vec{r}\,'-\vec{r}$ yields, bringing expression (2.4.14) in,

$$E_{xc}[n] = \frac{1}{2}\int n(\vec{r})\,d^3r \int_0^\infty R^2\,\frac{1}{R}\,dR \int \widetilde{n}_{xc}^{\mathrm{LDA}}(\vec{r},\vec{R})\,d\Omega\ , \tag{2.4.16}$$

with $\Omega\equiv\vec{r}\,'-\vec{r}$ defining the integration domain.

This expression shows that the exchange-correlation energy only depends on the spherical average of $\widetilde{n}_{xc}^{\mathrm{LDA}}(\vec{r},\vec{R})$. An immediate consequence being that the approximation for E_{xc} can give an ‘exact’ value even when the non-spherical part of $\widetilde{n}_{xc}^{\mathrm{LDA}}$ is quite inaccurate. Gunnarsson *et al.* (1979) were able to show, for the neon atom, that even though the LDA provides a poor description of the shape of the exchange-correlation hole, the resulting spherical average is quite acceptable.

Let us summarise some of the reasons behind the unexpected success of the local density approximation:

(*i*) Since $\widetilde{g}(\vec{r},\vec{r}\,')$ tends to unity when $|\vec{r}\,'-\vec{r}|$ tends to infinity, the interaction between the electron and its exchange-correlation hole is dominated by the total charge and *not* by its precise distribution;

(*ii*) The exchange-correlation hole in the local density approximation does not need to mimic precisely the true exchange-correlation hole to yield a good result for E_{xc}. It is sufficient that it provides a reasonable approximation for the spherical average of the exchange-correlation hole;

(*iii*) The satisfaction of the sum rule (2.4.15) implies that if $\widetilde{n}_{xc}^{\mathrm{LDA}}(\vec{r},\vec{R})$ has positive errors for some values of $\vec{R}$, it is bound to have negative errors for other values. Cancellation of errors has to be expected.

2.5 Spin polarised systems

The extension of the Kohn-Sham scheme to spin-polarised systems was first given by von Barth and Hedin (1972) and Rajagopal and Callaway (1973) together with a generalisation of the Hohenberg-Kohn theorems.

The spin-polarised Kohn-Sham self-consistent equation can be expressed as

$$\left\{-\frac{1}{2}\nabla^2 + v(\vec{r}) + U_{\mathrm{H}}([n];\vec{r}) + V_{xc}^{\sigma}([n_\uparrow;n_\downarrow];\vec{r})\right\}\psi_i^{\sigma}(\vec{r}) = \varepsilon_i^{\sigma}\,\psi_i^{\sigma}(\vec{r})\ , \tag{2.5.1}$$

with

$$n_\sigma(\vec{r}) = \sum_i \gamma_i^{\sigma}\,|\psi_i^{\sigma}(\vec{r})|^2\ . \tag{2.5.2}$$

Here $\sigma = \uparrow$ or $\downarrow$ is the z-component of spin. The occupancy numbers γ_i satisfy

$$\sum_i^{\infty} \gamma_i^{\sigma} = N_{\sigma} \quad \text{with} \quad N_{\uparrow} + N_{\downarrow} = N \ . \tag{2.5.3}$$

Here $N_{\uparrow}$ and $N_{\downarrow}$ are simply the number of electrons with spin 'up' and 'down' respectively.

The KS effective potential includes the classical Hartree term, $V^{\mathrm{H}}[n]$, defined as

$$V^{\mathrm{H}}([n];\vec{r}) = \int \frac{n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d^3r \tag{2.5.4}$$

with

$$n(\vec{r}) = n_{\uparrow}(\vec{r}) + n_{\downarrow}(\vec{r}) \ . \tag{2.5.5}$$

Using this expression for the total spin-polarised charge density, the relative spin-polarisation is defined as

$$\zeta(\vec{r}) = \frac{n_{\uparrow}(\vec{r}) - n_{\downarrow}(\vec{r})}{n(\vec{r})} \ , \tag{2.5.6}$$

ranging from 0 for a unpolarised system to ± 1 for a fully polarised system.

The other term in the KS effective potential is the exchange-correlation potential which now reads

$$V_{xc}^{\sigma}([n_{\uparrow}; n_{\downarrow}];\vec{r}) = \frac{\delta E_{xc}[n_{\uparrow}; n_{\downarrow}]}{\delta n_{\sigma}(\vec{r})} \ , \tag{2.5.7}$$

under the symmetry constraint

$$V_{xc}^{\uparrow}(\zeta) = V_{xc}^{\downarrow}(-\zeta) \ . \tag{2.5.8}$$

2.5.1 The local spin density approximation

The local spin density approximation (LSD) for the exchange-correlation energy is

$$E_{xc}^{\mathrm{LSD}}[n_{\uparrow}; n_{\downarrow}] = \int \epsilon_{xc}\big(n_{\uparrow}(\vec{r}), n_{\downarrow}(\vec{r})\big)\, \big[n_{\uparrow}(\vec{r}) + n_{\downarrow}(\vec{r})\big]\, d^3r \ , \tag{2.5.9}$$

where $\epsilon_{xc}(n_{\uparrow}, n_{\downarrow})$ is the exchange-correlation energy per particle for an electron gas of uniform spin densities $n_{\uparrow}$ and $n_{\downarrow}$ (Kohn & Sham, 1965; von Barth & Hedin, 1972; Vosko *et al.*, 1980; Perdew & Zunger, 1981).

2.6 Correlation effects

Finding the correlation energy per particle is a difficult problem, even for spin-compensated systems ($\zeta = 0$). Exact expressions are only known for two extreme limits. These are for high- and low densities.

In the case of high-density, $\epsilon_c(r_s, \zeta)$ — a function of the Wigner-Seitz radius now defined for the spin polarised case, as

$$r_s = \left[\frac{3}{4\pi}\big(\, n_{\uparrow} + n_{\downarrow}\,\big)\right]^{\frac{1}{3}} \ , \tag{2.6.1}$$

is given by

$$\epsilon_c(r_s,\zeta) = c_0(\zeta)\,\ln r_s - c_1(\zeta) + c_2(\zeta)\,r_s\,\ln r_s - c_3\,r_s + \cdots \qquad r_s \to 0\;. \tag{2.6.2}$$

This expression is known as the high-density expansion (Vosko *et al.*, 1980; Gell-Mann & Brueckner, 1957; Perdew & Zunger, 1981).

For the low-density limit (Vosko *et al.*, 1980; Perdew & Zunger, 1981), $\epsilon_c(r_s,\zeta)$ becomes

$$\epsilon_c(r_s,\zeta) = -\frac{d_0(\zeta)}{r_s} + \frac{d_1(\zeta)}{r_s^{3/2}} + \cdots \qquad r_s \to \infty\;. \tag{2.6.3}$$

Combining the results for the two density limits, the one-particle correlation can be written as (Perdew & Wang, 1992)

$$\epsilon_c(r_s,\zeta) = -2\,c_0\,(1+\alpha_1\,r_s)\,\ln\left[1 + \frac{1}{2\,c_0\,\left(\beta_1\,r_s^{1/2} + \beta_2\,r_s + \beta_3\,r_s^{3/2} + \beta_4\,r_s^2\right)}\right], \tag{2.6.4}$$

where

$$\beta_1 = \frac{1}{2\,c_0}\,\exp\left(\frac{c_1}{2\,c_0}\right) \quad \text{and} \quad \beta_2 = 2\,c_0\,\beta_1^2\;. \tag{2.6.5}$$

The coefficients entering in the above formulae were found by fitting to accurate quantum Monte Carlo correlation energies (Ceperley & Alder, 1980). Ceperley and Alder calculated the total energy for the uniform electron gas in spin-compensated and ferromagnetic states for several values of r_s. The correlation energy was obtained by subtracting the corresponding kinetic and exchange energies from the total energy,

$$E_c[n] = E[n] - T_s[r_s] - E_x[r_s]\;. \tag{2.6.6}$$

For an arbitrary spin polarisation ζ, $\epsilon_c(n_\uparrow,n_\downarrow)$ is given by (Vosko *et al.*, 1980)

$$\begin{aligned}\epsilon_c(n_\uparrow,n_\downarrow) \equiv \epsilon_c(r_s,\zeta) &= \epsilon_c(r_s,0) + \alpha_c(r_s)\,\frac{f(\zeta)}{f''(0)}\,(1-\zeta^4) + \left[\epsilon_c(r_s,1) - \epsilon_c(r_s,0)\right]\,f(\zeta)\,\zeta^4 \\ &= \epsilon_c(r_s,0) + \alpha_c(\zeta)\,\zeta^2 + \mathcal{O}(\zeta^4)\;,\end{aligned} \tag{2.6.7}$$

with

$$f(\zeta) = \frac{\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2\right]}{2(2^{1/3}-1)}\;. \tag{2.6.8}$$

note that $f(0) = 0$, $f(1) = 1$ and $f''(0) = 1.709921$. α_c is the correlation contribution to the spin stiffness being $\approx \epsilon_c(r_s,1) - \epsilon_c(r_s,0)$.

Spin-scaling relations can be used to convert density functional into spin density functionals (§1.5.2). For this, we write the exchange energy E_x as

$$E_x[n] = E_x^\uparrow + E_x^\downarrow \tag{2.6.9}$$

which is a negative number. Alternately, this expression can be expressed as

$$E_x[n_\uparrow,n_\downarrow] = E_x[n_\uparrow,0] + E_x[0,n_\downarrow] \tag{2.6.10}$$

Table 2.1: Parameters of the fit to the one-particle correlation energy, $\epsilon_c(r_s,\zeta)$ (2.6.4) and the high-density (2.6.2) and low-density (2.6.3) expansions for spin-compensated ($\zeta = 0$) and fully polarised ($\zeta = 1$) systems (Perdew & Wang, 1992).

	$\epsilon_c(r_s,0)$	$\epsilon_c(r_s,1)$	$-\alpha_c(r_s)$
c_0	0.031091	0.015545	0.016887
α_1	0.21370	0.20548	0.11125
β_1	7.5957	14.1189	10.357
β_2	3.5876	6.1977	3.6231
β_3	1.6382	3.3662	0.88026
β_4	0.49294	0.62517	0.49671
c_1	0.046644	0.025599	0.035475
c_2	0.00664	0.00319	0.00188
c_3	0.01043	0.00384	0.00521
d_0	0.4335	0.3287	0.2240
d_1	1.4408	1.7697	0.3969

The corresponding functional of a unpolarised system becomes

$$E_x[n] = E_x[n/2, n/2] = 2\,E_x[n/2, 0] \tag{2.6.11}$$

therefore, $E_x[n/2, 0] = \frac{1}{2}E_x[n]$. Now, $E_x[n_\uparrow, n_\downarrow]$ can be expressed as (Oliver & Perdew, 1979),

$$E_x[n_\uparrow, n_\downarrow] = \frac{1}{2}\,E_x[2n_\uparrow] + \frac{1}{2}\,E_x[2n_\downarrow]\,. \tag{2.6.12}$$

The exchange energy per particle can now be written as

$$\epsilon_x[n_\uparrow, n_\downarrow] = \epsilon_x(r_s, 0)\,\frac{\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right]}{2}\,, \tag{2.6.13}$$

using the spin-scaling relation (2.6.12). A similar expression can be found for the kinetic energy functional:

$$T_s[n_\uparrow, n_\downarrow] = T_s(r_s, 0)\,\frac{\left[(1+\zeta)^{5/3} + (1-\zeta)^{5/3}\right]}{2}\,. \tag{2.6.14}$$

2.7 Beyond the local spin density approximation

The work-horse approximation to the exchange-correlation energy has been the local spin density (LSD) approximation as defined in (§2.5.1):

$$E_{xc}^{\mathrm{LSD}}[n_\uparrow; n_\downarrow] = \int \epsilon_{xc}\big(n(\vec{r}), \zeta(\vec{r})\big)\, n(\vec{r})\, d^3r\,, \tag{2.7.1}$$

where $\epsilon_{xc}(n;\zeta)$ is the one-particle exchange-correlation energy in a uniform gas with density $n = n_\uparrow + n_\downarrow$ and spin polarisation $\zeta \equiv (n_\uparrow - n_\downarrow)/n$; $n(\vec{r})\, d^3r$ is the average number of electrons in the volume element d^3r.

The investigation of inhomogeneity corrections to the local density approximation was initiated by Hohenberg and Kohn (1964). The LSD approximation is the first term of an asymptotic Taylor expansion in gradients of the density (Kohn & Sham, 1965). However, it has been shown that the second-order gradient-expansion approximation (GEA),

$$E_{xc}^{\mathrm{GEA}}[n_\uparrow; n_\downarrow] = \int \epsilon_{xc}\big(n(\vec{r}), \zeta(\vec{r})\big)\, n(\vec{r})\, d^3r + \int C_{xc}\big(n(\vec{r}), \zeta(\vec{r})\big)\, |\vec{\nabla} n(\vec{r})|^2\, n(\vec{r})^{\frac{3}{4}}\, d^3r \;, \tag{2.7.2}$$

provides no systematic improvement.

More recently, generalised gradient approximations (GGA's) have become a popular tool among chemists (Perdew, 1985; Becke, 1988; Lee *et al.*, 1988; Perdew & Wang, 1992; Perdew *et al.*, 1996):

$$E_{xc}^{\mathrm{GGA}}[n_\uparrow; n_\downarrow] = \int f\big(n(\vec{r}), \zeta(\vec{r}), |\vec{\nabla} n(\vec{r})|\big)\, d^3r \;. \tag{2.7.3}$$

The input one-particle exchange correlation energy, $\epsilon(n_\uparrow, n_\downarrow)$ is in principle unique, since there is a possible system in which $n_\uparrow$ and $n_\downarrow$ are constant and for which LSD is exact. At least in this sense, there is no unique input $f(n, \zeta, |\vec{\nabla} n|)$. Despite this fact, a judicious choice of the form of f can improve the error of the LSD by a factor of 2 to 10.

Table 2.2 shows the typical errors found from self-consistent calculations within the LSD and GGA schemes (Perdew & Kurth, 1998).

Table 2.2: Typical errors for atoms, molecules and solids from self-consistent Kohn-Sham calculations within the LSD and GGA approximations of (2.7.1) and (2.7.3).

Property	LSD	GGA
E_x	5 % (not negative enough)	0.5 %
E_c	100 % (too negative)	5 %
bond lengths	1 % (too short)	1 % (too long)
structure	overly favours close packing	more correct

Very recently, Perdew *et al.* (1999) introduced the so-called *meta-generalised gradient approximation* (MGGA), which preserves the correct formal properties of LSD. According to this group, the argument of the function f in (2.7.3) is expand to include additional semi-local information, like the kinetic energy density τ_σ, defined as

$$\tau_\sigma(\vec{r}) = \frac{1}{2} \sum_\lambda |\vec{\nabla} \psi_\lambda^\sigma(\vec{r})|^2 \;, \tag{2.7.4}$$

with the summation over occupied KS orbitals. Although molecular atomisation energies are significantly improved over GGA, the calculated lattice constants are little changed.

2.8 Fractional occupancy

The idea of fractional occupancy is not new. Since Slater (1974), many authors have addressed this issue (Gilbert, 1975; Almbladh & von Barth, 1976; Parr *et al.*, 1978; Janak, 1978; Perdew & Zunger, 1981; Perdew *et al.*, 1982; Parr & Bartolotti, 1983; McHenry *et al.*, 1987; Kleinman, 1997; Perdew & Levy, 1997)

As deduced previously (§2.3.2), the KS total energy, as a function of the one-electron eigenvalues, can be written as

$$E[n] = \sum_{i=1}^{\infty} \gamma_i \, \varepsilon_i - U_{\mathrm{H}}([n]; \vec{r}) + E_{xc}[n(\vec{r})] - \int V_{xc}([n]; \vec{r}) \, n(\vec{r}) \, d^3r \; , \tag{2.8.1}$$

ignoring any spin degrees of freedom for simplicity. The charge density $n(\vec{r})$ is defined as

$$n(\vec{r}) = \sum_{i=1}^{\infty} \gamma_i \, |\psi(\vec{r})|^2 \; . \tag{2.8.2}$$

Here we have explicitly introduced occupation numbers γ_i, such that

$$\begin{cases} \gamma_i = 1 & : \quad \varepsilon_i < \mu \\ 0 \leqslant \gamma_i \leqslant 1 & : \quad \varepsilon_i = \mu \\ \gamma_i = 0 & : \quad \varepsilon_i > \mu \end{cases} \qquad \text{and} \qquad \sum_{i}^{\infty} \gamma_i = N \; . \tag{2.8.3}$$

If the above restriction is released so that all occupation numbers γ_i of levels below or above the Fermi level, are allowed to be arbitrary in the interval $0 \leqslant \gamma_i \leqslant 1$, the total energy functional $E[\psi_1, \psi_2 \ldots ; \gamma_1, \gamma_2 \ldots]$ can no longer be considered a density functional. In particular, no Hohenberg-Kohn variational principle applies to $E[\psi_i; \gamma_i]$, ψ_i and γ_i denoting the sets $\{\psi_i, i = 1 \ldots \infty\}$ and $\{\gamma_i, i = 1 \ldots \infty\}$, respectively.

Despite this, its possible to recover the general Kohn-Sham scheme by including the variation of *both* the single-particle orbitals *and* occupation numbers. In other words, the additional constraint,

$$\sum_{i=1}^{\infty} \gamma_i = N \qquad \text{with} \qquad 0 \leqslant \gamma_i \leqslant 1 \; , \tag{2.8.4}$$

imposing a fixed number of particles, is used in addition to the condition of orthonormality

$$\int \psi_i^*(\vec{r}) \, \psi_j(\vec{r}) \, d^3r = \delta_{ij} \; . \tag{2.8.5}$$

Including both conditions and making the substitution,

$$\gamma_i \equiv \cos^2(\theta_i) \; , \tag{2.8.6}$$

in order to account for the restriction (2.8.4), the variational problem can be posed as follows:

$$\Omega[\psi_n;\theta_n] = \widetilde{E}[\psi_n;\gamma_n] - \Big(\sum_{i,j}\lambda_{ij}\int \psi_i^*(\vec{r})\,\psi_j(\vec{r})\,d^3r + \mu\sum_i \gamma_i\Big)\ , \tag{2.8.7}$$

with respect to the wavefunction $\psi_k(\vec{r})$, as well as to θ_k. From the condition $\delta\Omega = 0$, where Ω is an auxiliary functional, the following set of equation may be deduced (Dreizler & Gross, 1990):

$$\gamma_k\,\widehat{h}_s(\vec{r})\,\psi_k(\vec{r}) - \sum_j \lambda_{kj}\,\psi_j(\vec{r}) = 0\ , \tag{2.8.8}$$

$$\sin(2\theta_k)\,\big[<\psi_k\,|\,\widehat{h}_s\,|\,\psi_k> - \mu\big] = 0\ . \tag{2.8.9}$$

Here, $\widehat{h}_s(\vec{r})$ is an effective one-electron Hamiltonian defined as

$$\widehat{h}_s(\vec{r}) \equiv -\frac{1}{2}\nabla^2 + v(\vec{r}) + V^{\mathrm{H}}([n];\vec{r}) + V_{xc}([n];\vec{r})\ . \tag{2.8.10}$$

The $\psi_k(\vec{r})$ are eigenvectors satisfying the eigenvalue problem

$$\hat{h}_s(\vec{r})\,\psi_k(\vec{r}) = \varepsilon_k\,\psi_k(\vec{r}) \qquad \text{with} \qquad \lambda_{kj} = \gamma_k\,\varepsilon_k\,\delta_{kj}\ , \tag{2.8.11}$$

satisfying (2.8.8), since the operator $\widehat{h}_s$ is clearly Hermitian.

Using (2.8.9), we can re-write (2.8.3) for the one-particle k-th level as

$$\begin{cases} \varepsilon_k = \mu & : \quad \theta_k \text{ arbitrary} \quad \Leftrightarrow \quad 0 \leqslant \gamma_i \leqslant 1 \\ \varepsilon_k \neq \mu & : \quad \theta_k = 0 \quad \Leftrightarrow \quad \gamma_i = 1 \\ \varepsilon_k \neq \mu & : \quad \theta_k = \frac{\pi}{2} \quad \Leftrightarrow \quad \gamma_i = 0 \end{cases}$$

Identifying μ with the Fermi energy, it can be seen that all the levels degenerate with the Fermi energy, can have fractional occupation, while the orbitals above μ are not occupied and those below are occupied.

In conclusion, the variation of $\widetilde{E}$ with respect to the occupancy numbers γ_i leads back to the ground-state occupational pattern (2.8.3) within the KS approach.

2.8.1 The Janak theorem

The functional $\widetilde{E}$ allows one to go beyond the limit of a 2-value occupancy ($\gamma_i = 0$ or 1) imposed by Kohn and Sham if the variation with respect to the one-electron orbital is carried out for an arbitrary but *fixed* set of occupation numbers γ_i.

The resulting variational equation (2.8.11) with the Hamiltonian (2.8.10) is formally identical to the KS counterpart. The set of occupation numbers γ_i only occurs in (2.8.2).

Assuming that a given set of self-consistent solutions,

$$\{\varepsilon_i(\vec{\gamma}),\ \psi_i^{(\vec{\gamma})},\ i = 1\ldots\infty\}$$

can be found for any given set of occupancy numbers $\vec{\gamma} = (\gamma_1, \gamma_2, \dots)$, the stationary point given by the KS energy functional is,

$$E(\gamma_1, \gamma_2, \dots) \equiv \widetilde{E}[\psi_1^{(\vec{\gamma})}, \psi_2^{(\vec{\gamma})}, \dots; \gamma_1, \gamma_2 \dots \gamma_N] \tag{2.8.12}$$

as a function of $\vec{\gamma}$.

Straightforward differentiation of this function with respect to γ_i yields the equation

$$\left.\frac{\partial E(\gamma_1, \gamma_2, \dots)}{\partial \gamma_j}\right|_{\vec{\gamma}} = \varepsilon_j(\vec{\gamma}) \tag{2.8.13}$$

known as Janak's theorem (Janak, 1978). Explicitly,

$$\left.\frac{\partial E(\gamma_1, \gamma_2, \dots)}{\partial \gamma_j}\right|_{\vec{\gamma}} = \left.\frac{\partial E([n], \vec{\gamma})}{\partial \gamma_j}\right|_{n(\vec{r})} + \int \frac{\delta E[n]}{\delta n(\vec{r})} \frac{\partial n(\vec{r})}{\partial \gamma_j} d^3 r \; . \tag{2.8.14}$$

Obviously, the second term on the right side of (2.8.14) vanishes as $E[n]$ is stationary to variations in $n(\vec{r})$, *i.e.*, $\delta E / \delta n = 0$. Re-calling (2.8.1), the first term can readily evaluated and the result (2.8.13) follows immediately.

The expression deduced by Janak is not original. It is just the LDF analogue to the relation deduced by Slater (1974) for the $X\alpha$ method. Both results are in close contrast with the HF method in which ε_i is given as a finite difference in energies for two states for which the occupancy, γ_i, of the i-th level differs by unity (1.7.24). This simply reflects the fact that a given level is either empty or filled with one electron. An immediate advantage of the $X\alpha$/LDF result over the HF one is that the former allows for the number of particles to vary continuously, so that the differentiation in (2.8.13) can be carried out straightforwardly.

Janak's theorem brings a whole new meaning to the KS eigenvalues, allowing one to go beyond Koopman's theorem (§1.7.1). It allows the calculation of excitation energies and not simply first ionisation energies. Refinements of Janak's theorem include those of Rajapogal (1980) and Perdew and Zunger (1981).

2.8.2 Slater's 'transition-state' argument

Assuming non-degeneracy, for simplicity, the ground state occupancy may be described as,

$$\vec{\gamma}_{\text{gs}} = \begin{cases} \gamma_i = 1 & : \quad i = 1 \dots N \\ \gamma_i = 0 & : \quad i > N \end{cases} \tag{2.8.15}$$

The exact ground-state energy of the interacting system can now be recovered,

$$E(\vec{\gamma}_{\text{gs}}) = E_0 \; . \tag{2.8.16}$$

Now consider the following occupancy:

$$\vec{\gamma}_{qp} = \begin{cases} \gamma_i = 1 & : \; i = 1, \dots, (q-1), (q+1), \dots, N \quad \text{and} \quad i (= p) > N \\ \gamma_i = 0 & : \; i > N \quad \text{and} \quad i = q \quad \text{with} \quad i \neq p \end{cases} \tag{2.8.17}$$

Clearly, this corresponds to a *electron-hole excitation* from the qth to the pth state. In other words, an electron is transfered from the q-th level—leaving a hole behind—to the p-th level above the Fermi level.

The self-consistent scheme described by equations (2.8.8)–(2.8.11) is identical to the Hartree-Fock-Slater method, if one replaces E_{xc} by Slater's exchange energy. In this case, the self-consistent Slater determinant corresponding to the occupation $\vec{\gamma}_{qp}$ can be considered an approximate excited state of the many-particle system in question so that,

$$\Delta E_{qp} = E(\vec{\gamma}_{qp}) - E(\vec{\gamma}_{\text{gs}}) > 0 \, , \tag{2.8.18}$$

represents an excitation energy.

It should be is noticed that the determination of the ΔE_{qp} involves two self-consistent calculations, one with $\vec{\gamma}_{qp}$ and one with $\vec{\gamma}_{\text{gs}}$. Additionally, for the case of a solid, ΔE_{qp} is a finite difference of two large but nearby equal quantities.

Instead of basing the calculations on the ground, Slater (1972; 1974) introduced the concept of *transition state*: a state in which the occupation numbers are exactly half-way between those of the initial and final state. Analytically,

$$\vec{\gamma}_{ts} = \begin{cases} \gamma_i = 1 & : \quad i = 1, \dots, (q-1), (q+1), \dots, N \\ \gamma_i = \dfrac{1}{2} & : \quad i = q \quad \text{and} \quad i = p \\ \gamma_i = 0 & : \quad i > q \quad \text{and} \quad i \neq p \end{cases} \tag{2.8.19}$$

The difference in total energies (2.8.18), describing the excited state of the system due to the transition of an electron between p- and q-th states, can now be expressed as

$$\begin{aligned} \Delta E_{qp} &\equiv E_q - E_p \\ &= E(\vec{\gamma}_{ts} - \Delta\vec{\gamma}) - E(\vec{\gamma}_{ts} + \Delta\vec{\gamma}) \, , \end{aligned} \tag{2.8.20}$$

with

$$\Delta\vec{\gamma} \equiv \begin{cases} \gamma_q = \dfrac{1}{2} & \\ \gamma_p = -\dfrac{1}{2} & \\ \gamma_i = 0 & : \quad i \neq p, q \end{cases} \tag{2.8.21}$$

Expanding both energies in power series (Slater, 1972; Slater, 1974), around $\vec{\gamma}_{ts}$, ignoring any structural change due to partial occupancies,

$$E_q = E(\vec{\gamma}_{ts}) + (\vec{\gamma}_{ts} - \Delta\vec{\gamma}) \left. \frac{\partial E}{\partial \gamma_q} \right|_{\vec{\gamma}_{ts}} + \cdots \tag{2.8.22}$$

$$E_p = E(\vec{\gamma}_{ts}) + (\vec{\gamma}_{ts} + \Delta\vec{\gamma}) \left. \frac{\partial E}{\partial \gamma_p} \right|_{\vec{\gamma}_{ts}} + \cdots \tag{2.8.23}$$

one obtains by means of Janak's theorem

$$\Delta E_{qp} \backsimeq \varepsilon_q(\vec{\gamma}_{ts}) - \varepsilon_p(\vec{\gamma}_{ts}) + \mathcal{O}((\Delta\vec{\gamma})^3) \tag{2.8.24}$$

Slater (1972) showed that the high order term $\mathcal{O}((\Delta\vec{\gamma})^3)$ is very small. In this case the excitation energy can finally be written as

$$\Delta E_{qp} = \varepsilon_q(\vec{\gamma}_{ts}) - \varepsilon_p(\vec{\gamma}_{ts}) \tag{2.8.25}$$

In conclusion, the excitation energy is well represented by the difference of single-particle eigenvalues, in a self-consistent calculation scheme like DFT theory. This last expression is the key formula for the work present in this thesis. As we will show in a forthcoming chapter, this result allows the evaluation of the position of a deep level of a particular defect within the material bandgap.

2.8.3 Ionisation potential and electron affinity

We now apply the transition state argument to calculate excitation energies like ionisation and affinity energies.

The first ionisation energy is the energy required to remove an electron from the highest occupied molecular, or defect, level. In other words, it corresponds to a bound-to-free transition ($p \to \infty$), defined by the occupancy

$$\vec{\gamma}_{\infty p} = \begin{cases} \gamma_i = 1 & : \ i = 1, \dots, (N-1) \\ \gamma_i = 0 & : \ i \geqslant N \end{cases} \tag{2.8.26}$$

simply reflecting the fact that the number of electrons in the system is not N but $N-1$. The cost in energy for the removal of the electron is then $\Delta E_{\infty p} = E(\vec{\gamma}_{\infty p}) - E(\vec{\gamma}_{\mathrm{gs}})$ with $E(\vec{\gamma}_{\mathrm{gs}})$ given by (2.8.15).

The ionisation energy can then be defined, invoking Janak's theorem, as

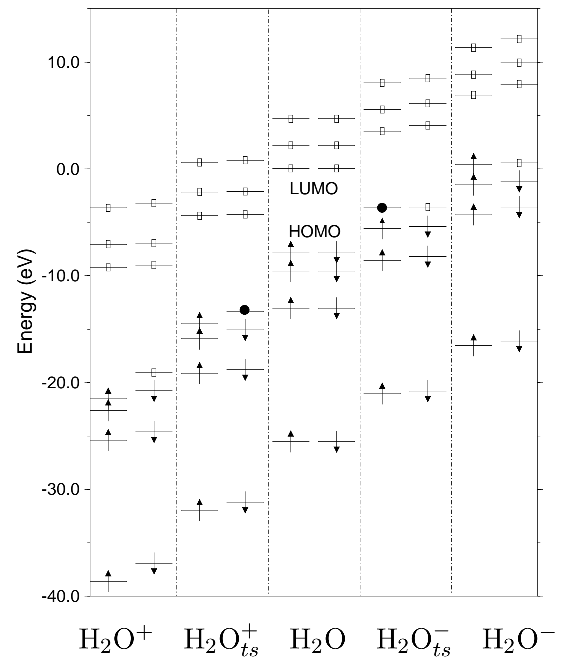


Figure 2.1: Optimised spin-polarised one-electron eigenvalue spectra for H_2O, its positive and negative ions, and the corresponding intermediate transition states. Arrows, boxes and '•'s denote occuppied, empty and partially filled orbitals, respectively.

$$I = \Delta E_{\infty p} \equiv E_{N-1} - E_N \backsimeq -\left.\frac{\partial E}{\partial \gamma_p}\right|_{\vec{\gamma}_{ts}} = -\varepsilon_p(\vec{\gamma}_{ts}) \, , \tag{2.8.27}$$

with

$$\vec{\gamma}_{ts} = \begin{cases} \gamma_i = 1 & : \quad i = 1, \dots, (N-1) \\ \gamma_i = \frac{1}{2} & : \quad i = p = N \\ \gamma_i = 0 & : \quad i > N \end{cases} \tag{2.8.28}$$

In a similar way, it can be shown that the electron affinity can approximated by

$$-A = \Delta E_{p\infty} \equiv E_{N+1} - E_N \backsimeq \varepsilon_p(\vec{\gamma}_{ts}) \; , \tag{2.8.29}$$

with $\vec{\gamma}_{ts}$ now defined as

$$\vec{\gamma}_{ts} = \begin{cases} \gamma_i = 1 & : \quad i = 1, \dots, N \\ \gamma_i = \dfrac{1}{2} & : \quad i = p = (N+1) \\ \gamma_i = 0 & : \quad i > (N+1) \end{cases} \tag{2.8.30}$$

Let us consider the case of the water molecule, and apply the transition-state argument to calculated its ionisation and potential energies. So far, we have not mencioned the effect of the structural optimisation upon Slater's calculation scheme. This issue is addressed in §5.7. For the calculated affinity of H_2O with the transition state, relaxation improved the result by ~ 0.1 eV. Although, considerable for an isolated molecule, this shift is much smaller for larger systems.

The results presented below were obtained from self-consistent pseudopotential local spin density theory calculations using a Gaussian basis (AIMPRO).[6] The following basis set was used: four and three Gaussian s and p orbitals with different widths centred on the oxygen and hydrogen atoms for the wave function; four and two s and p Gaussians, respectively, for oxygen and hydrogen, to fit the charge density. Three additional orbitals were placed at the centre of the O–H bonds. No charge-density gradient corrections were considered and both electronic and structural optimisations of the molecule were performed for all the charge states considered. This included the transition-state calculations of the molecule ionisation and affinity potentials in which the net charge states were $+\frac{1}{2}\,e$ and $-\frac{1}{2}\,e$ respectively.

The ionisation and affinity energies are now written as:

$$-I = E(\mathrm{H_2O}) - E(\mathrm{H_2O^+}) = \int_0^1 \varepsilon_{\mathrm{HOMO}}(\gamma)\, d\gamma \backsimeq \varepsilon_{\mathrm{HOMO}}(\gamma = \tfrac{1}{2}) \tag{2.8.31}$$

$$-A = E(\mathrm{H_2O^-}) - E(\mathrm{H_2O}) = \int_0^1 \varepsilon_{\mathrm{LUMO}}(\gamma)\, d\gamma \backsimeq \varepsilon_{\mathrm{LUMO}}(\gamma = \tfrac{1}{2}) \; , \tag{2.8.32}$$

respectively. Here γ represents the occupancy of the highest occupied molecular orbital (HOMO) and lowest occupied molecular orbital (LUMO) of H_2O. The occupancy of both orbitals is $\frac{1}{2}$ of an electron. This is depicted in Figure 2.1.

As shown, in the above mentioned figure, the self-consistent spin-up eigenvalue represented by the symbol '•' in the intermediate transition state of H_2O ($H_2O^+_{ts}$ for a net proton charge of $+\frac{1}{2}\,e$) *i.e.*, $-\varepsilon_{\mathrm{HOMO}}$, provides the first ionisation energy of that particular system. In other words, the ionisation potential of H_2O is the average of $-\varepsilon_{\mathrm{HOMO}}$ between H_2O^+ and H_2O, approximately the value of $-\varepsilon_{\mathrm{HOMO}}$ for the transition state $H_2O^+_{ts}$.

[6]Chapter 4 provides a detailed description of the AIMPRO method.

Table 2.3: Calculated and observed first ionisation energies, or ionisation potentials (I) and electron affinities (A) for the H_2O molecule (eV). Experimental values are taken from Moore (1970).

$-I \equiv E(\mathrm{H_2O}) - E(\mathrm{H_2O^+})$	$-I$ (TS)	$-I$ (Expt.)	Var. (%)
13.36	13.32	12.59	6.1/5.8
$-A \equiv E(\mathrm{H_2O^-}) - E(\mathrm{H_2O})$	A (TS)	A (Expt.)	Var. (%)
−3.53	3.54	3.61	2.2/1.9

The calculated values for the first ionisation potential and electronic affinity for the water molecule (H_2O) are shown in Table 2.3. The results from transition-state calculations are compared with those from total energy calculations, as well as with experimental values.

2.8.4 The band gap problem

In a semiconductor, the band gap is rigorously defined (Sham, 1966) as the difference between the lowest conduction-band energy and the highest valence-band energy, which in turn, are defined as the lowest energies to add and remove, respectively, an electron from the system. Analytically,

$$\Delta = \varepsilon_c - \varepsilon_v \;, \tag{2.8.33}$$

with

$$\varepsilon_c = E_N - E_{N-1} \equiv -I \quad \text{and} \quad \varepsilon_v = E_N - E_{N+1} \equiv A \;, \tag{2.8.34}$$

It is now necessary to relate these one-particle excitations with the self-consistent Kohn-Sham eigenvalues. Sham (1985) has shown that the band gap is

$$\Delta = \varepsilon^{\mathrm{KS}}_{N+1}(N+1) - \varepsilon^{\mathrm{KS}}_{N}(N) \tag{2.8.35}$$

in terms of the KS eigenvalues of a N-particle system. Note that this expression is different from the naive definition of the band gap

$$\Delta = \varepsilon^{\mathrm{KS}}_{N+1}(N) - \varepsilon^{\mathrm{KS}}_{N}(N) \tag{2.8.36}$$

which is only accurate for a system of non-interacting particles.

Perdew and Levy (1983), and independently Sham and Schlüter (1983), demonstrated qualitatively that the difference between (2.8.35) and (2.8.36) is due to exchange-correlation effects. Specifically, it arises from a discontinuity in the exchange-correlation potential when the number of particles in the system varies from N to $N+1$, *i.e.*,

$$\Delta_{xc} \equiv V^{(N+1)}_{xc}(\vec{r}) - V^{(N)}_{xc}(\vec{r}) \;, \tag{2.8.37}$$

where $V_{xc}(\vec{r})$ is the exact DFT exchange-correlation potential for a N-electron system. The band gap can now by defined as

$$\Delta = \Delta_{\rm KS} + \Delta_{xc} \tag{2.8.38}$$

The KS band gap evaluated within the LDA, $\Delta_{\rm KS}^{\rm LDA}$ is found to be to small by typically 40–50 % in comparison with experiment. Since (4.10.3) is an exact representation of the band gap, the error found for $\Delta_{\rm KS}^{\rm LDA}$ may have two distinct sources: (*i*) the discontinuity Δ_{xc} is neglected; (*ii*) the value $\Delta_{\rm KS}^{\rm LDA}$ obtained with the LDA is not equal to the exact KS gap $\Delta_{\rm KS}$.

A partial solution of this problem was given by Godby *et al.* (1986). These authors were able to show that the neglect of the discontinuity is the principal source of error. Δ_{xc} is responsible for over 80 % of the LDA-gap error.

3

Pseudopotential Theory

3.1 Introduction

It is a well known fact in chemistry that the properties of a molecule depends predominantly on the valence electrons and hardly at all on the core electrons. Core electrons occupy inner, inert, shells of atoms and their main role is to shield the outer electrons—valence electrons—from the nuclei. It is therefore tempting to eliminate them from an electronic structure calculation, replacing them instead by an effective potential acting only on valence electrons.

Due to the requirement of orthogonality between core and valence states, the wavefunctions of the valence electrons vary rapidly in the core region, as these always have a small spatial extent. Obviously, if the core electrons are removed, this constraint disappears. This suggests replacing the true valence wavefunctions by a *pseudo-wavefunction* varying smoothly near the nucleus (see Figure).

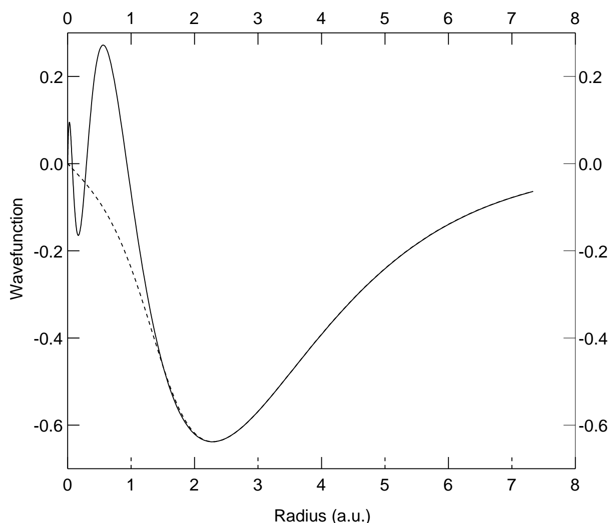


Figure 3.1: The real (full) and pseudo-radial (dashed) $4s$ wavefunction for the Ni atom.

3.2 Basic idea

In order to eliminate the core states and replace the valence-band wavefunction by a pseudo-wavefunction, we write

$$|\chi_v\rangle = |\phi_v\rangle + \sum_c \beta_c \, |\phi_c\rangle \quad \text{with} \quad \beta_c = \langle\phi_c \,|\, \chi_v\rangle \tag{3.2.1}$$

since $\langle\phi_c \,|\, \phi_v\rangle = 0$. Here $|\phi_c\rangle$ and $|\phi_v\rangle$ are the real conduction- and valence-band wavefunctions respectively and $|\chi_v\rangle$ is the pseudo-wavefunction replacing $|\phi_v\rangle$. In this way, the rapid variations in the near-core region, contained in $|\phi_v\rangle$, can be cancelled out.

It is now convenient to define a projection operator which projects any state onto the core states (Philips & Kleinman, 1959):

$$\widehat{P} = \sum_{c'} |\phi_{c'}\rangle\langle\phi_{c'}| \; . \tag{3.2.2}$$

satisfying $(1 - \widehat{P})\,|\phi_c\rangle = 0$, and more importantly

$$(1 - \widehat{P})\,|\chi_v\rangle = |\phi_v\rangle \; . \tag{3.2.3}$$

The energy eigenvalue equation satisfied by a valence-band state is

$$\widehat{T}\,|\phi_v\rangle + \widehat{V}(\vec{r})\,|\phi_v\rangle = \varepsilon\,|\phi_v\rangle \; , \tag{3.2.4}$$

is also satisfied by the core states, $|\phi_c\rangle$, with the same potential $V(\vec{r})$. Bringing in (3.2.2) and (3.2.3), this equation can be re-written as

$$(\widehat{T} + \widehat{V}^{ps})\,|\chi_v\rangle = \varepsilon\,|\chi_v\rangle \; , \tag{3.2.5}$$

with $\widehat{V}^{ps}$ the *pseudopotential* given by

$$\widehat{V}^{ps} = \widehat{V}(\vec{r}) + \sum_{c'} (\varepsilon - \varepsilon_{c'})\,|\phi_{c'}\rangle\langle\phi_{c'}| \; . \tag{3.2.6}$$

This pseudo-Schrödinger equation known as a pseudopotential equation provides the same eigenvalues as the original Schrödinger equation (3.2.4) *but* the core levels are missing. Despite this achievement, this equation is still not perfect: $\widehat{V}^{ps}$ involves an integral, non-local, operator, $\widehat{T}$ as well as the eigenvalue we are trying to find.

From (3.2.3), we have

$$\langle\chi_v\,|\,\chi_v\rangle - \langle\phi_{c'}|\,\widehat{P}\,|\phi_{c'}\rangle = \langle\phi_v\,|\,\phi_v\rangle \; , \tag{3.2.7}$$

which reveals another hole in this simple theory: if the pseudo-wavefunction is normalised then the true wavefunction is not. The difference between the true density and the pseudo charge density is termed *orthogonalisation hole* and is a consequence of the non-locality of $\widehat{V}^{ps}$.

Performing calculations with nodeless pseudo-wavefunctions instead of the rapidly varying true wavefunctions would seem advantageous; however, reliable pseudopotentials are relatively complex. Applying these seems useful only if these can be transferred from one system to another, *i.e.*, from a single atom to an atom in a molecule or solid, which implies that the core states must be considered *frozen*. These are not supposed to change in different chemical environment.

There are two ways of determining pseudopotentials: either by means of model potentials, or through *ab initio* techniques.

Traditionally, empirical pseudopotentials constructed by using some experimental data have been adopted (Austin *et al.*, 1962; Shaw Jr. & Harrison, 1967). These empirical pseudopotentials however have a weak point that the charge density does not coincide with that of the real atom even outside the core region ($r > r_c$).

3.3 *Ab initio* pseudopotentials

The majority of the pseudopotentials currently used in electronic-structure calculations are generated from all-electron atomic calculations. Within the density function theory, this is done by assuming a spherical screening approximation and solving the radial Kohn-Sham equation

$$\left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{l(l+1)}{2r^2}+V([n];r)\right]\phi_{nl}(r)=\varepsilon_{nl}\,\phi_{nl}(r)\ , \tag{3.3.1}$$

where

$$V([n];r)=-\frac{Z}{r}+V^{\mathrm{H}}([n];r)+V_{xc}^{\mathrm{LDA}}(n(r))\ , \tag{3.3.2}$$

with $n(r)$ the sum of the electron densities for the occupied wavefunctions $\phi_{nl}(r)=r\,R_l(r)$; l defining the quantum state. This method has been extented to include GGA corrections (Juan *et al.*, 1995; Fuchs *et al.*, 1998).

Most pseudopotentials are constructed such that they satisfy four general condition (Zunger & Cohen, 1978; Hamann *et al.*, 1979; Kerker, 1980; Bachelet *et al.*, 1982; Vanderbilt, 1985):

1. The generated valence pseudo-wavefunction generated from the pseudopotential should be nodeless;

2. *Norm conservation*: The normalised radial pseudo-wavefunction, $\phi_l^{ps}(r)$, is equal to the normalised all-electron wavefunction, $\phi_l^t(r)$, beyond a chosen cut-off radius r_{c_l} (Kerker, 1980):

$$R_l^{ps}(r)\Big|_{r=r_{c_l}}=R_l^t(r)\Big|_{r=r_{c_l}}\qquad\text{with}\qquad r>r_{c_l}\ , \tag{3.3.3}$$

   or converges rapidly to that value (we have omitted the principal quantum number n for simplicity);

3. In the core region ($r<r_{c_l}$) The charge density associated the two wavefunctions should be equal (Hamann *et al.*, 1979)

$$\int\limits_o^{r_{c_l}} r^2\,|R_l^{ps}(r)|^2\,dr=\int\limits_o^{r_{c_l}} r^2\,|R_l^t(r)|^2\,dr\ ; \tag{3.3.4}$$

   This restriction may pose a serious problem when expanding the total energy using a plane-wave basis. This is the case of first-row transition metals characterised by localised d-electrons. Some attempts have been made to generate *smooth* pseudopotential using a plane-wave set (Rappe *et al.*, 1990; Troullier & Martins, 1991). Vanderbilt (1990), and independently Blöchl (1990), dropping the norm-conservation condition, were able to overcome this difficult. Using lower kinetic energy cut-offs for the plane-wave expansion, these authors introduced the so-called *soft-pseudopotentials*.

4. The all-electron and pseudo valence eigenvalues are identical for a particular atomic configuration, *i.e.*, $\varepsilon_{ps} = \varepsilon_t$;

If a given pseudopotential meets the above conditions, it is commonly referred to as a *norm-conserving pseudopotential*.

Once the pseudo-wavefunction is obtained, the screened pseudoptential can be recovered by inverting the radial Schrödinger equation (3.3.1):

$$V_l^{ps}(r) = \varepsilon_l - \frac{l(l+1)}{2r^2} + \frac{1}{2r\, R_l^{ps}(r)} \frac{\mathrm{d}^2}{\mathrm{d}r^2} \left[r\, R_l^{ps}(r)\right] . \tag{3.3.5}$$

The pseudopotential concept requires that the scattering properties of the pseudo and all-electron atom are the same for the pseudo-wavefunction at the specific *matching radius*, *i.e.*, the core radius r_{c_l} (Hamann *et al.*, 1979). This radius controls the overall accuracy and 'transferability' of the pseudopotential:

Transferability criterion: The first order energy derivatives of the radial logarithmic derivative of the pseudo- and atomic-wavefunction coincide with each other at r_{c_l},

$$-\frac{1}{2}\left\{\frac{\partial}{\partial r}\left[\frac{\partial}{\partial \varepsilon} \ln R(r,\varepsilon)\right]_{\varepsilon=\varepsilon_l}\right\}_{r=r_{c_l}} = \frac{1}{r_{c_l}^2\, R^2(r_{cl},\varepsilon_l)} \int\limits_0^{r_{c_l}} r^2 \left|R_l^t(r,\varepsilon_l)\right|^2 dr \tag{3.3.6}$$

3.3.1 Bachelet, Hamann and Schlüter pseudopotentials

Bachelet, Hamman and Schlüter (1982), introducing a slight modification to their previous work (Hamann *et al.*, 1979), provided a complete tabulation of pseudopotentials from hydrogen to plutonium.

These are the pseudopotentials used in the electronic structure calculations described in this thesis. Due to this fact, a detailed description of these norm-conserving pseudopotentials, hereby referred to as BHS, is presented next.

The following is the process of constructing the BHS pseudopotential.

Step 1 The first step pseudopotential $V_1(r)$ is constructed from an all-electrons atomic potential, from which the Coulomb singularity at $r \approx 0$ has been removed, is written as

$$V_1(r) = V_{atom}(r)\left[1 - f_{cut}(r)\right] + c_l\, f_{cut}(x) , \tag{3.3.7}$$

where $V_{atom}(r)$ is an atomic potential. $f_{cut}(x)$ a smooth cut-off function, defined as

$$f_{cut}(x) = \exp\left(-x^{\lambda}\right) \quad \text{with} \quad x \equiv \frac{r}{r_{c_l}} , \tag{3.3.8}$$

satisfying

$$\lim_{r\to 0} f_{cut}(x) = 1 \quad \text{and} \quad \lim_{r\to\infty} f_{cut}(x) = 0 . \tag{3.3.9}$$

r_{c_l} a core radius (cut-off radius) and $\lambda = 3.5$. The adjustable parameter c_l in (3.3.7) is adjusted so that the lowest solution $R_1^{ps}(r)$ satisfying the radial Schrödinger equation (3.3.1) has an energy ε_{ps} identical to ε_t.

The first-stage pseudo-wavefunction $\chi_1(r)$ satisfies

$$\chi_1(r) = \gamma\,\phi(r) \tag{3.3.10}$$

is obtained by solving the eigenvalue problem with $V_1(r)$. $\phi(r)$ is the full core valence wavefunction for $r > r_{c_l}$, *i.e.*, r_{c_l} delimits the core of the pseudo-atom; γ an adjustable parameter.

Step 2 The second-stage pseudo-wavefunction is defined as

$$\chi_2(r) \equiv r\,R_2^{ps}(r) = \gamma\Big[\chi_1(r) + \delta\,r^{l+1}\,f_{cut}(x)\Big]\;, \tag{3.3.11}$$

where δ is and adjustable parameters taken to be the smaller solution of

$$\gamma^2 \int_0^\infty \Big[\chi_1(r) + \delta\,r^{l+1}\,f_{cut}(x)\Big]^2\,dr = 1\;, \tag{3.3.12}$$

which quantifies the condition of norm conservation of the pseudo-wavefunctions.

Step 3 The second-stage pseudoptential is obtained by inverting the Schrödinger equation, (3.3.5), as follows

$$\begin{aligned} V_2(r) &= V_1(r) + \frac{1}{2}\,\frac{\gamma\,\delta\,r^{l+1}}{\chi_2(r)}\,f_{cut}(x) \\ &\times\Big\{\frac{\lambda^2\,x^{2\lambda} - [2\lambda\,l + \lambda(\lambda+1)]x^{\lambda}}{r^2} + 2\varepsilon - 2V_1(r)\Big\} \end{aligned} \tag{3.3.13}$$

Step 4 The final, unscreened, pseudopotential $V^{ps}(r)$ is obtained by subtracting the Hartree potential, $V^{\mathrm{H}}(r)$, and exchange-correlation potential, $V^{xc}(r)$ due to the valence electrons from $V_2(r)$,

$$V^{ps}(r) = V_2(r) - \Big(\int \frac{n_{ps}(r')}{|r-r'|}\,d^3r' + \frac{\delta E_{xc}[n_{ps}]}{\delta n_{ps}}\Big)\;, \tag{3.3.14}$$

where

$$n_{ps}(r) = \sum_{occ} \frac{1}{r^2}\,|R_2^{ps}(r)|^2 \tag{3.3.15}$$

defines the valence pseudo charge density, with the sum over all occupied valence states.

Step 5 The $V^{ps}(r)$ is divided into two parts. One is a local part $V_{local}(r)$ and the other is a non-local part, $V^{l}_{non-local}(r)$ dependent on the angular momentum l:

$$V^{ps}(r) = V_{local}(r) + \sum_l |\,l\!>V_{non-local}(r)<\!l\,|\;. \tag{3.3.16}$$

To allow the use of these pseudopotentials, BHS fitted their pseudoptentials to a basis of convenient functions to obtain the required parameters for all the atoms.

Assuming that the core potential arises from Gaussian-like core charges, the local (l-independent) and non-local (l-dependent) components of the pseudopotentials were defined as

$$V_{local}(r) = -\frac{Z_v}{r}\Big[\sum_{i=1}^{2} c_i \,\mathrm{erf}(r\sqrt{\alpha_i})\Big] \quad \text{with} \quad c_1 + c_2 = 1 \,, \tag{3.3.17}$$

with Z_v being the valence charge, and

$$V^l_{non-local}(r) = \sum_{i=1}^{3} (A^l_i + r^2\, A^l_{i+3})\, \mathrm{e}^{-\alpha_i r^2} \,, \tag{3.3.18}$$

respectively. 'erf' is the error function and α_i is the inverse of the extent of the core charge density. The parameters c_i, α_i, and A_i are given in (Bachelet *et al.*, 1982) for the atoms H to Pu.

As mentioned previously, the first stage in the construction of a pseudopotential is the solution of the all-electron Kohn-Sham(-Dirac) equations for the atom to obtain its one-particle energy levels, ε_l, labelled by the angular momentum, l for light atoms, and total angular momentum $j = l \pm \frac{1}{2}$ for heavier elements.

This is done by choosing a configuration leading to a spherically symmetric charge density. For the carbon atom, for example, a possible configuration is obviously $1s^2\, 2s^2\, 2p^2\, 3d^0$. This is the tabulated configuration for the neutral, isolated, carbon atom, which excludes any d-electron. For the case of a solid, with the wavefunction made up of linear combinations of all atomic states, including d-levels, it is not clear if the influence of these states can be simply ignored.

If all these states, s, p and d, are taken into account, the resulting pseudopotential will then possess the same valence states as the all-electron atom. To accommodate this, two different configurations are used to solve the KS equation: $1s^2\, 2s^2\, 2p^2$ and $1s^2\, 2s^{0.75}\, 2p^1\, d^{0.25}$, for $l = 0, 1$ and $l = 2$ respectively. The fraction occupancy of the s-shell is chosen to avoid 'bumps' in the potential. Different configurations are used for different elements.

These configurations are then used to generate the all-electron spin density of the atom resulting from the solution of the Kohn-Sham equations. The associated all-electron potential is simply $V_{atom}(r)$ in (3.3.7), defining the first-stage pseudopotential V^{ps}_1.

Choice of a cut-off radius

The cut-off radius is chosen to lie between the outermost node and the outermost extremum of the all-electron radial wavefunction. Analitically,

$$r_{c_l} = \frac{r_{max}}{c_{c_l}} \,, \tag{3.3.19}$$

where r_{max} is the radius of the outermost extremum of the radial wavefunction; c_{c_l} is an adjustable parameter, typically $1.5 < c_{c_l} < 2.0$. (Bachelet *et al.*, 1982)

If r_{c_l} is close to the outermost node, then the generated pseudopotentials are strong and very l-dependent; if it is closer to the outer most extremum, the pseudopotentials are smoother and l-independent but include large errors in the pseudo-radial wavefunction.

4

The Real Space Cluster Method

4.1 Introduction

In Chapter 2, we described the foundations of density functional theory. In this chapter, we present an implementation of this theory, aiming the determination of the structural, vibrational, optical and electrical properties of defects in covalent semiconductors, in particular silicon.

Following the prescription of Kohn and Sham, we describe a numerical method to solve the set of one-electron equations offered by these authors. This includes the determination of the equilibrium structure of the electronic system, which proceeds via the minimisation of its total energy with respect to the atomic positions.

The conversion of Kohn-Sham formalism to a matrix representation, together with some numerical implementations, has allowed its translation into computer language: AIMPRO – an acronym for *ab initio modelling program*. All the matrix operations involved have been modified to enable the use of the code on massively parallel computers like the CRAY-T3E, using the PBLAS and SCALAPACK routines.

4.2 Methodology

The total electronic wavefunction is assumed to be a sum of single particle wavefunctions for each of the occupied valence-band states ψ_λ. The spin-polarised wavefunction of the λ-th electronic level is defined as a linear combination of localised orbitals $\phi_i(\vec{r} - \vec{R}_i)$:

$$\psi_\lambda(\vec{r}, \sigma) = \chi_\sigma \sum_i c_i^\lambda \, \phi_i(\vec{r} - \vec{R}_i) \quad \text{with} \quad c_i^\lambda \in \mathbb{R} \, . \tag{4.2.1}$$

In this way the Kohn-Sham differential equations are converted to matrix equations for c_i^λ. The localised orbitals are taken to be of the form:

$$\phi_i(\vec{r} - \vec{R}_i) \propto (x - R_{ix})^{n_1} \, (y - R_{iy})^{n_2} \, (z - R_{iz})^{n_3} \, \mathrm{e}^{-\alpha_i (\vec{r} - \vec{R}_i)^2} \quad \text{with} \quad n \in \mathbb{N} \, , \tag{4.2.2}$$

defining what it is called a set of Cartesian Gaussian orbitals. If all the exponents n are zero, $\phi_i(\vec{r} - \vec{R}_i)$ defines a s-orbital of spherical symmetry. Orbitals of p-symmetry

correspond to one of these integers being unity and the others zero, whereas five d-like and one s-like orbital can be generated if $n_1 + n_2 + n_3 = 2$.

The basis functions are real and hence so are all matrix elements. We can therefore drop complex conjugates from the equations. As shown before (§2.6), it is possible to convert density functionals into spin-density ones. This is straightforward for the case of the kinetic, Hartree and external potential terms, since these only depend on the total density and *not* on the particular spin state like the one-particle exchange and correlation term. Due to this, all quantities but the exchange-correlation are represented implicitly for a single spin state σ, *i.e.*, $\sigma = \uparrow$ or $\downarrow$.

The total charge density for the electronic system is

$$n_\sigma(\vec{r}) = \sum_\lambda \gamma_\lambda^\sigma \left|\psi_\lambda(\vec{r},\sigma)\right|^2 , \tag{4.2.3}$$

which is, in terms of the density matrix b_{ij}^σ, equivalent to

$$n_\sigma(\vec{r}) = \sum_{ij} b_{ij}^\sigma\, \phi_i(\vec{r}-\vec{R}_i)\, \phi_j(\vec{r}-\vec{R}_j) , \tag{4.2.4}$$

$$b_{ij}^\sigma = \sum_\lambda^{\text{occ}} \delta_{\sigma\sigma_\lambda}\, \gamma_\lambda^\sigma\, c_i^\lambda\, c_j^\lambda . \tag{4.2.5}$$

The sum is over occupied orbitals λ with spin σ. The total charge density $n(\vec{r})$ is defined as:

$$n(\vec{r}) = \sum_\sigma n_\sigma(\vec{r}) = \sum_{ij} b_{ij}\, \phi_i(\vec{r}-\vec{R}_i)\, \phi_j(\vec{r}-\vec{R}_j) , \tag{4.2.6}$$

$$b_{ij} = \sum_\sigma b_{ij}^\sigma . \tag{4.2.7}$$

Using the basis of localised orbitals defined above, the matrix elements of the Hamiltonian for the Kohn-Sham equations (2.3.16–2.3.17) can be written, for a spin up charge density ($\sigma = \uparrow$) , as

$$(\mathcal{H}_{ij})^\uparrow = T_{ij} + V_{ij}^{ext} + V_{ij}^{\mathrm{H}} + (V_{ij}^{xc})^\uparrow , \tag{4.2.8}$$

with

$$T_{ij} = -\frac{1}{2}\int \phi_i(\vec{r}-\vec{R}_i)\, \nabla^2\, \phi_j(\vec{r}-\vec{R}_j)\, d^3r \tag{4.2.9}$$

$$V_{ij}^{ext} = \int \phi_i(\vec{r}-\vec{R}_i)\, \overbrace{\sum_\alpha V_\alpha^{ps}(\vec{r}-\vec{R}_\alpha)}^{v_{ext}(\vec{r})}\, \phi_j(\vec{r}-\vec{R}_j)\, d^3r \tag{4.2.10}$$

$$V_{ij}^{\mathrm{H}} = \iint \phi_i(\vec{r}-\vec{R}_i)\, \frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\, \phi_j(\vec{r}\,'-\vec{R}_j)\, d^3r\, d^3r' \tag{4.2.11}$$

$$(V_{ij}^{xc})^\uparrow = \int \phi_i(\vec{r}-\vec{R}_i)\, v_{xc}^\uparrow\, \phi_j(\vec{r}-\vec{R}_j)\, d^3r , \tag{4.2.12}$$

where

$$v_{xc}^\uparrow = \epsilon_{xc}(n_\uparrow, n_\downarrow) + \frac{\partial \epsilon_{xc}(n_\uparrow, n_\downarrow)}{\partial n_\uparrow(\vec{r})} \Big[n_\uparrow(\vec{r}) + n_\downarrow(\vec{r}) \Big] \tag{4.2.13}$$

The evaluation of T_{ij} and V_{ij}^{ext} can be carried out straightforwardly. Unfortunately, the same is not true for the other two terms, (4.2.11) and (4.2.12).

For example, in terms of the localised basis (4.2.1), the Hartree energy may be written as

$$V_{ij}^{\mathrm{H}} = \frac{1}{2} \sum_{ijkl} c_i^\lambda \, c_j^\lambda \, c_k^\lambda \, c_l^\lambda \iint \frac{\phi_i(\vec{r} - \vec{R}_i) \, \phi_j(\vec{r} - \vec{R}_j) \, \phi_k'(\vec{r}' - \vec{R}_k) \, \phi_l'(\vec{r}' - \vec{R}_l)}{|\vec{r} - \vec{r}'|} \, d^3r \, d^3r' \ . \tag{4.2.14}$$

It is then obvious, that the evaluation of the Hartree energy requires $\mathcal{O}(N^4)$ integrals, where N is the number of basis functions, which is prohibitively large for a cluster where N might be 1000 or more.

In the next two sections, we present the approximations used to evaluate the Hartree and exchange-correlation integrals within AIMPRO. This is followed by its full matrix formulation, the approach to self-consistency and the structural optimisation procedure.

4.3 The Hartree energy

The exact Hartree energy is, from (1.10.4),

$$U_{\mathrm{H}}[n] = \frac{1}{2} \iint \frac{n(\vec{r}) \, n(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d^3r \, d^3r' \ . \tag{4.3.1}$$

It is obvious from (4.2.5) that the evaluation of the Hartree energy term, U_{H}, involves a four-centre integral of the wavefunction bases. Currently, there is no way to solve this double integral for clusters of useful size. It is therefore necessary to introduce an approximate, but analytic expression for the Hartree energy from which the corresponding potentials can be found.[1]

According to Dunlap *et al.* (1979) and Jones and Sayyash (1986), this can be achieved by expanding the charge density for each spin as,

$$\widetilde{n}_\sigma(\vec{r}) = \sum_k d_k^\sigma \, g_k(\vec{r}) \quad \text{for} \quad \widetilde{n}(\vec{r}) = \sum_\sigma \widetilde{n}_\sigma(\vec{r}) \ . \tag{4.3.2}$$

defining an intermediate fit to the charge density (Sambe & Felton, 1975).

The difference between the Hartree energy due to the true, $n(\vec{r})$ and intermediate fit charge density, $\widetilde{n}(\vec{r})$, can be expressed as,

$$\Delta U_{\mathrm{H}} = \frac{1}{2} \iint \frac{[n(\vec{r}) - \widetilde{n}(\vec{r})] \, [n(\vec{r}') - \widetilde{n}(\vec{r}')]}{|\vec{r} - \vec{r}'|} \, d^3r \, d^3r' \ , \tag{4.3.3}$$

which defines the error in estimating the Hartree term using (4.3.2). As a function of U_{H} and ΔU^{H}, the true Hartree energy can be written as,

$$U_{\mathrm{H}}[n] = \widetilde{U}_{\mathrm{H}}[n; \widetilde{n}] + \Delta U_{\mathrm{H}} \tag{4.3.4}$$

[1] A similar approach is persued for the treatment of the exchange-correlation contribution.

which allows us to define $\widetilde{U}_{\mathrm{H}}$ as

$$\widetilde{U}_{\mathrm{H}}[n;\widetilde{n}] = \iint \frac{n(\vec{r})\,\widetilde{n}(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3r\, d^3r' - \frac{1}{2}\iint \frac{\widetilde{n}(\vec{r})\,\widetilde{n}(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3r\, d^3r' \; . \tag{4.3.5}$$

The replacement is exact when $\widetilde{n} = n$, *i.e.*, when $\widetilde{n}(\vec{r}) \to n(\vec{r})$, $\Delta U_{\mathrm{H}} \to 0$.

Now, we expand the density in terms of a basis set $g_k(\vec{r})$ so that

$$\widetilde{n}(\vec{r}) = \sum_k c_k\, g_k(\vec{r}) \; , \tag{4.3.6}$$

and c_k is chosen to minimise ΔU^{H}. Replacing this density, and $n(\vec{r})$ by (4.2.7), in (4.3.5) results in

$$\widetilde{U}_{\mathrm{H}} = \sum_{ijk} c_k\, b_{ij}\, t_{ijk} - \frac{1}{2}\sum_{kl} c_k\, c_l\, G_{kl} \; , \tag{4.3.7}$$

with

$$t_{ijk} = \iint \frac{\phi_i(\vec{r}-\vec{R}_i)\,\phi_j(\vec{r}-\vec{R}_j)}{|\vec{r}-\vec{r}'|}\, g_k(\vec{r}')\, d^3r\, d^3r' \; , \tag{4.3.8}$$

$$G_{kl} = \iint \frac{g_k(\vec{r})\, g_l(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3r\, d^3r' \; . \tag{4.3.9}$$

Differentiating ΔU_{H} with respect to c_k to determine its minimum, using (4.3.7), *i.e.*,

$$\frac{\partial(\Delta U_{\mathrm{H}})}{\partial c_k} = \frac{\partial}{\partial c_k}\left[\frac{1}{2}\iint \frac{n(\vec{r})\, n(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3r\, d^3r' - \sum_{ijk} c_k\, b_{ij}\, t_{ijk} + \frac{1}{2}\sum_{kl} c_k\, c_l\, G_{kl}\right] = 0 \; , \tag{4.3.10}$$

yields,

$$\sum_l G_{kl}\, c_l = \sum_{ij} t_{ijk}\, b_{ij} \; . \tag{4.3.11}$$

Combining (4.3.7) and (4.3.11), we obtain the final expression for the approximation of the Hartree energy,

$$\widetilde{U}_{\mathrm{H}} = \frac{1}{2}\sum_{kl} c_k\, c_l\, G_{kl} \; . \tag{4.3.12}$$

We now proceed to calculate the matrix elements of the Hartree potential. The potential corresponding to a particular energy derivative $U[n]$ is by definition the functional derivative

$$V[n] = \frac{\delta U[n]}{\delta n(\vec{r})} \; . \tag{4.3.13}$$

with the corresponding matrix elements given by

$$V_{ij}[n] = \int \phi_i(\vec{r})\, V\, \phi_j(\vec{r})\, d^3r \; . \tag{4.3.14}$$

Inserting (4.3.13) into (4.3.14), we get

$$V_{ij}[n] = \int \frac{\delta U[n]}{\delta n(\vec{r})} \frac{\partial n(\vec{r})}{\partial b_{ij}} \, d^3r \; . \tag{4.3.15}$$

Differentiating $n(\vec{r})$, defined as (see 4.2.7)

$$n(\vec{r}) = \sum_{ij} b_{ij} \, \phi_i(\vec{r}) \, \phi_j(\vec{r}) \; , \tag{4.3.16}$$

with respect to b_{ij}, we have

$$\frac{\partial n(\vec{r})}{\partial b_{ij}} = \phi_i(\vec{r}) \, \phi_j(\vec{r}) \quad \Rightarrow \quad V_{ij}[n] = \frac{\partial U[n]}{\partial b_{ij}} \; . \tag{4.3.17}$$

Using (4.3.11) and (4.3.12), the matrix elements of the Hartree potential between the basis functions can be written as,

$$\begin{aligned} V_{ij}^{\mathrm{H}} &= \frac{\partial \widetilde{U}_{\mathrm{H}}}{\partial b_{ij}} \\ &= \frac{\partial}{\partial b_{ij}} \Big[\frac{1}{2} \sum_{kl} c_k \, c_l \, G_{kl} \Big] = \frac{\partial}{\partial b_{ij}} \Big[\sum_{ijk} c_k \, b_{ij} \, t_{ijk} \Big] \\ &= \sum_k c_k \, t_{ijk} \; . \end{aligned} \tag{4.3.18}$$

The full matrix formulation will be presented in the next section.

Having found a way to evaluate the Hartree term, its now necessary to choose the form of the intermediate fit to the charge density, $\widetilde{n}(\vec{r})$, or indirectly g_k. The charge density is taken to be of the form

$$\widetilde{n}(\vec{r}) = Z \left(\frac{\alpha}{\pi} \right) \mathrm{e}^{-\alpha(\vec{r}-\vec{R})^2} + \sum_k c_k \, \overbrace{\left[1 - \frac{2}{3} \beta_k \, (\vec{r} - \vec{R}_k)^2 \right] \mathrm{e}^{-\beta_k (\vec{r}-\vec{R}_k)^2}}^{g_k} \; , \tag{4.3.19}$$

defining a modified Gaussian fitting function.

The set of k-modified Gaussian functions gives rise to a potential which is short-ranged, *i.e.*,

$$\int \frac{g_k(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} \, d^3r' = \frac{3}{2\pi} \, \beta_k \, \mathrm{e}^{-\beta_k (\vec{r}-\vec{R}_k)^2} \; , \tag{4.3.20}$$

producing a charge of zero. Choosing fixed coefficients β, from an initialising total charge of an atom or ion, their contribution to the Hamiltonian remains unchanged during a self-consistent cycle. Consequently, the terms (4.3.20) can then be added to the external potential term—the pseudopotential term—leading to a considerable speed up in the code.

4.4 The exchange-correlation energy

As defined previously (2.5.9), the exchange-correlation energy is given, in the LSD approximation, by

$$E_{xc}[n_\uparrow; n_\downarrow] = \int \epsilon_{xc}(n_\uparrow, n_\downarrow) \; [n_\uparrow(\vec{r}) + n_\downarrow(\vec{r})] \; d^3r \; , \tag{4.4.1}$$

is replaced by an approximate one $\widetilde{E}_{xc}$ involving an approximate density, $\widetilde{n}_\sigma$. In an analogue way to the treatment of the Hartree term, an intermediate fit to the charge density is used,

$$\widetilde{E}_{xc} = \int \epsilon_{xc}(\widetilde{n}_\uparrow, \widetilde{n}_\downarrow) \, [\widetilde{n}_\uparrow + \widetilde{n}_\downarrow] \, d^3r \ . \qquad (4.4.2)$$

The intermediate fit is expanded in terms of a different set of Gaussian functions to those in (4.3.6),

$$\widetilde{n}_\sigma(\vec{r}) = \sum_k d_k^\sigma \, h_k(\vec{r}) \ , \qquad (4.4.3)$$

where d_k^σ is found from minimising

$$\int \left\{ n_\sigma(\vec{r}) - \widetilde{n}_\sigma(\vec{r}) \right\}^2 d^3r \ , \qquad (4.4.4)$$

defining the electrostatic energy associated with the error due to the intermediate fit to the charge density. A least-squares minimisation, *i.e.*,

$$\frac{\partial}{\partial d_k^\sigma} \left\{ \int \left[n_\sigma(\vec{r}) - \widetilde{n}_\sigma(\vec{r}) \right]^2 d^3r \right\} = 0 \ , \qquad (4.4.5)$$

leads to the set of equations:

$$\sum_l H_{kl} \, d_l^\sigma - \sum_{ij} u_{ijk} \, b_{ij}^\sigma = 0 \ , \qquad (4.4.6)$$

where,

$$H_{kl} = \int h_k(\vec{r}) \, h_l(\vec{r}) \, d^3r \ , \qquad (4.4.7)$$

$$u_{ijk} = \int \phi_i(\vec{r} - \vec{R}_i) \, \phi_j(\vec{r} - \vec{R}_j) \, h_k(\vec{r}) \, d^3r \ . \qquad (4.4.8)$$

Note that the integrals are the same for each spin-index σ and that u_{ijk} are simply proportional to t_{ijk} if h_k is chosen as in (4.3.19). Again, this also saves a considerable amount of computer time.

For the spin-averaged case,

$$\widetilde{E}_{xc}[\widetilde{n}] = \sum_k d_k \int h_k(\vec{r}) \, \epsilon_{xc}(\widetilde{n}) \, d^3r \ . \qquad (4.4.9)$$

If h_k is chosen to be a positive definite localised function such as a Gaussian,[2] then each integral is proportional to the average value of the exchange-correlation density under h_k,

$$\frac{\int h_k(\vec{r}) \, \epsilon_{xc}(\widetilde{n}) \, d^3r}{\int h_k(\vec{r}) \, d^3r} \equiv \langle \epsilon_{xc}(\widetilde{n}) \rangle_k \ . \qquad (4.4.10)$$

Since $\epsilon_{xc}(n)$ varies slowly with $n(\vec{r})$ (see Fig. 4.1), we can assume that

$$\langle \epsilon_{xc}(\widetilde{n}) \rangle_k \approx \epsilon_{xc}(\langle \widetilde{n} \rangle_k), \qquad (4.4.11)$$

[2] In other words, we consider $h_k(\vec{r})$ to be a non-normalised probability distribution function.

$$\langle \widetilde{n} \rangle_k = \frac{1}{I_k} \sum_l d_l\, H_{kl} \quad \text{for} \quad I_k \equiv \int h_k(\vec{r})\, d^3 r \; . \tag{4.4.12}$$

This approximation is equivalent to replace the exact exchange-correlation density at $\vec{r}$ by its homogeneous electron gas value for the average density $\langle \widetilde{n} \rangle_k$.

Assuming that the single particle exchange-correlation energy can be expanded in power series of $n(\vec{r})$, we write for the spin average case (Jones, 1988),

$$n(\vec{r})\, \epsilon_{xc}(n(\vec{r})) = n_0(\vec{r}) \left[\frac{n(\vec{r})}{n_0(\vec{r})} \right]^{1+s_0} \; . \tag{4.4.13}$$

For Ceperley-Alder exchange-correlation (Ceperley & Alder, 1980), $n_0 = 1$ a.u.$^{-3}$ and $s_0 = 0.30917$. For the Dirac-Slater exchange (2.8.2), s_0 is exactly $\frac{1}{3}$.

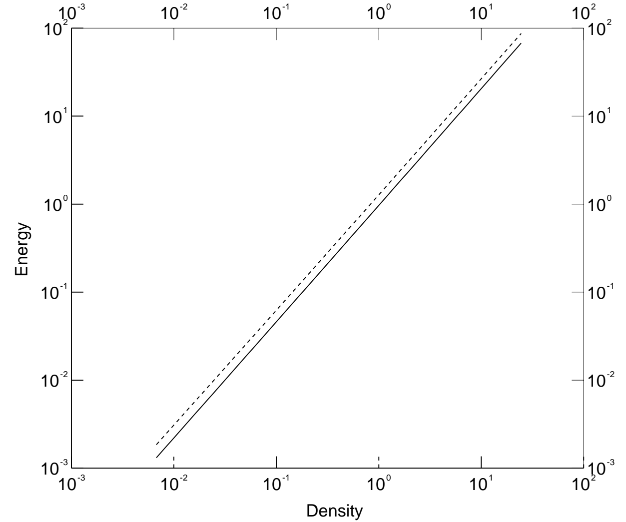


Figure 4.1: Variation of polarised (full) and non-polarised (dashed) exchange-correlation energy ($\times -1.0$ a.u.) per unit volume with density.

Let us replace the fitted value s_0 by a variable, say p, and consider the following function $f_k(p)$ defined as

$$f_k(p) = \ln \left[\frac{\sum_l c_l\, H_{kl}}{\langle \widetilde{n} \rangle_k^p\, I_k} \right] = \ln \left(\frac{\langle \widetilde{n}^p \rangle_k}{\langle \widetilde{n} \rangle_k^p} \right) \; , \tag{4.4.14}$$

with H_{kl} defined by (4.4.8).

The function $f_k(p)$ is determined at three points:

$$f_k(p) = \begin{cases} 0 & : \quad p = 0 \quad \text{and} \quad 1 \; ; \\ \ln \left(\dfrac{\langle \widetilde{n}^2 \rangle_k}{\langle \widetilde{n} \rangle_k^2} \right) & : \quad p = 2 \; , \end{cases} \tag{4.4.15}$$

with

$$\langle \widetilde{n}^2 \rangle_k = \frac{1}{I_k} \sum_{lm} d_l\, d_m \int h_k(\vec{r})\, h_l(\vec{r})\, h_m(\vec{r})\, d^3 r \; . \tag{4.4.16}$$

Applying the Lagrange formula for polynomial interpolation to fit a quadratic to evaluate $f_k(s_0)$, we have

$$f_k(s_0) = \frac{1}{2}\, s_0\, (s_0 - 1)\, f_k(2) \; . \tag{4.4.17}$$

The function f_k assumes its minumum value of ~ -0.1 for $p = 0.5$, *i.e.*, between $p = 0$ and $p = 1$ and close to s_0. This shows that the quadratic fit is sufficiently accurate.

The final expression for the one-particle exchange-correlation energy is as follows:

$$\widetilde{E}_{xc} = \sum_k d_k\, \epsilon_k \; , \tag{4.4.18}$$

Table 4.1: Parametrisation for the approximation to the spin-polarised exchange-correlation energy. The parameters A'_i, p'_i and q'_i are for larger density values.

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

where,

$$\epsilon_k = I_k\, \epsilon_{xc}(\langle\widetilde{n}\rangle_k)\, \exp\left[f_k(s_0)\right] \quad \text{and} \quad f_k(s_0) = \frac{1}{2}\, s_0\, (s_0 - 1) \ln\left(\frac{\langle\widetilde{n}^2\rangle_k}{\langle\widetilde{n}\rangle_k^2}\right) . \tag{4.4.19}$$

Re-calling (4.4.6), the set of equations (4.4.6) can now be written as

$$\sum_l H_{kl}\, d_l = \sum_{ij} b_{ij} \int \phi_i(\vec{r} - \vec{R}_i)\, \phi_j(\vec{r} - \vec{R}_j)\, h_k(\vec{r})\, d^3r\ . \tag{4.4.20}$$

The functions $h_k(\vec{r})$ are chosen to be similar to those introduced to quantify the Hartree contribution (§4.3).

The matrix elements of the exchange-correlation potential for the spin-average case are given by:

$$V_{ij}^{xc} = \sum_{kl} \left(H^{-1}\right)_{kl} t_{ijl} \Big[\epsilon_k\, I_k + \sum_m d_m \exp\left[f_m(s_0)\right] F_{km}\Big] \tag{4.4.21}$$

with

$$\begin{aligned} F_{kl} \;=\; & \left[\frac{\partial\epsilon_{xc}(\langle n\rangle_k)}{\partial\langle n\rangle_k} - \epsilon_{xc}(\langle\widetilde{n}\rangle_l)\, \frac{s_0\,(s_0-1)}{\langle\widetilde{n}\rangle_l}\right] H_{kl} \\ & + \epsilon_{xc}(\langle\widetilde{n}\rangle_l)\, \frac{s_0\,(s_0-1)}{\langle\widetilde{n}\rangle_l^2} \sum_m d_m \int h_k(\vec{r})\, h_l(\vec{r})\, h_m(\vec{r})\, d^3r\ . \end{aligned} \tag{4.4.22}$$

4.4.1 Spin polarisation

This theory has been extended to the spin-polarised case (Lister & Jones, 1988; Jones & Briddon, 1998). The spin-polarised exchange-correlation energy is written as,

$$E_{xc}[n_\uparrow; n_\downarrow] = \sum_{i\sigma} A_i \int n_\sigma^{p_i+1}\, n_{1-\sigma}^{q_i}\, d^3r\ , \tag{4.4.23}$$

and we replace n_σ on right hand side by $\widetilde{n}_\sigma$ obtaining:

$$\tilde{E}_{xc} = \sum_{k\sigma} d_k^\sigma \overbrace{\sum_i A_i\, I_k\, \langle\widetilde{n}_\sigma^{p_i}\, \widetilde{n}_{1-\sigma}^{q_i}\rangle_k}^{\epsilon_k^\sigma}\ , \tag{4.4.24}$$

with A_i, q_i and p_i for densities $n_\sigma < 1$ and A'_i, q'_i and p'_i for higher density values (vide Table 4.1).

Now, we define the quantity f_k by:

$$\langle \widetilde{n}_\sigma^p \, \widetilde{n}_{1-\sigma}^q \rangle_k = \langle \widetilde{n}_\sigma \rangle_k^p \, \langle \widetilde{n}_{1-\sigma} \rangle_k^q \, \exp\left[f_k(p,q) \right] , \tag{4.4.25}$$

$$f(p,q) = \ln \left(\frac{\langle \widetilde{n}_\sigma^p \, \widetilde{n}_{1-\sigma}^q \rangle_k}{\langle \widetilde{n}_\sigma \rangle_k^p \, \langle \widetilde{n}_{1-\sigma} \rangle_k^q} \right) . \tag{4.4.26}$$

We now approximate f_k by the formula:

$$f_k(p,q) = \frac{1}{2} \, p \, (p-1) \, f_k(2,0) + \frac{1}{2} \, q(q-1) \, f_k(0,2) + p \, q \, f(1,1) , \tag{4.4.27}$$

which interpolates f_k between the known integer values. In this way the spin-polarised exchange-correlation energy is evaluated.

4.5 Full matrix formulation

In terms of the approximate Hartree and exchange-correlation energies, the total energy can now be written:

$$E = \sum_{ij} \left\{ T_{ij} + V_{ij}^{ps} \right\} b_{ij} + \underbrace{\frac{1}{2} \sum_{kl} c_k \, c_l \, G_{kl}}_{\widetilde{U}_{\mathrm{H}}} + \overbrace{\sum_{k\sigma} d_k^\sigma \, \epsilon_k^\sigma}^{\widetilde{E}_{xc}} + V_{i-i} , \tag{4.5.1}$$

where the coefficients b_{ij} are given by (4.2.5) and V_{i-i} is the energy of the Coulombic interaction between the ions defined in (2.2.22) and known as the Madelung energy. The fitting coefficients c_k and d_k^σ are defined in terms of b_{ij}^σ by (4.3.11) and (4.4.6).

The total energy, E, is minimised subject to an orthonormal set of wavefunctions $\{\phi_i\}$, *i.e.*

$$\sum_{ij} c_i^\lambda \, c_j^\mu \, S_{ij} = \delta_{\lambda\mu} , \tag{4.5.2}$$

where the overlap matrix $\mathbf{S}$, is defined by:

$$S_{ij} = \int \phi_i(\vec{r} - \vec{R}_i) \, \phi_j(\vec{r} - \vec{R}_j) \, d^3r . \tag{4.5.3}$$

Again, this can be achieved by introducing Lagrange undetermined multipliers, ε_λ, so that we minimise without constraint,

$$\sum_{ij\lambda} \left\{ T_{ij} + V_{ij}^{ps} - \varepsilon_\lambda \, S_{ij} \right\} c_i^\lambda \, c_j^\lambda + \widetilde{U}_{\mathrm{H}} + \widetilde{E}_{xc} + V_{i-i} , \tag{4.5.4}$$

with respect to c_i^λ.

As shown before, the matrix elements of the Hartree potential can be written as:

$$\frac{\partial \widetilde{U}_{\mathrm{H}}}{\partial c_i^\lambda} = \sum_j \left(\frac{\partial \widetilde{U}_{\mathrm{H}}}{\partial b_{ij}} \right) \left(\frac{\partial b_{ij}}{\partial c_i^\lambda} \right) = \sum_j V_{ij}^{\mathrm{H}} \, c_j^\lambda . \tag{4.5.5}$$

A similar formula applies to V_{ij}^{xc}.

The matrix elements of the Hartree and exchange-correlation potentials can now be expressed as:

$$\frac{\partial \widetilde{U}_H}{\partial c_i^\lambda} = \sum_j V_{ij}^{\mathrm{H}}\, c_j^\lambda \ , \qquad V_{ij}^{\mathrm{H}} = \sum_{kl} G_{kl}\, c_l\, \frac{\partial c_k}{\partial b_{ij}} = \sum_k c_k\, t_{ijk} \ , \tag{4.5.6}$$

$$\frac{\partial \widetilde{E}_{xc}}{\partial c_i^\lambda} = \sum_j (V_{ij}^{xc})_{\sigma_\lambda}\, c_j^\lambda \quad \text{and} \quad (V_{ij}^{xc})_\sigma = \sum_k \frac{\partial d_k^\sigma}{\partial b_{ij}^\sigma} \Big\{ \epsilon_k^\sigma + \sum_l d_l^\sigma\, \frac{\partial \epsilon_l^\sigma}{\partial d_k^\sigma} \Big\} \ . \tag{4.5.7}$$

From (4.3.11) and (4.4.6), we find:

$$\sum_l G_{kl}\, \frac{\partial c_l}{\partial b_{ij}} = t_{ijk} \ ,$$

$$\sum_l H_{kl}\, \frac{\partial d_l^\sigma}{\partial b_{ij}^\sigma} = u_{ijk} \ ,$$

respectively.

Differentiating (4.5.4) with respect to c_i^λ we get the Kohn-Sham equations:

$$\sum_j \Big\{ T_{ij} + V_{ij}^{ps} + V_{ij}^{\mathrm{H}} + (V_{ij}^{xc})_{\sigma_\lambda} \Big\}\, c_j^\lambda = \varepsilon_\lambda \sum_j S_{ij}\, c_j^\lambda \ . \tag{4.5.8}$$

We note that the total number of electrons N and the spin S are given by:

$$N = \sum_{ij\sigma} S_{ij}\, b_{ij}^\sigma \ , \tag{4.5.9}$$

$$S = \sum_{ij} S_{ij}\, (b_{ij}^\uparrow - b_{ij}^\downarrow) \ . \tag{4.5.10}$$

We can now write (4.5.8) more compactly in matrix form. (4.5.8) is written in terms of two generalised eigenvalue problems, one for each spin, as:

$$\sum_j (H_{ij} - \varepsilon\, S_{ij})\, c_j = 0 \ , \tag{4.5.11}$$

or in matrix notation as:

$$(\mathbf{H} - E\,\mathbf{S})\,\mathbf{c} = \mathbf{0} \ . \tag{4.5.12}$$

For the cluster sizes and values of the exponents of the basis sets typically used, the matrices **H** and **S** come out to be quite dense. To make this problem treatable, we decompose **S** into an upper triangular matrix times its transpose—Choleski decomposition—as

$$\mathbf{S} = \mathbf{U}^t\, \mathbf{U} \ . \tag{4.5.13}$$

U and its inverse can be evaluated in $\mathcal{O}(N^3)$ operations. Substituting (4.5.13) into (4.5.11) and multiplying the left hand side by $(\mathbf{U}^{-1})^t$, the generalised eigenvalue problem can be posed as

$$\Big[(\mathbf{U}^{-1})^t\, \mathbf{H}\, \mathbf{U}^{-1} - E\,\mathbf{I} \Big]\, \mathbf{d} = \mathbf{0} \quad \text{with} \quad \mathbf{U}\,\mathbf{c} = \mathbf{d} \ . \tag{4.5.14}$$

The eigenvalues of this can be found by a standard Householder scheme which first reduces the matrix to tridiagonal form from which the eigenvalues can be found.

4.6 Approach to self-consistency

Self-consistency is the process by which charge is redistributed throughout the cluster until a minimum energy is reached, thereby hopefully producing an accurate simulation of the 'real' system. Strictly speaking, this charge density refinement process stops when the charge density produced gives rise to the same potential as was used in the equation that determined them within a pre-determined error. The self-consistency cycle may be summarised as follows:

1. Initialisation, by choosing sets of charge density coefficients c_k and d_k^σ taken from either neutral atoms, or a previous run;

2. The Kohn-Sham equation are then solved:

$$\sum_j \left\{ T_{ij} + V_{ij}^{ps} + V_{ij}^{\mathrm{H}} + (V_{ij}^{xc})_{\sigma_\lambda} \right\} c_j^\lambda = \varepsilon_\lambda \sum_j S_{ij}\, c_j^\lambda \ . \tag{4.6.1}$$

   to determine the coefficients b_{ij}^σ, defined as $\sum_\lambda \delta_{\sigma\sigma_\lambda}\, \gamma_\lambda^\sigma\, c_i^\lambda\, c_j^\lambda$.

3. Equations (4.3.11) and (4.4.6), *i.e.*,

$$\sum_l G_{kl}\, c_l = \sum_{ij} t_{ijk}\, b_{ij} \quad \text{and} \quad \sum_l H_{kl}\, d_l^\sigma = \sum_{ij} u_{ijk}\, b_{ij}^\sigma \ ,$$

   respectively, are then used to determine the output charge density coefficients, c_k^{out} and $d_{k\sigma}^{out}$. To achieve this, the components of t_{ijk} are evaluated using (4.3.8), and contracted with b_{ij} via (4.3.11) to form a vector q_k, defined as

$$c_k = \left(G^{-1}\right)_{kl} q_k \quad \text{and} \quad d_k = \left(H^{-1}\right)_{kl} \frac{3\beta_k}{2\pi}\, q_k \ , \tag{4.6.2}$$

   respectively.

4. The next step consists of selecting a new input charge density $c_k{}'$, defined in terms of c_k^{out} and c_k.

The last step is carried by using a weighted combination as in:

$$c_k{}' = c_k + w\,(c_k^{out} - c_k) \ . \tag{4.6.3}$$

The same weighting is used to define the new spin density coefficients $d_k^{\sigma}{}'$. A similar method to that of Brent (1973) is applied here. Using a suitable technique, the total energy is defined as a function of w, with w chosen to be the value which minimises the total energy.

Effectively, the above equation is solved for a particular value of w, say w_1, giving a specific output charge density c_{1k}^{out}. Then the deviation from the self-consistency of c_k can be written as

$$e_k = \frac{(1-w)\,(c_k - c_k^{out})}{w} - \frac{c_{1k}^{out} - c_k^{out}}{w_1} \,, \tag{4.6.4}$$

and choose w by minimising the electrostatic energy of a *pseudo-charge density* e_k defined as $\sum_k e_k\, g_k(\vec{r})$. w is then determined from the minimisation of the electrostatic energy

$$\frac{1}{2}\sum_{kl} e_k\, G_{kl}\, e_l \;. \tag{4.6.5}$$

It is possible to generalise this procedure so that the predicted charge density is built up from the coeficients from previous iterations.

In practice, the self-consistency cycle converges exponentially quickly, taking between four to ten iterations with the difference in the input and output Hartree energies typically becoming less than 10^{-5} a.u. Convergence is particularly rapid when there is a gap between the highest filled and lowest empty level but problems can arise when this gap is very small or vanishes. These are often related to an attempted crossing of an occupied and unoccupied energy level whereupon the charge density changes discontinuously. This effect is known as *charge sloshing.* This is discussed by Jones *et al.* (1995) for the case of substitutional Ni^- in Si.

This effect can be avoided by “smearing out” the occupation of levels by using Fermi statistics.

4.6.1 Fermi statistics

The simulation of a ‘finite temperature’ can be used to overcome convergency problems when optimising the charge density coefficients.

Let us assume that the Kohn-Sham level ε_λ is occupied by γ_λ electrons. This means that the energy to be be minimised now includes an entropy term as well as a term constraining the total number of electrons to N:

$$F = E - T\left\{-\,k_B \sum_\lambda \left[\gamma_\lambda \,\ln\gamma_\lambda + (1-\gamma_\lambda)\,\ln(1-\gamma_\lambda)\right] - \mu\left[\sum_\lambda \gamma_\lambda - N\right]\right\} . \tag{4.6.6}$$

using Stirling’s approximation for $x!$ to expand the entropy term, S (the expression in braces), and summing over all orbitals λ. Minimising the free energy F with respect to γ_λ and μ gives

$$N = \sum_\lambda \underbrace{\frac{1}{\mathrm{e}^{(\varepsilon_\lambda - \mu)/k_B T} + 1}}_{\gamma_\lambda} \;. \tag{4.6.7}$$

This equation, reflecting the conservation of the number of electrons in the λ-th state, is solved for μ using a numerical recipe described in Press *et al.* (1982), based on the Wijngaarden-Dekker-Brent method.

In practice, k_BT is taken to be about 0.04 eV. Often, where we have two energy levels separated by 0.1 eV that 'cross' in the approach to self-consistency, this will remove the discontinuous change, but when self-consistency is achieved, provided the final splitting is more than 0.04 eV one state is found to be fully occupied and the other empty. In this sense, we are using variable filling purely as a computational tool and are not attempting to simulate materials at finite temperatures.

4.7 Evaluation of atomic forces

Once the self-consistent charge density has been found, then the force acting on each atom can be evaluated. It is essential to determine the forces accurately in order to relax the cluster and calculate its vibrational modes. The force acting on atom i is given by:

$$\vec{f}_i = -\vec{\nabla}_{R_{i\alpha}} E \ . \tag{4.7.1}$$

Here $R_{i\alpha}$ is the component of displacement vector along one of the Cartesian directions α.

The atomic forces can be analytically evaluated by considering the change to each term in the energy in (4.5.4) when $\vec{R}^{k-1}_{i\alpha}$, from the $(k-1)$-th minimisation, is displaced by $\Delta\vec{R}^{k}_{i\alpha}$, defined as

$$\vec{R}^{k}_{i\alpha} = \vec{R}^{k-1}_{i\alpha} + \Delta\vec{R}^{k}_{i\alpha} \ . \tag{4.7.2}$$

Thus, the change in energy due to the displacement of an atom, *i.e.*, the force acting on that particular atom, is given by

$$\Delta E = \sum_{ij} b_{ij}\, \Delta(T_{ij} + V^{ps}_{ij}) + \sum_{ij} \left\{ T_{ij} + V^{ps}_{ij} \right\} \Delta b_{ij} + \Delta\widetilde{U}_{\mathrm{H}} + \Delta\widetilde{E}_{xc} + \Delta E_{i-i} \tag{4.7.3}$$

$$\Delta\widetilde{U}_{\mathrm{H}} = \sum_{kl} c_k\, G_{kl}\, \Delta c_l + \frac{1}{2}\sum_{kl} c_k\, c_l\, \Delta G_{kl} \tag{4.7.4}$$

$$\Delta\widetilde{E}_{xc} = \sum_{k\sigma} \epsilon^{\sigma}_{k}\, \Delta d^{\sigma}_{k} + \sum_{k\sigma} d^{\sigma}_{k}\, \Delta\epsilon^{\sigma}_{k} \ . \tag{4.7.5}$$

Δc_k can be evaluated from (4.3.11):

$$\sum_{l} G_{kl}\Delta c_l = \sum_{ij} \left\{ t_{ijk}\, \Delta b_{ij} + b_{ij}\, \Delta t_{ijk} \right\} - \sum_{l} c_l\, \Delta G_{kl} \ . \tag{4.7.6}$$

In the same way Δd^{σ}_{k} can be evaluated from (4.4.6):

$$\sum_{l} H_{kl}\Delta d^{\sigma}_{l} = \sum_{ij} \left\{ u_{ijk}\, \Delta b^{\sigma}_{ij} + b^{\sigma}_{ij}\, \Delta u_{ijk} \right\} - \sum_{l} d^{\sigma}_{l}\, \Delta H_{kl} \ . \tag{4.7.7}$$

$\Delta\epsilon^{\sigma}_{k}$ is found from (4.4.19) and given by

$$\begin{aligned} \Delta\epsilon^{\sigma}_{k} &= d^{\sigma}_{k} \frac{\partial\epsilon_{xc}(\langle n_\sigma\rangle_k)}{\partial\langle n_\sigma\rangle_k} I_k \exp\left[f_k(s_0)\right] \Delta\langle\widetilde{n}_\sigma\rangle_k \\ &+ \frac{s_0\,(s_0-1)}{2} d^{\sigma}_{k}\, \epsilon_{xc}(\langle n_\sigma\rangle_k)\, I_k \exp\left[f_k(s_0)\right] \left(\frac{\Delta\langle\widetilde{n}^2_\sigma\rangle_k}{\langle\widetilde{n}^2_\sigma\rangle_k} - \frac{2\,\Delta\langle\widetilde{n}_\sigma\rangle_k}{\langle\widetilde{n}_\sigma\rangle_k} \right) , \end{aligned} \tag{4.7.8}$$

with $\Delta\langle\widetilde{n}_\sigma\rangle_k$ and $\Delta\langle\widetilde{n}^2_\sigma\rangle_k$, defined from (4.4.12) and (4.4.16) as

$$\Delta\langle\widetilde{n}_\sigma\rangle_k = \frac{1}{I_k}\sum_l \left\{H_{kl}\,\Delta d^\sigma_l + d^\sigma_l\,\Delta H_{kl}\right\}, \tag{4.7.9}$$

$$\Delta\langle\widetilde{n}^2_\sigma\rangle_k = \frac{1}{I_k}\sum_{lm} \left\{2u_{klm}\,d^\sigma_l\,\Delta d^\sigma_m + d^\sigma_l\,d^\sigma_m\,\Delta u_{klm}\right\} \tag{4.7.10}$$

respectively.

Now, if we gather together the terms in Δb_{ij} and Δb^σ_{ij} we get,

$$\sum_{ij}\left\{T_{ij} + V^{ps}_{ij} + V^{\mathrm{H}}_{ij}\right\}\Delta b_{ij} + \sum_{ij\sigma}(V^{xc}_{ij})_\sigma\,\Delta b^\sigma_{ij}\ . \tag{4.7.11}$$

From the Kohn-Sham equation (4.5.8) this equals $\Delta(\sum_{ij\lambda}\varepsilon_\lambda S_{ij}c^\lambda_i c^\lambda_j)$, which can be written as:

$$\sum_\lambda \varepsilon_\lambda\,\Delta\overbrace{\left(\sum_{ij} c^\lambda_i c^\lambda_j S_{ij}\right)}^{N} - \sum_{ij\lambda}\varepsilon_\lambda\,c^\lambda_i\,c^\lambda_j\,\Delta S_{ij}\ . \tag{4.7.12}$$

As the first term vanishes, the force does not contain any derivatives in the wavefunction coefficients – as required by the Hellmann-Feynman theorem, depending only on the charge density with the atom in its unperturbed site.

Terms involving the matrix elements T_{ij} and S_{ij} depend on $R_{i\alpha}$ only through the basis functions $\phi_i(\vec{r}-\vec{R}_a)$, but the pseudopotential term has an additional dependence arising from $V^{ps}_a(\vec{r}-\vec{R}_a)$. This can be evaluated by integrating by parts:

$$\int \phi_i(\vec{r}-\vec{R}_i)\,\Delta V^{ps}_a(\vec{r}-\vec{R}_a)\,\phi_j(\vec{r}-\vec{R}_j)\,d^3r \tag{4.7.13}$$

$$= -\int\left\{\phi_j(\vec{r}-\vec{R}_j)\,\Delta\phi_i(\vec{r}-\vec{R}_i) + \phi_i(\vec{r}-\vec{R}_i)\,\Delta\phi_j(\vec{r}-\vec{R}_j)\right\}V^{ps}_a(\vec{r}-\vec{R}_a)\,d^3r\ . \tag{4.7.14}$$

Despite the complexity of the equations, the time taken to evaluate the forces is small in comparison with that taken to determine the self-consistent energy.

4.8 Structure optimisation

Having described how to calculate the self-consistent charge density, and hence the total energy and its associated tridimensional potential surface, of a cluster of atoms, the next step is the minimisation of the surface with respect to the atomic co-ordinates.

A conjugate gradient method is then used in order to generate the next search direction along which the minimisation is to be carried out. The equation used to generate the next search direction, $\vec{D}^k_{i\alpha}$, is,

$$\vec{D}^k_{i\alpha} = \vec{f}^k_{i\alpha} + \frac{\left(|\vec{f}^k_{i\alpha}|^2 - \vec{f}^k_{i\alpha}\cdot\vec{f}^{k-1}_{i\alpha}\right)}{|\vec{f}^{k-1}_{i\alpha}|^2}\,\vec{D}^{k-1}_{k\alpha}\ , \tag{4.8.1}$$

where $\vec{f}^k_{i\alpha}$ is the force vector in the k-th iteration of the minimisation. The vectors $\vec{f}^k_{i\alpha}$ and $\vec{D}^k_{i\alpha}$ have $3N$ components, where N is the number of atoms, being the force or search

direction along each of the Cartesian axes, α, for each atom, i. This equation is derived and discussed by Press *et al.* (1982).

The displacement vector $\Delta\vec{R}^k_{i\alpha}$ in the k-th iteration, as a function of the previous iteration vector can now be defined as

$$\Delta\vec{R}^k_{i\alpha} = \vec{R}^k_{i\alpha} - \vec{R}^{k-1}_{i\alpha} = k_\alpha \, \vec{D}^k_{i\alpha} \,. \tag{4.8.2}$$

The parameter k_α, quantifying the displacement along the α direction, is determined from the minimisation of the free energy of the cluster (4.6.6) by means of a quadratic or cubic interpolation.

By ensuring that $\vec{f}^k_{i\alpha}$ and $\vec{D}^k_{i\alpha}$ are orthogonal to $\vec{f}^{k-1}_{i\alpha}$ and $\vec{D}^{k-1}_{i\alpha}$ respectively, the condition that a given search direction is conjugated to all previous ones is ensured (Polak, 1971).

4.9 Determination of vibrational modes

4.9.1 Equations of motion in the harmonic approximation

Assuming that the potential energy of the cluster is a function of the instantaneous positions of all the atoms, it can be expanded as a Taylor series in powers of atomic displacement from the equilibrium configuration $\vec{R}^0$ as

$$\Phi(\vec{R}) = \Phi^{(0)}(\vec{R}^0) + \sum_{i\alpha} \Phi^{(1)}_{i\alpha} \, u_{i\alpha} + \frac{1}{2} \sum_{ij} \sum_{\alpha\beta} \Phi^{(2)}_{\substack{i\alpha \\ j\beta}} \, u_{i\alpha} \, u_{j\beta} + \cdots \,, \tag{4.9.1}$$

where $u_{i\alpha}$ represents the atomic displacement of atom i along the Cartesian direction α.

The first term in (4.9.1), the potential energy of the ions in the equilibrium configuration $\vec{R}^0 \equiv (\vec{R}^0_1, \ldots, \vec{R}^0_N)$, is unimportant for the dynamical problem and can be set to zero.

The coefficient of the second term is the derivative of the potential at equilibrium, *i.e.*, the force on that atom i, satisfies

$$\Phi^{(1)}_{i\alpha} = \left.\frac{\partial E}{\partial u_{i\alpha}}\right|_{\vec{R}^0} = 0 \,, \tag{4.9.2}$$

since the equilibrium position of an atom must be a minimum of the potential energy surface along the α direction.

The third term is the harmonic term and has its coefficients defined as

$$\Phi^{(2)}_{\substack{i\alpha \\ j\beta}} = \left.\frac{\partial^2 E}{\partial u_{i\alpha} \, \partial u_{j\beta}}\right|_{\vec{R}^0} \,, \tag{4.9.3}$$

which represents the negative of the linear force on atom i along the direction α (see §4.7).

The ionic Hamiltonian can now be written as

$$\mathcal{H} = \frac{1}{2} \sum_{i\alpha} m_i \, \dot{u}^2_{i\alpha} + \frac{1}{2} \sum_{ij} \sum_{\alpha\beta} \Phi^{(2)}_{\substack{i\alpha \\ j\beta}} \, u_{i\alpha} \, u_{j\beta} \,, \tag{4.9.4}$$

with m_i the mass of the i-th ion. The corresponding equation of motion reads

$$m_i \, \ddot{u}_{i\alpha} = -\sum_{j\beta} \Phi^{(2)}_{\substack{i\alpha \\ j\beta}} \, u_{j\beta} \,. \tag{4.9.5}$$

4.9.2 Evaluation of energy double derivatives

In order to determine $\Phi^{(2)}$, a numerical evaluation of the the double derivatives of the energy using the forces found analytically, is performed using

$$\frac{\partial^2 E}{\partial u_{i\alpha}\,\partial u_{j\beta}} \backsimeq -\frac{[f_{i\alpha}(\delta_\alpha) - f_{j\beta}(-\delta_\alpha)]}{2\delta_\alpha} \tag{4.9.6}$$

where $f_{j\beta}(\delta_\alpha)$ is the magnitude of force on ion j in the direction β when atom i is moved along the α direction by δ_α. It has been found that the value of $\delta_\alpha = 0.025$ a.u. is sufficient to ensure numerical stability.

Due to the fact that the derivative of the force is defined as a finite difference, it includes contributions from every even power of δ. Therefore, the calculated frequencies contain anharmonic terms. For this reason, the latter are sometimes called *quasi-harmonic* frequencies (Jones *et al.*, 1994).

To solve the equation of motion (4.9.5), we try a solution of the form

$$u_{i\alpha} = \frac{1}{\sqrt{m_i}}\, u'_{i\alpha}\, \mathrm{e}^{-i\omega t} \; . \tag{4.9.7}$$

By substituting this expression into the equation of motion (4.9.5), results in

$$\omega^2\, u'_{i\alpha} = \sum_{j\beta} \frac{1}{\sqrt{m_i\, m_j}}\, \Phi^{(2)}_{\substack{i\alpha\\ j\beta}}\, u'_{j\beta} \; . \tag{4.9.8}$$

Re-arranging, we obtain

$$\sum_{j\beta} \Big[D_{ij\alpha\beta} - \delta_{\alpha\beta}\, \delta_{ij}\, \omega^2 \Big] u'_{j\beta} = 0 \quad \text{with} \quad D_{ij\alpha\beta} = \frac{1}{\sqrt{m_i\, m_j}}\, \Phi^{(2)}_{\substack{i\alpha\\ j\beta}} \; , \tag{4.9.9}$$

that is,

$$\det \Big|\, D_{ij\alpha\beta} - \omega^2\, \delta_{\alpha\beta}\, \delta_{ij} \,\Big| = 0 \; . \tag{4.9.10}$$

This equation allows us to determine the vibrational frequencies of N vibrating atoms including those of the defect given by ω, as well as the corresponding normal mode displacement $u_{i\alpha}$.

Only some of the entries of the dynamical matrix with elements $D_{ij\alpha\beta}$, can be found in this way for a large cluster. In order to circuvent this difficulty—the evaluation of the double derivates is very time consuming— these are only calculated for the atoms surrounding a particular defect. The second derivatives of the remaining atoms are determined using a valence force potential.

Several types of potentials for a diamond/zincblende structure, depending on the number of degrees of freedom, can be found in the literature (Torres & Stoneham, 1985). A particularly useful one is that due to Musgrave and Pople (1962). This includes all possible bond lengths and bond angle distortions up to second order. The potential for atom i is:

$$V_i = \frac{1}{4}\sum_j k_r^{(i)}(\Delta r_{ij})^2 + \frac{1}{2}r_0^2\sum_{j>l} k_\theta^{(a)}\,(\Delta\theta_{jil})^2$$
$$+\sum_{l>j} k_{rr}^{(i)}\,\Delta r_{ij}\,\Delta r_{il} + r_0^2\sum_{m>l>j} k_{\theta\theta}^{(i)}\Delta\theta_{jil}\,\Delta\theta_{lim}$$
$$+\,r_0\sum_{l>j} k_{r\theta}^{(i)}(\Delta r_{ij}+\Delta r_{il})\,\Delta\theta_{jil}\,. \tag{4.9.11}$$

Here Δr_{ij} and $\Delta\theta_{jil}$ are the changes in the length of the i–j bond and angle between the i–j and i–l bonds, respectively. The sums are over the nearest neighbour atoms of atom i.

4.10 Practical considerations

4.10.1 Choosing a basis set

There are two different basis sets used in the method. The first is a basis used to describe the wavefunctions and the second for the description the charge density. These are invariably sets of Gaussian functions defined by an exponent α_i and sited on an atom or at the centre of a bond or some other location, R_i.

As described in §4.2, for the wavefunction expansion this Gaussian is multiplied by a polynomial in $x - R_{xi}$, $y - R_{yi}$, and $z - R_{zi}$ (vide 4.2.2). For spherically symmetric s-functions, the polynomial is trivially unity. For p-orbitals, the three possible polynomials are $x - R_{xi}, y - R_{yi}$ or $z - R_{zi}$. For d orbitals, all 6 polynomials of degree 2 are used, which generates a linear combination of five d- and one s-orbitals. f-orbitals can be included by considering third-order polynomials. The complete basis is then a linear superposition of these orbitals for different exponents α_i and centres $\vec{R}_i$.

The optimum exponents α_i, $i = 1, 2, \ldots m$, for a particular atom can be found by minimising the energy of the pseudo-atom as a function of α_i. This procedure also generates the coefficients of the wavefunction: c_i^λ, $i = 1, 2, \ldots m$ for each valence state λ. For example, it generates a set of coefficients for an s-orbital and a set for the three p-orbitals. When an application is made to a large cluster, the same fixed linear combination of the Gaussian orbitals with different exponents can then be used. This gives a basis of 4 orbitals for each Si atom for example and 10 for a transition element like Ni. This defines a *minimal basis* set. In many applications, the minimum atomic basis is used for atoms far away from the core. For other atoms, the coefficients which multiply the Gaussian orbitals are treated as variational parameters as described in section §4.2. In this case, the atoms are said to have a *big basis*.

For the treatment of the charge density, the set of modifed Gaussian functions described previously by (4.3.19) is used to expand the charge density. These functions are defined by the exponent β_k and centre $\vec{R}_k$. Again, the centres $\vec{R}_k$ can be chosen to lie at nuclei, bond-centres or other locations. The optimum basis consists of exponents and sites which *maximise* the estimated Hartree energy $\widetilde{U}_H$ as described in (§4.3).

It is always desirable to locate the Gaussian orbitals at a symmetrical site or the set of sites generated by symmetry, since otherwise the energy levels, vibrational modes *etc.*, will not possess the required degeneracy. It is expedient to define the basis in terms of N and M which means that a basis of N Gaussian s, p or d orbitals are placed at the location of each atom to describe the wavefunctions, while a basis of M Gaussian s-functions are used to describe the charge density. In addition the sites treated in terms of a minimal basis set need to be defined as well as any orbitals and fitting functions placed at bond centres. A minimal basis is often placed on the surface H atoms.

To summarise, the two of types of atomic basis currently used are as follows:

(*i*) *big basis*: The atomic wavefunction is generated as a linear combination of Cartesian Gaussian with different widths, *i.e.*, different exponents. The corresponding coefficients α_i are optimised to minimise the pseudo-atom energy;

(*ii*) *minimal basis*: In this case, the Gaussian coefficients α_i are not variational parameters. In other words, a fixed linear combination of orbitals is used. As mentioned above, the H-terminators are always treated in minimal basis.

Charge density input options

The intermediate fit to the charge density can be chosen to have different forms depending on the defect system being modelled.

The charge density can be written as (see Eq. 4.3.19)

$$\widetilde{n}(\vec{r}) = \sum_k c_k \, g_k(\vec{r}) + \sum_p c_p \, h_p(\vec{r}) \; , \tag{4.10.1}$$

with $g_k(\vec{r})$ defined by the following:

$$g_k(\vec{r}) = \left[1 - \frac{2}{3}\,\beta_k \,(\vec{r} - \vec{R})^2\right] e^{-\beta_k(\vec{r}-\vec{R}_k)^2} \; . \tag{4.10.2}$$

The possible forms of $\widetilde{n}(\vec{r})$ (`chden`) that can be used with AIMPRO are as follows:

- Type 0: Unconstrained charge density

$$h_p(\vec{r}) = Z_p \sqrt{\frac{\beta_p}{\pi}}\, e^{-\alpha_p(\vec{r}-\vec{R}_p)^2} \; ; \tag{4.10.3}$$

- Type 1: Constrained charge density:

$$h_p(\vec{r}) = \sum_{\lambda=1}^{2} c_{p\lambda}\, Z_p \sqrt[3]{\frac{\beta_{p\lambda}}{\pi}}\, e^{-\alpha_{p\lambda}(\vec{r}-\vec{R}_p)^2} \; ; \tag{4.10.4}$$

- Type 2: For using exponents of special functions for the Hartree density. This is the same as type 3 (below) but uses Gaussians with different widths (different values of α).

- Type 3: The whole of g_k is replaced with

$$g_k = e^{-\alpha_k(\vec{r}-\vec{R}_k)^2} \; . \tag{4.10.5}$$

The type-1 fit to the charge density is fastest, being used in almost all the runs performed during the course of our study. The exception are studies of defects involving oxygen, which possesing lone-pair orbitals require a larger atomic basis, as well as a different, slower, charge density fit — type 3.

4.10.2 Basis size

The basis size has a significant effect on calculated properties: with structures being least sensitive and energies and wavefunctions being most sensitive. It is not possible to converge total energies with the same degree of rigour as is occasionally obtained in plane-wave treatments. This is because simply increasing the number of exponents used to describe the basis eventually results in a numerical instability for the Choleski decomposition of the overlap matrix. However, in practice it is energy *differences* that are important as, for example, between a H atom at a bond centred and tetrahedral interstitial site. In this case the dependence of the total energy difference can be easily checked.

4.10.3 Termination effects

In dealing with defects within semiconductors, H terminated clusters have invariably been used. These saturate the dangling bonds at the surface of the cluster leading to widely separated filled and empty surface states for 'bulk' clusters, *i.e.*, clusters comprised with the same stoichiometry and atomic arrangement as the bulk semiconductor. If the surface H bond lengths are close to their equilibrium values, the band gaps are much greater than those of the bulk solids, with the exception of diamond. Values for representative Si clusters are given in Table 4.2. These were calculated for tetrahedral clusters with an $N = M = 8$ basis on the inner 5 atoms and a minimal basis on all the others. Two bond centred Gaussian basis functions with different exponents were sited on all the bond centres between host atoms.

Defect–surface interactions

In order to avoid defect-surface interactions and consequently erroneous results, it is necessary to investigate any variations of the properties of defect with cluster size. This interaction can occur in two different situations: when strong relaxation/distortion of the crystal surrounding the defect core occurs, which might involve a large volume increase (*e.g.*, Jahn-Teller systems like vacancy or transition-metal related defects), or when the defect wavefunction overlaps with the surface states. Although the former can be investigated straightforwardly, the latter might pose serious problems. This is related to the difficulties on pin-pointing states localised in reciprocal space (band-edge like) using a basis of functions localised in the real space.

The gap problem II

The LDF-LDA band gaps within a cluster approach like AIMPRO are much larger than those found using the same theory with supercells which are in turn smaller than experimental gaps. The origin of this problem is completely different from that described in §§2.8.4. The effect of the confining potential due to the terminating surface is to shift the occupied bonding *p*-orbitals of the Si atoms dowards in energy, which implies the upwards shift of the empty anti-bonding *s*-orbitals, resulting in a net increase of the bandgap.

The cluster band gap of Si varies only slightly with the basis size but becomes smaller if longer H bonds are allowed. It is not advisable to use long H bonds as this imposes a strain on inner bonds around defects and this can certainly modify their structure, seriously perturbing the local vibrational modes. The band gap also decreases slowly with cluster size.

Table 4.2: Lowest Kohn-Sham level, ε_1, highest occupied level ε_v, calculated (Δ_{KS}) and experimental band gap for various Si clusters (eV). The basis used is specified in §4.10.3.

					Bandgap width	
	Cluster size	Stoichiometry	ε_1	ε_v	Calc.	Expt.
	71	$Si_{35}H_{36}$	−16.31	−6.48	3.82	1.17
Silicon	131	$Si_{71}H_{60}$	−16.77	−6.41	3.13	1.17
	297	$Si_{181}H_{116}$	−16.91	−5.96	2.51	1.17

Despite the large band gaps, some information on the position of energy levels can be obtained. There are two common ways of 'correcting' the band gap to make allowance for the difference with experiment. The first is to simply scale defect levels by the band gap. Clearly this is simply pushing both valence and conduction band states closer together. The second is to use a *scissors* operator (Baraff & Schlüter, 1984). This is added to the Hamiltonian and displaces the unoccupied states of the 'perfect' cluster upwards by V. The scissor operator is

$$\Delta(\vec{r}, \vec{r}') = V \sum_{\lambda'} \psi_{\lambda'}(\vec{r})\, \psi_{\lambda'}(\vec{r}') \,, \tag{4.10.6}$$

where λ' represents unoccupied levels. It can also be expressed in terms of the occupied states and, for the spin-averaged case,

$$\Delta(\vec{r}, \vec{r}') = V \sum_{ij} (\delta_{i,j} - b_{ij})\, \phi_i(\vec{r} - \vec{R}_i)\, \phi_j(\vec{r}' - \vec{R}_j) \,. \tag{4.10.7}$$

V is chosen to give the correct band gap. This is then applied to a cluster containing a defect.

5

Deep Level Characterisation

5.1 Introduction

This chapter is intended to provide some insight into the properties of deep-level defects. We start with a comparison with shallow-level defects that can be accurately described by the effective-mass theory. Unfortunately, such a 'unified' theory is not available for deep-level centres. This is followed by a description of the emission and recombination processes that can occur between deep-levels and the band edges. Commonly used experimental techniques to characterise the structural, vibrational and electrical properties of deep-level defects are also described. This chapter ends with a summary of theoretical methods that have been applied to describe these defects, including a novel approach for the characterisation of their electrical properties.

Localised states, *i.e.*, states with decaying wavefunctions, can only appear if the translational symmetry of the perfect crystal is broken. For example, the crystal may contain a chemical impurity or an atom may be missing from the otherwise perfect array of atoms forming the crystal. These and many other types of imperfections are always present in any real semiconductor material.

Different charge states of the same defect are usually associated with different localised states. It is a well known fact that the presence of these levels in the band gap can, and often do, dramatically change the electrical and magnetic properties of the semiconductor material. This is simply due to the fact that these levels can exchange carriers with the conduction and valence bands, and between themselves by emission, capture or recombination of electrons or holes.

Several features associated with deep levels are very important for the optimisation design of silicon integrated devices. These are the relative position of the levels within the gap—activation energy, the capture rate of these levels, their concentrations, the identification of minority- and majority-carrier traps and their capture cross sections, *i.e.*, a measure of the ability of the deep level to trap carriers (Stoneham, 1975).

5.2 Shallow *versus* deep levels

A localised state within the bandgap could not ever exist if the entire charge of an impurity electron was localised in the atomic volume of the impurity. Obviously, in this case the binding energy could vary enormously, exceeding by far the magnitude of the band gap and a few impurity states could fall into the band gap. For example, the magnitude of gap of Si is $\sim$ 1 eV, constituting a small fraction of the binding energy of a valence electron at the free atom. It would appear that the localisation of these bound states states can only be such that a substantial portion of the charge density lies outside the range of the short-range, deep atomic potential of the impurity.

If the amount of localised charge located in the region of the short-range potential can be considered small, the full Schrödinger equation can be reduced to its hydrogenic form. This leads to the well-known effective-mass theory (Kohn, 1957; Pantelides, 1978).

Effective-mass theory was very successful in describing the electronic structure of shallow impurities. This theory made a very simple assumption about the form of the perturbation potential introduced by a shallow impurity (screened Coulombic) and a set of approximations which reduced the resulting Schrödinger equation into a hydrogenic form. This is usually equated by the following radial Schrödinger equation (in SI units).

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2m^*}{\hbar^2}\left(E + \frac{Zq^2}{\epsilon r}\right) - \frac{l(l+1)}{r^2}\right] R_l(r) = 0 \tag{5.2.1}$$

The solutions of this equation resembles the spectrum of an hydrogen atom immersed in a uniform dielectric medium (l is the angular momentum quantum number).

The large ionisation energies of deep levels implies a strong potential that localises the wavefunction of the carrier trapped to the defect. A consequence of localisation in real space is delocalisation in k-space (reciprocal space).

In the case of deep-level impurities the long-range character of the screened potential is the source of binding. The contraction of the impurity wavefunction, and consequently, the breakdown of the hydrogenic model, can only occur as a result of short-range interactions. The hydrogenic model views the localised states as 'superimposed' upon the electronic structure of the host crystal. An increase in the binding energy with respect to the relevant band edge can only mean a more localised wavefunction, and vice versa.

5.3 The effects of lattice relaxation

The introduction of a defect may involve a very abrupt change in the potential of the crystal. The valence wavefunctions of the host atoms may be modified substantially in the vicinity of a defect. If the impurity potential is weak, such modifications can be treated as a perturbation. Unfortunately, this is not always the case.

The form of the valence electron relaxation reflects the ability of the impurity potential to interact with valence electrons. The newly created electron distribution may substantially alter the lattice stability around the defect and the surrounding atoms may

be displaced. In this case, new equilibrium nuclear co-ordinates then exist which correspond to a lower total energy. The lattice relaxation alters the strength and often the symmetry of the impurity potential. In other words, the defect atom and its neighbours are displaced with respect to their perfect crystal positions.

The driving force for a distortion is often due to the Jahn-Teller effect (Jahn & Teller, 1937). This is the case of transition metal impurities and vacancy and vacancy-impurity complexes, which constitute the main topic of this thesis.

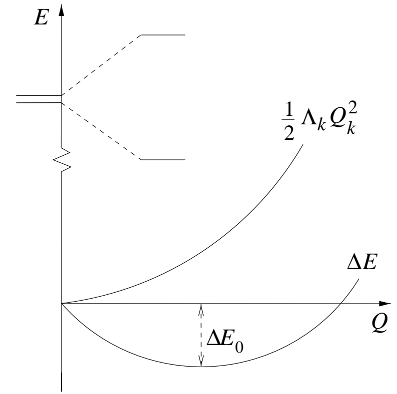


Figure 5.1: The electronic energy of a twofold degenerate localised state and the elastic energy $\frac{1}{2}\Lambda_k Q_k^2$ as functions of Q. ΔE is the total energy variation. ΔE_0 is the Jahn-Teller stabilisation energy.

Such a distortion occurs when the electronic state is orbitally degenerate, in which the nuclear configuration is unstable. Atomic displacements always exist which, by lowering the symmetry, split the the degenerate level. As the centroid remains fixed, there is a sub-level which is lowered and the system is thus stabilised by the spontaneous distortion. There are usually several equivalent new atomic configurations corresponding to distortions which are equivalent in symmetry.

We now discuss the origin of a Jahn-Teller distortion. Let us consider a localised state $\psi_\alpha(\vec{r},\vec{R})$, degenerate by symmetry for $\vec{R}=\vec{R}^0$, the equilibrium configuration, and discuss its possible splitting for $\vec{R}\neq\vec{R}_0$. For a nuclear configuration $\vec{R}$ close to $\vec{R}_0$, the ionic potential seen by an electron in the degenerate level, can be expanded as

$$\Phi(\vec{R}) = \Phi(\vec{R}_0) + \sum_k \frac{\partial\Phi(\vec{R}_0)}{\partial Q_k} Q_k + \frac{1}{2}\sum_k\sum_{k'} \frac{\partial^2\Phi(\vec{R}_0)}{\partial Q_k\,\partial Q_{k'}} Q_k\, Q_{k'} + \cdots\,. \tag{5.3.1}$$

Here Q_s represents a normal displacement, *i.e.*, a suitable linear combination of the Cartesian co-ordinates of the atomic displacement vectors from the perfect crystal positions.

The second and third terms represent a perturbation potential Φ'. The splitting of the degenerate levels can then be given to first order in Φ' by conventional perturbation theory, by the diagonalisation of the perturbation matrix, whose elements are given by

$$\Phi'_{\alpha\beta} = \sum_k Q_k <\psi_\alpha|\,\frac{\partial\Phi(\vec{R}_0)}{\partial Q_k}\,|\psi_\beta> + \frac{1}{2}\sum_k\sum_{k'} Q_k\,Q_{k'} <\psi_\alpha|\,\frac{\partial^2\Phi(\vec{R}_0)}{\partial Q_k\,\partial Q_{k'}}\,|\psi_\beta> \tag{5.3.2}$$

with ψ_α and ψ_β belonging to the sub-space of the degenerate level. The Q_k represent the normal modes of the system containing the defect, if we consider that only the motions of the impurity or defect and its immediate neighbours contribute appreciably to Φ — 'Quasi-molecular' model (Bourgoin & Lannoo, 1983).

The splitting is given by the diagonalisation of the first-order part since $\Phi'_{\alpha\beta}$ can generally be written as

$$\Phi'_{\alpha\beta} = \Big(\frac{1}{2}\sum_k \Lambda_k\, Q_k^2\Big)\, I_{\alpha\beta} + \sum_k Q_k < \psi_\alpha \,|\, \frac{\partial\Phi(\vec{R}_0)}{\partial Q_k} \,|\, \psi_\beta > , \tag{5.3.3}$$

where Λ_k are the force constants and $\mathbf{I}$ is the unit matrix.

According to the Hellmann-Feynman theorem (§1.5.1), the generalised force F_k acting on the nuclei in the sense to increase Q_k is

$$(F_k)_{\alpha\beta} = - < \psi_\alpha \,|\, \frac{\partial\Phi(\vec{R}_0)}{\partial Q_k} \,|\, \psi_\beta > \tag{5.3.4}$$

F_k is the analog of the classical $F_k = -\partial\Phi/\partial Q_k$. If F_k in (5.3.4) is non-zero for a certain nuclear configuration, the configuration is unstable and will spontaneously distort until F_k is zero.

The Jahn-Teller (JT) theorem states that if $|\psi_\alpha\rangle$ is a degenerate state when $Q_k = 0$, the matrix elements in (5.3.4) exist for some Q_k, with $k \neq 1$ (*i.e.*, for some non-totally symmetric coordinate) (Sturge, 1967). This can be proved by elementary group theory, as demonstrated by Jahn and Teller (1937).

In the simplest case, the change in total energy of a twofold degenerate level due to a JT distortion, to first order, is (Fig. 5.1)

$$\Delta E = \frac{1}{2}\,\Lambda_k\, Q_k^2 - F_k\, Q_k\ . \tag{5.3.5}$$

The minimum energy configuration is established at $Q_k^0 = F_k/\Lambda_k$ which corresponds to a net reduction of the total energy by $\Delta E_0 = F_k^2/2\Lambda_k$.

5.4 Carrier emission and recombination

Deep defect states are usually referred to as carrier traps, recombination centres or generation centres depending on the relative probabilities for the capture and emission of either electrons or holes.

In a neutral semiconductor an electron trap can be defined as a defect for which the electron capture rate, c_n, is much larger than the hole capture rate c_p, *i.e.*, $c_n \gg c_p$.

A recombination centre is characterised by large c_n and c_p. On the other hand, when the electron emission rate for electrons and holes, e_n and e_p respectively, are large defines a generation centre. The several types of centres are illustrated in Figure 5.2. The energy difference $E_c - E_T$ represents the binding energy for electrons and $E_T - E_v$ the binding energy for holes.

The electron and hole capture rates are given by (Hall, 1952; Shockley & Read Jr., 1952)

$$c_n = \sigma_n \,\langle v_n\rangle\, n \qquad \text{and} \qquad c_p = \sigma_p \,\langle v_p\rangle\, p\ , \tag{5.4.1}$$

with σ_n and σ_p being the cross sections for, respectively, capture of electrons and holes. Here n and p are the electron and hole concentrations in the conduction or valence band; $\langle v_n \rangle$ is a thermal mean electron velocity

$$\langle v_n \rangle \backsimeq \sqrt{\langle v_n^2 \rangle} = \sqrt{\frac{3k_B T}{m_n^*}} , \tag{5.4.2}$$

where m_n^* is the electron effective mass, k_B is Boltzmann's constant and T is the temperature in Kelvin. An analogous expression holds for $\langle v_p \rangle$.

Defects are also referred to as majority or minority carrier traps. Obviously an electron trap is majority carrier trap in n-type material and a minority carrier trap in p-type material.

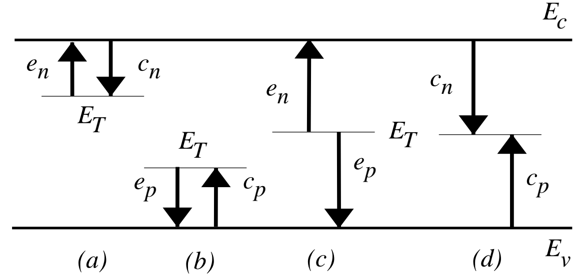


Figure 5.2: Thermal emission and capture processes from (a) electron and (b) hole traps, (c) generation and (d) recombination centres.

The thermal emission rates are proportional to a Boltzmann factor $\exp(-\Delta E/k_B T)$, where ΔE is the depth of the trap—free energy from the band edge to which the carrier is emitted. Therefore, the emission rate of an electron is

$$e_n \propto \exp\Big(-\frac{\Delta E_a}{k_B T}\Big) \tag{5.4.3}$$

with the activation energy $\Delta E_a = E_c - E_T$, where E_c and E_T are the energies of the conduction band and the trap, respectively (see Fig. 5.4).

Invoking the principle of detailed balance, the emission rate of electrons to the conduction band can be written as (Miller *et al.*, 1977)

$$e_n = \frac{\sigma_n \langle v_n \rangle N_c}{g} \exp\Big(-\frac{\Delta E}{k_B T}\Big) \quad \text{with} \quad \langle v_n \rangle N_c \propto T^2 , \tag{5.4.4}$$

where N_c is the effective density of states in the conduction band ($\propto T^{\frac{3}{2}}$) and g is the degeneracy of the defect level.[1] Again, a similar expression holds for the emission of holes to the valence band.

The electron and hole capture cross sections may be temperature dependent but generally they are not (Grimmeiss, 1977). For thermally activated cross-sections, $\sigma(T)$ seems to follow the law

$$\sigma(T) = \sigma^{\infty} \exp\Big(-\frac{E_\sigma}{k_B T}\Big) , \tag{5.4.5}$$

where E_b is the thermal activation energy of the capture cross section. This is the case of carrier capture by multi-phonon emission (Lang & Henry, 1975). This allows us to re-write

[1] Specifically, g is defined as the ratio of the ground-level degeneracies of the two ionisation states of the defect before and after the loss of the emitted carrier.

(5.4.4) as

$$e_n = \frac{\sigma_n^\infty \langle v_n \rangle N_c}{g} \exp\left(-\frac{\Delta E + E_\sigma}{k_B T}\right) . \tag{5.4.6}$$

As mentioned above, ΔE is the Gibbs free energy, *i.e.*,

$$\Delta E \equiv E_c - E_T = \Delta H - T \Delta S , \tag{5.4.7}$$

where ΔH and ΔS are the changes in enthalpy and entropy due to the emission of carriers and consequent change in the charge state of the defect. Inserting (5.4.7) into (5.4.4) gives

$$e_n = \frac{\sigma_n \langle v_n \rangle N_c X_s}{g} \exp\left(-\frac{\Delta H}{k_B T}\right) \quad \text{with} \quad X_s = \exp\left(\frac{\Delta S}{k_B}\right) . \tag{5.4.8}$$

It obvious that the slope of an Arrhenius plot yields the enthalpy of the deep level and *not* its free energy.

It important to note that only when σ^∞ and σ have been measured, ΔS may determined. Additionally, E_T the trap can only be determined if experimental data of emission and capture rates are available and that values for ΔH linearly extrapolated at $T = 0$ K are obtained for the case of temperature dependent ionisation enthalpies.

It is now convenient to write a more general expression for the thermal emission rate of electrons to the conduction from a trap characterised by a thermally activated cross-section as

$$e_n = \frac{A_n \sigma_n^\infty T^2 X_s}{g} \exp\left(-\frac{\Delta H - E_\sigma}{k_B T}\right) , \tag{5.4.9}$$

with X_s defined as in (5.4.8). For Si, $A_n = 3.42 \times 10^{21}$ s^{-1} K^{-2} cm^{-2} (Schmidt, 1998).

5.5 Detection and characterisation of deep levels

Capacitance methods like deep-level transient spectroscopy (DLTS) probe the variation of space-charge of a Schottky barrier or a *p*-*n* junction diode induced by sudden application of an external perturbation.

The space-charge layer for a one-sided abrupt p^+-n junction with uniform doping concentration and zero applied bias voltage is illustrated in Figure 5.3 (top).

The width of the space-charge layer, varies with the voltage applied to the diode, and thus, because free carriers are always swept out of this region by the electric field, it can be thought as a variable-width insulator.

The capacitance of a reversed-biased one-sided p^+-n junction, is given by

$$C = \epsilon_s \frac{A}{W} , \tag{5.5.1}$$

where A is the area of the junction, ϵ_s is the dielectric constant of the depleted semiconductor. W represents the width of the depleted zone as

$$W(V) = \sqrt{\frac{2 \epsilon_s}{q N_D} (V_b + V)} , \tag{5.5.2}$$

where V_b is the built-in voltage of the junction, or diffusion voltage, V is the externally applied voltage, q is the charge of the electron and N_D is the concentration of ionised dopant impurities in the lightly doped side of the junction.

Inserting (5.5.2) into (5.5.1), we can re-write the expression for the capacitance per unit area as

$$C = \sqrt{\frac{q\,\epsilon_s\,N_D}{2\,(V_b + V)}} \tag{5.5.3}$$

This equation forms the basis for capacitance-voltage (C-V) measurements. For uniformly doped material a plot of $1/C^2$ vs. V gives the doping density N_D from the gradient and V_b from the intercept of the slope.

Figure 5.3 shows the band diagram for a p^+-n junction having a uniformly doped n region at two different bias conditions (V_b and V). Two different regions of the space-charge region must be considered. The central region closest to the junction, where both the shallow donors and the deep levels lie above the Fermi, is depleted of all free carriers. Beyond this is a transition region (sometimes refereed to as the edge region), defined by the point where the Fermi level crosses the deep level, in which the deep levels are occupied. The width of this region is

$$\lambda = \sqrt{\frac{2\epsilon_s}{q^2\,N_D}\,(E_F - E_T)} \tag{5.5.4}$$

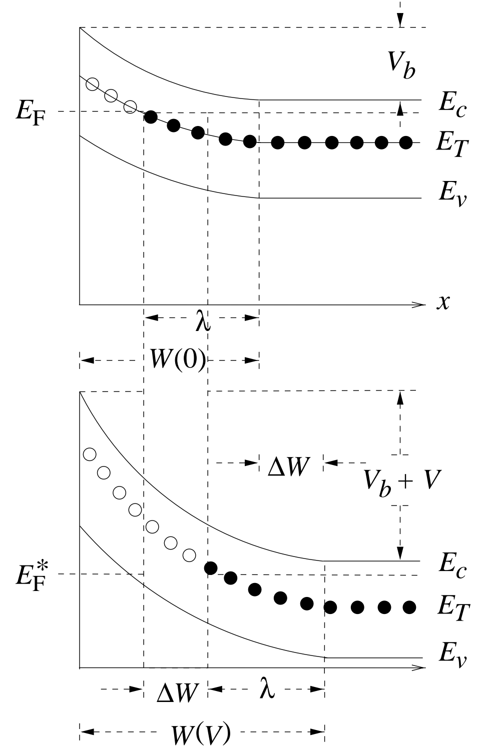


Figure 5.3: Distribution of filled and empty traps at steady state in a p^+-n step junction under zero bias (top) and reverse bias $V > V_b$ (From Lang (1979)).

Trap levels in the depletion region of diode can be detected by their effect on the depletion capacitance (Kimerling, 1974). It is evident from equations (5.5.1) and (5.5.2) that a change in the charge density in the depletion region will result in a corresponding change in the depletion layer width. Consequently, the junction capacitance is a direct measure of the total semiconductor charge.

If the concentration of carriers trapped at deep levels is changed by the thermal or optical emission of carriers to the appropriate band, this change can be monitored by measuring the variation in the junction capacitance at constant applied voltage.

The variation of the depletion region width of the diode with applied voltage makes possible other unique features of space-charge spectroscopy, namely the ability to separate carrier capture from carrier emission and to allow the capture of only majority carriers or both majority and minority carriers. The former is shown schematically in Figure 5.4 for a p^+-n junction having a deep level in the upper half of the band gap.

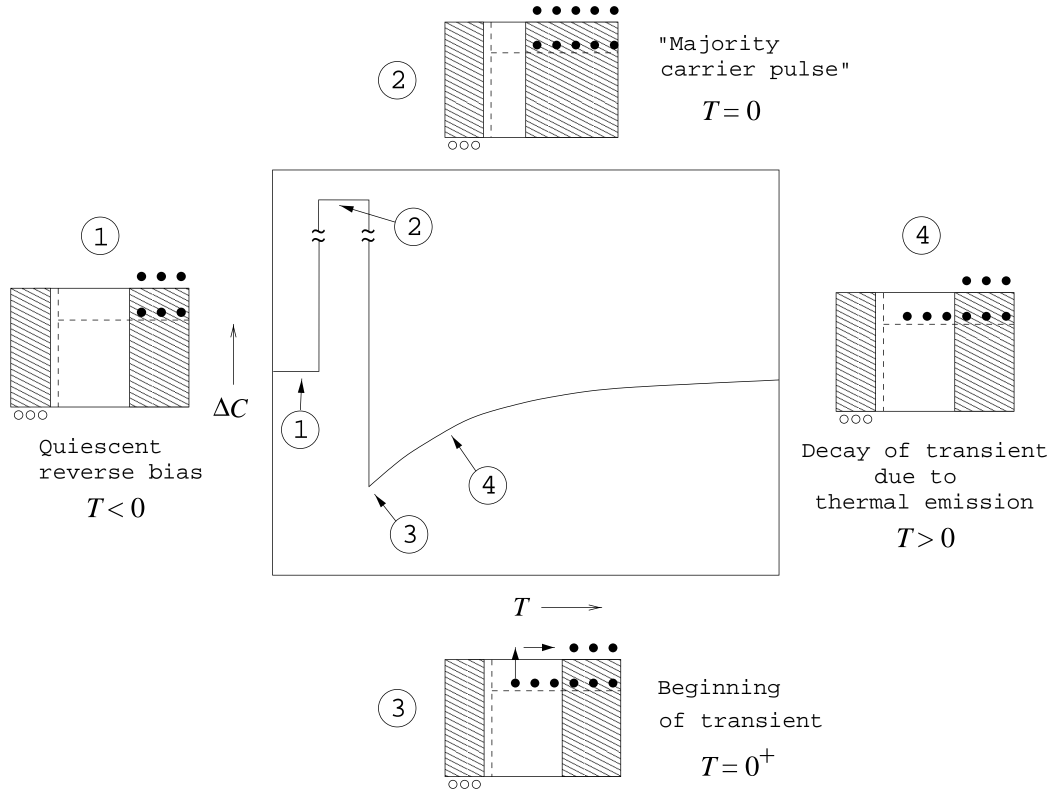


Figure 5.4: Isothermal capacitance regime evolution after the application of a reverse bias when the p^+-n junction is initially reversed biased. The insets show the occupation, space-charge layer width (unshaded) and free carrier concentrations during steps 1–4. Band bending omitted for clarity. Adapted from Lang (1979).

Under a steady-state reverse bias voltage (Fig. 5.4–1), traps in depletion region are empty. Reducing the applied voltage (Fig. 5.4–2), reduces the width of the depletion region and allows electrons to be trapped at the deep levels. Electron capture into an initially empty trap is given

$$N_T - n_T(t) = [N_T - n_T(0)]\ \exp\left(-c_n\, t\right) \tag{5.5.5}$$

where N_T is total trap density, n_T is the number of trapped electrons and c_n is the capture rate. In a sufficiently long period, all the centres captured an electron ($n_T(\infty) = N_T$). If the junction is then reversed-biased, the free carrier concentration in the transition region is nullified ($e_n \gg c_n \approx 0$). Traps release their captured electrons (Fig. 5.4–3). The transient response is then

$$n_T(t) = N_T\ \exp\left(-e_n\, t\right)\ , \tag{5.5.6}$$

where e_n is the thermal emission rate and N_T is the total trap concentration. Consequently the concentration of ionised centres, N_i, at time t, is

$$N_i(t) = N_D + [N_T - n_T(t)] = (N_D + N_T) - N_T\ \exp\left(-e_n\, t\right)\ , \tag{5.5.7}$$

showing that changes in the occupancy of these centres are directly proportional to the change in capacitance. Assuming that $N_D \gg N_T$, the capacitance of the junction per unit

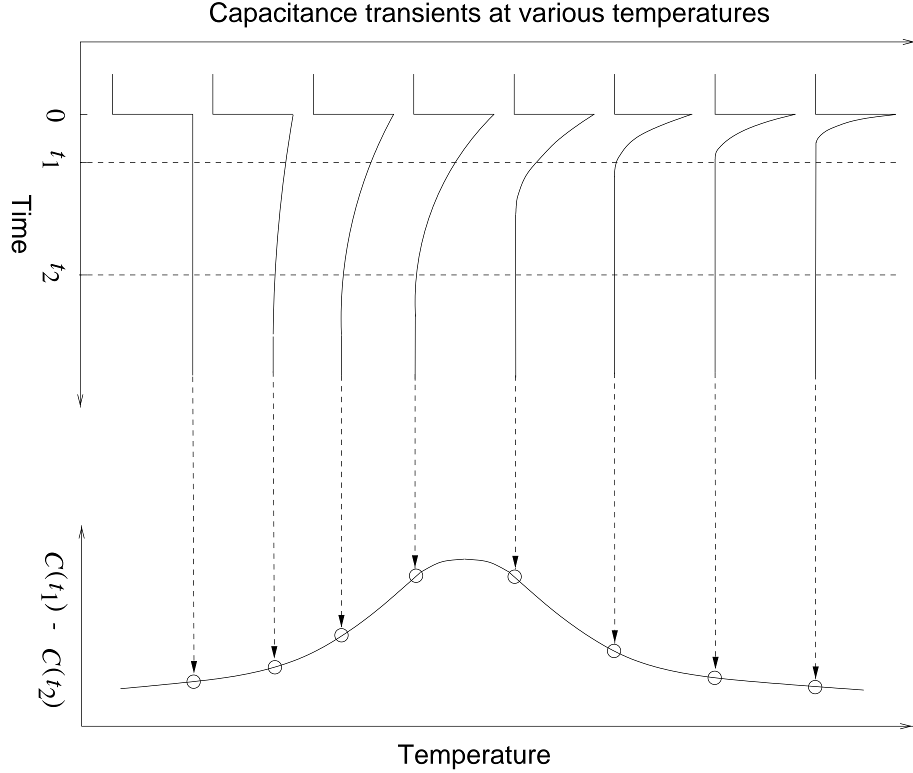


Figure 5.5: Implementation of a DLTS rate window by means of a boxcar integrator with gates set at times t_1 and t_2 (Lang, 1979).

area can be written as

$$C(t) \simeq C_0 \left[1 - \frac{N_T}{2\,N_D} \exp\left(-e_n\,t\right)\right] , \tag{5.5.8}$$

where C_0 is the quiescent capacitance under the applied voltage V given by (5.5.3). In other words, the transient capacitance $\Delta C(t) = C(t) - C_0$ decays by the exponential law

$$\Delta C \simeq -C_0\,\frac{N_T}{2\,N_D} \exp\left(-e_n\,t\right) , \tag{5.5.9}$$

with the electron emission rate given by (5.4.9).

The activation energy for thermal emission is then determined from measurements of the thermal emission rate at different sample temperatures. The trap concentration can be determined from the initial amplitude of the transient. Note that the sign of the resulting capacitance transient is negative for the emission of majority carriers (Fig. 5.4). Obviously, the capacitance transient is positive for minority carriers, simply due to the fact that their charge has the opposite sign (Miller *et al.*, 1977).

5.5.1 Deep level transient spectroscopy

If the variation of C with temperature from a transient capacitance experiment is processed in a way that a selected decay rate produces a maximum output, then a signal whose decay time changes monotonically with time reaches a peak when the rate passed through the rate window of *box-car* averager of the frequency of a *lock-in* amplifier.

When observing C-t transients through such a rate window while varying the decay time constant, τ ($\tau = 1/e_n$) by varying the sample temperature, a peak appears in the output versus temperature plot ($\Delta C/C = f(t)$). This is called a DLTS spectrum (Lang, 1974; Miller *et al.*, 1977; Lang, 1979).

Box-car DLTS

In the box-car system introduced by Lang (1974), depicted in Figure 5.5, the DLTS-signal ΔC is defined as the difference in capacitance at two sampling times t_1 and t_2

$$\Delta C = C(t_1) - C(t_2) \; . \tag{5.5.10}$$

The emission rate corresponding to the maximum of the DLTS signal is given by setting the derivative of (5.5.10) with respect to e_n equal to zero; thus

$$(e_n)_{\max} = \frac{1}{\Delta t} \ln\left(\frac{t_1}{t_2}\right) , \tag{5.5.11}$$

with $\Delta t = t_1 - t_2$. This equation is independent of the magnitude of the capacitance and a certain emission rate can be fixed by the choice of the rate window. Changing the values of t_1 and t_2 changes the rate window, and the temperature of the maximum of DLTS peak shifts.

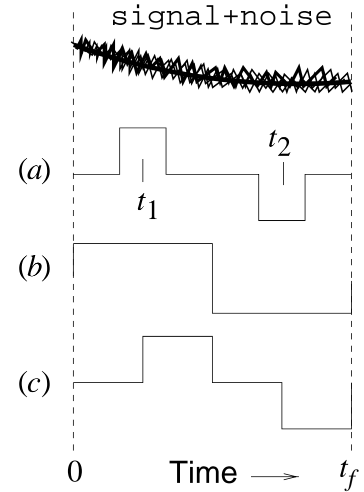


Figure 5.6: Diagrams of some rate-window operations, or weighting functions $w(t)$: (a) box-car, (b) lock-in amplifier and (c) two-phase lock-in amplifier.

Lang (1974) applied a double-gated box-car averaging method to detect the DLTS signal (Fig. 5.6). The weighting function was a follows:

$$w(t) = \delta(t - t_2) - \delta(t - t_1) \tag{5.5.12}$$

which provides a response

$$f(t) = \int_{t=0}^{t=t_f} C(t) \overbrace{\left[\delta(t - t_2) - \delta(t - t_1)\right]}^{w(t)} dt = C(t_2) - C(t_1) \; , \tag{5.5.13}$$

where t_f is the period of the filling pulse.

Lock-in DLTS

With a better signal/noise ratio, lock-in DLTS has replaced the original box-car system proposed by Lang (Miller *et al.*, 1977). A DLTS peak is observed when the frequency of the lock-in amplifier bears the proper relationship to the emission time constant. This frequency defines the period of a square-weighting function, $w(t)$, such as

$$\Delta C(t_i) = \frac{1}{t_i} \int_{t_D}^{t_D + t_i} \Delta C(t)\, w(t)\, dt \; , \tag{5.5.14}$$

where t_i represents the lock-in measuring time and t_D is a delay time.

Determination of activation energies and cross sections

In order to extract the activation energy and capture cross-section of a given trap, the rate of the window and thus the emission rate is varied several times, usually during the same temperature DLTS scan. The peak temperature and emission rate $(e_n, T_{\max})$ is extracted for each rate window.

An Arrhenius plot $(\ln{(e_n/T^2)}$ vs. $1/T)$, of the values of e_n and $T_{\max}$ taken from several rate windows, allows the determination of the activation energy and *apparent* capture-cross section: the activation energy corresponds to the slope and the cross-section to the y-axis intercept. Note that, as mencioned above, the intercept provides $\sigma_n \exp{(\Delta S/k_B)}$, providing the the cross-section implicitly, and not explicitly.

Depth profiling

The variation of the depletion region width with applied bias can be used to to measure the deep level concentration profile, if the shallow concentration level profile is known (Lang, 1979),

$$\delta\Big(\frac{\Delta C}{C_0}\Big) = \frac{\epsilon_s}{q\,W^2\,N_D}\,\frac{N_T(x)}{N_D(x)}\,\delta V \tag{5.5.15}$$

where x is the width of the space charge layer during the filling pulse and C, W and N_D are the values corresponding to the steady state bias between the pulses; $N_T(x)$ is usually measured using C-V profiling methods.

Influence of the electric field on emission rates

DLTS detects thermal emission of carriers from traps in a reverse biased junction, in which the built-in electric field can reach 10^8 V cm^{-1} for heavily doped semiconductors. In this case, the simple expressions derived previously for the emission rate no longer hold. Obviously, the dependence of the emission rates on the electric field has to taken into account. In this section, we describe one of the mechanisms of emission enhancement in an electric field: the Poole-Frenkel effect (Frenkel, 1938). For higher values of the electric field, phonon-assisted tunnelling or simple tunnelling are alternative processes.

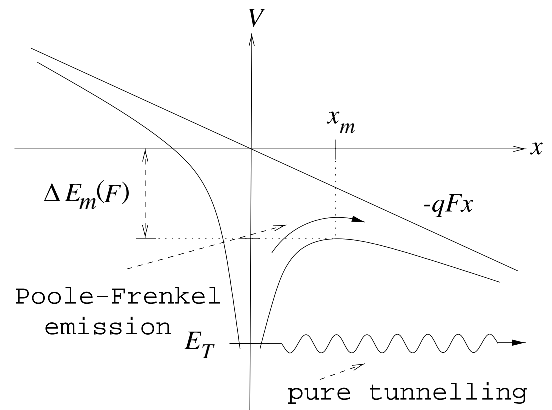


Figure 5.7: Schematic representation of a defect potential in an electric field.

In a region of an electric field $\vec{F}$, an electron bound to a defect experiences the defect potential plus the perturbative potential $-qFx$, where x is the electron coordinate along the field. The result is a field-dependent barrier lowering such that (Fig. 5.7)

$$E_T(F) = E_T(0) - \Delta E_m(F)\ , \tag{5.5.16}$$

where $E_T(F)$ and $E_T(0)$ are trap energies with and without the field respectively, and the energy ΔE_m enters in the Boltzmann factor of the thermal electron emission rate as

$$e_n(F) \propto \exp\Big(- \frac{E_T - \Delta E_m(F)}{k_B T}\Big) , \tag{5.5.17}$$

The main characteristic of the potential depicted in Figure 5.7 is that there is a finite probability for the defect electron to be thermally excited to states higher than E_m. Consequently, this results in the enhancement of the emission of carriers from the defect. This is obvious from (5.5.17).

For a Coulombic defect potential $(-q^2/\epsilon_s r)$, the quantity ΔE_m, representing the magnitude of the lowering of the defect potential, can be quantified as follows

$$\Delta E_m(F) = 2q\sqrt{\frac{q\,F}{\epsilon_s}} . \tag{5.5.18}$$

Therefore, the linearity of ΔE_m with the square root of the electric field provides the charge of the carrier leaving a centre of opposite sign. In other words, the study of ΔE versus F is a possible way of determining the acceptor or donor character of a defect level (Kimerling & Benton, 1981).

Minority carrier transient spectroscopy

Minority carrier transient spectroscopy (MCTS) is an optical technique (Brunwin *et al.*, 1979). It uses flooding illumination to generate excess carriers.

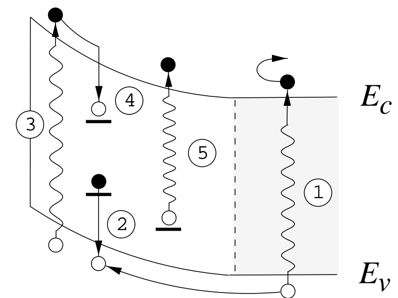


Figure 5.8: p^+-n junction diagram showing the dominant photon-semiconductor interactions during the fill pulse of a MCTS experiment (Davidson & Evans, 1996a).

A laser emitting an energy which is just greater than the bandgap energy illuminates the sample held under reverse bias. Carriers are generated in the bulk region of the semiconductor (Fig. 5.8-1: e−h pair generation) and those minority carriers within a diffusion length of the depletion region drift rapidly across the region under the influence of the field, whereas the majority carriers are excluded by the space charge.

The minority carriers contribute to the photocurrent but some are captured by minority carrier traps (Fig. 5.8-2). When the light pulse is terminated the carriers are thermally emitted at a rate dependent on the properties of the trap and give rise to a capacitance transient, similar to that of DLTS for majority carrier traps.

Other photon-electron interactions may occur during the light pulse. These can be photon absorption within the depletion region (Fig. 5.8-3), or the capture of majority carriers by deep traps (Fig. 5.8-4). Emision from these traps counteracts the capacitance produced by emission of minority carriers. Also possible is the direct photo-excitation of a carrier from a deep level to a band as in standard optical DLTS (Fig. 5.8-5)

Due to the fact that the trapped carriers have the opposite sign to that occuring in a DLTS experiment on the same sample, the peaks in an MCTS spectrum will be inverted relative to the majority carrier emission peaks of DLTS (Evans *et al.*, 1995).

Performing both standard dark DLTS and MCTS on the same sample provides a characterisation of the majority and minority traps present (Davidson & Evans, 1996b).

5.5.2 Laplace deep level transient spectroscopy

The peaks of a conventional DLTS spectrum are rather broad and the energy resolution is moderate, in the range of 10 meV. However, this disadvantage has been effectively circumvented using the Laplace transform method for the emission rate analysis in DLTS measurements (Dobaczewski *et al.*, 1994). An improvement by a factor of an order of magnitude in DLTS resolution can be achieved. Consequently, this method provides a better capability to establish a fine structure in the emission process producing the spectrum.

Laplace DLTS (LDTS) produces a plot of the spectral density as a function of emission rate rather than capacitance change versus temperature as in a conventional DLTS. The spectral density function $F(s)$, is given by the inverse Laplace transform of the recorded transient defined as

$$f(t) = \int_0^\infty F(s)\,\mathrm{e}^{-s\,t}\,ds\ .$$

The result is delta-like peak(s) in the spectra more or less broad depending on the exponential form of the transient (Fig. 5.9-b). The area under each peak is proportional to the charge release by the defect, which can be related to the concentration of the deep state. This contrasts with conventional DLTS in which the concentration of the deep level is given by the height of the peak and not by the integrated area under the peak as in LDTS.

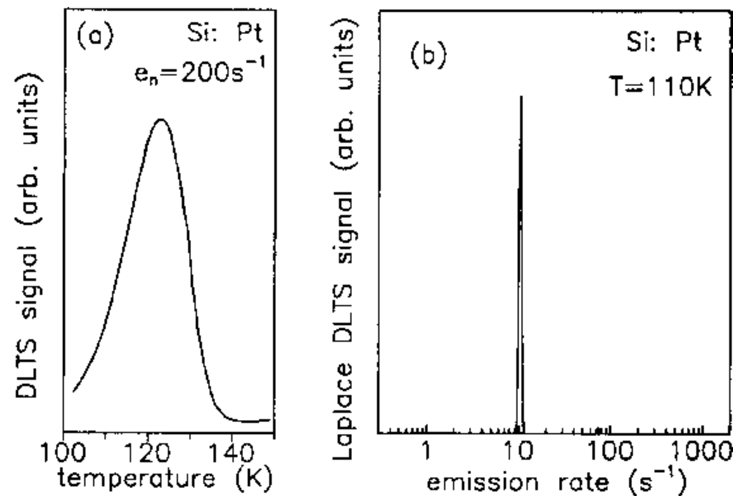


Figure 5.9: Comparison between a standard DLTS (a) and a Laplace spectrum (b) of Pt-doped Si. The spectra shown originate from the single acceptor level of substitutional Pt^- (Dobaczewski *et al.*, 1994).

To obtain the level activation energy, or the level enthalpy for temperature independent cross-sections, the LDLTS experiment is repeated at various temperatures. The slope of the corresponding Arrhenius plot gives the required energy.

5.5.3 Electron paramagnetic resonance

So far, we have described experimental techniques that can accurately described the electrical properties of a given defect. Unfortunately, neither of the previous techniques can provide any information about the microscopic structure, or chemical composition of the defect in question. This difficulty can be avoided by performing electron paramagnetic resonance (EPR) measurements, one of the most powerful techniques to characterise the microscopic structure of a paramgnetic centre.

The phenomenon of electron paramagnetic resonance is based on the ability of a unpaired electron orientated in a magnetic field to absorb electromagnetic radiation.

Principle

Consider a free paramagnetic ion with total angular momentum $\vec{J} = \vec{L} + \vec{S}$, with $\vec{L}$ and $\vec{S}$ defining the orbital and electronic spin momenta respectively. The magnetic dipolar moment due to $\vec{J}$ is

$$\vec{\mu} = -g\,\beta\,\vec{J}\ , \tag{5.5.19}$$

where $\beta \equiv eh/2mc$ is the Bohr magneton and g is a dimensionless factor given by Landé's formula

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2(J+1)}\ . \tag{5.5.20}$$

When this ion is subject to an external magnetic field, $\vec{B}$, the so-called Zeeman interactions are described by a Hamiltonian of the form

$$\mathcal{H} = -\vec{\mu}\cdot\vec{B} = g\,\beta\,\vec{J}\cdot\vec{B}\ . \tag{5.5.21}$$

For many defects in solids, the orbital momentum is strongly quenched.[2] Consequently, the contribution to $\vec{J}$, and $\vec{\mu}$ comes primarily from the spin momentum $\vec{S}$ of the electron bound to the defect. In this case, (5.5.19) is replaced by

$$\vec{\mu} = \beta\,\vec{S}\cdot\overleftrightarrow{g}\ , \tag{5.5.22}$$

with the tensor $\overleftrightarrow{g}$ defined as $\overleftrightarrow{g} = g_0 + \Delta\overleftrightarrow{g}$. $\Delta\overleftrightarrow{g}$ represents the departure from the isotropic spin-free value, $g_0 = 2.0023$, which depends on the orientation of $\vec{B}$ with the crystallographic axis. The Hamiltonian (5.5.21) can now be written as

$$\mathcal{H} = \beta\,\vec{S}\cdot\overleftrightarrow{g}\cdot\vec{B} \tag{5.5.23}$$

[2] The exceptions are rare-earth element impurities and, in special circumstances, $3d$ transition-metal ions.

which leads to the splitting of spin-degenerate energy levels into $(2S+1)$ sub-levels, equally separated by

$$E(m_S) = g\,\beta\,m_S\,. \tag{5.5.24}$$

m_S is the azimuthal quantum number, running from $-S$ to $+S$.

This departure from the free g-value originates from small orbital orbital momentum induced by the spin into surroundings (via spin-orbit interaction $\lambda\vec{S}\cdot\vec{L}$) at the atomic cores, including shielding and covalency effects.

Magnetic dipole transitions between to adjacent levels can take place ($\Delta m_S = \pm 1$) when the system absorbs energy quanta at the resonant frequency $h\,\nu = g\,\beta\,B$. Usually, this microwave frequency, ν_0, is tuned with that of the resonant cavity containing the sample. The resonant condition is searched by sweeping the applied external magnetic field. The resonant field,

$$B_0 = \frac{h\,\nu_0}{g\,\beta}\,, \tag{5.5.25}$$

can be detected and recorded by the EPR spectrometer, therefore serves as a direct measure of g, whose angular dependence is given by

$$g^2 = \sum_{i=1}^{3} g_i^2\,n_i^2\,. \tag{5.5.26}$$

Here g_i are the principal values of $\overleftrightarrow{g}$ and the n_i's the direction cosines of $\vec{B}$ with respect to the principal axes of $\overleftrightarrow{g}$.

If the spin-orbit interaction is considered a perturbation, the components of $\overleftrightarrow{g}$, g_{ij} have the form (Abragam & Pryce, 1951)

$$g_{ij} = g_0\,\delta_{ij} - 2\lambda \sum_{n\neq 0} \frac{<\Psi_0\,|\,L_i\,|\,\Psi_n><\Psi_n\,|\,L_j\,|\,\Psi_0>}{E_n - E_0} \tag{5.5.27}$$

where indices i, j denote the principal axes and λ is the spin-orbit coupling constant. ψ_0 and E_0 are the ground-state wavefunction and energy, respectively. The summation is over all excited states. Using (5.5.27), it is possible to correlate the g-shifts to the electronic structure of the defect and its excited states.

More details on the physics of magnetic resonance and its application in the study of paramagnetic systems can be found in many texts and books like Abragam and Bleaney (1970), Pake and Estle (1973), Weil *et al.* (1994) or Watkins (1998).

Hyperfine interactions

In the case of a nucleus, or nuclei, having nuclear spin $I \neq 0$, interactions with nuclear magnetic moment

$$\vec{\mu}_n = g_N\,\beta_N\,\vec{I}\,, \tag{5.5.28}$$

where β_N the nuclear magneton, may also be important due to the fact that the nuclear spins also quantise in the magnetic field.

The interaction of the weak magnetic field, $\vec{B}_{\rm loc}$ induced by the presence of nuclear magnetic dipoles, with the electronic spin $\vec{S}$ is called hyperfine interaction.

The Hamiltonian now reads,

$$\mathcal{H} = \beta\, \vec{S} \cdot \overleftrightarrow{g} \cdot \vec{B} + \vec{S} \cdot \underbrace{\Big(\sum_j \overleftrightarrow{A}_j \cdot \vec{I}_j \Big)}_{\beta \overleftrightarrow{g} \cdot \vec{B}_{\rm loc}} \tag{5.5.29}$$

in which the sum is over all nearby nuclei j. The tensor $\overleftrightarrow{A}_j$ contains information about the amplitude, distance and angular-dependent parts of the nuclear magnetic dipoles responsible for $\vec{B}_{\rm loc}$.

The corresponding eigenvalues are given by

$$E(m_S; m_I) \cong \Big(g\, \beta\, B + \sum_j A_j\, m_I^j \Big)\, m_S \;, \tag{5.5.30}$$

where m_I is the nuclear azimuthal quantum number. The resonant field (5.5.25) can now be re-written to include hyperfine interactions as

$$B_0 = \frac{1}{g\, \beta} \Big\{ h\, \nu_0 - \sum_j A_j\, m_I^j \Big\} \;, \tag{5.5.31}$$

with $\Delta m_I = 0$ as an additional selection rule. In analogy to the expression (5.5.26) for g, we write

$$A_j^2 = \sum_\alpha A_{j\alpha}^2\, n_{j\alpha}^2 \tag{5.5.32}$$

where $A_{j\alpha}$ and $n_{j\alpha}$ are, respectively, the principal values of $\overleftrightarrow{A}_j$ and the direction cosines of its corresponding principal axis with respect to the quantisation direction of $\vec{S}$, *i.e.*, $\overleftrightarrow{g} \cdot \vec{B}$. The effect of the hyperfine interactions is to split the resonant into $(2I_j + 1)$ lines separated by $A_j/g\beta$. The study of angular dependence of these splittings leads directly to the evaluation of the principal values and axes of $\overleftrightarrow{A}_j$ trough (5.5.32).

The interaction with further neighbouring atoms provides further information about the structure of the defect. This results in a set of weak satellite lines around the main hyperfine line, due to the less abundant isotopes of Si, like the 4.7 % abundant ^{29}Si with $I = \frac{1}{2}$. Again, the analysis of the angular dependence of these satellite lines provided information about the local symmetry of the defect, from the referential of the nearby nucleus. Their intensity determines the isotopic abundance of the j-th nucleus involved in the defect.

The interaction between the magnetic moment of the electron $\vec{\mu}_e$ and with the nuclear moment $\vec{\mu}_n$ is composed by two terms:

(i) the anisotropic dipolar interaction term defined by

$$H_d = \frac{\vec{\mu}_e \cdot \vec{\mu}_n}{r^3} - 3\, \frac{(\vec{\mu}_e \cdot \vec{r})(\vec{r} \cdot \vec{\mu}_n)}{r^5} \;; \tag{5.5.33}$$

(*ii*) and the 'Fermi contact term':

$$H_c = -\frac{8\pi}{3}\,(\vec{\mu}_e \cdot \vec{\mu}_n)\,\delta(\vec{r}) \tag{5.5.34}$$

which is isotropic. This term accounts for the fact that when $r \to 0$, the point-dipole approximation used to deduce (5.5.33) is no longer valid.

Substituting $\vec{\mu}_e = -g\,\beta\,\vec{S}$ and $\vec{\mu}_n = +g_N\,\beta_N\,\vec{I}$, the hyperfine interaction is

$$\mathcal{H} = \vec{S} \cdot \overleftrightarrow{A} \cdot \vec{I}\ , \tag{5.5.35}$$

with the tensor components of $\overleftrightarrow{A}$ given by

$$A_{ij} = g\,g_N\,\beta\,\beta_N \left\{ \left\langle \frac{3\,x_i\,x_j}{r^5} - \delta_{ij}\,\frac{1}{r^3} \right\rangle + \frac{8\pi}{3}\,|\Psi(0)|^2\,\delta_{ij} \right\} \tag{5.5.36}$$

$|\Psi(0)|^2$ represents the amplitude to the wavefunction at the nucleus, coming directly from the Dirac delta function $\delta(\vec{r})$. The angular brackets indicate a expectation value or matrix element of the enclosed function for the electronic ground-state.

Information about the electronic distribution and the character of the wavefunction in the defect is usually obtained by applying the conventional linear combination of atomic orbitals (LCAO) approximation. In this approximation, the wavefunction is written as superposition of the atomic *s*-, *p*- and *d*-orbital on a particular site j in the following way:

$$\Psi = \sum_j \eta_j\,\overbrace{(\alpha_j\,\psi_s + \beta_j\,\psi_p + \gamma_j\,\psi_d)}^{\psi_j}\ , \tag{5.5.37}$$

with $\alpha_j^2 + \beta_j^2 + \gamma_j^2 = 1$. η_j^2 represents the the fraction of the total wavefunction that is localised in the atomic orbital ψ_j centred on each atomic site j.

Once we know the electronic wavefunction, (5.5.36) allows the evaluation of the hyperfine interaction tensor. It is customary to decompose $\overleftrightarrow{A}_j$ as

$$\overleftrightarrow{A}_j = a_j\,\overleftrightarrow{1} + \overleftrightarrow{B}_j\ . \tag{5.5.38}$$

so that $a_j = \frac{1}{3}\,\mathrm{Tr}\,\overleftrightarrow{A}_j$ and $\overleftrightarrow{B}_j$ is a traceless tensor, *i.e.*, $\mathrm{Tr}\,\overleftrightarrow{B}_j = 0$.

For an atomic *s*-wavefunction, $\overleftrightarrow{A}_j$ is isotropic and the interaction with the central nucleus comes only from the $|\psi_s(0)|^2$—the remaining terms average to zero over the spherically symmetric *s*-state—and given, for the j-th nucleus, by

$$A_{11} = A_{22} = A_{33} \equiv a_j = \frac{8\pi}{3}\,g\,\beta\,\frac{\mu_n}{I_j}\,\eta_j^2\,\alpha_j^2\,|\psi_s(0)|^2 \tag{5.5.39}$$

where μ_n and I_j are the magnetic moment and the spin of the j-th nucleus;

Since only the atomic orbitals centred on the relevant sites are considered, the wavefunction expansion (5.5.37) gives rise to axial hyperfine interactions, *i.e.*, $\overleftrightarrow{B}$ is axial and its principal values are

$$\overleftrightarrow{B}_j = \begin{pmatrix} 2b_j & 0 & 0 \\ 0 & -b_j & 0 \\ 0 & 0 & -b_j \end{pmatrix} \quad \text{with} \quad b_j = \frac{2}{5}\,g\,\beta\,\frac{\mu_n}{I_j}\,\eta_j^2\,\beta_j^2\,\langle r^{-3}\rangle_p \tag{5.5.40}$$

The atomic wavefunction parameters $|\psi_s(0)|^2$ and $\langle r^{-3}\rangle_p$ for most atoms are available from Hartree-Fock calculations (Morton & Preston, 1978).

The deviation from axiality by $\overleftrightarrow{B}$ can be expressed with an extra parameter c. In that case, the principal values of $\overleftrightarrow{B}$ become $(2b, -b+c, -b-c)$. For large values of c in comparison with b the LCAO approximation is no longer valid.

ENDOR

The Zeeman effect also occurs when the applied external magnetic field interacts with the nuclear magnetic dipoles $\vec{\mu}_n$. This interaction is usually described as

$$-\beta_N \vec{I} \cdot \overleftrightarrow{g}_N \cdot \vec{B} \tag{5.5.41}$$

The effect of this interaction is usually not perceptible in EPR experiments. Although, it can be resolved in electron-nuclear double-resonance (ENDOR).

The Hamiltonian now includes two extra terms:

$$-\sum_j g_N \beta_N \vec{I}_j \cdot \vec{B} + \sum_j \vec{I}_j \cdot \overleftrightarrow{Q}_j \cdot \vec{I}_j \tag{5.5.42}$$

The first term represents the direct interaction of the nucleus with $\vec{B}$ (5.5.41) and the second the interaction between an electric quadrapole moment with the electronic charge distribution around the j-th atom for $I_j > \frac{1}{2}$. The analysis of the latter term, provides useful information about the type of bonding, *i.e.*, percentage of s- and p-character, around the j-atom. Here g_N is considered to be scalar due to the limited spatial extent of the nuclear wavefunction. This is not always true as the coupling between $\vec{I}$ and $\vec{B}$ may occur indirectly – pseudo-Zeeman effect.

The nuclear Zeeman interaction provides an unambiguous chemical identification of the j-th atom involved in the centre.

To first order in $A/g\beta B$, the effective Hamiltonian can be written as

$$\mathcal{H}_{N_j} \cong -g_{N_j} \beta_N \vec{I}_j \cdot \vec{B}_j^{\text{eff}} + \vec{I}_j \cdot \overleftrightarrow{Q} \cdot \vec{I}_j \;, \tag{5.5.43}$$

with

$$\vec{B}_j^{\text{eff}} = \left[1 - \frac{m_S}{g\, g_N\, \beta\, B} \overleftrightarrow{A}_j \cdot \overleftrightarrow{g}\right] \cdot \vec{B} \;.$$

Here $\vec{B}_j^{\text{eff}}$ is a net magnetic field seen by the j-th nucleus.

When $Q_j \ll g_{N_j}\beta_N B_j^{\text{eff}}$, (5.5.43) leads, for the ENDOR transitions $\Delta m_I = \pm 1$ and $\Delta m_S = 0$, to (Watkins, 1998)

$$h\,\nu(m_I \to m_I - 1) \cong g_{N_j}\, \beta_N\, B_j^{\text{eff}} \pm \frac{3}{2}\,(2m_I - 1)\, Q_j^2 \;, \tag{5.5.44}$$

with $Q_j^2 = \sum_\zeta Q_{\zeta j}^2\, n_{j\zeta}^2$, where $Q_{\zeta j}$ and $n_{j\zeta}$ are, respectively, the principal values of $\overleftrightarrow{Q}_j$, and their direction cosines with respect to the direction of B_j^{eff}.

Fine structure terms for high-spin centres

When a defect has two or more paramagnetic electrons coupled together to give a resultant spin $S \geqslant \frac{1}{2}$, may result in further splitting of the lines from direct and indirect interactions between the different paramagnetic electrons averaged over their spatial distribution. If these interactions are significant, m_S is no longer a good quantum number and states within the spin multiplet will have different energies in the absence of a magnetic field or hyperfine interactions.

The fine structure in the EPR spectrum arising from the above effects—known as zero-field splitting—is included in spin Hamiltonian as

$$\vec{S} \cdot \overleftrightarrow{D} \cdot \vec{S} \ . \qquad (5.5.45)$$

The $\overleftrightarrow{D}$ tensor describes the interaction between electronic magnetic dipoles located at distant neighbours. The classical expression describing a dipole-dipole interaction (5.5.33) applies here.

The effect of this extra term is to destroy the equal energy spacings produced by the $\beta\vec{S}\cdot\overleftrightarrow{g}\cdot\vec{B}$ term, producing a $2S$ equally separated group of transitions for each $m_s \to m_s+1$ transition.

Final remarks

From a study of the angular dependence of the EPR spectrum is a single crystal, $\overleftrightarrow{g}$, $\overleftrightarrow{A}_j$ and $\overleftrightarrow{D}$ can be determined; $\overleftrightarrow{g}$ and $\overleftrightarrow{D}$ reflect the overall character and symmetry of the defect. The spin of the defect is simply determined from the number of spectral fine-structure lines.

$\overleftrightarrow{A}_j$ reflects the symmetry and character of the spin-containing electronic wavefunction at each of the j atom sites in the immediate vicinity of the defect. This analysis is usually performed using the 'defect-molecule' model within the LCAO-MO approximation.

The intensity of the hyperfine satellites determines the isotopic abundance of the j nucleus involved. The nuclear Zeeman interaction provides an unambiguous chemical identification of the j atom involved. From its quadrapole interaction, the total charge density surrounding the j-th atom can be probed.

From the microspic information described about, a detailed model of the defect can be constructed.

Auxiliary techniques

A number of auxiliary techniques are often used to complement the information obtained from the EPR/ENDOR experiment (Watkins, 1998). Two of them are particularly important:

(a) *Optical illumination*

In situ illumination of the sample can also provide a wealth of additional information. Excited states can be produced, provided their lifetime is long enough to allow their detection. The defect alignment can also be achieved by polarised light.

At sufficiently low temperatures, with the phonon-assisted transitions heavily reduced, electronic equilibrium times can be extremely long. As a result, different charge states of the same defect can be produced by selective generation with near-bandgap monochromatic light. From the wavelength dependence of this process, the electrical level positions of these defects can be estimated. This method is not ideal as some of the excited states of the defect are necessarily not paramagnetic and therefore invisible to EPR.

(*b*) *Temperature effects*

Annealing studies are another alternative tool to characterise a paramagnetic centre. For example, changes in the EPR spectrum can be monitored, at cryogenic temperatures between isochronal, or isothermal higher thermal anneals.

These changes are usually analysed by monitoring the variation of spectral linewidth with temperature. The effect usually responsible for these changes in the linewidth is usually a thermally activated hoping of the paramagnetic electron from one Jahn-Teller direction to another.

5.5.4 Local vibrational mode spectroscopy: infra-red absorption

The incorporation of impurities into the crystal that are lighter than the host atoms, gives rise to new vibrational modes with frequencies higher than the host's vibrational modes. The vibrational spectroscopy of these so-called *local modes* has become an important tool for the study of defects in solids. Despite not being tailored for the sole study of deep-level defects, this technique has been successfully applied to study the vibrational properties of H-containing deep-level defects. Since it can probe both paramagnetic and non-paramagnetic centres, its application becomes of paramount importance when combined with EPR.

Background

Let us consider an infinitely long chain of $2N$ atoms forming N unit cells of length $2a$, with a basis of two atoms with masses m and M with $(m < M)$. The spring constant is Λ. The frequencies of the associated normal modes are given by

$$\omega^2 = \Lambda \left(\frac{1}{m} + \frac{1}{M}\right) \pm \Lambda \left[\left(\frac{1}{m} + \frac{1}{M}\right)^2 - \frac{4\sin^2(q\,a)}{m\,M}\right]^{\frac{1}{2}} \tag{5.5.46}$$

where q is the wave vector. Corresponding to the two signs in (5.5.46) there are two branches of the phonon dispersion curve of the diatomic linear chain (Fig. 5.10).

The amplitudes of vibration A_1 and A_2 of alternate atoms along the chain satisfy the relation

$$\frac{A_1}{A_2} = \frac{2\Lambda\, \cos{(q\,a)}}{2\Lambda - m\,\omega^2} = \frac{2\Lambda - M\,\omega^2}{2\Lambda\, \cos{(q\,a)}} \,.$$

The maximum lattice frequency $\omega_{\rm max}$ which occurs at $q = 0$ is given by

$$\omega_{\rm max}^2 = 2\Lambda \left(\frac{1}{m} + \frac{1}{M}\right) . \tag{5.5.47}$$

At the Brillouin zone boundary where $q = \pm\,\pi/2a$, it obvious that $\omega_1^2 = \sqrt{2\Lambda/M}$ and $\omega_2^2 = \sqrt{2\Lambda/m}$ for the acoustic and optical branches respectively. If $m = M$, as for silicon, the optical and acoustic branches are degenerate at the zone boundary but otherwise there is a 'gap' between the two bands of frequencies.

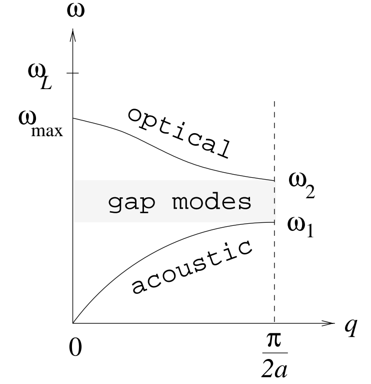


Figure 5.10: The phonon dispersion spectrum of a diatomic linear chain plotted in the first Brillouin zone.

The effects due to the presence of impurities with low masses, m', like H in silicon, are of great importance. In this case, a new vibrational mode emerges at a frequency ω_L greater than $\omega_{\rm max}$. To first order, the local inter-atomic forces Λ' may be assumed to be little changed from those of the host crystal and we can write (Newman, 1973)

$$\omega_L^2 = 2\Lambda' \left(\frac{1}{m'} + \frac{1}{\chi\, M_N}\right) , \tag{5.5.48}$$

where M_N is the mass of the nearest neighbours to the impurity and χ is a parameter that depends on the local angle bending and bond stretching force constants. Since $\omega_L > \omega_{\rm max}$, vibrations cannot propagate throughout the crystal and the impurity mode is said to be *localised*.

Infra-red absorption spectroscopy

Not all the impurity vibrational modes are infra-red (IR) active. The transition probability for an allowed electric-dipole transition between vibrational states is proportional to the square of the matrix element of the electric dipole moment of the centre, which is a function of the coordinates of the defect atoms.

Within the harmonic approximation, a mode is IR active if the motion of the defect atom(s) gives rise to a change in the dipole moment of the centre. In other words, if there is a non-zero dipole-moment derivative with respect to the vibration's normal mode coordinates. The vibrational mode will only be excited by incident light that has a component of its electric vector polarised along the direction of the oscillating dipole moment.

The integrated intensity of absorption, I, due to a defect can be defined, in CGS units, as (Leigh & Szigeti, 1967)

$$I \propto \int \alpha(w)\, dw = \frac{2\pi^2\, \rho}{n\, c\, m'}\, \eta^2 \ , \tag{5.5.49}$$

with the summation over all degenerate modes. Here η is the effective charge, c is the velocity of light, n the refractive index of the material, m' and ρ are the mass and concentration of the impurity respectively; α is the absorption coefficient at the angular frequency w.

The square of the 'effective charge' of the vibration, η^2, is given by

$$\eta^2 = \sum_k m' \left(\frac{\partial \mu}{\partial Q_k}\right)^2 , \quad (5.5.50)$$

where μ is the dipole moment in the direction of the polarisation of the electromagnetic field. Q_k is the normal coordinate of the k-th mode.

Apart from the defect concentration (ρ) and the effective charge (η), all terms in (5.5.49) are known. To determine the former, experimental calibrations have to be performed by estimating the concentration using an alternative method like secondary ion mass spectroscopy.[3] Note that this calibration has to be performed for each and every defect.

Auxiliary techniques

Like for EPR, a number of auxiliary techniques are available which effectively increase the strength of vibrational spectroscopy. These can be summarised as follows:

(*a*) *Isotopic shifts*

The effect of isotopic substitutions on the vibrational frequencies of a defects often makes possible the identification of the vibrating defect atoms. The isotopic shifts can also provide detailed information about the arrangement of the atoms in the defect, like bond angles. Recently, the ability to engineer the isotopic constitution of the host crystal has been achieved (Haller, 1995).

The following commonly used expression provides a simple way to account for the effect that both the mass of the light impurity, m, and one of its immediate neighbours, of mass M_N, have upon the vibrational frequency (Leigh & Newman, 1982; Leigh *et al.*, 1994)

$$\omega^2 = \Lambda \left(\frac{1}{m} + \frac{1}{\chi M_N}\right) , \quad (5.5.51)$$

where χ is a parameter.

(*b*) *Stress alignment*

When used in conjunction with vibrational spectroscopy, uniaxial stress perturbations provide additional information about the symmetry, reorientation kinetics and ground-state energy shift of a defect (Stavola, 1998b).

[3] This technique involves the bombardment of the wafer with primary ions and the detection of secondary ions sputtered from the sample surface. This is a quite destructive and complex experimental procedure, usually used to detect the presence of dopants and their depth distribution (Clegg, 1990).

The stress induced shift, Δ, of the transition energy of a specifically orientated defect can be written in terms of a symmetric piezospectroscopic tensor, $\overleftrightarrow{A}$ with components A_{ij}, as

$$\Delta = \sum_{ij} A_{ij}\,\sigma_{ij} \qquad \text{with} \qquad \sigma_{ij} = s\,n_i\,n_j\ . \tag{5.5.52}$$

Here σ_{ij} are the components of the stress tensor in a Cartesian referential, s is the magnitude of the stress, and n_i and n_j are the direction cosines of the stress vector along axes i and j.

With the results of Kaplyanskiĭ (1994), who tabulated the stress split components, their shift rates and their relative intensities for transitions between non-degenerate levels for the all the seven symmetry types, it is possible to identify the symmetry of the defect responsible for the observed stress pattern.

5.6 Photoluminescence

Photoluminescence (PL) is one of many techniques that has been widely applied to investigate intrinsic electronic transitions and electronic transitions at impurities and defects. Other related techniques are also available, like cathodoluminescence (CL), carrier injection (electroluminescence) or absorption. PL means luminescence excited by photons, as distinct from accelerated electrons (CL) or carrier injection (electroluminescence).

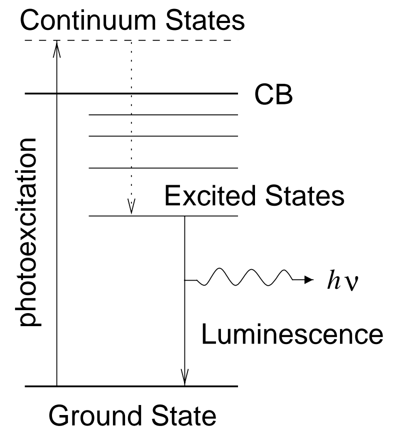


Figure 5.11: Schematic representation of low temperature photoluminescence. Adapted from Lightowlers (1976).

During a PL experiment, the sample is illuminated with light of a frequency such that $h\nu > E_g$, where E_g is the bandgap energy, which creates electron-hole pairs. The electrons rapidly relax, radiatively or more likely non-radiatively, into lower, excited, energy states, eventually reaching defect states. The electrons ultimately recombine with a hole in the valence band or at another defect related state, which results in a photon with an energy equal to the difference in energies between the ground and excited states. If there is a multiplicity of excited states (Fig. 5.11), only transitions from the lowest excited state can generally be observed at low temperatures because of rapid thermalisation.

PL and absorption differ in that absorption can reveal most of the excited states, provided that the optical transition is allowed by symmetry. Under absorption, a continuum of optical frequencies is incident on the sample, and the 'transmitted' spectrum is recorded.[4] The difference between the incident and transmitted spectra is the absorption.

[4]The change in intensity of the transmitted light, I_t, is proportional to the number of absorbing centres: $I_t = I_0\,\mathrm{e}^{-\alpha t}$, where α is the absorption coefficient.

Despite being difficult in the case of thin samples or those with low defect concentrations, absorption measurements can be used to determine the concentration of centres (Dexter, 1962), as in EPR or IR absorption.

Along with the zero-phonon line (ZPL), defined as the transition energy associated with the luminescence process with *no* phonon emission, the above optical techniques reveal transitions that are mediated by one or more phonons, termed phonon side bands. The difference in energy between the ZPL and one-phonon side-band is equivalent to the phonon energy $\hbar\omega$. Hence, this can be employed to determine the local vibration modes of the defect responsible for the ZPL.

It is interesting to discuss the microscopic interpretation of the measured transitions for the case of a deep centre. Figure 5.12 depicts total energies as a function of the atomic positions described by a configuration co-ordinate Q. The lowest state $E[\mathrm{V}^n\mathrm{D}^m\mathrm{C}^0]$ represents the equilibrium ground state of a neutral defect with fully occupied valence (n electrons) and defect (m electrons) levels and with an empty conduction band (CB).

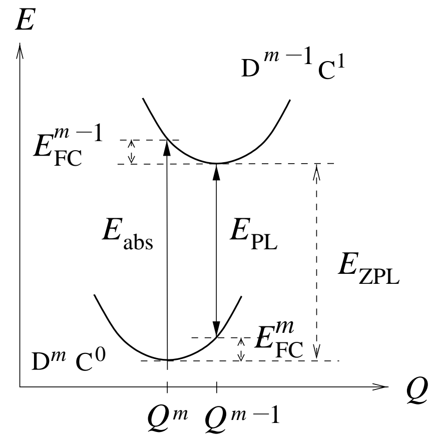


Figure 5.12: Schematic configuration co-ordinate (Q) diagram of the luminescence transitions for a given defect.

Upon excitation with above-bandgap energy photons, an electron is removed from a VB state, or the impurity level, and promoted to the CB. The energy cost is E_{abs} (Fig. 5.12).

In a second stage, if the electron was excited from the valence band, the hole left behind in the VB lowers its energy by relaxing into the defect level. At this point, the system is in its excited configuration, denoted $E[\mathrm{V}^n\mathrm{D}^{m-1}\mathrm{C}^1]$, corresponding to the electron in the CB, bound to the positively charged, $m-1$ electron defect (*impurity bound-exciton*).

After the vertical electronic transition into this excited state, the system relaxes via phonon emission. The Frank-Condon shift in energy associated with this process in the charged ($m-1$ electrons) defect is denoted by E_{FC}^{m-1} in the above figure.

In the relaxed excited state the system has a finite probability to make a radiative transition to the ground state, which results in the PL emission with energy E_{PL}.

Since the electronic transition is much faster than the lattice response, the true ground state is reached only after phonon emission (E_{FC}^m). The E_{ZPL} of the ZPL can be thought of as the electronic transition where the ionic configuration can follow the electronic transition adiabatically, so E_{ZPL} can be approximated as (Mattila & Zunger, 1998)

$$E_{\mathrm{ZPL}} = E[\mathrm{V}^n\mathrm{D}^{m-1}\mathrm{C}^1] - E[\mathrm{V}^n\mathrm{D}^m\mathrm{C}^0]\ , \tag{5.6.1}$$

where the total energies correspond to equilibrium configurations, *i.e.*, minima of the two parabola in Fig. 5.12.

A key to assigning PL features to atomic species are isotopic effects. The zero-point energy of atom differs for different isotopes and the optical transition energy therefore is

also different.

Uniaxial stress measurements can determine the symmetry of the centre, as well as the symmetries of the electronic states involved in the optical transition. The energy separation between nearby gap-states can be determined by monitoring the changes in the ZPLs with temperature.

5.7 Modelling deep levels

A well-known example of a deep-level defect is a transition metal impurity. Any theoretical method capable of describing this type of defect, has to describe accurately the localised character of these centres, and more importantly, any distortions.

Pioneering studies of TM defects in Si were carried out by Cartling (Cartling, 1975) and Hemstreet (Watkins, 1983; Hemstreet, 1976; Hemstreet & Dimmock, 1979; Hemstreet, 1980) using a self-consistent $X\alpha$-cluster formalism. Although the clusters used were very small and no structural relaxation was included, the method served to demonstrate that the impurities introduce gap levels, and that their wavefunctions were mainly localised on Si atoms in the defect core.

Next, more sophisticated methods were introduced to determine the energy levels of TM substitutional defects. These included the use of methods based on Green's functions within the density-functional formalism (Alves & Leite, 1986; DeLeo *et al.*, 1982; Pêcheur & Toussaint, 1983; Watkins *et al.*, 1983; Picoli *et al.*, 1984; Lindfelt & Zunger, 1984; Alves & Leite, 1984; Katayama-Yoshida & Zunger, 1985; Fazzio *et al.*, 1985; Delerue *et al.*, 1989; Beeler *et al.*, 1990; Dreyhsig *et al.*, 1993).

The primary advantage of Green's functions is that they treat an isolated defect in an otherwise infinite perfect crystal with the same accuracy that one chooses to treat the corresponding perfect crystal. The key to its success is the fact that one first treats the infinite crystal, taking advantage of the periodicity and Bloch's theorem. The main disadvantage is that this method is more difficult to implement than cluster methods.

All these different approaches have provided information on the properties of TM impurities in Si but with a relatively large error bar. The source of this error varies from one study to another, either being due to the use of a very small clusters, lack of spin polarisation, poor description of correlation effects or, the absence of structural optimisation. The inclusion of the latter is of course essential in the understanding of JT distortions (Watkins, 1992; Lannoo, 1987; DeLeo *et al.*, 1988; Öğüt *et al.*, 1997).

Of particular interest here are the donor and acceptor levels associated with TM-related defects. The position of these levels within the bandgap are usually calculated as the energy difference between a defect state and the band-edge. Consider the process involving the capture of an electron in the conduction band by a defect level D of a neutral impurity centre. The corresponding change in the total energy (E_T) of the system leads to the

$(-/0)$ level given by[5]

$$E_c - E(-/0) \equiv E_T[D^\gamma; (\mathrm{CB})^1] - E_T[D^{\gamma+1}; (\mathrm{CB})^0] \,. \tag{5.7.1}$$

A practical disadvantage in using this expression to evaluate the $(-/0)$ level lies in the fact that the energy change due to the change in occupancy of two levels is very small compared with the total energy. However, Slater's transition-state method (§2.8.2) allows a much more accurate calculation of this energy difference. This assumes that E_T is a continuous function of the occupation numbers of γ electrons in the defect level D and in the conduction band (CB) respectively, and therefore each of the above expressions can be found from a Taylor expansion about the transition state defined by the configuration

$$[D^{\gamma+1/2}; (\mathrm{CB})^{1/2}] \,. \tag{5.7.2}$$

The difference in total energies of the two configurations is then related to the difference in the derivatives of E_T with occupancy of the defect level D and the CB. These derivatives, by Janak's theorem (§2.8.1) are simply the Kohn-Sham energy levels. Thus,

$$E(-/0) \backsimeq \varepsilon_c - \varepsilon_D \,, \tag{5.7.3}$$

where ε_D is the Kohn-Sham eigenvalue corresponding to the defect level, with the electronic configuration of the transition state. The energy level of the CB is taken to be independent of configuration.

Similarly, the donor level can be related to the difference in ε_D and ε_v in the transition state configuration corresponding to an occupancy $D^{\gamma-1/2}$. A problem is that none of the *ab initio* methods are able to describe simultaneously the localised defect state and the extended CB and in particular the energy gap. To circumvent this, the band gap calculated by the method is either scaled or a scissors operator (§4.10.3) is used to bring it into alignment with the experimental one. This method has been applied to calculate the levels of a number of substitutional isolated TM centres in Si with considerable success (Fazzio *et al.*, 1985; Beeler *et al.*, 1990).

5.7.1 Electrical level calculations with supercells

Despite some advantages, due to the fact that most supercell methods use a plane-wave basis, the supercell approximation has also obvious drawbacks. This relates to the interaction between the defect and its periodic replicas. If the defect-defect distance is not large enough the electronic structure of an isolated defect becomes distorted due to a defect energy band in the bandgap with a finite dispersion and localisation of the deep level wavefunction may change considerably. The size of the supercell also restricts the ionic relaxation. The relaxation pattern is truncated midway between a defect and its nearest periodic image.

Within a supercell method, the total energies are used to extract the position of a given deep level within the bandgap, by finding the value of the electron chemical potential that

[5]The notation (i/j) means that the defect has charge i if the level is filled with an electron, j if empty.

minimises the energy (Pöykkö *et al.*, 1996; Puska *et al.*, 1998; Pesola *et al.*, 1998). This is done indirectly from the formation energy of a given defect.

The formation energy of the defect in the charge state Q is calculated as (Qian *et al.*, 1988; Zhang & Northrup, 1991; Mattila & Nieminen, 1996)

$$E_f^Q = E_d(Q) - Q\,(\mu - E_v^Q) - \sum_s n_s\,\mu_s\ , \tag{5.7.4}$$

where E_d is the total energy of the defect supercell, E_v^Q is the position of the VB top for the defect supercell with charge Q, and μ is the Fermi energy. The constituents of the last summation, n_s and μ are the number of type s atoms in the cell and the atom chemical potential, respectively.

The position of the deep level, or ionisation level (Q'/Q), is given by the position of the Fermi level, when the defect charge state changes from Q' to Q. To determine μ, the equality between the formation energy of defect in the two charge states Q' and Q is imposed. Analytically,

$$E_d(Q) - Q\,(\mu - E_v^Q) = E_d(Q') - Q'\,(\mu - E_v^{Q'}) \tag{5.7.5}$$

To perform this calculation is necessary to align the energy levels in order to get the position of the VB maximum in the defect supercell. This mismatch arises from finite size effects, *i.e.*, the variation of the supercell volume with charge state. The line up of the levels is performed using the average potential correction, as proposed by García and Northup (1995). The value of E_v used is the corresponding value for the bulk supercell corrected by the difference between the difference between the potential in a bulk-like environment of the defect supercell and the average potential in the ideal bulk supercell, that is to write (Pöykkö *et al.*, 1996)

$$E_v = E_v^{\text{bulk}} + (V_{ave}^{\text{defect}} - V_{ave}^{\text{bulk}})\ . \tag{5.7.6}$$

5.7.2 Novel approach to the calculation of deep levels

Very recently, a new approach to the calculation of electrical levels within a cluster approach has been introduced (Resende *et al.*, 1999). To circumvent the difficulty of describing the band gap, *i.e.* the exact position of valence band maximum and conduction band minimum, the electron affinity of the defect A_d is calculated. This is simply the difference in energies between a charged and neutral defect which again can be calculated by a Taylor expansion about a relaxed transition state configuration $D^{\gamma+1/2}$. If the equilibrium structures of the neutral, charged and transition states are denoted by $\vec{R}^0$, $\vec{R}^-$, and $\vec{R}^t$ respectively, then

$$\begin{aligned} A_d &= E_T\big(D^\gamma, \vec{R}^0\big) - E_T\big(D^{\gamma+1}, \vec{R}^-\big) \\ &\backsimeq -\varepsilon_D - \frac{\partial^2 E_T\big(D^{\gamma+1/2}, \vec{R}^t\big)}{\partial\gamma\,\partial\vec{R}} \cdot \big(\vec{R}^- - \vec{R}^0\big). \end{aligned} \tag{5.7.7}$$

Here ε_D is the Kohn-Sham eigenvalue associated with the defect level in the transition state. The second term can be ignored if the difference in the structures of the neutral and charged states of the defect are usually small. This term accounts for changes in structure due to partial occupancy.

The affinity is then related to the Kohn-Sham level of the defect in a transition state. However, the calculation of this quantity is subject to the same difficulties as described above. For example, the electron affinity of bulk Si is simply the band gap which cannot be calculated with sufficient precision. However, we can compare the electron affinities for two defects having localised states within the band gap. It was already shown for the case of the water molecule, that ionisation potentials and affinity energies can be accurately calculated using Janak's theorem with Slater's transition state concept (Table 2.8.3). Their difference being total energy differences, are simply their relative acceptor levels. In practice, the carbon interstitial defect, C_i, which possesses $(-/0)$ and $(0/+)$ levels at $E_c - 0.1$ and $E_v + 0.28$ eV respectively (Song *et al.*, 1990) is used as a standard. For second acceptor levels, the $(=/-)$ level of PtH_2 is used, which is taken to lie at $E_c - 0.16$ eV (Weinstein & Stavola, 1999; Sachse *et al.*, 1999b).

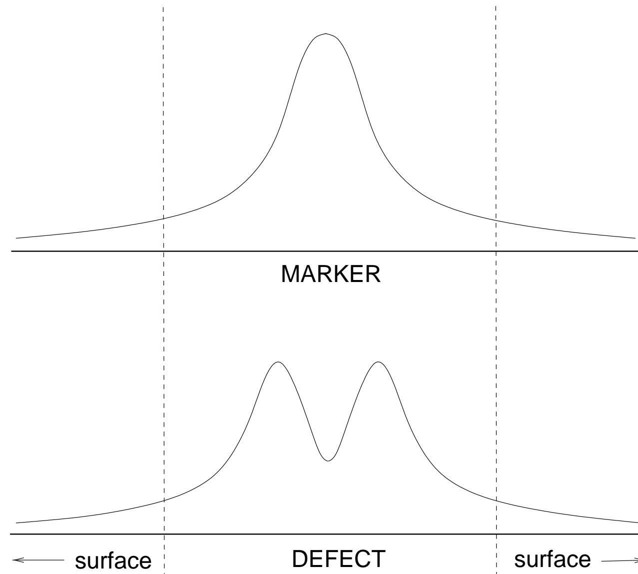


Figure 5.13: A schematic of the localised wavefunction of the ionised defect and that of the marker. Note that the decaying tail of two wavefunctions is identical in the near-surface region, which is shown delimited by the dotted lines.

Now, a deep donor level, $(0/+)$, with respect to the top of the VB, is the difference between the ionisation energy of the defect, I_d, and that of bulk Si. Since the latter cannot be accurately accounted for, for the reasons discussed above, we write the donor level as,

$$(0/+)_d = I_d - \overbrace{(I_s - (0/+)_s)}^{I_{\text{bulk}}} \,, \tag{5.7.8}$$

where I_d and I_s are the calculated ionisation energies of the defect and the standard, and $(0/+)$ is the observed donor level position of the standard or 'marker' defect. For an

acceptor level, $(-/0)$, we have an analogous expression,

$$(-/0)_d = A_d - A_s + (-/0)_s \ , \tag{5.7.9}$$

with A_d, A_s and $(-/0)$ being, respectively, the calculated electron affinities of the defect and the standard, and observed position of the marker.

The error associated with this method depends on how well the asymptotic part of the wavefunction of the ionised defect, which is related with the depth of the level from the band edges, compares with that of the standard defect (see Fig. 5.13). This assumes that the cluster is large enough to avoid any interactions between defect and surface states. Consequently, the method works best when the defect levels are close to those of the standard.

The shift in the calculated levels caused by the surface is

$$\int \Delta V(\vec{r})\, |\psi_\lambda(\vec{r})|^2 \, d^3 r \ , \tag{5.7.10}$$

where ΔV represents the difference between the potential of an infinite solid and a finite cluster, being identical for both the defect and the standard if each were simulated using similar sized clusters and whose localised wavefunctions were asymptotically equal (see Fig. 5.11).

In general the calculated electrical levels of deep-level impurities are within about 0.2 eV of the experimental results where available. In the next chapter, we present the results of a comparative study of the electrical properties of a number of well known deep centres in irradiated and implanted Si.

This method supersedes an earlier one based on the scaling of the band gap although the two give rather similar results for substitutional Au and Ag complexed with hydrogen (Resende *et al.*, 1997).

6

Radiation Defects

6.1 Introduction

The physics of radiation damage in semiconductor materials like crystalline silicon is now a matter of great interest. This is due to the desirability of understanding, and ultimately controlling, damage produced by high-energy ion irradiation. This is of vital importance for space-born Si electronics, and high-resistivity Si particle detectors in 'rad-hard' environments (Watts (1999) and references therein). For both applications, there is a need to tackle problems in the understanding of defect-defect and defect-impurity interactions at both the experimental and theoretical front.

Ion implantation is a well established technique in semiconductor technology with many applications (Williams *et al.*, 1993). This is the case of silicon integration techniques like the SmartCut® process (Bruel, 1995; Freund, 1997) or the wafer exfoliation mechanism (Weldon *et al.*, 1997; Weldon *et al.*, 1998).

Damage to the crystal occurs when an incoming high-energy particle hits the nucleus of the host Si atoms. Obviously, the damage is highly dependent on the energy of the incoming ions, depending also on their numbers. For sufficiently low energies the damage to the crystal is minute, resulting mainly on lattice heating. But above a certain threshold energy, atoms are knocked-out from their crystal equilibrium sites, producing mainly vacancy-interstitial (Frenkel) pairs. The recoiling atoms are free to move across the crystal or interact with newly formed vacancies and interstitial defects, as well as with available impurities or dopants. The motion of the recoiling atom, as well as the impinging particle proceeds via successive atomic collisions, which subsequently reduce their kinetic energy, being absorbed by the lattice. If the energy of the incoming particle is much higher than the displacement energy for one atom, than a large number of Frenkel pairs and recoiling atoms are generated during the collision cascade (Fig 6.1). The envelope of a single collision cascade is usually to as *displacement cluster*. The size of the displacement cluster, *i.e.*, a localised agglomerate of point defects, depends on many parameters, like incident particle mass, energy, mass of the target ions, and temperature of the target material. Energy losses occur via two processes, *i.e.*, nuclear and electronic collisions. For high-energy particles in the MeV range, energy loss due the electronic scattering dominates towards the end of

ion trajectory. This results in non-uniform distribution of energy in the samples after ion implantation, and consequently a non-uniform defect distribution with a peak at the end of the ion range.

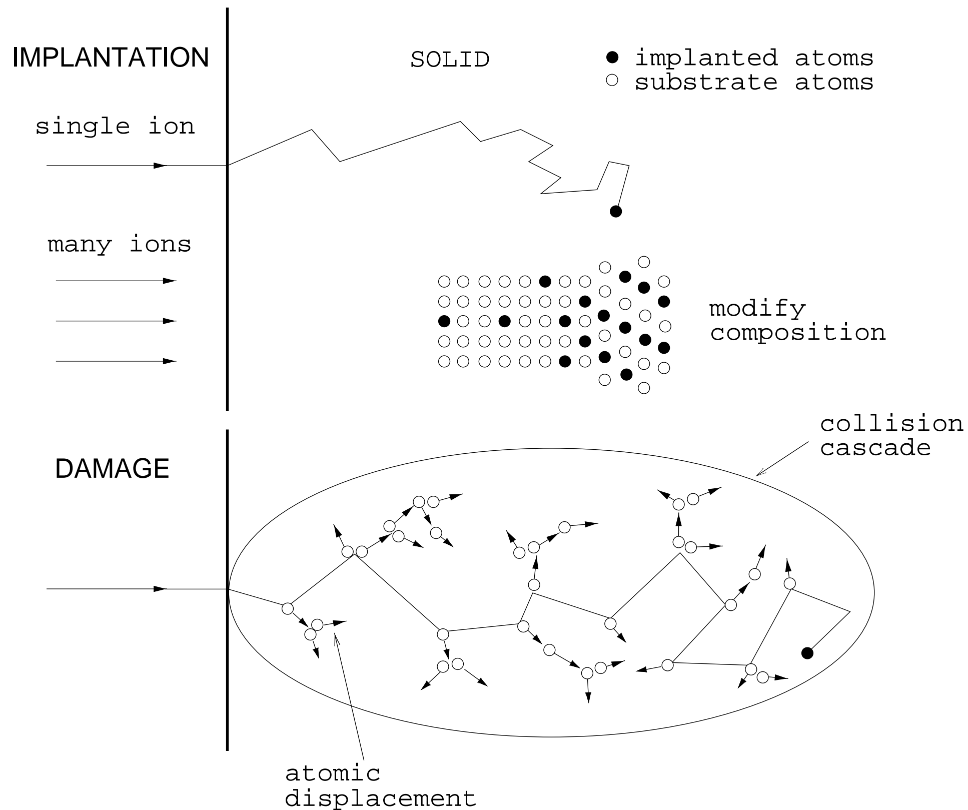


Figure 6.1: Schematic of the implantation and damage processes using energetic ion beams, after Williams and Poate (1984).

In principle, irradiation of protons or α-particles should result in a similar 'damage pattern' to that from ion implantation. Slightly different patterns are formed after irradiation of MeV-electrons, sub-atomic particles—muons or pions, or even γ-rays. Despite this fact, most of the point defects created are common to both irradiation and implantation processes.

The former group of particles, atomic-like particles exhibiting shell structure, give rise to high-order agglomerates of vacancies and interstitials. The so-called 'rod'-like defects, or {311} defects, are a perfect example of an interstitial type of defect (Stolk *et al.*, 1995; Eaglesham, 1995). {311} defects are intimately linked to the problem of transient enhanced diffusion or TED. One of the first clear evidence of TED was published by Michel *et al.* (1987). Apparently, TED occurs due to anomalous diffusion of boron, or arsenic, in low-energy Si implanted material due to the pairing with fast Si self-interstitials, released from large reservoirs of Si interstitials, possibly {311} defects, as a result of thermal annealing at temperatures above 800 °C (Cowern *et al.*, 1994; Eaglesham *et al.*, 1997; Agarwal *et al.*, 1997; Gossmann *et al.*, 1997; Agarwal *et al.*, 1999).

Large vacancy clusters, *i.e.*, cavities or voids (Myers & Petersen, 1998), and stacking faults or dislocations, like Frank partials (Frank, 1949), can also be produced.

The latter group can be responsible for smaller-scale vacancy aggregation, depending on the irradiation dose. Evidence for this type of defects are low-order clusters of vacancies, like V_2 (Watkins & Corbett, 1965; Sieverts *et al.*, 1978a) and possibly V_3, V_4 or V_5 (Brower, 1971a; Lee & Corbett, 1973; Lee & Corbett, 1974). The formation of small aggregates of Si self-interstitials, like the di-interstitial (Lee & Corbett, 1976) cannot be excluded. Due to an almost non-existent optical and electrical activity, due to the lack of dangling-bonds, and the very high mobility of the Si self-interstitial (even at 4.2 K in p-type material!), these defects are very difficult to detect.

Very recently, Au 'labelling' has been used to map the concentration of vacancy clusters in MeV Si-implanted silicon (Venezia *et al.*, 1998). The attraction of substitutional transition-metal impurities like gold to voids, created by irradiation, and located in the inactive regions of the device, can be used to control the effects produced by TM impurities. This 'proximity gettering' is being intensively investigated (Mohadjeri *et al.*, 1995; Schmidt *et al.*, 1998; Kinomura *et al.*, 1998). Interstitial TM impurities, like Fe and Cu, are also known to precipitate at Frank-type partial dislocations (Shen *et al.*, 1996; Shen *et al.*, 1997).

Oxygen and carbon are two major impurities always present in the silicon crystal (Table 6.1). Consequently, it is only natural that both influence radiation induced defects. While oxygen captures mono-vacancies, thus leading to narrower defect profiles of vacancy defect in Cz-Si as opposed to epitaxial material (Libertino *et al.*, 1998), carbon is known to interact with self-interstitials, thus reducing thermally enhanced diffusion (TED) (Stolk *et al.*, 1995; Rücker *et al.*, 1998)

Table 6.1: Oxygen and carbon content in different silicon materials (Schmidt, 1998).

		Cz-Si	Fz-Si	epitaxial-Si
Oxygen content	$[\mathrm{cm}^{-3}]$	10^{18}	10^{16}	$< 10^{15}$
Carbon content	$[\mathrm{cm}^{-3}]$	10^{16}	$< 10^{16}$	$< 10^{15}$

Oxygen is present in silicon in its interstitial configuration (O_i)—leading to lattice expansion, while carbon takes up the substitutional position C_i (lattice shrinkage). These are both electrically inactive. However, a wealth of electrically active defect involving these two impurities are known to be formed as a result of radiation damage.

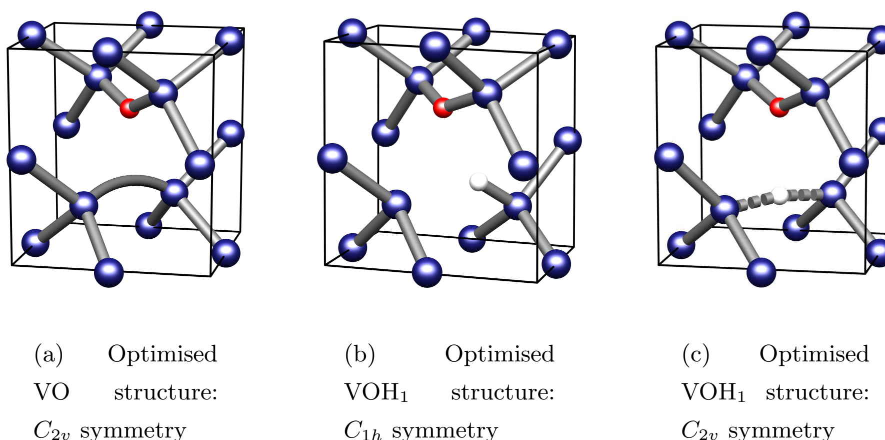
In the next section of this chapter, we present the results of a study on the electrical properties of a number prominent defects in damaged Si material by either implantation or irradiation. Due to the technological and scientific importance of proton implantation, the interaction of these radiation defects with atomic hydrogen is also investigated.

6.2 Electrical properties characterisation

6.2.1 The V–O pair: the *A*-centre

The vacancy-oxygen pair (or *A*-centre) is the most prominent defect in either irradiated or implanted Si material. It is formed in large concentrations, independently of the dopant, and material type, *i.e.*, independently of the oxygen and carbon contents (Table 6.1). Very recent DLTS annealing experiments have shown that the thermal stability of this centre is independent of the mass of the implanting ion, with its concentration increasing with increasing ion mass after moderate annealing ($T \leqslant 200\,^{\circ}$C) (Pellegrino *et al.*, 1999).

This defect is formed by trapping of a highly mobile mono-vacancy by an interstitial oxygen atom (O_i). Its microscopic structure has been established by EPR (Watkins & Corbett, 1961) and its vibrational properties by IR-absorption spectroscopy (Abou-el Fotouh & Newman, 1974; Lindström & Svensson, 1986). This consists of an oxygen atom bridging a pair of Si neighbours of the vacancy (see Fig. 6.2.1a). This structure has been confirmed by theory (DeLeo *et al.*, 1984; Hjalmarson & Jennison, 1985; Snyder & Corbett, 1986; van Oosten *et al.*, 1994; Ewels *et al.*, 1995; Pesola *et al.*, 1999). The formation of a V–O pair allows the RT migration of the otherwise immobile O_i—the barrier for migration was found experimentally to be 1.46 ± 0.29 eV, according to Londos *et al.* (1996) or 1.86 eV by Svensson *et al.* (1985). This defect is only stable when neutral and singly negatively charged. Consequently, only a single level is expected to appear in the band gap: a deep single acceptor level or electron-trap at 0.17 eV below the conduction band minimum ($E_c - 0.17$ eV) (Watkins & Corbett, 1961; Kimerling, 1977).



(a) Optimised VO structure: C_{2v} symmetry

(b) Optimised VOH_1 structure: C_{1h} symmetry

(c) Optimised VOH_1 structure: C_{2v} symmetry

Figure 6.2: Lowest-energy structures of the VO and VOH defects in the transition-state (net proton charge $-\frac{1}{2}\,e$) for the calculation of the defect's electronic affinity. The cube axes indicate the $\langle 100 \rangle$ crystal directions. The VOH defect is shown in its low (b) and high-temperature configurations (c).

Assuming that the oxygen impurity atom is 'healing' two of the four dangling bonds of the vacancy forming an Si–O–Si molecule, the electrical activity of this centre can only arise from the strained, slightly reconstructed, bond between the remaining pair of Si

atoms. That this reconstruction occurs, is already known since the EPR study of Watkins and Corbett (1961). These authors observed a large s character of the one-electron orbital localised on the dangling-bond atoms, consistent with a sp^2 hybridised system, which can only mean a reduction of the s character of the bonds between these atoms and their back-bonded neighbours, and consequently a decrease of the corresponding bond angles with the two Si being pulled from their three immediate neighbours.

Obviously, any serious attempt to simulate the electrical properties of the A-centre should take the reconstruction into account. The magnitude of this reconstruction is intuitively expected to influence of the calculated $(-/0)$ level.

VO C_{2v}	Si–Si	$(-/0)$
Calc.	2.92	$E_c - 0.12$
	3.56	$E_c - 0.32$
Obs.		$E_c - 0.18$

Table 6.2: Calculated $(-/0)$ level (eV) versus Si–Si separation (Å) for the VO pair. The observed value for the VO$(-/0)$ level is given for comparison.

The electrical levels are found by embedding the defect in a tetrahedral 131 ($Si_{71}H_{60}$) H-terminated atom cluster. The wavefunction basis consists of N Cartesian s, p Gaussian orbitals sited on each atom. The charge density is fitted to M Gaussian functions. In this study, (N, M) are: O(5,5), Si(4,5) and H(2,3). Three extra functions, located midway between each bonded pair of atoms, excluding the H-atom terminators, were added to the basis for the wavefunction, and the charge density. The intermediate fit for the latter was of type 3 (§§4.10.1). The clusters were relaxed, as described previously, in either the $+1/2e$, $-1/2e$, or $-3/2e$ charge state to obtain the defect's ionisation and affinities and allow the the evaluation of the $(0/+)$, $(-/0)$ and $(=/-)$ level positions. To force the reconstruction, a pre-relaxation was performed in which the five core atoms remained fixed in the reconstructed configuration, while all the other atoms of the cluster, including terminators were allowed to move. The full optimisation of the pre-relaxed structure was then performed. This procedure is used whenever the defect in question possesses reconstructed bonds, like the E-centre, di-vacancy, VH_1, VH_2, etc.

If we force a strong reconstruction, the $(-/0)$ level associated the C_{2v} configuration is calculated to lie at $E_c - 0.12$ eV. The length of the reconstruction across the two Si atoms, for a transition-state occupancy of $-\frac{1}{2}e$, was 2.92 Å. The magnitude of this reconstruction is considerably larger than the previously reported values of 3.77 for VO^0 by Ewels *et al.* (1995) or 3.37 Å by Pesola *et al.* (1999). If the H-surface is kept fixed during optimisation—like in the calculations of Ewels *et al.* using AIMPRO with a similar sized cluster to that used in the calculations being described here, the reconstruction is lost due to the outwards movement of the second shell of host atoms. In this case, the $(-/0)$ level comes out at $E_c - 0.32$ eV, for a Si–Si length of 3.56 Å (Table 6.2). It is important to note that the Si–O length, the distance of the oxygen impurity atom to the vacant lattice site and the $\measuredangle$ Si–O–Si vary very little from one structure to another. This means that degree of reconstruction is the only factor influencing the calculated $(-/0)$ level, *i.e.*, the position of the electron trap within the bandgap. In other words, the O impurity seems

to be completely passivating two of the vacancy dangling-bond atoms.

It interesting to mention that Pesola *et al.* (1999) calculates the $(-/0)$ level of VO to lie at $E_c - 0.76$ eV. Furthermore, their calculations also predict a $(=/-)$ level, and consequently the stability of VO^{2-}, in disagreement with experiment. The plane-wave method used in the electrical level calculations by these authors was described previously (§5.7.1).

6.2.2 Interaction of hydrogen with the *A*-centre

VOH

Very recent EPR experiments on proton- and deuteron-implanted Cz-grown Si (Johannesen *et al.*, 1999) have confirmed the expectation that the hydrogen atom in VOH^0 is strongly bonded to one of the Si atoms of the mono-vacancy that are not bonded to Si, as previously suggested by theory (Artacho & Ynduráin, 1989) and in contradiction with the results of previous experiments (Gutsev *et al.*, 1989). If the experimental assignment is correct, the resulting structure should be similar to that of VH^0, with the Si–O–Si unit of VOH replacing the reconstructed Si–Si bond of VH^0. Consequently, VHO^0 is expected to exhibit monoclinic-I (C_{1h}) symmetry with the dangling bond and the Si–H bond lying in a {110} mirror plane perpendicular to the plane of the Si–O–Si unit (Fig 6.2.1b). This seems to be the case. A classical analysis of the interaction between the proton and the paramagnetic electron, at low temperature, located the proton at ~ 2.5 Å from the centre of gravity of the electron spin distribution. This value is consistent with the structure described above. It is obtained from the spatial dependence of the anisotropic component of the hyperfine tensor.

The analysis of the angular dependence of EPR signal observed by Johannesen *et al.* (1999), revealed a transition from monoclinic-I to orthorhombic-I ($C_{1h} \rightarrow C_{2v}$) at ~ 240 K. This was interpreted as reflecting thermally activated 'jumps' of the lower symmetry defect between equivalent crystal directions having in common the {100} planes. The activation energy for this motionally-averaged effect was found to be 0.18 ± 0.01 eV. The annealing behaviour of this defect was found to be similar to that of the *A*-centre, as the EPR activity ceases after heat treatments at 588 K.

It is now known that the VOH defect, containing a single hydrogen impurity, possesses a $(-/0)$ at $E_c - 0.31$ eV, as a result of comparative EPR/DLTS annealing studies (Johannesen *et al.*, 1999; Bonde Nielsen *et al.*, 1999). The study of the effect of uniaxial stress on the DLTS peaks of this centre, at 160 K, revealed orthorhombic-I symmetry, confirming the observations of Johannesen *et al.* (1999) at the same EPR measuring temperature.

This seems to be a common defect in proton-implanted material, previously reported by several authors, but only tentatively assigned (Svensson *et al.*, 1989; Tokuda & Shimada, 1998; Lalita *et al.*, 1997; Peaker *et al.*, 1999), appearing also in *e*-irradiated Si preceded by a wet-chemical etching treatment (Feklisova & Yarykin, 1997).

For the orthorhombic-I (C_{2v}) configuration of VOH_1, we calculate the $(-/0)$ to be

Table 6.3: Structural parameters for the optimised structure of neutral VOH_1 and VOH_2 defects (Å).

Complex	Symmetry	O–Si	H–Si	H–H
VOH_1	C_{1h}	1.74	1.49	···
	C_{2v}	1.74	1.79	···
VOH_2	C_{2v}	1.72	1.48	1.73

$E_c - 0.22$ eV. To investigate the high-temperature behaviour of this centre, another configuration was considered. In this, the hydrogen impurity atom sat midway between two of the Si neighbours of the vacancy, along the [100] direction, within C_{2v} symmetry (see Fig. 6.2.1c). The calculations revealed that the $(-/0)$ level of VOH drops within the bandgap by less then 0.1 eV to lie at $E_c - 0.29$ eV, to compare with the experimental value of $E_c - 0.31$ eV (Bonde Nielsen *et al.*, 1999; Peaker *et al.*, 1999).

The energy barrier for the hoping of the hydrogen impurity atom was found to be 0.33 eV, in reasonable agreement with the experimental value of 0.18 eV. This barrier was calculated adiabatically, *i.e.*, as the difference in total energies of the fully relaxed C_{1h} and C_{2v} atomic arrangements of VOH.

The distance between the hydrogen impurity atom and the Si dangling-bond atom was found to be 2.72 Å for neutral (C_{1h}) VOH_1, resembling closely that of neutral VH_1 of 2.78 Å (Table 8.4.2). Due to the presence of the hydrogen atom, the oxygen impurity goes slightly off-site (Fig. 6.2.1b) along the $[0\bar{1}1]$ direction by ~ 0.6 Å.

VOH_2

Despite a recent claim by Feklisova and Yarykin (1997) that VOH_2 is an electrically active defect, there is no direct experimental evidence for this fact. Supported by a simple theoretical analysis based on the relative penetration depth of the DLTS centres, these authors suggested that $E_c - 0.32$ eV DLTS (E4) trap is due to VOH_2. Despite having being able to provide strong evidence for the involvement of hydrogen in the centre,[1] as well as to correlate the E4 with the *A*-centre, the only evidence supporting the suggestion that E4 involves two and *not* one hydrogen atom, comes from the rapid decay of the concentration of the E4 signal with depth, when compared with other signals. The fact that hydrogen concentration is highest in the near-surface region, suggests that this centre might involve more hydrogen atoms than all the other centres.

Pearton (1982) observed a decrease of the DLTS signal of VO as a result of exposure of the diode to atomic and molecular hydrogen. This observation was recently confirmed by Tokuda and Shimada (1998), providing further arguments against Feklisova and Yarykin assignment.

[1]As recently proved by Bonde Nielsen *et al.* (1999).

Our calculations provide no support to the hypothesis of the electrical activity of VOH_2, as the corresponding Kohn-Sham spectrum revealed no levels in the bandgap. The structural parameters for the lowest-energy configuration of neutral VOH_2 are given in Table 6.3. Additionally, no donor levels, (0/+), were found for the *A*-centre or any of its complexes with atomic hydrogen.

6.2.3 The di-vacancy (V_2)

The lattice di-vacancy is another example, like the *A*-centre, of a very important defect in radiation damaged Si material. Due its importance, a whole chapter of this thesis is dedicated to it (Chapter 7).

6.2.4 The V_2O complex

Very little is known about the structural and electrical properties of this defect. To our knowledge no experimental information on the structure of this defect is available. Despite this fact, some information is available as a result of theoretical work (Ewels *et al.*, 1995; Pesola *et al.*, 1999). Concerning its electrical properties, Kimerling *et al.* (1977) has suggested that an electron trap at $E_c - 0.30$ eV is related to this defect, annealing out at $\sim 300\,^\circ$C.

The electrical properties of this defect were simulated using a trigonal 134-atom cluster centred at a bond. The termination was kept fixed and the atomic basis was the following, for a intermediate fit to the charge density of type 0: O(4,5), Si(4,5) and H(2,3); all atoms but the H-terminators in big basis and three extra function were placed midway between each pair of bonded atoms to the description of the wavefunction as well as the charge density. The transition-state calculations revealed an acceptor level, (−/0), at $E_c - 0.47$ eV.

6.2.5 The VP pair: the *E*-centre

The phosphorus-vacancy pair or *E*-centre is known to exhibit structural and electrical properties very similar to VH^0 (Watkins & Corbett, 1964; Bech Nielsen *et al.*, 1997; Bonde Nielsen *et al.*, 1999). As pointed out previously as a result of EPR measurements (Bech Nielsen *et al.*, 1997), the VH^0 defect can be considered a group-V–vacancy like defect, with the phosphorus atom replacing one of the neighbours of the vacancy. This defect is stable up to $\sim 150\,^\circ$C, not being formed in Si material with a phosphorus concentration $\lesssim 10^{14}$ cm^{-3} (Watkins & Corbett, 1964; Kimerling *et al.*, 1975; Schmidt, 1998). Since the structural and electrical properties of the Jahn-Teller distorted *E*-centre are the subject of a forthcoming chapter of this thesis (Chapter 8), the results will not be presented here.

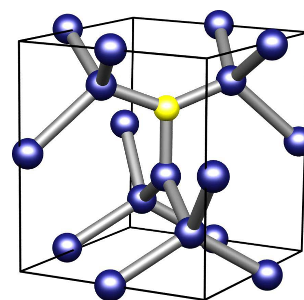


Figure 6.3: Optimised structure of the isolated carbon interstitial C_i with a partial occupancy of $-\frac{1}{2}\,e$.

6.2.6 The isolated C_i impurity

In untreated Si material, carbon exists as a substitutional defect (C_s), being a common and therefore important impurity in this material, typically occurring in concentrations of $\sim 10^{15}$–10^{17} cm^{-3}. Any radiation damage provoked upon the crystal, results in the formation of more complex carbon related defects. The simplest is the isolated carbon interstitial (C_i). This defect is formed when C_s captures a fast diffusing Si self-interstitial (Si_i or I)—a direct product of irradiation/implantation, produced in the primary damage event, releasing the strain associated with both defect structures. A very recent DLTS study on the generation vs. dose rate of Si_i has estimated its migration energy to be as low as 0.065 ± 0.015 eV (Hállen *et al.*, 1999).

Concerning its microscopic structure, the C_i defect can be considered a $\langle 100 \rangle$ Si–C interstitialcy centred at a single substitutional lattice site and therefore having C_{2v} symmetry (Fig. 6.3). This defect has the same structure in all three stable charge states, annealing out at around room temperature. Combined EPR and DLTS measurements were able to show that C_i is in fact an amphoteric defect giving rise to a $(-/0)$ level at $E_c - 0.1$ eV and a $(0/+)$ level at $E_v + 0.28$ eV (Kimerling *et al.*, 1975; Watkins & Brower, 1976; Song *et al.*, 1990).

As mentioned previously, this defect is used as the standard in electrical level calculations.

6.2.7 The C_iP complex

The carbon-phosphorus pair (C_iP) is a bistable defect, having a total of four metastable configurations. This defect was first observed by DLTS in *e*-irradiated Si material which was annealed at 340 K for 30 minutes (Song *et al.*, 1986; Chantre & Kimerling, 1986). It is formed when a substitutional phosphorus impurity captures a C_i, annealing out at $T > 125\,^\circ$C. The ground state configuration gives rise to donor and acceptor levels at $E_c - 0.38$ and $E_v + 0.48$ eV respectively.

The atomic basis used for carbon and phosphorus were C(4,4) and P(4,5), respectively, with a type-0 charge density fit. The hydrogen atoms of the surface were kept fixed during the transition-state relaxations. The calculated levels lay at $E_c - 0.53$ and $E_v + 0.37$ eV for, respectively the $(-/0)$ and $(0/+)$ levels, for a energy separation between the two levels of 0.26 eV, in very good agreement with the observed value of 0.3 eV. When the deep acceptor level lies above the donor one, this energy separation is considered to be positive, by definition. The $(-/0)$ level is expected to lie above the donor level because the second electron, although possible trapped into a similar orbital as the first, and experiencing therefore the same attractive interaction to the core of the defect, is *repelled* by the Coulombic

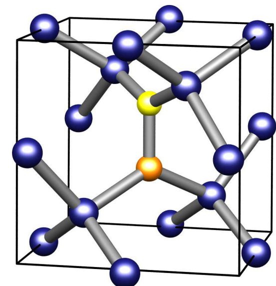


Figure 6.4: Lowest-energy configuration of the C_iP defect.

interaction with the first electron. This added Coulomb repulsion is defined as U and given by the acceptor–donor level separation.[2]

6.2.8 The C_iO_i pair: the K-centre

Annealing at room temperature of *e*-irradiated Si allows the diffusion of C_i and subsequently trapping by O_i to form a C_i–O_i complex. DLTS studies revealed a single level at $E_v + 0.38$ eV (Mooney *et al.*, 1977) and EPR, C_{1h} symmetry (Watkins, 1965). Additional information on the microscopic structure of this defect was provided by Trombetta and Watkins (1988), by means of EPR. Despite this fact, the suggested model could not explain the large shift on the observed vibrational modes of this complex when ^{18}O replaces ^{16}O (Davies *et al.*, 1986). *Ab initio* calculations with AIMPRO have revealed a somewhat unusual configuration for the defect (Jones, 1992). Their lowest-energy structure can be explained as a $\langle 100 \rangle$ C_i split-interstitial (Fig. 6.3) plus a second-shell bond-centred O atom as illustrated in Figure 6.4. The interesting point is that these authors found that as a result of a dative bond between Si_i and the O_i impurity, with the latter becoming over-coordinated, and *frustrated* (Jones, 1992) due to the presence of the carbon dangling-bond.

To model the K-centre, we have used a similar basis to that used to extract the electrical levels of VO, VOH and V_2O, with a tetrahedral 131-atom cluster. The relaxation of the transition-state for the calculation of the defect's ionisation energy, provided such a value that the (0/+) level of the C_iO_i pair was found at $E_v + 0.18$ eV to compare with the experimental value of $E_v + 0.35$ eV.

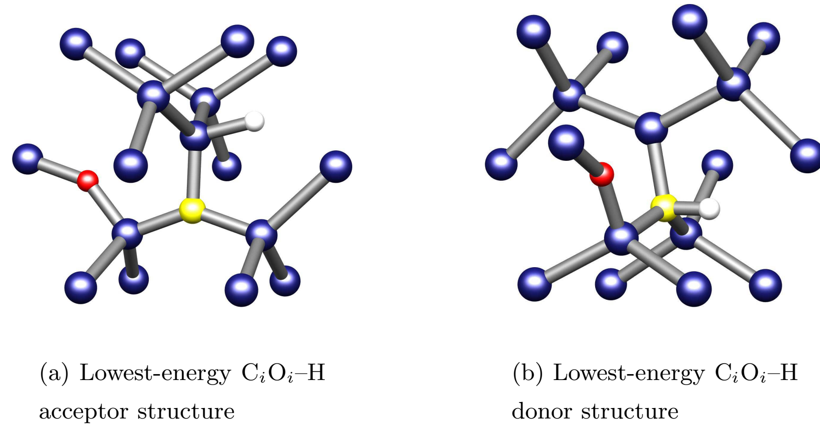
6.2.9 Interaction of hydrogen with the K-centre

To our knowledge no information exists on the electrical properties of C_iO_i–H defects. The only exception is the work by Feklisova and Yarykin (1997). According to these authors, this defect is responsible for an electron trap at $E_c - 0.36$ eV, with an electron capture cross-section of 3×10^{-15} cm^2.

Bearing in mind the structure of the C_iO_i pair, there are obviously two possible configurations for the C_iO_i–H complex. Both the interstitial carbon and the silicon atoms, lying along the [100] direction and possessing dangling-bonds, can readily accomodate an incoming hydrogen impurity atom and form a C–H or a Si–H bond.

We found that this defect can actually assume two different configurations depending on the charge state. When singly positively charged, hydrogen attaches itself to the carbon, while for the singly negatively charge state, the hydrogen atom is bonded to the interstitial silicon atom. The corresponding transition-state structures are shown in Figure 6.5. The corresponding levels were calculated to be $E_v + 0.94$ and $E_c - 0.59$ eV, respectively, for the (0/+) and (−/0) levels, which means that the acceptor level is predicted to lie below the donor level. This is a remarkable result, since it suggests that the present is a *negative-U* centre ($U = -0.36$ eV).

[2] U is often referred to as the Hubbard 'correlation energy' (Hubbard, 1963).



(a) Lowest-energy C_iO_i–H acceptor structure

(b) Lowest-energy C_iO_i–H donor structure

Figure 6.5: Low energy structures of C_iO_i–H complex in two transition states: singly negatively $(-\frac{1}{2})$ and (b) positively charged $(+\frac{1}{2})$ structures.

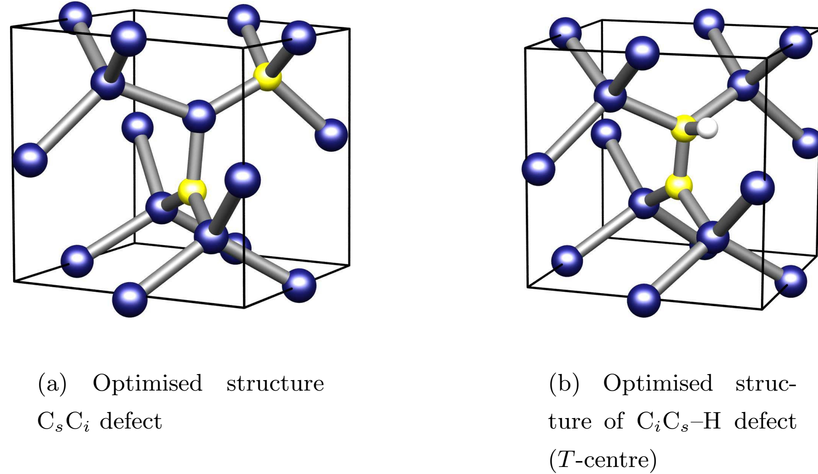
The concept of negative-U in solids was first proposed by Anderson (1975). This author postulated that if the energy (correlation) gain in pairing two electrons, coupled with a possible large relaxation, overcomes the Coulombic repulsion between these two electrons, the levels will cross over, resulting in a negative-U system. In this case, a net attraction between the electrons occurs, resulting in the trapping of electrons by pairs. Known examples of negative-U centres in silicon are the interstitial boron (Watkins & Troxell, 1980b; Troxell & Watkins, 1980; Harris *et al.*, 1987), the lattice mono-vacancy (Baraff *et al.*, 1979b; Watkins & Troxell, 1980b; Newton *et al.*, 1983) and isolated hydrogen (Holm *et al.*, 1991; Johnson *et al.*, 1994; Hitti *et al.*, 1999).

6.2.10 The C_sC_i pair

The di-carbon defect is bi-stable assuming two different configurations when charged (A-form) and neutral (B-form), differing only by a simple bond-switching transformation (Song *et al.*, 1988; Song *et al.*, 1990; Leary *et al.*, 1997). The conversion from one to another can be achieved by near-bandgap illumination, or by externally applying stress. EPR and DLTS measurements have shown that the $(C_i–C_s)^{\pm}$ defects possess C_{1h} symmetry with $(-/0)$ and $(0/+)$ levels at $E_c - 0.17$ and $E_v + 0.09$ eV respectively (Song *et al.*, 1988; Song *et al.*, 1990). It is interesting to note that these levels were previously reported as belonging to the A-centre (Jellison, 1982).

In the A-form, the structure of C_iC_s is similar that of the isolated C_i, with the symmetry being lowered due to the presence of a second carbon impurity at a substitutional lattice site. The relaxed structure of the defect in this form is shown in Figure 6.6a. When the Fermi level is around midgap, the defect assumes the B-form. In this structure, the two carbon atoms lie close to neighbouring lattice sites, with a Si_i close to the bond-centred (BC) site, lying between them.

The affinity and ionisation runs within the transition-state were set for the form-A of the C_iC_s defect. The calculated levels were as follows: the $(-/0)$ at $E_c - 0.07$ and the



(a) Optimised structure C_sC_i defect

(b) Optimised structure of C_iC_s–H defect (T-centre)

Figure 6.6: Optimised structures of the C_sC_i defect (A-form) and the C_iC_s–H defect (T-centre) with a transition-state occupancy of $-\frac{1}{2}\,e$.

(0/+) at $E_v + 0.04$ eV. These are to be compared with the observed levels $E_c - 0.17$ and $E_v + 0.09$ eV, for the (−/0) and (0/+) levels respectively. The structure resulting from the optimisation of the C_iC_s is shown in Fig. 6.6a. Only the defect core and host atoms were allowed to relax.

6.2.11 The C_sC_i–H defect: the T-centre

The C_sC_i–H consists of a ⟨100⟩-oriented C–CH pair, or C–CH split interstitial, occupying a single lattice site (Safonov *et al.*, 1996; Leary *et al.*, 1998). This centre gives rise to the 0.9351 eV luminescence system (T-line), created in either Cz- or Fz-Si material subjected to radiation damage, followed by a 400–600 °C thermal treatment (Irion *et al.*, 1985; Lightowlers *et al.*, 1994). Uniaxial stress and magnetic field measurements have shown that this centre has monoclinic-I C_{1h} symmetry (Fig. 6.6b) and possesses a (−/0) level at $E_c - 0.20$ eV (Safonov *et al.*, 1996). The neutral ground state is paramagnetic, with $S = \frac{1}{2}$, being isoelectronic with the C_iP defect.

The optimisation of the structure of the ⟨100⟩-oriented C–CH defect in the transition-state for the affinity energy resulted in a (−/0) level at $E_c - 0.30$ eV, in good agreement with the experimental value of $E_c - 0.20$ eV. The (0/+) level, never observed experimentally, was calculated to lie 0.38 eV above E_v.

6.3 Summary

Table 6.4: Summary of calculated and observed level positions of common deep-level defects in irradiated and/or implanted crystalline silicon.

Defect	Level	Energy level position (eV)		$\sigma_{n/p}$	Comments
		Calc.	Obs.	(cm^2)	
VO	(−/0)	$E_c - 0.12$	$E_c - 0.17$	$\sim 10^{-14}$	out $\gtrsim 300\,°$C
VOH	(−/0)	$E_c - 0.29^{\natural}$	$E_c - 0.31$	10^{-15}	out 400 °C
V_2	(=/−)	$E_c - 0.21^{\S *}$	$E_c - 0.23$	10^{-15}–10^{-16}	out 200–300 °C
V_2	(−/0)	$E_c - 0.37^{\S}$	$E_c - 0.42$	$\sim 10^{-15}$	”
V_2	(0/+)	$E_c - 0.41^{\S}$	$E_c - 0.42$	$\sim 10^{-15}$	”
V_2O	(−/0)	$E_c - 0.47$	···	···	
VP	(−/0)	$E_c - 0.48$	$E_c - 0.44$	10^{-14}–10^{-15}	out 150 °C
C_i	(−/0)	$E_c - 0.10^{\dagger}$	$E_c - 0.10$	···	out $\gtrsim$ RT
C_i	(0/+)	$E_v + 0.28^{\dagger}$	$E_v + 0.28$	···	”
C_iO_i	(0/+)	$E_v + 0.18$	$E_v + 0.35$	$\sim 10^{-16}$	out $\gtrsim 300\,°$C
C_iO_iH	(−/0)	$E_c - 0.58$	···	···	Negative-U
C_iO_iH	(0/+)	$E_v + 0.94$	···	···	
C_iP_s	(−/0)	$E_c - 0.53$	$E_c - 0.38$	···	metastable
C_iP_s	(0/+)	$E_v + 0.37$	$E_v + 0.48$	···	out $\sim 125\,°$C
C_iC_s	(−/0)	$E_c - 0.07^{\ddagger}$	$E_c - 0.17$	···	bistable
C_iC_s	(0/+)	$E_v + 0.04^{\ddagger}$	$E_v + 0.09$	···	out $\sim 200\,°$C
C_sC_iH	(−/0)	$E_c - 0.30^{\sharp}$	$E_c - 0.20$	···	out $\gtrsim 600\,°$C
C_sC_iH	(0/+)	$E_v + 0.38^{\sharp}$	···	···	

† Markers for the calculation of single acceptor and donor levels.

‡ A-form (C_{1h}).

§ C_{2h} configuration.

* AuH (=/−) as the marker.

♯ C_s–$(C_iH)_{(100)}$

♮ C_{2v} symmetry configuration (see text).

7

The Lattice Divacancy

7.1 Introduction

Being one of the fundamental defects in silicon, the lattice di-vacancy is a well studied defect. Knowledge of its properties is a basic step towards a better understanding of many electron phenomena, defect-defect interactions, diffusion and many other properties.

The fact that this defect is stable, and immobile, at room temperature with a unexpectedly high activation energy for diffusion has allowed its study by a whole range of experimental techniques.

Controlled production of di-vacancies is usually achieved by MeV e-irradiation of the material at cryogenic temperatures (Corbett & Watkins, 1965). The atoms are displaced by Rutherford scattering of the high energy electron, with the low mass of the electron assuring simple damage since the recoiling nuclei obtain little excess kinetic energy, discouraging further atomic displacements by it. Furthermore, at this low temperatures, the displacement products are heavily reduced. Despite this fact, this is not by any means the only process to produce di-vacancies. These are inevitably, unintentionally, created after low temperature or RT irradiation and/or implantation of protons, α-particles, neutrons, heavy ions, γ-rays, pions or muons.

The production rate of V_2 seems to be impurity dependent, increasing with increasing concentrations of carbon, boron and oxygen (Cheng *et al.*, 1966; Lindström *et al.*, 1982).

Di-vacancies are also created during device processing. This defect assumes critical importance in micro-electronic integration, *e.g.*, VLSI (Very Large Scale Integration) processing. Ion implantation, plasma etching or even beam-lithography are major sources for the formation of di-vacancies and vacancy aggregates.

These higher-order vacancies can act as gettering centres. An example, are micro-voids which have been observed to getter transition-metal impurities (Wong-Leung *et al.*, 1995; McHugo *et al.*, 1996; Koveshnikov & Kononchuk, 1998; Myers & Petersen, 1998). This is a very important feature, since it allows one to control the concentration of electrically active centres. This is particularly useful for the optimisation of silicon-integrated devices, like high-resistivity particle detectors in high-radiation (*rad-hard*) environments (Kurokawa *et al.*, 1995), or space-born electronics (Dale *et al.*, 1994; Hopkinson *et al.*, 1996). The

nature of the levels responsible for the enhancement of the leakage current, effective doping changes or the so-called *reverse-annealing* effect, in heavily irradiated diodes is still not yet fully understood. (see *e.g.* Watts (1999) and references therein). It has been recently proposed that a di-vacancy-oxygen defect may be the origin of such detrimental effects (Gill *et al.*, 1997).

In this chapter, we present a review of the experimental and theoretical work performed to date. This is followed by a complete study of the structural, electrical and optical properties of the lattice di-vacancy in crystalline silicon. It addresses two main problems: the microscopic structure of the defect and its electrical properties. These two issues are still matter of great controversy. Even more, when these properties are intimately related.

The lattice di-vacancy (V_2) is formed by the removal of two neighbouring silicon atoms from the crystal. This results in six dangling bonds which induce two doubly degenerate one-electron levels in the silicon band gap. These are labelled e_u and e_g according with the irreducible representations of the point-symmetry group D_{3d}. The singly positively charge state, the anti-symmetric e_u doublet becomes partially occupied with one electron. In this situation of a partially occupied orbitally degenerate orbital, the JT theorem predicts that the symmetry of the defect will be lowered via a lattice distortion, with a consequent splitting of the e_u and e_g manifolds and lowering of the total energy of the system. Obviously, the system is also JT-unstable for V_2^0 and V_2^-.

7.2 Experimental background

From an experimental point of view, the di-vacancy in silicon has been the subject of study for more than 30 years. Different experimental techniques like EPR (Watkins & Corbett, 1965; Corbett & Watkins, 1965; Ammerlaan & Watkins, 1972; Sieverts *et al.*, 1978b), ENDOR (de Wit *et al.*, 1976; Sieverts *et al.*, 1978a; Sieverts *et al.*, 1990), photoconductivity (Kalma & Corelli, 1968; Young & Corelli, 1972; Carton-Merlet *et al.*, 1982), FTIR (Fan & Ramdas, 1959; Cheng *et al.*, 1966; Svensson *et al.*, 1988), positron-annihilation spectroscopy (Kauppinen *et al.*, 1997; Kauppinen *et al.*, 1998) and DLTS (Evwaraye & Sun, 1976; Mooney *et al.*, 1977; Kimerling, 1977; Awadelkarim *et al.*, 1986; Svensson & Willander, 1987; Svensson *et al.*, 1991; Lindström *et al.*, 1982; Hallén *et al.*, 1990; Svensson *et al.*, 1997) have been applied. Despite this, the understanding of the properties of the di-vacancy remains incomplete.

7.2.1 Structural properties

In a series of elegant EPR experiments supported by a whole range of auxiliary techniques, Watkins and Corbett (1965) found the stabilising structure post-JT distortion to have C_{2h} symmetry and spin $S = \frac{1}{2}$ for both V_2^+ and V_2^- charge states, with a g close to the free-spin value. Based on their findings, it was postulated that the behaviour of the six electrons accommodated in the dangling-bond orbitals determines all the relevant properties of V_2. Assuming a six-particle problem within a LCAO picture, Watkins and Corbett proposed

a one-electron model to predict the electronic properties of the di-vacancy. According to this picture, an a_g singlet resulting from the splitting of the π-like molecular mono-vacancy e_g doublet lies below the level arising from the splitting of the $e_u(= a_u + b_u)$ manifold to achieve a configuration $a_g^{\uparrow\downarrow}\, b_u^{\uparrow}$ for V_2^-. In their experiments, the analysis of the hyperfine interactions showed that $\backsimeq$ 60% of the unpaired-electron wavefunction is located on the dangling-bond atoms. Additionally, it was shown that the highest occupied state for V_2^+ and V_2^- structures has amplitude on the $\langle 111 \rangle$ mirror plane of the di-vacancy.

Concerning the microscopic structure, Watkins and Corbett proposed that a JT-driven strong reconstruction of the bonds between four of the dangling-bond atoms, which move inwards towards the vacant lattice sites. The remaining two being responsible for the paramagnetism exhibited by the defect. It was observed as a result of uniaxial stress measurements that V_2 re-orientates preferentially along one of the $\langle 111 \rangle$ JT orientations. This process occurs with an activation energy of $\sim$ 0.06 eV, even at 30 K. Based on this simple observation, the possibility of the involvement of a nearby impurity was rejected. These stress measurements also provided the sense of the distortion. The fact that two pairs of Si atoms were 'pulled together' was evident from the preferential alignment of the defect when the compressional stress was applied in a manner to favour the reconstruction.

Due to a strong JT distortion, a large splitting of the e_u and e_g levels is expected. Assuming that JT effects are more important than those of the crystal field (accounting for the e_g–e_u energy separation), a 'cross-over' of singlet levels resulting from the splitting of the e-manifolds was invoked to explain the spin density localisation monitored by EPR. The magnitude of the JT distortion has to be large enough to provoke level cross-over, which cannot be explained by a modest distortion (see Fig. 7.2).

However, although invoking a large distortion, Watkins and Corbett could not predict the ordering of the one-electron levels since, both, $b_u^{\uparrow\downarrow}\, a_g^{\uparrow}$ and $a_g^{\uparrow\downarrow}\, b_u^{\uparrow}$ configurations agree with their experimental findings. Both a_g and b_u orbitals have amplitude in the di-vacancy mirror plane, *i.e.*, a non-vanishing Fermi contact interaction resulting from the non-zero electron density at the nuclei.

7.2.2 Electrical and optical properties

It has been suggested that the di-vacancy can exist in four different charge states, which give rise to three levels lying deep in the Si band gap. Resulting from DLTS studies (Evwaraye & Sun, 1976; Kimerling, 1977; Awadelkarim *et al.*, 1986; Peaker *et al.*, 1999) the following assignments have been made: a double acceptor level, (=/−), at $E_c - 0.23$ eV; a single acceptor $E_c - 0.42$ (−/0) and a single donor level (0/+) at $E_v + 0.19$ eV.

Kalma and Corelli (1968) have suggested a different picture, confirming an early EPR study (Watkins & Corbett, 1965). The (=/−) level lies deeper at $\sim E_c - 0.4$ eV, locating a single acceptor at $E_c - 0.54$ eV.

There is some discussion around the reported values for capture cross-section of the electron traps at $E_c - 0.23$ and $E_c - 0.42$ eV. Evwaraye and Sun (1976), have reported an unexpected large capture cross-section for the (=/−) trap (see Table 7.1), which has casted

Table 7.1: Observed capture cross-sections (σ_n/σ_p) for the levels of V_2 (cm^2) from DLTS measurements on irradiated silicon.

$(=/-)$	$(-/0)$	$(0/+)$	References
0.6×10^{-16}	1.6×10^{-16}	···	Evwaraye & Sun (1976)
2.0×10^{-16}	4.0×10^{-15}	2.0×10^{-16}	Kimerling (1977)
···	···	3.0×10^{-16}	Mooney *et al.* (1977)
0.6×10^{-16}	1.6×10^{-16}	···	Brotherton & Bradley (1982a)*
4.0×10^{-16}	2.2×10^{-15}	···	Hallén *et al.* (1990)†
7.2×10^{-15}	···	0.1×10^{-16}	Asghar *et al.* (1993)‡
$\sim 10^{-14}$	1.5×10^{-15}	···	Fretwurst *et al.* (1999)♮

† MeV H^+ and He^{2+} irradiated material; ‡ 5-MeV He^{2+} irradiated *p*-type junctions; ♮ Measurements on high-resistivity *n*-tye Si irradiated with MeV/GeV H^+ and MeV neutrons and pions;* Thermally activated capture-cross section for $(=/-)$: $\sigma_n = 4 \times 10^{-16} \exp(-0.017/k_BT)$ (cm^2).

some doubts on the assignment of the $(=/-)$ to V_2. Kimerling (1977) has suggested that this large capture cross-section, meaning a considerable delocalisation of the deep level, is probably due to the fact that the microscopic structure of V_2^- is quite different for V_2^{2-}, which might decrease the magnitude of the Coulombic force upon the electron defining the transition. A theoretical study by Lindfelt and Yong-Liang (1988)—a semi-self consistent method using a unrelaxed supercell of 686 atoms—has supported the idea of two different configurations for V_2^{2-}.

Following the pioneer work of Watkins and Corbett (1965) and posterior ENDOR studies of Sieverts and co-workers (1976; 1978a), it has been suggested that V_2 can be found in its JT-distorted state (C_{2h}) at low temperatures $T < 90$ K, and D_{3d} for higher temperatures. The latter is usually explained as a result of a *motionally averaged* state due to a high electronic jump rate between equivalent JT directions, not involving atomic motion.[1] This means the lifting of the Jahn-Teller distortion, resulting in an apparent increase of the symmetry of the defect's point group. Svensson *et al.* (1991) have argued that the V_2^{2-} charge state of the di-vacancy is only possible at higher temperatures at which the motional effects are expected to be strong.

If one looks at the results of generation vs. bombardment energies, annealing kinetics and depth concentration profiles in DLTS studies, it can be seen that for electron-irradiated material, there is a perfect 1:1 correlation between the concentrations of the $E_c - 0.23$ and $E_c - 0.42$ eV levels (Svensson *et al.*, 1991). Although, for specimens bombarded with α-particles or heavier ions, this picture changes drastically (Hallén *et al.*, 1990). The level

[1] This effect, known as electronic *bond-switching* (Watkins & Corbett, 1964) is usually monitored by the analysis of the variation of the spectral line-width with temperature (see §5.5.3).

$E_c - 0.23$ eV disappears in heavily bombarded areas of the crystal, with the $E_c - 0.42$ eV level assuming its highest concentration at the surface region.

Three IR absorption peaks are tentatively associated with the di-vacancy, at 0.69 (1.8 μm), 0.34 (3.6 μm) and 0.31 eV (3.9 μm), occurring for different charge states of the defect. The 0.69-eV peak has been described as arising from an internal transition between defect states in a neutral charge state of V_2 (Cheng & Vadja, 1969). Carton-Merlet *et al.* (1982), have suggested that the 0.34-eV absorption occurs when V_2 is in its single negatively charged state, V_2^-; Svensson and co-workers (1988) using optical interference filters, confirmed the latter assignment. Furthermore, this group has tentatively suggested that the 0.34-eV peak arises from an internal a_g–a_u transition, with the a_u lying close to the conduction band minimum. Despite this, one experimental fact remains to be explained. The 0.34-eV peak has appeared when the pseudo-Fermi level is well above $E_c - 0.23$ eV. If this electron trap acts like a double acceptor, the defect must be in the double positively charge state, which contradicts all previous experimental results.

7.3 Previous theoretical work

From a theoretical point of view, the modelling of the lattice di-vacancy in silicon is a remarkable and challenging task. The proof of this is the number of studies that can be found in the literature (Lee & McGill, 1973; Fazzio *et al.*, 1983; Humphreys *et al.*, 1983; Lindefelt & Yong-Liang, 1988; Lee & Corbett, 1973; Song *et al.*, 1993; Sugino & Oshiyama, 1990; Saito & Oshiyama, 1994; Smargiassi & Carr, 1996; Seong & Lewis, 1996; Pesola *et al.*, 1998).

Recently, a different model was suggested for V_2^- as a result of first-principles periodic pseudopotential LDA simulations (Saito & Oshiyama, 1994). Saito and Oshiyama argued that there is a gain in energy due to a different sense of the JT distortion as discussed by Watkins and Corbett (1965). As clarified by Watkins (1995), there is nothing new about this sense of the distortion.

C_{2h}	E	C_2	σ_h	i
a_g	1	1	1	1
b_g	1	−1	−1	1
a_u	1	1	−1	−1
b_u	1	−1	1	−1

Table 7.2: Character table for the C_{2h} point-symmetry group.

The system resulting from the coupling between doubly-degenerate ,electronic and vibrational lattice, modes ($e \times E$) giving rise to the Jahn-Teller effect for $V_2^{\pm}$ is well studied (Sturge, 1967). The surface of the ionic potential seen by the unpaired electron in two dimensional E-mode space, defines the so-called *mexican hat*. The electronic doubly-degenerate states transform in a similar way to the two normal modes, Q_1 and Q_2. These are the modes responsible for the $D_{3d} \to C_{2h}$ symmetry reduction in V_2 (Fig. 7.1). Within the harmonic approximation, the corresponding ionic Hamiltonian, can be written as

$$\mathcal{H} = -A \begin{bmatrix} -Q_2 & Q_1 \\ Q_1 & Q_2 \end{bmatrix} + \frac{1}{2}\,\Lambda\,(Q_1^2 + Q_2^2)\,, \qquad (7.3.1)$$

neglecting nuclear kinetic energy terms. In this case, both distortion 'directions' are equally probable. The orbital degeneracy is replaced by a vibrational degeneracy. It is only when anharmonic restoring terms are considered, that these two directions become inequivalent. Unfortunately, it is very difficult to account quantitatively for these higher-order terms.

The inclusion of anharmonic terms destroys the symmetry of the potential, leaving a threefold symmetry as required by the cubic symmetry of the Hamiltonian. As a result, the potential surface contains now three wells, separated by saddles, along the three equivalent distortion directions. The height of these saddles, although small compared with the JT energy, is considerably greater than $k_B T$ at low temperatures. Consequently, the system is expected to be found frozen into one these valleys, leading to a permanent distortion. The raising of the temperature, results in the electronic activation over the potential barrier that separates these valleys, and the system has then a time-averaged symmetry. This seems to explain well the findings of Watkins and Corbett (1965) and Sieverts (1978a), as described previously.

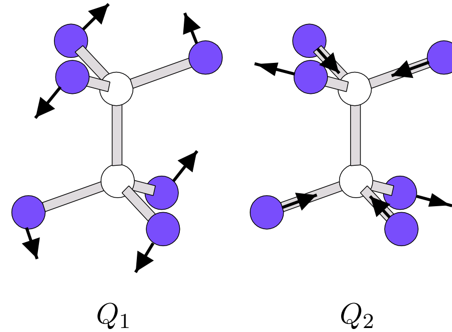


Figure 7.1: The two normal modes of the Jahn-Teller distortion (Q_1 and Q_2) responsible for the lowering of symmetry of V_2 from D_{3d} to C_{2h}.

For this alternative distortion, four of the dangling-bond atoms are expected to move outwards away from the vacant sites (Q_1 mode), as opposed to an inward movement (Q_2 mode), as previously proposed. Consequently, the six atoms constituting the divacancy become almost five-fold coordinated. In analogy with well known over-coordinated structures in chemistry, Saito and Oshiyama introduced the *resonant-bond* (RB) model.

These authors found the 'reconstruction-by-pairs' structure to be metastable. The energy required to change the sense of the JT distortion was calculated to be 2.4 meV, with the RB structure resulting in a $b_u^{\uparrow\downarrow}\, a_u^{\uparrow}$ configuration for V_2^-. As modern local-density functional theory is limited to an accuracy of 0.1–0.2 eV for an energy difference (Kohn, 1997), it is therefore not possible to argue based on energetics to favour either of the competing structures.

Since no reconstruction takes place between pairs of Si atoms according to the RB model, the resulting structure is much closer to that of the undistorted (D_{3d}) defect. Accordingly, no 'cross-over' between the a_g and b_g levels should be expected for such a modest distortion. Again, like in the case of the pairing distortion, the cross-over can only occur if the outward distortion is sufficiently large (Fig. 7.2). This is not obviously

the case, as it is evident from the reported structural and electronic configurations. The distances between the nearest-neighbouring Si atoms of the vacant atoms, labelled l_{ab}, l_{bc} and l_{ac}, were found to be 3.75, 3.75, 3.61 Å for V_2^+, and 3.60, 3.60 and 3.69 Å respectively (Table 7.5). Accordingly, for V_2^+ we have $l_{ab} = l_{bc} > l_{ac}$, and $l_{ab} = l_{bc} < l_{ac}$ for V_2^-, as required by the pairing and resonant-bond models, respectively.

This is also confirmed by a very small separation between the a_u and b_u singlet states of ~ 0.02 eV (Saito & Oshiyama, 1994)–which can be used to quantify the JT distortion, as pointed out by Watkins (1995).

Further evidence for a strong distortion was provided by ENDOR. De Witt *et al.* (1976) and Sieverts *et al.* (1978a) confirmed a considerable deviation from sp^3 for V_2^-. The same ENDOR experiments have also revealed some s character of the wavefunction (only 10% for V_2^- and 13% for V_2^+). This piece of evidence was used by Saito and Oshiyama to reject their RB model for V_2^+, which was found to provide a $b_u^{\uparrow}$ one-electron configuration.[2] For V_2^-, the highest occupied one-electron level can have either a_g or a_u symmetries.

It is interesting to note that Sugino and Oshiyama (1990), performing Green's-functions based calculations, reported a configuration $b_u^{\uparrow\downarrow}\, a_g^{\uparrow}$ for V_2^- arguing that a weak JT distortion occurs. This cannot be correct due to the reported admixture of *gerade* and *ungerade* levels for V_2^-, which can only occur in the case of a large distorion. In the latter, the simple valence-force model, combined with the results from the Green's functions calculation for the perfect crystal, was used to simulate the relaxation of a 32-atom cluster containing the di-vacancy in a semi-quantitative approach.

The model described above has been recently confirmed as a result of first-principles pseudopotential LDF/LDA theory based calculations (Seong & Lewis, 1996; Pesola *et al.*, 1998).

Seong and Lewis (1996) using a 64-atom unit cell to expand the Si lattice, with a kinetic-energy cut-off of 8 Ry, found the relaxed structure of V_2^0 to undergo a RB-like distortion, with $l_{ab} = l_{bc} = 3.40$ and $l_{ac} = 3.71$ Å.

Very recently, Pesola *et al.* (1998) have employed a similar method to that used by Seong and Lewis to study the structural and electrical properties of V_2. Using a fixed cut-off of 15-Ry, the defect was simulated embedded in supercells containing from 64 up to 216 atoms. Different k-point sets were used for the sampling of the Brillouin zone.

These workers found the resonant-bond configuration of V_2^- to be the stablest, while for the other two charge states, V_2^+ and V_2^0, the lowest-energy structures exhibited a pairing-like distortion. The result for V_2^0 does not provide any support to the conclusions of Seong and Lewis (1996). The symmetry of V_2^+ and V_2^- was found to be triclinic (S_2) and not monoclinic-I (C_{2h}) as found experimentally for the low temperature configuration. To explain their results, Pesola *et al.* invoked the lack of quantum zero-point energy motion—not included in the LSD calculations.

[2]The conclusion that only a_g- and b_u states, and not a_u and b_g, are compatible with an s content at the atoms on the C_{2h} reflection plane follows from inspection of the character table for the group C_{2h} (Table 7.2): only the a_g and b_u irreducible representations have even parity under reflection.

Additionally, Pesola *et al.* (1998), have also calculated the electrical levels associated with V_2, using the method described previous in §5.7.1. Their results will presented and discussed in a forthcoming section of this chapter.

7.4 Results

7.4.1 Cluster and basis

The calculations described here were performed using a trigonal, bond-centred, cluster containing 244 atoms ($Si_{148}H_{96}$). To allow the reconstruction of the Si–Si bonds, the cluster was pre-relaxed keeping the atoms in defect core fixed in the required configuration and allowing the relaxation of all the surrounding atoms. The resulting structure was then fully optimised, including hydrogen terminators. These were all symmetry constrained optimisations.

Table 7.3: Atomic basis and basis expansion type used for the calculations on V_2. N and M denoted the number of orbitals used to describe the wavefunction and charge density respectively.

Cluster	Atom	N	M	Basis expansion	`chden`
$Si_{148}H_{96}$	Si	4	5	24 big/122 minimal	0
	H	2	3	all minimal	
Bond-centres	all bonds	2	2	big	

In order to ensure the convergency of our approach with cluster size, the defect structures were also optimised embedded in a larger cluster, containing 346 atoms. The structures obtained, mainly the reconstruction for the paramagnetic charge states V_2^+ and V_2^-, were similar to those obtained using the 244 atom cluster. The surface of this larger cluster was kept fixed during relaxation.

The basis used is summarised in Table 7.3. Bond-centres, *i.e.*, additional Gaussian basis functions, with fixed exponents, were placed midway between each and every bonded pair of atoms, but excluding surface terminators.

7.4.2 The undistorted di-vacancy

For V_2^{2+}, there are four electrons to accommodate in the one-electron levels localised mainly on the dangling-bond atoms of the di-vacancy. Due to crystal field splitting, the six-fold degenerate level associated with the broken-bonds splits into two doublet and two singlet levels (Fig. 7.2), with the latter lying lower in energy. The symmetry and degeneracy of these levels is easily found from group theory by reducing the representation of the full octahedral group, O_h, into the irreducible representations of the point-symmetry group

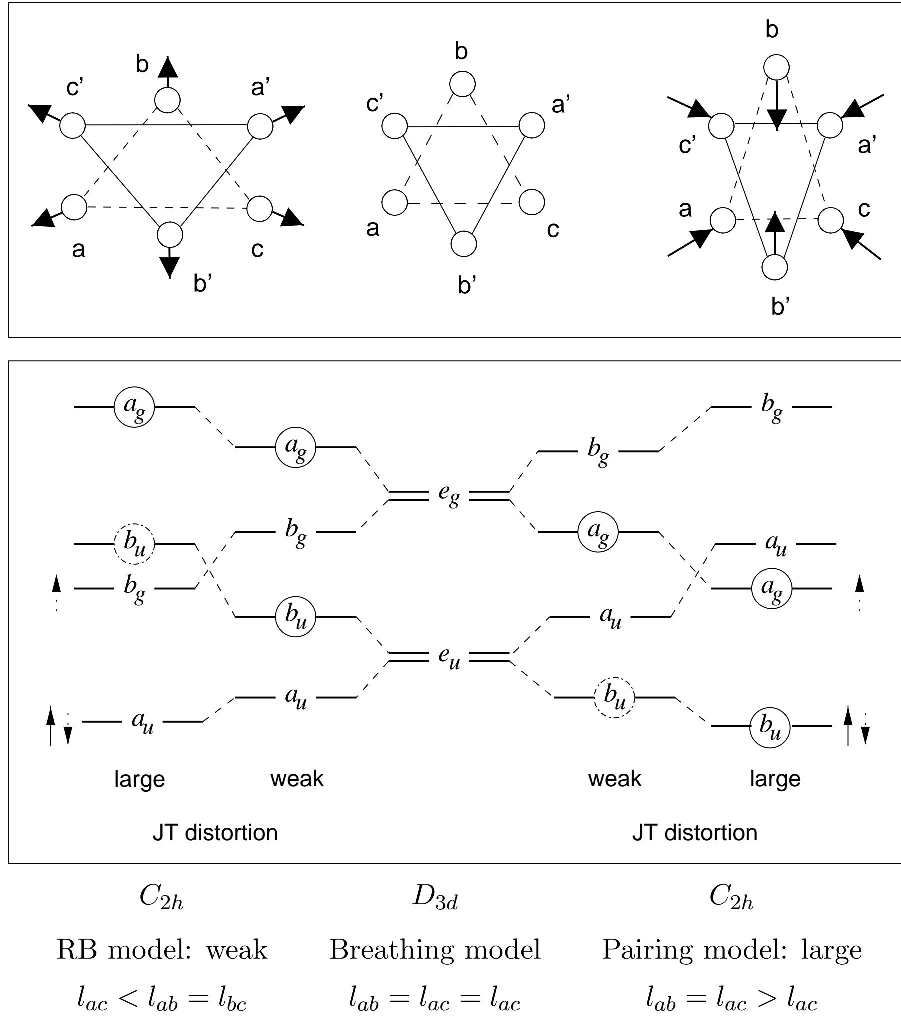


Figure 7.2: The effect of symmetry lowering for the two possible Jahn-Teller distortions on the one-electron levels of the undistorted di-vacancy. For clarity, circles indicate the orbitals with finite amplitude on the defect mirror plane. From the left, a small and large resonant-bonding (RB) *outward* distortions, the undistorted D_{3d} structure and a small and large pairing *inward* distortions, viewed along the [111] crystal direction. The solid arrows denote the electrons, and their spins for V_2^+. The dotted arrows denote the additional electrons for V_2^-.

D_{3d}. There is some ambiguity in the labelling of the resulting singlet levels, arising from the fact that the group D_{3d} can be generated by either $C_{3v} \times i$ or $D_3 \times i$. The singlet levels are labelled a_g and b_u for the former case, and a_{1u} and a_{2u} for the latter. We will adopt the a_u/b_g representation throughout this chapter. Obviously, an analysis based on group theory does not provide the ordering of the one-electron levels, *i.e.*, their relative energies.

For V_2^{2+}, the number of active electrons is only enough to completely fill the lower two singlets, leaving the two doublets unoccupied. Consequently, no symmetry-breaking distortions can occur. The six atoms of V_2 are then expected to relax outwards away from the the vacant sites, in a breathing motion that leaves the symmetry of the ideal di-vacancy (trigonal D_{3d}) unaltered.

The distance between the six Si atoms constituting the defect was calculated to be 3.75 Å, for the optimised structure of V_2^{2+}. This corresponded to a small increase of the open volume associated with the ideal di-vacancy. The Si were displaced ~ 0.11 Å from their positions in the ideal Si lattice, in a symmetric outward breathing motion. Concerning its electronic structure, only two one-electron levels were found to exist in the band gap. These are an e_g and an e_u doublets, anti-bonding combinations of the broken-bond orbitals, with the symmetric (g) lying above the anti-symmetric (u) manifold.[3] The other two singlet levels arising from the LCAO-molecular orbital picture, being occupied, are buried in the valence band. This one-electron level structure confirms the recent results of first-principles plane-wave based calculations of Saito and Oshiyama (1994) and Pesola *et al.* (1998), who have also found the two doublets to be the only gap levels in D_{3d} symmetry.

At this point, we note that the extended Hückel calculation of Lee and McGill (1973) found the e_u level to sink into the valence band for the undistorted di-vacancy. On the other hand, the cluster calculations of Kirton *et al.* (1984) the a_g level was located within the band gap. Only Humphreys *et al.* (1983), using a Green's functions method with a model defect potential, and Fazzio *et al.* (1983) employing a muffin-tin multiple-scattering method within the $X\alpha$-LDF approximation on unsaturated Si clusters of 20 atoms at an arbitrary value for the distortion co-ordinates, were able to correctly describe the electronic structure of undistorted di-vacancy.

7.4.3 The V_2^+ and V_2^- paramagnetic states

For V_2^+, there is now an electron occupying the lower band-gap doublet. The orbitally degenerate e-level is now partially occupied, which means that a symmetry breaking distortion is bound to occur.

To simulate the distorted structure of V_2^+, two configurations of of the defect were considered. For the first structure, the arrangement of the Si atoms of V_2^+ in C_{2h} symmetry prior to optimisation was such that $l_{ab} = l_{bc} = 3.75$ and $l_{ac} = 2.75$ Å, as prescribed by the pairing model of Watkins and Corbett (1965). The second candidate structure corresponded to the other sense of the distortion. Accordingly, the atoms were pulled away from the two vacant sites, with the distances between the Si atoms being $l_{ab} = l_{bc} = 3.15$ and $l_{ac} = 3.77$ Å. As it is evident, the initial structures for both models were strongly distorted, either due to an inward or outward movement of the defect atoms.

The relaxation of the first structure confirmed its stability, with the Si atoms defining the reconstructed bonds and the dangling-bond atoms moving slightly inwards, toward each other by ~ 0.2 Å. Its lowest energy structure was achieved for the relative Si distances of $l_{ab} = l_{bc} = 3.92$ and $l_{ac} = 2.94$ Å (Table 7.5). The inspection of the corresponding spin-polarised Kohn-Sham eigenvalue spectrum revealed a b_u level, occupied by a single electron, as the highest occupied one.

[3]By symmetric and anti-symmetric, we mean odd and even with respect to the rotations of $\frac{\pi}{3}$ followed by reflection around the defect main rotation axis (C_3).

The second structure was found to be unstable, relaxing back to a weakly-reconstructed pairing-like structure. The final distances between the Si atoms in the defect core were $l_{ab} = l_{bc} = 3.86$ and $l_{ac} = 3.68$ Å. It is interesting to note that these distances are quite similar to those of the optimised trigonal version of V_2^+, for which $l_{ab} = l_{bc} = l_{ac} = 3.77$ Å.

The (meta)stable atomic configuration that resulted from the relaxation of the second candidate structure—a weakly distorted structure of pairing character—was found to be 0.24 eV higher in energy than that exhibiting strong pairing. Nonetheless, it important to mention that both, strong and weakly, distorted structures give rise to an electronic configuration $b_u^{\uparrow}$. Consequently, both agree with the spin-density localisation found by Watkins and Corbett from the analysis of the G6 EPR spectrum (V_2^+). In conclusion, the strength of the JT distortion for V_2^+ is not an issue, in what concerns the localisation of the unpaired spin, as both reproduce the experimental findings. Despite this fact, the pairing structure characterised by a strongly reconstructed bond across two pairs of Si atoms, is favoured by an argument based on energetics.

We now present the results for the other paramagnetic state, V_2^-. In a similar fashion to V_2^+, the structure of V_2^- is also expected to undergo a JT distortion, which results in the lowering of the defect's symmetry from D_{3d} to C_{2h}. Despite the obvious similarities between the structural and electronic properties of the two paramagnetic charge states, it has been proposed that the sense of this distortion is different for V_2^- (Saito & Oshiyama, 1994). Based on the low energy barrier of $\sim$ 0.06 eV from one distortion to another reported by Watkins and Corbett, and that a strong reconstruction by paring "results in two isolated dangling-bonds with some frustration", Saito and Oshiyama introduced the resonant-bond model, as already mentioned in the introductory sections of this chapter.

Again, like in the case of V_2^+, the RB model structure was found to unstable against a weak, paring like, reconstruction, with $l_{ab} = l_{bc} = 3.59$ and $l_{ac} = 3.48$ Å. This structure provided an electronic configuration $b_u^{\uparrow\downarrow}\, a_u^{\uparrow}\, a_g^0$. In this case, the highest occupied level (a_u), does not satisfy the requirements set by EPR, as this level does not posses any appreciable magnitude on the defect's mirror plane (Fig. 7.3).

The lowest energy structure for V_2^- was found for the strong-pairing case. This structure was 0.36 eV lower in energy than that with weakly reconstructured bonds (see Ta-

Table 7.4: Total cluster energy difference (eV) for different configurations and charge states of V_2.

	Energy differences (eV)			
	V_2^{2-}	V_2^-	V_2^0	V_2^+
strong pairing	0	0	0	0
undistorted	+0.50	+0.83	+0.38	+0.24
weak pairing	+0.48	+0.48	+0.73	+0.57

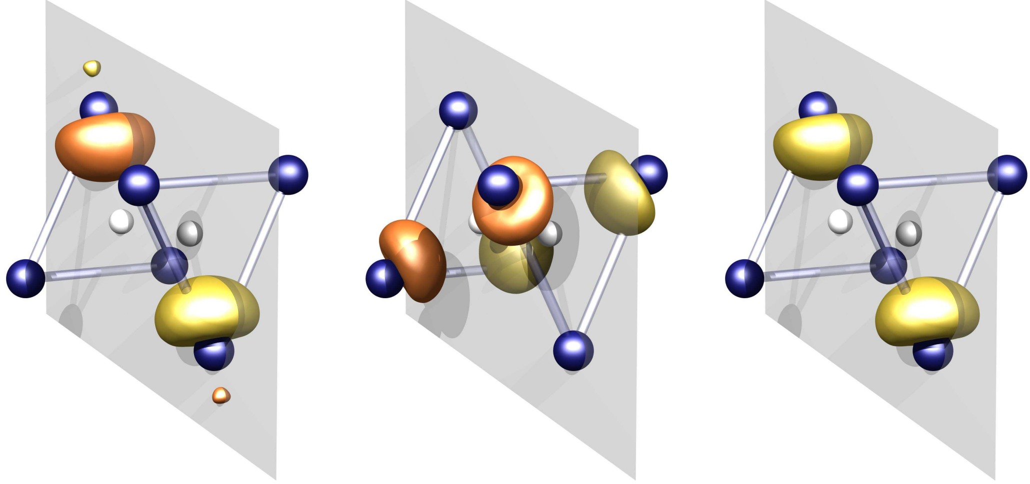
Table 7.5: Calculated geometries and distortion types for several charge states of V_2. The l_{ij} indicate the distance in Å between primed and unprimed i and j atoms neighbouring a vacant site. B, P, R and M indicate the type of distortion: breathing, pairing, resonant and mixed respectively.

	l_{ab}	l_{bc}	l_{ac}	Type	Symm.	References
V_2^{2+}	3.38	3.38	2.38	B	D_{3d}	Pesola *et al.* (1998)
	3.75	3.75	3.75	B	D_{3d}	This study
$V_2{}^{+}$	3.75	3.75	3.61	P	C_{2h}	Saito & Oshiyama (1994)
	3.46	3.41	2.99	P	$\sim C_{2h}$	Pesola *et al.* (1998)
	3.92	3.92	2.94	P	C_{2h}	This study
$V_2{}^{0}$	3.40	3.40	3.71	R	C_{2h}	Seong & Lewis (1996)
	3.45	3.26	2.89	M(P)	$\sim C_{2h}$	Pesola *et al.* (1998)
	3.78	3.78	2.92	P	C_{2h}	This study
$V_2{}^{-}$	3.60	3.60	3.69	R	C_{2h}	Saito & Oshiyama (1994)
	3.26	3.13	3.38	M(R)	$\sim C_{2h}$	Pesola *et al.* (1998)
	3.71	3.71	2.76	P	C_{2h}	This study
V_2^{2-}	3.23	3.26	2.84	B	$\sim D_{3d}$	Pesola *et al.* (1998)
	3.66	3.66	3.66	B	D_{3d}	This study
	3.62	3.62	2.68	P	C_{2h}	This study

ble 7.4. The length of the reconstructed bond across two pairs of Si atoms was calculated to be 2.76 Å, *i.e.*, 0.76 Å shorter than its counterpart in the weakly distorted structure of V_2^-. Relatively to the ideal crystal positions, the atoms participating in the reconstructions moved inwards towards the vacant sites by ~ 0.5 Å.

The main question that needs to be addressed concerns the strength of this strong(er) reconstruction. Is this strong enough to result in symmetry level-crossing? Let us look at the resulting one-electron configuration. This was found to be $b_u^{\uparrow\downarrow}\,a_g^{\uparrow}\,a_u^{0}$, which means that the answer is affirmative, as it can be seen from Figure 7.4.3. Both b_u and a_g states have magnitude on the σ_h reflection plane, in perfect agreement with experiment.

For the V_2^0 and V_2^{2-} charge states, our calculations have also confirmed the pairing model as the preferred one, with the distorted structures being more stable than corresponding D_{3d} versions, by 0.38 and 0.50 eV respectively (Table 7.4). The structural parameters for both structures are given in Table 7.5. It is interesting to compare the lengths of strongly-reconstructed bonds for both configurations of V_2^{2-}. According to our results, the reconstruction is strongest for V_2^{2-}. This can be understood within an one-electron picture. By further charging the defect, from V_2^+ and V_2^-, a spin compensated



(a) b_u symmetry: this wavefunction has, at least, a nodal surface on the plane containing four of the six Si atoms of V_2 and another between two Si atoms and their immediate back-bonded neighbours.

(b) a_u symmetry—highest occupied defect state: a nodal surface exists which is parallel to the plane shown and includes two of the Si neighbours of vacant lattice sites.

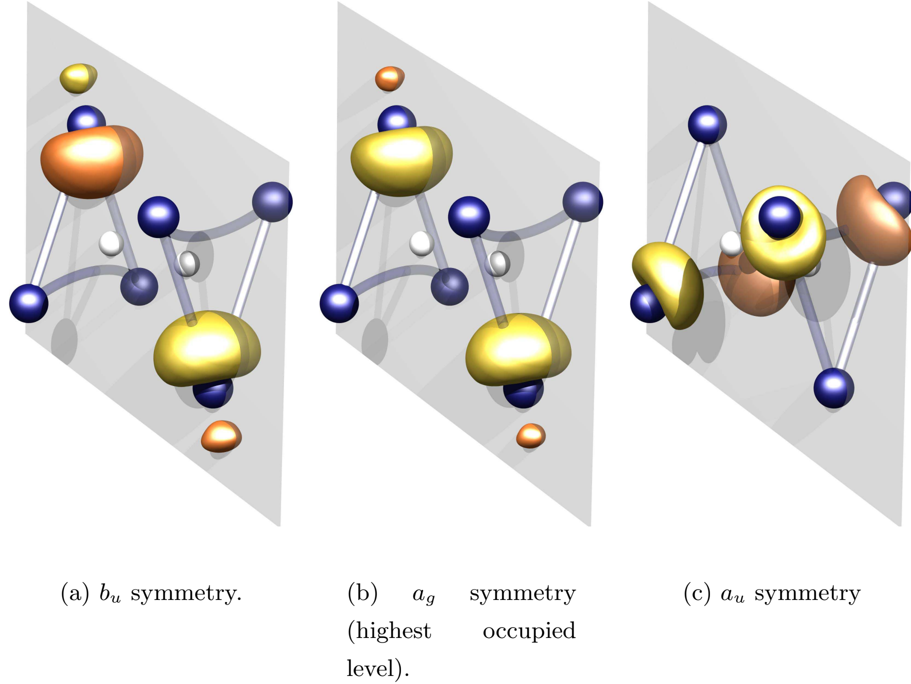
(c) a_g symmetry: the wavefunction for this level vanishes in the same regions as in the case of the b_u state (a); it has peaks between two Si neighbours and the corresponding vacant sites.

Figure 7.3: 3D-isosurface plot of the wavefunction corresponding to the two highest occupied and the lowest unoccupied spin-up Kohn-Sham levels of the weakly distorted structure of V_2^- (pairing like distortion). Note that no symmetry level cross-over occurs for such a distortion.

system is achieved with the extra electron heavily reducing the Coulombic interaction between the electron already occupying the 'broken-bond' orbital, and the electrons localised on the reconstructed bond.

The fact that the C_{2h} structure of V_2^{2-} was found to be slightly more stable than that with D_{3d}, it is not surprising. Both D_{3d} and C_{2h} configurations of V_2^{2-} are in principle possible. Let us examine how these can be achieved.

Starting from the neutral defect, and adding two electrons to the e_u level, we obtain a totally symmetric charge density with respect to the D_{3d} point-symmetry group, to achieve a $S = 0$ configuration. Consequently, the electronic force has the symmetry of the lattice, which implies that no symmetry-breaking forces act on the neighbouring atoms that might start a JT event. The other alternative mechanism, it is the capture of an electron by the defect when its distorted C_{2h} configuration. This also results in a totally symmetric charge density distribution, but now with respect to the group C_{2h}. It then natural to assume that the V_2^{2-} (D_{3d}) state can be accessed if the concentration of electrons is high enough to allow the capture of two electrons in a time interval much shorter than the lattice relaxation time. The other state, V_2^{2-} (C_{2h}), however, can appear whenever the



(a) b_u symmetry.

(b) a_g symmetry (highest occupied level).

(c) a_u symmetry

Figure 7.4: 3D-isosurface plot of the wavefunction corresponding to the two highest occupied and the lowest unocupied spin-up Kohn-Sham levels of distorted (C_{2h}) V_2^-. The 'banana'-like bonds illustrate the strong reconstruction between four of the six nearest neighbours of the vacant sites (white), in agreement with the pairing model proposed by Watkins and Corbett (1965).

defect, in its single minus charge state, captures an electron. Obviously, phonon-assisted transitions can occur between these two states, in which the V_2^{2-} (C_{2h}) configuration is likely to be lower in energy than that with D_{3d}, as suggested by a less restrictive symmetry in the former case, being less influenced, for example, when in damaged areas of heavily irradiated material.

7.4.4 Optical properties

As mentioned previously, several IR bands have associated with the di-vacancy. One of these bands, with a peak at 0.34 eV (3.6 μm) (Fan & Ramdas, 1959; Cheng *et al.*, 1966; Carton-Merlet *et al.*, 1982) is of particular importance. This band has been related to V_2^-. The photo-ionisation energy threshold of this IR spectrum was found to be very similar to that of the G7 EPR spectrum, of approximately 0.4 eV (Carton-Merlet *et al.*, 1979). This assignment was confirmed by Svensson *et al.* (1988), who performed FTIR measurements to monitor this particular absorption, with the aid of optical filters. As a result, and based on the theoretical work of Lindfelt and Yong-Liang (1988), this group suggested that this peak results from an $a_g \rightarrow a_u$ transition within the distorted V_2^- charge state. This absorption peak disappeared when the measuring temperature was raised above 90 K.

This intra-defect level transition can be calculated using the Slater's transition argument, as described in §2.8.2. Recalling (2.8.25), the corresponding energy can be calculated

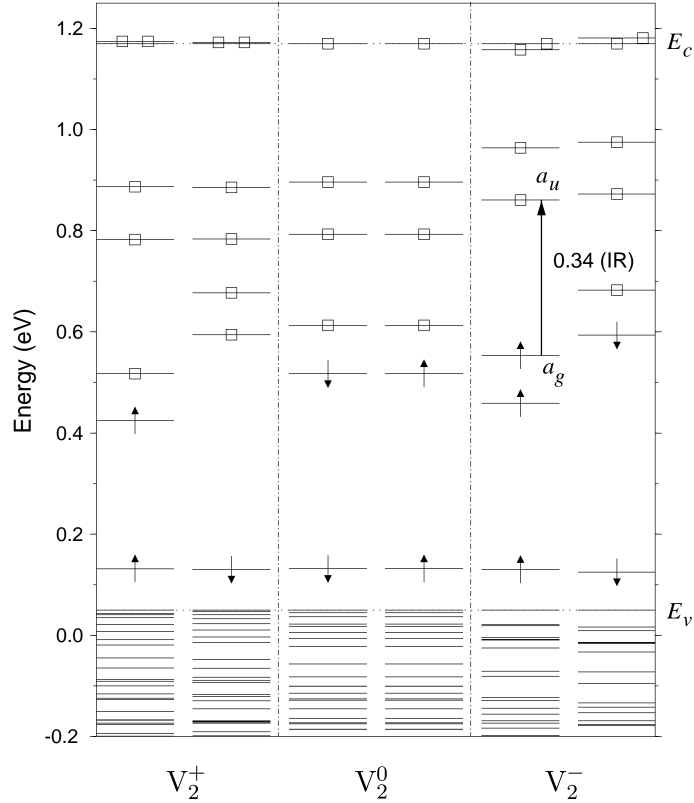


Figure 7.5: Scaled Kohn-Sham eigenvalue spectra for the lowest-energy structures of V_2 (C_{2h}) in three different charge states, in the vicinity of the band gap. Arrows and boxes indicate filled and empty levels respectively. The suggested a_g–a_u transition in V_2^- associated with the 0.34 eV IR peak (Svensson *et al.*, 1988) is also shown.

as

$$\Delta E_{a_g-a_u}(V_2^-) = \varepsilon_{a_u}(\vec{\gamma}_{ts}) - \varepsilon_{b_u}(\vec{\gamma}_{ts}) \; , \tag{7.4.1}$$

with

$$\vec{\gamma}_{ts} = \begin{cases} \gamma_{a_u} = +\dfrac{1}{2} \\ \gamma_{a_g} = -\dfrac{1}{2} \end{cases} \tag{7.4.2}$$

Here, ε_{a_u} and ε_{a_u} are the KS eigenvalues corresponding to highest occupied spin-up level and first, otherwise, unoccupied level of V_2^-, respectively (see Fig. 7.5). These are extracted from a single full optimisation run for the defect in the transition state (7.4.2), with a net proton charge of $-1\,e$. The self-consistent ε_{a_u} and ε_{a_u} eigenvalues were calculated to be -2.6260 and -2.2606 eV respectively. The transition is then simply given by the difference between these two eigenvalues, *i.e.*, ~ 0.36 eV, in excellent agreement with the observed transition.

7.4.5 Electrical levels

It is now accepted, as a result of photocondutivity (Kalma & Corelli, 1968; Young & Corelli, 1972), infrared absorption (Cheng *et al.*, 1966; Cheng & Vadja, 1969), and DLTS

(Evwaraye & Sun, 1976; Mooney *et al.*, 1977; Kimerling, 1977; Svensson & Willander, 1987; Svensson *et al.*, 1991) that V_2 can exist in four different charge states, as a defect in crystalline Si. As a result of electron and hole transitions between the several charge states of the defect, three levels exist in the bandgap. These have been assigned to the following transitions:

$$V_2^- + e^- \rightarrow V_2^{2-} \ : E_c - 0.23 \text{ eV} , \tag{7.4.3}$$

$$V_2^0 + e^- \rightarrow V_2^- \ : E_c - 0.42 \text{ eV} , \tag{7.4.4}$$

$$V_2^0 + h^+ \rightarrow V_2^+ \ : E_v + 0.31 \text{ eV} , \tag{7.4.5}$$

Calculation details

Using the relaxed structures from the full charge runs for V_2^{2-}, V_2^- and V_2^+ in both, D_{3d} and C_{2h} arrangements, the transition-state configurations were set as shown in Table 7.6. The underlined values indicate the one-electron levels with partial occupancy, *i.e.* $\frac{1}{2}e$ as required by Slater's transition argument. Note that for the D_{3d} structure, due to the degeneracy of the highest occupied level, the fractional charge was spread out equaly over the e-manifold levels in order to avoid self-consistency problems. Next, the clusters were fully relaxed to extract the required self-consistent eigenvalues. The net proton charge of the corresponding defect clusters was $-1\frac{1}{2}$, $-\frac{1}{2}$ and $+\frac{1}{2}$, respectively for the evaluation of the electron affinity of V_2^{2-}, V_2^- and the first ionisation potential of V_2^+.

Table 7.6: One-electron occupancies for the transition-state used to calculate the ionisation and affinity of the defect in two different configurations. The underlined values represent the occupancy of the highest occupied defect levels, in the transition-state, providing the defect's affinity and ionisation energies.

Config.	e_u occup. (D_{3d}) spin up	spin down	$b_u\,a_g$ occup. (C_{2h}) spin up	spin down	Cluster net charge	Run type
V_2^{2-}	$\underline{0.25}$ $\underline{0.25}$	0.5 0.5	1.0 $\underline{0.5}$	1.0 1.0	−1.5	affinity
V_2^-	$\underline{0.75}$ $\underline{0.75}$	0.5 0.5	1.0 $\underline{0.5}$	1.0 0.0	−0.5	affinity
V_2^+	$\underline{0.75}$ $\underline{0.75}$	0.0 0.0	$\underline{0.5}$ 0.0	0.0 0.0	+0.5	ionisation

Calculated electrical level structure

The calculated levels are given in Table 7.7. The single acceptor level of distorted V_2 was calculated to lie at $E_c - 0.37$ eV, which compares well the DLTS level at $E_c - 0.42$ eV. For the D_{3d} configuration, the calculated $(-/0)$ was found to lie deeper at $E_c - 0.52$ eV. This difference in the positioning of the $(-/0)$ level in D_{3d} and C_{2h} symmetries is expected, resulting from the splitting of the e_g doublet, which implies a downwards shift of the highest

occupied level energy and therefore a lowering value of the defect's affinity. Obviously, the calculated position of the (0/+) of V_2 should also reflect the fact the e-level is largely split for the C_{2h} configuration. This was confirmed as the donor level was located at $E_v + 0.61$ and $E_v + 0.41$ eV, for the D_{3d} and C_{2h} configurations of the defect, respectively. Interestingly, a similar value for the double acceptor level was found for both configurations of the defect. This result was confirmed with the use of a different marker for the calculation of the double acceptor. Despite a difference of $\sim$ 0.2 eV between the calculated value using the two different markers, PtH_2 and AuH, the calculations seem to indicate that the position of the (=/−) is independent of the degree of distortion of the defect.

Now, it is important to correlate these results with the available information on the electrical activity of V_2. As described previously, it is a fact, that at the DLTS scanning temperatures (90–200 K in the case of the two acceptor levels of V_2), the defect should exist in a symmetry-averaged state due to fast electronic bond switching. The jump rate between equivalent JT distortion directions is many orders of magnitude higher that the rate for electron emission from the traps at $E_c - 0.23$ and $E_c - 0.42$ eV. Svensson and Willander (1987) and Svensson *et al.* (1991) used this fact to explain the disappearance of the (=/−) level in heavily irradiated material. According to these authors, the lattice strain associated with 'highly' damaged regions in the crystal, quenches the electronic hoping motion and the V_2 remains in its C_{2h} configuration. However, under this assumption seems to be difficult to explain our results and the EPR results of Watkins and Corbett (1965), who observed the G7 spectrum (V_2^-) even when the Fermi level was above $E_c - 0.23$ eV, and the defect therefore should be in its V_2^{2-} charge state as indicated by DLTS. This can only be explained by invoking the instability of this charge state at low temperatures ($T < 80$ K). An alternative explaination was suggested by Lindfelt and Yong-Liang (1988), as mentioned previously. Depending on the electron emission process, these workers postulated that V_2^{2-} can exist in both D_{3d} and C_{2h} configurations. However, our calculations do not confirm the expected difference between the position of the (=/−) level for the two configurations as suggested by Svensson *et al.* (1988), that might explain the EPR results, *i.e.*, the observation of the G7 spectrum when the Fermi level is above $E_c - 0.23$ eV.

It is worthwhile noting that the reduction of the (=/−) activity of V_2 with increasing ion mass after irradiation and/or implantation is also accompanied by an increase of the concentration of the (−/0) level, together with a slight shift of the corresponding DLTS peak towards lower temperatures (Svensson *et al.*, 1991; Svensson *et al.*, 1997) This seems to suggest that another defect structure is being formed, as a result of the damage being inflected upon the crystal. Obviously, this different approach to this problem requires further investigation.

7.5 Summary

To conclude, we have successfully modelled the distortion required to explain the experimental observations for V_2. Calculations confirm the strong-reconstruction by-pairing

Table 7.7: Comparison between calculated and observed electrical levels of the lattice di-vacancy. The single and double acceptor levels and the single donor level are given with respect to E_v and E_c respectively (eV). The defects, and their relevant experimental levels, used as markers in the calculations are also given.

	Calculated				Observed
Level	Pesola†	marker	D_{3d}	C_{2h}	DLTS
(=/−)	0.73	PtH_2 (0.16)	0.38	0.39	0.23
		AuH_1 (0.19)	0.22	0.21	
(−/0)	0.78	C_i (0.10)	0.52	0.37	0.42
(0/+)	0.04	C_i (0.28)	0.61	0.41	0.25

† Pseudopotential self-consistent plane-wave DFT/LSD method with 216-atom supercells (Pesola *et al.*, 1998).

model of Watkins and Corbett for *both* positive and negative charge states. The localisation of the spin density found for V_2^+ and V_2^- and the calculated optical transition within V_2^-—a measure of the magnitude of the JT distortion—confirm the accuracy of the calculations presented here. The alternative resonant-bond model introduced by Saito and Oshiyama was found to be unstable for all the charge states investigated.

The calculated level structure does not support previous experimental explanations for the breakdown of the perfect 1:1 correlation between the concentrations of the DLTS peaks at $E_c - 0.24$ and $E_c - 0.42$ eV, assigned to the (=/−) and (−/0) levels of V_2 in ion-irradiated and implanted material.

In addition, the results of our calculations do not agree with those from recently published work. This relates to the sense of the distortion for both paramagnetic states of V_2, as well its electrical level structure. An example of the latter, is the work by Pesola *et al.* (1998), who used a plane-wave method with large supercell to describe the electrical properties of this defect by *ab initio* methods.

Common to all the previous calculations on the lattice di-vacancy is the fact that all were performed using plane wave-based codes with periodic boundary conditions. It is now clear to us that in order to model the occurring strong reconstruction within our cluster approach, its necessary to allow the H-surface to relax to achieve a proper description of the defect's potential energy surface. Unfortunately, the volume of the supercells used to describe the crystal in all the plane-wave methods cannot be allowed to vary within the relaxation process. Another problem inherent to this type of calculation is that it is necessary to achieve a reasonable compromise between the value of the kinetic energy cut-of, *i.e.*, the number of plane waves, and the size of the supercell.

8

Vacancy–Hydrogen Defects

8.1 Introduction

Hydrogenation is usually thought of as a process to neutralise the optical and electrical activity of shallow and deep-level defects in silicon (Pankove *et al.*, 1983; Johnson *et al.*, 1986; Pearton *et al.*, 1992; Estreicher, 1995). It can be also used to 'decorate' various defects, thereby facilitating their detection and characterisation. This is the case for the lattice mono-vacancy. With an extraordinary low barrier for diffusion, it is very difficult to detect this defect as an isolated species. It is a highly distorted centre (Jahn-Teller), with the spin-singlet states V^{2+}, V^0 and V^- having T_d, D_{2d} and C_{2v} symmetry respectively (Watkins & Troxell, 1980a; Newton *et al.*, 1983). It is a 'negative-U' centre (Baraff *et al.*, 1979b; Baraff *et al.*, 1979a), due to the metastability of the singly positively charged state, with the double donor level lying above the single donor level. The single and double acceptor levels have never been directly observed (Watkins, 1992).

These remarkable properties are only matched by its ability to form complexes with dopants and impurities, and with itself (Watkins & Corbett, 1965; Lee & Corbett, 1973; Lee & Corbett, 1974).

It is now a well known fact that the single vacancy can be 'decorated' with up to 4 hydrogen atoms (Roberson & Estreicher, 1994; Bech Nielsen *et al.*, 1995; Bech Nielsen *et al.*, 1996; Stallinga *et al.*, 1998). Apparently, the interaction with hydrogen it is not sufficient to immobilise the single vacancy. This was first reported by Corbett *et al.* (1988), to explain SIMS profiles of H and D in sub-surface regions of Si, and latter confirmed by Sopori *et al.* (1992). Sopori *et al.*, monitoring the *I-V* characteristics of ion-irradiated photo-voltaic devices, suggested that hydrogen was able to reach and passivate defects in the active regions of the devices much faster in vacancy-rich material than in high-quality (vacancy-poor) material.

Proton implantation at cryogenic temperatures has been successfully performed to create multi-hydrogen–vacancy complexes (Bech Nielsen *et al.*, 1995; Bech Nielsen *et al.*, 1996). The ability of H to form strong covalent bonds with Si atoms, which are stronger than Si–Si bonds, is well known from chemistry. It is then only natural that H will attack 'dangling-bond' atoms whenever possible in order to stabilise the defect structure and

therefore reduce its total energy.

8.2 Background

The identification of VH^0 by means of electron paramagnetic resonance (EPR) measurements was recently reported by Bech Nielsen *et al.* (1997). Recorded at $\sim$ 45 K, the EPR spectrum was assigned to a paramagnetic $S = \frac{1}{2}$ centre with monoclinic-I (C_{1h}) symmetry. The comparison between the spectra recorded from proton- and deuteron-implanted samples revealed the involvement of hydrogen in the centre. Furthermore, the changes of the hyperfine satellites with the applied field confirmed the presence of a single ^{29}Si nucleus. The distance between the unpaired electron and the proton was estimated to be $\sim$ 2.7 Å from the anisotropic contribution to the hyperfine tensor $\overleftrightarrow{A}$, within a point-dipole approximation. Bech Nielsen and co-workers also pointed out the striking similarities between the spin-Hamiltonian parameters of this newly recorded EPR spectrum and that of the *E*-centre (Watkins & Corbett, 1964). The analysis of the variation of the spectral linewidth with temperature revealed broadening and motional narrowing effects, with the centre becoming trigonal at higher temperatures ($\gtrsim$ 60 K) with an activation energy $E_a \simeq 0.06$ eV. As in the case of the *E*-centre, this was interpreted as an electronic 'hopping' between the three equivalent Jahn-Teller directions in the double degenerate electronic system.

Fourier-transform infra-red (FTIR) absorption spectroscopy is a powerful technique to detect local vibrational modes (LVMs) arising from defects involving light impurities, like hydrogen. Several IR bands have been assigned to VH^0. Infrared absorption measurements by Xie *et al.* (1991) on H_2-grown neutron-doped Si by radioactive transmutation, and independently Meng on neutron-irradiated float-zone (FZ) silicon (Meng, 1991), reported a IR band at 1839 cm^{-1} as being due to VH. This assignment is inconsistent with recent IR studies by Holbech *et al.* (1993), which identified the line at 1839 cm^{-1} as a LVM of the so-called H_2^* defect (Fig. 8.1). Bech Nielsen *et al.* reported a different value for the Si–H stretch mode frequency of VH at 2067.5 cm^{-1}, identified as arising from a centre with C_{1h} symmetry, as a result of uniaxial stress measurements (Bech Nielsen *et al.*, 1996; Bech Nielsen *et al.*, 1995). The assignment to VH^0 was tentative, simply based on previous theoretical calculations (Deák *et al.*, 1989; Estreicher, 1995). Very recently, Stallinga *et al.* have reported a new value for the H-stretch frequency of VH^0 as a result of combined EPR/FTIR annealing measurements on proton-implanted Si samples, which showed that the EPR signal of VH^0 and the mode at 2038 cm^{-1} both disappear at $\sim$ 480 K (Stallinga *et al.*, 1998). The corresponding hydrogen- and deuterium-related vibrational bands were located, respectively, at 2038.5 and 1494.6 cm^{-1}. The mode previously attributed to VH at $\sim$ 2068 cm^{-1} is now assigned to a neutral complex between hydrogen and the lattice di-vacancy containing a single H atom (V_2H^0) (Stallinga *et al.*, 1998).

Two modes at around 2121 and 2144 cm^{-1} have been identified as two H-stretch frequencies of the same defect—a di-hydrogen–vacancy defect, VH_2 (Bech Nielsen *et al.*, 1996). These two modes have shown an identical annealing behaviour and the same inten-

sity ratio in samples implanted with different proton doses (Bech Nielsen *et al.*, 1995). A single new mode appeared at 2134.2 cm^{-1} in samples containing both H and D (deuterium), revealing the involvement of two equivalent H atoms in the defect. The corresponding decoupled D-stretching mode was found at 1556.8 cm^{-1}

Uniaxial stress measurements combined with theoretical modelling, determined the symmetry of the defect to be C_{2v} (Bech Nielsen *et al.*, 1995). In addition, Budde (1998) has shown that the bands at 1987 and 1990 cm^{-1}, previously attributed to VH_2 (Xie *et al.*, 1991; Meng, 1991) originate from a defect with monoclinic-II (C_2) symmetry and not orthorhombic-I (C_{2v}). This author has also identified two bands at 1547 and 1564 cm^{-1}, as the deuterium counterpart of VH_2, by implanting deuterons instead of protons and analysing their relative annealing behaviour and intensity.

Due to the fact that VH_2 is non-magnetic ($S = 0$) in its neutral charge state, it has never been observed by electronic resonance techniques. One way to detect paramagnetic activity from VH_2 is to excite optically the defect, by promoting an electron to the first excited state within the neutral state to achieve a spin-triplet state ($S = 1$). According to Chen *et al.* such activity has been detected in Cz-grown, oxygen-rich, material subsequently irradiated with 2-MeV electrons, analysed by ODMR (Chen *et al.*, 1990). Based on its hyperfine structure, the ODMR spectrum was described as an excited state $S = 1$ of VH_2 with orthorhombic-I (C_{2v}) symmetry. Recently, Stallinga and Bech Nielsen (1998), noting the similarities between the reported spectrum and that $S = 1$ excited state of the A-centre (VO^*) (Brower, 1971b), and unresolved hyperfine interactions, have refuted this assignment. In reply to this comment, Chen *et al.* (1998) argued that the observed weak hyperfine interaction provides the strongest support for their assignment, as it allows one to distinguish between two, expectedly, similar EPR spectra, *i.e.*, that of VO^* and $S = 1$ VH_2, only distinguishable due to a different behaviour under thermal annealing.

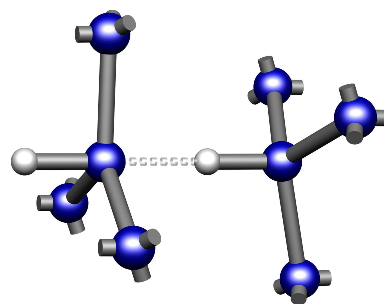


Figure 8.1: The H_2^* defect: the H atoms at the near 'bond-centred' (BC) and anti-bonding (AB) sites in the lattice. The line joining the two H atoms defines the [111] direction (Holbech *et al.*, 1993).

The identification of the local vibrational modes of VH_3^0 has proven to be a difficult task (Budde, 1998). Two pairs of H-stretch frequencies have been measured and attributed to a trigonal defect, resulting from splitting and frequency shifts induced by applied uniaxial stress (Bech Nielsen & Grimmeiss, 1989; Bech Nielsen *et al.*, 1996) These are the pairs {2155, 2185} and {2166, 2191} cm^{-1} (Bech Nielsen *et al.*, 1995; Bech Nielsen *et al.*, 1996). The 2166-cm^{-1} line corresponds to an E mode and the 2191-cm^{-1} to an A_1 mode, as predicted for a trigonal centre. Despite the fact that theory was able to shown that the lowest-energy structure of VH_3 is achieved for C_{3v} symmetry (Bech Nielsen *et al.*, 1995; Deák *et al.*, 1991; Xu, 1992; Roberson & Estreicher, 1994; Park *et al.*, 1995), experiment

has not being able to differentiate between the defects responsible for the two pairs of bands. From an experimental point of view, there is obviously more than one possible structural model which satisfies the requirement of trigonal symmetry. These are the cases of interstitial SiH_3 (C_{3v}), V_2H_6 (D_{3d}) and obviously VH_3 (C_{3v}).

The lines 2166 and 2191-cm^{-1} disappear upon annealing at $\sim$ 800 K being therefore related to one of the most stable hydrogen-related complex in proton-implanted Si. Deák *et al.* (1991) showed that the structure with a single Si self-interstitial plus three isolated H atoms is more stable that the *i*-SiH_3 configuration. This provides strong evidence against the assignment of the {2166, 2191}-cm^{-1} bands to SiH_3. Furthermore, infra-red measurements (Cardona, 1983) on molecular disilane (Si_2H_6) have show that order of the E and A_1 modes is the reverse of that found for the {2166, 2191}-cm^{-1} bands. Very recently, Hourahine *et al.* (1999) calculated the LVMs associated with different charge states of four possible configurations of interstitial silyl (*i*-SiH_3) by *ab initio* methods (AIMPRO). These configurations corresponded to the possible arrangements of the H atoms with the self-interstitial at T_d or hexagonal interstitial sites (Fig. 8.2). The calculated vibrational frequencies differed from the those at {2166, 2191} by more than 300 cm^{-1} for all the structures but one. Although with modes closer to experiment at 2317 (A_1) and 2162-cm^{-1} (E), this structure (Fig. 8.2-d) was found to $\sim$ 4.2 eV higher than the energy of the stablest configuration.

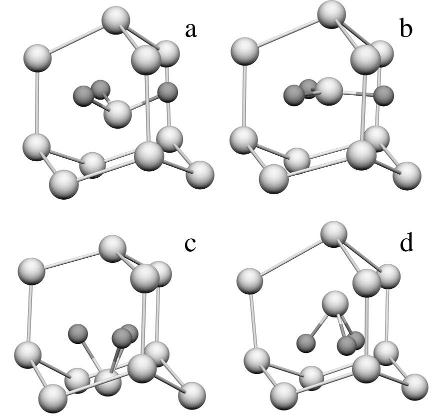


Figure 8.2: Relaxed structures for the neutral form of the four distinct C_{3v} configurations of *interstitial* SiH_3 (Hourahine *et al.*, 1999).

The recent observation of the paramagnetic $S = \frac{1}{2}$ state of neutral VH_3 has brought a whole new perspective to this problem (Herstrøm, 1998). The components of the associated g-tensor, consistent with a centre with trigonal symmetry, revealed all the characteristics of a dangling-bond orbital in a vacancy-like defect. Furthermore, the EPR signal exhibited strong hyperfine interactions with three equivalent H atoms, together with a strong trigonal hyperfine interaction with a single ^{29}Si nucleus. Consequently, this EPR centre was attributed to VH_3. The comparison of the annealing behaviour of the EPR and FTIR centres, under isochronal annealing, supports the assignment of the absorption lines {2155, 2185}-cm^{-1} to VH_3^0; the EPR signal disappears at $\sim$ 490 K, while the bands {2166, 2191}-cm^{-1} are stable up to 800 K (Table 8.1).

If the absorption bands {2155, 2185}-cm^{-1} arise from VH_3, it reasonable to assume that the other pair of bands at 2166 and 2191 cm^{-1} are the H-modes of the hydrogen-saturated di-vacancy (V_2H_6). With trigonal D_{3d} symmetry, this defect is expected to give rise to four H-modes with A_{1g}, E_g, A_{2u} and E_u symmetry. Experimentally, only two modes can be observed: the IR-active A_{2u} and E_u modes. If V_2H_6 is considered to be made from

two weakly-interacting C_{3v} VH_3 units, the bands at {2166, 2191}-cm^{-1} can well be its IR active H-stretch modes. This argument simply assumes that the coupling between H atoms is larger for V_2H_6, despite the existence of a second vacant lattice site nearby, resulting in a upward shift compared to isolated VH_3.

Beside a different behaviour under annealing, the defect structures responsible the two pairs of IR bands should also reveal distinct electrical properties. While the former is bound to be electrically active, the latter is electrically inactive. A study of the sensitivity of the LVMs at 2166 and 2191 cm^{-1} to bandgap light revealed that this defect has no levels in the gap as no dependence whatsoever was detected (Bech Nielsen, 1999).

Like for the case of VH_3, the identification of the vibrational modes of VH_4 is still matter of some discussion. Early IR absorption studies (Bai *et al.*, 1985) on Si grown in H_2- and D_2-atmospheres, found that a line at around 2222 cm^{-1} was consistent with a tetrahedral defect containing four H atoms bonded to Si. The IR peak at $\sim$ 2222 cm^{-1}, observed at RT, was observed to shift to 2210 cm^{-1} at 10 K. At this low temperature, the peak width was reduced and a fine structure appeared, with the main peak splitting into three separate peaks. The ratios of the intensity of these peaks remained always proportional to the natural abundance of the three isotopes of Si, under isochronal annealing treatments. A shift of $\sim \sqrt{2}$ [$\approx (m_{\rm H}/m_{\rm D})^{1/2}$] was observed, establishing that this is a H-related mode. This is the expected frequency shift due to the isotopic substitution of H by D, as noted by Stein (1975).

This line was also observed in proton implanted samples (Mukashev *et al.*, 1985). The assignment of T_d symmetry to the defect responsible for this line was subsequently confirmed by the uniaxial stress studies of Bech Nielsen *et al.* (1989). IR studies have also shown that this line anneals out at $\sim 525\,^\circ$C – a characteristic of a very stable defect (Bech Nielsen *et al.*, 1995). Although, confirming the symmetry of the defect giving rise to the 2222-cm^{-1} band, it was not possible to determine whether VH_4, or i-SiH_4 as suggested by Shi *et al.* (1985), is responsible for this degenerate vibrational 2221.9-cm^{-1} mode.

However, the assignment of the 2222-cm^{-1} mode to VH_4 has been recently questioned by Suezawa (1997; 1998a). IR measurements on Si grown in a H-atmosphere have revealed that the integrated intensity of the absorption of this line increases when carbon, boron or even gold are present in the crystal. These impurities are assumed to create Si self-interstitials (I), via the so-called *kick-out* mechanism. This can happen directly in the case of C, and indirectly for B and Au, which are known to lock vacancies, therefore increasing the number of interstititials by reducing the probability of V–I pair annihilation events. Although being consistent with the idea that the 2222-cm^{-1} line is due to i-SiH_4, this observation does not provide irrefutable evidence.

The growth in the intensity of the 2222-cm^{-1} band at $\sim 180\,^\circ$C has also been correlated (Suezawa, 1998b; Suezawa, 1998c) with a decrease in the intensity of an optical absorption band at 0.34 eV usually associated with V_2^- (Cheng *et al.*, 1966; Svensson *et al.*, 1988). Again, this observation might suggest an interstititial-like defect as responsible for the 2222-cm^{-1} band. Nevertheless, it is important to remember that the di-vacancy is itself a

Table 8.1: Relative thermal stability of the several neutral vacancy-hydrogen defects. $T_{\rm ann}$ is the temperature at which the defects anneal out (K) and the vibrational frequencies are given in cm^{-1}.

Defect	$T_{\rm ann}$ (K)	Techniques	
VH	~ 480	FTIR (2038)	EPR
V_2H	~ 530	FTIR (2073)	EPR
VH_2	~ 475	FTIR (2121–2145)	
VH_3	490–500	FTIR (2155–2185)‡	EPR
V_2H_6	~ 900	FTIR (2166–2191)†‡	
VH_4	~ 850	FTIR (2223)	

† The pair of vibrational bands at 2191 and 2166 cm^{-1} is formed at ~ 560 K, attaining its maximum amplitude at ~ 670 K to anneal out at ~ 800 K (Budde, 1998). ‡ Tentative experimental assignment.

reservoir of single vacancies, which can then be trapped by hydrogen to form VH_4.

Apparently, theory seems to support the assignment of the 2222-cm^{-1} line to VH_4 (Deák *et al.*, 1991; Bech Nielsen *et al.*, 1995). The T_2 H-stretch frequency was calculated by both groups to lie at ~ 110 cm^{-1} above the experimental value. The Raman active (IR inactive) A_1 singlet mode of VH_4 has been predicted to lie *above* the T_2 at 2222-cm^{-1}. Further support to this model has been provided recently by *ab initio* calculations (Hourahine *et al.*, 1998) on the vibrational modes of silane (SiH_4) in and outside the Si lattice. The H-frequencies of trapped silane were found to lie at ~ 1800 cm^{-1}. This provides further evidence to the assignment of the 2222-cm^{-1} to VH_4.

8.2.1 Electrical properties

Very little is known about the electrical activity of vacancy-hydrogen defects in silicon. The involvement of hydrogen is usually indirectly confirmed from comparative DLTS studies on implanted samples with protons (H^+) and α-particles (He^{2+}).

Two electron traps at $E_c - 0.32$ and $E_c - 0.45$ eV have been assigned to defects involving hydrogen formed as a result its the interaction with radiation induced defects (Irmscher *et al.*, 1984; Svensson *et al.*, 1989; Hallén *et al.*, 1990; Lalita *et al.*, 1997; Feklisova & Yarykin, 1997).

Irmscher *et al.* (1984) have correlated the trap at $E_c - 0.32$ eV to another level detected in p-type Si at $E_v + 0.28$ eV, as a result of DLTS measurements on low-dose H^+- and α-implanted material. These were identified as a single acceptor and single donor level of the same H-containing defect—a vacancy-hydrogen complex containing two hydrogen atoms, with their similar concentrations changing simultaneously under isochronal annealing. No

comments on the possible involvement of oxygen in the centre were made.

Svensson *et al.* (1989) have also detected these two traps in H^+-irradiated (RT) diodes. These two defects were not formed in α-irradiated material. The corresponding capture cross-sections were 3×10^{-15} and 2×10^{-17} cm^2, respectively, for the $E_c - 0.32$ and $E_c - 0.45$ eV DLTS peaks. The production rate of these two peaks decreases at higher doses, especially that of the $E_c - 0.32$ eV trap. This might suggest that low-order complexes are responsible for these two traps, disappearing for higher doses as a result of interaction with radiation induced defects, or other impurities that become mobile at RT. The $E_c - 0.32$ eV-trap was tentatively attributed to partially hydrogenated, VO or V_2 defects.

A similar DLTS study by Hállen *et al.* (1990) has confirmed the conclusions of Svensson *et al.*. The fact that the depth profiles for these two traps have showed a smaller skewness compared to the other main peaks (VO$(-/0)$, $V_2(=/-)$ and $V_2(-/0)$), resembling the simulated profile of implanted hydrogen,[1] provided further confirmation of the presence of hydrogen in the $E_c - 0.32$ and $E_c - 0.45$ eV defect structures.

These two traps were also produced as a result of an annealing treatment at 250 °C for 30 min in nitrogen ambient of self-implanted Si material at RT (Lalita *et al.*, 1997). The observed capture cross-section of $\sim 10^{-15}$ cm^2 for the $E_c - 0.32$ eV DLTS peak, suggests that this level is indeed similar to that reported by Svensson *et al.* (1989). These authors have also reported that the concentration of this trap growths at the expense of the $(=/-)$ level of V_2 at $E_c - 0.23$ eV during annealing. This observation was explained as resulting from the interaction of previously trapped hydrogen, which becomes mobile at ~ 150 °C, with the lattice di-vacancy. Consequently, these authors proposed a V_2H defect as responsible for the $E_c - 0.32$ eV level.

Even assuming that the above assignments are correct, it is not possible to determine the number of hydrogen impurity atoms involved. While it is generally accepted for the case of VOH defects that only the one containing a single H atom is electrically active, for the case of V_2H there is no obvious reason why multi-H di-vacancy complexes are not active as well. Recently, Feklisova and Yarykin (1997), have proposed that the $E_c - 0.32$ eV trap arises from a VOH defect containing not one *but* two hydrogen atoms. The corresponding capture cross-section was 1.0×10^{-15} cm^2, measured in e-irradiated samples in which hydrogen was incorporated by wet chemical etching. Recent theoretical (Jones *et al.*, 1999a) and experimental studies (Johannesen *et al.*, 1999; Bonde Nielsen *et al.*, 1999) do not seem to confirm the hypothesis of the electrical activity of VOH_2.

Bruni *et al.* (1994) have tentatively assigned two hole traps to VH_2 and VH_3. The two DLTS signals, arising from two levels at $E_v + 0.67$ and $E_v + 0.33$ eV respectively, were observed in high-dose (1.6×10^{16} cm^{-2}) H^+-implanted B-doped silicon. The $E_v + 0.33$ eV level appeared after annealing at 100 °C, whereas that at $E_v + 0.67$ eV was already present in as-implanted samples. The concentration of these traps was strongly reduced at 300 °C, increasing for $T > 300$ °C, which was explained with the formation of the electrically inac-

[1] The depth profile simulations were performed using the Monte Carlo code TRIM-89 (Developed at IBM, Yorktown, NY by J. P. Biersack and J. F. Ziegler).

tive VH_4 defects at $T < 300\,^{\circ}C$, and by the release of H from previously H-saturated VH complexes ($T \gtrsim 400\,^{\circ}C$), inducing partial vacancy depletion, and consequently restoring their electrical activity. Further annealing at $800\,^{\circ}C$ resulted in the complete neutralisation of the electrical activity of two defects originating the hole traps.

The results of the annealing studies of Bruni *et al.* (1994) were later confirmed by combined positron annihilation/DLTS measurements on H-implanted Si (Brusa *et al.*, 1994).

Very recently, Bonde Nielsen *et al.* (1999), and independently, Peaker *et al.* (1999) have performed DLTS and LDLTS measurements on hydrogenated (by H^+-implantation for the former and wet etching, after *e*-irradiation, for the latter study) Si. According to Bonde Nielsen *et al.*, the activation enthalpies of the $(-/0)$ of VH and that of V_2H are 0.443 and ~ 0.43 eV, respectively. These two Laplace DLTS peaks were not observed in α-implanted samples, which provides strong evidence for the involvement of atomic hydrogen in the electrically active defects. Furthermore, the annealing behaviour of these traps showed a perfect correlation with EPR annealing data for VH (Bech Nielsen *et al.*, 1997) and V_2H (Stallinga *et al.*, 1998). Both H^+ and He^{2+} implantations were performed at low temperature (60 K), followed by a reverse-bias anneal (RBA) at 400 K.

It is interesting to note that the reported values for the activation energy of these two traps is virtually indistinguishable from those of the $(-/0)$ of V_2 and the *E*-centre. Additionally, the value reported for VH is quite identical to that the level at $E_c - 0.45$ eV reported by Svensson *et al.* (1989), and tentatively attributed to VO or possibly V_2.

8.3 Previous theoretical studies

Despite the number of theoretical studies that can be found in the literature on multi-hydrogen–vacancy complexes in crystalline silicon (Singh *et al.*, 1977; Pickett, 1981; DiVincenzo *et al.*, 1983; Grekhov *et al.*, 1983; DeLeo *et al.*, 1984; Frolov & Mukashev, 1988; Bai *et al.*, 1988; Gutsev & Myakenkaya, 1989; Roberson & Estreicher, 1994), only a few authors performed total energy calculations, allowing for lattice relaxation and distortion effects.

Deák *et al.* (1989; 1991) applied a semi-empirical HF method of modified intermediate neglect of differential overlap (MINDO/3) in cyclic clusters made of unit cells containing 31 host atoms. VH was found to have C_{1h} symmetry (very close to trigonal C_{3v}), 1.499 Å, C_{2v} for VH_2, C_{3v} for VH_3 and VH_4 tetrahedral T_d. The corresponding Si–H lengths were 1.499 Å for VH, 1.489 Å for VH_2, 1.485 Å for VH_3, and finally for the fully H-saturated vacancy, VH_4 this length was 1.490 Å. The associated stretching modes were calculated to be, after scaling,[2] 2057, 2080, 2106 and 2109 cm^{-1}.

Park *et al.* (1995) used a first-principles spin-averaged (non-self-consistent), pseudo-atomic-orbital molecular dynamics method ($T = 0$ K). The optimised Si–H lengths were

[2]The scaling was performed in such a way that theory reproduces a known vibrational frequency such as that of silane.

Table 8.2: Comparison between calculated and observed local vibrational modes of multi-hydrogen–vacancy defects (cm^{-1}).

	Symm.	Mode	Deák [a]	Park [b]	Jones [c]	Expt.
VH_4	T_d	A_1	2364	2334	2404	N-IR
		T_2	2347		2319	2222
VH_3	C_{3v}	A_1	2360	2301	2318	2185
		E	2347		2256	2155
VH_2	C_{2v}	A_1	2331	2268	2316	2144
		B_1	2327		2267	2121
VH	C_{1h}	A'	2305	2168	2248	2038

[a] Deák *et al.* (1989); [b] Park *et al.* (1995); [c] Jones *et al.* (1995).

1.508, 1.504, 1.498 and 1.475 Å, for neutral VH, VH_2, VH_3 and VH_4 respectively. VH was found to have trigonal C_{3v} symmetry. The length of the Si–Si bond in the plane perpendicular to the Si–H bond was calculated to be 3.880 Å, revealing a very weak reconstruction. This is very similar to the Si–Si bond length of 3.929 Å, reported in the same paper for the neutral mono-vacancy. To explain the low symmetry obtained, these authors argued that, since the calculations did not account for spin-polarisation effects, the dangling-bond could never be 'occupied' by a single electron, resulting in a structure very close to that VH^+, which was already known to be trigonal (Deák *et al.*, 1989; Deák *et al.*, 1991).

Jones *et al.* (1995) employed pseudopotential local density functional cluster theory (AIMPRO) on a cluster derived from a 71-atom tetrahedral cluster, ($Si_{35}H_{36}$), representing bulk Si. All atoms, including H terminators, were allowed to relax.

These calculations confirmed that the symmetry is C_{1h} for VH and C_{3v} for VH^+. Accordingly, the one-electron eigenvalue spectrum for VH revealed the splitting of the highest occupied e-manifold into two singlet levels, due to reconstruction of the bond across two of the 'dangling' Si atoms. The Si–H length was 1.511 Å with the bond direction departing 5° from the [111] crystal direction. The vibrational frequency of H was calculated to be 2293 cm^{-1}, increasing 11 cm^{-1} if anharmonicity is considered.

The lowest energy structure for VH_2 was found to have C_{2v} symmetry with the Si–H lengths equal to 1.505 Å. The H atoms moved 0.1 Å away from the [111] alignment, which can be explained as resulting from H-H interactions. The bond between the other two Si neighbours of the vacancy was 2.555 Å, being 1.2 Å shorter than that reported by Roberson and Estreicher (1994) using HF techniques. The separation between symmetric and anti-symmetric H-stretch modes was calculated to be 49 cm^{-1}, with the former mode lying higher at 2316 cm^{-1}. The fact that the calculations neglected any coupling between stretch and other modes, *e.g.* bend modes at $\sim$ 800 cm^{-1}, was invoked to explain such a

large separation.

The structure of VH_3 was analysed in the neutral and singly positively charged states. The symmetry of VH_3 and VH_3^+ was found to be the same, *i.e.* C_{3v}, with the Si–H length varying very little with charge state, being 1.514 Å for the neutral species. Again, like in the case of VH_2 the separation between the A_1 and E modes was overestimated by a factor of 2.5. The quasi-harmonic frequencies of the singly- and doubly-degenerate modes were 2283 and 2209 cm^{-1} respectively. VH_3 possesses a fully occupied e level lying above a singlet level a_1 in the gap, in fair agreement with the semi-empirical tight-binding simulations of Xu (1992).

This study confirmed a tetrahedral symmetry for VH_4, with the four equivalent Si–H bonds calculated to be 1.489 Å long. The volume expansion of the vacancy cage was found to be 67% from that of the ideal vacancy. All the experimental trends for the H and D modes of this vacancy-hydrogen complex were well reproduced, with the absolute frequencies being $\sim$ 4% too high. Confirming previous theoretical studies (Deák *et al.*, 1989), the IR inactive A_1 mode was calculated to lie above the triply-degenerate mode, at 2222 cm^{-1}. The electrical inactivity of VH_4 was confirmed, as no KS levels were found in the Si band gap by these authors.

Despite all the different theoretical techniques employed up to date to study the structure of vacancy–hydrogen defects in Si, the results seem to agree on the following points:

(*i*) The interaction of atomic hydrogen with the lattice mono-vacancy leads to the formation of strong covalent Si–H bonds, with the hydrogen atoms pointing approximately towards to the vacant site;

(*ii*) The Si–H bond lengths get shorter as the number of hydrogen atoms in the vacancy increases, with the vibrational frequencies increasing correspondingly;

(*iii*) The IR peak at 2222 cm^{-1} should be assigned to VH_4 rather than i-SiH_4;

(*iv*) With the exception of VH_4, all VH_n complexes with $n < 4$ are expected to be electrically active.

The work presented in this chapter is directly related to point (*iv*). The present aimed the study of the electrical properties of vacancy–hydrogen defects in crystalline silicon. In order to guarantee the quality of the structures, the present included a detailed investigation of the structural properties of the relevant defects. Furthermore, results on the vibrational properties of di-vacancy–hydrogen defects, namely V_2H and V_2H_6, are also described.

8.4 Results

8.4.1 Cluster and basis

The calculations were performed on a tetrahedral cluster containing $(296 + n)$ atoms $(H_nSi_{180}H_{116})$, with $n = 1\ldots4$, centred at the vacant site. Further calculations involv-

ing the di-vacancy were performed using larger trigonal, bond-centred, clusters containing $(244+n)$ and $(345+n)$ atoms in the configurations $H_nSi_{146}H_{96}$ and $H_nSi_{212}H_{144}$ respectively.

Table 8.3: Atomic basis and basis expansion type used for the calculations on vacancy-hydrogen defects. N and M denoted the number of orbitals used to describe the wavefunction and charge density respectively.

Cluster	Atom	N	M	Basis expansion	`chden`
$Si_{180}H_{116}$	Si	5	4	all big	
	H	3	2	all minimal	0
Bond-centres	all bonds	2	2	big	
$Si_{148}H_{96}$	Si	5	4	24 big/122 minimal	
	H	3	2	all minimal	0
Bond-centres	all bonds	2	2	big	

In order to allow for the expected reconstruction in VH and VH_2, similar pre-relaxations to those described previously for the study of V_2, with the exception that only two shells of atoms surrounding the first neighbours of the vacany were allowed to move. The length of the reconstructed bond across two Si bordering the vacancy was set to 3.36 Å, for both VH and VH_2. No pre-relaxation was performed for the study of VH_3 and VH_4. For the undistorted C_{3v} structure of VH, this length was set to 4.14 Å.

All these defects structures were then relaxed, keeping the hydrogen atoms in the surface fixed. The results of the calculations are presented and discussed next.

8.4.2 VH and V_2H

In its undistorted configuration, VH consists of a Si–H bond and three Si dangling-bonds pointing towards the vacant site. Since the latter are symmetrically equivalent and can thus be symmetrised according to the irreducible representations of the point-symmetry group C_{3v}, an a_1 and e symmetric states are expected to appear within the bandgap region (Fig. 8.3). The s orbital of hydrogen interacts strongly with one of the Si hybrid orbitals of the otherwise isolated vacancy, giving bonding and anti-bonding a_1 states with large separation. The formation of the Si–H bond results from the occupation of two electrons on the a_1 bonding state, which interacts weakly with the remaining three dangling-bond orbitals. The e gap manifold of C_{3v} VH stems from these three Si dangling bonds, therefore being very close to the t_2 triplet gap state of the isolated lattice vacancy.

It is evident that the VH defect in its neutral charge state should undergo a JT distortion due to fact that the e gap state (four-fold degenerate, including spin) is occupied

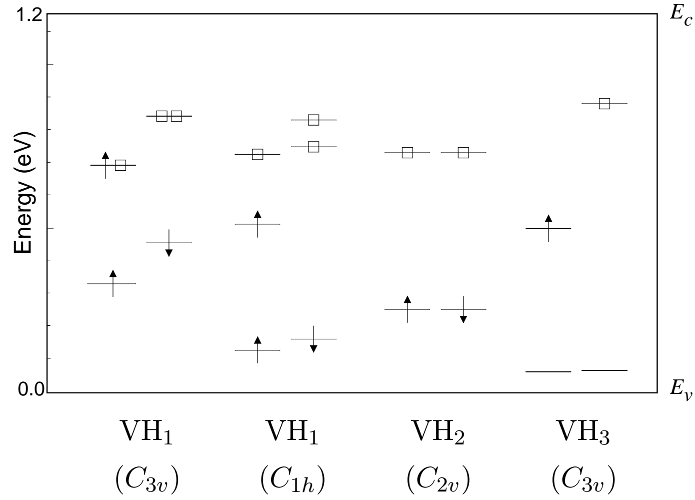


Figure 8.3: Scaled spin-polarised Kohn-Sham eigenvalue spectra of optimised VH_n defects. Those for the C_{3v} and C_{1h} forms of VH^0 are also shown for sake of comparison.

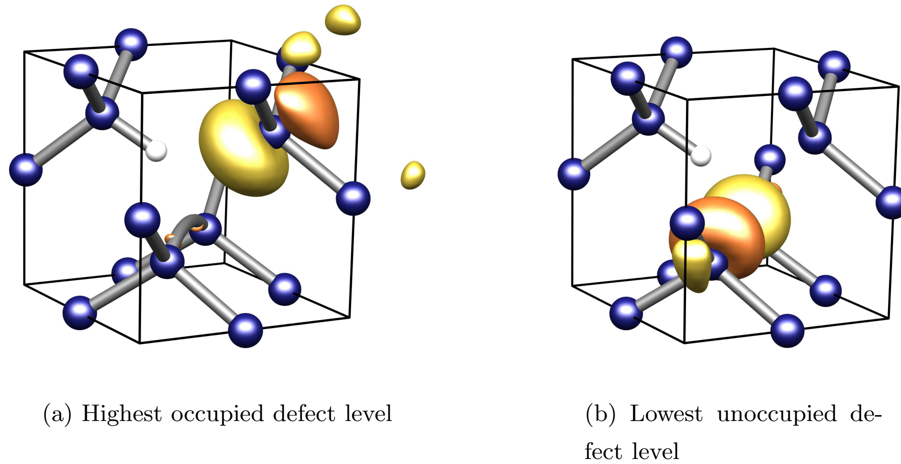
by a single electron (Fig. 8.3). A possible distortion corresponds to movement of Si atoms inwards to form a pair-bonding state, in a similar way to the reconstruction that occurs for V_2, with the remaining one moving away from the centre of the vacancy cage. This distortion will lower the defect symmetry from C_{3v} to C_{1h}, causing the splitting of the e according to $e = a' + a''$, with the unpaired electron occupying the a' singlet for VH^0 (see Fig. 8.3)

The optimisation of the C_{3v} and C_{1h} configurations of VH^0 has revealed the distorted structure to be marginally lower in energy than the C_{3v} counterpart, in close agreement with the previous theoretical studies (Deák *et al.*, 1989; Deák *et al.*, 1991; Bech Nielsen *et al.*, 1995). The corresponding spin-polarised KS eigenvalue spectra for two configuration of VH^0 are shown in Figure 8.3. The optimised structure of (C_{1h}) VH^0 can be seen in Figure 8.5, in two different views. The optimum Si–H length was found to be 1.501 Å,

Table 8.4: Structural parameters of the optimised neutral vacancy–hydrogen defects containing up to four hydrogen atoms (Å).

Complex	Symmetry	Si–H	Si–Si	H–Si †	H–H
VH	C_{1h}	1.501	3.299	2.784	···
VH_2	C_{2v}	1.477	3.249	···	1.733
VH_3	C_{3v}	1.499	···	2.870	1.867
VH_4	T_d	1.491	···	···	1.834

† Distance between the hydrogen atom(s) and the dangling-bond Si atoms in the defect core.



(a) Highest occupied defect level

(b) Lowest unoccupied defect level

Figure 8.4: 3D-isosurface plot of the wavefunction corresponding to the highest occupied (a' symmetry) and the lowest unoccupied (a'' symmetry) spin-up Kohn-Sham levels of C_{1h} VH^0.

with the bond direction making an angle of $\sim 5°$ with the [111] crystal direction, along the defect's mirror plane, which compares well with the observed value of 3° (Bech Nielsen *et al.*, 1997). The distance between the hydrogen atom and the dangling-bond Si atom was found to 2.784 Å, which confirms the EPR point-dipole estimative of $\sim$ 2.7 Å (Bech Nielsen *et al.*, 1997). For the optimised structure of C_{3v} VH^0, the distance between the three equivalent Si atoms was 4.07 Å (Table 8.4).

The barrier for the jump of the H atom between equivalent $\langle 111 \rangle$ directions for C_{1h} VH^0 was calculated adiabatically to be 0.35 eV. This value must be considered an upper limit for this barrier, as expected from a LDF-based method (Artacho & Falicov, 1991). Experimentally, the barrier for the jump of the danging bond between equivalent configurations was found to be 60 meV (Bech Nielsen *et al.*, 1997). This strongly points to phonon-assisted, or even pure quantum, tunnelling as the mechanism responsible for such a low, thermally activated, barrier for hoping.

The quasi-harmonic H-stretch frequency arising from the Si-H unit was calculated to 2126.3 cm^{-1}, shifting to 1528.5 cm^{-1} when hydrogen was replaced by deuterium. The relative error associated with this calculation with respect to the corresponding experimental values is 4.3 and 2.3%. This corresponds to a considerable improvement on the calculated LVMs compared with previous calculations (Table 8.2). Using a similar methodology to that used in the calculations described in this thesis (AIMPRO) but smaller clusters, Jones *et al.* (Bech Nielsen *et al.*, 1995), calculated the H- and D-stretch modes to lie at 2247.6 and 1612.6 cm^{-1}, respectively, which implies an error of 10.2 and 7.9%.

The $S1$ EPR centre

As pointed out by Stallinga *et al.* (1998), the electronic properties of VH^0 and V$_2$H^0 should indeed be similar. Based on the similarities between the EPR spectrum of VH^0 and the $S1_a$ EPR spectrum, arising from a defect with a $S = \frac{1}{2}$ ground state with C_{1h} symmetry

(Lütgemeier & Schnitzke, 1967; Gorelkinskiĭ & Nevinnyi, 1991; Kleinhenz *et al.*, 1979; Stallinga *et al.*, 1998), these authors assigned the latter to a neutral divacancy-hydrogen defect, containing a single hydrogen atom (V_2H^0).

To model V_2H, we have used the lowest-energy C_{2h} structure of V_2. One hydrogen atom was then attached to one of the dangling-Si atoms in such a way that the symmetry of defect became C_{1h}, *i.e.*, with the Si–H bond direction along the mirror plane of V_2. This structure was then fully optimised. Details of the cluster and basis used are given in Table 8.3.

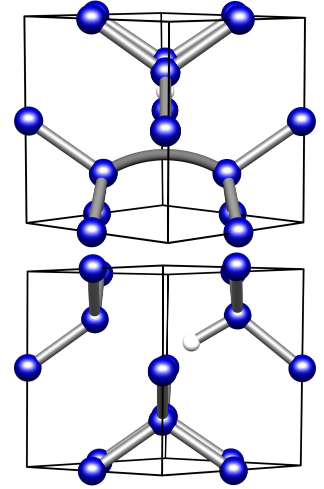


Figure 8.5: Two views of the lowest-energy structure of VH^0 (C_{1h}). The cube axes indicate the ⟨100⟩ crystal directions.

The calculated vibrational frequency of hydrogen was found to increase by ~ 35 cm^{-1} when this impurity was incorporated into the larger, second-order, vacancy defect to form V_2H^0. This value is in excellent agreement with the reported shift of ~ 30 cm^{-1}. The Si–H length was 1.50 Å. Due to the presence of H, the reconstructed Si–Si bonds are no longer equivalent, with that closer to the H atom being 2.86 Å long, and the other one, 2.82 Å. Before optimisation, the length of these reconstructed bonds was the same and equal to 2.83 Å. The frequency shifts due to the isotopic substitution of hydrogen by deuterium are also found in good agreement with the observed ones (Table 8.5).

It is interesting to note that despite the fact that the calculated and observed frequencies show an increase of the H-stretch frequency when going from VH^0 to V_2H^0, the Si–H bond is longer in V_2H than in VH. Even a small variation of ~ 0.006 Å of the Si–H length, should be more than enough to cause a 30 cm^{-1} shift. The main difference between VH and V_2H, is that the Coulombic interaction between the unpaired electron localised at the dangling bond of V_2H and the reconstructed bond is much weaker than that for VH. One way to test the influence of the dangling bonds on the calculated LVMs is to passivate one of the six dangling bonds of undistorted V_2 (D_{3d}), with hydrogen. In the optimised structure of V_2H, with $\sim D_{3d}$ symmetry, the proton lay 2.95 Å away from two nearest two dangling-bond Si atoms, for a Si–Si distance of 3.95 Å. The calculated H-stretch frequency was found to shift downwards by ~ 182 cm^{-1}, for a Si–H length of 1.53 Å. Unfortunately, the overall effect of the dangling bonds on the vibrational frequency of hydrogen seem to be masked by that arising from a release of the strain on the Si–H bond due to outward movement of the broken-bond atoms, which represent the main reason for the frequency drop from VH to V_2H.

8.4.3 VH_2

In the VH_2 defect, two of the four dangling bonds are saturated by hydrogen impurity atoms. This structure is now invariant under the operations of the point-symmetry

Table 8.5: Comparison between the calculated and observed H-stretch frequencies of VH_1 and V_2H (cm^{-1}). Experimental values taken from Stallinga *et al.* (1998).

Defect	Symm.	Mode	Calc.	Isot. shift	Obs.	Isot. shift
VH	C_{1h}	A'	2126.3	597.8	2038.5	543.9
VD			1528.5		1494.6	
V_2H	C_{1h}	A'	2161.6	606.7	2068.1	560.5
V_2D			1554.9		1507.6	

group C_{2v}. Consequently, no JT distortion can occur since this group can only have one-dimensional irreducible representations.

This point group contains two reflection operations, σ_v and $\sigma_{v'}$, with respect to a mirror plane containing the two Si–H units and another containing the non-hydrogenated Si atoms, respectively.

The relaxation of C_{2v} VH_2 resulted in a one-electron configuration of $a_1^{\uparrow\downarrow}\,b_2^0$. As expected, these orbitals were found to be localised on the 'dangling-bond' atoms, which are pulled together to achieve a distance of 3.37 Å after relaxation. In order to explain this reconstruction for VH_2, one has to consider the formation of this defect as resulting from

Table 8.6: Variation of the calculated (quasi-harmonic) hydrogen and deuterium local vibrational modes of VH_3 with charge state (cm^{-1}).

		Net proton charge	
Config.	Mode	0	–
VH_3	A_1	2175.4	2190.2
	E	2153.9	2164.1
VH_2D	A'	2168.3	2181.6
	A''	2153.9	2164.1
	A'	1588.4	1562.0
VHD_2	A'	2161.1	2172.9
	A'	1557.2	1567.3
	A''	1549.9	1556.8
VD_3	A_1	1561.4	1572.7
	E	1549.0	1556.8

the capture of an hydrogen atom by an already distorted, C_{1h} VH^0 centre. Note that the reconstructed-bond length is 0.07 Å longer that the corresponding bond in VH^0. The other four electrons are accommodated in the valence band a_1 and b_1 (resonant) states formed from the hydrogen s orbitals and the H-saturated Si hybrids. The two equivalent H atoms lay 1.49 Å away from their Si neighbours and 1.80 Å from each other.

The A_1 and B_1 vibrational modes of VH_2 are calculated to lie at 2211.2 and 2204.7 cm^{-1}, respectively. The calculated A_1–E separation between symmetric and anti-symmetric modes of 6.5 cm^{-1} is in fair agreement with the observed value of 22.5 cm^{-1}. The symmetric H-D mixed mode was found at 2207.9 cm^{-1}, being $\sim$ 620 cm^{-1} above the anti-symmetric mode, in excellent agreement with the experimental value of $\sim$ 500 cm^{-1}.

The $S = 1$ excited state of VH_2

Calculations on the neutral spin-triplet ($S = 1$) state of VH_2 have confirmed the model proposed by Chen *et al.* (1990). The lowest energy structure was found to have C_{2v} symmetry. The possibility of a lower symmetry structure was also investigated. None of the several distorted structures considered were found to be unstable, with the defect evolving to an orthorhombic (C_{2v}) arrangement as a result of relaxation.

8.4.4 VH_3 *versus* V_2H_6

As it can be seen from Figure 8.3, two KS levels appear in the bandgap for neutral C_{3v} VH_3. These are a fully occupied e-manifold with four electrons (including spin), and a higher singlet level (a_1), occupied by a single spin-up electron. Consequently, this configuration leads to a paramagnetic spin-doublet state for the neutral defect. The e-level originates from the three H-saturated sp^3 hybrid orbitals, while the a_1 singlet arises from the remaining dangling-bond. This is obvious from a three-dimensional plot of the wavefunction corresponding to this level—the highest occupied KS orbital for the optimised structure of VH_3^0 (Fig. 8.6).

The symmetry of VH_3^0 is confirmed to be C_{3v}, with the three equivalent Si–H bonds having a length of 1.50 Å. The relative distance between hydrogen atoms was 1.87 Å.

The quasi-harmonic hydrogen-related LVMs are calculated to be 2175.4 (A_1) and 2153.9 cm^{-1} (E). These shift to 1561.0 and 1549.9 cm^{-1}, respectively, when hydrogen is replaced by deuterium. The mixed H-D stretch frequencies are given in Table 8.8. All the calculated modes are within 0.1% of the experimental ones.

As mentioned previously, so far the only direct experimental observation that has allowed a distinction between the two pairs of bands in question, {2155,2185} and {2166,2191}-cm^{-1}, relates to their relative stability. It is known from EPR that VH_3 anneals out at $\sim$ 490 K, while the {2166,2191}-cm^{-1} modes only appear at $T \gtrsim 560$ K annealing out at $\sim$ 800 K. Consequently, and assuming that the disappearance of the EPR-signal is not due to a Fermi-level effect, it is obvious that {2166,2191}-cm^{-1} modes cannot originate from VH_3. It seems very unlikely that the ceasing of the EPR activity is due to a $VH_3^0 \longrightarrow VH_3^-$ transition since this would required the Fermi level to be close to the

Table 8.7: Calculated H-stretch frequencies (cm^{-1}) of neutral VH_3 and V_2H_6 defects. A cluster of 348 atoms, centred at a bond, was used to extract the LVMs of V_2H_6.

Defect	Sym.	Mode	Calc.	Obs.	Defect	Sym.	Mode	Calc.	Obs.
VH_3	C_{3v}	A_1	2175	2185	V_2H_6	D_{3d}	A_{2u}	2113	2191
		E	2154	2155			E_u	2098	2166
							A_{1g}	2089	N-IR
							E_g	2084	N-IR

conduction band edge. This is unlikely as the Fermi level is probably being pinned by the $(-/0)$ of V_2 at $E_c - 0.42$ eV, which is known to be a prominent trap in H^+-implanted material existing in large concentrations (Bonde Nielsen *et al.*, 1999).

Our calculations indicate that VH_3 is only stable in the neutral and singly negatively charged state. We predict a single acceptor at E_c−0.27 eV, but no donor level. A detailed account of the results on the electrical properties of vacancy-hydrogen defects will given later in this chapter. The calculated H- and D-related modes for these two charge states are given in Table 8.6. It can be seen that the calculated frequencies increase by ~ 10 cm^{-1} for all H and D modes, as a result of the change in the defect's charge state.

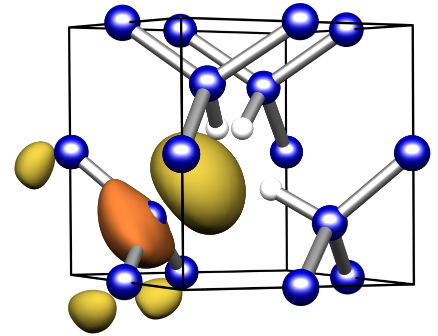


Figure 8.6: 3D-isosurface plot of the wavefunction corresponding to the highest occupied spin-up Kohn-Sham level of VH_3 (see Fig. 8.3). Note the localization on the dangling-bond atom.

The structure used for the study of V_2H_6 was optimised in a 348-atom cluster, keeping the hydrogen surface fixed. After relaxation, the length of the six equivalent Si–H bond in D_{3d} symmetry was 1.50 Å. This value is identical to that for VH_3 up to the third decimal place. The shortest distance between H atoms in different VH_3 units is 3.37 Å, with the distance between H atoms in the same unit being 1.97 Å. Compared to VH_3, the hydrogen atoms are now further away form each other by ~ 0.1 Å. The two A_{2u} and E_u IR-active modes for V_2H_6 are 2113 and 2098 cm^{-1}.

8.4.5 VH_4

For VH_4, the a_1 and t_2 one-electron levels arising from the four silicon hybrids surrounding the vacancy are now fully occupied, resulting in the formation of four covalent Si–H bonds. Consequently, no levels should exist in the bandgap, which results in the electrical neutralisation of the electrical activity of the lattice monovacancy.

Confirming previous studies (Park *et al.*, 1995; Bech Nielsen *et al.*, 1995), the KS

Table 8.8: A summary of the calculated and observed hydrogen and deuterium stretch frequencies for multi-hydrogen–vacany defects in silicon (cm^{-1}).

	Mode	Calc.	Obs.
VH_4	A_1	2251.7	N-IR
	T_2	2239.8	2221.9
VH_3D	A_1	2248.8	2250.4
	E	2239.8	2223.5
	A_1	1610.9	1620.3
VH_2D_2	A_1	2245.8	2243.5
	B_1	2239.8	2225.4
	A_1	1611.8	1628.3
	B_2	1610.1	1614.6
VHD_3	A_1	2242.8	2236.1
	A_1	1612.2	1636.1
	E	1610.1	1615.6
VD_4	A_1	1613.7	N-IR
	T_2	1610.1	1616.1

	Mode	Calc.	Obs.
VH_3	A_1	2175.4	2185
	E	2153.9	2155
VH_2D	A'	2168.3	···
	A''	2153.9	···
	A'	1588.4	···
VHD_2	A'	2161.1	···
	A'	1557.2	···
	A''	1549.9	···
VD_3	A_1	1561.4	···
	E	1549.0	···

	Mode	Calc.	Obs.
VH	A'	2126.3	2038.5
VD	A'	1528.5	1494.6

	Mode	Calc.	Obs.
VH_2	A_1	2211.2	2143.8
	B_1	2204.7	2121.3
VHD	A'	2207.9	2134.2
	A'	1586.9	1554.5
VD_2	A_1	1590.0	1564.0
	B_1	1583.9	1547.0

eigenvalue spectrum of the lowest-energy T_d VH_4 structure exhibited no levels in the band gap. The length of the four equivalent Si–H bonds was 1.49 Å, for a separation between the hydrogen atoms in the defect core of 1.83 Å.

The highest calculated vibrational mode, at 2251.7 cm^{-1}, was found to be A_1. It corresponds to an in-phase 'breathing' motion of the four hydrogen atoms resulting in a zero dipole moment, and being therefore non IR-active. The triply-degenerate T_2 was found lower in frequency at 2239.8 cm^{-1}, in excellent aggreement with observed value of 2221.9 cm^{-1}. The same is true for the calculated D-stretch modes of VD_4. The calculated mixed H/D modes are given in Table 8.8.

On the convergence with basis and cluster size

As it obvious from a comparison between the calculated frequencies presented here and those reported by Jones *et al.* (1995), using precisely the same method (AIMPRO), that these represent a considerable improvement on the calculational accuracy. Jones *et al.* based their calculations on a clusters containing 131 atoms, with a fairly small number of basis functions. To investigate the convergence of the calculations with cluster and basis size, VH_4 was used, with the observed modes as a reference. The VH_4 defect was then embedded in three different ideal clusters containing 71, 131 and 297 atoms, optimised using differently sized basis sets.

One immediate result was that the accuracy of the calculations converge slowly with cluster size. Another important result concerns the choice of basis. For the 71-atom cluster the calculated frequencies are always found to be higher than the observed ones, independently of the 'quality' of the basis, with the A_1 lying at ~ 2330 cm^{-1}.

For the two other clusters, the calculated LMVs are found within less than 80 cm^{-1} of the experimental values—above these for a small basis and below for a large basis. For the largest cluster, containing 297 atoms, the same trend was observed but now with a smaller error. The results present in this chapter for VH_n were performed using a fairly large basis with the 297-atom cluster.

Further tests were done to investigated the influence of different approaches to the treatment of the terminating hydrogen surface. These were either kept fixed or allowed to relax. For the latter case, full relaxations were also performed using a simple spring-ball model to describe the terminating surface potential with the Si–H length set close to the ideal value (1.47 Å). Within this approach, several different spring constants were also used. For the two larger clusters, we found that a different treatment of the surface produces only a small difference on the calculated values, of $\sim$ 20 cm^{-1}, between the relaxations with a fixed surface and the other alternative treatments.

8.4.6 Electrical properties of multi-hydrogen–vacancy defects

Calculations were carried using the same basis and cluster used for the structure optimisations (Table 8.3).

The calculated levels of the several VH_n defects seem to be determined by the shift and lowering of the vacancy t_2 level. This effect is monitored by starting with the undistorted structure of VH^0, with C_{3v} symmetry, to then look at the influence of the distortion that lowers the symmetry to C_{1h} on the calculated ionisation levels, followed by the result of adding one and two extra hydrogen atoms for VH_2 and VH_3, respectively. In each case, we have assumed that a low spin results from the addition of hydrogen.

The single acceptor level of C_{3v} VH, $(-/0)$, was located 0.11 eV below the conduction band minimum ($E_c - 0.11$ eV), with the single donor level at 0.74 eV above the valence band top ($E_v + 0.74$ eV). For VH, with C_{1h} symmetry, the $(-/0)$ level is at $E_c - 0.43$ and the $(0/+)$ level at $E_v + 0.40$ eV. This is understood as resulting from the splitting of

the e-manifold due to reconstruction across two Si atoms, which means that the highest occupied level is now a singlet a' level, considerable lower in energy than the partially filled one-electron doublet of C_{3v} VH. For the E-centre, we place the $(-/0)$ at $E_c - 0.48$ eV, in very good agreement with the observed value of $E_c - 0.44$ eV (Kimerling *et al.*, 1975). This provides further confirmation of the similarities between this and the VH centre, *i.e.*, VH reveals all the properties of a group-V–vacancy defect.

Although in good agreement with experiment for the $(-/0)$ level, the same is not true for the $(0/+)$ level. We predict a single donor for both VH and VP centre. To our knowledge, no such level has ever been reported.[3] It is evident that increasing the reconstruction, meaning further splitting of the e-level, would result in a shallower donor and acceptor levels.

Table 8.9: Calculated electrical levels of VH_n defects. $(0/+)$ is referred to valence band maximum (E_v) and $(-/0)$ to the conduction band minimum (E_c) and given in eV's. The calculated electrical levels of the E-centre are also shown. This centre is known to posses a single acceptor level at $\sim E_c - 0.4$ eV (Watkins & Corbett, 1964).

	VH (C_{3v})	VH (C_{1h})	VP (C_{1h})	VH_2 (C_{2v})	VH_3 (C_{3v})
$(-/0)$	0.11	0.54	0.48	0.53	0.27
$(0/+)$	0.74	0.40	0.39	0.00	0.00

For neutral VH_2, there are now two electrons (including spin) occupying the a' level, for a $S = 0$ configuration. As a result, the singlet one-electron level is shifted downwards, being now much closer to the valence band top. As a direct consequence of this fact, the donor level is now buried in the valence band. The $(-/0)$ of VH_2 was positioned at $E_c - 0.53$ eV, therefore being very close to that of VH. The transition-state calculation of the affinity of VH_2 for the calculation of the $(-/0)$ level, involved the filling the b_1-singlet lying above the a' level, with half of an electron. The latter becomes partially occupied for neutral VH_3, resulting in a shallower $(-/0)$ at $E_c - 0.27$ eV. For VH_4, the singlet level is completelly filled, which results in the lowering of its position within the bandgap. This results in a electrically passive defect, as no levels were found in the band gap.

8.5 Summary

Despite an improvement of the calculated local vibrational modes for VH_n, $n = 1, \ldots 4$, our calculations confirm the results of Jones *et al.*. The vibrational properties of VH are

[3] Note that the observation of a level in the lower part of the band gap in P-doped Si material is only possible by performing DLTS on p-n junctions, and not metal-semiconductor Schottky barrier devides. This difficulty can be circunvented by the MCTS technique (§§5.5.1).

confirmed to be very similar to those of V_2H. Despite an increase of the calculational accuracy, it is not possible to confirm the experimental assignment of the two pairs of IR bands at {2155,2185} and {2166,2191}-cm^{-1}, to VH_3 and V_2H_6 respectively.

The electrical properties of VH are confirmed to be similar to those of VP, with the $(-/0)$ calculated to lie at ~ 0.50 eV below the conduction band minimum. According to our results, VH_2 should have an acceptor level lying close to those of VH and VP. For VH_3, the $(-/0)$ level is predicted to be shallower at $E_c - 0.27$ eV. As expected, no activity was found for VH_4.

9

Platinum and Palladium Defects

9.1 Introduction

The presence of transition-metal impurities in Si-integrated devices it is not always detrimental. This is the case for platinum. Intentional incorporation of Pt into Si, has been performed to control the lifetime of charge carriers in fast-switching devices. This is of critical importance since the lifetime of minority carriers characterises the quality of the silicon wafer. Although widely used, intentional incorporation of Pt is carried out with little knowledge of the properties of Pt defects giving rise to the deep levels responsible for the adjustment of the carriers lifetime.

Early studies suggested that this control proceeds due to carrier capture and emission from an acceptor level located 0.23 eV below the conduction band minimum ($E_c - 0.23$ eV) and a donor level located at 0.33 eV above the conduction band top ($E_v + 0.33$ eV) (Miller *et al.*, 1976; Chen & Milnes, 1980). Other levels, which are now known to belong to substitutional platinum, have been proposed as responsible for the lifetime control of carriers (Evwaraye, 1976; Lisiak & Milnes, 1976; Braun *et al.*, 1977; Brotherton *et al.*, 1979).

Recently, a new interpretation of electrical properties of Pt has been put forward (Stöffler & Weber, 1986; Lemke, 1987; Kwon *et al.*, 1987; Gill *et al.*, 1990). The control is achieved due to the presence of a single dominant recombination level—arising from the so-called 'midgap' centre—observed in n-type material and located at $E_c - 0.5$ eV. Despite being often indirectly identified as substitutional Pt, there is no consensus about the structure or the chemical composition of this Pt-related defect from which the midgap level originates. For example, defect structures like a platinum-platinum pair (Höhne, 1992; von Bardeleben *et al.*, 1988) or a platinum-oxygen pair (Kwon *et al.*, 1987; Höhne & Juda, 1995; Juda *et al.*, 1996) have been proposed as the centres responsible for the midgap level.

This difficulty in the identification of the defect responsible for the levels illustrates a weakness of DLTS. Despite being a very powerful technique for the detection of deep levels, it does not provide any information about either the defect's structure or its chemical composition.

The interaction of hydrogen with transition-metal impurities is usually thought as a process to eliminate the electrical activity of the TM centre by removing the deep-lying levels from the gap as a consequence of hydrogen chemical-rebonding, achieving what is usually termed *passivation* (Benton *et al.*, 1980; Pearton *et al.*, 1992).

The electrical passivation of Pt (n-type) and Pd (n- and p-type) was first reported by Pearton and Haller (1983) subjecting the Si wafers to a remote high-frequency H-plasma at 300 °C. The disadvantages of this technique are low penetration depths, the creation of native defects since the samples are subject to ultra-violet radiation and considerable particle bombardment (Lightowlers, 1995) and possible surface damage. Additionally to this neutralisation effect, hydrogenation can, and often does, create new hydrogen–TM-related defects, or it can simply shift the existent levels across the bandgap (Johnson *et al.*, 1987; Sachse *et al.*, 1999a). Obviously, the formation, or not, of TM–H defects is strongly dependent on the hydrogenation process.

This is evident when hydrogen is deliberately incorporated into platinum-doped Si (Sachse *et al.*, 1997c; Sachse *et al.*, 1997b), by a wet chemical etching at room temperature (Tavendale *et al.*, 1988), which results in the formation of a number of hydrogen-platinum defects, containing several H atoms. A similar picture has been reported for wet-etched Pd-doped material (Sachse *et al.*, 1997a).

Soaking in H_2 gas at high temperatures ($\gtrsim$ 900 °C) followed by rapid quenching to room temperature is another method of incorporating H. This hydrogenation process has been tailored for the study of TM-hydrogen defects in Si by spectroscopic techniques like EPR or FTIR which require an uniform density throughout the Si bulk, typically of $\approx 10^{15}$ cm^{-2} (Veloarisoa *et al.*, 1991; Williams *et al.*, 1993).

The outline of the present chapter is as follows. First we summarise the available experimental information, combining the results of several spectroscopic studies on Pt- and PtH-related defects and discussing the several available models for the electronic structure of substitutional Pt^- and Pd^-. Secondly, we report the results of a theoretical study on the structural properties of defects involving the TM impurities as well as hydrogen. These results are compared with experiment, *e.g.*, EPR and FTIR spectroscopy. Finally, we extend the study to the electrical properties of the defects, *i.e.*, we calculate the position of the levels associated with the substitutional Pt centre, and investigate the effect on these of bringing H atoms close the impurity ion. The results are compared with relevant experimental findings, and the role of a PtH defect as a recombination centre in silicon is also discussed.

9.2 Experimental work

It is known since the pioneering work of Woodbury and Ludwig (1962) that the platinum impurity sits substitutionally in the Si lattice, being singly negatively charged in its paramagnetic state with ground-state spin of $S = \frac{1}{2}$. The corresponding EPR spectrum ($T \lesssim 12$ K) was interpreted as arising from a defect with orthorhombic symmetry (C_{2v}),

due to a displacement of the TM impurity along the dihedral [100] direction; the g-tensor components for Pt^- were $g_{xx} = 1.3867$, $g_{yy} = 1.4266$ and $g_{zz} = 2.0789$, revealing a large departure from the free-spin value, g_0, of 2.0023. This picture has been confirmed by later EPR studies (Henning & Egelmeers, 1983; Milligan *et al.*, 1984; Omling *et al.*, 1987).

9.2.1 Microscopic models for the electronic structure

Two main models exist to explain the electronic structure of substitutional transition-metal (TM) impurities, like Pt or Pd. These are the ionic model and the vacancy model.

The ionic model was introduced by Ludwig and Woodbury (1960) as an attempt to explain the electronic structure of interstitial and substitutional TM impurities in Si, studied by EPR (Ludwig & Woodbury, 1962). In the case of a substitutional impurity, this model postulates the transfer of electrons from the d shell into valence states to achieve a strong covalent hybridization between the impurity ion and its four Si nearest-neighbours, resulting in a undistorted, tetrahedral, centre. It assumes the validity of the Hund's rule, as in the case of an isolated ion, *i.e.*, zero crystal-field. In the case of a substitutional Pt impurity (Pt_s^-), the tetrahedral crystal-field splits the fivefold orbitally-degenerate d level of the TM ion, into a triplet (t_2) and a doublet (e) with the later lying below the t_2 level.

The vacancy model is due to Watkins (1983) based on EPR experiments on the single-vacancy (Watkins, 1976), A-centre (Watkins & Corbett, 1961) and the X_α cluster calculations of Hemstreet on a number of substitutional TM impurities (Hemstreet, 1976). According to this model, the properties of substitutional impurities can be understood by considering a filled d-shell lying deep in the valence band, with the remaining electrons occupying the gap-lying t_2, vacancy-like (Hemstreet, 1976; Cartling, 1975; Alves & Leite, 1986), orbitals originating from the Si first-neighbours (Fig. 9.1). In other words, it is assumed that all the properties of a substitutional TM impurity are simply due to the Si sp^3 hybrid orbitals, and not from the many-electron character of the TM d-shell, as in the ionic model.

Interpretation of the experimental data

According to the vacancy model, the electronic configuration of substitutional Pt^-, Pt_s^- *i.e.*, the number of electrons in the d shell, is $5d^{10}$. The Pt^- impurity, as well as Pd_s^- behave like a closed $5d$-shell sitting on a single-charged vacancy, V^-. For Pt^-, three electrons occupy the anti-bonding-like triplet lying deep in the gap ($t_2^{\uparrow\uparrow\uparrow}$). This partial occupancy of the t_2-manifold provides the basis for a JT distortion.

A different electronic structure is predicted by the ionic model for Pt^-: $5d^7$. More generally, this model predicts an electronic configuration d^{q-4}, for substitutional defects in Si, with q being the number of valence electrons of the TM atom. Assuming that electron-electron effects are more important than the crystal-field or Jahn-Teller effects, this model predicts a high-spin configuration of $S = \frac{3}{2}$ for Pt^-, in clear disagreement with experiment (Woodbury & Ludwig, 1962; Milligan *et al.*, 1984; Anderson *et al.*, 1992b).

The vacancy model, assuming that lattice interactions (JT effect) are more important than those between electrons, allows a low-spin configuration.

At this point, it is interesting to mention the case of the Ni_s^- centre. In diamond, this centre has a ground-state spin of $S = \frac{3}{2}$ (Isoya *et al.*, 1990; Goss *et al.*, 1995), whereas in silicon, experiment and theory have confirmed that Ni_s^- assumes the low spin configuration of $S = \frac{1}{2}$ (Vlasenko *et al.*, 1990; Són *et al.*, 1990; Watkins *et al.*, 1991; Watkins & Williams, 1995; Jones *et al.*, 1995). It is still not fully understood how the same centre can have different spin configurations. Despite the same crystalline structure, exchange-correlations effects are known to be more important in diamond than in silicon. In addition, silicon is also 'softer' than diamond, which makes it more susceptible to JT-driven distortions. This relates to the relative importance between crystal-field, electron-electron and JT effects. While many-body effects are more important in diamond, it seems that static Jahn-Teller effects dominate in silicon.

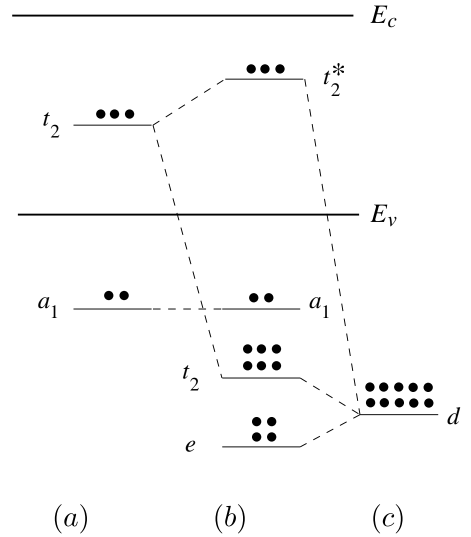


Figure 9.1: One-electron energy levels for (*a*) the negative silicon vacancy in full tetrahedral symmetry, (*b*) the d^{10} states of the transition-element impurity, and (*c*) the substitutional impurity in silicon (without Jahn-Teller distortion).

For the microscopic structure, the vacancy model predicts a distorted centre similar to V^-; the electronically degenerate t_2 level, being partially occupied, is therefore unstable against a symmetry-lowering JT distortion to first-order. Like in the case of V^- (Watkins, 1976; Watkins, 1992), two static distortions occur ($T_d \rightarrow D_{2d} \rightarrow C_{2v}$) to stabilise the structure of the TM centre. This is easier to understand if one considers the formation of the substitutional Pt^- as a two-step process (Hemstreet, 1976): the first step is the formation of the vacancy and the second is the accommodation of Pt impurity into the vacancy, which originates the interaction of the metallic $d(t_2)$ states and the t_2 vacancy orbitals, with the later defined as linear-combinations of orbitals localised at the Si dangling-bonds. These two steps may be combined for the formation of the defect, in what is usually referred to as a *kick-out* mechanism (Gösele *et al.*, 1980), *i.e.*, the vacancy is created by the impurity itself, releasing a Si self-interstitial (Pt $\Rightarrow$ $Pt_s + Si_i$). Experiments have confirmed the kick-out mechanism as the preferred one for the diffusion of Pt in Si for $T > 850\,^\circ$C (Kwon *et al.*, 1987; Mantovani *et al.*, 1986; Schmidt *et al.*, 1998).

In their original paper, Ludwig and Woodbury (1962) considered a closed-shell $5d^8$ configuration, with a hole ($j = \frac{3}{2}$) localised around the impurity Pt ion, to justify the departure from T_d symmetry and its paramagnetism. Lowther (1980) supported this idea, proposing a different structure for Pt^- and arguing that the resonance originates from a hole localised at p-type bonding-orbital outside the Pt $5d$-shell. According to this model,

the substitutional Pt^- centre shows a similar structure to that of the Si-A centre (Watkins & Corbett, 1961), *i.e.*, the Pt ion is directly bonded to two of its four Si neighbours through its $6s$ and $6p$ atomic orbitals.

It is now obvious that only one of the models described above can account for the properties of the Pt^- centre. Experiment (Milligan *et al.*, 1984; Stöffler & Weber, 1986; Anderson *et al.*, 1992a; Watkins & Williams, 1995) and theory (Alves & Leite, 1986; Beeler *et al.*, 1985; Anderson *et al.*, 1991; Jones *et al.*, 1995) have confirmed the ground-state spin and symmetry, as predicted by the vacancy model. Despite this fact, Ammerlaan and von Oosten (1989) have suggested that this model cannot account for the observed g values. These authors introduced the *dihedral* model, a modified version of ionic model, resulting in a structure similar to that of Lowther (1980). This results in a $5d^9$ configuration, $S = \frac{1}{2}$ and $L = 1$, with the unpaired electron in the $5d$ shell, for Pt^-, being this open-shell structure responsible for the large departure of $g_\perp$ from g_0.

Using a perturbation theory approach, which included spin-orbit interactions together with JT electron-lattice coupling, Anderson *et al.* (Anderson *et al.*, 1992a) have shown that the vacancy model can indeed account for the properties of the Pt^- centre, like g values, hyperfine parameters and its behaviour under stress. The later relates to a difficulty of this model to predict the sense of the JT distortions for both, tetragonal and trigonal modes.

Uniaxial stress measurements (Williams *et al.*, 1994) on the EPR lines have shown that the t_2 gap-levels split into a configuration $b_1^{\uparrow\downarrow}\, b_2^{\uparrow}\, a_1^0$, which is the reverse of that of the V^-, *i.e*, $a_1^{\uparrow\downarrow}\, b_1^{\uparrow}\, b_2^0$ (Watkins, 1992). Anderson *et al.* (1992a) have demonstrated that if non-linear JT coupling terms are considered, the vacancy model can account qualitatively for the form of the JT distortion undergone by the Pt_s^- centre. Additionally, this group were able to explain the departure of the $g_\perp$ from g_0, in light of the vacancy model, as a contribution to the orbital g factor from the metallic $5d$ orbital of the impurity atom. This admixture was found to be $\approx 10\%$, consistent with a vacancy-like character, in disagreement with the value reported by Ammerlaan and von Oosten of $\approx 70\%$ (Ammerlaan & van Oosten, 1989).

Recently, Watkins and Williams (1995) have proved that there is definitely a close competition between spin-orbit and JT effects. To illustrate this fact, this group investigated the case of Au_s^0, comparing it with Pt_s^-, Pd_s^- and Ni_s^- (Watkins *et al.*, 1991; Kleverman *et al.*, 1995) for which they postulated that the increasing importance of spin-orbital effects—which work against the JT driving force—going from Pt_s^- to Au_s^0, may be responsible for the fact the Au_s^0 centre is almost tetrahedral. It was then argued that this is the reason why the Au_s^0 defect was never observed by EPR. Due to the fact that the Au_s^0 is characterised by $g_\perp \simeq 0$, showing an even larger departure from the free-spin value for Pt_s, the microwave transition ($\Delta m_s \neq 1$) becomes very small, being difficult to detect by EPR.

Interaction with atomic hydrogen

Due to its diamagnetic properties, $(PtH_1)^-$, containing a single H atom was never observed directly by EPR. To overcome this difficulty, infra-red absorption spectroscopy has been used to investigate the structural properties of such complexes. Thermal treatment, the analysis of deuterium-induced isotopic shifts and behaviour under stress are additional tools that allow the characterisation of the defects being observed by IR absorption spectroscopy. The behaviour of the vibrational bands during annealing may correlate the bands to a particular defect and stress-induced splittings assign the defect's symmetry. The Fermi dependence of the bands can be used to determine the charge state(s) of defect responsible for the vibronic band(s).

Several local vibrational modes have been identified as arising from different charge states of a trigonal (C_{3v}) PtH_1 defect (Uftring *et al.*, 1995). These were the bands at 1880.7 and 1897.2 cm^{-1}, with the lower frequency associated with the unoccupied charge state. The fact that no shifts were detected due to ^{29}Si, or ^{30}Si, which might have provided some evidence for a Si–H bond, prevented the identification of the lattice site occupied by the H atom. (Evans *et al.*, 1997) If one compares the observed LVMs of PtH_1 and AuH_1 (Evans *et al.*, 1997; Evans *et al.*, 1999), it is than obvious that their vibrational properties are indeed similar (70 cm^{-1} separation between PtH_1 and AuH_1 modes), suggesting a similar microscopic structure for both defects. Although, the symmetry assignment is different for the two complexes. AuH_1 has monoclinic (C_{1h}) symmetry (Evans *et al.*, 1999), in −, 0 and + charge states, while the PtH_1 is a trigonal or near-trigonal defect (Uftring *et al.*, 1995). To explain the lowering in symmetry for AuH_1, Evans *et al.* (1999) have suggested that an off-site displacement occurs along the [100] direction in a similar way to that of Pt_s^-.

Combined EPR and LVM spectroscopy measurements (Uftring *et al.*, 1995) have allowed the identification of additional H-related stretch bands at 1888.2 and 1901.6 cm^{-1}, as antisymmetric and symmetric modes of PtH_2^-, stable up to 600 K. Additional H-vibrations have been detected and assigned to other charge states of the same PtH_2 defect. These are given in Table 9.3.4. The involvement of two equivalent H atoms in the complex was undoubtedly proved by the analysis of the angular dependence of the g tensor and the ^{195}Pt hyperfine tensors in the EPR measurements, which revealed a C_{2v} symmetry. Furthermore, the analysis of the anisotropic part of the hyperfine tensor, located the H atom at a distance of ~ 4.2 Å from the central Pt ion. This distance is consistent with a model for the PtH_2 defect, in which the H atoms lie at anti-bonding sites (AB), being back-bonded to two of the Si neighbours of Pt^-, as in the case of $(NiH_2)^-$ and $(AuH_2)^0$ (Jones *et al.*, 1995; Resende *et al.*, 1999). The EPR measurements of Höhne and colleagues (1994) have confirmed the above assignment (Uftring *et al.*, 1995). As a point of detail, it is important to mention that in their EPR experiments, hydrogenation proceeds by subjecting the samples to a H_2O-vapour atmosphere in the temperature range of 800–1200 °C.

Electrical levels

It is now accepted, as a result of optical (Armelles *et al.*, 1986; Omling *et al.*, 1987; Olajos *et al.*, 1989) and capacitance measurements (Brotherton *et al.*, 1979; Stöffler & Weber, 1986; Zimmermann & Ryssel, 1991; Lemke, 1995; Sachse *et al.*, 1997b), that Pt can exist in four different charge states, as a substitutional impurity in Si. As a result of electron and hole transitions between the several charge states of the defect, three levels exist in the bandgap. These are given in Table 9.1.

Table 9.1: Activation energy ΔE, majority capture cross sections σ, activation energy of the capture cross section E_σ, the enthalpy ΔH and the entropy factor X_s for Pt and Pd energy levels. Values taken from Sachse *et al.* (1997c).

	Level	ΔE (eV)	$\sigma_{n,p}$ (cm^2)	E_σ (eV)	ΔH (eV)	X_s
	$(-/0)$	$E_c - 0.23$	5×10^{-15}	0	$E_c - 0.23$	0.5
Pt	$(0/+)$	$E_v + 0.33$	8×10^{-16}	0	$E_v + 0.33$	16
	$(+/^+_+)$	$E_v + 0.09$	2×10^{-17}†	0.03	$E_v + 0.06$	10
	$(-/0)$	$E_c - 0.22$	5×10^{-15}	0	$E_c - 0.22$	0.8
Pd	$(0/+)$	$E_v + 0.33$	8×10^{-16}	0	$E_v + 0.33$	21
	$(+/^+_+)$	$E_v + 0.14$	1×10^{-16}†	0.03	$E_v + 0.11$	7

† Temperature dependent.

It is obvious from the above table, that the level structure of Pd_s is very similar to that of Pt, which strongly suggests a similar microscopic structure for all charge-states of substitutional Pd. The $(-/0)$ of Pd occurs at $E_c - 0.22$, the $(0/+)$ at $E_v + 0.31$, and the $(+/^+_+)$ level at $E_v + 0.14$ eV (Lemke, 1984; Zimmermann & Ryssel, 1991; Gill *et al.*, 1993; Sachse *et al.*, 1997a). The activation energies of the donor levels of Pd and Pt were found to be electric-field dependent. A decrease in the activation energies of the levels, having as a consequence the enhancement of their emission rates, was observed, as predicted by Poole and Frenkel (1938).

Wang *et al.* (1988) as a result of stress measurements on the DLTS bands observed in Pd-doped Si, have suggested that the above $(-/0)$ and $(0/+)$ levels arise from a tetrahedral (T_d) defect while a level at $E_c - 0.18$ eV level comes from a defect exhibiting C_{2v} symmetry. No second donor level was reported.

Apart form the levels attributed to isolated Pt, many other levels were reported in the literature (see Brotherton (1979) and references therein). For example, Evwaraye and Sun (1976) reported a total of six levels in Pt-doped Si, interpreting two of them as the single acceptor and donor levels of Pt. Performing DLTS experiments on *np*-junction devices with different doping levels for *n*- and *p*-type structures (phosphorus- and boron-doped, respectively), Evwaraye and Sun were able to assign some levels to complexes of

platinum with shallow dopants, based on the fact some of those levels were only present in p-type structures, while others were only observed in n-type devices. An example is the level at $E_v + 0.41$ eV, previously reported by Woodbury and Ludwig (1962) which was not present in p-type material, suggesting a Pt-P complex as responsible for the level. One of the levels detected by Evwaraye and Sun, located at $E_c - 0.34$ eV, was proposed as a candidate for the dominant centre. This level disappeared from the DLTS thermal scan when holes were injected into the depletion region of the n-type junction device. Despite this strong evidence of an efficient recombination centre, its origin remains unknown.

Table 9.2: Activation energy ΔE, majority capture cross sections σ, activation energy of the capture cross section E_σ and the enthalpy ΔH for a number of PtH-related electrical levels.

Level	ΔE (eV)	$\sigma_{n,p}$ (cm^2)	E_σ (eV)	ΔH (eV)
E(90)	$E_c - 0.18$	$1 \times 10^{-17\dagger}$	0.02	$E_c - 0.16$
E(250)	$E_c - 0.50$	2×10^{-16}	0	$E_c - 0.50$
H(150)	$E_v + 0.30$	9×10^{-16}	0	$E_v + 0.30$
H(210)	$E_v + 0.40$	3×10^{-17}	0	$E_v + 0.40$

† Temperature dependent. The values of ΔE and ΔH for the single donor level are given from extrapolation to zero field. These and the other values are taken from Sachse *et al.* (1997b).

Recent DLTS experiments (Stöffler & Weber, 1986; Lemke, 1987; Kwon *et al.*, 1987; Gill *et al.*, 1990) have detected a midgap level at $\sim E_c - 0.50$ eV observed near the surface region of the n-type Schottky diodes. The fact that this level was found in low concentrations, was invoked to explain why it was not observed previously. According to the Shockley-Read-Hall (SRH) model (Shockley & Read Jr., 1952; Hall, 1952), this 'midgap' level presents itself as a better recombination centre than that reported by Evwaraye and Sun (1976). According to SRH, the rate for successive capture of electrons and holes increases exponentially with energy depth.

Sachse *et al.* (1997c) presented a different interpretation of dominant recombination centre in Pt-doped Si. According to this group, the midgap level arises from a defect involving Pt and hydrogen, introduced by room temperature wet etching, and not from an isolated platinum impurity, as suggested by Gill *et al.* (1990). Sachse *et al.* (1997c) interpreted the level at $E_c - 0.55$ eV, observed by Gill and colleagues (Gill *et al.*, 1990), as the single acceptor level of substitutional gold, due to a residual Au contamination.

The DLTS spectrum of n-type hydrogenated Si samples exhibited two distinct peaks, (E90) and (E250), at $E_c - 0.18$ and $E_c - 0.50$ eV, respectively (Sachse *et al.*, 1997c; Sachse *et al.*, 1997b), The corresponding electron-cross sections were 1×10^{-17} and 2×10^{-16} cm^2. The smaller capture cross section of the first is consistent with a $(=/-)$ level. Depth

profiling and thermal annealing on the above DLTS peaks have suggested that E(90) and E(250) are levels of two different Pt-hydrogen defects stable up to 600 K. This study has also provided some evidence for the passivation of Pt in their n-type samples. Near the sample surface, the concentration of the three levels, E(90), Pt($-$/0) and E(250), did not add up to the concentration of the Pt($-$/0) in 'H-free' (cleaved) samples. This suggests that a passive defect has been formed, in the H-rich near-surface region, as a result of the interaction of atomic hydrogen with the Pt ion. Minority-carrier transient spectroscopy (MCTS) measurements (Sachse *et al.*, 1997c) on p-type samples revealed that the concentration trap E(250) is higher than that of H(210), suggesting that these do not belong to the same centre. A similar situation was observed for wet-etched Pd-doped n-silicon, between 450 and 600 K in the DLTS scan after annealing (Table 9.3). The ($-$/0) level of Pd_s is recovered for temperatures above 600 K.

Table 9.3: Activation energy ΔE, majority capture cross sections σ, activation energy of the capture cross section E_σ and the enthalpy ΔH for a number of PdH-related electrical levels.

Level	ΔE (eV)	$\sigma_{n,p}$ (cm^2)	E_σ (eV)	ΔH (eV)
E(60)	$E_c - 0.10$	$1 \times 10^{-18\dagger}$	0.03	$E_c - 0.07$
E(160)	$E_c - 0.29$	2×10^{-16}	0	$E_c - 0.29$
E(200)	$E_c - 0.43$	2×10^{-17}	0	$E_c - 0.43$
H(280)	$E_v + 0.55$	5×10^{-17}	0	$E_v + 0.55$
H(140)	$E_v + 0.24$	5×10^{-16}	0	$E_v + 0.24$
H(55)	$E_v + 0.08$	2×10^{-17}	0	$E_v + 0.08$
H(45)	$E_v + 0.07$	1×10^{-18}	?	$E_v + 0.07$

† Temperature dependent. The values of ΔE and ΔH for the single donor level are given from extrapolation to zero field. These and the other values are taken from Sachse *et al.* (1997d; 1997a).

The passivation by hydrogen was observed by Pearton and Haller (1983), who were able to remove the Pt($-$/0) level from the upper part of the bandgap. The Pt(0/+) suffered not change with hydrogenation (low-pressure 300 °C H-plasma). Unexpectedly, no H-related levels were reported by these authors.

In a following paper, Sachse *et al.* (1997b) reported four DLTS signals in p-type Si, including the single and donor levels of isolated Pt. The concentration of the levels was shown to be strongly dependent upon the temperature and duration of the annealing procedure. The concentration of the (0/+) and ($+/^+_+$) levels were identified in all samples, clearly confirming that these are levels of the same defect. A peak at $T = 210$ K, detected right after etching, was enhanced after a annealing at 400 K, with the simultaneous reduction of the concentration of (0/+) and ($+/^+_+$) DLTS peaks. This peak, H(210), was

identified as arising from a level located at $E_v + (0.40 \pm 0.03)$ eV (Table 9.2). Longer annealing at $T = 490$ K has revealed another peak, H(150), at $E_v + (0.30 \pm 0.03)$ eV ($\sigma_p = 9 \times 10^{-16}$ cm^2). Both H(150) and H(210) disappeared after annealing at higher temperatures ($T > 650$ K). The observation of a Poole-Frenkel effect for the H(150) level strongly suggests that this is an *acceptor* $(-/0)$ level detected in the lower part of the bandgap (Weber, 1999).

Depth profiling under reverse bias annealing (RBA) suggested that the two levels, H(150) and H(210), arise from PtH defects involving a different number of hydrogen atoms. The latter, H(210), is detected immediately after etching suggesting a defect with few hydrogen atoms.

This can be understood as a direct effect of the RBA on the stability of B-H pairs. This was confirmed by depth profiling of the net acceptor activity in the p-type samples, by means of capacitance-voltage (CV) measurements. It was then evident from the CV profiles, that the RBA results in a drift of atomic hydrogen (H^+), being released from previously stable B–H pairs (stable up to 400 K), towards higher depths and away from the sample surface. The CV profiling also revealed the character of the H(150) level. The CV profiling showed an increase of the net concentration of acceptors at depths where that of H(150) is maximum, revealing the acceptor-like character of the level.

Correlation between DLTS and EPR results

Uftring and co-workers (1995) as a result of EPR and LVM spectroscopy studies on PtH defects, were also able to identify elegantly some of the electrical levels of PtH_2. Varying the pseudo-Fermi level position by either using samples with different Pt concentrations, or through band-gap light illumination, and using the levels of C_i produced by 2.5 MeV-electron irradiation as markers (Williams *et al.*, 1993), these workers were able to locate two levels of PtH_2 within a certain energy range. A second acceptor was found between the (0/+) level of P and that of the (0/+) of C_i, *i.e.*, in the range 0.045–0.1 eV below the conduction band minimum, and a single acceptor was located between the single acceptor and donor levels of substitutional Pt (Fig. 9.2).

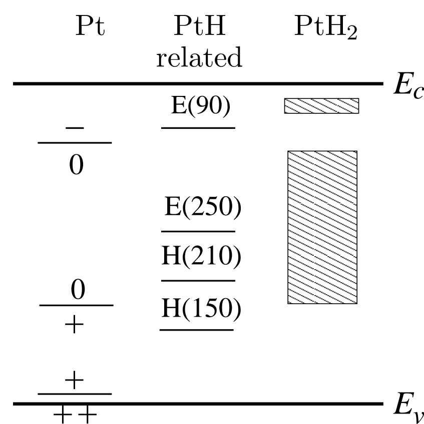


Figure 9.2: Energy-level diagram depicting the levels of Pt and PtH as observed by DLTS, and PtH_2 from EPR. The filled rectangles indicate that the $(-/0)$ of PtH_2 is between $E_c - 0.045$ and $E_c - 0.1$ eV and that the $(-/0)$ is between $E_c - 0.23$ and $E_v + 0.33$ eV.

The question now is how do the levels observed by Sachse *et al.* (Sachse *et al.*, 1997c; Sachse *et al.*, 1997b) correlate with the findings of Uftring and co-workers (1995; 1993) Simply looking at their thermal stability, either three levels lying in the upper part of the bandgap—E(90), E(250) or H(210) — are

possible candidates for the two levels of PtH_2 detected by Uftring *et al.* (Uftring *et al.*, 1995; Williams *et al.*, 1993). However, despite an apparent similarity in thermal stability between the DLTS peaks and the EPR signals, this assignment cannot be made. E(90) and E(250) are very likely levels of a different PtH complex and E(90), at $E_c - 0.18$ eV, seems to be too deep to be the double acceptor of PtH_2 (by $\sim$ 80 meV), observed by EPR and LVM spectroscopy. This difference is larger larger than the error in determining ionisation enthalpies in the DLTS experiments (typically of $\sim$ 10 meV). This discrepancy does not seem to be understood (Sachse *et al.*, 1999b).

Determination of the number of H atoms in TM–H complexes from DLTS

The analysis of the annealing behaviour and depth profiling characteristics of a DLTS spectrum are two powerful techniques that are usually used to tentatively assign two or more peaks to a particular defect, or alternatively, to refute that same assumption. An example of the later is the case of the E(90) and E(250) DLTS peaks.

Depth profiling under RBA strongly suggested that these two levels do not arise from the same defect. Similarly, H(210)—an acceptor level detected in *p*-type material—is a different defect from that responsible for E(250). Assuming that all these defects involve a single Pt ion and H, these must contain a different number if H atoms. Unfortunately, DLTS cannot provide any information about the number of H atoms in the PtH complexes.

Recently, Feklisova and Yarykin (1997) have proposed that the variation of the defect concentration with depth can be used to identify the chemical composition of the defect using DLTS, in particular defects containing several hydrogen atoms. The foundations of Feklisova and Yarykin's conjecture are based on the observed superlinear dependence on the hydrogen concentration of the TM–H defects as a function of the penetration depth.

The concentration profiles of mobile hydrogen and hydrogen can be described by the following diffusion equation (Feklisova & Yarykin, 1997):

$$\frac{\partial[\mathrm{H}]}{\partial t} = D\,\frac{\partial^2[\mathrm{H}]}{\partial x^2} + v\,\frac{\partial[\mathrm{H}]}{\partial x} - \frac{[\mathrm{H}]}{\tau}\ , \tag{9.2.1}$$

where [H] is the concentration of mobile hydrogen, x is the distance to surface, v is the etch rate, D and τ are the diffusion coefficient and lifetime of mobile hydrogen, respectively. Assuming a deep enough layer, all these coefficients can be considered constant and the steady-state solution of the above equation can be written as

$$[\mathrm{H}] = H_0\,\exp\left(-x\,L\right)\ , \tag{9.2.2}$$

where (Feklisova & Yarykin, 1997),

$$L = \sqrt{\frac{1}{D\,\tau} + \left(\frac{v}{2D}\right)^2} + \frac{v}{2D}\ . \tag{9.2.3}$$

Assuming that all the H-related complexes formed by subsequent addition of H atoms are thermally stable, the concentration of the complexes containing n hydrogen atoms, $[\mathrm{A}_n]$, can be described as

$$\frac{\partial[\mathrm{A}_n]}{\partial t} = v\,\frac{\partial[\mathrm{A}_n]}{\partial x} + 4\,\pi\,D\left(r_{n-1}\,[\mathrm{A}_{n-1}] - r_n\,[\mathrm{A}_n]\right)[\mathrm{H}] \tag{9.2.4}$$

Table 9.4: Penetration lengths L_n and the corresponding experimental assignments for PtH and PdH complexes (Sachse, 1997; Sachse *et al.*, 1997d).

Level	L_n (μm)	L_1/L_n	n	Assignment
E(250)	1.55 ± 0.05	1	1	PtH_1
E(90)	0.80 ± 0.02	1.95 ± 0.1	2	PtH_2
H(210)	0.78 ± 0.02	2.0 ± 0.1	2	PtH_2
H(150)	0.48 ± 0.05	3.2 ± 0.4	3	PtH_3
E(60)	0.71 ± 0.08	2.3 ± 0.4	2–3	$PdH_{2,3}$
E(160)	0.52 ± 0.06	3.2 ± 0.3	3	PdH_3
E(200)	1.62 ± 0.03	1	1	PdH_1
H(45)/H(55)	0.57 ± 0.06	2.9 ± 0.3	3	PdH_3
H(140)	0.49 ± 0.06	3.3 ± 0.5	3–4	$PdH_{3,4}$
H(280)	0.82 ± 0.03	2.0 ± 0.1	2	PdH_2

where r_n is the capture radius of hydrogen to the TM–H complex.

Assuming that $[\mathrm{A}_n] \gg [\mathrm{A}_{n+1}]$, we can ignore the second term in parenthesis in (9.2.4), to write its solution as

$$[\mathrm{TM{-}H}_n] \propto \exp\left(-x\, L_n\right) \quad \text{and} \quad L_n = \frac{L}{n}\,, \tag{9.2.5}$$

with L_n being the characteristic penetration depth of the TM–H defect, defined by (9.2.3).

The slope resulting from the linearisation of the concentration–depth curve provides L_n for a particular trap. The ratio L_1/L_n between a reference penetration depth, L_1, and that of a given defect is then simply the number n of H involved in the defect. The reference trap is usually taken to be that which shows a larger penetration depth, and consequently should involve only one H atom. The relative penetration depths for Pt– and Pd–H defects are given in Table 9.4. The levels E(250) and E(200) are the reference traps (L_1) for Pt and Pd respectively. The proposed assignments for the observed DLTS level of Pt– and Pt– are given in Table 9.4.

Applying this method to the DLTS traps E(90) and E(250), for which L is equal to 0.80 and 1.55 μm respectively (Sachse, 1997), and using latter as the reference, one obtains $L(90)/L(250) \simeq 1.95$. In other words, E(90) involves two times more H atoms than E(250), *i.e.*, assuming that the complex responsible for E(250) contains a single H atom, E(90) should be assigned to a PtH_2 complex.

9.3 Results

9.3.1 Cluster and basis

The calculation were performed on a tetrahedral cluster containing 131 atoms ($Si_{71}H_{60}$), centred at an atom site. In the defect cluster, the impurity replaced the central Si host atom. For the optimisation runs, the hydrogen terminators were kept fixed. The atomic basis used is described below (Table 9.5).

Table 9.5: Atomic basis and basis expansion type used for the calculations on Pt and Pd–H defects. N and M denoted the number of orbitals used to describe the wavefunction and charge density respectively. One additional basis function was placed midway between all bonded pairs for the electrical level calculations.

Cluster	Atom	N	M	Basis expansion	`chden`
	Pt/Pd	8	14	both big	
$Si_{71}H_{60}$	Si	5	4	all big	0
	H	3	2	all minimal	
Bond-centres	all bonds	2/3	2/3	big	

9.3.2 Microscopic structure as isolated species

To model the Pt centre, two different configurations have been considered: an undistorted T_d structure and the distorted orthorhombic C_{2v} structure. For both, Pt was embedded in a negatively charged tetrahedral $Si_{70}H_{60}$ centred around the Pt substitutional site. For the initial structure of Pt_s^- (C_{2v}), the TM ion was displaced along the [100] direction by 0.25 Å, resulting in the lowering of symmetry of the otherwise T_d defect. To avoid self-consistency problems during spin optimisation (numerical instabilities due to *charge-sloshing*) (Jones *et al.*, 1995) and allow the JT distortion, the levels were filled according to Fermi statistics at 0 K, in a low-spin configuration for a ground-state spin $S = \frac{1}{2}$. Similar structures were considered for the isoelectronic Pd_s^- centre.

The self-consistent energy of the cluster and the inter-atomic forces acting on the inner 71 atoms for the above two configurations of Pt_s^- and Pd_s^- were found as a result of the *ab initio* optimisation of the cluster atomic positions.

The spin-polarised Kohn-Sham eigenvalue spectrum corresponding to the optimised structure of undistorted Pt_s^- revealed a single one-electron gap level: a partially occupied t_2 manifold. No other levels were found in the Si bandgap. For the C_{2v} configuration, three closely spaced singlet levels lay in the bandgap as a result of splitting of the t_2 manifold, a consequence of the off-site displacement of the Pt ion and the simultaneous re-arrangement

of the Si atoms surrounding the ion.

The optimisation of the undistorted Pt_s^-, resulted in an outwards breathing motion of the Si nearest-neighbours of the impurity. The four Pt–Si bonds were equivalent with lengths equal to 2.48 Å, being 4% longer than a relaxed Si–Si bond (2.38 Å) in a pure Si cluster. For Pd_s^-, there is a larger outward movement of the first shell. Accordingly, the Pd–Si lengths were 2.64 Å for an increase of ≈ 11% relatively to a bulk Si–Si bond.

For both Pt^- and Pd^-, the relaxation of the distorted C_{2v} structure revealed a very small departure from the central lattice site, with the TM ion sitting 0.05 and 0.02 Å away from the T_d site, respectively. The Si atoms moved away and towards each other by pairs, as in the case of the dominant tetragonal (D_{2d}) distortion of V^-. The Si–Pt lengths were 2.44 and 2.48 Å (Table 9.6). This slight distortion is consistent with the ease of reorientation of the TM impurity around the central lattice site, observed experimentally even at very low temperatures ($T = 4.2$ K) (Anderson *et al.*, 1992b). The adiabatic barrier for the off-site displacement, defined as the difference in total energy between T_d and C_{2v} configurations, was calculated to be 0.15 eV for Pt_s^- and 0.14 eV for Pd_s^-.

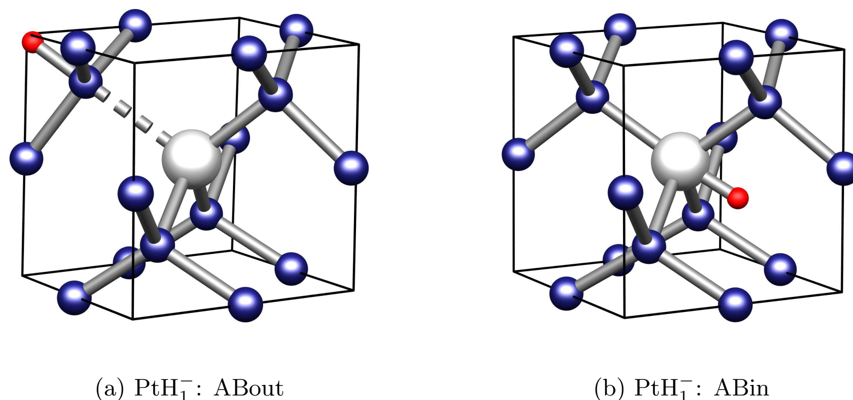
To test the dihedral model of Ammerlaan and van Oosten, (Ammerlaan & van Oosten, 1989) a third configuration for Pt_s^- was considered. Following the prescription of this model, the structure was constructed by moving the impurity Pt ion 1.0 Å away from the central lattice site and along the [100] direction, being now much closer to two of its Si neighbours. Additionally, the bond between the pair of Si atoms in the (110) plane was slightly reconstructed. This configuration was found to be unstable, as the Pt ion moved back considerably to the central T_d site as a result of relaxation. The Pt found a stable position at 0.19 Å away from the T_d and along the [100] direction. This structure was 0.12 eV higher in energy than that of orthorhombic Pt_s^- defect described previously.

9.3.3 Transition-metal–hydrogen defects

The effect of H on the structure of the Pt and Pd centres was investigated using a tetrahedral $PtSi_{70}H_{60+n}$ cluster, with n being the number of H atoms added to the centre, running from 1 to 4.

There are two obvious, high-symmetry, lattice sites which can accommodate atomic hydrogen: bond-centred (BC), midway between the central TM ion and one of its immediate neighbours; and anti-bonding (AB), back-bonded to first-shell Si atoms. Following a suggestion by Evans and co-workers (Evans *et al.*, 1997), an additional configuration was considered. For this, the H atom lay inside the vacancy cage at an AB position to the TM ion along the trigonal axis. At first sight, it may seem that this configuration can only be possible for TM–H defects, involving a single H atom. We will show that this intuitive statement seems to be indeed correct. For all the PtH_n defects modelled ($n = 1$ to 4), our results, together with available experimental information show that the structures for which the H atoms lie outside the vacancy cage, at AB sites, are the most probable ones.

The fact that the TM impurities are characterised by large covalent radii, provoking *per se* a considerable relaxation of the vacancy cage on trying to reproduce their environment



(a) PtH_1^-: ABout (b) PtH_1^-: ABin

Figure 9.3: Optimised structure of C_{1h} PtH_1^- complexes with the hydrogen atom at anti-bonding (AB) sites, outside (a) and inside the vacancy (b). Cube axes indicate the $\langle 100 \rangle$ directions.

as isolated species, makes the probability of sharing it with more than one H very small. The presence of multiple H atoms would represent further perturbation to the defect's structure arising from the Coulombic repulsive interaction between H atoms and between these and the electrons in the Si dangling-bonds orbitals.

To study PtH_1 and PdH_1 defects, we have considered mainly two configurations: AB with the H atom lying outside the vacancy, which we will refer to as configuration (a), and a second AB configuration, with the H atom lying inside the vacancy cage. These two configurations, (a) and (b) respectively, are shown in Figure 9.3. Both were investigated in trigonal (C_{3v}) and near-trigonal (monoclinic C_{1h}) symmetries. For completeness, a third configuration was able considered (c), in which the hydrogen impurity atom sat at the bond-centred (BC) site.

The lower symmetry structure was set by either moving the TM ion off-site or by reconstructing the bond across Si atoms in the (100) plane. These structures were investigated in three different charge states. The BC configuration (c) was found to be unstable as a result of optimisation under C_{1h} symmetry. The H atom moved from its initial BC lattice position to sit at an AB position back-bonded to Pt along the trigonal axis, *i.e.*, configuration (b) (see Fig. 9.3).

Energetically, our calculations have confirmed the C_{1h} structure of $(PtH_1)^-$ in configuration (a), with a slight reconstruction of the Si–Si bond in the plane perpendicular to trigonal axis with the TM occupying the central cluster site as the lowest energy one, when compared with the trigonal C_{3v} structure or that in which the Pt ion was initially moved off-site. Quantitatively, the former and the latter were found to be degenerate in energy.

The structure optimisation of $(PtH_1)^-$ in configuration (a) showed a collective movement of the atoms lying along the [111] direction, with an outwards movement of the Si atom to which the H atom is bonded and a slight displacement of 0.13 Å by the Pt ion along the trigonal axis away from the central T_d site (Figure 9.3a). The Pt–H distance was 4.53 Å. The outwards relaxation of the Si–H dimer can be understood as result of sp^2

bonding between the Si atom and its three Si nearest-neighbours. Back-bond angles of 119.4 and 119.8 and 119.4ø supported this argument. Despite being four-fold coordinated, the Si atom directly bonded to H (Si–H length of 1.58 Å), and its three immediate Si neighbours formed an almost planar system, further confirming a sp^2 hybridization.

For the relaxed structure of $(PtH_1)^-$ in configuration (b), the impurity atom moved 0.21 Å away from the central site along the $[\bar{1}\bar{1}\bar{1}]$ direction. This is a consequence of the strength of the Pt–H bond. Its length was 1.58 Å. The distance between the Pt ion and the two Si atoms lying in the (110) plane reflected the distortion of the defect. These were equal to 2.54 and 2.48 Å. The Pt–Si length of the bonds between the Pt ion and the other two Si atoms were 2.53 Å. The Pt–H length was 1.58 Å.

Unfortunately, experiment has not been able to make a distinction between the candidate configurations (a) and (b) (Uftring *et al.*, 1995). Our calculations have revealed that the two structures are almost degenerate in energy, with (a) being 12 meV lower than configuration (b) of PtH_1^-. It is, therefore, not possible to invoke energetics and comment on this energy difference, since it is smaller than the predicted error of an LDF-based method (Kohn, 1997). Interestingly, the adiabatic re-orientation barrier is very similar for both configurations. These energy barriers for the hoping of H between equivalent <111> directions were approximately 38 and 50 meV for configurations (a), for which the H atom is back-bonded to a Si atom, and (b), with the H atoms directly bonded to the Pt ion, respectively.

This situation changes for PtH_2, for which there is a wealth of experimental information (Williams *et al.*, 1993; Williams *et al.*, 1994; Uftring *et al.*, 1995; Höhne *et al.*, 1994). For PtH_2^-, we found the C_{2v} AB structure with two equivalent H atoms back-bonded to two Si nearest-neighbours of the TM ion to be the preferred structure (Fig. 9.5). A similar

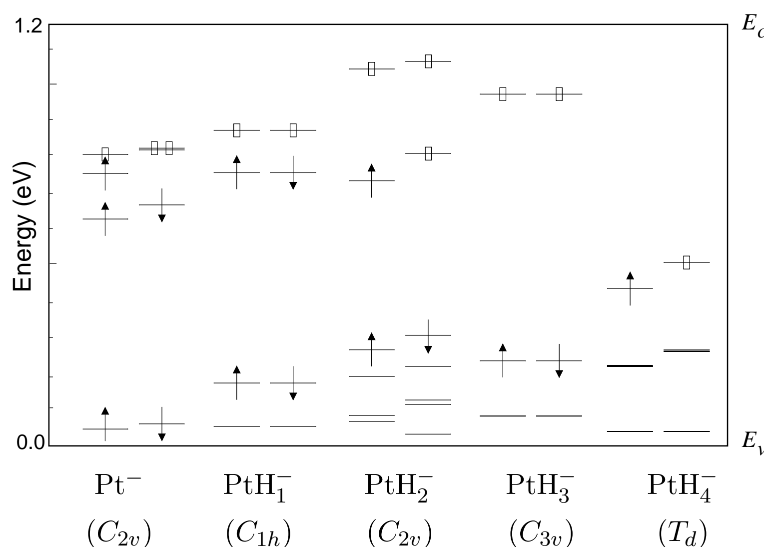


Figure 9.4: Scaled spin-polarised Kohn-Sham eigenvalue spectra of optimised singly negatively charged Pt and PtH defects. Only gap levels are shown.

Table 9.6: Structural parameters for the Pt_s^- and neutral PtH_n centres. Those for PtH_1^- values are given for the two competing configurations. d is the distance of the Pt atom of the centre of the cluster, *i.e.*, a measure of the off-site displacement of the Pt ion from the T_d site. Values in Å.

	Sym.	Pt–Si				Pt–H	H–Si	d
Pt_s^-	T_d	2.477	2.477	2.477	2.477	···	···	0.000
Pt_s^-	C_{2v}	2.441	2.441	2.476	2.476	···	···	0.051
PtH_1^- (a)	C_{1h}	2.483	2.483	2.488	3.008	4.531	1.523	0.134
PtH_1^- (b)	C_{1h}	2.528	2.528	2.480	2.540	1.578	···	0.218
PtH_2^-	C_{2v}	2.451	2.451	2.929	2.929	4.450	1.521	0.000
PtH_3^-	C_{3v}	2.419	2.968	2.968	2.968	4.494	1.526	0.016
PtH_4^-	T_d	2.838	2.838	2.838	2.838	4.367	1.529	0.000

situation was found for PdH_2^-.

The relaxed structure of the di-hydrogen–Pt complex showed that the ion remains on site as a consequence of adding an extra H at the AB position, sitting 1.52 Å away from the Si to which it is bonded and 4.45 Å from the Pt ion. Note that Uftring *et al.* (1995) estimated a value of $\approx$ 4.2 Å for the distance between the Pt ion the H atom, from an analysis of the anisotropic contribution to the hyperfine tensor. These experimental results, together with the results described above, provide the final confirmation of the AB sited model (configuration (a) – Fig. 9.5).

Uftring *et al.* (1995) claimed that for PtH_2^-, the ion should move off-site, due to a residual JT distortion from the isolated Pt_s^- centre, further re-enforced by the presence of two equivalent H atoms. To investigate this suggestion, an extra H was added to the optimised structure of PtH_1^- with monoclinic symmetry, with the Pt ion $\sim$ 0.13 Å away from the T_d site (vide Table 9.6), which was then relaxed, releasing all symmetry constraints. The resulting structure was very similar to the relaxed structure of (C_{2v}) PtH_2^-, with the Pt ion returning spontaneously to the central lattice site. This suggests that the two equivalent H atoms do not enhance the JT distortion responsible for the off-site displacement of the TM ion, but rather inhibit it.

Obviously, the above structure for PtH_2^- is not the only one that might satisfy the defect's symmetry as found experimentally. This is the case of an analogue structure to PtH_1^- in configuration (b) (Figure 9.3). The two H atoms are still equivalent in the (110) plane but now lie inside the vacancy cage. The Pt–H length was 1.71 Å. The H atoms are now much closer and a strong coupling should be expected. Experimentally, the frequency separation between symmetric and anti-symmetric modes is only 13.4 cm^{-1}, strongly suggesting a large separation between the H atoms. This was confirmed by our calculations. For structure (a), in which the two atoms sit outside the first shell cage and

Table 9.7: Calculated and observed hydrogen and deuterium stretch frequencies (cm^{-1}) for C_{2v} $(PtH_2)^q$ complexes in three charge states, q.

Charge state		$q=0$		$q=-1$		$q=-2$	
Config.	Sym.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.
PtH_2	A_1	2081.4	1891.9	1986.2	1901.9	1959.8	1898.0
	B_1	2077.5	1873.1	1986.1	1888.2	1943.5	1889.0
PtHD	A_2	2079.5	1880.3	1986.2	1894.6	1951.7	1893.9
	B_2	1495.2	1361.0	1429.2	1366.9	1400.9	1367.5
PtD_2	A_1	1496.6	1365.2	1429.2	1370.7	1406.9	1363.3
	B_1	1493.8	1352.4	1429.1	1362.5	1395.1	···

are separated by 7.24 Å, the A_1–B_1 mode separation was 0.1 cm^{-1}, while in the case of (*b*) this separation was $\sim$ 74 cm^{-1}, with the two H atoms separated by 2.80 Å (Table 9.7).

The vibrational properties of these PtH defects are discussed below.

9.3.4 Vibrational properties

For the PtH_1 defect, in either configuration (*a*) or (*b*), it is reasonable to assume that only two atoms are involved in the local vibrational mode of the defect, with these being the H atom and its immediate neighbour. Hence, the effect of both the light hydrogen atom and its nearest neighbour, which is a Si atom in configuration (*a*) and the Pt ion for (*b*), assuming that the vibrational modes are harmonic and the neglecting the motion of second and further neighbours of H, can be described by the expression (Leigh & Newman, 1982; Leigh *et al.*, 1994)

$$\omega^2 = \Lambda \left(\frac{1}{m} + \frac{1}{\chi M_N} \right) , \qquad (9.3.1)$$

where m is the mass of hydrogen, M_N is the mass of the nearest-neighbour atom and χ a parameter. Using equation (9.3.1), it is possible to extract the value of M_N from the ratio of the vibrational frequencies for the isotopic substitution of hydrogen by deuterium (D). Analytically,

$$M_N \simeq \frac{\omega_H^2 - \omega_D^2}{2\,\omega_D^2 - \omega_H^2} . \qquad (9.3.2)$$

Here, we have used a typical value for χ, *i.e.*, $\chi = 2$ (Newman, 1973), with m_H and m_D equal to 1 and 2, respectively. This expression will allow us to identify the defect responsible for the vibrational spectra observed by Uftring *et al.* (1995) by comparing the experimental and calculated values of M_N. Assuming that the calculated quasi-harmonic

Table 9.8: Calculated and observed hydrogen and deuterium stretch frequencies (cm^{-1}) for two configurations (see Fig. 9.3) of PtH_1 (C_{1h}) in three charge states, q. The corresponding values of M_N and ω_H/ω_D, are also shown.

	$q=0$			$q=-1$			$q=-2$		
	(a)	(b)	Obs.	(a)	(b)	Obs.	(a)	(b)	Obs.
PtH_1	2034.1	2098.9	1880.7	2022.2	1961.5	1897.2	1910.3	2174.8	···
PtD_1	1464.3	1486.8	1358.5	1453.2	1389.2	1368.8	1373.4	1540.3	···
M_N	13.2	139.2	11.0	14.7	156.3	11.7	14.3	154.1	···
ω_H/ω_D	1.389	1.412	1.384	1.392	1.412	1.386	1.391	1.413	···

vibrational frequencies are indeed harmonic, anharmonicity can be brought into this picture via the expression (Davidson *et al.*, 1993)

$$\nu = \omega - \frac{B}{m}\,, \tag{9.3.3}$$

defining the *anharmonic* frequency. B is a measure of the anharmonicity and ω the 'harmonic' frequency. Fitting this expression to the calculated and observed H-stretch frequencies of PtH_2^-, we found the frequency correction due to anharmonicity to be ~ 40 cm^{-1}. Despite this, the inclusion of such a correction does not qualitatively change the results of this simple method, which are presented next.

It is convenient to first apply equation (9.3.2) to the case of PtH_2 before proceeding to PtH_1. As previously discussed (§9.3.3), the two H atoms in the di-hydrogen–Pt complex sit at anti-bonding lattice positions being back-bonded to two Si neighbours of the Pt ion. Applying the above expression to calculate the mass of the atom directly bonded to the H atom, one should obtain the exact mass of Si, *i.e.*, 28 a.m.u. In a similar way, we should obtain a much larger value for M_N, 195 a.m.u., for PtH_2 in configuration (b). The values of M_N for both the calculated and observed Si–H stretch frequencies of PtH_2 (A_1 and B_1 modes) are given in Table 9.9.

From the values of M_N obtained from the observed stretch modes, it is not obvious that a Si atom is attached to H, since M_N varies between 11 and 15 a.m.u. for the different charge states of PtH_2. Although, if one compares these values with the calculated ones for both structure (a) and (b), it is evident that whatever is the atom responsible for the vibrational modes, it must be the same in both structure (a) and in the defect giving rise to the observed H-related modes. The calculated values of M_N for structure (b) of PtH_2^-, are obtained within $\approx 10\%$ of the mass of Pt.

To overcome the errors due to anharmonicity in the calculation of LVMs, we compare the ratio of the frequencies for corresponding H and D stretch modes for the several defect configurations. These ratios are then used to identify the atomic species other than H, responsible for the vibration. This ratio, ω_H/ω_D, for the isolated Si–H and Pt–H dimers

was calculated to be 1.390 and 1.410 respectively. This small difference in the ω_H/ω_D ration of isolated dimers, is well reproduced by the defect structures. As it can be seen in Table 9.9 for PtH_2 modes and in Table 9.8 for PtH_1, the results point for the (a) structure as responsible for the observed IR modes.

The comparison between the calculated and observed values of M_N and ω_H/ω_D for the PtH_1 defect in different configurations and charge states, clearly points for configuration (a) as that responsible for the observed LVMs of PtH_1 (Table 9.8).

9.3.5 Electrical properties

It is now a fact that the interaction of atomic H with substitutional TM centres, like Pt or Pd, results in a dramatic change of their electrical properties. This can be understood by monitoring the splitting of the gap-lying t_2 manifold of the otherwise undistorted TM centre, assuming a low-spin configuration within a simple one-electron picture, as H atoms are added to the substitutional centre (Fig.9.4).

As described before, the undistorted T_d TM centre gives rise to a single level in the bandgap: a t_2 vacancy-like manifold partially occupied with two electrons for Pt_s^0 or Pd_s^0. For the C_{2v} structure, these two levels are slightly split into a doublet (e) and singlet (a_1). The e level lies above the a_1 one, with their separation reflecting the magnitude of the JT distortion. The electronic configuration is $a_1^{\uparrow\downarrow}\,e^0$ for the C_{2v} defect and $t_2^{\uparrow\uparrow}$ for the undistorted T_d defect. Despite a small a_1–e energy separation, a calculation of the energy levels should reflect this difference in electronic configurations. It is obvious that the ionization of the defect is, energetically, more costly in the case of a T_d structure than it is for the orthorhombic defect.

Our calculations confirm this interpretation as both calculated $(-/0)$ and $(0/+)$ levels of the T_d TM defect lay deeper that those of the distorted structure. For Pt, the $(-/0)$ of the T_d defect is 0.43 eV deeper than that of the C_{2v} configuration, while for $(0/+)$, this separation is 0.12 eV. The calculated levels for the C_{2v} Pd_s were $E_c-0.25$ and $E_v+0.54$ eV

Table 9.9: Calculated and observed hydrogen and deuterium stretch frequencies (cm^{-1}) of the symmetric and anti-symmetric modes for two configurations (see text) of C_{2v} $(PtH_2)^q$ complexes, the corresponding values of M_N and ω_H/ω_D ratios.

	$q=0$			$q=-1$			$q=-2$		
	(a)	(b)	Obs.	(a)	(b)	Obs.	(a)	(b)	Obs.
PtH_2	2081.4	1969.6	1891.9	1986.2	2134.5	1901.6	1959.8	1985.8	1898.0
PtD_2	1496.6	1394.2	1365.2	1429.2	1511.0	1370.7	1406.9	1405.8	1363.3
M_N	14.2	234.2	11.6	13.6	223.9	12.3	15.8	215.1	15.2
ω_H/ω_D	1.391	1.413	1.386	1.390	1.413	1.387	1.393	1.412	1.392

for the $(-/0)$ and $(0/+)$ respectively. It is interesting to note, that despite not predicting a $(+/^+_+)$ for Pt_s in C_{2v} symmetry, our calculations predict a $(+/^+_+)$ for its undistorted T_d configuration at $E_v + 0.09$ eV. Here, we used the $Pd(+/^+_+)$ level as a marker, which is known to possess a level at $E_v + 0.14$ eV (Sachse *et al.*, 1997a).

Adding one H atom to the distorted Pt_s^0 defect, further splits the e doublet filled for $S = \frac{1}{2}$. Due to the fact that the overall lowering of the t_2 manifold is compensated by the additional splitting of the e level due to the extra proton being added to the defect, the difference from the levels of PtH_1, or PdH_1 and those of substitutional T_d defects is quite small. The consequent deepening of the $(-/0)$ level could make the existence of a second acceptor level possible. This does seem to apply to PdH_1 and PtH_1, as no $(=/-)$ levels were found. The corresponding donor levels are located at $E_v + 0.45$ and $E_v + 0.42$ eV for PtH_1 and PdH_1, respectively. However, so far, no donor levels have been detected for the PtH and PdH defects containing a single hydrogen atom.

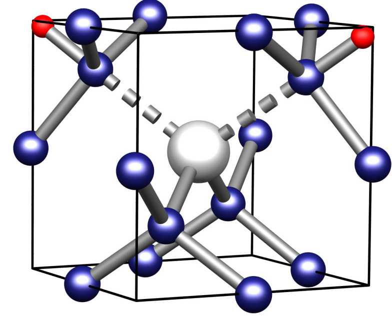


Figure 9.5: Optimised structure of C_{2v} PtH_2^- defect with the hydrogen atom at anti-bonding (AB) sites. The Si–H length is 1.52 Å and the distance between the TM ion and the hydrogen impurity atom is 4.53 Å.

A second H atom results in an additional electron occupying the e-manifold, which is now completely split, with one of the resulting singlet levels being fully occupied and the other empty and consequently higher in energy. Thus there is only a slight downwards shift in energy of the acceptor levels from PtH_1 to PtH_2, a spin-compensated $S = 0$ system. The $(0/+)$ level lies now in the valence band and the $(-/0)$ drops by ~ 0.1 eV to lie at $E_c - 0.42$ eV for PtH_2, and 0.04 eV for PdH_2 being now at $E_c - 0.36$ eV. It is interesting to note that the position of the $(=/-)$ level does not change appreciably going from tetrahedral Pd_s, to PdH_1 and PdH_2 (Table 9.10). All these levels are quite shallow, with that of T_d Pt_s located 0.11 eV below E_c, very close to the CB for monoclinic PtH_1, in configuration (b), and $E_c - 0.05$ eV for PdH_2.

For PtH_3^0, the highest spin-up defect singlet is now occupied. A transition $PtH_3^0 \rightarrow PtH_3^-$ would mean the filling of the e level, making impossible the capture of a second electron by the defect. A second acceptor can only arise from a new state entering in the bandgap. A similar situation occurs for the single acceptor level, $(-/0)$ of PtH_4. For PtH_3 and PdH_3, the $(-/0)$ level are calculated to lie 1.35 and 1.15 eV below E_c, revealing a high electron affinity to fill the last defect level. Note that the error inherent to our method increases with depth.

The Kohn-Sham eigenvalue spectrum of PtH_4^0 revealed an almost level-free bandgap, with the exception of a singlet level lying close to the conduction-band bottom. The defect-related levels lay very close to the valence band edge. Obviously, again like in the case of PtH_2 and PtH_3, no donor levels were found. The shallow one-electron level found in the

upper part of the bandgap is probably a delocalised 5s and 6s levels of the TM impurities. According to results, this s-level is responsible for very deep, $(=/-)$ and $(-/0)$ levels, of PdH_3 and PdH_4 respectively (see Table 9.10). This means that both PtH_4 and PdH_4 are *not* electrically passive.

Comparison with previous assignments

As a result of the DLTS studies on hydrogenated Pt- and Pd-doped silicon described above, tentative assignments have been made for the observed deep levels (Sachse *et al.*, 1997b). Very recently, some of these assignments were refined using Yarykin's model (Sachse *et al.*, 1999a). With a few exceptions, our results agree well with the experimental predictions (see Table 9.10 and Figure 9.3).

The calculations show that not only does PtH_1 possess a $(-/0)$ level at $E_c - 0.56$ eV in good agreement with the 0.50 eV E(250) level, but also a deep donor level at $E_v + 0.45$ eV. There is no evidence for this level which might be obscured by the Pt $(0/+)$ level at $E_v + 0.30$ eV. Unfortunately, experiment does not confirm this explanation as the $(0/+)$ of Pt showed a similar depth profile to that of the Pd$(+/{}^{+}_{+})$ level. Alternatively, the

Table 9.10: Comparison between calculated and experimental electrical levels of Pt and Pd–H defects (eV). $(0/+)$ and $(+/{}^{+}_{+})$ are referred to E_v and, $(-/0)$ and $(=/-)$, to E_c.

	$(+/{}^{+}_{+})$		$(0/+)$		$(-/0)$		$(=/-)$	
	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.
Pt (C_{2v})	0	0.09	0.22	0.33	0.32	0.23	0.12	···
Pt (T_d)	0.09		0.65		0.44		0.24	
PtH_1 (b)	0	···	0.45	···	0.55	0.50 (E250)	0.03	···
PtH_1 (a)	0		0.45		0.56		0.24	
PtH_2	0	···	0	···	0.42	0.76 (H210)	0.15‡	0.16 (E90)
PtH_3	0	···	0	···	1.35†	0.86 (H150)	0	···
PtH_4	0	···	0	···	0.97	···	0.61	···
Pd (C_{2v})	0.14‡	0.14	0.54	0.33	0.25	0.22	0.11	···
PdH_1	0	···	0.42	···	0.45	0.43 (E200)	0	···
PdH_2	0	···	0	···	0.36	0.61 (H280)	0.05	0.18 E(60)
PdH_3	0	0.07 (H45)	0	0.08 (H55)	1.15†	0.92 (H140)	0	0.29 E(160)♮
PdH_4	0	···	0	···	0.71	···	0.34	···

† The method has a larger error when the predicted level is further from that of the standard.

‡ This is a standard level, as discussed in the text.

♮ This assignment is tentative.

calculated level may be in error by ~ 0.2 eV and the actual level falls below E_v.

The injection of electrons monitored by MCTS in p-type material (Sachse *et al.*, 1997b) has showed that E(250) and H(210) do not belong to the same defect, which points for a PtH_2 complex as responsible for the midgap level E(250). Although, it is important to note that our calculations indicate that PtH_1 also possesses a single acceptor level, in close proximity to E(250). This might suggest that the E(250) DLTS peak is due to electron emission from two close-by levels originated by two distinct PtH defects involving one and two hydrogen atoms. A similar situation has been observed by Laplace DLTS (LDLTS) for the case of the G4 trap in hydrogenated gold-doped silicon (Sveinbjörnsson & Engström, 1995). Laplace DLTS was able to separate three defects with very similar carrier-emission characteristics (Deixler *et al.*, 1998; Rubaldo *et al.*, 1999).[1]

It is therefore necessary that a LDLTS analysis be performed on the electron trap E(250). This is of paramount importance since theory and experiment strongly suggest that this level may be responsible for the control of the lifetime of charge carriers.

The $(=/-)$ level at $E_c - 0.18$ eV due to PtH_2 is used as a standard by the theoretical calculations. Experimentally, it seems unclear what is the character of the E(90) trap. Early DLTS studies have identified this level as a *donor* observed in n-type samples (Sachse *et al.*, 1997b), while recently this same level was reported as a double acceptor (Sachse *et al.*, 1999a). Despite this fact, combined EPR and LVM spectroscopic studies (Williams *et al.*, 1994; Uftring *et al.*, 1995) do not support the argument that E(90) is the double acceptor of PtH_2, as this level lies deeper than $E_c - 0.1$ eV — the position of the $(-/0)$ level of C_i — at 0.18 eV below E_c.

H(210) at $E_c - 0.76$ eV is now assigned to the $(-/0)$ level. The calculated level at $E_c - 0.42$ eV is about 0.3 eV too shallow and this discrepancy originally suggested to us that H(210) was the $(0/+)$ level of PtH_1 (Jones *et al.*, 1999b). There should be an associated $(=/-)$ level but the lack of any firm experimental evidence for the $(=/-)$ level of PtH_1 from the IR-absorption studies (Stavola, 1998a), appears to rule this out. Nevertheless, there should be a Poole-Frenkel effect associated with H(210) which has not, as far as we are aware, been reported. There is no donor level associated with PtH_2.

The theoretical results show clearly that PtH_3 has a deep acceptor level in the lower half of the band gap in agreement with the depth and character of H(150). Thus the assignment of this defect agrees with the conclusions of the DLTS profiling studies (Sachse, 1997).

As previously observed experimentally (Sachse *et al.*, 1997a), our calculations confirm a similar level structure for the several PdH complexes to those of PtH. In a similar way to platinum in hydrogenated samples, Pd can also act as a recombination as a result of its interaction with atomic hydrogen. As described above for Pt, two PdH complexes involving a different number of H atoms have levels around midgap. These lie at $E_c - 0.45$ and $E_c - 0.36$ eV, for PdH_1 and PdH_2 respectively (Table 9.10). No donor was found for PdH_2 in our calculations.

[1] The results of an extensive study on the electrical properties of gold-hydrogen defects are presented and discussed in detail in the next chapter of this thesis.

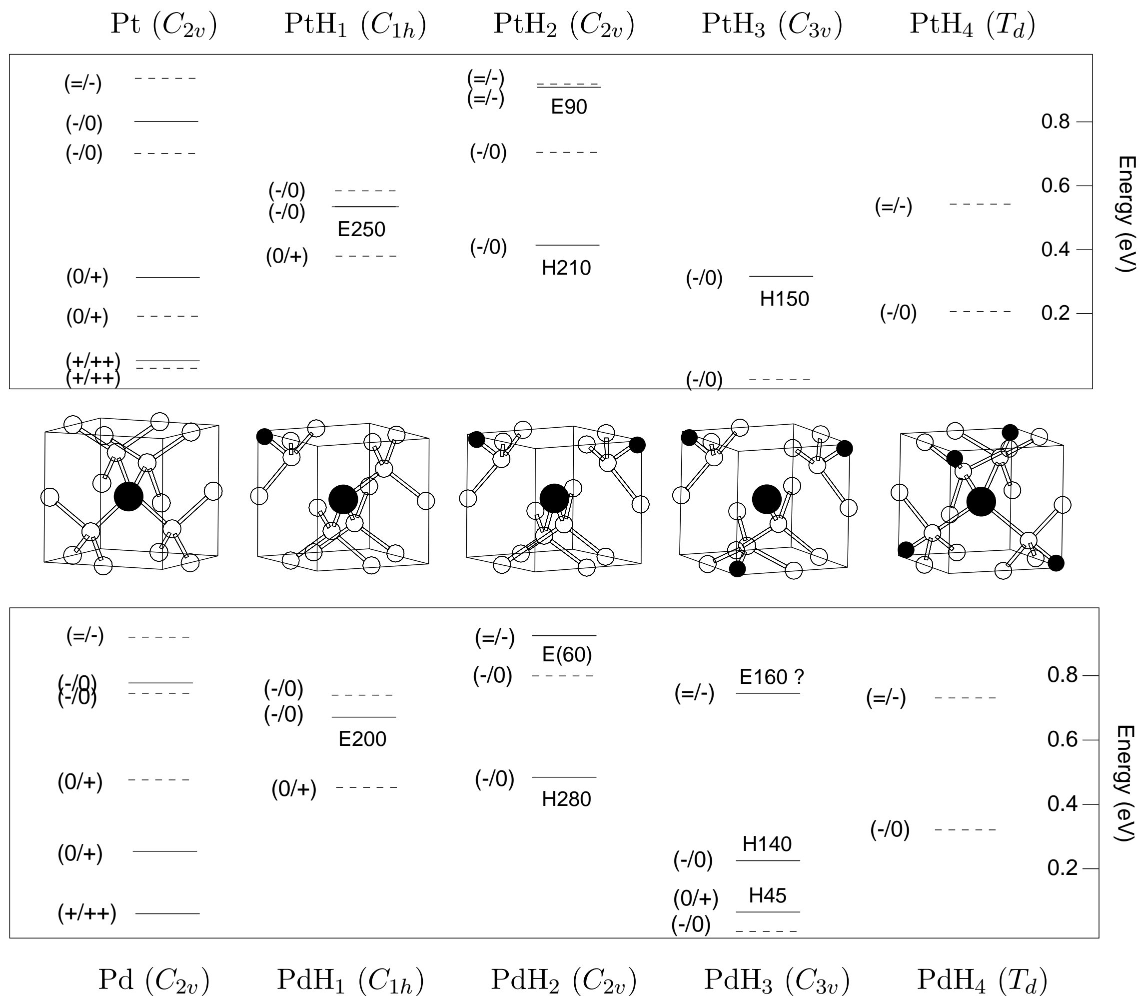


Figure 9.6: Structure and corresponding electrical levels of platinum and palladium and their complexes with hydrogen. Experimental and calculated levels are represented by solid and dashed lines, respectively.

9.3.6 Electrical passivation of Pt and Pd centres

As described previously, it is now clear that the electrical activity of Pt and Pd cannot be neutralised by the simple interaction with atomic hydrogen. However, our calculations have shown that molecular di-hydrides of Pt and Pd are passive when inserted into lattice hydrogen-induced micro-voids. In the present case, these voids were simulated by a deca-vacancy ($V_{10}H_{16}$) embedded by 148-atom Si H-terminated cluster.

Transition-metal ions have large atomic radius being readily attracted to voids in the crystal. This characteristic has been used to improve gettering techniques to remove TM from active regions of a device (Mohadjeri *et al.*, 1995; Meyers *et al.*, 1996; Kinomura *et al.*, 1998).

This model for the electrical passivation of Pt and Pd explains well the results of Pearton and Haller (Pearton *et al.*, 1983). It is known that the hydrogenation process used by these workers is quite aggressive, resulting in the creation of aggregates of vacancies in the near-surface region of the samples. For the case of incorporation of H via a

chemical wet-etching, this interpretation is more difficult. To our knowledge, there is no experimental evidence for the formation of aggregates of vacancies as a result of chemical etching. Despite this fact, the formation of lattice vacancies as a consequence of a Frank-Turnbull mechanism, cannot be ruled out.

9.4 Final remarks

Our calculations strongly suggest that two distinct PtH complexes, involving one and two H atoms give rise to two close-by levels around midgap (Fig. 9.6). These two levels, and not one as previously reported, may be responsible for the control of the lifetime of both electrons and holes in power devices. For both defects, the H atoms lie at anti-bonding sites, being back-bonded to Si nearest-neighbours of the Pt ion. Our calculations also show that the proposed Pt–Pt and Pt–O defects do not give rise to any levels around midgap. A model for the passivation of Pt and Pd centres is proposed. The similarity between the electrical properties of PtH and PdH defects is confirmed.

10

Gold and Silver Defects

10.1 Introduction

Gold as an impurity in crystalline silicon is many respects one of the most throughly studied defect system in any semiconductor. It is one of the most common elements in Si technology, used for electrical contacts. Additionally, it is also used, together with silver, as an effective recombination centre. This is the case of special devices, like fast-switching diodes, into which the TMs are intentionally incorporated for the control of the lifetime of charge carriers. Despite this fact, little is known about the microscopic properties of the defects responsible for electrical activity showed by Au- and Ag-doped Si material. The fact that the literature on the electrical properties of silver in silicon it rather limited relates to the very high diffusion temperature, which is needed to achieve to incorporated noticeable concentrations of electrically active silver. Nowadays, gold or silver contamination is not really a problem, when compared to other common impurities like iron, copper or nickel, since its concentration rarely exceeds 10^{12} cm^{-3} (Graff, 1995).

The interaction of hydrogen with transition metals, and in particular gold, has recently become of great interest (Pearton & Tavendale, 1982; Veloarisoa *et al.*, 1991; Williams *et al.*, 1993; Sveinbjörnsson & Engström, 1995; Evans *et al.*, 1997; Evans *et al.*, 1999; Weinstein & Stavola, 1999). The presence of hydrogen has three effects on the electronic levels of the TM impurity. It can shift these levels, introduce additional ones, or it can remove them completely from the band gap achieving passivation (Pearton & Tavendale, 1982). However, in no case are these effects completely understood. For example, capacitance transient studies have shown that, either electrically active or inactive defects can be generated when hydrogen interacts with Au. The former with the work of Sveinbjörnsson *et al.* (1995) by incorporating H into the samples by RT wet-etching and the latter, with the work of Pearton and Tavendale (1982) by exposing the samples to a remote H-plasma at very high temperatures ($\sim 1200\,^{\circ}$C). Little is known about the defect structures responsible for such distinct, and technological and scientifically important, features revealed by these transition-metal–hydrogen complexes.

10.2 Experimental background

10.2.1 Structural properties as isolated centres

In the framework of the vacancy model (Watkins, 1983; Anderson *et al.*, 1992b), the gold centre should show similar properties to those of isoelectronic and substitutional Pt_s^- or Ag_s^0. Again, similarly to Pt_s^-, its structural properties arise from the vacancy-like t_2 partially filled t_2 manifold with three electrons (Fig. 10.1). The consequent Jahn-Teller distortion was found to drive the impurity atom off-site along the [100] direction, giving rise to an intense C_{2v} EPR signal, which has been fully understood (Anderson *et al.*, 1992a; Williams *et al.*, 1993). An orthrhombic-I EPR spectrum ($g_\perp \approx g_\parallel \approx 2.0$) has been report as due to Ag_s^0 (Hai *et al.*, 1997). Despite these facts, the study of Au_s^0 by EPR, one of the most powerful tool to determine the defect's structure, has not produced any relevant results. Several gold-related defects have observed using EPR, but all the conclusions refer to a complex structure, with gold probably pairing off with another impurity, and not to the isolated species (Woodbury & Ludwig, 1962; Höhne, 1980; Lang *et al.*, 1980; Höhne, 1982; Kleinhenz *et al.*, 1981; Alteheld *et al.*, 1994).

Figure 10.1: Band-gap states for Au^0 (a) T_d symmetry, (b) after tetragonal JT distortion, (c) plus a trigonal t_{2z} JT distortion and (d) plus spin-orbit interactions, after Watkins and Williams (1995).

Overcoming this difficulty, spectroscopic Zeeman studies have shown that the ground state of substitutional gold splits according to an effective spin $S = \frac{1}{2}$ (C_{2v}), with $g_\parallel \approx 2.8$ and $g_\perp \approx 0$ (Watkins *et al.*, 1991; Kleverman *et al.*, 1995). Due to this unexpectedly high anisotropy ($g_\perp \approx 0$), the transition probability for microwave transitions becomes very low and the EPR spectrum is difficult to be observed. This makes this centre invisible to EPR. Recently, this vanishing of the diagonal g_{xx} and g_{yy} components of the hyperfine g tensor was explained as due to the increasing importance of the spin-orbit interaction, which tend to cancel the off-centre JT displacement of the Au_s impurity, when compared

with that of atomic Pt^-.

10.2.2 Interaction with atomic hydrogen

For substitutional gold, Au_s^0, and the isoelectronic silver centre, Ag_s^0, there are three electrons in the one-electron t_2 manifold, which makes this defect Jahn-Teller active. The ground-state configuration is $b_2^{\uparrow\downarrow}\, b_1^{\uparrow}\, a_1^0$ (Resende *et al.*, 1999) resulting from the splitting of the triply degenerate t_2 level, in agreement with the vacancy model.

10.2.3 FTIR measurements

Several hydrogen-vibrational bands have been identified around ~ 1800 cm^{-1} in Si containing hydrogen and gold impurities (Evans *et al.*, 1997; Evans *et al.*, 1999; Weinstein & Stavola, 1999).

Two bands at 1785.6 and 1803.3 cm^{-1} were identified as H-stretch vibrations of a AuH_2 complex containing two equivalent hydrogen atoms (Evans *et al.*, 1997; Evans *et al.*, 1999). Showing the characteristic frequency shift of $\sim \sqrt{2}$, two other bands at 1292.1 and 1304.4 cm^{-1} were observed, and assigned to the anti-symmetric and symmetric-stretching modes of AuD_2. This pair of bands revealed the expected behaviour of a C_{2v} structure, under externally applied uniaxial stress (Weinstein & Stavola, 1999). Based on the variation of their relative intensities with the position of the Fermi level, the lines at 1827.1–1813.3 cm^{-1} were found to arise from different charge states of gold-hydrogen complex containing a single hydrogen atom. Uniaxial stress studies of the splittings of the transition energies and the stress induced alignment of this AuH_1 complex were consistent with a defect with C_{1h} or lower. It was then proposed that Au ion goes off-site along the [100] direction to explain the departure from the expected C_{3v} symmetry (Evans *et al.*, 1999).

As pointed by the authors, all the detected hydrogen and deuterium modes in hydrogenated Au-doped Si have been found to be very similar to those assigned to the PtH corresponding structures, with the latter being slighty higher in frequency by ~ 70 cm^{-1} for all the observed modes, as reported by Uftring *et al.* (1995).

10.2.4 Electrical properties

It is now an accepted fact that isolated gold and silver are amphoteric subsitutional impurities giving rise to two deep levels within the Si bandgap. These are a $(-/0)$ level at $E_c - 0.55$ and $(0/+)$ at $E_v + 0.35$ eV for Au_s (Collins *et al.*, 1957; Tasch & Sah, 1970; Brotherton & Bradley, 1982b; Mesli *et al.*, 1987; Petersen & Nielsen, 1990), and $(-/0)$ level at $E_c - 0.55$ and $(0/+)$ at $E_v + 0.38$ eV for Ag_s (Yau *et al.*, 1984; Lemke, 1984; Sachse, 1997). A major contribution to the current knowledge of the level stucture of isolated gold came from the radioactive transmutation study of Petersen and Nielsen (1990). It was demonstrated that in the radioactive decay of ^{195}Au into ^{195}Pt, the $E_c + 0.55$ and $E_v + 0.35$ eV levels convert in 1:1 fashion to an $E_c - 0.22$ eV acceptor level, which has been well established to arise from isolated Pt_s.

Table 10.1: Activation energy ΔE, majority capture cross sections σ, activation energy of the capture cross section E_σ, the enthalpy ΔH and the entropy factor X_s for Au and Ag energy levels. Values taken from Mesli *et al.* (1987) and Sachse (1997), for Au and Ag respectively.

	Level	ΔE (eV)	$\sigma_{n,p}$ (cm^2)	E_σ (eV)	ΔH (eV)	X_s
Au	$(-/0)$	$E_c - 0.55$	1×10^{-16}	0	$E_c - 0.23$	51
	$(0/+)$	$E_v + 0.35$	2×10^{-15}	0	$E_v + 0.35$	26
Ag	$(-/0)$	$E_c - 0.55$	1×10^{-16}	0	$E_c - 0.55$	35
	$(0/+)$	$E_v + 0.38$	4×10^{-15}	0	$E_v + 0.38$	49

The accepted values for the activation energy, majority capture cross sections, activation energy of the capture cross section, enthalpy and the entropy factor for Au and Ag energy levels are given in Table 10.1.

Table 10.2: Activation energy ΔE, majority capture cross sections σ, activation energy of the capture cross section E_σ and the enthalpy ΔH for a number of AuH and AgH electrical levels. Note the similarities between the reported levels of Au- and Ag-H defects.

	Level	ΔE (eV)	$\sigma_{n,p}$ (cm^2)	E_σ (eV)	ΔH (eV)
AuH related	G1	$E_c - 0.20$	$8 \times 10^{-18\dagger}$	0.02	$E_c - 0.18$
	G4	$E_c - 0.54^\ddagger$	?	0	$E_c - 0.56$
	G2	$E_v + 0.21$	0.20×10^{-15}	0	$E_v + 0.21$
	G4$'$	$E_c - 0.58^\ddagger$	?	0	$E_c - 0.58$
	G3	$E_v + 0.47$	50×10^{-15}	0	$E_v + 0.47$
AgH related	E3	$E_c - 0.09$	$0.67 \times 10^{-15\dagger}$	0.0019	$E_c - 0.09$
	E2	$E_c - 0.45$	0.24×10^{-15}	0	$E_c - 0.45$
	H2	$E_v + 0.28$	6.6×10^{-15}	0	$E_v + 0.28$
	E6	$E_c - 0.50$	?	0	$E_c - 0.50$
	H3	$E_v + 0.38$	2.1×10^{-15}	0	$E_v + 0.40$

† Temperature dependent. The values of ΔE and ΔH for the G1 and H3 levels are given from extrapolation to zero field. These and the other values are taken from Sveinbjörnsson and Engström (1995) and Sachse (1997) for the AuH levels, and Yarykin *et al.* (1999) for AgH related ones; ‡ Laplace DLTS measurements of Rubaldo *et al.* (1999).

Sveinbjörnsson and co-workers (Sveinbjörnsson *et al.*, 1994; Sveinbjörnsson & Engström, 1995), confirming the amphotheric behaviour of the Au centre, have reported four *new* deep traps in Au-doped Si into which H was introduced by wet etching. Stable up to 150 °C, these were labelled G1–G4 in the DLTS spectra. G1, G4 and G2 were tentatively assigned to the *same* AuH complex, involving a single H atom. G1 is an electron trap located at $E_c - 0.19$ eV and G2 a hole trap at $E_v + 0.21$ eV. G4 was reported in both n- and p-type material, exhibiting similar electron capture and emission characteristics to the single acceptor of Au (Table 10.2). Due to this fact the position of the G4 trap could not be determined accurately. This picture was confirmed by MCTS studies (Davidson & Evans, 1996b) combined with DLTS in p-type Si. This optical transient capacitance technique monitors majority and minority carrier emission and capture processes within the same type of sample. The character of the G3 level (observed in p-type Si at $E_v + 0.47$ eV) is still matter of discussion (Davidson & Evans, 1996b; Sachse *et al.*, 1999a). Davidson and Evans (Davidson & Evans, 1996b) have argued that the G3 peak might arise from two levels, lying closely within the bandgap, exhibiting similar hole capture cross sections but different values for electron capture cross section.

Very recently, a high resolution DLTS study (Laplace DLTS) was able to successfully resolve three distinct levels in the region of the Au(−/0)/G4 activity (Deixler *et al.*, 1998; Rubaldo *et al.*, 1999). Detected as a single peak by DLTS, LDLTS was able to decompose G4 into three close-by delta-like peaks. These were the (−/0) level of Au_s, G4, a (−/0) level at $E_c - 0.54$ eV, and a third, G4′, at $E_c - 0.58$ eV. The acceptor character of G4 was confirmed. Tentatively, it was suggested that G4′ may be related to a AuH complex involving more H atoms than that responsible for G1, G2 and G4.

The DLTS study of Yarykin *et al.* (1999) on Ag-doped Si into which hydrogen was introduced via wet-etching, has provided strong evidence for the similarity between AuH and AgH complexes. This resemblance can be understood in terms of Watkins' vacancy model. According to this model, these two impurities, obviously being chemically similar, should both behave like a *black-box* being rather insensitive to their crystalline surroundings.

10.3 Results

10.3.1 Cluster and basis

The cluster and basis used for the study of the gold- and silver-related centres were similar to those used for the investigation of the properties of Pt and Pd defects, described in the previous chapter. The exception is obviously the basis used for the gold and silver atoms. Details are given in Table 10.3.

10.3.2 Isolated gold and silver centres: structural properties

Neutral Au and Ag centres have a electronic configuration $t_2^{\uparrow\uparrow\uparrow}$ and are slightly distorted T_d defects with the impurity moving 0.03 Å along [100] for Au and 0.01 Å for Ag. The resulting C_{2v} symmetry is consistent with EPR experiments on Ag (Són *et al.*, 1992; Hai

Table 10.3: Atomic basis and basis expansion type used for the calculations on Au and Ag–H defects. N and M denoted the number of orbitals used to describe the wavefunction and charge density respectively. One additional basis function was placed midway between all bonded pairs for the electrical level calculations.

Cluster	Atom	N	M	Basis expansion	`chden`
$Si_{71}H_{60}$	Au	6	12	big	0
	Ag	7	14	big	
	Si	5	4	all big	
	H	3	2	all minimal	
Bond-centres	all bonds	2/3	2/3	big	

et al., 1997) and optical absorption studies on Au (Watkins *et al.*, 1991; Kleverman *et al.*, 1995). The Au–Si and Ag–Si lengths are all between 2.6 and 2.7 Å. We also find a small adiabatic reorientation barriers of 0.11 eV for Au_s^0, and 0.14 eV for Ag_s^0. These are upper limits to the barrier and tunnelling probably plays an important role as reorientation can occur at cryogenic temperatures (Watkins *et al.*, 1991).

It is interesting to note that Au_0 and Ag_0 seem to produce a larger enlargement of the vacancy cage as substitutional impurities, due to outwards breathing movement of the TM impurity's Si neighbours, than the Pt^- or Pd^- (Table 10.4).

10.3.3 Gold and silver multi-hydrogen defects: Structural and vibrational properties

We now apply this theory to the complexes of substitutional TM impurities with hydrogen. There are three likely positions for the H atom. These are (a) when H sits at an anti-bonding, AB, site to a Si neighbour of the TM impurity, (b) AB sited to the TM impurity, and (c) bond centred, BC, sited between the impurity and Si. Configurations (a) and (b) were similar to those considered in the case of the platinum mono-hydride (see Fig. 9.3). For the AuH_1 defect, configurations (b) and (c) are less stable than (a) by 0.23 and 0.47 eV respectively. This result is sensitive to basis size and a smaller basis reverses this ordering (Resende *et al.*, 1997). The H stretch vibrational frequencies for the three configurations are given in Table 10.5 along with experimental results (Evans *et al.*, 1997; Evans *et al.*, 1999). Agreement is best for the configuration (a) but the sense of the small shifts arising with different charge states are not reproduced. The H-reorientation barrier among the equivalent $\langle 111 \rangle$ directions is 0.41 eV for (a) and 0.23 eV for (b). The observed barrier is athermal and presumably must proceed by a tunnelling mechanism.

For structures (a) and (c), the shift in frequency on replacing ^{28}Si with ^{29}Si is $\lesssim 1\ cm^{-1}$ for each charge state. For configuration (b) where H is bonded to the impurity, the

Table 10.4: Structural parameters for the Au_s and neutral AuH_n centres. Those for AuH_1 values are given for the two competing configurations. d is the distance of the Au atom of the centre of the cluster, *i.e.*, a measure of the off-site displacement of the Au atom from the T_d site. Values in Å.

	Sym.	Au–Si				Au–H	H–Si	d
Au_s	T_d	2.702	2.702	2.702	2.702	···	···	0.000
Au_s	C_{2v}	2.687	2.678	2.732	2.732	···	···	0.036
AuH_1 (a)	C_{1h}	2.718	2.719	2.719	3.250	4.750	1.500	0.183
AuH_1 (b)	C_{1h}	2.706	2.706	2.629	2.751	1.750	···	0.023
AuH_2	C_{2v}	2.657	2.657	3.145	3.145	4.647	1.503	0.175
AuH_3	C_{3v}	2.628	3.094	3.094	3.094	4.594	1.501	0.146
AuH_4	T_d	2.997	2.997	2.997	2.997	4.499	1.502	0.000

shift is negligible. Thus, although the energy and vibrational mode calculations favour configuration (a), the observed low reorientation barrier favours configuration (b).

For AuH_2 and AgH_2 defects with C_{2v} symmetry, configuration (a) is more stable than (b) or (c). This structure is the same as that suggested by magnetic resonance on PtH_2 defects (Williams *et al.*, 1993). As shown in the previous chapter, our theoretical calculations provide strong support for the model consistent with the anti-bonding arragement of the two hydrogen impurity atoms (see Fig. 9.5). AuH_2 defects have been detected by FTIR absorption studies (Evans *et al.*, 1997; Evans *et al.*, 1999), as described previouly in this chapter. For the configuration (a), Table 10.6 shows that the separation in frequencies

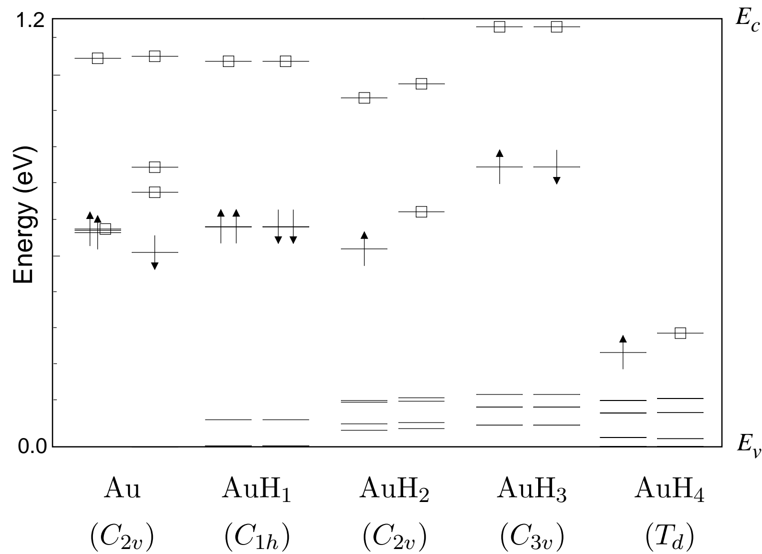


Figure 10.2: Scaled spin-polarised Kohn-Sham eigenvalue spectra of optimised neutral Au and AuH defects. Only gap levels are shown.

Table 10.5: Calculated and observed hydrogen stretch modes for three configurations (see text) of C_{1h} AuH_1 in three different charge states. Values in cm^{-1}.

Charge state	$q = 0$		$q = -1$		$q = -2$	
Config.	AuH	AuD	AuH	AuD	AuH	AuD
Expt.	1787.7	1292.9	1813.3	1310.9	1827.1	1319.4
Config. (a)	1947.4	1401.1	1890.5	1359.6	1893.2	1361.4
Config. (b)	2014.5	1427.3	1980.0	1402.8	2118.0	1500.2
Config. (c)	2419.0	1726.2	2500.8	1783.5	2454.3	1748.9

between the symmetric (A_1) and anti-symmetric (B_1) modes is much closer to experiment than the separations in configurations (b) and (c).

10.3.4 Electrical levels

The electrical levels are determined assuming that a low spin state results from the addition of each H atom. The addition of one H atom lowers and splits the t_2 level into a filled a_1 level lying below a half-filled e level. The lowering of the t_2 manifold is off-set by the splitting so that the differences in the donor and acceptor levels from those of the TM impurity are quite small. Table 10.7 gives these levels for configuration (a). The $(=/-)$ levels of AuH_1 and AgH_1 lie at $E_c - 0.22$ and $E_c - 0.36$ eV respectively. The levels found for configuration (b) are very similar. For example, the $(0/+)$, $(-/0)$ and $(=/-)$ levels of AuH_1 lie at $E_v + 0.37$, $E_c - 0.76$ and $E_c - 0.36$ eV respectively.

A second H atom, added in configuration (a), results in an additional electron occupying the e manifold (Fig. 10.2) which is pushed downward and splits with the upper level being occupied. Thus once again there are only small shifts in the donor and acceptor levels on

Table 10.6: Calculated and observed hydrogen and deuterium stretch frequencies (cm^{-1}) for three configurations (see text) of AuH_2 (C_{2v}) complexes.

	AuH_2		AuHD		AuD_2	
Mode	A_1	B_1			A_1	B_1
Expt.	1803.3	1785.6	1792.5	1298.6	1304.4	1292.1
Config. (a)	1974.7	1970.5	1972.6	1419.2	1420.6	1417.8
Config. (b)	1988.4	1743.0	1884.6	1303.6	1406.2	1237.5
Config. (c)	2060.6	2013.6	2037.6	1461.9	1475.5	1449.2

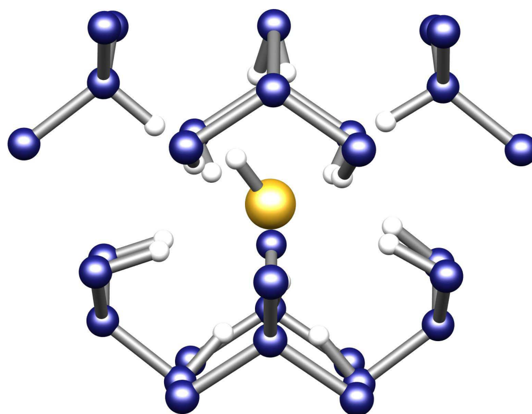


Figure 10.3: A gold mono-hydride molecule in a hydrogen-induced microvoid. The void was simulated by a H-terminated deca-vacancy ($V_{10}H_{16}$) embedded in a 148-atom cluster.

going from say AgH_1 to AgH_2. The e-manifold is filled for $(AuH_2)^-$ and $(AgH_2)^-$, and a second acceptor level can only arise from a new state entering the gap. The calculations give no hint for such states and hence the di-hydrogen defects do not possess $(=/-)$ levels.

Adding a third H fills the e manifold which is now pushed below E_v. Thus AgH_3 and AuH_3 do not possess any donor levels. However, it appears that an empty level, due to the 5s and 6s levels of Ag and Au respectively, creeps into the band gap. We place the resulting $(-/0)$ levels of AgH_3 and AuH_3 at 0.13 and 0.26 eV below E_c.

In conclusion, the calculations show that AuH_n and AgH_n, $n = 1, 2$ defects, possess $(0/+)$ and $(-/0)$ levels close to Au and Ag, while only the monohydrides possess $(=/-)$ levels. The trihydride defects possess shallow acceptor levels.

DLTS studies have led to a number of levels assigned to AuH_n and AgH_n defects. There are two uncertainties in identifying these levels. Firstly, the number of H atoms associated with each level and secondly its character. Annealing and defect profiling studies have shown that the G1 (at $E_c - 0.19$ eV), G4 and G2 levels due to Au-hydrogen arise from the same defect (Sveinbjörnsson *et al.*, 1994). The location of the levels (see Table 10.2) and their emission cross-sections suggest that they correspond to $(=/-), (-/0)$ and $(0/+)$ respectively. This has been supported by minority carrier transient spectroscopic measurements (Evans *et al.*, 1997). The deep penetration of these defects is taken to imply that they possess one hydrogen impurity atom. Similar considerations show that in the Ag case the E3 (at $E_c - 0.09$ eV), E2 and H2 levels are due to $(=/-)$, $(-/0)$ and $(0/+)$ levels of AgH_1 (Yarykin *et al.*, 1999). Such assignments agree with our calculations (Table 10.7). We can exclude AuH_2 and AgH_2 as being responsible for these levels as these defects do not possess a $(=/-)$ level.

It is obvious from Table 10.7 that the levels of the Au- and Ag-hydrogen defects G3 and H3 are close to the calculated donor levels of AuH_2 and AgH_2 respectively. Using the model for the determination of the number of hydrogen impurity based on the depth dependence of the concentration of the defects (§9.2.1), Yarykin *et al.* (1999) argued that these levels should be assigned to AuH_2 and AgH_2. There has been some changes of view

as to their character (Sachse *et al.*, 1999a) but the most recent opinion (Yarykin *et al.*, 1999) is that they are donor (0/+) levels. This agrees with our calculations. Our results also imply that the corresponding (−/0) levels are very close to those of AuH_1 and AgH_1 respectively. Evidence for these (−/0) levels has been more difficult to obtain. However, an early DLTS (Parakhonskií *et al.*, 1996) and a recent Laplace-DLTS study (Deixler *et al.*, 1998) on Au-doped Si, show a level G4′ at $E_c - 0.58$ eV while the difference in the concentrations of E2 and E3 in the surface region is taken to imply the existence of the (−/0) level of AgH_2 labeled E6 (Yarykin *et al.*, 1999). The G4′ and E6 levels are close to the calculated (−/0) levels.

Table 10.7: Comparison between calculated and experimental electrical levels of Au– and Ag–H defects (eV). (0/+) and $(+/^{+}_{+})$ are referred to E_v and, (−/0) and (=/−), to E_c.

	(0/+)		(−/0)		(=/−)	
	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.
Au	0.21	0.35	0.66	0.56		
AuH_1	0.36	0.21 (G2)†	0.62	0.54 (G4)†‡	0.29	0.19 (G1)†
AuH_2	0.28	0.47 (G3)†	0.62	0.58 (G4′)‡	0	
AuH_3	0.00	···	0.26	0.28		
AuH_4	···	···	1.40$^{\sharp}$	···		
Ag	0.46	0.37	0.60	0.55		
AgH_1	0.36	0.28 (H2)$^{\natural}$	0.45	0.45 (E2)$^{\natural}$	0.45	0.09 (E3)$^{\natural}$
AgH_2	0.33	0.38 (H3)$^{\natural}$	0.50	0.5 (E6)$^{\natural}$	0	
AgH_3	0.00	···	0.13	···		
AgH_4	0.00	···	0.97$^{\sharp}$	···		

† Sveinbjörnsson and Engström (1995); ‡ Deixler *et al.* (1998); $^{\natural}$ Yarykin *et al.* (1999); $^{\sharp}$ The method has a larger error when the predicted level is further away from that of the standard.

10.3.5 On the passivation of ioslated transition metal centres

Both Au and Ag can be passivated by hydrogen (Pearton & Tavendale, 1982; Pearton & Tavendale, 1984) and our calculations suggest that AgH_3 and AuH_3 are candidates given our errors can be around 0.2 eV. However, in the case of gold, a level at $\approx E_c -$ 0.28 eV (Wang & Evwaraye, 1976) arises after a long room temperature anneal (Rubaldo *et al.*, 1999) and this is close to the calculated (−/0) level of AuH_3. If this level arises from this defect, then the passive defect cannot be due to substitutional Au complexed with H atoms.

Molecular hydrides AuH or AgH are chemically stable with dissociation energies ∼ 3.1 eV (Herzberg, 1945) and large ionization energies. If these formed within a void whose surface is passivated by hydrogen, then the defects would be electrically inactive. Calculations were carried out on a deca-vacancy, containing AuH_1 and AgH_1, and whose inner surface was passivated by hydrogen (Hourahine *et al.*, 1998) as described in the previous chapter for the Pt and Pd di-hydrides. The resulting ionization and electron affinity were such that no gap levels are present and such defects are then passive. However, whereas it is known that voids form in plasma treated or H-implanted material, it is unclear whether they, or a multi-vacancy complex containing Au and Ag, can be produced by wet chemical etching.

10.4 Concluding remarks

In summary, the calculations have allowed us to determine the structure and electrical activity of transition-metal–hydrogen defects (see Fig. 10.4). This has allowed us to assign the

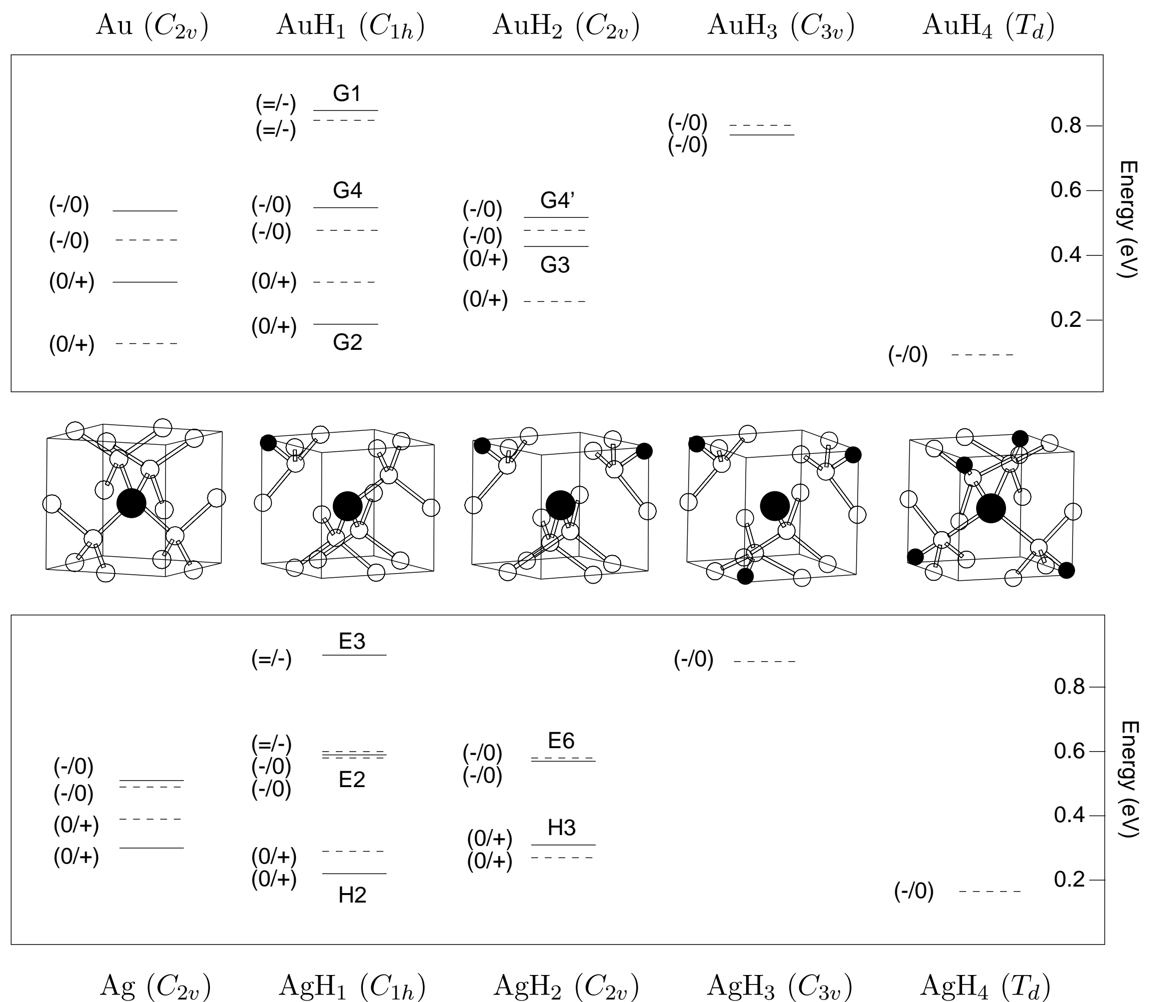


Figure 10.4: Structure and corresponding electrical levels of gold and silver and their complexes with hydrogen in Si. Experimental and calculated levels are represented by solid and dashed lines, respectively.

observed DLTS levels to specific defects. AuH_3 and AgH_3 defects have filled t_2-manifolds lying in the valence band and lack donor activity. They possess shallow $6s$ and $5s$ acceptor levels. Molecular hydrides of the metal impurities are passive when inserted into voids or passivated multi-vacancy centres.

11

Conclusions and Future Work

As mentioned in the abstract of this thesis, the main objective of the work described here was the investigation of the electrical properties of point deep-levels defect in crystalline silicon. It is crucial to understand the influence of this defects on the electrical and optical properties of a semiconductor like crystalline silicon.

A novel approach to the calculation of the electrical levels within a cluster approach was developed and successfully applied to the study of a number of defects.

Obviously, a calculation of the electronic affinity and ionisation energies of a given defect by means of Slater's transition-state scheme, despite being straightforward, is highly dependent on the atomic arrangement of the defect structure. Consequently, an accurate description of the ground-state configuration of the defect is fundamental. To describe the structure of the defects, we have used the AIMPRO code, an *ab initio* pseudopotential spin-polarised local density functional method applied to large hydrogen terminated atomic clusters containing up to 346 atoms. This method is highly suited for the modelling of covalent material, proving a wealth of information on the microscopic properties of such systems.

Since the method is based on a comparison between the calculated affinities of the defect in question and a standard one, it is imperative to verify the convergence with basis and cluster size in order to reduce the error of the calculation to a minimum. We found these error to be always less than about 0.2 eV.

Because of such an error, the method does not allow, in all cases, a clear distinction between the defects responsible for two closed spaced energy levels. Despite this, the method has been successful in the identification of a number of transition-metal–hydrogen defects containing several hydrogen atoms—a problem of immense technological and scientific importance. Ultimately, the control of levels would allow a considerable improvement in the performance of Si integrated device by supplying information about the impact of a defect on its electrical and optical properties.

The method has also allowed us to study the influence of strong Jahn-Teller driven distortions on the electrical properties of deep-levels. The vacancy-oxygen pair and the lattice di-vacancy are excellent examples of this effect.

So far, this novel approach to the calculation of deep levels has only been applied to

the study of defects in silicon. Obviously, there is no impediment to the application of the present method to the characterisation of the electrical properties of defects in other semiconducting materials.

Bibliography

Abou-el Fotouh F. A. and Newman R. C. (1974): Sol. State Commun. **15**, 1409

Abragam A. and Bleaney B. (1970): *Electron Paramagnetic Resonance of Transition Ions* – Clarendon Press, Oxford

Abragam A. and Pryce M. H. L. (1951): Proc. R. Soc. (London), Ser A **205**, 135

Agarwal A., Gossmann H.-J., Eaglesham D. J., Herner S. B., Fiory A. T., and Haynes T. E. (1999): Appl. Phys. Lett. **74**, 2435

Agarwal A., Gossmann H.-J., Eaglesham D. J., Pelaz L., Jacobson D. C., Haynes T. E., and Erokhin Y. E. (1997): Appl. Phys. Lett. **71**, 3141

Almbladh C. O. and von Barth U. (1976): Phys. Rev. B **13**, 3307

Alteheld P., Greulich-Weber S., Spaeth J.-M., Overhof H., and Höhne M. (1994): Mater. Sci. Forum **143–147**, 1173

Alves J. L. A. and Leite J. R. (1984): Phys. Rev. B **30**, 7284

Alves J. L. A. and Leite J. R. (1986): Phys. Rev. B **34**, 7174

Ammerlaan C. A. J. and van Oosten A. B. (1989): Phys. Scripta. **T25**, 342

Ammerlaan C. A. J. and Watkins G. D. (1972): Phys. Rev. B 5

Anderson F. G., Delerue C., Lannoo M., and Allan G. (1991): Phys. Rev. B **44**, 10925

Anderson F. G., Ham F. S., and Watkins G. D. (1992a): Phys. Rev. B **45**, 3287

Anderson F. G., Milligan R. F., and Watkins G. D. (1992b): Phys. Rev. B **45**, 3279

Anderson P. W. (1975): Phys. Rev. Lett. **34**, 953

Armelles G., Barrau J., Brousseau M., and Noguier J. P. (1986): Phys. Rev. B **33**, 1243

Artacho E. and Falicov L. M. (1991): Phys. Rev. B **43**, 12507

Artacho E. and Ynduráin F. (1989): Sol. State Commun. **72**, 393

Asghar M., Iqbal M., and Zafar N. (1993): J. Appl. Phys. **73**, 4240

Austin B. J., Heine V., and Sham L. J. (1962): Phys. Rev. **127**, A276

Awadelkarim O. O., Weman H., Svensson B. G., and Lindström J. L. (1986): J. Appl. Phys. **60**, 1974

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

Bai G. R., Qi M. W., Xie L. M., and Shi T. S. (1985): Sol. State Commun. **56**, 277

Bai G. R., Zhou J., Chen J. M., Shi T., Qi M. W., and Xie L. M. (1988): Sci. Sinica A **31**, 630

Baraff G. A., Kane E. O., and Schlüter M. (1979a): Phys. Rev. B **21**, 5662

Baraff G. A., Kane E. O., and Schlüter M. (1979b): Phys. Rev. Lett. **43**, 956

Baraff G. A. and Schlüter M. (1984): Phys. Rev. B **30**, 3460

Bech Nielsen B. (1999): (Private communication).

Bech Nielsen B. and Grimmeiss H. G. (1989): Phys. Rev. B **40**, 12403

Bech Nielsen B., Hoffmann L., and Budde M. (1996): Mater. Sci. & Eng. B **36**, 259

Bech Nielsen B., Hoffmann L., Budde M., Jones R., Goss J., and Öberg S. (1995): Mater. Sci. Forum **196–201**, 933

Bech Nielsen B., Johannesen P., Stallinga P., and Bonde Nielsen K. (1997): Phys. Rev. Lett. **79**, 1507

Bech Nielsen B., Olajos J., and Grimmeiss H. G. (1989): Phys. Rev. B **39**, 3330

Becke A. D. (1988): Phys. Rev. A **38**, 3098

Beeler F., Andersen O. K., and Scheffler M. (1990): Phys. Rev. B **41**, 1603

Beeler F., Anderson O. K., and Scheffler M. (1985): Phys. Rev. Lett. **55**, 1498

Benton J. L., Doherty C. J., Ferris S. D., Flamm D. L., Kimerling L. C., and Leamy H. J. (1980): Appl. Phys. Lett. **36**, 670

Blöchl P. E. (1990): Phys. Rev. B **41**, 5414

Bonde Nielsen K., Dobaczewski L., Goscinski K., Bensensen R., Andersen O., and Bech Nielsen B. (1999): in *Proceedings of the 20th ICDS*, 26-30 July, Berkeley, USA (To be published in Physica B).

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

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

Braun S., Grimmeiss H. G., and Spann K. (1977): J. Appl. Phys. **48**, 3883

Brent R. P. (1973): *Algorithms for Minimisation Without Derivatives*, Chapt. 5 – Prentice-Hall, New Jersey

Brotherton S. D. and Bradley P. (1982a): J. Appl. Phys. **53**, 57230

Brotherton S. D. and Bradley P. (1982b): J. Appl. Phys. **53**, 1543

Brotherton S. D., Bradley P., and Bicknell J. (1979): J. Appl. Phys. **50**, 3396

Brower K. L. (1971a): in *Radiation Effects in Semiconductors* edited by J. W. Corbett and G. D. Watkins (Gordon and Breach, New York), p. 189

Brower K. L. (1971b): Phys. Rev. B **4**, 1968

Bruel M. (1995): Electron. Lett. **70**, 3519

Bruni M., Bisero D., Tonini R., Ottaviani G., Quierolo G., and Bottini R. (1994): Phys. Rev. B **49**, 5291

Brunwin B., Hamilton B., Jordan P., and Peaker A. R. (1979): Electron. Lett. **15**, 349

Brusa R. S., Naia M. D., Zecca A., Nobili C., Ottaviani G., Tonini R., and Depasquier (1994): Phys. Rev. B **49**, 7271

Budde M. (1998): *Ph.D. thesis*, Aarhus Center for Atomic Physics, University of Aarhus, Denmark

Cardona M. (1983): Phys. Status Solidi B **118**, 463

Cartling B. G. (1975): J. Phys. C **8**, 3183

Carton-Merlet F., Pajot B., Ton D. T., Porte C., Clerjaud B., and Mooney P. M. (1982): J. Phys. C: Solid State Phys. **15**, 2239

Carton-Merlet F., Pajot B., and Vadja P. (1979): in *Defects and Radiation Effects in Semiconductors*, Inst. Phys. Conf. Proc. Serv. No. 46 (Institute of Physics, Bristol), p. 311

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

Chantre A. and Kimerling L. C. (1986): Appl. Phys. Lett. **48**, 1000

Chen J. W. and Milnes A. G. (1980): Annu. Rev. Mater. Sci. **10**, 157

Chen W. M., Awadelkarim O. O., Monemar B., Lindström J. L., and Oehrlein G. S. (1990): Phys. Rev. Lett. **64**, 3042

Chen W. M., Awadelkarim O. O., Monemar B., Lindström J. L., and Oehrlein G. S. (1998): Phys. Rev. Lett. **80**, 423

Cheng L. J., Corelli J. C., Corbett J. W., and Watkins G. D. (1966): Phys. Rev. **152**, 151

Cheng L. J. and Vadja P. (1969): Phys. Rev. **186**, 816

Clegg J. B. (1990): *Depth Profiling of Semiconductor Materials by Secondary Ion Mass Spectroscopy* in Growth and Characterisation of Semiconductors, Stradling, R. A. and Klipstein, P. C. (eds.), Adam Hilger, Bristol

Collins C. B., Carlson R. O., and Gallager C. J. (1957): Phys. Rev. **105**, 1168

Corbett J. W., Lindström J. L., and Pearton S. J. (1988): in *Defects in Electronic Materials* edited by M. Stavola, S. J. Pearton and G. Davies, MRS Symposia Proceedings No. 104 (Materials Research Society, Pittsburgh), p. 229

Corbett J. W. and Watkins G. D. (1965): Phys. Rev. A55

Cowern N. E. B., van de Walle G. F. A., Zalm P. G., and Vandenhoudt D. W. E. (1994): Appl. Phys. Lett. **65**, 2981

Dale C. J., Chen L., McNulty P. J., Marshall P. W., and Burke E. A. (1994): IEEE Trans. Nucl. Sci **NS41**, 1974

Davidson B. R., Newman R. C., Bullough T. J., and Joyce T. B. (1993): Phys. Rev. B **48**, 17106

Davidson J. A. and Evans H. (1996a): J. Appl. Phys. **81**, 251

Davidson J. A. and Evans J. H. (1996b): Semicond. Sci. Technol. **11**, 1704

Davies G., Oates A. S., Newman R. C., Woolley E C L. E. C., Binns M. J., and Wilkes J. G. (1986): J. Phys. C **19**, 841

de Wit J. G., Sieverts E. E., and Ammerlaan C. A. J. (1976): Phys. Rev. B **14**, 3494

Deák P., Heinrich M., Snyder L. C., and Corbett J. W. (1989): Mater. Sci. & Eng. B **4**, 57

Deák P., Snyder L., Heinrich M., Ortiz C. R., and Corbett J. W. (1991): Physica B **170**, 253

Deixler P., Terry J., Hawkins I. D., Evans-Freeman J. H., Peaker A. R., Rubaldo L., Maude D. K., Portal J.-C., Dobaczewski L., Nielsen K. B., and Larsen A. N. (1998): Appl. Phys. Lett. **73**, 3126

DeLeo G. G., Fowler W. B., and Watkins G. D. (1984): Phys. Rev. B **29**, 1819

DeLeo G. G., Watkins G. D., and Fowler W. B. (1982): Phys. Rev. B **25**, 4972

DeLeo G. G., Watkins G. D., and Fowler W. B. (1988): Phys. Rev. B **37**, 1013

Delerue C., Lannoo M., and Allan G. (1989): Phys. Rev. B **39**, 1669

Dexter D. L. (1962): Vol. 6 of *Solid State Physics* – Academic Press, New York

Dirac P. A. M. (1930): Proc. Cambridge Phil. Soc. **26**, 376

DiVincenzo D. P., Bernholc J., and Brodsky M. H. (1983): Phys. Rev. B **28**, 3246

Dobaczewski L., Kaczor P., Hawkins, and Peaker A. R. (1994): J. Appl. Phys. **76**, 194

Dreizler R. M. and Gross E. K. U. (1990): *Density Functional Theory* – Springer, Berlin

Dreyhsig J., Gumlich H.-E., and Allen J. W. (1993): Phys. Rev. B **48**, 15002

Dunlap B. I., Connolly W. J., and Sabin J. R. (1979): J. Chem. Phys. **71**, 4993

Eaglesham D. (1995): Physics World p. 41

Eaglesham D., Haynes T. E., Gossmann H.-J., Jacobson D. C., Stolk P. A., and Poate J. M. (1997): Appl. Phys. Lett. **70**, 3281

Englisch H. and Englisch R. (1983): Physica **121A**, 253

Estreicher S. K. (1995): Mater. Sci. & Eng. R **14**, 317

Evans J. H., Davidson J. A., Saritas M., Vandini M., Qian Y., and Peaker A. R. (1995): Mater. Sci. Forum **11**, 696

Evans M. J., Gornstein M. G., and Stavola M. (1997): Mat. Res. Soc. Symp. Proc. **442**, 275

Evans M. J., Stavola M., Weinstein M. G., and Uftring S. J. (1999): Mater. Sci. & Eng. B **58**, 118

Evwaraye A. O. (1976): J. Appl. Phys. **47**, 3172

Evwaraye A. O. and Sun E. (1976): J. Appl. Phys. **47**, 3776

Ewels C. P., Jones R., and Öberg S. (1995): Mater. Sci. Forum **196-201**, 1297

Fan H. Y. and Ramdas A. K. (1959): J. Appl. Phys. **30**, 1127

Fazzio A., Caldas M. J., and Zunger A. (1985): Phys. Rev. B **32**, 934

Fazzio A., Leite J. R., and Caldas M. J. (1983): Physica B **116**, 90

Feklisova O. V. and Yarykin N. A. (1997): Semicond. Sci. Technol. **12**, 742

Fermi E. (1928): Z. Phys. **48**, 73

Feynman R. P. (1939): Phys. Rev. **56**, 340

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

Frank F. C. (1949): Proc. Phys. Soc. (London) Sect. A **62**, 202

Frenkel J. (1938): Phys. Rev. **54**, 647

Fretwurst E., Kuhnke M., Moll M., and Lindström G. (1999): in *Proceedings of the First ENDEASD Workshop*, edited by C. Clays (April 21-22, Santorini, Greece), p. 39

Freund L. B. (1997): Appl. Phys. Lett. **70**, 3519

Frolov U. V. and Mukashev B. N. (1988): Phys. Status Solidi B **148**, K105

Fuchs M., Bockstedte M., Pehlke E., and Scheffler M. (1998): Phys. Rev. B **57**, 2134

Fulde P. (1995): *Electron Correlations in Molecules and Solids* – Springer, Berlin

García A. and Northrup J. E. (1995): Phys. Rev. Lett. **74**, 1131

Gell-Mann M. and Brueckner K. A. (1957): Phys. Rev. **106**, 364

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

Gill A. A., Baber N., and Iqbal M. Z. (1990): J. Appl. Phys. **67**, 1130

Gill A. A., Iqbal M. Z., and Zafar N. (1993): Semicond. Sci. Technol. **8**, 675

Gill K., Hall G., and McEvoy B. (1997): J. Appl. Phys. **82**, 126

Godby R. W., Schlüter M., and Sham L. J. (1986): Phys. Rev. Lett. **56**, 2415

Gorelkinskiĭ Y. V. and Nevinnyi (1991): Physica B **170**, 115

Gösele U., Frank W., and Seeger A. (1980): Appl. Phys. **23**, 361

Goss J., Resende A., Jones R., Öberg S., and Briddon P. R. (1995): Mater. Sci. Forum **196**, 67

Gossmann H.-J., Haynes T. E., Stolk P. A., Jacobson D. C., Gilmer G., Poate J. M., Luftman H. S., Mogi T. K., and Thomson M. O. (1997): Appl. Phys. Lett. **71**, 3862

Graff K. (1995): *Metal Impurities in Silicon-Device Fabrication* – Springer, Berlin

Grekhov A. M., Gun'ko V. M., Klapchenko G. M., and Tsyashchenko Y. P. (1983): Sov. Phys. Semicond. **17**, 1186

Grimmeiss H. G. (1977): Ann. Rev. Mater. Sci **7**, 341

Gross E. K. U. and Dreizler R. M. (1979): Phys. Rev. A **20**, 1798

Gross E. K. U., Runge E., and Heinonen O. (1991): *Many-Particle Theory* – Adam Hilger, Bristol

Gunnarsson O., Jonson M., and Lundqvist B. I. (1979): Phys. Rev. B **20**, 3136

Gunnarsson O. and Lundqvist B. I. (1976): Phys. Rev. B **13**, 4274

Gutsev G. L., Myakenkaya G. S., Frolov V. V., and Glazman V. B. (1989): Phys. Status Solidi B **153**, 659

Gutsev G. L. and Myakenkaya G. S. M. (1989): Phys. Status Solidi B **156**, 319

Hai P. N., Gregorkiewicz T., Ammerlaan C. A. J., and Don D. T. (1997): Mater. Sci. Forum **258-263**, 491

Hall R. N. (1952): Phys. Rev. **86**, 600

Hállen A., Keskitalo N., Josyula L., and Svensson B. G. (1999): J. Appl. Phys. **86**, 214

Hallén A., Sundqvist B. U. R., Paska Z., Svensson B. G., Rosling M., and Tirén J. (1990): J. Appl. Phys. **67**, 1266

Haller E. E. (1995): J. Appl. Phys. **77**, 2857

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

Harris J. (1984): Phys. Rev. A **29**, 1648

Harris J. and Jones R. O. (1974): J. Phys. F **4**, 1170

Harris R. D., Newton J. L., and Watkins G. D. (1987): Phys. Rev. B **36**, 2094

Hartree D. R. (1928): Proc. Cambridge Phil. Soc. **24**, 89

Hemstreet L. A. (1976): Phys. Rev. B **15**, 834

Hemstreet L. A. (1980): Phys. Rev. B **22**, 4590

Hemstreet L. A. and Dimmock J. O. (1979): Phys. Rev. B **20**, 1527

Henning J. C. M. and Egelmeers E. C. J. (1983): Phys. Rev. B **27**, 4002

Herstrøm S. (1998): *Master's thesis*, Aarhus Center for Atomic Physics, University of Aarhus, Denmark

Herzberg G. (1945): *Molecular Structure and Molecular Structure. II. Infrared and Raman Spectra of Polyatomic Molecules* – D. van Nostrand Rheinhold, New York

Hitti B., Kreitzman S. R., Estle T. L., Bates E. S., Dawdy M. R., Head T. L., and Litchi R. L. (1999): Phys. Rev. B **59**, 4918

Hjalmarson H. P. and Jennison D. R. (1985): Phys. Rev. B **31**, 1208

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

Höhne M. (1980): Phys. Status Solidi **99**, 651

Höhne M. (1982): Phys. Status Solidi **109**, 525

Höhne M. (1992): Phys. Rev. B **45**, 5883

Höhne M. and Juda U. (1995): Mat. Sci. Technol. **11**, 680

Höhne M., Juda U., Martynov Y. V., Gregorkiewicz T., Ammerlaan C. A. J., and Vlasenko L. S. (1994): Phys. Rev. B **49**, 13423

Holbech J. D., Bech Nielsen B., Jones R., Sitch P., and Öberg S. (1993): Phys. Rev. Lett. **71**, 875

Holm B., Bonde Nielsen K., and Bech Nielsen B. (1991): Phys. Rev. Lett. **66**, 2360

Hopkinson G. R., Dale C. J., and Marshall P. W. (1996): IEEE Trans. Nucl. Sci **NS43**, 614

Hourahine B., Jones R., Öberg S., and Briddon P. R. (1999): Phys. Rev. B **59**, 15729

Hourahine B., Jones R., Öberg S., Newman R. C., Briddon P. R., and Roduner E. (1998): Phys. Rev. B **57**, 666

Hubbard J. (1963): Proc. R. Soc. (London), Ser. A **276**, 238

Humphreys R. G., Brand S., and Jaros M. (1983): J. Phys. C **16**, L337

Irion E., Burger N., Thonke K., and Sauer R. (1985): J. Phys. C **18**, 5083

Irmscher K., Klose H., and Maass K. (1984): J. Phys. C **17**, 6317

Isoya J., Kanda H., Norris J. R., Tang J., and Bowman M. K. (1990): Phys. Rev. B **41**, 3905

Jahn H. A. and Teller E. (1937): Proc. R. Soc. (London), Ser A **161**, 220

Janak J. F. (1978): Phys. Rev. B **18**, 7165

Jellison G. E. (1982): J. Appl. Phys. **53**, 5715

Johannesen P., Bech Nielsen B., and Byberg J. R. (1999): in *Proceedings of the 20th ICDS*, 26-30 July, Berkeley, USA (To be published in Physica B).

Johnson N. M., Herring C., and Chadi D. J. (1986): Phys. Rev. Lett. **56**, 769

Johnson N. M., Herring C., and Van de Walle C. G. (1994): Phys. Rev. Lett. **73**, 130

Johnson N. M., Ponce F. A., Street R. A., and Nemanich R. J. (1987): Phys. Rev. B **35**, 4166

Jones R. (1988): J. Phys. C **21**, 5735

Jones R. (1992): Proc. R. Soc. (London), Ser. A **341**, 351

Jones R. and Briddon P. R. (1998): *Identification of Defects in Semiconductors*, Vol. 51A of *Semiconductors and Semimetals*, Chapt. 6, p. 287 – Academic Press, Boston

Jones R., Coomer J., Resende A., Öberg S., and Briddon P. R. (1999a): (To be published).

Jones R., Goss J., Ewels C., and Öberg S. (1994): Phys. Rev. B **50**, 8378

Jones R., Öberg S., Goss J., Briddon P. R., and Resende A. (1995): Phys. Rev. Lett. **75**, 2734

Jones R., Resende A., Öberg S., and Briddon P. R. (1999b): Mater. Sci. & Eng. B **58**, 113

Jones R. and Sayyash A. (1986): J. Phys. C **19**, L653

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

Juan Y.-M., Kaxiras E., and Gordon R. G. (1995): Phys. Rev. B **51**, 9521

Juda U., Scheerer O., Höhne M., Riemann H., Schilling H.-J., Donecker J., and Gerhardt A. (1996): J. Appl. Phys. **80**, 3435

Kalma A. H. and Corelli J. C. (1968): Phys. Rev. **173**, 734

Kaplyanskiĭ A. A. (1994): Opt. Spectrosc. (USSR) **16**, 329

Katayama-Yoshida H. and Zunger A. (1985): Phys. Rev. B **31**, 7877

Kauppinen H., Corbel C., Nissilä J., Saarinen K., and Hautojärvi P. (1998): Phys. Rev. B **57**, 12911

Kauppinen H., Corbel C., Skog K., Laine T., Hautijärvi P., Desgardin P., and Ntsoenzok E. (1997): Phys. Rev. B **55**, 9598

Kerker G. P. (1980): J. Phys. C **13**, L189

Kimerling L. C. (1974): J. Appl. Phys. **45**, 1839

Kimerling L. C. (1977): in *Radiation Effects in Semiconductors*, edited by N. B. Urli and J. W. Corbett, Inst. Phys. Conf. Proc. Serv. No. 31 (Institute of Physics, Bristol), p. 221

Kimerling L. C. and Benton J. L. (1981): Appl. Phys. Lett. **39**, 410

Kimerling L. C., DeAngelis H. M., and Diebold J. W. (1975): Sol. State Commun. **16**, 171

Kinomura A., Williams J. S., Wong-Leung J., and Petravic M. (1998): Appl. Phys. Lett. **72**, 2713

Kirton M. J., Banks P. W., Lian D. D., and Jaros M. (1984): J. Phys. C **17**, 2487

Kleinhenz R. L., Lee Y. H., Corbett J. W., Sieverts E. G., Müller S. H., and Ammerlaan C. A. J. (1981): Phys. Status Solidi A **108**, 363

Kleinhenz R. L., Lee Y. H., Singh V. A., Mooney P. M., Jaworowski A., Roth L. M., Corelli J. C., and Corbett J. W. (1979): Inst. Phys. Conf. Proc. Serv. No. 46 (Institute of Physics, Bristol), p. 200

Kleinman L. (1997): Phys. Rev. B **56**, 12042

Kleverman M., Thilderkvist A., Grossman G., Grimmeiss H. G., and Watkins G. D. (1995): Sol. State Commun. **93**, 383

Kohn W. (1957): *Shallow Impurity States in Silicon and Germanium*, Vol. 5 of *Solid State Physics* – Academic Press, New York

Kohn W. (1983): Phys. Rev. Lett. **51**, 1596

Kohn W. (1997): *Fundamentals of Density Functional Theory* in Proceedings of the Tenth Chris Engelbrecht Summer School in Theoretical Physics, Meerensee, South Africa, 19–29 January 1997, edited by D. Joubert (Springer-Verlag, Berlin)

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

Koopmans T. A. (1933): Physica **1**, 104

Koveshnikov S. and Kononchuk O. (1998): Appl. Phys. Lett. **73**, 2340

Kurokawa M., Motobayashi T., Ieki K., Shimoura S., Murakami H., Ikeda Y., Moriya S., Yanagisama Y., and Nomura T. (1995): IEEE Trans. Nucl. Sci **NS42**, 163

Kwon Y. K., Ishikawa T., and Kuwano H. (1987): J. Appl. Phys. **61**, 1055

Lalita J., Svensson B. G., Jagadish C., and Hallén A. (1997): Nucl. Instr. Meth. B **127/128**, 69

Lang D. V. (1974): J. Appl. Phys. **45**, 3023

Lang D. V. (1979): Topics in Appl. Phys. **41**, 93

Lang D. V., Grimmeiss H. G., Meijer E., and Jaros M. (1980): Phys. Rev. B **22**, 3917

Lang D. V. and Henry C. H. (1975): Phys. Rev. Lett. **35**, 1525

Langreth D. C. and Perdew J. P. (1975): Sol. State Commun. **17**, 1425

Langreth D. C. and Perdew J. P. (1977): Phys. Rev. B **15**, 2884

Lannoo M. (1987): Phys. Rev. B **36**, 9355

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

Leary P., Jones R., Öberg S., and Torres V. J. (1997): Phys. Rev. B **55**, 2188

Lee C., Yang W., and Parr R. G. (1988): Phys. Rev. B **37**, 785

Lee T. F. and McGill T. C. (1973): J. Phys. C **6**, 3438

Lee Y.-H. and Corbett J. W. (1973): Phys. Rev. B **8**, 2810

Lee Y.-H. and Corbett J. W. (1974): Phys. Rev. B **9**, 4351

Lee Y.-H. and Corbett J. W. (1976): Phys. Rev. B **13**, 2653

Leigh R. S. and Newman R. C. (1982): J. Phys. C **15**, L1045

Leigh R. S., Newman R. C., Sangster M. J. L., Davidson B. R., Ashwin M. J., and Robbie D. A. (1994): Semicond. Sci. Technol. **9**, 1054

Leigh R. S. and Szigeti B. (1967): Proc. R. Soc. (London), Ser A **301**, 211

Lemke H. (1984): Phys. Status Solidi A **86**, K39

Lemke H. (1987): Phys. Status Solidi A **101**, 193

Lemke H. (1995): Mater. Sci. Forum **196-201**, 683

Levy M. (1982): Phys. Rev. A **26**, 1200

Levy M. (1996): *Recent Developments and Applications of Modern Density Functional Theory* – Elsevier, Amsterdam

Levy M. and Perdew J. P. (1985): *Density Functional Methods in Physics*, Chapt. The constrained search formulation of density functional theory, p. 11 – Plenum, New York

Libertino S., Coffa S., Benton J. L., Halliburton K., and Eaglesham D. J. (1998): Nucl. Instr. Meth. B **148**, 247

Lieb E. H. (1981): Phys. Rev. Lett. **46**, 457

Lieb E. H. (1983): Int. J. Quantum Chem. **24**, 243

Lightowlers E. C. (1976): in *Growth and Characterisation of Semiconductors*, ed. by R. A. Stradling and P. C. Klipstein (IOP Publishing, Bristol), p. 135

Lightowlers E. C. (1995): Mater. Sci. Forum **196-201**, 817

Lightowlers E. C., Newman R. C., and Tucker J. H. (1994): Semicond. Sci. Technol. **9**, 1370

Lindefelt U. and Yong-Liang W. (1988): Phys. Rev. B **38**, 4107

Lindfelt U. and Zunger A. (1984): J. Phys. C **17**, 6047

Lindström J. L. and Svensson B. G. (1986): Mater. Sci. Soc. Symp. Proc. **59**, 45

Lindström L. J., Oehrlein G. S., Jaworowski A. E., and Corbet J. W. (1982): J. Appl. Phys. **53**, 8686

Lisiak K. P. and Milnes A. G. (1976): J. Appl. Phys. **46**, 5229

Lister G. M. S. and Jones R. (1988): Unpublished

Londos C. A., Sarlis N. V., and Fytros L. G. (1996): NATO ARW 'Oxygen 96', ASI series, edited by R. Jones **17**, p. 477

Lowther J. E. (1980): J. Phys. C **13**, 3665

Ludwig G. W. and Woodbury H. H. (1960): Phys. Rev. Lett. **5**, 98

Ludwig G. W. and Woodbury H. H. (1962): *Electron Spin Resonance in Semiconductors*, Vol. 13 of *Solid State Physics* – Academic Press, New York

Lütgemeier H. and Schnitzke K. (1967): Phys. Lett. A **25**, 232

Mantovani S., Nava F., Nobili C., and Ottaviani G. (1986): Phys. Rev. B **33**, 5536

March N. H. (1983): *Theory of the Inhomogeneous Electron Gas*, Chapt. 1, p. 1 – Plenum, New York

Mattila T. and Nieminen R. M. (1996): Phys. Rev. B **54**, 16676

Mattila T. and Zunger A. (1998): Phys. Rev. B **58**, 1367

Shaw Jr. R. W. and Harrison W. A. (1967): Phys. Rev. **163**, A604

von Bardeleben H. J., Stiévenard D., Brousseau M., and Barrau J. (1988): Phys. Rev. B **38**, 6308

McHenry M. E., O'Handley R. C., and Johnson K. H. (1987): Phys. Rev. B **35**, 3555

McHugo S. A., Weber E. R., Myers S. M., and Petersen G. A. (1996): Appl. Phys. Lett. **69**, 3060

Meng X. T. (1991): Physica B **170**, 249

Mesli A., Courcelle E., Zundel T., and Siffert P. (1987): Phys. Rev. B **36**, 8049

Meyer A., Wang G. C., and Young W. H. (1976): Z. Naturforsch **31a**, 898

Meyers S. M., Petersen G. D., and Seager C. H. (1996): J. Appl. Phys. **80**, 3717

Michel A. E., Raush W., Ronsheim P. A., and Kastl R. H. (1987): Appl. Phys. Lett. **50**, 416

Miller G. L., Lang D. V., and Kimerling L. C. (1977): Ann. Rev. Mater. Sci **7**, 377

Miller M. D., Schade H., and Nuese C. J. (1976): J. Appl. Phys. **47**, 2569

Milligan R. F., Anderson F. G., and Watkins G. D. (1984): Phys. Rev. B **29**, 2819

Mohadjeri B., Williams J. S., and Wong-Leung J. (1995): Appl. Phys. Lett. **66**, 1889

Mooney P. M., Cheng L. J., Süli M., Gerson J. D., and Corbett J. W. (1977): Phys. Rev. B **15**, 3836

Moore C. (1970): National Standard Reference Data Series **34**

Morton J. R. and Preston K. F. (1978): J. Magn. Res. **30**, 577

Mukashev B. N., Tamendarov M. F., Tokmoldin S. Z., and Frovol V. (1985): Phys. Status Solidi A **91**, 509

Myers S. M. and Petersen G. A. (1998): Phys. Rev. B **57**, 7015

Newman R. C. (1973): *Infrared Studies of Crystal Defects* – Taylor and Francis, London

Newton J. N., Chatterjee A. P., Harris R. D., and Watkins G. D. (1983): Physica B **116**, 219

Öğüt S., Kim H., and Chelikowsky J. R. (1997): Phys. Rev. B **56**, 11353

Olajos J., Kleverman M., and Grimmeiss H. G. (1989): Phys. Rev. B **40**, 6196

Oliver G. L. and Perdew J. P. (1979): Phys. Rev. B **20**, 397

Omling P., Emanuelson P., and Grimmeiss H. G. (1987): Phys. Rev. B **36**, 6202

Pake G. E. and Estle T. L. (1973): *The Physical Principles of Electron Paramagnetic Resonance* – W. A. Benjamim, Inc., Reading

Pankove J. I., Carlson D. E., Berkeyheiser J. E., and Wance R. O. (1983): Phys. Rev. Lett. **51**, 1983

Pantelides S. K. (1978): Rev. Mod. Phys. **50**, 797

Parakhonskiĭ A. L., Feklisova O. V., Karelin S. S., and Yarykin N. A. (1996): Semiconductors **30**, 362

Park Y. K., Estreicher S. K., Myles C. W., and Fedders P. A. (1995): Phys. Rev. B **52**, 1718

Parr R. G. and Bartolotti L. J. (1983): J. Phys. C **87**, 2810

Parr R. G., Donelly R. A., Levy M., and Palke W. (1978): J. Chem. Phys. **68**, 3801

Parr R. G. and Yang W. (1989): *Density-Functional Theory of Atoms and Molecules* – Oxford University Press, Oxford

Peaker A. R., Evans-Freeman J. H., Kan P. Y. Y., Rubaldo L., Hawkins I. D., Vernon-Parry K. D., and Dobaczewski L. (1999): in *Proceedings of the 20th ICDS*, 26-30 July, Berkeley, USA (To be published in Physica B).

Pearton S. J. (1982): Phys. Status Solidi A **72**, K73

Pearton S. J., Corbett J. W., and Stavola M. (1992): *Hydrogen in Crystalline Semiconductors*, Chapt. 3 – Springer-Verlag, Berlin

Pearton S. J., Kahn J. M., Hansen W. L., and Haller E. E. (1983): J. Appl. Phys. **55**, 1464

Pearton S. J. and Tavendale A. J. (1982): Phys. Rev. B **26**, 7105

Pearton S. J. and Tavendale A. J. (1984): J. Phys. C **17**, 6701

Pêcheur P. and Toussaint G. (1983): Physica B **116**, 112

Pellegrino P., Kuznetsov A. Y., and Svensson B. G. (1999): *Impurity-assisted annealing of point defect complexes in ion implanted silicon* (To be published)

Perdew J. P. (1985): Phys. Rev. Lett. **55**, 1665

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

Perdew J. P. and Kurth S. (1998): in D.Joubert (ed.), *Proceedings of the Tenth Chris Engelbrecht Summer School in Theoretical Physics, Meerensee, South Africa, 19–29 January 1997*, Lecture Notes in Physics, Springer, Berlin

Perdew J. P., Kurth S., Zupan S., and Blaha P. (1999): Phys. Rev. Lett. **82**, 2544

Perdew J. P. and Levy M. (1983): Phys. Rev. Lett. **51**, 1884

Perdew J. P. and Levy M. (1997): Phys. Rev. B **56**, 16021

Perdew J. P., Parr R. G., Levy M., and Balduz Jr. J. L. (1982): Phys. Rev. Lett. **49**, 1691

Perdew J. P. and Wang Y. (1992): Phys. Rev. B **45**, 13244

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

Pesola M., von Boehm J., Mattila T., and Nieminen R. M. (1999): Submitted to PRB

Pesola M., von Boehm J., Pöykkö S., and Nieminen R. M. (1998): Phys. Rev. B **58**, 1106

Petersen J. W. and Nielsen J. (1990): Appl. Phys. Lett. **56**, 1122

Philips J. C. and Kleinman L. (1959): Phys. Rev. B **116**, 287

Pickett W. E. (1981): Phys. Rev. B **23**, 6603

Picoli G., Chomette A., and Lannoo M. (1984): Phys. Rev. B **30**, 7138

Pines D. (1963): *Elementary Excitations in Solids* – W. A. Benjamin, Inc., New York

Polak E. (1971): *Computational Methods in Optimisation* – Academic Press, New York

Pöykkö S., Puska M. J., and Nieminen (1996): Phys. Rev. B **53**, 3813

Press W. H., Flannery B. P., Teukolsky S. A., and Vetterling W. T. (1982): *Numerical Recipes* – Cambridge University Press, Cambridge

Puska M. J., Pöykkö S., Pesola M., and Nieminen R. M. (1998): Phys. Rev. B **58**, 1318

Qian G.-X., Martin R. M., and Chadi D. J. (1988): Phys. Rev. B **38**, 7649

Rajagopal A. K. and Callaway J. (1973): Phys. Rev. B **7**, 1912

Rajapogal A. K. (1980): Adv. Chem. Phys. **41**, 59

Rappe A. M., Rabe K. M., Kaxiras E., and Joannopoulos J. D. (1990): Phys. Rev. B **41**, 1227

Resende A., Goss J. P., Öberg S., Briddon P. R., and Jones R. (1997): Mater. Sci. Forum **258-263**, 295

Resende A., Jones R., Öberg S., and Briddon P. R. (1999): Phys. Rev. Lett. **82**, 2111

Roberson M. A. and Estreicher S. K. (1994): Phys. Rev. B **49**, 17040

Rubaldo L., Deixler P., Hawkins I. D., Terry J., Maude D. K., Portal J.-C., Evans-Freeman J. H., Dobaczewski L., and Peaker A. R. (1999): Mater. Sci. & Eng. B **58**, 126

Rücker H., Heinemann B., Röpke W., Kurps E., Krüger D., Lippert G., and Osten H. J. (1998): Appl. Phys. Lett. **73**, 1682

Sachse J.-U. (1997): *Ph.D. thesis*, Max-Planck-Institut für Festkörperforschung, Stuttgart

Sachse J.-U., Jost W., Weber J., and Lemke H. (1997a): Appl. Phys. Lett. **71**, 1379

Sachse J.-U., Sveinbjörnsson E. Ö., Jost W., and Weber J. (1997b): Phys. Rev. B **55**, 16176

Sachse J.-U., Sveinbjörnsson E. Ö., Jost W., and Weber J. (1997c): Appl. Phys. Lett. **70**, 1584

Sachse J.-U., Sveinbjörnsson E. Ö., Yarykin N., and Weber J. (1999a): Mater. Sci. & Eng. B **58**, 134

Sachse J.-U., Weber J., and Lemke H. (1997):d Mater. Sci. Forum **258-263**, 307

Sachse J.-U., Weber J., and Sveinbjörnsson E. Ö. (1999b): Phys. Rev. B **60**, 1474

Safonov A. N., Lightowlers E. C., Davies G., Leary P., Jones R., and Öberg S. (1996): Phys. Rev. Lett. **77**, 4812

Saito M. and Oshiyama A. (1994): Phys. Rev. Lett. **6**, 866

Sambe H. and Felton R. H. (1975): J. Chem. Phys. **61**, 3862

Schmidt D. C. (1998): *Ph.D. thesis*, Université de Poitiers, France

Schmidt D. C., Svensson B. G., Keskitalo N., Godey S., Ntsoenzok E., Barbot J., and Blanchard C. (1998): J. Appl. Phys. **84**, 4214

Seong H. and Lewis L. J. (1996): Phys. Rev. B **53**, 9791

Sham L. J. (1966): Phys. Rev. **150**, 720

Sham L. J. and Schlüter M. (1983): Phys. Rev. Lett. **51**, 1888

Sham L. J. and Schlüter M. (1985): Phys. Rev. B **32**, 3883

Shen B., Zhang X. Y., Chen P., Zhang R., Shi Y., Zheng Y. D., Sekiguchi T., and Sumino K. (1996): Appl. Phys. Lett. **68**, 214

Shen B., Zhang X. Y., Chen P., Zhang R., Shi Y., Zheng Y. D., Sekiguchi T., and Sumino K. (1997): Appl. Phys. Lett. **70**, 1876

Shi T. S., Xie L. M., Bai G. R., and Qi M. W. (1985): Phys. Status Solidi B **131**, 511

Shockley W. and Read Jr. W. T. (1952): Phys. Rev. **87**, 835

Sieverts E. G., Muller S. H., and Ammerlaan C. A. J. (1978a): Phys. Rev. B **18**, 6834

Sieverts E. G., Muller S. H., and Ammerlaan C. A. J. (1978b): Sol. State Commun. **28**, 221

Sieverts E. G., Sprenger M., and Ammerlaan C. A. J. (1990): Phys. Rev. B **41**, 8630

Singh V. A., Weigel C., Corbett J. W., and Roth L. M. (1977): Phys. Status Solidi A **81**, 637

Slater J. C. (1928): Phys. Rev. **32**, 339

Slater J. C. (1930): Phys. Rev. **35**, 210

Slater J. C. (1972): Adv. Quantum Chem. **6**, 1

Slater J. C. (1974): *The Self-Consistent Field for Molecules and Solids*, Vol. IV – McGraw-Hill, Auckland

Smargiassi E. and Carr R. (1996): Phys. Rev. B **53**, 9760

Snyder L. C. and Corbett J. W. (1986): Mat. Res. Soc. Symp. Proc. **59**, 207

Són N. T., Kustov V. E., Gregorkiewicz T., and Ammerlaan C. A. J. (1992): Phys. Rev. B **46**, 4544

Són N. T., von Oosten A. B., and Ammerlaan C. A. J. (1990): Sol. State Commun. **80**, 439

Song E. G., Kim E., Lee Y. H., and Hwang Y. G. (1993): Phys. Rev. B **48**, 1486

Song L. W., Zhan X. D., Benson B. W., and Watkins G. D. (1986): Phys. Rev. B **33**, 1452

Song L. W., Zhan X. D., Benson B. W., and Watkins G. D. (1988): Phys. Rev. Lett. **60**, 460

Song L. W., Zhan X. D., Benson B. W., and Watkins G. D. (1990): Phys. Rev. B **42**, 5765

Sopori B. L., Jones K., and Deng X. J. (1992): Appl. Phys. Lett. **61**, 2560

Stallinga P. and Bech Nielsen B. (1998): Phys. Rev. Lett. **80**, 422

Stallinga P., Johannsen P., Herstrøm S., Bonde Nielsen K., and Bech Nielsen B. (1998):

Phys. Rev. B **58**, 3842

Stavola M. (1998a): (Private communication).

Stavola M. (1998b): *Identification of Defects in Semiconductors*, Vol. 51B of *Semiconductors and Semimetals*, Chapt. 3, p. 153 – Academic Press, Boston

Stein H. J. (1975): J. Electro. Mater. **4**, 159

Stöffler W. and Weber J. (1986): Phys. Rev. B **33**, 8892

Stolk P. A., Eaglesham D. J., Grossman H.-J., and Poate J. M. (1995): Appl. Phys. Lett. **66**, 1370

Stoneham A. M. (1975): *Theory of Defects in Insulators and Semiconductors* – Clarendon Press, Oxford

Sturge M. D. (1967): *The Jahn-Teller Effect in Solids*, Vol. 20 of *Solid State Physics* – Academic Press, New York

Suezawa M. (1997): Mater. Sci. Forum **258**, 229

Suezawa M. (1998a): Jap. J. Appl. Phys. **37**, L259

Suezawa M. (1998b): Jap. J. Appl. Phys. **37**, L806

Suezawa M. (1998c): J. Appl. Phys. **83**, 1958

Sugino O. and Oshiyama A. (1990): Phys. Rev. B **42**, 11869

Sveinbjörnsson E. Ö., Andersson G. I., and Engström O. (1994): Phys. Rev. B **49**, 7801

Sveinbjörnsson E. Ö. and Engström O. (1995): Phys. Rev. B **52**, 4884

Svensson B. G., Hállen A., and Sundqvist B. U. R. (1989): Mater. Sci. & Eng. B **4**, 285

Svensson B. G., Jagadish C., Hallén A., and Lalita J. (1997): Phys. Rev. B **55**, 10498

Svensson B. G., Lindström J. L., and Corbett J. W. (1985): Appl. Phys. Lett. **47**, 841

Svensson B. G., Mohadjeri B., Hallén A., Svensson J. H., and Corbett J. W. (1991): Phys. Rev. B 43

Svensson B. G. and Willander M. (1987): J. Appl. Phys. **62**, 2758

Svensson J. H., Svensson B. G., and Monemar B. (1988): Phys. Rev. B **38**, 4192

Tasch A. F. and Sah C. T. (1970): Phys. Rev. B **1**, 800

Tavendale A. J., Williams A. A., and Pearton S. J. (1988): Mater. Res. Soc. Symp. Proc **104**, 285

Thomas L. H. (1927): Proc. Cambridge Phil. Soc. **23**, 542

Tokuda Y. and Shimada H. (1998): in *hydrogen in Semiconductors and Metals*, Ed. by N. K. Nickel, W. B. Jackson and R. C. Bowman, Mater. Res. Soc, **513**, p. 363

Torres V. J. B. and Stoneham A. M. (1985): *Handbook of Interatomic potentials. III. Semiconductors* – Harwell Laboratory, Theoretical Physics Division

Trombetta J. M. and Watkins G. D. (1988): Mater. Res. Soc. Symp. Proc. **104**, 93

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

Troxell J. R. and Watkins G. D. (1980): Phys. Rev. B **22**, 921

Uftring S. J., Stavola M., Williams P. M., and Watkins G. D. (1995): Phys. Rev. B **51**, 9612

van Oosten A. B., Frens A. M., and Schmidt J. (1994): Phys. Rev. B **50**, 5239

Vanderbilt D. (1985): Phys. Rev. B **32**, 8412

Vanderbilt D. (1990): Phys. Rev. B **32**, 8412

Veloarisoa I. A., Stavola M., Kozuch D. M., and Watkins G. D. (1991): Appl. Phys. Lett. **59**, 2121

Venezia V. C., Englesham D., Haynes T. E., Agarwal A., Jacobson D. C., Gossmann H.-J., and Baumann F. H. (1998): Appl. Phys. Lett. **73**, 2980

Vlasenko L. S., Són T. S., von Oosten A. B., and Ammerlaan C. A. J. (1990): Sol. State Commun. **73**, 393

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

von Weizsäcker C. F. (1935): Z. Physik **96**, 431

Vosko S. H., Wilk L., and Nusair M. (1980): Can. J. Phys **58**, 1200

Wang K. L. and Evwaraye A. O. (1976): J. Appl. Phys. **47**, 4574

Wang L., Yao X. C., Zhou J., and Qin G. G. (1988): Phys. Rev. B **38**, 13494

Watkins G. and Troxell J. R. (1980a): Phys. Rev. Lett. **44**, 593

Watkins G. D. (1965): in *Radiation Damage in Semiconductors* (Dunod, Paris), p. 97

Watkins G. D. (1976): in *Defects and Their Structure in Non-metallic Solids*, ed. by B. Henderson and A. E. Hughes – Plenum, London

Watkins G. D. (1983): Physica B **117-118**, 9

Watkins G. D. (1992): *Deep Centers in Semiconductors*, Chapt. 3 – Gordon and Breach, Switzerland, 2nd edition

Watkins G. D. (1995): Phys. Rev. Lett. **74**, 4353

Watkins G. D. (1998): *Identification of Defects in Semiconductors*, Vol. 51A of *Semiconductors and Semimetals*, Chapt. 1, p. 1 – Academic Press, Boston

Watkins G. D. and Brower K. L. (1976): Phys. Rev. Lett. **36**, 1329

Watkins G. D. and Corbett J. W. (1961): Phys. Rev. **121**, A1001

Watkins G. D. and Corbett J. W. (1964): Phys. Rev. **134**, A1359

Watkins G. D. and Corbett J. W. (1965): Phys. Rev. **138**, A543

Watkins G. D., DeLeo G. G., and Fowler W. B. (1983): Phys. Status Solidi B+C **116**, 28

Watkins G. D., Kleverman M., Thilderkvist A., and Grimmeiss H. G. (1991): Phys. Rev. Lett. **67**, 1149

Watkins G. D. and Troxell J. R. (1980b): Phys. Rev. Lett. **44**, 593

Watkins G. D. and Williams P. M. (1995): Phys. Rev. B **52**, 16575

Watts S. (1999): in *Proceedings of the First ENDEASD Workshop*, edited by C. Clays (April 21-22, Santorini, Greece), p. 93

Weber J. (1999): (Private communication).

Weil J. A., Bolton J. R., and Wertz J. E. (1994): *Electron Paramagnetic Resonance* – Wiley, New York

Weinstein M. G. and Stavola M. (1999): (Unpublished).

Weldon M. K., Collot M., Chabal Y. J., Venezia V. C., Agarwal A., Haynes T. E., Eaglesham D. J., Christman S. B., and Chaban E. E. (1998): Appl. Phys. Lett. **73**, 3721

Weldon M. K., Marsico V. E., Chabal Y. J., Agarwal A., Eaglesham D. J., Sapjeta J.,

Brown W. L., Jacobson D. C., Caudano Y., Christman S. B., and Chaban E. E. (1997): J. Vac. Sci. and Technol. B **15**, 1065

Williams J. S. and Poate J. M. (eds.) (1984): *Ion Implantation and Beam Processing* – Academic Press, Sydney

Williams P. M., Watkins G. D., Uftring S., and Stavola M. (1993): Phys. Rev. Lett. **70**, 3816

Williams P. M., Watkins G. D., Uftring S., and Stavola M. (1994): Mater. Sci. Forum **143–147**, 891

Wong-Leung J., Ascheron C. E., Petravic M., Elliman, and Williams J. S. (1995): Appl. Phys. Lett. **66**, 1231

Woodbury H. H. and Ludwig G. W. (1962): Phys. Rev. **126**, 466

Xie L. M., Qi M. W., and Chen J. M. (1991): J. Phys. C **3**, 8519

Xu H. (1992): Phys. Rev. B **46**, 1403

Yarykin N., Sachse J.-U., Lemke H., and Weber J. (1999): Phys. Rev. B **59**, 5551

Yau L. D., Smiley C. F., and Sah C. T. (1984): Phys. Status Solidi A **13**, 457

Young R. C. and Corelli J. C. (1972): Phys. Rev. B **5**, 1455

Zhang S. B. and Northrup J. E. (1991): Phys. Rev. Lett. **67**, 2339

Zimmermann H. and Ryssel H. (1991): Appl. Phys. Lett. **58**, 499

Zunger A. and Cohen M. L. (1978): Phys. Rev. Lett. **41**, 43