First Principles Modelling of Dislocations in BCC Iron

Preetma Kaur Soin

Submitted for the fulfllilment of the degree of Doctor of Philosophy

Materials

Imperial College London

June 2012

Preetma Kaur Soin: *First Principles Modelling of Dislocations in BCC Iron,*
Submitted for the fulfliilment of the degree of Doctor of Philosophy, ©
June 2012

ABSTRACT

This work centers around employing magnetic tight binding to study defects in body centred cubic iron, a material of potentially great importance as a structural component in hydrogen fusion power plants. An existing d-band tight binding model was extended to include charge. The first task was to develop a working scheme for calculating spins and charges self consistently. This was achieved using an extended form of the Harris-Foulkes functional and implemented with a generalisation of the Newton-Raphson minimisation procedure. Having established such a scheme it was tested on bulk structures, point defects and straight dislocations. The motion of dislocations and the opposition to this was also considered though calculation of Peierls barriers.

DECLARATION

I declare that the work presented here is my own, conducted whilst in the Department of Materials, Imperial College London. Some of the material presented has been included in the two publications listed.

London, June 2012

Preetma Kaur Soin

for

Jee and Dadi Ma

my Soin grandparents

Sardar Baldip Singh

A true Gentleman of the old school

and

Sardarni Govinder Kaur

A most gracious and courageous Lady

PUBLICATIONS

- Soin P, Horsfield AP, Nguyen-Manh D, Efficient self-consistency for magnetic tight binding, COMPUTER PHYSICS COMMUNICATIONS, 2011, Vol:182, Pages:1350-1360.
- Preetma K. Soin, Andrew P. Horsfield and Duc Nguyen-Manh (2011). Magnetic Tight-Binding Simulations of Defects in Iron. MRS Proceedings, 1298 , mrsf10-1298-q06-10 doi:10.1557/opl.2011.428

ਪਤਿ ਪਤਿ ਗਡੀ ਲਦੀਅਹਿ

A man may load carts with books;

ਪਤਿ ਪਤਿ ਭਰੀਅਹਿ ਸਾਥ ॥

He may load men with books to take with him;

ਪਤਿ ਪਤਿ ਬੇਤੀ ਪਾਈਐ

Books may be put on boats;

ਪਤਿ ਪਤਿ ਗਡੀਅਹਿ ਖਾਤ ॥

Pits may be filled with them.

ਪਤੀਅਹਿ ਜੇਤੇ ਬਰਸ ਬਰਸ

A man may read books for months;

ਪਤੀਅਹਿ ਜੇਤੇ ਮਾਸ ॥

He may read them for years;

ਪਤੀਐ ਜੇਤੀ ਆਰਜਾ

He may read them for life;

ਪਤੀਅਹਿ ਜੇਤੇ ਸਾਸ ॥

He may read them while he hath breath·

ਨਾਨਕ ਲੇਖੈ ਇਕ ਗਲ

Nanak, only one word, God's name, would be of account;

ਹੋਰੁ ਹਉਮੈ ਝਖਣਾ ਝਾਖ ॥੧॥

All else would be the senseless discussion of pride.

ACKNOWLEDGMENTS

I would like to thank my supervisors Dr Andrew Horsfield and Dr Duc Nguyen-Manh for the vast amount of help they have given me over the course of my work. I would also like to thank Prof Adrian Sutton for his input during the early part of my project. This PhD research was conducted with the support of EPSRC and the Culham Centre for Fusion Energy and I gratefully acknowledge this.

This work could not have been completed without my family and friends. I would like to thank my numerous cousins for all the fun we have had together which provided a good balance with my PhD work. The support of my friends from Imperial has been invaluable, in particular Wei li, Hannah and Angela and the hundreds of lunches we have enjoyed together. My thanks also go to my oldest friend Louise and the Kenny family of Adams farm who put up with me for weeks at a time when I needed to get away from London. Most of all I want to thank my parents, Pavitra and Randeep for their unconditional and selfless support.

CONTENTS

1 INTRODUCTION 17
- 1.1 The problem? - Hydrogen Fusion Power Plants 17
- 1.2 Summary of thesis 21

2 COMPUTATIONAL METHODOLOGY OF STUDYING DEFECTS IN BCC IRON 24
- 2.1 What is known about defects in bcc Iron? 24
  - 2.1.1 Point Defects 24
  - 2.1.2 Dislocations 27
- 2.2 Methodology 33
  - 2.2.1 Techniques Part 1: Electronic Structure Methods 34
  - 2.2.2 Tight Binding 39
  - 2.2.3 Techniques Part 2: Empirical potentials and linear elasticity 42

3 EFFICIENT SELF CONSISTENCY FOR MAGNETIC TIGHT BINDING 48
- 3.1 The Model 48
- 3.2 The Problem 53
- 3.3 The Solution 55
  - 3.3.1 Charge/Spin self consistency using the extended Harris-Foulkes functional 55
  - 3.3.2 Newton-Raphson Mixing 57
- 3.4 Calculating the Linear response matrix 60
- 3.5 Pulay mixing and the linear response matrix 65
- 3.6 The Newton-Raphson mixing scheme in action 66
- 3.7 Conclusions 70

4 POINT DEFECTS 71
- 4.1 Bulk Calculations 71
  - 4.1.1 Variation of energy with volume 71
  - 4.1.2 Variation of magnetic moment with volume 73
  - 4.1.3 Variation of energy with magnetic moment 75
  - 4.1.4 Elastic Constants 79
- 4.2 Point Defects 81
  - 4.2.1 The Vacancy and the effect of the Hubbard U 82
- 4.3 Interstitials 85
  - 4.3.1 Energetic ordering of interstitials 85
  - 4.3.2 A close look at the spins 86
  - 4.3.3 Technical difficulties presented by the interstitials 90
- 4.4 Conclusions 96

5 $\frac{1}{2} < 111 >$ SCREW DISLOCATIONS 97
- 5.1 Simulating the $\frac{1}{2} < 111 >$ Screw Dislocation in BCC iron 97
- 5.2 Easy and hard core dislocations 99
- 5.3 Is the easy core structure degenerate or non-degenerate? 101
- 5.4 The hard core structure 105
- 5.5 Further Gamma Surface Calculations 106
  - 5.5.1 $(1\bar{1}0)$ Gamma surface 108
  - 5.5.2 $(11\bar{2})$ Gamma Surface 111
- 5.6 The simulation geometry 115
- 5.7 1-D Electrostatics 120

5.8 Variation in spins around dislocation centre 125
5.9 Variation in formation energy of easy core dislocation with radius of wire 126
5.10 Convergence of hard and easy core energies 127
5.11 Conclusions 128
6 PEIERLS BARRIERS 129
6.1 Peierls Barrier Calculations 129
6.1.1 What is a Peierls Barrier? 129
6.1.2 Initial Setup and the drag method 130
6.1.3 The Shape of the Peierls Barrier 132
6.1.4 DD plots - a closer look at what is happening 133
6.2 Parameters that can be varied 138
6.2.1 The effect of freezing 138
6.2.2 The effect of k point sampling 139
6.3 Minimised Barriers 142
6.3.1 A closer look at what is happening 143
6.4 Motion of the five significant atoms 146
6.5 Magnetic contribution to the energy of the five significant atoms 149
6.6 What is the structure? 152
6.7 Calculation of Peierls barrier using new parameters 153
6.8 Conclusions 154
7 CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK 156

BIBLIOGRAPHY 161

LIST OF FIGURES

Figure 1 Diagram of ITER fusion machine [1]. 19
Figure 2 Idealised Stress-Strain Graph 21
Figure 3 Schematic diagram of Screw Dislocation. The purple arrow is an atom to atom path around the centre of a dislocation. The red cross indicates the centre of the dislocation; the dark blue arrow indicates the dislocation line and the small red arrow the Burgers vector. . Adapted from [2]. 28
Figure 4 Schematic diagram of double kink nucleation and propagation. The green arrows indicate the direction in which the kinks expand. 30
Figure 5 Explaining Stoner model of magnetism - a simplistic picture. The right hand section of the figure indicates spin up density of states, and the left hand section the spin down. μ labels the chemical potential. 49
Figure 6 The three bond integrals 52
Figure 7 Repulsive Energy 53
Figure 8 Linear mixing 54
Figure 9 The left panel shows the variation of the first and second derivatives of energy with magnetic moment for the low spin state of bulk ferromagnetic fcc iron at a volume per atom of 10.65$\AA^3$. The first derivative is labelled E' and the second derivative E''. The right panel shows the spin residue which determines by how much the input spin is adjusted for one update. First published in [82] 67
Figure 10 Newton Charge mixing verses Pulay charge mixing for a 137 atom bulk bcc iron wire 69
Figure 11 The axes in an hcp structure 72
Figure 12 Binding energy curves for fcc and bcc iron for a variety of magnetic structures. First published in [82] 75
Figure 13 Binding energy curves of for hcp and bcc iron for a variety of magnetic structures. First published in [82] 76
Figure 14 Variation of magnetic moment per atom with atomic volume for fcc and bcc iron. First published in [82] 77
Figure 15 Variation of magnetic moment per atom with atomic volume of high spin hcp and bcc iron. First published in [82] 78

Figure 16 The left panel shows the variation of the first and second derivatives of energy with magnetic moment for bulk ferromagnetic bcc iron. The first derivative is labelled E' and the second derivative E''. The right panel shows the ratio $-E'/E''$ which determines by how much the input spin is adjusted for one update. The red and green vertical asymptotes in the right panel separate regions where the Newton-Raphson iteration tends to different solutions. First published in [82] 79
Figure 17 Left Panel: Variation in energy with magnetic moment for bulk bcc ferromagnetic iron. The black dots represent the real curve and the red line is the fourth order fit to this. Right Panel: The first and second derivatives of the fourth order.First published in [82] 80
Figure 18 Variation in charge around a single vacancy in a 53 atom cell. The neighbour index indicates the shell around the vacancy with increasing radius. First published in [82] 84
Figure 19 Percentage Displacement around the 53 atom vacancy 84
Figure 20 Magnetic moment around the 53 atom vacancy 85
Figure 21 Variation in magnetic moment with distance from each of the three dumbbell interstitials 93
Figure 22 Jmol illustration of $\langle 110 \rangle$ interstitial. The dumbbell atoms are show in blue, the atoms corresponding to the sharp dip in spins corresponds to the pink atoms. All other atoms are shown in orange. The x, y and z axis form a right handed cartesian coordinate system. 94
Figure 23 Jmol illustration of $\langle 111 \rangle$ interstitial. The dumbbell atoms are show in blue, the atoms corresponding to the sharp dip in spins corresponds to the pink atoms. All other atoms are shown in orange. The x, y and z axis form right handed cartesian coordinate system. 95
Figure 24 [111] view of bcc iron. The pink atoms correspond to a coordinate of 0 along the $\frac{1}{2}$ [111] axis, the blue to a coordinate of $\frac{1}{3}$ and the green to a coordinate of $\frac{2}{3}$. The white arrows indicate the direction of spiral of the atoms in adjacent triangles. A clockwise arrow indicates spiralling out of the page and an anticlockwise arrow spiralling into the page. 98
Figure 25 Bulk, easy and hard core dislocations profile along the $[1\bar{1}0]$ direction. 100
Figure 26 The equivalent $< 112 >$and $< 110 >$ directions in bcc crystals. Dashed lines represent one possible set of directions for spreading to occur and the bold lines the other. 103

Figure 27 Degenerate core structure DD diagram Wang et al. [94]. The shades of gray represent the three displacements in the [111] direction of the bulk as described in section 3.1. 103
Figure 28 Easy Core (DD) plot for non-degenerate structure generated by simulations conducted in this work. The colours represent the three displacements in the [111] direction of the bulk as described in section 3.1. Axis are as in Figure 25. 103
Figure 29 Schematic picture of 2D projection of cells sliding over each other to produce gamma surfaces 104
Figure 30 $(1\bar{1}0)$ gamma surface in the [111] direction 105
Figure 31 Hard Core differential displacement (DD) plot. The colours represent the three displacements in the [111] direction of the bulk as described in section 3.1. 106
Figure 32 (100) gamma surface. The x axis in in the [010] direction and the y axis in the [001] direction. 107
Figure 33 $(1\bar{1}0)$ gamma surface - $[11\bar{2}]$ direction 109
Figure 34 The magnetic moments across the gamma surface for an atom on the surface of the simulation cell (S1) and for an inner atom (S2) in both unrelaxed and relaxed cases. The coordinate labels the step in the gamma surface. 110
Figure 35 Variation of first nearest neighbour bond lengths and the averge of these bond lengths for a surface atom at each reaction coordinate of the fully relaxed $(1\bar{1}0)$ gamma surface in the $[11\bar{2}]$ direction. The coordinate labels the step in the gamma surface. 111
Figure 36 Local density of states at each reaction coordinate of the fully relaxed $(1\bar{1}0)$ gamma surface in the $[11\bar{2}]$ direction for an inner atom of the simulation cell with spin S2. The red curve is the up spin density of states and the black curve the down spin density of states. 112
Figure 37 $(11\bar{2})$ gamma surface in the $[1\bar{1}0]$ direction 113
Figure 38 $(11\bar{2})$ gamma surface in the $[1\bar{1}0]$ direction relaxed 114
Figure 39 $(11\bar{2})$ gamma surface [111] direction 114
Figure 40 Cylindrical shape whisker. Orange atoms are allowed to relax, purple atoms are frozen. 117
Figure 41 Hexagonal Cross sectional whisker. Orange atoms are allowed to relax, purple atoms are frozen. 118
Figure 42 Variation in easy core formation energy with outer radius (defined as half the distance between two diagonally opposite corners) 118
Figure 43 Variation in hard-easy core formation energy difference with outer radius (defined as half the distance between two diagonally opposite corners) 119
Figure 44 1-D Boundary conditions 120
Figure 45 Change in the charge on an atom in a cylindrical whisker as inter-site calculation progresses 122

Figure 46 Change in magnetic moment on an atom in a cylindrical whisker as inter-site calculation progresses 123
Figure 47 Inter-site local densites of states using 1-D boundary conditions 124
Figure 48 On-site local densities of states using 1-D boundary conditions 125
Figure 49 Spins around an easy core dislocation. Excluding orange atoms, all other atoms with the same colour have the same magnetic moment. Centre of dislocation in centre of triangle of pale pink atoms. 126
Figure 50 Variation in easy core dislocation energy with ln of outer radius (R) of hexagonal whisker 127
Figure 51 Hard/easy core energy convergence with ln of outer radius (R) of whisker 128
Figure 52 Jmol picture showing ellipse configuration. Dislocation is dragged between the centres of the left and right most purple triangles. 133
Figure 53 Jmol picture showing hexagon configuration. Dislocation is dragged between the centres of the left and right most purple triangles 134
Figure 54 Comparison of the Peierls barrier of elliptical and hexagonal shapes 135
Figure 55 DD plot of configuration 0 for original ellipse Peierls barrier 135
Figure 56 DD plot of configuration 1 for original ellipse Peierls barrier 136
Figure 57 DD plot of configuration 2 for original ellipse Peierls barrier 136
Figure 58 DD plot of configuration 3 for original ellipse Peierls barrier 136
Figure 59 DD plot of configuration 4 for original ellipse Peierls barrier 137
Figure 60 Contributions to the shape of the Peierls barrier for the original elliptical shape 138
Figure 61 Method of freezing atoms in ellipse calculations 140
Figure 62 Effect of Freezing for ellipse calculations 141
Figure 63 Peierls barrier for 169 atom hexagon with both 16 and 32 k points in the $< 111 >$ direction. 141
Figure 64 Minimised barrier using ellipse shape 142
Figure 65 Minimised barrier using hexagon shape 143
Figure 66 DD plot of configuration 0 144
Figure 67 DD plot of configuration 1 144
Figure 68 DD plot of configuration 2 144
Figure 69 DD plot of configuration 3 145
Figure 70 DD plot of configuration 4 145
Figure 71 Relative change in displacement of atoms of each atomic configuration along the Peierls barrier for the ellipse shape whose endpoints have been minimised. 147

Figure 72 Relative change in displacement of atoms of each atomic configuration along the Peierls barrier for the hexagon shape whose endpoint have been minimised. 148
Figure 73 Ellipse - Magnetic contribution to the energy of the five significant atoms 150
Figure 74 Hexagon - Magnetic contribution to the energy of the five significant atoms 151
Figure 75 Dislocation in the middle of a bond. The bright pink dot marks the centre of a bond. 152
Figure 76 DD plot for dislocation placed initially in the middle of a bond 153
Figure 77 Peierls barrier using latest tight binding parameters. 154

LIST OF TABLES

Table 1 Number of loops to converge bulk 67
Table 2 Finding an efficient mix factor. The number of loops (min) refers to the number of loops required to leave the local energy minimum. 68
Table 3 Comparison of equilibrium atomic volumes and magnetic moment per atom for various bulk structures. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82] 74
Table 4 Comparison of Elastic Constants in GPaDFT values from [41] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82]. 81
Table 5 Comparison of Formation Energy of Single Vacancy. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82] 83
Table 6 Variation in Formation Energy of Vacancy with value of Hubbard U. First published in [82] 83
Table 7 Formation Energies of Interstitials. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82] 87
Table 8 Charge and Spin on Dumbbell atoms of interstitials. Intersite refers to the inclusion of the long range Coulomb interaction for charge, this is not inluded in onsite calculations.. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82] 88
Table 9 Sensitivity to initial spin for the three interstitials. All calculations carried out using the tight binding inter-site description of charge. 91

Table 10 Initial energy per atom 108
Table 11 Success of both Newton and Pulay Mixing with different boundary conditions. Yes = successful calculation 121
Table 12 Change in magnetic moments from bulk values around an easy core dislocation 126
Table 13 Bond lengths in the negative region of the pair potential for all possible configurations of the original Peierls barrier from the elliptical shape 137
Table 14 Equivalent atoms in hexagonal and elliptical shapes 146
Table 15 Comparison of relaxed energies for ellipse with dislocations at different initial positions 153

1

INTRODUCTION

1.1 THE PROBLEM? - HYDROGEN FUSION POWER PLANTS

This work is motivated by its applications to hydrogen fusion power plants and their aim to provide clean energy. The problem of radiation damage caused to materials used in the fusion reactor in this process is a complex one. Seeking a better understanding of this problem through simulation is extremely necessary as it has important consequences for the mechanical properties of the materials [22]. There is a limit to what can be learnt experimentally at the current time because of the need to recreate the extreme conditions in a fusion power plant. The fusion process being considered is indicated in eqn 1.1.

$$^{2}_{1}H + ^{3}_{1}H \rightarrow ^{4}_{2} He + n \tag{1.1}$$

Essentially the isotopes of hydrogen (deuterium and tritium) fuse to form helium nuclei and emit fast neutrons (14.1 MeV) as happens in the sun. The helium nuclei remain inside the plasma heating it up and allowing the reaction to continue. The neutrons, however, can escape and are primarily responsible for the structural damage to the walls of the reactor. They can also produce hydrogen and helium which contribute to the damage. The neutrons do have a useful function which is to breed a supply of tritium from a lithium blanket as described in eqn 1.2.

$$^{1}_{0}n + ^{7}_{3} Li \rightarrow ^{4}_{2}He + ^{3}_{1}H + ^{1}_{0} n \tag{1.2}$$

Much thought needs to be given to the choice of material used in the fusion reactor. The plasma is to be confined magnetically for a significant amount of time (around 400s) in the toroidal vacuum which is common to the Tokomak design of reactors. This is important

because for the fusion power plant to function the energy produced by the reaction including losses must be great enough to keep the reaction going without any external power. The Lawson criterion [42] for this to occur depends on the product of three things: the density of electrons n_e, the time for which the plasma can be confined t_p and the temperature of the plasma T_p as indicated in eqn 1.3.

$$n_e T_p t_p \geqslant \frac{T_p^2}{\langle \sigma v \rangle} \frac{12 k_B}{E_{cp}} \tag{1.3}$$

where k_B is the Boltzmann constant, E_{cp} is the energy of the charged particles, v is the relative velocity of the particles and σ is the fusion cross section.

Body centred cubic (bcc) transitions metals are prime candidates to be used in the plasma facing components [11, 12] as illustrated in Figure 1 and the divertor in particular. The function of the divertor is to remove heat and helium from the reactor and it is placed around the bottom of the vessel. The bcc transition metals have a high atomic density without being too heavy. This means that they can stop the neutrons from escaping and still be used realistically as structural materials. The Joint European Torus (JET) is the current experimental device of this type and the ITER reactor is planned for the future. High temperatures (>600^{o}C) and stresses (300GPa) are experienced by the plasma facing components and the materials employed must be able to withstand these extreme conditions. The materials of a reactor also feel stress under their own weight. To prevent deformation in response to this stress, a high creep strength is required. The specific class of materials which are currently being investigated for use in the reactor are ferritic/martensitic steels with 9% chromium added.

The damage to these materials from the fast neutrons is measured in displacements per atom (dpa). This refers to the number of times on average that an atom has been displaced from its equilibrium position from the beginning of the reaction. This is estimated to be 150 dpa for a fusion reactor. This results in both point defects and line defects including edge and screw dislocations. The point defects produced can move and come together with the helium produced by the neutron damage, leading to embrittlement of the material. An edge dislocation

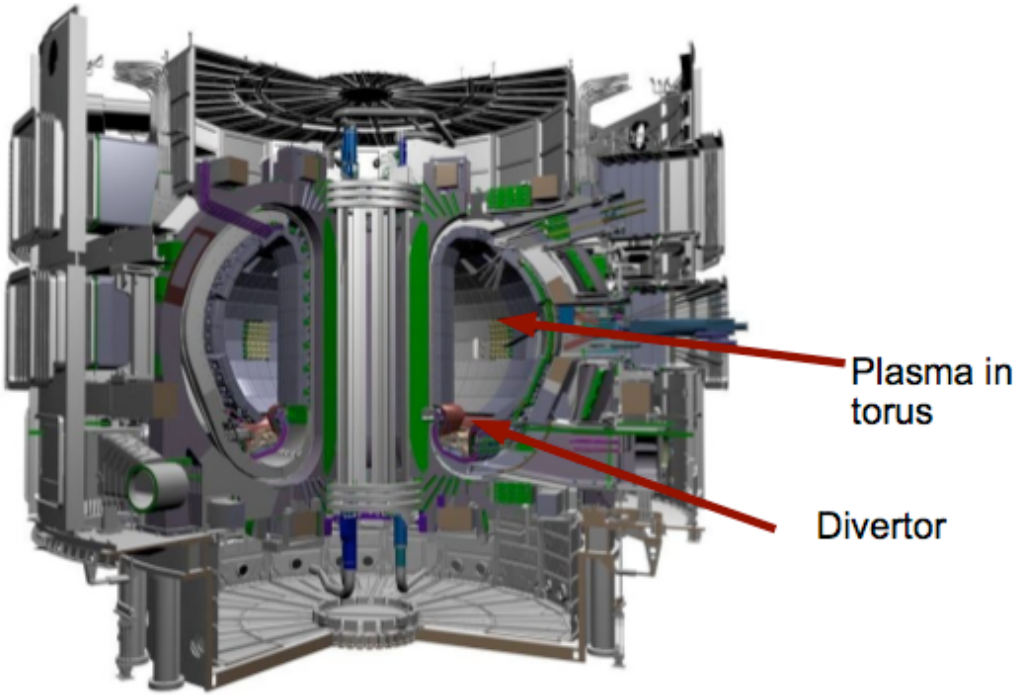


Figure 1: Diagram of ITER fusion machine [1].

is created when an extra half plane of atoms is inserted in the crystal. Irradiation produces dislocation loops which have the same effect. The transition between the brittle and ductile regime is of utmost importance in choosing the material. It is desired to have the lowest brittle to ductile transition temperature that is possible, as brittle fracture of a power plant would be catastrophic. For this reason chromium is added to the steel, as at a certain concentration (9%) it lowers the brittle-ductile transition temperature under irradiation [58]. It is also added as it has the property of making steel stainless, thus reducing the propensity to rust and corrode.

This is where dislocations become important in the general picture. It is the motion of screw dislocations which controls the plasticity of the material, as edge dislocations are much more mobile in bcc materials. Without the presence of dislocations the applied stress required for slip is very high. Slip is the process whereby applying an external force causes planes of the crystal to glide over each other. It is the motion of screw dislocations via the mechanism of double kink nucleation (Chapter 2) which allows slip to occur and hence affects the plasticity.

Dislocations can interact with defects caused by irradiation and other alloying elements in the iron such as the added chromium. Screw dislocations can become pinned by these defects, while edge dislocations remain more mobile, emphasising why screw dislocations are the crucial factor in controlling the plasticity. The questions to be answered ultimately relate to the motion of dislocations and how this is affected by their interaction with defects.

Iron is a magnetic material and it is necessary to account for this. This is achieved using a tight binding framework for magnetic iron using a d-band Stoner model. The short range covalent type bonds in iron make it a good candidate for tight binding. In iron there is a narrow d-band and a broad s-p band. The d-band dominates the magnetism and variation of bonding with atomic coordination, justifying our use of the d band model. Point defects such as vacancies and self interstitials and static $\frac{1}{2}\langle 111\rangle$ screw dislocations are considered in detail in this work. However, the goal for the future is to understand what affects the motion of dislocations, so values for the formation and migration energy of kinks will also be important.

Figure 2 is an idealised plot of stress against strain. As indicated, at low strains it is expected that the stress and strain will be linearly proportional to each other (red region of graph). This is simply an expression of Hooke's law familiar within everyday life, i.e the greater the load applied the greater the extension. However, at a certain value of strain this is is no longer the case and two situations can result. Either the material can snap, a phenomenon known as brittle fracture (green cross on graph) or it can deform plastically i.e. in a non-linear fashion (blue region of graph). Materials which are ductile deform plastically to a large extent before they fracture whereas brittle materials demonstrate very little plastic deformation before this occurs.

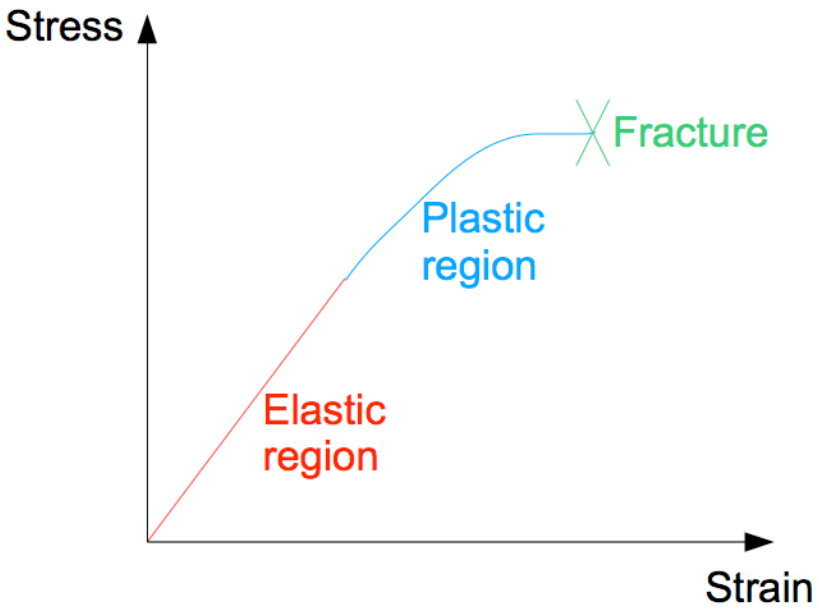


Figure 2: Idealised Stress-Strain Graph

1.2 SUMMARY OF THESIS

This work can be split into two broad areas. The first relates to developing a methodology for finding the atomic charges and spins self consistently using an existing d-band tight binding model for iron, extended to include charge as well as spin. The second involves utilising this novel method to investigate defects in iron, both point defects and the $\frac{1}{2} < 111 >$ screw dislocation.

Initially a brief literature review is conducted which surveys existing work in the area of both point defects and the structure of $\frac{1}{2} < 111 >$ screw dislocations and their movement. Studies using a variety of methods are considered, from the most accurate density functional theory methods (DFT) and tight binding through to interatomic potentials and elasticity theory. The final section of this review deals with with these methods in detail. In particular the derivation of non-self consistent tight binding from DFT is illustrated. The Finnis-Sinclair interatomic potential is looked at as an example of empirical potentials and the merits of this simple method are considered. Finally recent applications of elasticity theory are described.

Chapter 3 describes the technical details of the new self consistency method. Initially the parts of the tight binding model are discussed, in particular the magnetic Stoner model and the description of atomic charge using the Hubbard U. The problems with the original method of convergency using linear mixing are highlighted. The Harris-Foulkes functional initially described in the literature review in Chapter 2 is extended to include both charge and spin. This is done in such as way as to include input and output spin and charge. Minimising the total energy given by this functional, correctly gives a solution in which the input and output charges and input and output spins are equal to each other. In this way achieving self consistency now becomes a variational problem which can be placed on a sound mathematical footing using a generalisation of the Newton-Raphson procedure.

Chapter 4 essentially tests the work from Chapter 3 on various bulk structures and point defects. For the bulk structures the variation in energy and magnetic moment with volume are ascertained. The point defects considered are the mono-vacancy and the three self interstitials, and it is found that there is good agreement with DFT in all cases. Finally any further difficulties with convergence due to spin in particular are investigated.

Having established the efficacy of the method on small test problems, the $\frac{1}{2} < 111 >$ screw dislocation is then considered in chapter 5. There are two geometrically different type of dislocation possible termed hard and easy core, the easy being the lower in energy. Initial calculations establish the atomic core structure of the easy core dislocation both visually by using differential displacement maps and though use of gamma surfaces. It is found that the easy core structure obtained upon relaxation is non-degenerate in agreement with DFT. The philosophy behind using the geometry which we set-up is discussed, resulting in the choice of a hexagonal whisker with free surfaces. The atoms are displaced isotropically and the three outer layers are frozen in an attempt to eliminate the affect of the surface. The use of a whisker requires a modification of the method described in Chapter 3 to deal with a one dimensional system. The difficulties of using 1-D electrostatics with

a full inter-site description of charge are discussed. The relationship between the radius of the whisker and the easy core formation energy is considered. Also the relationship between the radius and the difference in the hard core and easy core energy is investigated.

Chapter 6 investigates the barrier to dislocation motion. Whiskers are constructed such that an easy core dislocation can be placed in two geometrically equivalent positions and the energy at these two positions is found to be identical. The dislocation is then moved between these positions and the energy at a series of intermediate positions is noted. Two shapes of whisker are considered, one hexagonal and one roughly elliptical. The shape of the barrier is not what would be expected from DFT, however. This is possibly due to the parameterisation of the pair potential in the tight binding model used; further details of this are given in the chapter 6.

The final chapter sums up the major conclusions of the project and provides suggestion for future work, both in terms of improving the scaling with number of atoms of the method described in Chapter 3 and in considering more complicated systems for analysis.

2

COMPUTATIONAL METHODOLOGY OF STUDYING DEFECTS IN BCC IRON

Having motivated the study of ferritic-martensitic steels as structural materials for hydrogen fusion power plants in Chapter 1, a more detailed discussion of the possible defects that can occur in iron and the methods of modelling them is now embarked upon. Initially point defects are considered and various conclusions about their stability and migration using different methods of analysis are drawn. The discussion is then extended to screw dislocations. In particular the nature of their core structure and the mechanism of their motion are considered. Finally the computational methodology of calculation of these defects is dealt with going from the most accurate to the least accurate.

2.1 WHAT IS KNOWN ABOUT DEFECTS IN BCC IRON?

2.1.1 *Point Defects*

These are the simplest defects in bcc iron. They are of two types, known as vacancies and interstitials. Vacancies involve the removal of an existing atom, whilst interstitials involve removing an atom and placing two atoms equidistant on either side of the position of the original atom. These two atoms are known as a dumbbell and three types of interstitials are considered depending on the line along which the two atoms making up the dumbbell lie. In this work the atoms that make up the dumbbell are both iron, so the interstitials are sometimes referred to as self interstitials. Formation energies are important quantities that can be calculated for point defects from both experiment and simulation. The formation energies of these dumbbells have a specific order which

is agreed upon by experiment [88] and simulation carried out using both DFT (density functional theory) [37] and tight binding [59]. The ⟨100⟩ interstitial is the highest in energy, followed by the ⟨111⟩, with the lowest being the ⟨110⟩. There is a certain range in the value of the vacancy formation energy depending on the method used. The experimental, interatomic potential [10] and tight binding method give values around 1.5eV to 1.6eV and DFT a higher value of roughly 2eV.

Where DFT simulations becomes particularly useful is in gaining information about the magnetic structure of the defects. In the case of the dumbbell atoms it is indicated by DFT that the magnitude of the magnetic moment on all dumbbell atoms is much less than in the bulk, around $0.2\mu_b$ compared to $2.62\mu_b$. In addition the moment on the ⟨110⟩ atoms is antiparallel to that of the bulk [20] surrounding it. Tight binding simulations also register the same effect [59].

Whilst mainly vacancies are present in equilibrium conditions in bcc iron, interstitials become as common under irradiation. This is because Frenkel pairs consisting of a single vacancy and single interstitial are produced. The movement of these defects created by irradiation is extremely pertinent to the use of iron as a structural material for hydrogen fusion power plants as described in the introductory chapter. Studies to determine the migration energy of single vacancies are conducted using both experiment [79, 90] and simulation. There is some argument between the two approaches as experimentally it is difficult to obtain a completely pure sample of iron without the presence of impurities. These impurity atoms can trap the vacancies and contribute to a fairly wide range of migration energies. A study by Ackland et al. [4] considered the introduction of copper atoms in particular, using an interatomic potential for an Fe-Cu system. The study of copper precipitates is important as they can contribute to hardening of the metal. Ackland found that in dilute solutions the vacancy and the copper atom are attracted to each other, as is the interstitial to the copper atom. However, for the interstitial the binding energy is found to be less than half of that of the vacancy.

In the case of interstitials it is found that their migration energy depends on the energy difference between the ⟨111⟩ and ⟨110⟩ interstitials. The actual steps in the migration process are complicated but rotations between these two types of interstitial strongly affect the path taken. Whilst the energetic ordering of the interstitials is well established, there is some disagreement over the magnitude of the ⟨111⟩/ ⟨110⟩ energy difference. Experiment seems to obtain a far larger value than empirical potentials [69]. The work from DFT fits in with what is observed experimentally.

As well as existing as single defects, vacancies and self interstitials are attracted to each other and migrate to form cluster of defects as documented by TEM (Transmission Electron Microscopy) [56]. Molecular dynamics studies of Soneda and de la Rubia [83] and Osetsky et al. [73] show that under irradiation, vacancy clusters are found to be less mobile than self interstitial ones but much more mobile than single vacancies. Experiment is not in complete agreement on this issue, with work by Yao et al. [100] suggesting that motion of vacancy clusters was relatively common and that of Arakawa et al. [7] concluding that self interstitial clusters migration was far more significant.

The actual nature of the self interstitial motion can be broadly split into two categories. One suggests a one dimensional scheme gaining support from work using interatomic potentials [23, 9] and the second a three dimensional migration indicated by DFT calculations [37]. This is linked to the earlier discussion about the disparity in the energy difference of the two interstitials using DFT and empirical potentials. The effect on cluster migration of adding copper atoms to pure iron is considered in the molecular dynamics calculations of [62] using the Ackland potential. This finds that the motion of interstitial clusters in pure iron is only three dimensional in character for a small number of constituent interstitials in a cluster (< 5) and above this number is found to be purely one dimensional. However, the introduction of copper is found to increase the three dimensional nature of the mechanism.

It seems that an approach is needed which uses the accuracy of ab initio calculations and the ability to consider many atoms and their dynamics to obtain an accurate picture of cluster motion. Recent work by Fu et al. [38] attempts to do this, by taking ab initio data and feeding it into kinetic Monte Carlo simulations. This approach agrees with DFT calculations whilst enabling more atoms to be considered.

2.1.2 *Dislocations*

Before launching in to a detailed discussion of calculations performed on dislocations it is necessary to understand the answer to the basic question: what is a dislocation? Dislocations are essentially line defects within a crystal of which there are two types: edge and screw. This work is primarily concerned with screw dislocations, the definition of which is illustrated in Figure 3. The dislocation line is perpendicular to the top crystal face. Any atom to atom path which encloses this line is referred to as a Burgers circuit. A similar path drawn in a perfect crystal would not close and the amount by which it does not close is called the Burgers vector. In the case of a screw dislocation this vector is parallel to the dislocation line, whilst for an edge dislocation the two are perpendicular.

Static Dislocations

Studies of dislocations in iron have been primarily conducted using simulation. These concentrate first on establishing the core structure of straight dislocations and then consider dislocation motion via kink propagation. One possible definition of what constitutes the core region of a dislocation relates to where linear elasticity theory breaks down, i.e. the atomic displacements become too large for the linear approximation. Images of dislocations have been obtained using transmission electron microscopy (TEM). However, it is not possible to resolve the core structure using this method.

There are two possible geometrical configurations of the core of the $\frac{1}{2}\langle 111\rangle$ screw dislocation. These two dislocation core types are referred

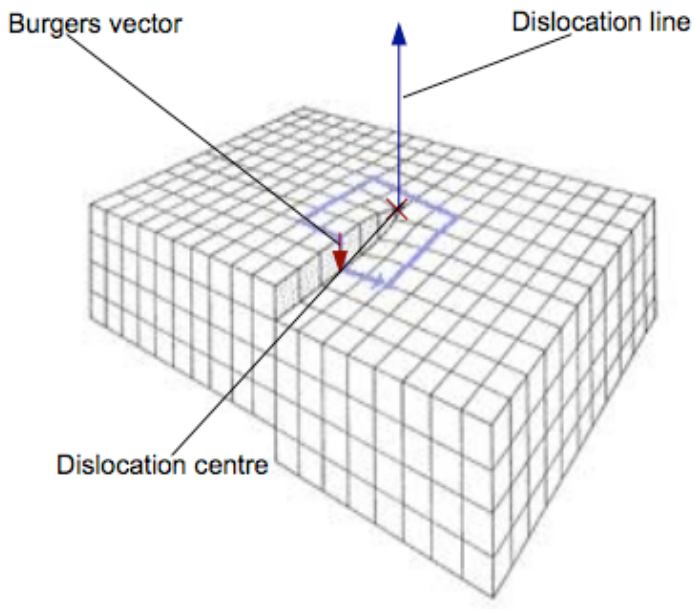


Figure 3: Schematic diagram of Screw Dislocation. The purple arrow is an atom to atom path around the centre of a dislocation. The red cross indicates the centre of the dislocation; the dark blue arrow indicates the dislocation line and the small red arrow the Burgers vector. . Adapted from [2].

to as hard and easy [99]. The structure of the easy core, which has been found to be consistently lower in formation energy than the hard [36, 17], has been a subject of some debate. There are two possibilities: a degenerate structure predicted by empirical potentials in which there are two atomic configurations that have the same formation energy or a symmetric non-degenerate core resulting from DFT and tight binding calculations. This is discussed fully in Chapter 5.

It is possible for the dislocation core to interact with impurity atoms. such as chromium as documented in [89], using a Fe-Cr interatomic potential [72] which predicts the non-degenerate core structure. This suggests that there is a weak interaction between the core and a single chromium atom and a strong repulsion between the core and a pair of chromium atoms. Earlier work of Farkas et al. [29] using an embedded atom potential which predicts a degenerate core also considers the introduction of impurity chromium atoms to dislocations in pure bcc iron. This mentions that the three fold symmetry of the core in pure iron is lost by the introduction of chromium. In terms of application to the

material properties of iron for use in fusion power plants the effect on double kink nucleation by the chromium atoms is of greater importance and is discussed in the next section on dislocation motion. Other studies consider the interaction of the dislocation core with copper. In particular the work of Harry and Bacon [44], which again uses a potential which predicts the degenerate core structure and suggests that the dislocation formation energy is reduced by the addition of copper precipitates.

Dislocation Motion

An understanding of the movement of screw dislocations in bcc metals is essential in a discussion about their plasticity. Dislocation motion occurs by the process of glide in which the dislocation moves in a plane containing both its Burgers vector and line. However, a significant amount of applied stress is required before the $\frac{1}{2}\langle 111\rangle$ screw dislocations move [46]. This high degree of immobility is due to their non-planar structure and means that they are the controlling factor in the plastic deformation of the iron used as a structural material in hydrogen fusion power plants. The precise mechanism of glide is provided by a double kink nucleation and propagation. Long straight dislocations do not glide as a whole as this would be energetically unfavorable. Instead as a result of random thermal fluctuations short segments of the dislocation move to a neighbouring position forming a double kink. This double kink then expands out dragging the rest of the dislocation with it as illustrated in Figure 4. The move is from an energetically favourable easy core position to a double kink in an energetically unstable hard core position. The same process then repeats itself to put the dislocation back in an easy core position. In this way the kink is propagated as indicated in Figure 4.

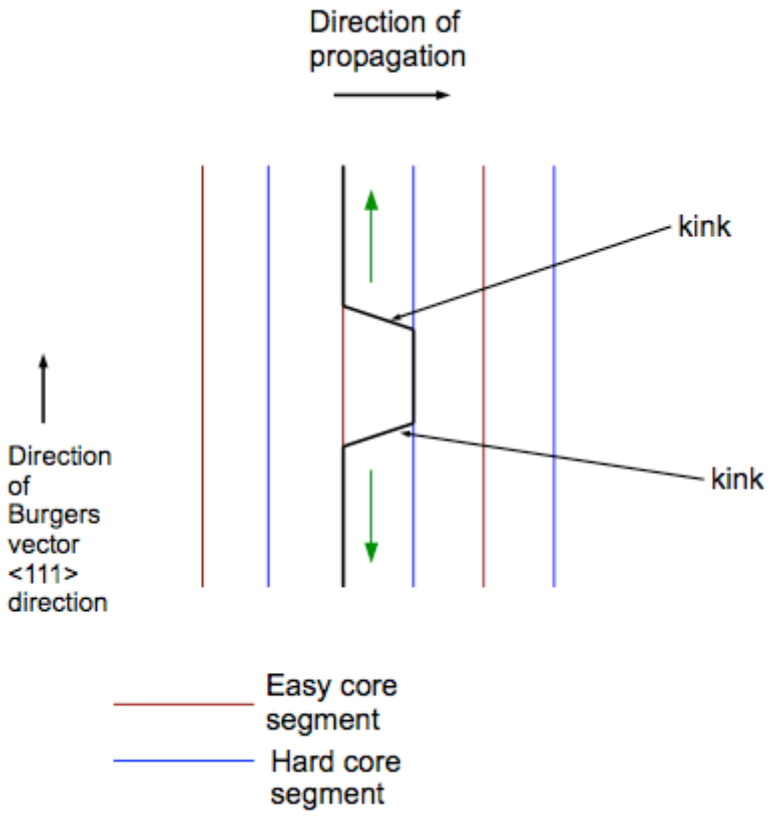


Figure 4: Schematic diagram of double kink nucleation and propagation. The green arrows indicate the direction in which the kinks expand.

Before beginning a discussion on kink motion via this mechanism it is instructive to consider static single kinks. Initial studies on these structures were conducted using interatomic potentials that result in a degenerate core structure. The two possible degenerate atomic configurations feed in to the possible types of kink. This results in a complicated situation where there were six possible types of kink [95, 68]. However, assuming a non-degenerate core structure, two types of kink can occur which are referred to as vacancy and interstitial. These can be differentiated by their effect on the row of atoms in the [111] direction, which are affected by the kink. In the case of the interstitial kink these atoms become closer by an amount that is a third of the Burgers vector and in that of the vacancy kink they become further apart by the same amount. A recent study by Ventelon et al. [92] has been carried out using the Ackland-Mendelev [5] empirical potential, and calculates the formation energies of both of these types of kink. It is found that the interstitial kink has a much higher energy than the vacancy, i.e. 0.57

eV in comparison to 0.08eV, the sum of these giving a total formation energy for a double kink of 0.65eV per kink which fits in well with experiment [14]. Use of the Ackland-Mendelev potential for this purpose does, however, have its limitations. Magnetism is not explicitly included and the shape of the Peierls barrier is double humped as opposed to single peaked. This results in a kink width which is significantly lower than experiment.

As well as the more sophisticated methods discussed above, there is still some scope for elasticity theory in the calculation of the formation energy of double kinks. The work of Terentyev and Malerba [89] considers the formation energy of double kinks in two situations, one where the connecting segment between the two single kinks is easy core in character and the other hard using an interatomic potential [72]. Now let E represent a section of easy core dislocation and H represent a section of hard core dislocation. Then it is found that the E-E-E double kink structure agrees well with elasticity theory, while the E-H-E does not because of the inability of elasticity theory to take proper account of the difference in core structures.

Whilst the nature of single kinks has been established, there are still questions to be answered about the exact mechanism of the motion of kinks. What is not known is the route the dislocation takes and where it ends up. Experimental observation of the accepted double kink nucleation mechanism is difficult. Recent work by Marian et al. [63] using molecular dynamics does however provide support for this model for dislocation movement. The Ackland potential [4] is used which incorrectly makes the easy core structures degenerate. This work is however significant because it extends this well established notion of dislocation motion over a range of stress and temperature regimes. Separate stress regimes are defined. Low stresses correspond to the established nucleation and propagation model. This flow of dislocations is defined as laminar (smooth) whilst at a higher range of stresses there is a very rough flow, resulting in phenomena such as cross kinks and debris loops. This work seeks to establish the qualitative concepts,

but is lacking a quantitative description of the effects of stress and temperature.

Extensions to these molecular dynamics (MD) simulations are provided by Domain and Monnet [21] who attempted to quantify the link between MD simulations and experimental results. Contrary to the study of Marian et al. [62], the Mendelev potential [64] is employed instead of that of Ackland. This predicts the correct non-degenerate structure of the easy core as this potential was partly designed to fit ab initio results. Differences may occur between the ab initio and Mendelev potential as it does not take into account the magnetic properties of iron. These are crucial to measuring accurately quantities such as the formation and migration energies of kinks. Domain and Monnet [21] found that screw dislocations only glide on $\{110\}$ planes over the whole range of considered temperatures. Marian et al. [63] also predicted glide on the $\{110\}$ planes. The motion of the dislocations described in Domain and Monnet [21], uses the traditional picture of double kink nucleation. However, further insights are provided concerning the path of the dislocation in relation to the correct hard and easy core structures. It is proposed that the double kink would move from an easy position to an unstable hard core position, resulting in the immediate nucleation of a second double kink to another easy core position.

The simulation set up in Domain and Monnet [21] consists of a screw dislocation in the centre of the simulation cell, periodic in the $[111]$ direction. The novelty in this work comes from the attempt made to quantify the effect of having free surfaces in the $\left[11\bar{2}\right]$ direction. Calculations with free surfaces do not have the disadvantage of those with dislocation dipoles and the interactions introduced by these; this is discussed in Chapter 5. Fixed planes of atoms were used in the upper and lower $\left(\bar{1}10\right)$ sufaces. Moving from the centre of the simulation box the dislocation would be attracted by the free surfaces. This is included in the simulation by adding in an extra stress component. A constant strain rate is applied, producing an initially elastic region in the stress strain curve followed by a plastic region in which there are several local maxima due to the probabilistic nature of double kink nucleation.

This probability is proportional to the length of the dislocation line and hence the number of nucleation sites. This approach to free surfaces seems reasonable and produces good agreement with compressive tensile tests carried out in experiment.

Chaussidon et al. [16] consider the differences between this type of boundary condition, known as rigid, in which free surfaces are only kept in one direction and those where all surfaces are kept free. A stress is applied parallel to the {112} planes and dislocation motion observed. In the case of rigid boundary conditions, glide occurred on the {112} planes in contrast to non-rigid boundary conditions which suggests glide on the {110} planes. The Peierls stress (stress required for a straight dislocation to overcome the energy barrier stopping it moving to a neighbouring equivalent lattice position) is found to be lower than in the non rigid boundary condition case. More recent work by Gilbert et al. [39] using MD simulations also suggests that the movement on the {112} planes is an artifact of the boundary conditions. In actuality the kink nucleation and propagation is only occurring on the {110} planes.

Impurity atoms can have an effect on dislocation motion. For example introduction of impurity elements such as chromium is reported by Terentyev and Malerba [89] to favour the formation of kinks as a result of the core splitting from interaction of the dislocation with chromium atoms. Earlier work by Farkas et al. [29] finds that the mobility of kinks is decreased in the presence of high chromium concentration although this conclusion may not be completely valid as the interatomic potential used predicts the degenerate core structure.

Whilst the story for static dislocations is relatively clear, the motion of dislocations over different stress regimes and in impure iron still requires more investigation for the potential use of iron as a structural material in fission and fusion power plants.

2.2 METHODOLOGY

The approach to the examination of the techniques is discussed in two broad categories. The first deals with electronic structure methods; that is DFT (density functional theory) and the derivation of tight bind-

ing from this via the Harris-Foulkes functional. The second considers both potentials and elasticity theory. The philosophy is to move systematically from the most accurate ab initio methods which can be used to study dislocations through to empirical potentials and to look at the relative merits of these.

2.2.1 *Techniques Part 1: Electronic Structure Methods*

The essential aim of density functional theory and tight binding is to solve the time independent Schrödinger equation for a system of atoms.

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle \tag{2.1}$$

$$\begin{aligned}\hat{H} &= \hat{T}_e + \hat{T}_n + V_{n-n} + V_{e-n} + V_{e-e} \\ &= \sum_i -\frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_I -\frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{4\pi\epsilon_0\left|\vec{R_I}-\vec{R_J}\right|} \\ &- \sum_{Ii}\frac{Z_I e^2}{4\pi\epsilon_0\left|\vec{R_I}-\vec{r_i}\right|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0\left|\vec{r_i}-\vec{r_j}\right|}\end{aligned}$$

where $\hat{H}$ is the hamiltonian operator and $\left|\Psi\right\rangle$ the wavefunction for all the electrons and nuclei in the system. The terms in the Hamiltonian are the kinetic energy of the electrons $\hat{T}_e$, the kinetic energy of the nuclei $\hat{T}_n$, the nucleus-nucleus interaction V_{n-n}, the nucleus-electron interaction V_{e-n} and the electron-electron interaction V_{e-e}. $\vec{R}_I$ is the position vector of a nucleus, Z_I is the nuclear charge on atomic site I, $\vec{r_i}$ is the position vector of electron i, e is the charge on an electron, m_e is the mass of an electron, M_I is the mass of nucleus I and ϵ_0 is the permittivity of free space. Certain helpful approximations can be made, the first being the Born-Oppenheimer approximation. As the nuclei are so much more massive than electrons, their motion can be ignored as the electrons adapt almost instantaneously to changes in the nuclear position. This is also known as the adiabatic approximation and means the kinetic energy of the nuclei can be ignored. The simplified Hamiltonian just for the electrons $\hat{H}_e$ can now be written as:

$$\hat{H}_e = \hat{T}_e + \sum_{Ii} -\frac{e^2}{4\pi\epsilon_0}\frac{Z_I}{\left|\vec{R_I}-\vec{r_i}\right|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0}\frac{1}{\left|\vec{r_i}-\vec{r_j}\right|}$$

$$= \hat{T}_e + \sum_i V_{ext}(\vec{r_i}) + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\vec{r_i}-\vec{r_j}|} \tag{2.2}$$

$V_{ext}(\vec{r}_i)$ is the external potential from the nuclei acting on an electron.

It is possible to put an upper bound on the energy of the system, from the Ritz variational principle. This states that the energy E of any trial wavefunction $|\psi\rangle$ is greater than or equal to the exact ground state energy.

$$E = \frac{\langle\psi|\hat{H}_e|\psi\rangle}{\langle\psi|\psi\rangle} \geqslant E_0 \tag{2.3}$$

So in principle this problem can be approximately solved by choosing an appropriate trial wave function and minimising the energy with respect to it. The first attempts at expressing this can be grouped into single particle theories in which the electrons are considered as independent particles in an effective potential. This is made up of the contribution from the external potential and a mean field contribution of all of the electrons. In the Hartree approximation this effective potential is described in eqn 2.4

$$V_{eff}(\vec{r}) = V_{ext}(\vec{r}) + \int \frac{e^2}{4\pi\epsilon_0}\frac{n(\vec{r'})}{|\vec{r}-\vec{r}'|}d^3\vec{r}' \tag{2.4}$$

where $n(\vec{r})$ is the electron number density. The term including $n(\vec{r}')$ in eqn 2.4 approximates the Coulomb potential experienced by one electron due to the presence of the others. The trial many body wavefunction is written as the product of the wavefunctions of all the N electrons in the system as illustrated in eqn 2.5.

$$|\psi\rangle = \phi_1(\vec{r_1})\phi_2(\vec{r_2})\phi_3(\vec{r_3})\ldots\ldots\ldots\phi_N(\vec{r_N}) \tag{2.5}$$

Using the trial ansatz in eqn 2.5, $n(\vec{r})$ can be written as eqn 2.6.

$$n(\vec{r}) = \sum_{i=1}^{N}|\phi_i(\vec{r})|^2 \tag{2.6}$$

The electrons stack from the lowest energy upwards according to the Pauli exclusion principle, with one electron of each spin in each orbital. The index i is defined to include spin as well and N is the total number of electrons.

The next stage is to minimise the energy E subject to the constraint that the wavefunctions of the single electrons are orthogonal to each other as in eqn 2.7.

$$\frac{\delta}{\delta\phi_i}\left(\langle\psi|\,\hat{H}_e\,|\psi\rangle - \sum_{ij}^{N}\mu_{ij}\left[\int\phi_i^*(\vec{r})\phi_j(\vec{r})d^3\vec{r} - \delta_{ij}\right]\right) = 0 \tag{2.7}$$

Inserting the expression for the Hamiltonian in eqn 2.2 into eqn 2.7, results in N single particle Schrödinger equations which have to be solved simultaneously for the N orbitals.

$$(\hat{T}_e + V_{ext}(\vec{r}) + \sum_{j\neq i}\int\left|\phi_j(\vec{r'})\right|^2\frac{1}{|\vec{r}-\vec{r}'|}\,d^3\vec{r}')\phi_i(\vec{r}) = \epsilon_i\phi_i(\vec{r}) \tag{2.8}$$

This approach has several salient omissions. Firstly it does not consider fully that electrons are fermions and hence their wavefunctions are antisymmetric upon exchange of a pair of particles and secondly it does not consider correlation. A more sophisticated method is to consider a trial ansatz for the many body wavefunction that is antisymmetric under exchange. This is achieved using the properties of determinants where the exchange of two rows or two columns changes the sign of the determinant. This is given in the expression for the Slater determinant in eqn 2.9.

$$|\psi\rangle = \frac{1}{(N!)^{\frac{1}{2}}}\begin{Vmatrix} \phi_1(r_1) & \phi_1(r_2) & \phi_1(r_3) & \cdots & \phi_1(r_N) \\ \phi_2(r_1) & \phi_2(r_2) & \phi_2(r_3) & \cdots & \vdots \\ \phi_3(r_1) & \phi_3(r_2) & \phi_3(r_3) & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_N(r_1) & \cdots & \cdots & \cdots & \phi_N(r_N) \end{Vmatrix} \tag{2.9}$$

It is now possible to go through a similar minimisation process to the Hartree case, resulting in a series of N single particle Schrödinger equations as detailed in eqn 2.10.

$$\left[\hat{T}_e + V_{ext}(\vec{r}) + \int\frac{e^2}{4\pi\epsilon_0}\sum_j\left|\phi_j(\vec{r}')\right|^2\frac{1}{|\vec{r}-\vec{r}'|}d^3\vec{r}'\right]\phi_i(\vec{r}) \tag{2.10}$$

$$-\int\frac{e^2}{4\pi\epsilon_0}\sum_j\phi_j^*(\vec{r'})\phi_j(\vec{r})\frac{1}{|\vec{r}-\vec{r}'|}\phi_i(\vec{r'})d^3\vec{r}' = \epsilon_i\phi_i(\vec{r})$$

The first term of the single particle Schrödinger equation in this case looks similar to the Hartree case, the only difference being that the sum over j does not exclude the possibility of an electron interacting with itself. The second term is new and is a result of the exchange interaction introduced by taking account of the spin statistics of fermions. This is still not a complete description as, despite taking into account antisymmetry, it does not include the Coulomb repulsion between the electrons. As mentioned previously both effects are required for an accurate inclusion of the correlation of electrons.

The question arises: what if it were possible to describe all the electrons using the electron number density $n(\vec{r})$ instead of in terms of the many body wavefunction? The two Hohenberg-Kohn theorems provide a basis for the answer to this. The first theorem states that the external potential and the ground state electron number density uniquely determine each other. So if the electron number density is known, then the external potential is known and it should in principle be possible to solve the Schrödinger equation to find the many particle wavefunction. Therefore the expectation value of the kinetic energy and electron-electron interaction can be written as a functional of the electron number density $n(\vec{r})$ as described in eqn 2.11.

$$F[n(\vec{r})] = \langle\psi|\,\hat{T}_e + V_{e-e}\,|\psi\rangle \tag{2.11}$$

The second Hohenberg-Kohn theorem derives from the variational principle and the first theorem and states that the energy as a function of the density is greater than or equal to the ground state energy E_0.

$$E[n(\vec{r})] = F[n(\vec{r})] + \int V_{ext}(\vec{r})n(\vec{r})d^3\vec{r} \geqslant E_0 \tag{2.12}$$

This should then just be an energy minimisation problem as encountered previously, subject to the constraint that the total number of electrons remains constant.

$$\delta\left(F[n(\vec{r})] + \int V_{ext}(\vec{r})n(\vec{r})d^3\vec{r} - \mu\left(\int n(\vec{r})d^3\vec{r} - N\right)\right) = 0 \tag{2.13}$$

Carrying out the minimisation leads to the expression in eqn 2.14.

$$\frac{\delta}{\delta n(\vec{r})}F[n(\vec{r})] + V_{ext}(\vec{r}) = \mu \tag{2.14}$$

The problem is that the form of the functional $F[n(\vec{r})]$ is not known. The pioneering approach of Kohn and Sham [57] is to consider a non-interacting system of electrons in an effective potential whose charge density is exactly the same as that of the fully interacting system. For this auxiliary system it is well known how to write the kinetic energy and the Hartree term (representing the Coulomb interaction) using Slater determinants as discussed above. The remainder of the energy for the interacting system that is not included in the non-interacting solution can be grouped into a term known as the exchange-correlation energy. The Euler-Lagrange eqn 2.14 is then reduced to eqn 2.15.

$$\frac{\delta T_e}{\delta n(\vec{r})} + V_H(\vec{r}) + V_{ext}(\vec{r}) + V_{xc}(\vec{r}) = \mu \tag{2.15}$$

where $V_H(\vec{r}) = \int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|} d^3\vec{r}'$ is the Hartree potential and $V_{ext}(\vec{r})$ is the external potential. $V_{xc}(\vec{r})$ is the exchange correlation potential, which is the functional derivative of the exchange correlation energy: $V_{xc}(\vec{r}) = \delta E_{xc}/\delta n(\vec{r})$. The form of the exchange correlation energy is not known, so approximations must be made. The form of eqn 2.15 is essentially describing a system of non-interacting electrons in an effective potential, where

$$V_{eff}(\vec{r}) = V_H(\vec{r}) + V_{ext}(\vec{r}) + V_{xc}(\vec{r}) \tag{2.16}$$

All that now needs to be done is to solve the single particle Schrödinger equations for a system of non-interacting electrons with Hamiltonian $\hat{H}_{ks} = \hat{T}_e + V_{eff}$. The solution gives the total energy for the interacting system by construction. The single-particle wave functions are the Kohn-Sham eigenstates and they represent the electron number density through eqn 2.6.

However, there is a complexity to consider. It can be seen that the effective potential depends on the density, and that the Schrödinger equations have to be solved to get the density. This results in the need for a self consistent iterative process to get the best possible density. This process can be computationally expensive and does not always scale well with system size [6]. There is a way in which it is possible not to have to calculate the solution iteratively. This is provided by the ideas of Harris and Foulkes which were worked out independently

[43, 34]. The total energy of the interacting system using the Kohn-Sham eigenstates is expressed in eqn 2.17.

$$E_{total} = \sum_i \langle \phi_i | \hat{H}_{ks} | \phi_i \rangle - \frac{1}{2} \int\int \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r} d^3\vec{r}' \tag{2.17}$$
$$- \int V_{xc}\left[n\left(\vec{r}\right)\right] n\left(\vec{r}\right) d^3\vec{r} + E_{xc}\left[n\left(\vec{r}\right)\right] + E_{nuc}$$

If we evaluate eqn 2.17 using a guess for the density $n_0\left(\vec{r}\right)$ we get eqn 2.18

$$E_{total} \approx \sum_i \langle \phi_i | \hat{H}_0 | \phi_i \rangle - \frac{1}{2} \int\int \frac{n_0\left(\vec{r}\right) n_0\left(\vec{r}'\right)}{|r - r'|} d^3r d^3r' \tag{2.18}$$
$$- \int V_{xc}\left[n_0\left(\vec{r}\right)\right] n_0 d^3r + E_{xc}\left[n_0\right] + E_{nuc}$$

where $\hat{H}_0$ is $\hat{H}_{ks}$ evaluated with density $n_0\left(\vec{r}\right)$ and the functions ϕ_i are eigenfunctions of $\widehat{H_0}$. This is the Harris-Foulkes functional. It can be shown that the error in the Harris-Foulkes expression for the total energy is second order in the error in the electron density.

2.2.2 *Tight Binding*

Tight binding can be derived from the Harris-Foulkes functional. Firstly the total energy eqn 2.17 is split into the attractive band energy of the first term and the remaining terms are grouped together into a repulsive energy term. The approach then used by tight binding is to express the Kohn-Sham eigenstates as a linear combination of atomic orbitals (LCAO) [49, 48, 47, 87, 40, 35]. The trial ansatz is indicated in eqn 2.19

$$|\psi\rangle = \sum_{I\alpha} a_{I\alpha} |I\alpha\sigma\rangle \tag{2.19}$$

where I is an atomic index, α is an orbital index and σ is a spin index. If this is substituted into the single-particle Schrödinger equation (eqn 2.1) it results in eqn 2.20.

$$\hat{H}_0 \sum_{I\alpha} a_{I\alpha} |I\alpha\sigma\rangle = E_n \sum_{I\alpha} a_{I\alpha} |I\alpha\sigma\rangle \tag{2.20}$$

Multiplying both sides of eqn 2.20 by $\langle J\beta\sigma|$ and then integrating over the spatial coordinate of the atomic orbitals gives eqn 2.21, which is known as the secular equation:

$$\sum_{I\alpha} a_{I\alpha}\langle J\beta\sigma|\hat{H}|I\alpha\sigma\rangle = E_n \sum_{I\alpha} a_{I\alpha}\langle J\beta\sigma|I\alpha\sigma\rangle \tag{2.21}$$

The term $\langle J\beta\sigma|\hat{H}|I\alpha\sigma\rangle$ is referred to as the hopping integral when $I \neq J$ and an on-site integral when $I = J$. The term $\langle J\beta\sigma|I\alpha\sigma\rangle$ is called the overlap integral. As the electrons are basically confined to quantum wells (the influence of the electron-nucleus attraction), the overlap integral decreases exponentially with distance.

The hopping integral determines the rate at which an electron moves from atom to atom. Hopping integrals are restricted to neighbouring atoms only in most tight binding approaches and the overlap between atomic orbitals is often ignored (orthogonal tight binding). Classically electrons on the individual atoms would be unable to move between atoms in this way; however, quantum mechanically, these electrons tunnel out of their individual quantum wells.

Solving the Schrödinger equation for the many-atom system it is found that the energy levels are split with respect to the energy levels of the isolated atoms. The extent of the splitting depends on the magnitude of the hopping integral. Electrons inhabit the lowest energy or bonding states as far as the Pauli exclusion principle will allow.

The first term in the total energy of the system as given in eqn 2.22 is a sum of the occupied eigenvalues (E_n), with a factor of 2 for spin degeneracy. This term (known as the band energy) is attractive (otherwise bonding would not occur), whilst the second term is described by a sum of pair potentials and is repulsive. R_{IJ} is the distance between a pair of atoms I and J.

$$E_{total} = 2 \sum_{n(occ)} E_n + \frac{1}{2} \sum_{I \neq J} \phi_{IJ}(R_{IJ}) \tag{2.22}$$

At this stage it is necessary to include the effect of magnetism as the tight binding method is going to be used to model iron which is a magnetic material. This is a simple generalisation of the non-magnetic case in which the starting point is also the single particle Schrödinger

equation. As in the non-magnetic case the band energy part of the total energy is given by the sum of the eigenvalues of the occupied orbitals. However, in this case it is split into a sum over up-spins and sum over down-spins. This is illustrated in eqn 2.23.

$$E_{band} = \sum_{n\uparrow} E_{n\uparrow} + \sum_{n\downarrow} E_{n\downarrow} \tag{2.23}$$

The matrix elements of the spin-up or spin-down Hamiltonian can be expressed in terms of the non-magnetic Hamiltonian matrix elements

$$\langle J\beta\sigma|H_\sigma|I\alpha\sigma\rangle = \langle J\beta\sigma|H_0|I\alpha\sigma\rangle + \Delta E_{I\sigma}\delta_{IJ}\delta_{\alpha\beta} \tag{2.24}$$

where the index σ can indicate either spin-up or spin-down. The term $\Delta E_{I\sigma}\delta_{IJ}\delta_{\alpha\beta}$ represents the splitting in the energy on a single site due to an imbalance in the number of up and down spins. It is assumed that this splitting only affects on-site terms (hence the presence of the δ_{IJ} delta function). Hopping can only occur from spin-up to spin-up or spin-down to spin-down states.

The energy splitting on each site is expressed in eqn 2.25

$$\Delta E_I = \Delta E_{I\uparrow} - \Delta E_{I\downarrow} \tag{2.25}$$

where $\Delta E_{I\uparrow}$ and $\Delta E_{I\downarrow}$ are the changes in the spin-up and spin-down on-site energies relative to the non-spin polarised case. These shifts are given by eqns 2.26 and 2.27.

$$\Delta E_{I\uparrow} = -\frac{1}{2} I_s^I (n_{I\uparrow} - n_{I\downarrow}) \tag{2.26}$$

$$\Delta E_{I\downarrow} = -\Delta E_{I\uparrow} \tag{2.27}$$

I_s^I is a positive constant known as the Stoner integral on at atom I [87] and $q_{I\uparrow}$ and $q_{I\downarrow}$ are the numbers of electrons with spin up and down respectively.

The work included in Chapter 3 seeks to extend the existing non-self consistent Harris-Foulkes tight binding scheme in order to calculate both charge and spin self consistently in an efficient manner. The nature of the Stoner model is also described in more detail.

2.2.3 *Techniques Part 2: Empirical potentials and linear elasticity*

There are limits placed on the number of atoms in a metallic system that can be studied using DFT and tight binding, in the former case around 100 atoms and in the latter around 1000. An approach that bridges the gap between tight binding and empirical potentials is that of bond order potentials (BOP's). These attempt to derive potentials from the principles of tight binding [76, 66]. Empirical potentials can deal with millions of atoms and are thus of great use in considering extended defects.

Empirical Potentials

The need for interatomic potentials is motivated by a need for algorithms to calculate energies and forces on atoms in a quick and efficient way. It is then possible to minimise the energy and allow the forces to go to zero in order to determine the relaxed structures of defects in crystals. The functional forms of the potentials are based on approximations to tight binding and density functional theory. Atoms are treated as point-like and loosely speaking the potential is the "glue" that binds them together. To parametrize the model, it is necessary to fit to physical properties obtained experimentally such as elastic constants and vacancy formation energies. A simple but extremely effective interatomic potential is that of Finnis and Sinclair [31], which is used to model transition metals.

There are two parts to the form of the energy of a Finnis-Sinclair potential. The first term is a many body attraction term and the second term is a pair potential part which describes the nucleus-nucleus repulsion between atoms as described in eqn 2.28, where R_{IJ} is the distance between a pair of atoms I and J, $\eta(R_{IJ})$ is a pair term that can be related to tight binding hopping integrals and $\phi(R_{IJ})$ is the repulsive pair potential. The mathematical forms of the cohesive and many-body potentials are fitted by comparison with cohesive energy, equilibrium

volume and the elastic moduli for various transition metals. The square root in the many body term is important in correctly reproducing the bonding in solids. If the coordination number of an atom is Z it can be seen from eqn 2.28 that there will be a $\sqrt{Z}$ multiplying the energy from the many body term. If this is divided by Z, the energy per bond obtained is proportional to $1/\sqrt{Z}$. This means that atoms at the surface of a material with a lower coordination number will be more strongly bonded than atoms deep in the material further from the surface. In this way Finnis-Sinclair potentials take into account the local environment of an atom.

$$E = -A \sum_I \sqrt{\sum_{J \neq I} \eta(R_{IJ})} + \frac{1}{2} \sum_{I,J} \phi(R_{IJ}) \tag{2.28}$$

The form of the potential means that there is a simple algorithm to calculate the forces on atoms. Differentiating the expression for the energy in eqn 2.28 the final expressions for the x components of the attractive and repulsive parts of the total forces are

$$\begin{aligned} F_{rep,x} &= -\frac{\partial}{\partial x_M} \frac{1}{2} \sum_{IJ} \phi(R_{IJ}) \\ &= -\frac{1}{2} \sum_I \sum_J \frac{\phi'(R_{IJ})}{R_{IJ}} (x_I - x_J)(\delta_{IM} - \delta_{JM}) \\ &= -\frac{1}{2} \sum_{J \neq M} \frac{\phi'(R_{JM})}{R_{JM}} (x_M - x_J) - \frac{1}{2} \sum_{I \neq M} \frac{\phi'(R_{IM})}{R_{IM}} (x_M - x_I) \end{aligned}$$

$$= -\sum_{I \neq M} \frac{\phi'(R_{IM})}{R_{IM}} (x_M - x_I) \tag{2.29}$$

$$\begin{aligned} F_{att,x} &= -\frac{\partial}{\partial x_M} \left[-A \sum_I \sqrt{\sum_{J \neq I} \eta(R_{IJ})} \right] \\ &= \frac{A}{2} \sum_I \left(\sum_{L \neq I} \eta(R_{IL}) \right)^{\frac{-1}{2}} \sum_{J \neq I} \eta'(R_{IJ})(x_I - x_J)(\delta_{IM} - \delta_{JM}) R_{IJ}^{-1} \\ &= \frac{A}{2} \left(\sum_{L=M} \eta(R_{ML}) \right)^{\frac{-1}{2}} \sum_{J \neq M} \eta'(R_{MJ})(x_M - x_J) R_{MJ}^{-1} \\ &+ \frac{A}{2} \sum_{I=M} \left(\sum_{L \neq I} \eta(R_{IL}) \right)^{\frac{-1}{2}} \eta'(R_{MI})(x_M - x_I) R_{MI}^{-1} \end{aligned}$$

$$= \frac{A}{2} \sum_{J \neq M} \eta'(R_{JM}) \left(\frac{1}{\sqrt{\nu_M}} + \frac{1}{\sqrt{\nu_J}} \right) \frac{(x_M - x_J)}{r_{JM}} \tag{2.30}$$

where $\nu_M = \sum_{L \neq M} \eta(R_{ML})$. The attractive force on the central atom m depends not only on its environment as illustrated in the $\frac{1}{\sqrt{\nu_M}}$ term but also on the environment of the other atoms J in its cutoff radius through $\frac{1}{\sqrt{\nu_J}}$. In the model of Finnis and Sinclair the cutoff is chosen between second and third nearest neighbours.

Variations on this early potential came later with the Ackland [4] and Mendelev [64] potentials. The Mendelev potential for iron does predict a non-degenerate dislocation core in agreement with DFT calculations whilst the Ackland potential does not. The main problem with such potentials is that they underestimate the value of the Peierls barrier and do not recreate its characteristic single hump shape, which has been determined by DFT calculations [91].

The above mentioned works do not treat magnetism which is important for studies on iron to be used in fusion power plants. This is considered in recent tight binding models, such as those that are for pure iron [59] and modified versions which are for Fe-Cr systems [71], in which calculations are carried out using a spin polarised tight binding framework using the Stoner model. Magnetic bond order potentials are also employed in recent studies [67], to reporduce the structure of the easy core provided by DFT.

It is necessary to mention briefly the Dudarev-Derlet potential [21] as this is an interatomic potential which attempts to take into account magnetism by using a combination of a Stoner model (a simple band model for 3d transition metal, described in Chapter 3) and the Ginzburg-Landau approach to the study of second-order phase transitions [75]. A second order phase transition occurs when there is a discontinuity in the second derivative of the free energy of a system as happens in ferromagnetic iron. The Ginzburg Landau model correctly predicts the distribution of densities of atomic states, which is related to the magnetism of a material. Using a combination of these two models a many-body empirical potential is formed like the earlier Finnis-Sinclair potential. The embedding part of it is based on a combination of the

two models mentioned. A good agreement with DFT and tight binding for the magnetization curves of iron is obtained.

Elasticity theory

Whilst elasticity theory is at the bottom of the hierarchy of methods discussed above it can still provide much useful information with very little computational effort. For example it can give the displacements of atoms outside the core and general trends in energies with system size can be established. It can also manage more complicated situations such as the interaction between a dislocation and a point defect and the interaction between a pair of dislocations. The groundbreaking paper of Eshelby et al. [28] derives a method for calculating the fully anisotropic displacement solution for an extended body and then deals with a particular screw and edge dislocation to illustrate this formalism. More recent work considers defects called dislocation loops created by radiation damage using anisotropic elasticity theory. Work by Fitzgerald and Yao [33] considers the shape of these loops and compares them to images obtained using transmission electron microscopy and finds good agreement. The formation energy of these loops is also considered by Dudarev et al. [25] and found to be in good agreement with values obtained using the Dudarev-Derlet potential. Straight dislocation segments are also considered [32] and their self force using full anisotropic elasticity theory calculated.

In this work the isotropic elastic solution is used. This is described in eqn 2.31

$$u_{z'} = \frac{b\theta}{2\pi} \tag{2.31}$$

where $u_{z'}$ is the displacement in the [111] direction, b is the magnitude of the Burgers vector and θ is the angle between a general point and an arbitrarily defined line from the centre of the dislocation to a point in the $x'y'$plane.

The form of the displacement can be derived from the Navier Lamé equation (eqn 2.32), which is the basic equation of elasticity theory.

This is essentially the equation of motion for the components of a displacement $\vec{u}$ in a cartesian coordinate system. The $\sigma_{ij,j}$ term is the divergence of the stress tensor, F_i a body force and $\rho\frac{\partial^2 u_i}{\partial t^2}$ the time dependence of the displacement. The Einstein summation convention is used throughout this discussion.

$$\sigma_{ij,j} + F_i = \rho\frac{\partial^2 u_i}{\partial t^2} \tag{2.32}$$

The first index of the stress tensor σ_{ij} is the ith component of a force vector per unit area (where $i\epsilon\{1,2,3\}$) acting on a plane whose outward pointing normal is parallel to the positive x_jth direction. u_i is the component of the displacement vector in the ith direction, F_i is the component of the body force per unit volume in the ith direction and ρ is the mass density. In this problem time dependence can be ignored as equilibrium conditions are imposed.

It is necessary to couch the governing equation in terms of the displacement to be able to solve for it. To do this the concept of strain must be introduced. The strain tensor e_{kl} is a purely geometrical concept which describes changes in shape to a material upon deformation.

$$e_{kl} = \frac{1}{2}(u_{k,l} + u_{l,k}) \tag{2.33}$$

where u_k is the displacement in the kth direction. The relationship between stress and strain is linear and is expressed as

$$\sigma_{ij} = c_{ijkl}e_{kl} \tag{2.34}$$

c_{ijkl} is known as the stiffness tensor. It represents the elastic constants and has the following symmetries:

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{jilk} \tag{2.35}$$

In the isotropic case under consideration the direction in which a force is applied makes no difference to the resulting displacements. This enables certain simplifications to be made to the form of the stiffness tensor. The relationship between stress and strain can then be written as:

$$\sigma_{ij} = \lambda\delta_{ij}e_{kk} + 2\mu e_{ij} \tag{2.36}$$

where λ and μ are known as Lamé parameters. Inserting the expression for the strain from eqn 2.33 into eqn 2.36 provides an expression for the stress tensor in terms of displacement.

$$\sigma_{ij} = \lambda\delta_{ij}u_{k,k} + \mu(u_{i,j} + u_{j,i}) \tag{2.37}$$

Differentiating eqn 2.37 and inserting into eqn 2.32 provides a soluble equation for the displacement.

$$(\lambda + \mu)u_{j,ij+}\mu u_{i,jj} + F_i = 0 \tag{2.38}$$

The general framework for obtaining the displacement field of an isotropic body has now been fully developed. Applying this to the case of a screw dislocation is fairly simple as the displacement depends only on the angle θ. There are also no body forces to consider so F_i can be ignored. Eqn 2.38 now reduces to a second order ordinary differential equation for the displacement $u_z = u(\theta)$. It is assumed there is no displacement in the $x-y$ plane, so only u_z needs to be considered.

$$\frac{d^2u_z}{d\theta^2} = 0 \tag{2.39}$$

The solution is of the form:

$$u_z = A\theta + B \tag{2.40}$$

Applying the two conditions that at $\theta = 0$, $u_z = 0$ and at $\theta = 2\pi$, $u_z = b$ (where b is the magnitude of the Burgers vector), the final expression for the displacement field of an isotropic screw dislocation is then $\frac{b\theta}{2\pi}$.

3

EFFICIENT SELF CONSISTENCY FOR MAGNETIC TIGHT BINDING

3.1 THE MODEL

The parameters employed to investigate the defect properties of ferromagnetic bcc iron are obtained from a d-band Stoner tight binding model first published in Liu et al. [59]. Whilst the original model assumed local charge neutrality (LCN), this work seeks to extend this by adapting it to include the process of charge transfer. The expectation is that this would provide a more realistic picture of the behaviour of defects in iron. The constituents of the energy in the original scheme are the band energy, the repulsive contribution and the Stoner exchange energy [86].

The d band in iron is relatively narrow in comparison to the s-p band. A narrower band implies that the interaction between atoms is less strong and that the orbitals on atoms are highly localised making it a prime candidate for a tight binding description. In this case the s and p electrons are included in the pair potential as they are strongly delocalised and their effect on the structure is thought to be small. There are other models which include s, p and d orbitals. For example, the model of Paxton and Finnis [74], also considers orbital overlap and multipolar charges on each site. Spin is dealt with similarly to the model used in this work by a simple Stoner description. More complicated magnetic descriptions can be considered. Cohen and Mukherjee [18] included non-collinear magnetism in iron using tight binding.

The basis for the Stoner band model of magnetism reduces to a competition between the kinetic energy lost and exchange energy gained when a certain number of electrons of one spin swap their spin spontaneously to the opposite orientation. This is best understood from

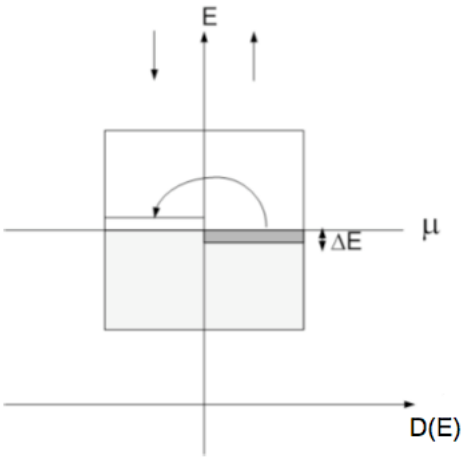


Figure 5: Explaining Stoner model of magnetism - a simplistic picture. The right hand section of the figure indicates spin up density of states, and the left hand section the spin down. μ labels the chemical potential.

Figure 5 where the energy change ($\triangle W$) from the change in spin of the electrons in the shaded area can be written as:

$$\begin{aligned} \triangle W &= kinetic + exchange \\ \Delta W &= D(\mu)\Delta E\Delta E - \frac{1}{4} I_s (2D(\mu)\triangle E\}^2 \end{aligned} \tag{3.1}$$

where $D(\mu)$ labels the density of states of the electrons at the chemical potential μ, $\triangle E$ is the range of energies of the electrons which change spin and I_s is the Stoner parameter which measures the strength of the exchange interaction . The sign of $\triangle W$ dictates whether a system is magnetic or not.

$$\triangle W < 0 \Rightarrow magnetism \tag{3.2}$$

$$\triangle W > 0 \Rightarrow non - magnetic \tag{3.3}$$

Setting eqn 3.1 equal to zero, and canceling through by the common factors $\triangle E$ and $D(E)$, in conjunction with the conditions specified in the inequalities 3.2 and 3.3 leads to the following condition for ferromagnetism behaviour

$$I_s \, D\,(\mu) > 1 \tag{3.4}$$

The total energy of the system is written as an extended form of the Harris-Foulkes functional; extended to include spin and charge.

$$
\begin{aligned}
E_{tot} &= \frac{1}{N_k}\sum_{\sigma n\vec{k}}\left\langle \sigma n\vec{k}\middle|\hat{H}^0\middle|\sigma n\vec{k}\right\rangle f_{\sigma n\vec{k}} + E_{pair} \\
&+ \sum_{IJ}(n_I - z_I)V_{IJ}q_J - \frac{1}{2}\sum_{IJ} q_I V_{IJ} q_J \\
&- \frac{1}{2}\sum_{I} I_s^I(n_{I\uparrow} - n_{I\downarrow})m_I + \frac{1}{4}\sum_{I} I_s^I m_I^2
\end{aligned}
\tag{3.5}
$$

$\widehat{H}^0$ is the reference Hamiltonian which does not include spin or charge, σ labels the spin, n is the band index, $\vec{k}$ is a Bloch vector, N_k is the number of k-points used in the Brillouin zone sum and $\vec{R}$ is the translation vector for a periodically repeated computational cell. Certain simplifications with regard to the hopping integrals [40, 87] are made at this stage. Only on-site and two centre hopping integrals are considered, crystal field terms and three centre integrals are neglected. This was considered to be sufficient for the task required.

The Bloch states are represented as a linear combination of atomic orbitals (LCAO) [40] as described in eqn 3.6.

$$
|\,\sigma n\vec{k}\rangle = \frac{1}{\sqrt{N_k}}\sum_{i\vec{R}} c_i^{\sigma n\vec{k}} \exp(i\vec{k}\cdot\vec{R})\,|\,\sigma i\vec{R}\rangle \tag{3.6}
$$

The atomic orbital $|\,\sigma i\vec{R}\rangle$ for an electron of spin σ in cell with position $\vec{R}$ is indexed by i, with the orbital being centered on atomic site I. The expansion coefficients $c_i^{\sigma n\vec{k}}$ are normalised such that $\left\langle \sigma n\vec{k}\,|\,\sigma n\vec{k}\right\rangle = 1$, and $f_{\sigma n\vec{k}}$ is the occupation number for state $|\,\sigma n\vec{k}\rangle$. E_{pair} is a repulsive pairwise potential. The various charges and magnetic moments (spins) on site I are: z_I is the core charge, q_I is the input charge, m_I is the input magnetic moment, and $n_I = n_{I\uparrow} + n_{I\downarrow}$ is the total number of output electrons, where $n_{I\uparrow}$ and $n_{I\downarrow}$ are the output number of up and down spin electrons respectively. At self-consistency, when the input and output spin and charge densities are identical, $q_I = n_I - Z_I$ and $m_I = n_I^{\uparrow} - n_I^{\downarrow}$.

The output number of electrons n_I is obtained from the wavefunctions.

$$
n_I = \frac{1}{N_k}\sum_{i\epsilon I}\sum_{n\vec{k}\sigma} |\,c_i^{\sigma n\vec{k}}\,|^2\, f_{\sigma n\vec{k}} \tag{3.7}
$$

I_S^I is the Stoner integral [86] for the atom on site I, and V_{IJ} is the potential between the charges on sites I and J.

The extension implemented in this present work to the d band stoner model of Liu et al. [59] is to include charge explicitly instead of assuming LCN. We assume monopolar charges on each atomic site and describe these charges using Gaussian distributions. At long range this reduces the inter-site interactions ($I \neq J$) to Coulomb's law whilst at short range the width of the Gaussian ensures that the on-site energy ($I = J$) is equal to the value of the Hubbard U [51] for a particular atomic site.

$$E_{charge} = \frac{1}{2}\sum_{IJ} q_I V_{IJ} q_J$$
$$V_{II} = U \tag{3.8}$$

$$V_{IJ} = \sum_{\vec{R}} \gamma_{IJ}(\left|\vec{r}_I - \vec{r}_J - \vec{R}\right|) \tag{3.9}$$

γ_{IJ} is defined below (eqn 3.10). It is assumed that there are spherical Gaussian shaped charge distributions $\rho_I(\vec{r})$ on each atomic site allowing the program PLATO to compute the inter-site terms [47, 55]. This charge distribution is given by $\rho_I(\vec{r}) = q_I N_I \exp(-\alpha_I(\vec{r} - \vec{r}_I)^2)$, and is normalised such that $N_I = (\alpha_I/\pi)^{3/2}$ guaranteeing the integral of ρ_I is q_I. $\vec{r}_I$ is the position of site I and α_I is a constant selected to make certain the on-site terms are correct i.e. $V_{II} = U$. For periodic systems, V_{IJ} includes a sum over periodic images as described in eqn 3.9 where $\vec{R}$ is the translation vector.

The interaction between a pair of atoms ($\gamma_{IJ}(r)$) can be calculated using a Gaussian integral [13], resulting in eqn 3.10,

$$\gamma_{IJ}(r) = \frac{e^2}{4\pi\epsilon_0 r}\mathrm{erf}\left(\sqrt{\frac{\alpha_I\alpha_J}{\alpha_I+\alpha_J}}r\right) \tag{3.10}$$

where e is the charge on an electron and ϵ_0 is the permittivity of free space. For the on-site terms it is known that $\gamma_{II}(0) = U_I$ and hence $\alpha_I = \frac{\pi}{2}(4\pi\epsilon_0 U_I/e^2)^2$.

There are five types of d orbital on each iron atom which form sigma, pi and and delta bonds. The distance dependence of these bond

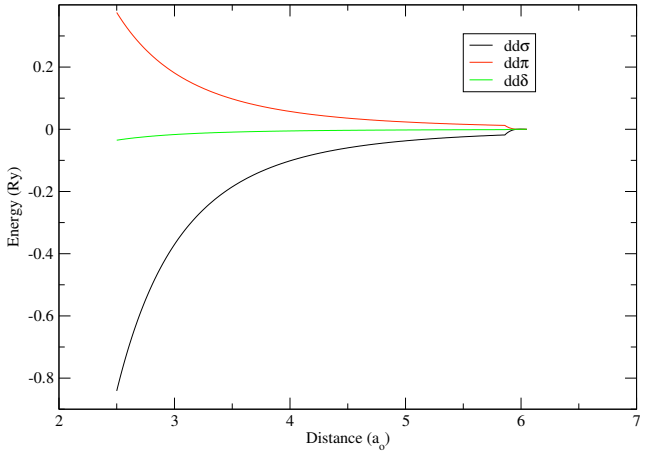


Figure 6: The three bond integrals

integrals are all described by a power law relation, with different values of the exponent.

$$t(r) = t(r_0)\left(\frac{r}{r_0}\right)^{-s}$$

It was found necessary to add a cubic spline to these functions in order for them to decay smoothly to zero at a suitable distance r. Plots of the bond integrals are provided in Figure 6.

The repulsive energy E_{pair} in eqn 3.11 is comprised of both a long range exponential term and short range Yukawa term which tails off much faster than the long range term due to the $\frac{B}{r_{IJ}^m}$ factor. The relationship is plotted in Figure 7. A,B,a,b and m are all fitted parameters provided by the original model.

$$E_{pair} = \frac{1}{2}\left(\frac{B}{r_{IJ}^m}\exp(-ar_{IJ}) + A\exp(-br_{IJ})\right) \tag{3.11}$$

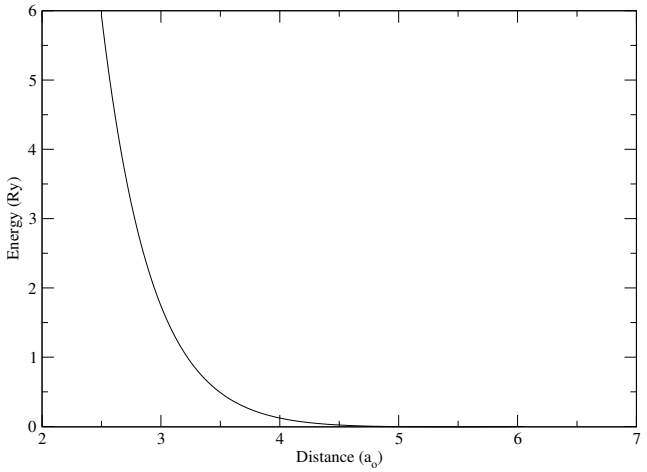


Figure 7: Repulsive Energy

3.2 THE PROBLEM

A set of initial simulations for bulk iron encountered two unexpected difficulties. The first relates to the slow speed of convergence. The speed of convergence is itself dependent on two factors: the first is the number of computational loops required and the second is the time taken for each loop. The second and more serious problem is the possibility that on some occasions convergence can never be reached. To begin with simulations were conducted using an extremely simplistic linear mixing scheme indicated in eqn 3.12. The input and output refer to either spin or charge and M_f is the mixfactor .

$$input_{new} = input_{old} + M_f(output_{old} - input_{old}) \tag{3.12}$$

Conducting tests on a 3 x 3 x 3 cell of bulk bcc ferromagnetic iron it is found that convergence is not attained and that the simulation becomes numerically unstable. This is illustrated in Figure 8. The bulk is uncharged by symmetry so only spin convergence is considered in this case. The residue plotted is the difference in input and output spin which ideally we want to go to zero. It can be seen that for approximately the first 30 self consistent iterations the residue remains extremely low and decreases slowly. This situation then abruptly changes and the residue increases rapidly before oscillating indefinitely.

It can be supposed there will be a certain amount of statistical noise and that once the simulation is 'nudged' slightly from its correct course there is no mechanism for bringing the convergence back on track.

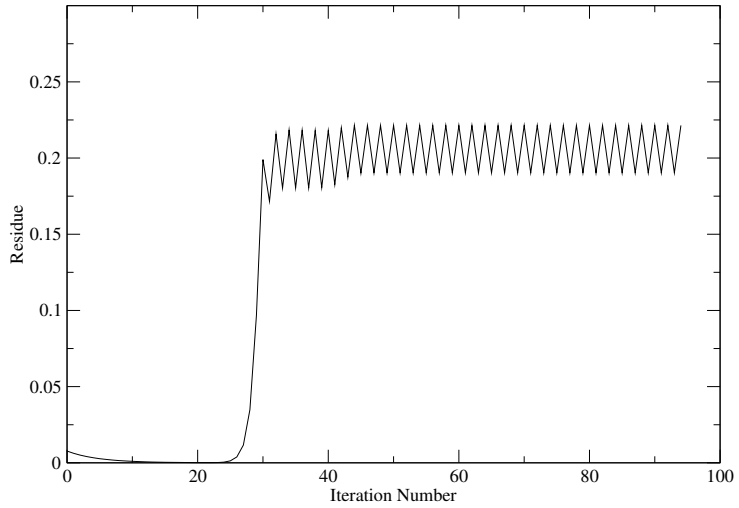


Figure 8: Linear mixing

Other possible alternatives involved using a Pulay type mixing scheme [77] for charge and spin. Whilst this is not plagued by the instabilities of linear mixing, it is quite slow as it requires a large number of loops for convergence. Improvements were made to the original Pulay scheme, in that it was found it was better to mix charge and spin separately. That is, to keep updating the charge until completely converged (with a tolerance of 10^{-7}e) and only then to allow charge and spin to be mixed together. If the convergence on charge goes below the tolerance the simulation reverts back to mixing charge only.

This is still not ideal, particularly not with larger numbers of atoms hence motivating the development of the Newton mixing scheme described in the next section.

3.3 THE SOLUTION

3.3.1 *Charge/Spin self consistency using the extended Harris-Foulkes functional*

The novelty in this scheme arises from the need to achieve self consistency for both spin and charge. Methods for implementing this in density functional theory (DFT) have received much attention. However, the problem has been largely neglected in tight binding. This novel approach is motivated by the difficulties encountered in using existing DFT methods (Pulay and linear mixing) for self consistency in tight binding. The result is a robust and efficient scheme anchored on firm mathematical algorithms.

The original form of the Harris-Foulkes functional is obtained (as explained in Chapter 2) by expanding to second order the exact DFT energy functional about a certain input charge density. If the second order terms in this expansion are then ignored, the magnitude of the error in the energy is second order in the error in the density. The energy can then be calculated non self-consistently using a suitable input charge density. The Stoner magnetic tight binding model used here, extended to include charge transfer between atoms, provides an opportunity to formulate the method of achieving self consistency as a variational problem. The two terms which describe the electrostatic and magnetic contributions along with their double counting terms are added on to the original form of the Harris-Foulkes functional. These terms include both input and output spin and charge. Upon minimisation of this extended functional it is found to be stationary with respect to changes in the wavefunction given that Schrödinger's equation is satisfied. It is also found to be stationary with respect to input and output spins and charges when they are equal. That is, the problem of reaching self consistency has now been reduced to an energy minimisation problem. In the following sections the implementation of

this minimization via a Newton-Raphson [78] type scheme is discussed in detail.

Upon minimising the energy functional in eqn 3.5 with respect to the expansion coefficients, the Schrödinger equation for the system results. The solution of this equation allows the output spins and charges to be calculated from the eigenstates. The values of output spins and charges and in turn the energy are now known unequivocally, given fixed values of input spins and charges. That is, the energy is now just a function of the input charges and spins.

$$\hat{H}^{\sigma} \mid \sigma n k\rangle = \epsilon^{\sigma}_{nk} \mid \sigma n k\rangle \tag{3.13}$$

where ϵ^{σ}_{nk} is the eigenvalue and the spin dependent Hamiltonian $\hat{H}^{\sigma}$ is given in terms of its matrix elements $H^{\sigma}_{i\vec{R},j\vec{R}'} = \langle \sigma i\vec{R} \mid \hat{H}^{\sigma} \mid \sigma j\vec{R}'\rangle$ by eqn 3.14

$$H^{\sigma}_{i\vec{R},j\vec{R}'} = H^{0}_{i\vec{R},j\vec{R}'} + \delta_{ij}\delta_{\vec{R}\vec{R}'}\left[U_I q_I + \sum_{M(\neq I)} V_{IM} q_M - \frac{1}{2} I^I_s m_I (1 - 2\delta_{\sigma\uparrow})\right] \tag{3.14}$$

where $H^{o}_{i\vec{R},j\vec{R}'} = \langle \sigma i\vec{R} \mid \hat{H}^{0} \mid \sigma j\vec{R}'\rangle$. If $\sigma = \uparrow$ then $\delta_{\sigma\uparrow} = 1$; if $\sigma = \downarrow$ then 0.

The energy function that uses the optimised expansion coefficients is labelled $E_M(q_I, m_I)$. Input charge and spin can now be treated as variational parameters with respect to which the minimised functional must be optimised. Derivatives with respect to these parameters are calculated by differentiating eqn 3.5, which results in eqns 3.15 and 3.16. As the expansion coefficients have already been optimised their derivatives disappear.

$$\partial E_M / \partial q_I = \sum_J (n_J - z_J) V_{IJ} - \sum_J V_{IJ} q_J \tag{3.15}$$

$$\partial E_M / \partial m_I = -\frac{1}{2} I^I_s (n_{I\uparrow} - n_{I\downarrow}) + \frac{1}{2} I^I_s m_I \tag{3.16}$$

Setting the derivatives in eqns 3.15 and 3.16 to zero and canceling through by common factors results in:

$$\begin{aligned} q_I &= n_I - z_I \\ m_I &= n^{\uparrow}_I - n^{\downarrow}_I \end{aligned}$$

It is now evident that finding a stationary point E_M is equivalent to achieving self consistency. This is an extremely useful result.

After the charge and spins are fully converged and the energy calculated it is necessary to calculate a value of the atomic forces. These are given by Hellman-Feyman forces [30].

$$\vec{F_K} = \sum_{ij\vec{R}\vec{R}'} \rho_{i\vec{R},j\vec{R}'} \frac{\partial H^0_{i\vec{R},j\vec{R}'}}{\partial \vec{r}_K} + \frac{\partial E_{pair}}{\partial \vec{r}_K} + \frac{1}{2} \sum_{IJ} q_I q_J \frac{\partial V_{IJ}}{\partial \vec{r}_K} \tag{3.17}$$

Here $\rho_{i\vec{R},j\vec{R}'} = \frac{1}{N_k} \sum_{\sigma n \vec{k}} f_{\sigma n \vec{k}} c_i^{\sigma n \vec{k}*} c_j^{\sigma n \vec{k}} \exp(i\vec{k} \cdot [\vec{R}' - \vec{R}])$ are the on-site elements of the density matrix. Forces are minimised to a prescribed tolerance to optimise the atomic coordinates.

In the context of energy minimisation, the relationship between this scheme and the local charge neutrality of the original model can now be understood. If the value of the Hubbard U is increased, then in order for the energy to be minimised the charges must become very small, in effect recreating local charge neutrality.

3.3.2 *Newton-Raphson Mixing*

To utilize this variational property the usual techniques for locating the minima and maxima of functions are employed. First the minimised energy is expanded as a Taylor expansion about some initial values of charge and spin (q_I^0 and m_I^0) up to and including quadratic terms as described in eqn 3.18. Ideally it would be better to expand to higher order than this. However, the constraint is the excessive computational expense.

$$\begin{aligned} &E_M(q_I^0 + \Delta q_I, m_I^0 + \Delta m_I) \approx E_M(q_I^0, m_I^0) \\ &+ \sum_I \left(\left.\frac{\partial E_M}{\partial q_I}\right|_0 \Delta q_I + \left.\frac{\partial E_M}{\partial m_I}\right|_0 \Delta m_I \right) \\ &+ \frac{1}{2} \sum_{IJ} \left(\left.\frac{\partial^2 E_M}{\partial q_I \partial q_J}\right|_0 \Delta q_I \Delta q_J + \left.\frac{\partial^2 E_M}{\partial m_I \partial m_J}\right|_0 \Delta m_I \Delta m_J + 2 \left.\frac{\partial^2 E_M}{\partial q_I \partial m_J}\right|_0 \Delta q_I \Delta m_J \right) \end{aligned} \tag{3.18}$$

Minimising this expanded functional with respect to input charge and spin produces a linear relationship between the difference in the input and output spins and charges (spin and charge residues) and the input spins and charges. These are related by a linear response matrix, the calculation of which is discussed in the next section. It is the calculation of this which takes the majority of the computational effort. This is achieved using Green's functions. In order to reach self consistency values of input charges and spins $q_I = q_I^0 + \Delta q_I$ and $m_I = m_I^0 + \Delta m_I$ must be found that satisfy the conditions $\partial E_M/\partial q_I = \partial E_M/\partial m_I = 0$. Setting the derivatives of eqn 3.18 to zero results in eqn 3.19.

$$\begin{aligned} \left.\frac{\partial E_M}{\partial q_I}\right|_0 + \sum_J \left.\frac{\partial^2 E_M}{\partial q_I \partial q_J}\right|_0 \Delta q_J + \sum_J \left.\frac{\partial^2 E_M}{\partial q_I \partial m_J}\right|_0 \Delta m_J &= 0 \\ \left.\frac{\partial E_M}{\partial m_I}\right|_0 + \sum_J \left.\frac{\partial^2 E_M}{\partial m_I \partial m_J}\right|_0 \Delta m_J + \sum_J \left.\frac{\partial^2 E_M}{\partial m_I \partial q_J}\right|_0 \Delta q_J &= 0 \end{aligned} \tag{3.19}$$

At this stage it makes sense to simplify the notation and adopt a combined notation for both spin and charge. New quantities are defined as: $\lambda_I \in \{m_I, q_I\}$ and $\Delta\lambda_I \in \{\Delta m_I, \Delta q_I\}$. The first ($g_I^\lambda$) and second ($X_{IJ}^{\lambda\lambda'}$) derivatives of E_M are now defined. We call $X_{IJ}^{\lambda\lambda'}$ the linear response matrix; it has dimensions $2N \times 2N$, and as will be seen below, it can be computed straightforwardly. From this notation the form of a Newton-Raphson type equation immediately becomes apparent (eqn 3.21).

$$g_I^\lambda = \left.\frac{\partial E_M}{\partial \lambda_I}\right|_0 \qquad X_{IJ}^{\lambda\lambda'} = \left.\frac{\partial^2 E_M}{\partial \lambda_I \partial \lambda_J'}\right|_0 \tag{3.20}$$

$$g_I^\lambda + \sum_{\lambda' J} X_{IJ}^{\lambda\lambda'} \Delta\lambda_J' = 0 \tag{3.21}$$

This linear matrix equation can easily be solved. Differentiating eqn 3.5, eqns 3.22 and 3.23 result.

$$\frac{\partial E_M}{\partial q_I} = \sum_J V_{IJ}(n_J - z_J - q_J) = \sum_J V_{IJ} R_J^q \tag{3.22}$$

$$\frac{\partial E_M}{\partial m_I} = -\frac{1}{2} I_s^I (n_I^\uparrow - n_I^\downarrow - m_I) = -\frac{1}{2} I_s^I R_I^m \tag{3.23}$$

where $R_J^q = (n_J - z_J - q_J)$ is defined as the charge residue and $R_J^m = (n_J^{\uparrow} - n_J^{\downarrow} - m_J)$ the spin residue. To obtain the second derivatives, eqns 3.22 and 3.23 are differentiated with respect to both charge and magnetic moment to give

$$\frac{\partial^2 E_M}{\partial q_I \partial q_K} = \sum_J V_{KJ} \frac{\partial R_J^q}{\partial q_I} \tag{3.24}$$

$$\frac{\partial^2 E_M}{\partial q_I \partial m_K} = -\frac{1}{2} I_s^K \frac{\partial R_K^m}{\partial q_I} = \sum_J V_{IJ} \frac{\partial R_J^q}{\partial m_K} \tag{3.25}$$

$$\frac{\partial^2 E_M}{\partial m_I \partial m_K} = -\frac{1}{2} I_s^K \frac{\partial R_K^m}{\partial m_I} \tag{3.26}$$

Substituting eqns 3.22, 3.23, 3.24, 3.25 and 3.26 into Eq. 3.21 and factoring out V_{IJ} results in the final expressions:

$$0 = R_J^q + \sum_K \left(\frac{\partial R_J^q}{\partial q_K} \Delta q_K + \frac{\partial R_J^q}{\partial m_K} \Delta m_K \right) \tag{3.27}$$

$$0 = R_J^m + \sum_K \left(\frac{\partial R_J^m}{\partial q_K} \Delta q_K + \frac{\partial R_J^m}{\partial m_K} \Delta m_K \right) \tag{3.28}$$

These equations are easily interpreted. Together they constitute a linear extrapolation to find where the residues go to zero, i.e when self-consistency is achieved. To illustrate this point, it is instructive to consider a 'one-dimensional' situation in which the atoms in a configuration are charged but non magnetic. The problem then reduces to the simplest case of Newton-Raphson minimisation and the charge residue will reach zero swiftly in a very few computational steps. When both charge and spin are considered, the simple Newton-Raphson numerical method can be generalised. So instead of minimising a curve which is a function of one variable only there is a surface which is a function of two variables to which a plane will touch tangentially. If the energy surface were exactly quadratic in both charge and spin, self consistency would be reached in one step. However, the tight binding energy surface is naturally rough, making it necessary to carry out several iterations before self-consistency is reached. To construct an overall measure of the total residue at each step it is convenient to

define a quantity that is called $\bar{R}$. For each atom the charge and spin residues are calculated at each step and N is the number of atoms. The length of the vector formed of these quantities is then calculated.

$$\bar{R} = \left\{ \frac{1}{N} \sum_{J} \left(\left(R_J^q \right)^2 + \left(R_J^m \right)^2 \right) \right\}^{0.5} \tag{3.29}$$

3.4 CALCULATING THE LINEAR RESPONSE MATRIX

The Newton-Raphson method appears to be a transparent and straightforward method for reaching spin and charge self consistency. However, there is the difficult problem of calculating the linear response matrix $X_{IJ}^{\lambda\lambda'}$ to tackle. As a minimal basis set is used, it is expected that simulations containing a larger number of atoms than in similar DFT calculations can be completed. However a bottleneck in the system arises from the dependence of the linear response matrix on the number of atoms. Following through with the algebra it can be seen that the dependence is actually on the fourth power of the number of atoms. There are four derivatives that must be calculated to obtain the full form on the linear response matrix as both λ and λ' represent both magnetic moment and charge. These are detailed in eqn 3.30.

$$\begin{aligned} \frac{\partial R_J^q}{\partial q_I} &= \frac{\partial n_J^{\uparrow}}{\partial q_I} + \frac{\partial n_J^{\downarrow}}{\partial q_I} - 1 \qquad \frac{\partial R_J^q}{\partial m_I} = \frac{\partial n_J^{\uparrow}}{\partial m_I} + \frac{\partial n_J^{\downarrow}}{\partial m_I} \\ \frac{\partial R_J^m}{\partial q_I} &= \frac{\partial n_J^{\uparrow}}{\partial q_I} - \frac{\partial n_J^{\downarrow}}{\partial q_I} \qquad \frac{\partial R_J^m}{\partial m_I} = \frac{\partial n_J^{\uparrow}}{\partial m_I} - \frac{\partial n_J^{\downarrow}}{\partial m_I} - 1 \end{aligned} \tag{3.30}$$

It it apparent that the job of calculating the linear response matrix reduces to finding the derivative $\partial n_I^{\sigma}/\partial \lambda_J$, where $\lambda_J \in \{q_J, m_J\}$. It can be demonstrated that the non-trivial aspects of the process of calculating this derivative can be reduced to the calculation of one particular derivative. This is defined in eqn 3.31.

$$K_{IJ}^{\sigma} = \partial n_I^{\sigma}/\partial \epsilon_J^{\sigma} \tag{3.31}$$

where ϵ_J^σ labels the diagonal matrix elements of the Hamiltonian $H^\sigma_{i\vec{R},i\vec{R}}$. The most straightforward method of doing this is through the use of Green's functions.

The number of output electrons depends only on the chemical potential and the Hamiltonian. The starting assumption is that electron levels are thermally populated according to Fermi-Dirac statistics. So the occupation number is given by: $f_{\sigma n\vec{k}} = f(\epsilon^\sigma_{nk} - \mu)$ and $f(\epsilon - \mu) = 1/[1 + \exp((\epsilon - \mu)/k_B T)]$ where k_B is Boltzmann's constant, T labels the temperature, μ the electron chemical potential.

$$\frac{\partial n_I^\sigma}{\partial \lambda_J} = \frac{\partial n_I^\sigma}{\partial \mu}\frac{\partial \mu}{\partial \lambda_J} + \sum_{\vec{R}\vec{R}'ij} \frac{\partial n_I^\sigma}{\partial H^\sigma_{i\vec{R},j\vec{R}'}} \frac{\partial H^\sigma_{i\vec{R},j\vec{R}'}}{\partial \lambda_J} \tag{3.32}$$

where $\lambda_J \in \{q_J, m_J\}$.

The dependence of the Hamiltonian on charge and spin is completely contained in its on-site elements $\epsilon_I^\sigma = H^\sigma_{i\vec{R},i\vec{R}}$ (see eqn 3.14) . As the parameters from the tight binding model used are for d band orbitals only, there is only one on-site energy per atom. (Whilst the details here are only given for d-orbitals, this scheme is equally applicable to models which include s, p and d orbitals.) However the algebra is not as simple. Therefore consider the derivative of the Hamiltonian with respect to charge and spin in eqn 3.32, the only derivative required is given in eqn 3.33.

$$W_{IJ}^{\sigma\lambda} = \partial \epsilon_I^\sigma / \partial \lambda_J \tag{3.33}$$

In the same way finding the derivative of the number of electrons n_I^σ with respect to the Hamiltonian reduces to $\partial n_I^\sigma / \partial \epsilon_J^\sigma$. Differentiating the expression for the Hamiltonian given in eqns 3.14, the derivatives in eqns 3.34 and 3.35 result.

$$W_{IJ}^{\sigma m} = -\frac{I_s^I}{2}(1 - 2\delta_{\sigma\uparrow}) \tag{3.34}$$

$$W_{IJ}^{\sigma q} = U_I \delta_{IJ} + V_{IJ}(1 - \delta_{IJ}) \tag{3.35}$$

There are now only the terms in eqn 3.32 containing the chemical potential to consider. It will now be demonstrated that these quantities

can be completely described in terms of K^{σ}_{IJ} and $W^{\sigma\lambda}_{IJ}$. The total number of electrons n^{σ}_{I} can be expressed using the expression for the occupation number in eqn 3.36.

$$n^{\sigma}_{I} = \sum_{i \in I} \langle i\sigma\vec{R}|f(\hat{H} - \mu)|i\sigma\vec{R}\rangle \tag{3.36}$$

Considering the derivative of this with respect to the chemical potential it becomes apparent how this depends on the crucial derivative K^{σ}_{IJ}.

$$\begin{aligned}
\frac{\partial n^{\sigma}_{I}}{\partial \mu} &= \sum_{j l \vec{R}'} \frac{\partial n^{\sigma}_{I}}{\partial (H^{\sigma}_{j\vec{R}l\vec{R}'} - \mu\delta_{j\vec{R}l\vec{R}'})} \frac{\partial (H_{j\vec{R}l\vec{R}'} - \mu\delta_{j\vec{R}l\vec{R}'})}{\partial \mu} \\
&= -\sum_{j} \frac{\partial n^{\sigma}_{I}}{\partial (H^{\sigma}_{j\vec{R}j\vec{R}} - \mu)} \\
&= -\sum_{J} \frac{\partial n^{\sigma}_{I}}{\partial \epsilon^{\sigma}_{J}} \\
&= -\sum_{J} K^{\sigma}_{IJ}
\end{aligned} \tag{3.37}$$

where the sums over j and l are independent of each other. The last remaining term to evaluate in eqn 3.32 is $\partial\mu/\partial\lambda_J$, which can be found from the obvious condition that the total number of electrons (N_e) does not change.

$$\frac{\partial N_e}{\partial \lambda_L} = \sum_{I\sigma} \frac{\partial n^{\sigma}_{I}}{\partial \lambda_L} = 0 \tag{3.38}$$

Using the total derivative to expand the expression for $\frac{\partial n^{\sigma}_{I}}{\partial \lambda_L}$ results in eqn 3.39

$$\begin{aligned}
\sum_{I\sigma} \frac{\partial n^{\sigma}_{I}}{\partial \lambda_L} &= \sum_{I\sigma} \left\{ \left(\frac{\partial n^{\sigma}_{I}}{\partial \lambda_L} \right)_{\mu} + \frac{\partial n^{\sigma}_{I}}{\partial \mu} \frac{\partial \mu}{\partial \lambda_L} \right\} \\
&= \sum_{I\sigma} \left\{ \sum_{J} \left(\frac{\partial n^{\sigma}_{I}}{\partial \epsilon^{\sigma}_{J}} \right)_{\mu} \frac{\partial \epsilon^{\sigma}_{J}}{\partial \lambda_L} + \frac{\partial n^{\sigma}_{I}}{\partial \mu} \frac{\partial \mu}{\partial \lambda_L} \right\} \\
&= \sum_{JI\sigma} K^{\sigma}_{IJ} W^{\sigma\lambda}_{JL} - \frac{\partial \mu}{\partial \lambda_L} \sum_{JI\sigma} K^{\sigma}_{IJ}
\end{aligned} \tag{3.39}$$

using previous definitions of K^{σ}_{IJ} and $W^{\sigma\lambda}_{JL}$. Applying the condition in eqn 3.38 and performing some rearrangement the form of the required derivative results.

$$\frac{\partial \mu}{\partial \lambda_L} = \frac{\sum_{IJ\sigma} K^{\sigma}_{IJ} W^{\sigma\lambda}_{JL}}{\sum_{IJ\sigma} K^{\sigma}_{IJ}} \tag{3.40}$$

Having systematically demonstrated how the four terms in eqn 3.14 depend on K_{IJ}^{σ}, all that remains is to find an exact expression for it. The primary step is to re-write eqn 3.36 for the number of output electrons using Green's functions. Expressing the occupation number as $f(\epsilon_{nk}^{\sigma} - \mu) = \int f(E-\mu)\delta(E-\epsilon_{nk}^{\sigma})\,dE$ and inserting a complete set of the Hamiltonian eigenstates , eqn 3.36 now gives:

$$n_I^{\sigma} = \int f(E-\mu)\frac{1}{N_k}\sum_{n\vec{k},i\epsilon I}\left|c_i^{\sigma n\vec{k}}\right|^2\delta(E-\epsilon_{nk}^{\sigma})\,dE \tag{3.41}$$

A more useful form of the delta function is given by:

$$\delta(E-\epsilon_{nk}^{\sigma}) = -\frac{1}{\pi}\mathrm{Im}\left\{\lim_{\eta\to 0}\frac{1}{E+i\eta-\epsilon_{nk}^{\sigma}}\right\} \tag{3.42}$$

and the Green's function in this case is written as eqn 3.43

$$G_{\sigma i\vec{R}j\vec{R}'}(E+i0^{+}) = \lim_{\eta\to 0}\sum_{n\vec{k}}\frac{1}{N_k}\frac{c_i^{\sigma n\vec{k}}c_j^{\sigma n\vec{k}*}\exp\left(i\vec{k}\cdot(\vec{R}-\vec{R}')\right)}{E+i\eta-\epsilon_{nk}^{\sigma}} \tag{3.43}$$

Inserting the expressions from eqns 3.42 and 3.43 into 3.41 and rearranging results in the eqn 3.44 for the number of output electrons.

$$n_I^{\sigma} = \sum_{i\in I} -\frac{1}{\pi}\mathrm{Im}\int G_{\sigma i\vec{R}i\vec{R}}(E+i0^{+})f(E-\mu)\,dE \tag{3.44}$$

It is necessary to find the derivative of the Green's function with respect to the on-site Hamiltonian terms in order to reach an expression for K_{IJ}^{σ}. The first order change in the Green's function is sufficient to calculate this. From the definition of a Green's function G_0 it is known that from an arbitrary complex energy Z.

$$(Z-H_0)G_0 = 1 \tag{3.45}$$

Defining the change in a Hamiltonian matrix as ν means that:

$$(Z-H_0-\nu)(G+\triangle G) = 1 \tag{3.46}$$

Using eqn 3.45 and ignoring second order terms leaves the expression $\Delta G_0 = G_0\nu G_0$ for the change in the Green's function. So the change in n_I^{σ} is:

$$\Delta n_I^{\sigma} = \sum_{J\vec{R}',i\in I,j\in J} -\frac{1}{\pi}\mathrm{Im}\int G_{\sigma i\vec{R}j\vec{R}'}(E+i0^{+})\Delta\epsilon_j^{\sigma}G_{\sigma j\vec{R}'i\vec{R}}(E+i0^{+})f(E-\mu)\,dE \tag{3.47}$$

Differentiating eqn 3.47 with respect to the on-site energy results in eqn 3.48:

$$K^{\sigma}_{IJ} = \sum_{i\in I, j\in J} -\frac{1}{\pi}\mathrm{Im}\int f(E-\mu)\sum_{\vec{R}'} G_{\sigma i\vec{R} j\vec{R}'}(E+i0^{+})G_{\sigma j\vec{R}' i\vec{R}}(E+i0^{+})\,dE \quad (3.48)$$

Calculating the product of the Green's functions in eqn 3.48 using the definition from (3.43) results in eqn 3.49.

$$\begin{aligned}
&\sum_{\vec{R}'} G_{\sigma i\vec{R} j\vec{R}'}(E+i0^{+})G_{\sigma j\vec{R}' i\vec{R}}(E+i0^{+}) \quad (3.49)\\
&= \frac{1}{N_k^2}\sum_{n\vec{k}} \frac{c_i^{\sigma n\vec{k}} c_j^{\sigma n\vec{k}*}}{E+i\eta-\epsilon^{\sigma}_{n\vec{k}}} \sum_{m\vec{k}'} \frac{c_j^{\sigma m\vec{k}'} c_i^{\sigma m\vec{k}'*}}{E+i\eta-\epsilon^{\sigma}_{m\vec{k}'}} \sum_{\vec{R}'} \exp\left(i(\vec{k}-\vec{k}')\cdot(\vec{R}-\vec{R}')\right)\\
&= \frac{1}{N_k}\sum_{\vec{k}nm} \frac{c_i^{\sigma n\vec{k}} c_j^{\sigma n\vec{k}*} c_j^{\sigma m\vec{k}} c_i^{\sigma m\vec{k}*}}{(E+i\eta-\epsilon^{\sigma}_{n\vec{k}})(E+i\eta-\epsilon^{\sigma}_{m\vec{k}})}
\end{aligned}$$

The full expression for K^{σ}_{IJ} is thus:

$$K^{\sigma}_{IJ} = -\frac{1}{\pi}\mathrm{Im}\int f(E-\mu)\lim_{\eta\to 0}\frac{1}{N_k}\sum_{\vec{k}nm} \frac{\sum_{i\in I} c_i^{\sigma n\vec{k}} c_i^{\sigma m\vec{k}*}\sum_{j\in I} c_j^{\sigma n\vec{k}*} c_j^{\sigma m\vec{k}}}{(E+i\eta-\epsilon^{\sigma}_{n\vec{k}})(E+i\eta-\epsilon^{\sigma}_{m\vec{k}})}\,dE \quad (3.50)$$

which upon evaluation of the integral leads to the final expression in eqn 3.51.

$$K^{\sigma}_{IJ} = \frac{1}{N_k}\sum_{\vec{k}nm} \frac{f(\epsilon^{\sigma}_{n\vec{k}}-\mu)-f(\epsilon^{\sigma}_{m\vec{k}}-\mu)}{\epsilon^{\sigma}_{n\vec{k}}-\epsilon^{\sigma}_{m\vec{k}}}\Big(\sum_{i\in I} c_i^{\sigma n\vec{k}} c_i^{\sigma m\vec{k}*}\Big)\Big(\sum_{j\in J} c_j^{\sigma n\vec{k}*} c_j^{\sigma m\vec{k}}\Big) \quad (3.51)$$

From the expression for K^{σ}_{IJ} in eqn 3.51, it can be seen that there is a double sum over the band indices and sums over the orbitals i and j. The double sum over the bands is for each I and J. If N is the number of atoms, there are N^2 terms, and each term requires N^2 operations to compute. The sums over i and j are over the number of orbitals per site, and so are independent of the number of atoms. The overall dependence on the calculation of K^{σ}_{IJ} (the most crucial quantity in the calculation of the linear response matrix) is therefore dependent on the fourth power of the number of atoms.

3.5 PULAY MIXING AND THE LINEAR RESPONSE MATRIX

Pulay mixing can be slotted into the formalism of the linear response matrix quite neatly. Instead of calculating this matrix exactly an assumption is made about its form. The starting assumption of Pulay mixing is that input quantities are optimized by updating them as a linear combination of a specified number of previous inputs. This is described using magnetic moments as an example in eqn 3.52.

$$\triangle m_I^{(p+1)} = \sum_{r=0}^{N_s-1} a_r \triangle m_I^{(p-r)} \tag{3.52}$$

where p is the iteration number, N_s is the number of previous magnetic moments that are to be included, $m_I^{(p)}$ is the input magnetic moment on a site I , $\Delta m_I^{(p)} = m_I^{(p)} - m_I^{(p-1)}$, and a_r unknown coefficients to be calculated. A first order Taylor expansion of the residue with respect to the change in magnetic moments gives eqn 3.53. Note that this gives eqn 3.28 if we take the final residue to be zero.

$$R_I^{m(p+1)} = R_I^{m(p)} + \sum_J \frac{\partial R_I^{m(p)}}{\partial m_J} \triangle m_J^{(p+1)} \tag{3.53}$$

Inserting eqn 3.52 into eqn 3.53 gives

$$R_I^{m(p+1)} = R_I^{m(p)} + \sum_{r=0}^{N_s-1} a_r \sum_J \frac{\partial R_I^{m(p)}}{\partial m_J} \triangle m_I^{(p-r)} \tag{3.54}$$

Assuming that $\partial R_I^{m(p)}/\partial m_J$ is independent of the iteration number, eqn 3.53 can be substituted into eqn 3.54 giving:

$$R_I^{m(p+1)} = R_I^{m(p)} + \sum_{r=0}^{N_s-1} a_r \left(R_I^{m(p-r)} - R_I^{m(p-r-1)} \right) \tag{3.55}$$

The residues $R_I^{m(p+1)}$ form a vector whose magnitude is minimised with respect to the coefficients a_r.

$$\frac{\partial}{\partial a_r} \sum_I (R_I^{m(p+1)} R_I^{m(p+1)}) = 0 \tag{3.56}$$

Inserting Eq. 3.55 into Eq. 3.56 and following through with the differentiation results in a set of simultaneous linear equations.

$$
\begin{aligned}
0 &= \sum_I R_I^{m(p)}(R_I^{m(p-r)} - R_I^{m(p-r-1)}) \\
&+ \sum_{r'=0}^{N_s-1} a_{r'} \sum_I (R_I^{m(p-r)} - R_I^{m(p-r-1)})(R_I^{m(p'-r')} - R_I^{m(p'-r'-1)})]
\end{aligned}
$$

Solving this set of simultaneous equations provides the values of the coefficients a_r and in turn the now optimized input magnetic moments. To make the residue still lower and obtain information about the output magnetic moments, one step of linear mixing is required:

$$\triangle m_I^{(p+1)'} = \triangle m_I^{(p+1)} + C R_I^{m(p+1)} \tag{3.57}$$

where $R_I^{m(p+1)}$ is found from Eq. 3.55 and C simply a constant. The time taken for calculations should scale linearly with the system size.

3.6 THE NEWTON-RAPHSON MIXING SCHEME IN ACTION

The efficacy of the Newton mixing scheme is exemplified in a comparison of the various mixing schemes used on a 3x3x3 cell of bulk bcc ferromagnetic iron (Table 1). It can be seen that the Newton Mixing scheme can achieve convergence in only 4 loops. From this test it seems that Newton mixing is definitely the answer to reaching a speedy convergence.

However there are situations in which Newton mixing cannot achieve convergence even with an indefinite number of loops. This is the case for a cell of bulk fcc ferromagnetic iron of atomic volume $10.65 Å^3$ using a 24 x 24 x 24 mesh of k-points and an electron temperature of 0.001Ry. Despite all efforts the residue does not go to zero and at a certain stage the residues at each step begin to form a repeating pattern. Starting from an unfortunate choice of input spin no solution is found, instead the algorithm just oscillates. The mystery is elucidated by considering the magnetic moment residue against magnetic moment. This analysis is possible as the charge on all four atoms on the fcc basis cell is zero by symmetry, immensely simplifying the picture and allowing the problem to be reduced to one dimension. In more complex simulations

Method	Number of scf Loops
Linear	infinite
Pulay	7
Newton	4

Table 1: Number of loops to converge bulk

where there is both charge and spin, a more complicated version of the problem will occur. The energy surface as a function of both spin and charge is 'bumpy' leading to similar difficulties in convergence.

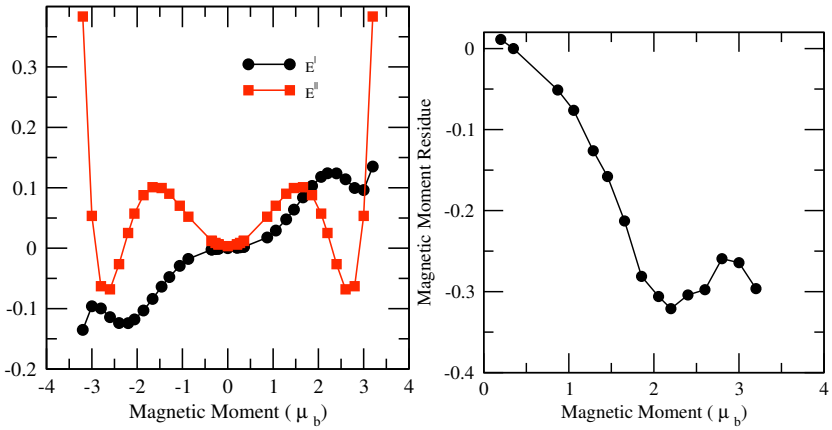


Figure 9: The left panel shows the variation of the first and second derivatives of energy with magnetic moment for the low spin state of bulk ferromagnetic fcc iron at a volume per atom of 10.65Å^3. The first derivative is labelled E' and the second derivative E''. The right panel shows the spin residue which determines by how much the input spin is adjusted for one update. First published in [82]

From the plot of the magnetic moment residue against magnetic moment per atom in the right panel of Figure 9 the reason for the difficulties is illuminated. A minimum point exists before the residue reaches zero. This is at the root of the problem. Considering applying the graphical application of the Newton-Raphson procedure it is obvious that drawing tangents at points near the minimum of the graph will result in a process which can carry on indefinitely if left. From eqn 3.22 it is understood that the magnetic moment residue R_J^q is proportional to the first derivative of the energy $\frac{\partial E_M}{\partial m_I}$ as show in Figure 9. Any difficulties should therefore also occur in the first derivative of the energy as well as the residue and sure enough there are minima present in Figure 9.

Mix factor	No. Loops (min)	No. Loops Total
0.2	42	69
0.1	24	70
0.05	20	53
0.02	20	56

Table 2: Finding an efficient mix factor. The number of loops (min) refers to the number of loops required to leave the local energy minimum.

To extricate simulations from this predicament it is necessary to employ Pulay mixing [77] for spin. The basic premise of Pulay mixing is that it uses several previous input spins allowing widely spaced points on either side of the local minimum to be chosen. It then interpolates between them avoiding becoming trapped by the local minimum. With more complicated defects the same explanation still holds except extended to consider multi-dimensional, 'rough' surfaces instead of curves. An empirical approach is taken to find the most efficient spin mix factor for Pulay mixing. The quantity taken into account is the number of loops taken for the residue to 'climb out' of the local minimum in Figure 9. During these tests the electron temperature is kept at 0.001 Ry and the number of k points is 24x24x24. From Table 2 it can be seen that the highest values of the mix factor are the least efficient but that this effect levels off at a value of about 0.05. The total number of loops is also the lowest at this mix factor.

It is instructive to consider the differences between Newton charge mixing and Pulay charge mixing, whilst using Pulay spin mixing in both cases. It can be seen that there is a dramatic improvement in the number of loops needed for convergence using Newton charge mixing in comparison with Pulay charge mixing as can be seen in Figure 10. The shallow regions in the graph are where the charge is being slowly converged and the steeper sections where charge and spin are being mixed together.

Deciding which combination of methods to employ in general reduces to a competition between the number of loops and the time taken for each loop. In most small cell bulk calculations it is quickest to use

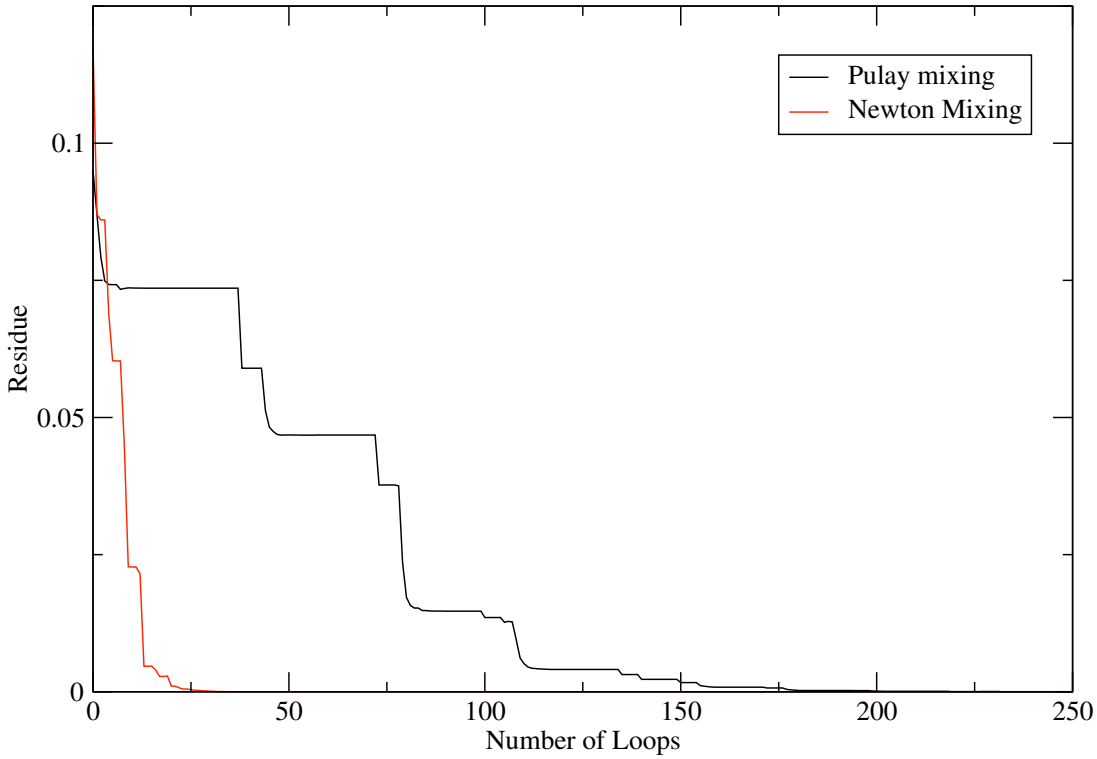


Figure 10: Newton Charge mixing verses Pulay charge mixing for a 137 atom bulk bcc iron wire

Newton mixing for both charge and spin as the number of loops will be the minimum achievable and the time taken for each loop will be negligible. In more complex simulations of point defects with approximately 150 atoms or less, we reach a compromise and use Newton mixing for charge and Pulay mixing for spin. This allows us to be sure of reaching convergence, which is achieved within around 50 loops while the time for each loop is still reasonable at 35s. As we get to larger configurations of atoms (150-500 atoms) containing line defects such as dislocations, we find that using Pulay mixing for both spin and charge is preferable given the time taken for each loop.

Obtaining a fast and reliable method of reaching charge and spin convergence for a variety of defect simulations has provided the driving motivation behind the development of the Newton-Raphson scheme.

3.7 CONCLUSIONS

Charge transfer has successfully been added to an existing tight binding model which previously worked on the assumption of local charge neutrality. This has been implemented using an efficient mixing scheme based on the simple idea of the Newton-Raphson minimisation procedure. Difficulties achieving convergence with this method have been investigated and can be attributed to the shape of the tight binding energy surface. Computationally, the main challenge faced is calculating the linear response matrix, and it is this which provides the time and memory constraints on simulations that can be carried out. In the following chapter this method is applied to simple bulk structures and point defects.

4

POINT DEFECTS

4.1 BULK CALCULATIONS

Having established a lucid and efficient scheme for achieving charge and spin self consistency, the next logical step is to test it on all manner of bulk calculations. The significant relationships to establish for the various bulk structures are between energy and volume, and magnetic moment and volume. Comparison with the existing tight binding model which uses local charge neutrality (LCN) [59] (extended in this work to include charge) and DFT calculations [53, 45] is favourable. In addition consideration of the variation in energy with magnetic moment for bulk bcc ferromagnetic iron highlights some of the imperfections and potential pitfalls in the Newton-Raphson method. Establishing the values of certain elastic constants gives further verification of the ability of the method to be used to study iron and its defects.

4.1.1 *Variation of energy with volume*

The binding energy curves for all the bcc (body centered cubic) and fcc (face centered cubic) cases required are produced by varying the cell vector, and then at each cell vector allowing the atoms to fully relax. For the hcp (hexagonal close packed) case, there are two cell vectors a and c as indicated in Figure 13. To begin with the value of a is kept static and c is varied methodically. A curve is then fitted to the plot of energy against c. The optimum c depends on a, so it is necessary to optimise c for a range of a. To be fully confident of the accuracy of the results it was necessary to apply much tighter convergence conditions for self-consistency than in previous tight binding calculations which assumed LCN [59] and used the OXON code [50]. A tolerance in the

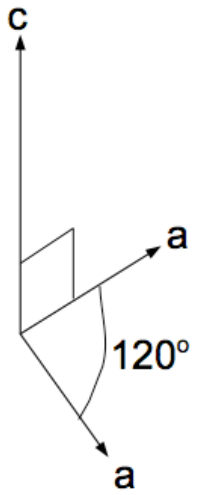


Figure 11: The axes in an hcp structure

spin of $10^{-8}\mu_b$ and in the charge of $10^{-8}e$ was considered suitable. The binding energy curves are constructed for non-magnetic, ferromagnetic and antiferromagnetic cases. The k point mesh is extremely fine. For all bulk calculations this is set to $24 \times 24 \times 24$. Atoms in the perfect bulk are of course uncharged by symmetry and so the value of the Hubbard U parameter is not relevant. The electron temperature is always 0.001Ry; a low temperature is chosen as we do not want temperature to have a significant effect.

The binding energy curves of bcc, fcc and hcp iron are shown in Figures 12 and 13 with bcc curves included in both figures for ease of comparison. The ferromagnetic bcc structure has been found to be the most stable from previous studies [84, 45, 85, 53]. In the bcc structure it is possible to have non magnetic (nm), ferromagnetic (fm) and antiferromagnetic (afm) states. The bcc antiferromagnetic case has a cubic primitive cell with a basis of two atoms at positions $(0,0,0)$ and $(\frac{a}{2},\frac{a}{2},\frac{a}{2})$ which have equal and opposite spins; a is the lattice constant. For both the fcc and hcp structures there is a non magnetic state, an antiferromagnetic state and two ferromagnetic states (high spin and low spin) characterised by different lattice constants (high spin occurs at the larger lattice constant as it reduces the overlap between the d-orbitals). In the hcp case the low spin ferromagnetic case is very unstable and it was not possible to perform complete calculations on it.

For the antiferromagnetic states the spins are arranged with the two atoms in the unit cell at positions $(0,0,0)$ and $(\frac{2a}{3}, \frac{a}{3}, \frac{c}{2})$ which have equal and opposite spins.

The curves obtained agree qualitatively with those computed using DFT [53, 45], reproducing the high and low spin ferromagnetic curves of fcc and hcp. A comparison of the equilibrium atomic volumes for the different structures with those found from DFT [45] and tight binding [59] is given in Table 3. The tight binding values compared to are those obtained using the same tight binding model but with LCN imposed. For perfect solids, symmetry ensures the atoms are neutral, so these results should be identical to those obtained. Any differences must be due to $k-$point sampling, electron temperature, self-consistency convergence criteria or the technique for finding the minimum energy lattice constant. The PLATO code used for the Newton-Raphson method automatically searches for the zero pressure configuration. The equilibrium atomic values for the various structures are close to both DFT and OXON values, highlighting the success of both the tight binding model and the new mixing scheme.

Instructive comparisons can also be gained by considering the minima of the above graphs, that is the volume of the basis cell at equilibrium as detailed in Table 3.

4.1.2 *Variation of magnetic moment with volume*

The curves of magnetic moment against atomic volume are all in good agreement with previous studies using both DFT and charge neutral tight binding. These are illustrated in Figure 14 for fcc structures and Figure 15 for hcp with bcc structures included in both figures. The most striking features of the ferromagnetic curves is the fcc high spin case,

	DFT		Tight Binding with LCN		Tight binding with charge	
	Eq. At.Vol ($\dot{A}^3$)	Mag. Mom/atom (μ_b)	Eq. At.Vol ($\dot{A}^3$)	Mag. Mom/atom (μ_b)	Eq. At.Vol ($\dot{A}^3$)	Mag. Mom/atom (μ_b)
bcc nm	10.59	0.00	10.73	0.00	10.72	0.00
bcc fm	11.44	2.17	11.57	2.61	11.50	2.62
bcc afm	10.97	1.25	-	-	11.16	1.93
fcc nm	10.34	0.00	10.70	0.00	10.50	0.00
fcc fm high spin	12.03	2.57	-	-	12.37	3.20
fcc fm low spin	10.62	1.02	-	-	10.54	0.29
fcc fm	10.69	1.59	10.84	1.30	10.77	1.30
hcp nm	10.32	0.00	10.39	0.00	10.40	0.00
hcp afm	11.40	2.04	11.04	0.37	11.00	2.16
hcp fm high spin	12.03	2.55	12.41	3.07	12.36	3.20

Table 3: Comparison of equilibrium atomic volumes and magnetic moment per atom for various bulk structures. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82]

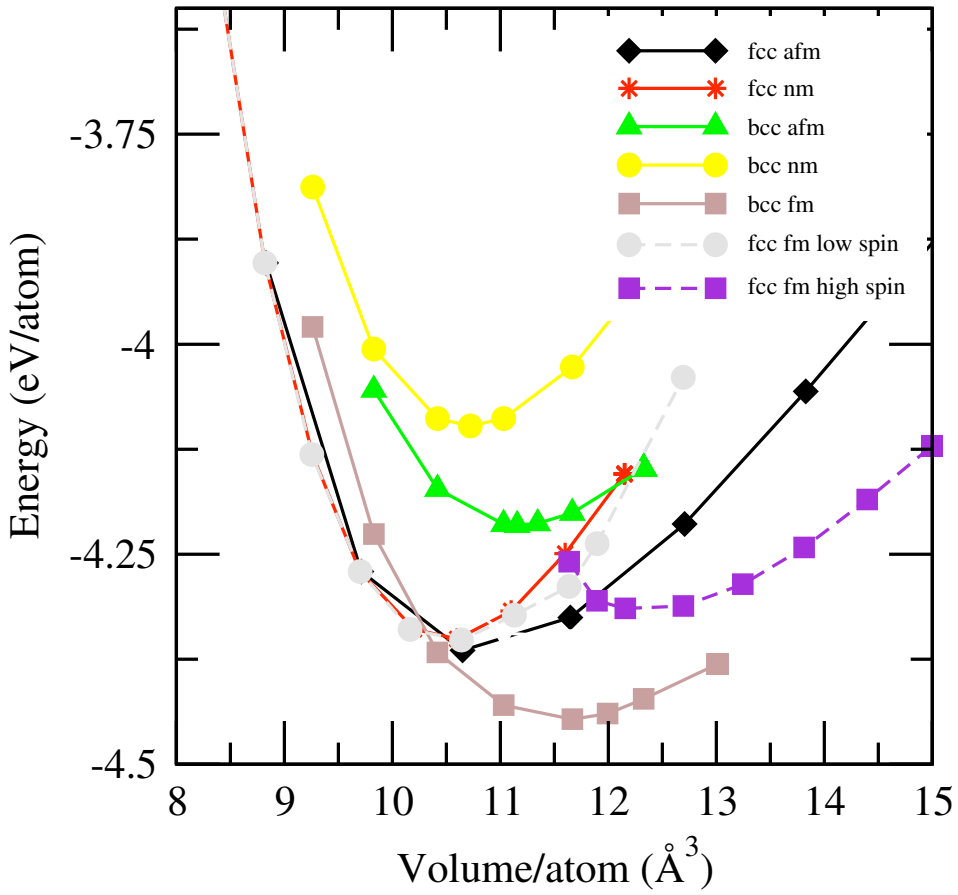


Figure 12: Binding energy curves for fcc and bcc iron for a variety of magnetic structures. First published in [82]

which has a sharp step between the non-magnetic and magnetic. This is also true for the hcp ferromagnetic and antiferromagnetic structures.

4.1.3 *Variation of energy with magnetic moment*

It is instructive to consider the variation in energy with the magnetic moment of a bulk system to gain an understanding of the obstacles to obtaining spin convergence. Theoretically self consistency could be achieved in a single step using the Newton-Raphson procedure providing that the energy surface of the test system were exactly quadratic. From experience with bulk calculations it is found that this is not so and that a number of iterations are necessary. These iterations must be conducted in a specific manner. In the case of magnetic systems it is discovered that mixing charges first (at fixed input moments) until they are converged is both more efficient and reliable than updating the

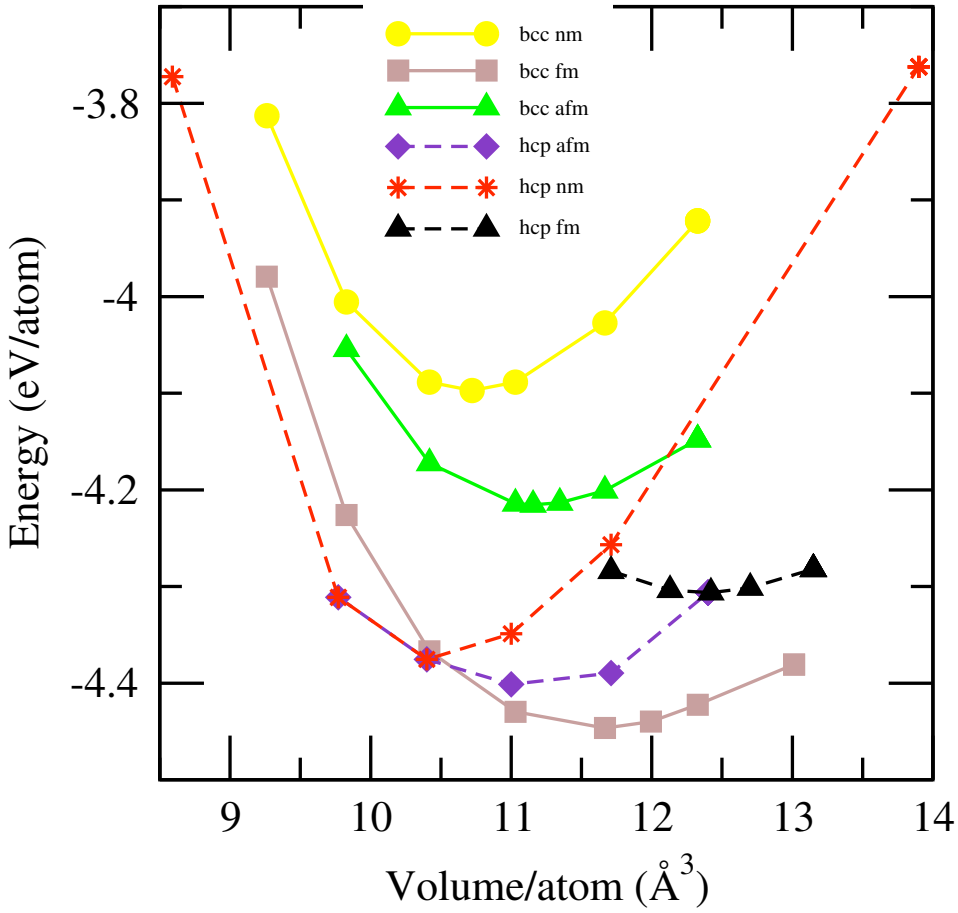


Figure 13: Binding energy curves of for hcp and bcc iron for a variety of magnetic structures. First published in [82]

charge and moments in a combined step. A charge tolerance is defined so that if the residue exceeds this the code returns to updating charges only. To prevent the changes in spin and charge becoming too great in one step a maximum change in the charge of $0.01e$ and in the spin of $0.1\mu_b$ is imposed.

Body centred cubic BCC

The example of bulk ferromagnetic bcc iron is considered in detail to illustrate the origin of some of the complexities. As it is a bulk system, symmetry ensures charge transfer between atoms does not occur and the magnetic moment is the same on every atom. The magnetic moment per atom is therefore the only remaining degree of freedom that can change through the self-consistency cycles. All calculations are con-

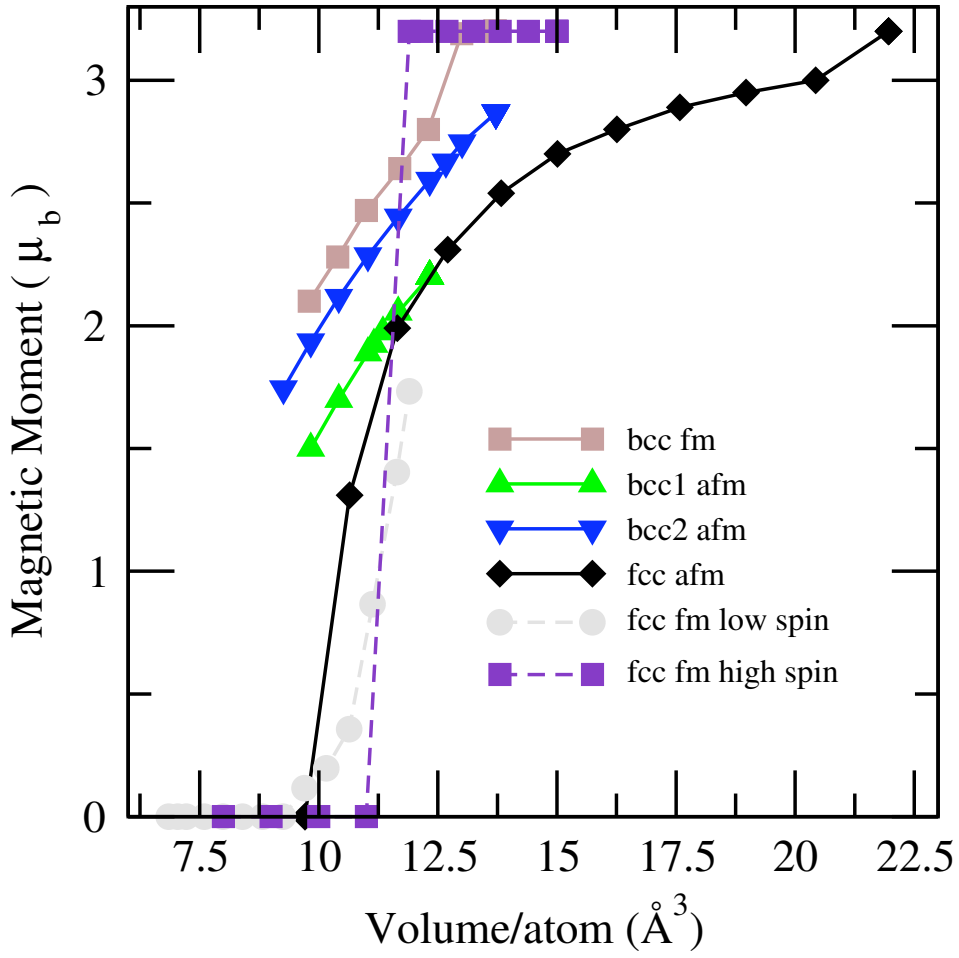


Figure 14: Variation of magnetic moment per atom with atomic volume for fcc and bcc iron. First published in [82]

ducted at the experimental equilibrium lattice parameter 2.8665Å with a 24 x 24 x 24 k-point mesh and an electronic temperature of 0.001Ry.

The standard Newton-Raphson update equation introduced in eqn 3.21 can be specialised to the situation where there is only one magnetic moment that can be varied leading to eqn 4.1.

$$\Delta m_I = -\left.\frac{\partial E_M}{\partial m_I}\right|_0 \Big/ \left.\frac{\partial^2 E_M}{\partial m_I^2}\right|_0 \tag{4.1}$$

In the left panel of Figure 16 are plotted the first and second derivatives of E_M and in the right panel the negative of their ratio. The four magnetic moments where $\partial^2 E_M/\partial m_I^2 = 0$ are all points where Δm_I diverges. The values of magnetic moment at which the divergences occur indicate boundaries between regions that asymptotically take the calculations to different solutions as indicated in the right panel of Figure 16. Two important technical points emerge from this. Firstly

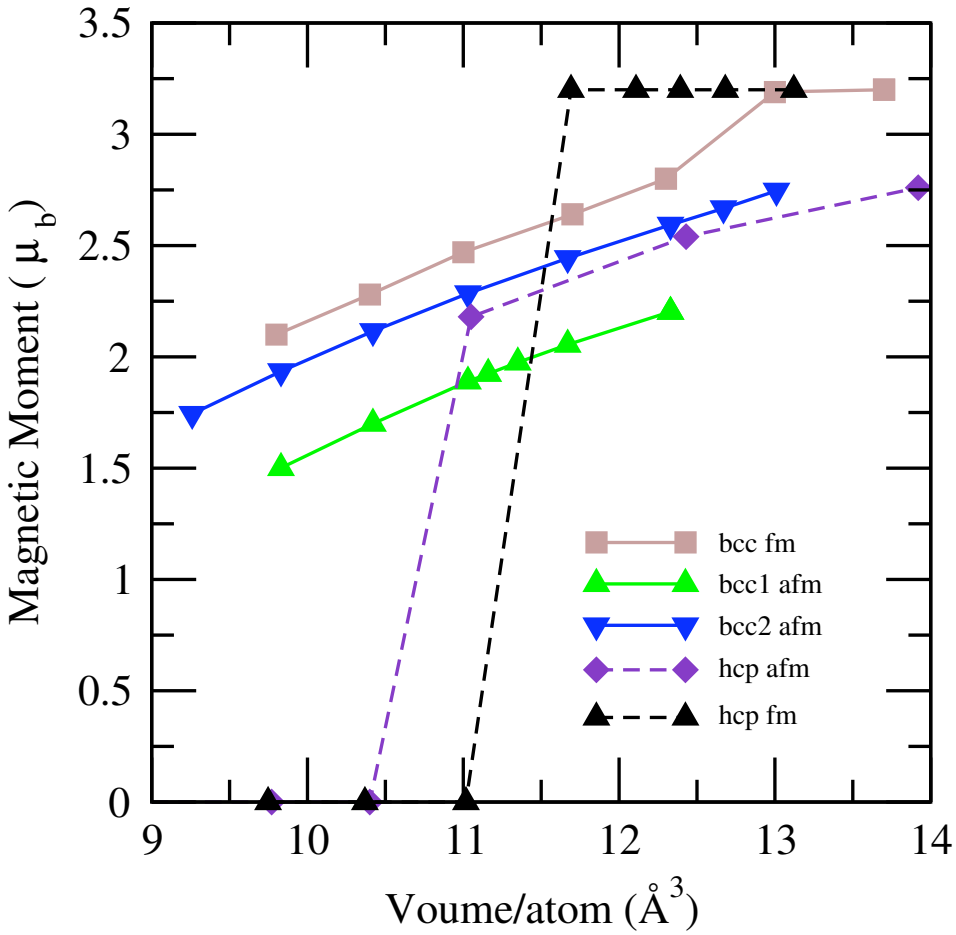


Figure 15: Variation of magnetic moment per atom with atomic volume of high spin hcp and bcc iron. First published in [82]

that the initial choices for the input moments must be close to the final solution and secondly that an upper bound to the magnitude of any single update be imposed to guard against massive variations in the input magnetic moments.

To the naked eye the variation of total energy with input magnetic moment per atom looks like a quartic polynomial as can be seen from the black dotted curve in the left panel of Figure 17. However, it is necessary to use a polynomial of higher order to obtain an acceptable fit to the computed data points. A 10th order polynomial of which the coefficients of the odd powers are negligible is sufficient. A 4th order polynomial is fitted to the actual curve in the left panel of Figure 17 of which the first and second derivatives are calculated in the right panel of Figure 17. This illustrates how simple they are in comparison with those of the 10th order fit in the left panel of Fig. 16. Employing

the quartic fit it can be seen that the second derivative of the energy with respect to magnetic moment is just a simple parabola and that it is therefore the higher order corrections to the quartic that introduce the difficulties with the convergence of magnetic moment.

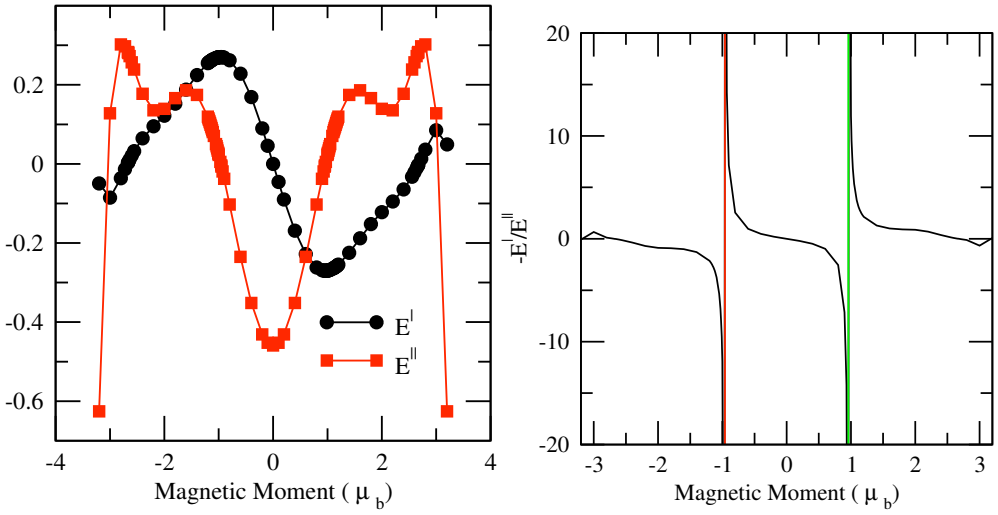


Figure 16: The left panel shows the variation of the first and second derivatives of energy with magnetic moment for bulk ferromagnetic bcc iron. The first derivative is labelled E' and the second derivative E''. The right panel shows the ratio $-E'/E''$ which determines by how much the input spin is adjusted for one update. The red and green vertical asymptotes in the right panel separate regions where the Newton-Raphson iteration tends to different solutions. First published in [82]

4.1.4 *Elastic Constants*

The remaining bulk quantities to verify are the elastic constants. These provide information about how a material behaves under strain and how it returns to its original shape. The transformation matrix $T_{\beta\eta}$ acts on the basis vectors $a_{j\eta}$ of a simulation cell by an amount γ to produces the new basis vectors $a_{j\beta'}$. The transformation matrix defines the nature of the applied strain.

$$a_{j\beta'} = a_{j\beta} + \gamma \sum_{\eta} T_{\beta\eta} a_{j\eta} \tag{4.2}$$

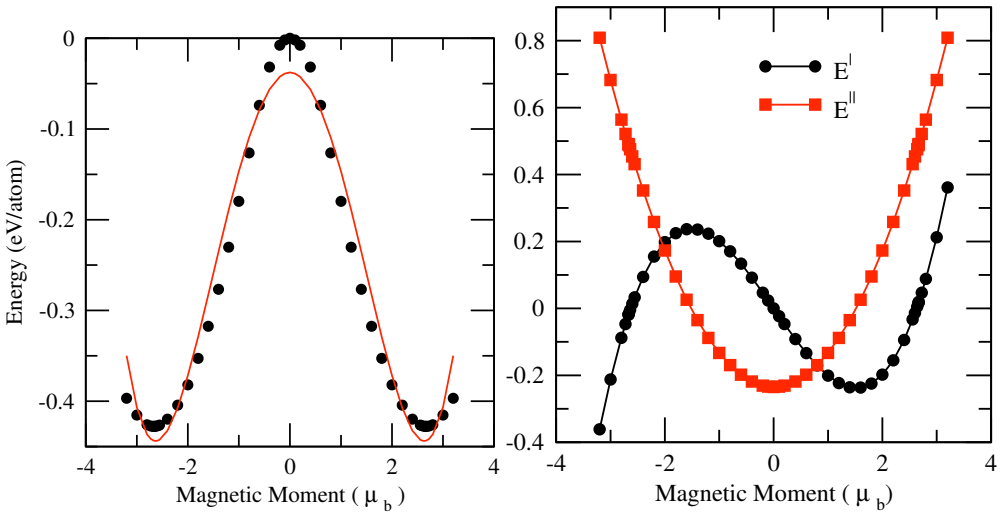


Figure 17: Left Panel: Variation in energy with magnetic moment for bulk bcc ferromagnetic iron. The black dots represent the real curve and the red line is the fourth order fit to this. Right Panel: The first and second derivatives of the fourth order.First published in [82]

Bulk modulus

The bulk modulus is obtained by varying the volume of the cell according to the matrix T which is the unit matrix.

$$B = \frac{1}{9V_a}\frac{\partial^2 E_{total}}{\partial \gamma^2} \tag{4.3}$$

To evalate B, the value of γ is varied and the energy at each point plotted. A polynomial is then fitted to the resulting curve and the second derivative calculated from the coefficient of the second order term at $\gamma = 0$. V_a is the volume of the minimised cell in $\dot{A}^3$.

C_{44}

This describes a shear on the (100) plane.

$$T = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.4}$$

$$C_{44} = \frac{1}{4V_a}\frac{\partial^2 E_{total}}{\partial \gamma^2} \tag{4.5}$$

C'

This describes a shear on the (110) plane.

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.6}$$

$$C' = \frac{1}{4V_a}\frac{\partial^2 E_{total}}{\partial \gamma^2} \tag{4.7}$$

The resulting elastic constants are then compared to corresponding values from tight binding with LCN and DFT. From Table 4 it can be seen that the comparison is quite favourable.

Elastic Constant	DFT (GPa)	Tight Binding with LCN (GPa)	Tight Binding with charge (GPa)
B	174	168	162
C_{44}	69	36	35
C'	99	118	123

Table 4: Comparison of Elastic Constants in GPaDFT values from [41] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82].

4.2 POINT DEFECTS

The results from bulk calculations being encouraging, the next test for both the Newton-Raphson mixing scheme and the charge modifications to the original tight binding model is point defects. Calculations on these defects challenge the method far more rigorously than simple bulk structures as the condition of LCN does not automatically follow from symmetry considerations. It is expected that the inclusion of monopolar charges on every atomic site will result in some difference in the atomic configuration and electronic structure and hence the defect energies in comparison with LCN. The types of defects considered are vacancies and interstitials.

As the original tight binding model employed is constructed for LCN, a large value of the Hubbard U is selected. This is chosen to be 1Ry in agreement with [74]. The variation of the energy of the vacancy with the value of U is considered and it is found to change very little. A smaller value of U actually pushes the energy marginally further away from the value provided by DFT [20]. The complete effect of U is discussed in detail in the following section.

4.2.1 *The Vacancy and the effect of the Hubbard U*

The most simple point defect to consider is the single vacancy. This is created by removing the central atom in a cubic cell and then allowing both the atomic positions and the cell vectors to relax, keeping the shape of the cell fixed. The energies of a vacancy in 54 and 128 atom cell are calculated and then compared with values from DFT and tight binding using LCN. The charge self-consistent TB results agree well with the LCN TB results but less well with the DFT results. The TB results with and without charge self-consistency are almost unchanged when the cell size is increased from 53 to 127 atoms, although the DFT results change significantly. This is detailed in Table 5. The agreement with tight binding using LCN is closest to our results indicating that the effect of monopolar charge is small. To keep the density of the k-point mesh the same, $8 \times 8 \times 8$ k-points were used in calculations on the 54 atom cell and $6 \times 6 \times 6$ k-points for the 128 atom cell. The electron temperature is kept at 0.001Ry and a value for Hubbard U of 1Ry is used in all calculations.

From Table 6 it can be seen that there is a very small change in vacancy formation energy for three values of U. It is also worth considering the relationship between charge and distance from the vacancy for the three values of U chosen. The radial shells of atoms around the central vacancy are labelled by neighbour index, with 1 being the shell closest to the vacancy site. This is illustrated in Figure 18.

Method	No. Atoms	E^f (eV)
DFT	53	1.95
Tight Binding with LCN	53	1.68
Tight Binding with charge	53	1.63
DFT	127	2.07
Tight Binding with LCN	127	1.68
Tight Binding with charge	127	1.63

Table 5: Comparison of Formation Energy of Single Vacancy. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82]

Value of U (Ry)	No. Atoms	E^f (eV) Tight Binding with charge (Inter-site description)
1.0	53	1.63
0.5	53	1.62
0.25	53	1.61
1.0	127	1.63

Table 6: Variation in Formation Energy of Vacancy with value of Hubbard U. First published in [82]

The results are as expected. The greatest fluctuations in charge occur with the lowest value of U considered (0.25Ry). For the largest value of U (1Ry) the charge is almost exactly zero across all radial shells, whilst the intermediate value of U displays a behaviour somewhere between these two extremes. Overall, however, the magnitude of the charges is small for all values of U. This is unsurprising when it is borne in mind that the original tight binding model was built for local charge neutrality (LCN).

As a sanity check on the behaviour of the vacancy, the displacement of the atoms surrounding the vacancy was detailed. For each of these shells the percentage displacement relative to the perfect crystal positions of the atoms is calculated. As expected further away from the vacancy the atoms return to their undisturbed positions. As can be seen from Figure 19 it is only the two first nearest shells that are affected. Reassuringly,

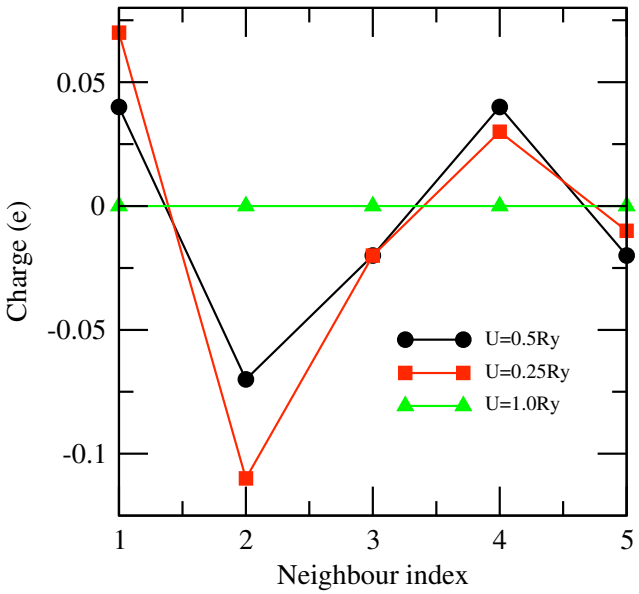


Figure 18: Variation in charge around a single vacancy in a 53 atom cell. The neighbour index indicates the shell around the vacancy with increasing radius. First published in [82]

there is also little difference indicated by changing the value of the Hubbard U

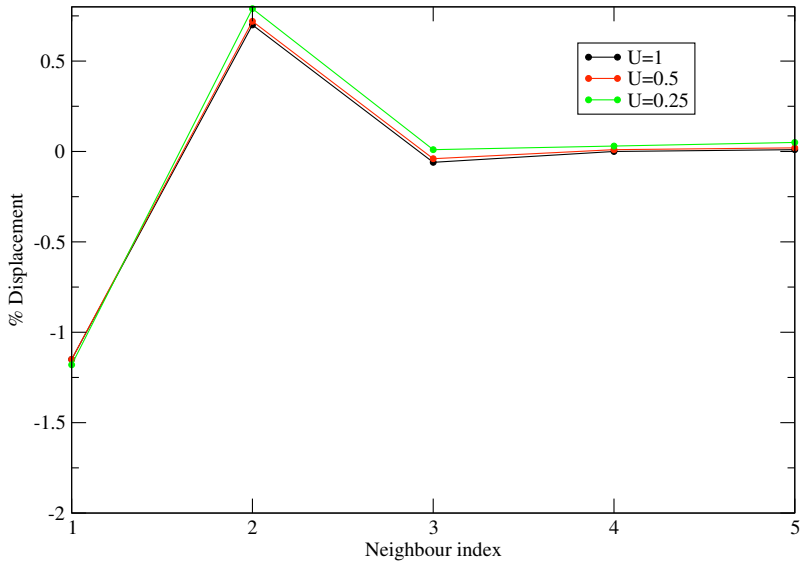


Figure 19: Percentage Displacement around the 53 atom vacancy

The relationship between the magnetic moment and the neighbour index as shown in Figure 20, tallies well with that of the percentage displacement. For the third neighbour shell and beyond, the magnetic moment levels off to the bulk value as expected. Close to the vacancy it has the highest value which decreases sharply between shells one

and three. This is readily understood by appreciation of the fact that around the vacancy the atoms have fewer neighbours, meaning that the density of states will be taller and narrower and which in turn means that magnetism is favoured.

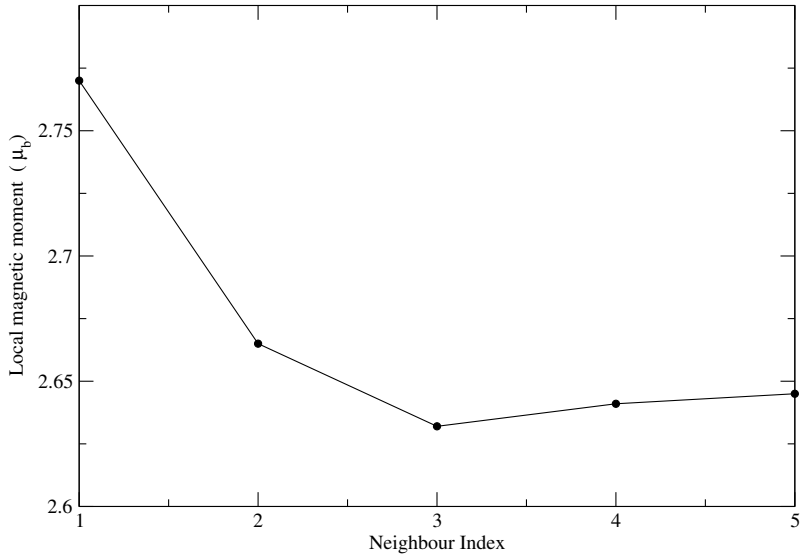


Figure 20: Magnetic moment around the 53 atom vacancy

4.3 INTERSTITIALS

An interstitial is an entirely more challenging form of defect to simulate. It consists of removing an atom and replacing it by two atoms which are equidistant from either side of the position of the original atom. It is possible to place the pairs of atoms in different directions, enabling different interstitials to be produced. The three interstitials considered in this study are the $\langle 100 \rangle$, $\langle 110 \rangle$ and $\langle 111 \rangle$.

4.3.1 *Energetic ordering of interstitials*

The same numbers of k-points, electron temperature and relaxation strategy are used as for the vacancy. As well as including the monopole interactions as discussed above, the effect of only allowing Coulomb interactions on-site was considered; that is a finite value of U_I but $V_{IJ} = 0$ when $I \neq J$. The energies found using the two different descriptions (on-

site and inter-site) of V_{IJ} are similar, though including all the inter-site terms generally lowers the energy slightly. The same relative ordering of the formation energies of the interstitials is obtained compared with tight binding using LCN [59] and DFT [20] as indicated in Table 7. As was the case for the vacancy, the charges on the atoms around the interstitial are small in all cases. There is, however, some disagreement between our results and those obtained from tight binding using LCN, especially in the magnetic moments as can be seen in Table 8. This requires further investigation, but the finer k point sampling may be the reason for this. The magnetic moment is sensitive to the density of states at the Fermi surface which is better sampled when more k points are used. For the 55 atom interstitials $8 \times 8 \times 8$ k points are used and for the 129 atom interstitials $6 \times 6 \times 6$ k points, thus keeping the density of the k point mesh the same. This is significantly more than those used in the tight binding calculations using LCN which use $6 \times 6 \times 6$ k points for the 55 atoms interstitials and $5 \times 5 \times 5$ k points for the 129 atom interstitials. The DFT calculations used a $5 \times 5 \times 5$ k point mesh for 55 atom calculations, and a $3 \times 3 \times 3$ k point mesh for 129 atom calculations. It is also found that in the case of the on-site description of a $\langle 111 \rangle$ intersitial, the interstitial seeks to become a $\langle 110 \rangle$ interstitial upon cell relaxation. Hence the value for the energy before this occurs is chosen. On-site only Coulomb interactions are not considered separately to the full inter-site description for the 129 atom interstitials. This is because from Table 7 it can be seen that there is little difference in the results in the formation energies of the 55 atom interstitials between the two descriptions. Therefore inter-site calculations for the 129 atoms intersitials would not be profitable.

4.3.2 *A close look at the spins*

Variation in spin with distance from the interstitial

The variation in magnetic moment with distance from the position of the removed central atom is plotted for the three interstitials. The

Interstitial	No. Atoms	E^f (eV) Tight binding with charge (On-site description)	E^f (eV) Tight binding with charge (Inter-site description)	E^f (eV) Tight binding with LCN	E^f (eV) DFT
<100>	55	5.75	5.71	5.37	4.37
<110>	55	4.30	4.23	4.36	3.41
<111>	55	5.28	5.31	4.72	4.11
<100>	129	-	5.67	5.13	4.64
<110>	129	-	4.48	4.28	3.64
<111>	129	-	5.13	4.55	4.34

Table 7: Formation Energies of Interstitials. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82]

Interstitial	No. Atoms	μ_b/atom Tight binding with charge (On-site description)	q/atom Tight binding with charge (Inter-site description)	μ_b/atom Tight binding with charge (On-site description)	q/atom Tight binding with charge (Inter-site description)	μ_b/atom Tight binding with LCN
<100>	55	-1.59	0.03	-1.60	0.05	-
<110>	55	-1.26	0.02	-1.36	0.00	-
<111>	55	-1.96	0.03	-1.95	0.03	-
<100>	129	-	-	-1.39	0.04	0.17
<110>	129	-	-	-1.16	0.02	-0.20
<111>	129	-	-	-1.68	0.00	0.18

Table 8: Charge and Spin on Dumbbell atoms of interstitials. Intersite refers to the inclusion of the long range Coulomb interaction for charge, this is not inluded in onsite calculations.. DFT values from [45] and tight binding with LCN values from [59]. Tight binding with charge values first published in [82]

dumbbell atoms are the ones closest to this position. In the case of the ⟨100⟩ interstitial it can be seen that the closest atoms to the dumbbell pair have a negative magnetic moment (i.e displaying antiferromagnetic behaviour). The magnetic moment of subsequent shells increases steeply until leveling off to the bulk value of $2.6\mu_b$. This is illustrated in Figure 21. The behaviour of the ⟨111⟩ and ⟨110⟩ interstitials is less simple. The atoms furthest from the dumbbell are roughly bulk in character and the magnetic moments of those nearest to it are negative in sign as in the ⟨100⟩ case. However, instead of increasing quickly to bulk values, there is a significant 'dip' in the magnetic moment at about 3.1Å from the interstitial where this becomes almost as negative as the atoms closest to the dumbbell. The reason for this is not obvious. However, by considering the illustrations of the atomic positions of the dumbbell and surrounding atoms in Figures 22 and 23 certain displacements could contribute to the magnetic moment againt distance characteristic. The pink atoms illustrated in these figures show a larger displacement from their equilibrium position than all of the rest of the atoms (excluding the dumbell atoms) indicated in orange. The two dumbbell atoms are included in blue. It is these pink atoms that correspond to the dip in the magnetic moments. Figures 22 and 23 are orientated to show the positions of the pink atoms clearly.

Sensitivity to initial spin

Whilst the energetic ordering is consistent with DFT calculations and tight binding calculations conducted using LCN, there is some discrepancy between the final values of the magnetic moments on the dumbbell atoms of all three of the 129 atom interstitials as indicated in Table 8. It is unknown if this is the case for the 55 atom interstitials as there is no data available for these examples. An insight into the root cause of the difference in spins is gained by using the final atomic positions of the interstitial calculations and making these the starting point for further simulations. These calculations are then restarted but with the magnetic moments on the dumbbell atoms not being the same as on the other atoms in the configuration. In these further calculations

there are noticeable differences in the final magnetic moments on the dumbbell atoms, depending on the initial values as detailed in Table 9. The most significant difference is for the ⟨110⟩ interstitial where there is a change of $0.46\mu_b$ in the magnetic moment. Whilst the reason for this extreme sensitivity to the initial spin is unclear, it can possibly be attributed to the 'roughness' of the energy surface as a function of spin and charge. As well as changes in the final magnetic moments there are necessarily changes in the formation energies. It can be seen that there are reasonably large differences in energy corresponding to the different final spin configurations.

4.3.3 *Technical difficulties presented by the interstitials*

The relaxation of interstitials presented one main technical difficulty. After the atomic forces have reached their prescribed tolerance the minimisation of the cell vector begins. Whilst the atomic relaxation was relatively straight forward, relaxing the cell vector resulted in subsequent cell relaxations being "bounced" from one extreme value to the other. This problem was removed by introducing a simple preconditioning scheme.

The minimisation of the energy of the cell E_c with respect to the cell vector a, is achieved using the basic Newton-Raphson technique, where the change in the cell vector is given by $\tilde{a} = -\frac{1}{\frac{\partial^2 E_c}{\partial a^2}} \frac{\partial E_c}{\partial a}$. The PLATO code originally calculated this using the expression $-\frac{\partial E_c}{\partial a}$ only. It was therefore necessary to calculate $\frac{\partial^2 E_c}{\partial a^2}$ using the bulk modulus B which had already been calculated. The expression for the bulk modulus in terms of the energy is described in eqn 4.8.

$$B = -V\frac{\partial P}{\partial V} = V\frac{\partial^2 E_c}{\partial V^2} \tag{4.8}$$

where P represents the pressure. The aim is to find the relationship between the bulk modulus and the second derivative of the energy with respect to the cell vector.

$$\frac{\partial E_c}{\partial V} = \frac{\partial a}{\partial V}\frac{\partial E_c}{\partial a} \tag{4.9}$$

Interstitial	initial magnetic moment on all atoms (μ_b/atom *no. atoms)	final magnetic moment (μ_b) on dumbbell atoms only	E^f (eV)
⟨100⟩	(1.0*53)(0*2)	-1.66	6.26
⟨100⟩	(2.6*53)(0.2*2)	-1.68	6.19
⟨100⟩	2.6*55	-1.60	5.71
⟨110⟩	(1.0*53)(0*2)	-0.90	4.42
⟨110⟩	2.6*55	-1.36	4.23
⟨111⟩	(1.0*53)(0*2)	-1.84	5.50
⟨111⟩	(2.6*53)(0.2*2)	-1.84	5.50
⟨111⟩	(1.0*53)(2.6*2)	-1.68	6.14
⟨111⟩	2.6*55	-1.95	5.31

Table 9: Sensitivity to initial spin for the three interstitials. All calculations carried out using the tight binding inter-site description of charge.

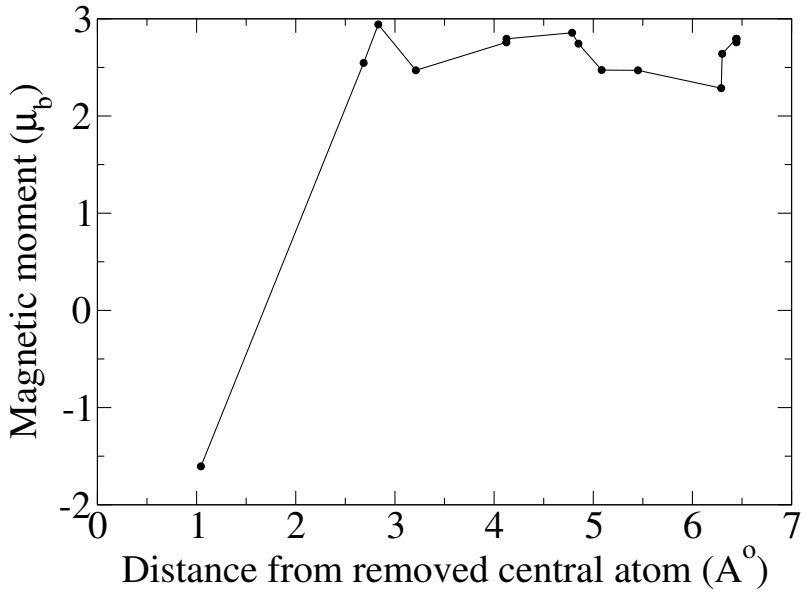
The relationship between the volume of the cell and the cell vector is of course $V = a^3$, for the cubic cell. Differentiating eqn 4.9 using this relationship results in eqn 4.10.

$$\frac{\partial^2 E_c}{\partial V^2} = \left(-\frac{2}{3a^2} \frac{\partial E_c}{\partial a} + \frac{\partial^2 E_c}{\partial a^2} \frac{1}{3a^2} \right) \frac{1}{3a^2} \tag{4.10}$$

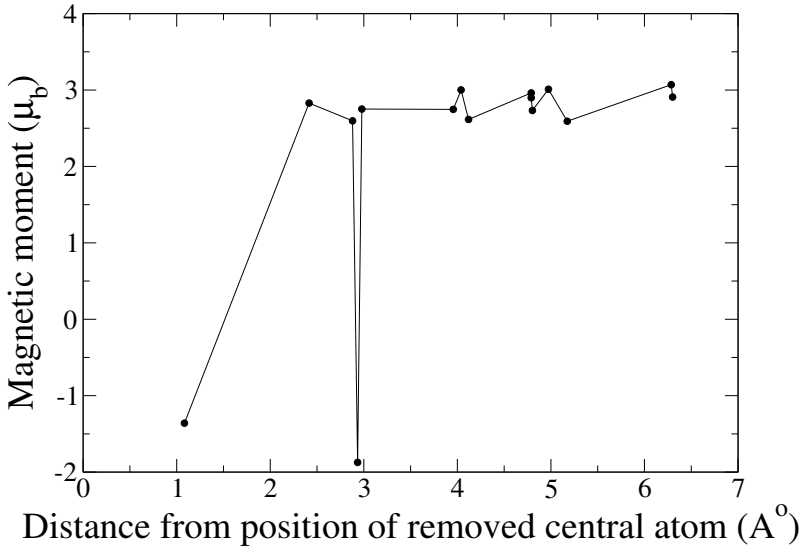
Assuming that $\frac{\partial E_c}{\partial a} = 0$, the final relationship required is given by eqn 4.11

$$\frac{\partial^2 E_c}{\partial a^2} = 9aB \tag{4.11}$$

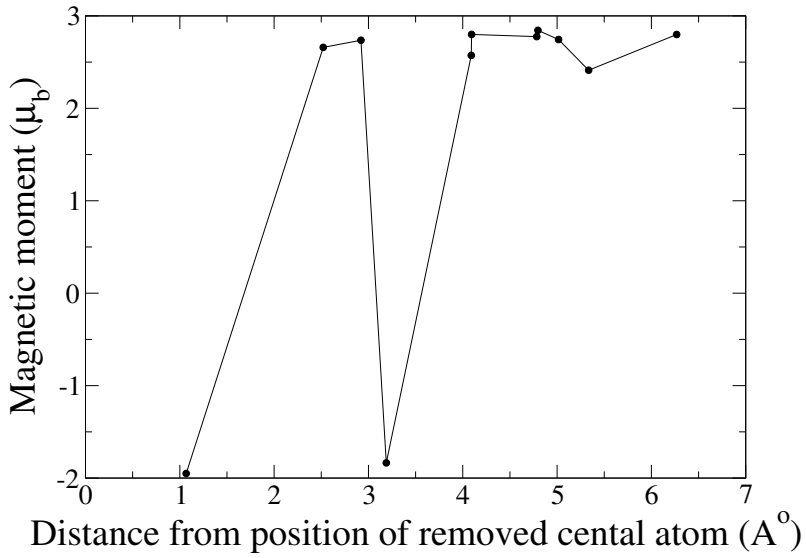
By including the factor $\frac{\partial^2 E_c}{\partial a^2}$ calculated from the value of the bulk modulus, it is found that the energy relaxes smoothly to a value within the specified tolerance for the change in the cell vector.



(a) ⟨100⟩ intersitial



(b) ⟨111⟩ interstitial



(c) ⟨110⟩ interstitial

Figure 21: Variation in magnetic moment with distance from each of the three dumbbell interstitials

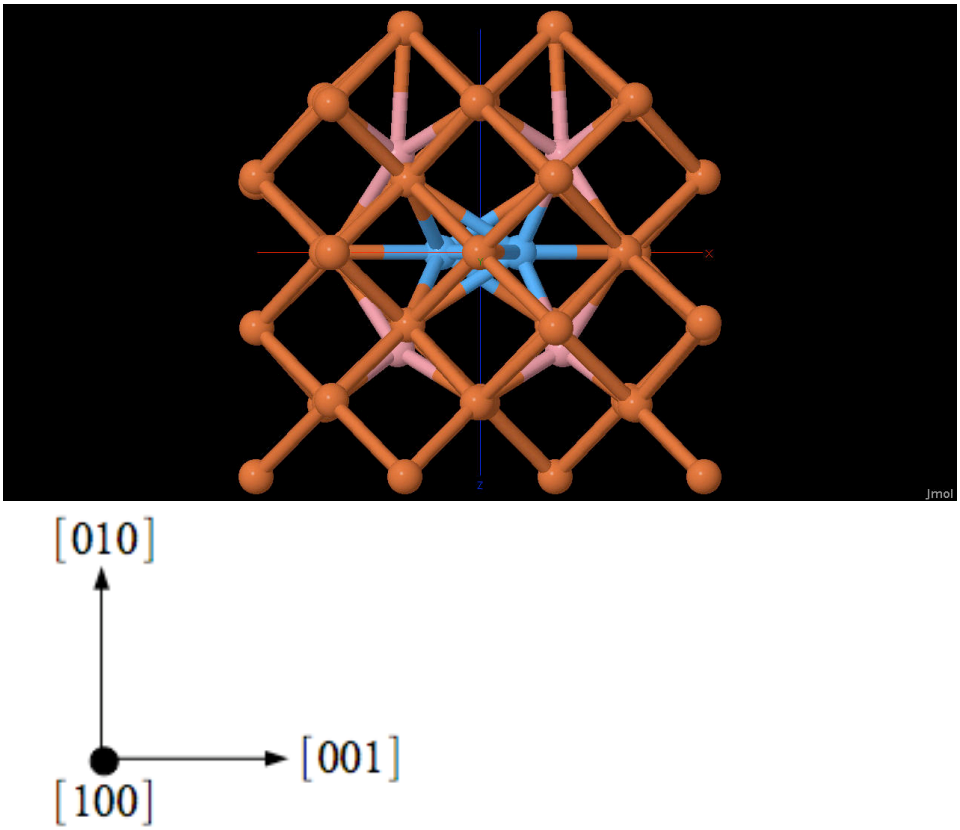


Figure 22: Jmol illustration of ⟨110⟩ interstitial. The dumbbell atoms are show in blue, the atoms corresponding to the sharp dip in spins corresponds to the pink atoms. All other atoms are shown in orange. The x, y and z axis form a right handed cartesian coordinate system.

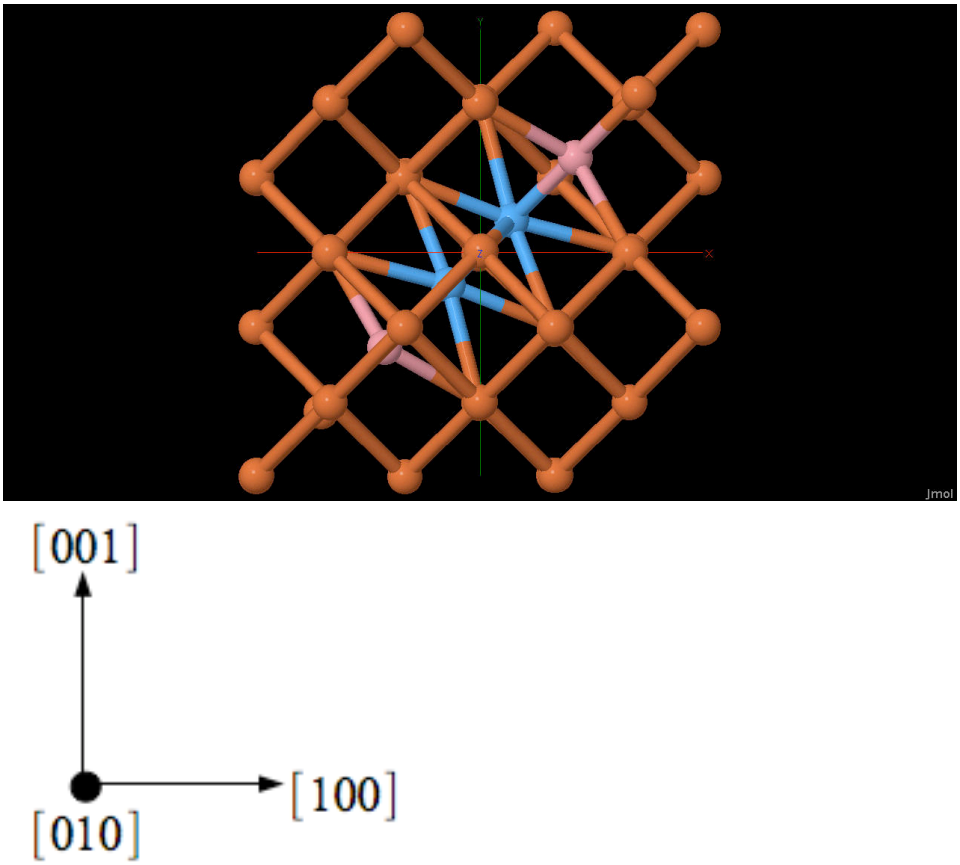


Figure 23: Jmol illustration of ⟨111⟩ interstitial. The dumbbell atoms are show in blue, the atoms corresponding to the sharp dip in spins corresponds to the pink atoms. All other atoms are shown in orange. The x, y and z axis form right handed cartesian coordinate system.

4.4 CONCLUSIONS

The Newton-Raphson scheme developed in Chapter 3 has successfully been applied to bulk calculations and point defects, providing results in good agreement with both DFT and tight binding using LCN. The main difficulties encountered can be attributed to the 'roughness' of the energy surface with spin. The obvious next step is to consider line defects and in particular the $\frac{1}{2}\langle 111\rangle$ screw dislocation. The success of this method in dealing with static dislocations is discussed in the following chapter.

5

$\frac{1}{2}<111>$ SCREW DISLOCATIONS

5.1 SIMULATING THE $\frac{1}{2}<111>$ SCREW DISLOCATION IN BCC IRON

The first task is to establish the geometry required for the simulation. The unit cell vectors (x, y, z) orientated with the [100], [010] and [001] directions used in the previous chapter on point defects are not conducive to an investigation of the $\frac{1}{2}\langle 111\rangle$ screw dislocation. The Burgers vector in this case is $\frac{1}{2}a\,[111]$, where a is the lattice parameter of iron. It therefore makes more sense to define the z axis along the direction of the Burgers vector. The axes must therefore be rotated to another orthogonal system (x', y', z') along the $[11\bar{2}]$, $[1\bar{1}0]$ and [111] directions. The basis of the bcc unit cell in this new coordinate system is found by applying the condition given in eqn 5.1.

$$(x, y, z) = x'(11\bar{2}) + y'(1\bar{1}0) + z'(111) \tag{5.1}$$

The positions of the atoms in this new coordinate system are:

x'	y'	z'
0	0	0
0	0	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$
$\frac{1}{6}$	$\frac{1}{2}$	$\frac{5}{6}$
$\frac{2}{3}$	0	$\frac{1}{3}$
$\frac{2}{3}$	0	$\frac{5}{6}$
$\frac{1}{3}$	0	$\frac{1}{6}$
$\frac{1}{3}$	0	$\frac{2}{3}$
$\frac{5}{6}$	$\frac{1}{2}$	$\frac{1}{6}$
$\frac{5}{6}$	$\frac{1}{2}$	$\frac{2}{3}$

All coordinates are multiplied by the lattice vector a. It is clear that there are six values of the coordinate z' along the [111] direction:

$0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}$. It can also be seen that there are always two atomic positions with the same x' and y' coordinates, whose z' coordinates differ by $\frac{1}{2}$. This results in a repeated layer half way along the [111] direction. It is therefore only necessary to consider one of these layers and consider the z' coordinates of the atoms in this layer as fractions of $\frac{1}{2}$ [111] axis, which are $0, \frac{1}{3}, \frac{2}{3}$. Looking along this direction the view is of a hexagonal arrangement of triangles as illustrated in Figure 24. The three values of z' are differentiated by colour .

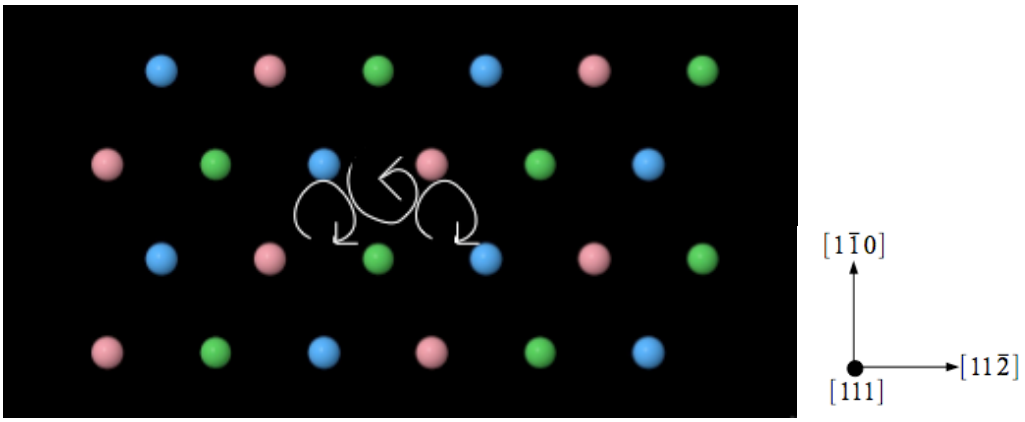
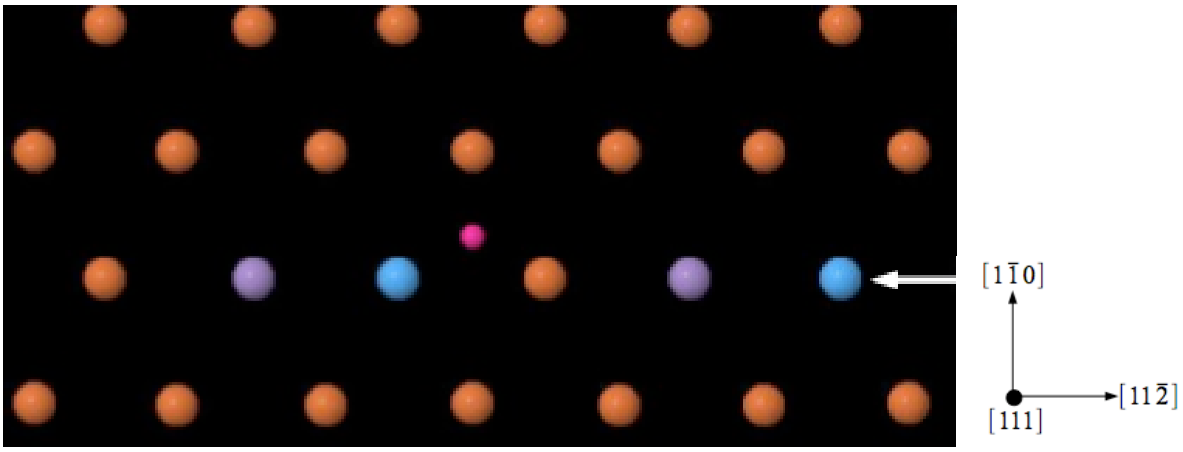


Figure 24: [111] view of bcc iron. The pink atoms correspond to a coordinate of 0 along the $\frac{1}{2}$ [111] axis, the blue to a coordinate of $\frac{1}{3}$ and the green to a coordinate of $\frac{2}{3}$. The white arrows indicate the direction of spiral of the atoms in adjacent triangles. A clockwise arrow indicates spiralling out of the page and an anticlockwise arrow spiralling into the page.

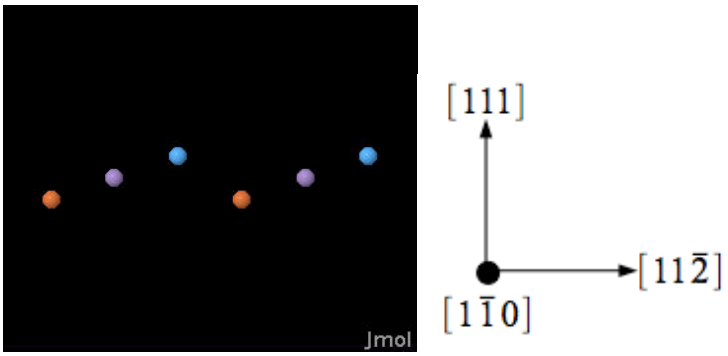
The next issue to resolve is where to insert the dislocation. This is placed in the exact centre of one of these triangles so as to minimise interaction between the dislocation and the surrounding three columns of atoms [27]. In this work a dislocation is introduced using the isotropic solution expressed $u_{z'} = \frac{b\theta}{2\pi}$ where $u_{z'}$ is the displacement in the [111] direction, b is the magnitude of the Burgers vector and θ is the angle between a general point and an arbitrarily defined line from the centre of the dislocation to a point in the $x'y'$plane. The full derivation of this was given in Chapter 2.

5.2 EASY AND HARD CORE DISLOCATIONS

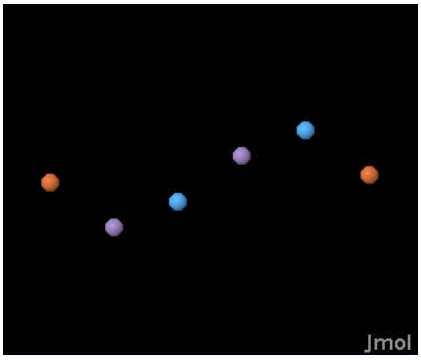
There are two possible types of dislocation core that can result when dislocation is inserted into the centre of a triangle [26]. The atoms which are the vertices of the triangles shown in Figure 24 can either spiral in to or out of the page, in the [111] direction. As the Burgers vector can be either parallel or antiparallel to the direction in which the atoms making up the triangle in the bulk spiral there are two alternative cores. These can be created by either keeping the centre of the triangle fixed and exchanging the sign of the Burgers vector or by holding the sign of the Burgers vector fixed and moving the centre of the dislocation to an adjacent triangle in which the atoms spiral in the opposite direction. The two resulting kinds of dislocation core are referred to as easy and hard [98]. The easy is consistently lower in energy. Geometrically the stacking of the atoms close to the core is different and can be understood by considering the profile of the core in the $[1\bar{1}0]$ direction. For clarity just one row of atoms along the $[11\bar{2}]$ direction is shown (marked by the white arrow in Figure 25a). The centre of the dislocation is indicated by the bright pink dot. Beginning with the bulk in Figure 25b the repeating pattern of three atoms with three different values of the [111] coordinate is apparent. However, introducing the screw dislocation results in two very different stacking sequences close to the dislocation core.



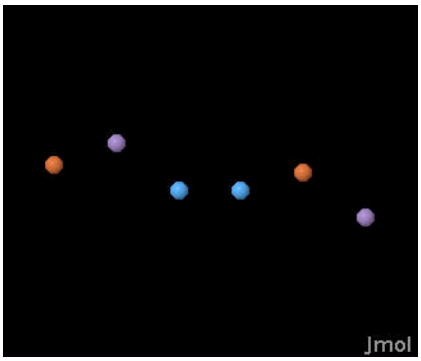
(a) View in the [111] direction of bulk bcc iron. Bright pink dot marks the centre of the dislocation. White arrow: row of interest.



(b) Bulk stacking sequence **abcabc**



(c) Easy core stacking sequence **abcbca**



(d) Hard core stacking sequence **abccab**

Figure 25: Bulk, easy and hard core dislocations profile along the [1$\bar{1}$0] direction.

5.3 IS THE EASY CORE STRUCTURE DEGENERATE OR NON-DEGENERATE?

The answer to this question provides the most crucial test of the ability of the tight binding method under consideration to model screw dislocations. For many years the conventional wisdom was that the easy core of bcc metal atoms did exhibit degeneracy. This was used as an explanation of the high degree of immobility of the screw dislocations in bcc metals [26] and was supported by studies [98, 99] using different empirical potentials. However, more recent density functional theory studies [36, 52, 97, 96] have dispelled this notion and showed that the core structure is in fact non-degenerate with no preferential spreading in the $\langle 112 \rangle$ directions in the $\{110\}$ planes. Recent bond order potential (BOP) studies agree with DFT about the non-degenerate nature of the core [65, 70]. This work is significant as it allows relaxation to occur, which the early DFT studies did not. The idea of degeneracy can be visualised using a differential displacement diagram (DD diagram). A DD diagram essentially plots the displacement in the [111] direction relative to the bulk structure and projects this onto the plane perpendicular to this. That is the length of the arrows between atoms is proportional to the extra displacement between the atoms in the [111] direction. If an arrow between two atoms covers the complete distance between them, then by normalisation the displacement between the atoms is $\frac{\vec{b}}{3}$ in the [111] direction. When a dislocation is introduced into the geometric centre of one of the triangles according to the elastic solution, the atoms are displaced in such a way that the total displacement vector around the triangle is equal to the Burgers vector, each side of the triangle contributing $\frac{\vec{b}}{3}$.

The easy core structure obtained from calculations using interactomic potentials [3, 4, 60, 24, 19, 31] has a threefold rotational symmetry about the centre of its triangle in the [111] direction. The plane perpendicular to the [111] direction is defined by the $[1\bar{1}0]$ and $[11\bar{2}]$ directions. There are six possible symmetrically equivalent $\langle 112 \rangle$ directions along which the core could spread. In the case of the degenerate core only three

of the directions are selected. There are two possible ways in which this can occur. The three directions of spreading in the two cases are rotated by $\frac{\pi}{3}$ relative to each other as indicated in Figure 26. The DD plots for the two atomic configurations with the same formation energy are given in Figure 27. Comparison with the left panel of Figure 27 and the bold lines in Figure 26 and the right panel of Figure 27 and the dashed lines in Figure 26 indicate how the spreading occurs. However, using more accurate ab initio techniques it is possible to observe that spreading does not occur on any of the equivalent $\langle 112 \rangle$ directions. It is this core which is obtained in the present tight binding calculations by relaxing the initial atomic configuration as illustrated in Figure 28, which is consistent with DFT.

The condition for degeneracy can be formulated more exactly in terms of the useful concept of gamma surfaces as given in inequality 5.2. This is known as the Duesbery-Vitek Criterion [26]. The notation $\gamma(1\bar{1}0)\left(\frac{b}{3}\right)$, refers to the energy at a point $\frac{b}{3}$ along the cross section of the $(1\bar{1}0)$ gamma surface in the [111] direction. The energy at $\frac{b}{3}$ and $\frac{b}{6}$ along this curve can be inserted into the condition to determine whether a certain core structure is degenerate. The basis for this idea comes from the differences in symmetry between the two types of core. The non-degenerate core has a six fold symmetry, whereas the degenerate core has a three fold symmetry. This means the displacement field in the [111] direction in the non-degenerate case is in six small steps in the displacement per Burgers vector and in the degenerate case in three larger steps.

$$\gamma(1\bar{1}0)\left(\frac{b}{3}\right) < 2\gamma(1\bar{1}0)\left(\frac{b}{6}\right) \tag{5.2}$$

It is first necessary to gain an understanding of the nature of gamma surfaces. A gamma surface is constructed by sliding two parallel planes over each other and noting how the energy changes as a function of the relative displacement. In practice, the two-dimensional energy landscape is investigated by calculating the energy at every point on a two-dimensional grid of relative displacements in the sliding plane. Figure 29 illustrates this idea in two dimensions. The angle between the cell vectors a_1 and a_2 is varied to create sliding. The gamma surface

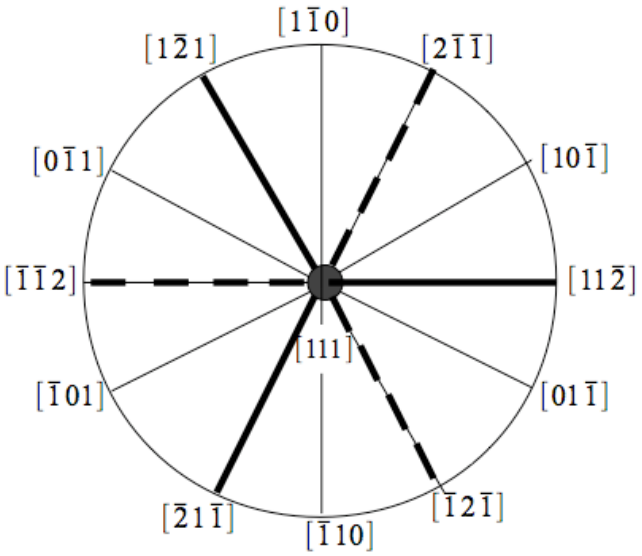


Figure 26: The equivalent $< 112 >$and $< 110 >$ directions in bcc crystals. Dashed lines represent one possible set of directions for spreading to occur and the bold lines the other.

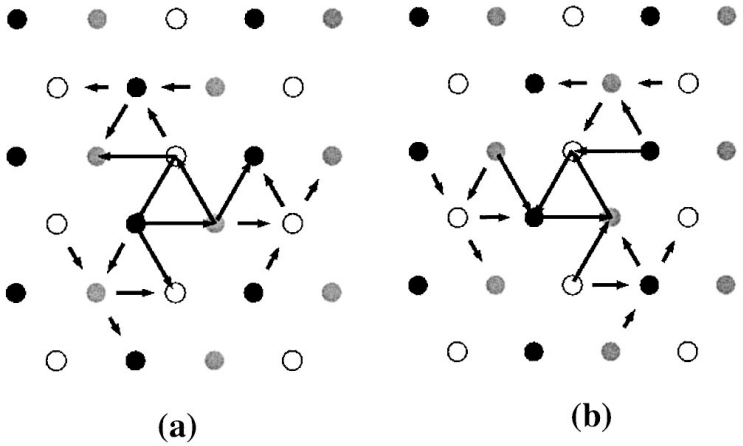


Figure 27: Degenerate core structure DD diagram Wang et al. [94]. The shades of gray represent the three displacements in the [111] direction of the bulk as described in section 3.1.

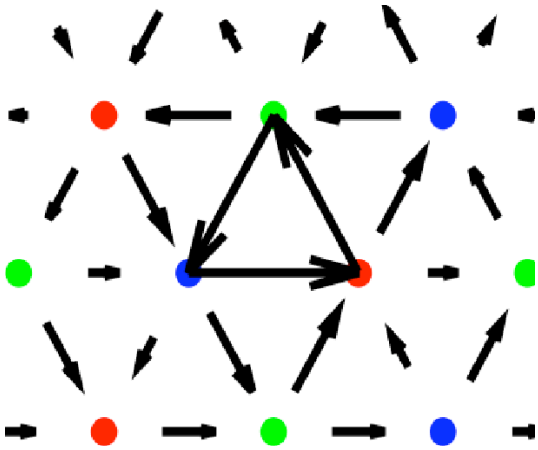


Figure 28: Easy Core (DD) plot for non-degenerate structure generated by simulations conducted in this work. The colours represent the three displacements in the [111] direction of the bulk as described in section 3.1. Axis are as in Figure 25.

sampling points are on a fine 16x16 grid. Each of the points on the grid is labelled by a pair of coordinates between 0 and 15, one for each of the two perpendicular axes of the surface.

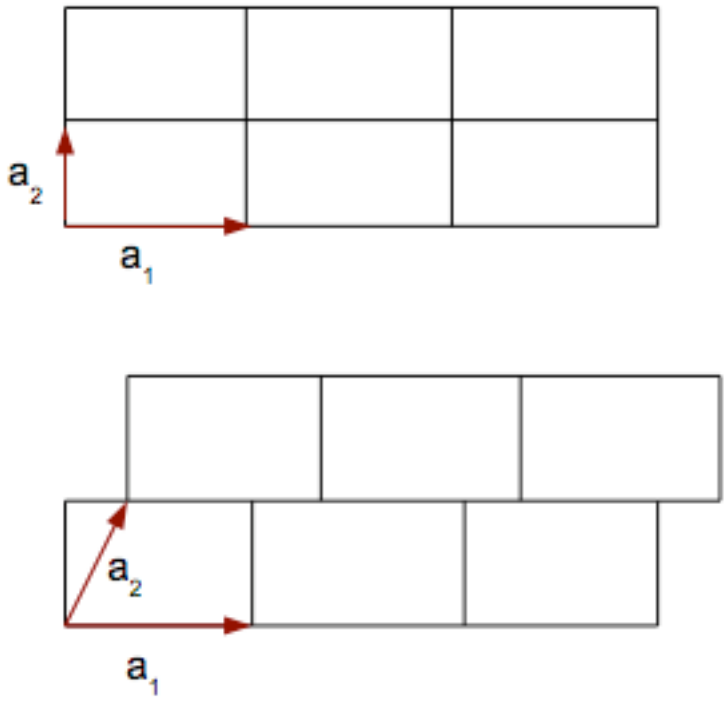


Figure 29: Schematic picture of 2D projection of cells sliding over each other to produce gamma surfaces

The number of k-points used in each direction is selected to be inversely proportional to the length of the cell vector in that direction. Cross sections of this surface along certain key directions can be obtained, and it is these that are referred to in eqn 5.2. This equation is concerned with the $(1\bar{1}0)$ gamma surface in the [111] direction. The values of the energy at $\frac{b}{3}$ and $\frac{b}{6}$ along this curve can be inserted into the condition to determine whether a certain core structure is degenerate. In the calculations conducted relaxation was allowed in the [111] direction only. The reason for the unrelaxed and relaxed energies being so similar along the [111] direction is unclear and could possibly be an artifact of the model.

The relaxed gamma surface cross section obtained from the tight binding model used indicates that the core is definitely non-degenerate

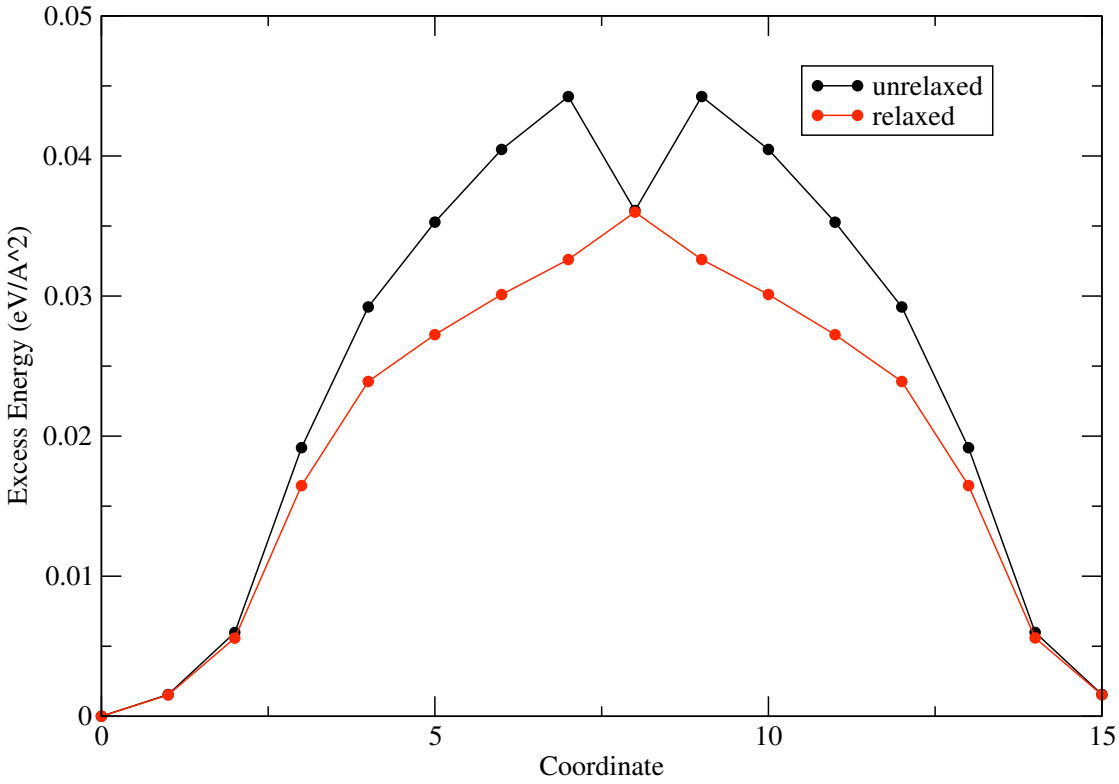


Figure 30: $(1\bar{1}0)$ gamma surface in the [111] direction

which is consistent with the DD plot. From Figure 30 the values of the excess energy in $(eV/\dot{A}^2)$ are:

$$0.013 \quad at \quad \frac{b}{6}$$
$$0.028 \quad at \quad \frac{b}{3}$$

Inserting these numbers into eqn 5.2, it is clear that the condition for degeneracy is not met.

5.4 THE HARD CORE STRUCTURE

This hard core dislocation has essentially the same symmetries as the non-degenerate easy core structure when relaxed. From Figure 31 obtained in this work it can be seen that the three arrows around the

centre of the dislocation point in the opposite direction to the easy core in Figure 28.

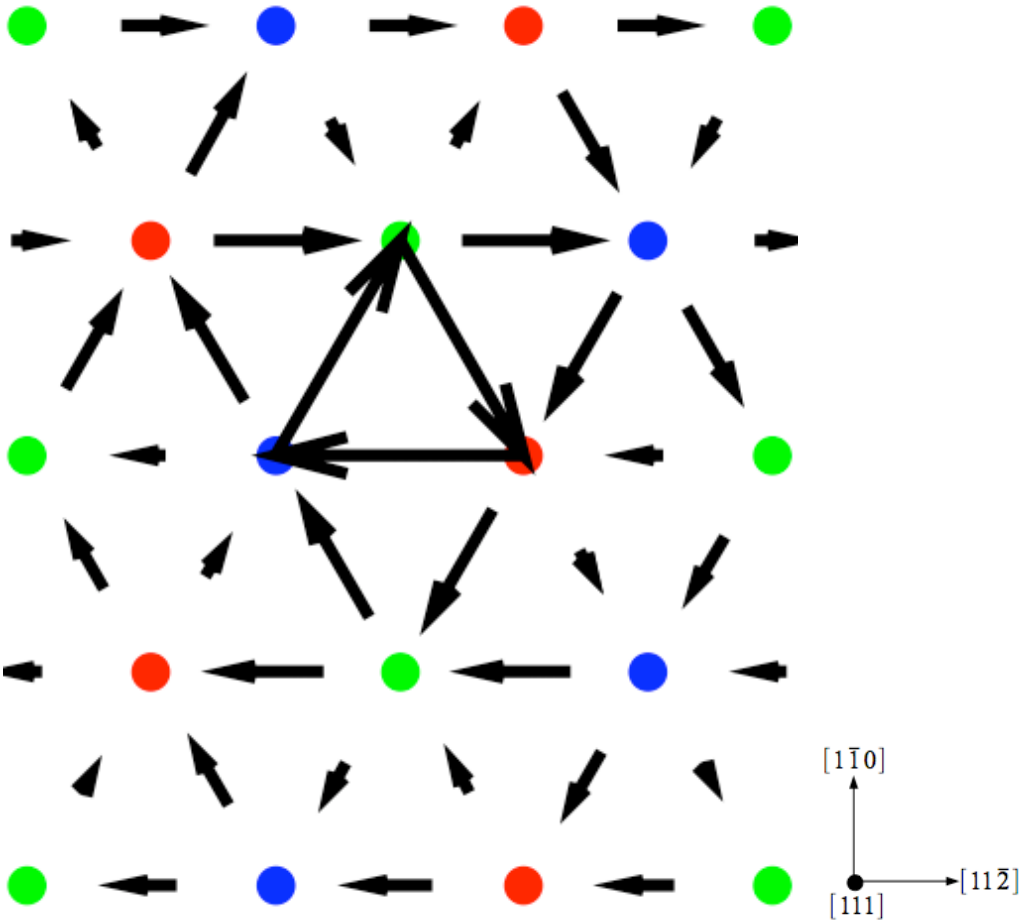


Figure 31: Hard Core differential displacement (DD) plot. The colours represent the three displacements in the [111] direction of the bulk as described in section 3.1.

5.5 FURTHER GAMMA SURFACE CALCULATIONS

Whilst the (1$\bar{1}$0) gamma surface is directly useful in establishing the non-degenerate state or the easy core screw dislocation, it is also interesting to consider other gamma surface calculations as they provide more information about the relaxed structure of dislocations. To begin with the simplest case of the (001) gamma surface is considered. The axes perpendicular to this surface are in the [100] and [010] directions and it is expected that there will be a maximum in the energy when the

two sliding surfaces of the cell are displaced half way along each of these axes relative to each other. At this point the atoms of the surfaces will be closest together. This is confirmed by Figure 32 below. The calculations are conducted by allowing the cell to relax, but constraining the nature of this relaxation to be perpendicular to the (001) plane. In this case only 8x8 sampling points are used as this is sufficient for the shape of the gamma surface to become apparent.

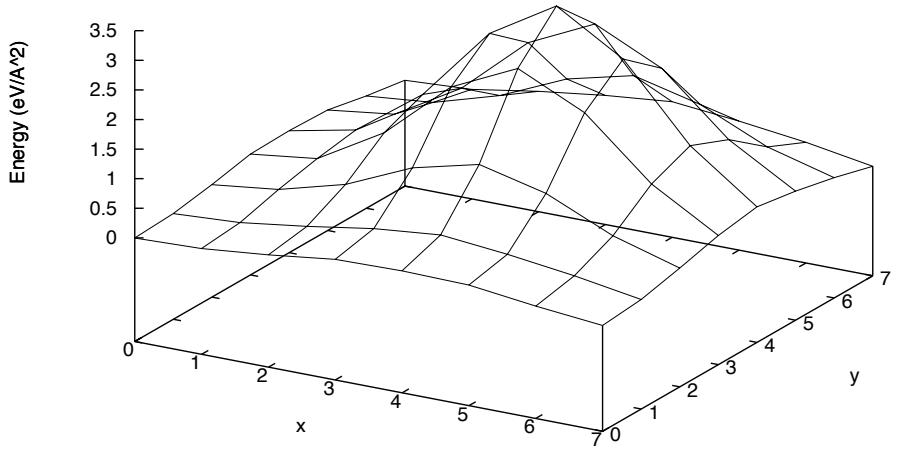


Figure 32: (100) gamma surface. The x axis in in the [010] direction and the y axis in the [001] direction.

It is necessary to check the validity of the k points sampling used by considering the effect on the initial energy per atom in various gamma surface calculations. Calculating this for the (001), $(1\bar{1}0)$ and $(11\bar{2})$ gamma surfaces it can be seen from Table 10 that there is an agreement of 3 d.p. The number of k points is chosen to be inversely proportional to the length of the cell vectors. In the case of the $(1\bar{1}0)$ gamma surface the k point mesh 15x10x9 for the [111], $[11\bar{2}]$ and $[1\bar{1}0]$ directions respectively is used. The number of k points is therefore sufficient at the initial calculation when the cell vectors are still all mutually perpendicular. However, the same notion of k points sampling does not apply as soon as one surface begins to slide over the other and the angles between the cell vectors begin to change. To be certain there are enough k points it is necessary to increase the density of the k point mesh and observe any significant change in the energy for all points in the mesh. Thus the number of k points in all directions for the $(1\bar{1}0)$ gamma surface is doubled to 30x20x18 and all energies agree to within 3 d.p. The k point sampling is therefore completely sound and puts all calculations on a solid foundation.

Gamma Surface	Total Energy (eV)	No. Atoms	Energy/Atom (eV) (3.d.p)
001	-17.78711	4	-4.447
110	-106.72966	24	-4.447
112	-106.72671	24	-4.447

Table 10: Initial energy per atom

5.5.1 $(1\bar{1}0)$ *Gamma surface*

So far the cross section in the [111] of the $(1\bar{1}0)$ gamma surface has been considered to ascertain that the easy core dislocation structure is non-degenerate. However, it is also interesting to plot the cross section for this surface in the $[11\bar{2}]$ to understand any differences in the behaviour as illustrated in Figure 33 for both the unrelaxed and relaxed

calculations. This energy is significantly different in the unrelaxed case but the overall shape of the curve is not altered.

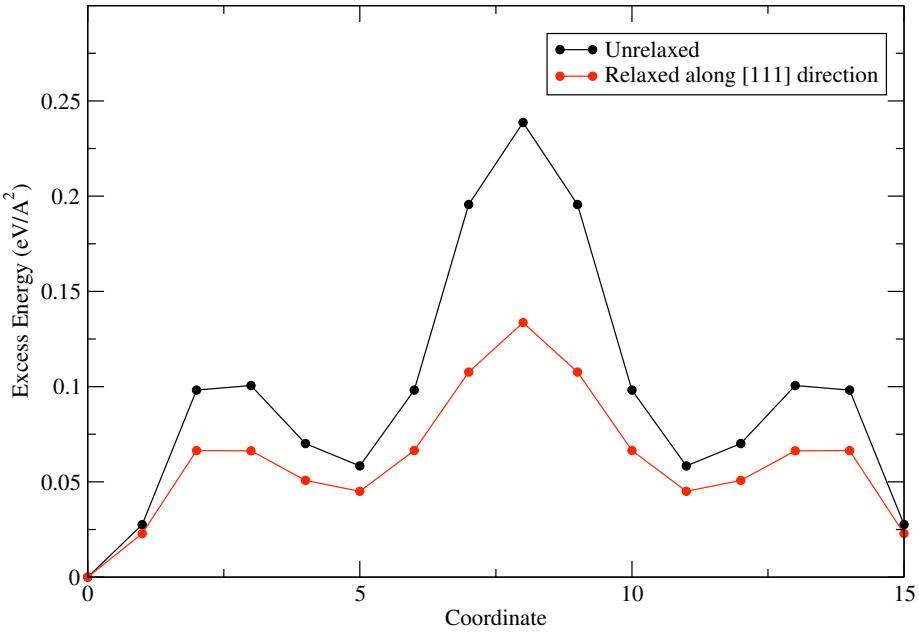


Figure 33: $(1\bar{1}0)$ gamma surface - $[11\bar{2}]$ direction

Explaining the shape of the cross section in the $[11\bar{2}]$ *direction of the* $(1\bar{1}0)$ *gamma surface*

It can be seen that there are minima either side of the expected central maximum. That is where the atoms on the two sliding planes are closest to each other. These minima are not observed in the cross section in the [111] direction illustrated in Figure 30. Looking in detail at the individual spins of all the atoms it was noticed that at each coordinate the atoms displayed one of two possible spins.

The basic cell used to create the gamma surface has four layers in the $[1\bar{1}0]$ direction. The atoms which have the magnetic moment S1 are on the surface of this cell, whilst the atoms with the magnetic moment S2 are the inner two layers of atoms. The variation of these magnetic moments for both the unrelaxed and the relaxed case are plotted below. Considering the unrelaxed case in Figure 34 it is observed that S2 remains more or less constant around the bulk value of $2.6\mu_b$, which make sense as the unrelaxed, inner atoms will display bulk properties.

In contrast the variation in the value of S1 is far more dramatic and in shape resembles a mirror image of the variation in the gamma curve energy in Figure 33. The atoms on the surface form fewer bonds, favouring magnetism, which could possibly explain why the energy curve and the plots of spins are mirror images of each other.

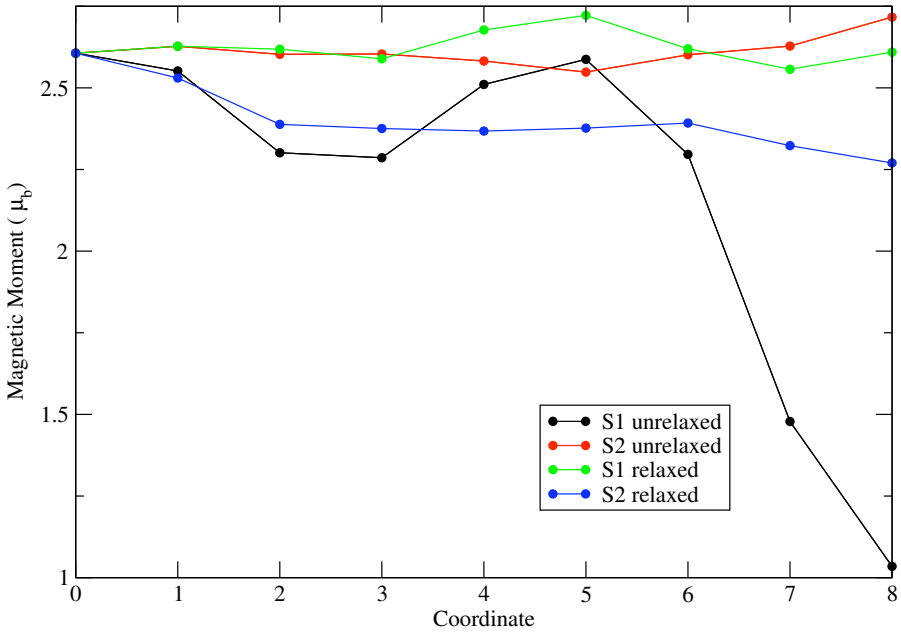


Figure 34: The magnetic moments across the gamma surface for an atom on the surface of the simulation cell (S1) and for an inner atom (S2) in both unrelaxed and relaxed cases. The coordinate labels the step in the gamma surface.

The explanation in the relaxed case is very different as the value of S1 is consistently greater than that of S2. There is also significant variation in the magnetic moment with configuration, but it is not as obvious a comparison to make as with the energy of the gamma surface. The shape of the magnetic moment curve for the surface atom in a cell that has been relaxed in Figure 34 (S1 relaxed) is possibly explained by considering bond length. The bonds the surface atom forms vary in length along the cross section of the gamma surface as illustrated in Figure 35. Bonds 1 and 2 refer to the two lengths of bonds that are made by a surface atom at each coordinate. However, as a surface atom will form a different number of bonds at different points along the curve, it is necessary to consider the average bond length as shown in Figure 35. It is a combination of these two effects that result in the S1 relaxed curve in Figure 34, with the peak at coordinate 5. Longer bond lengths

favour magnetism, which explains the similarities in shape between S1 relaxed curve in Figure 34 and the average bond length characteristic in Figure 35.

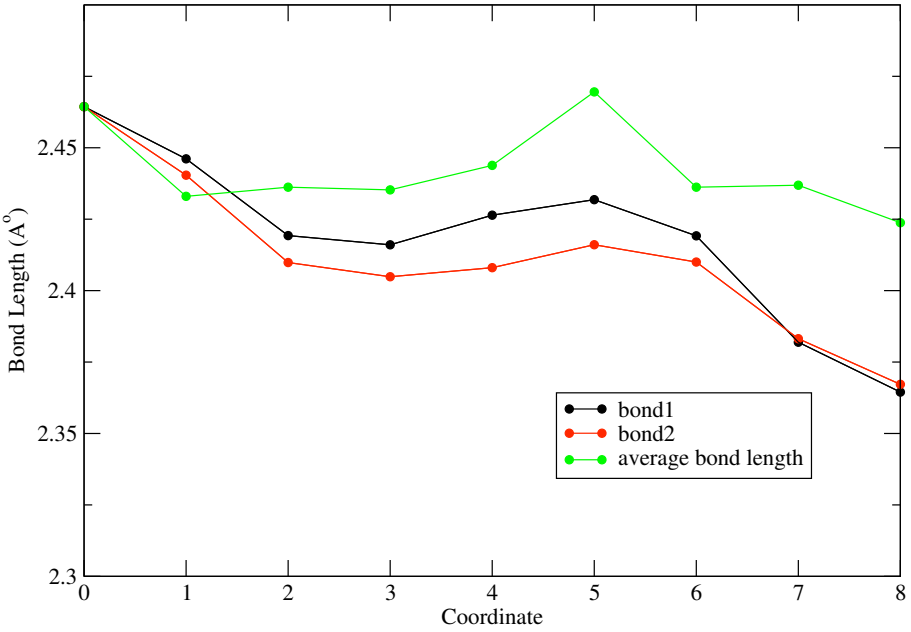
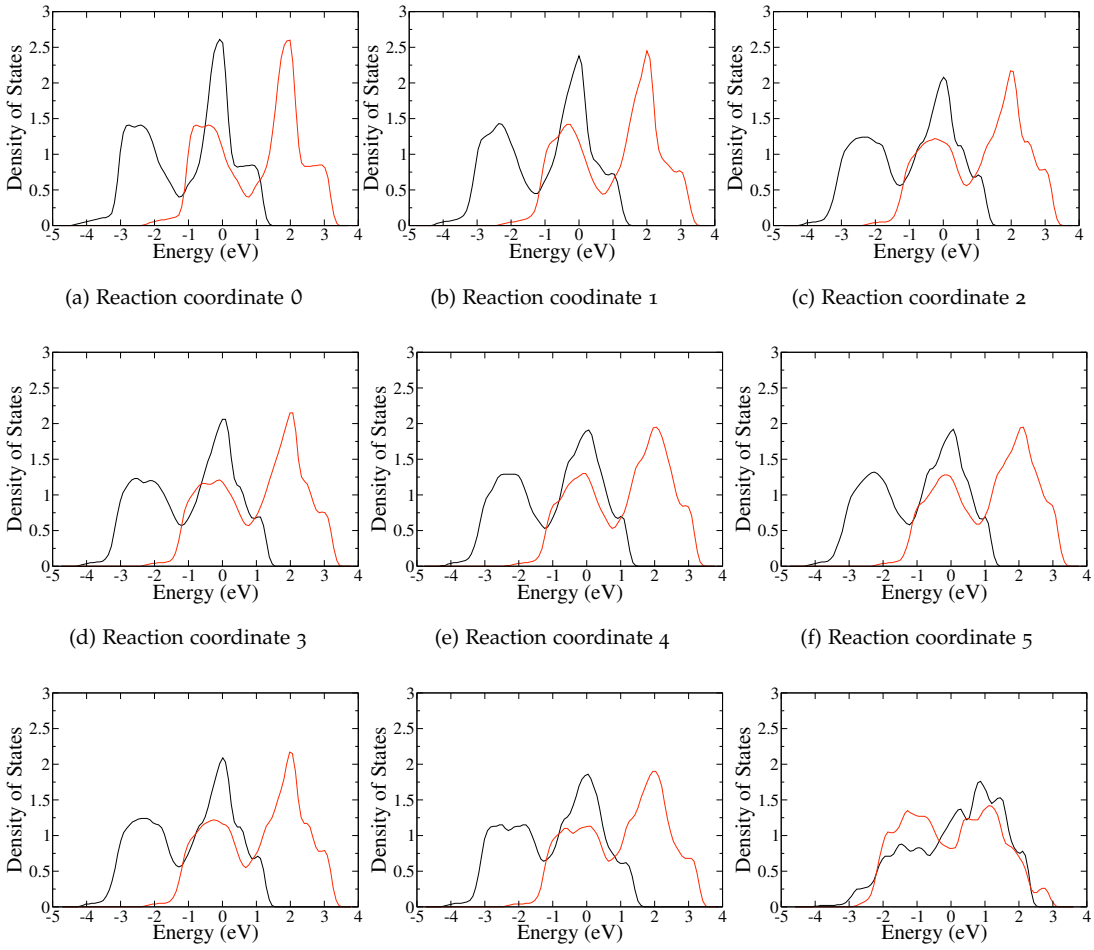


Figure 35: Variation of first nearest neighbour bond lengths and the averge of these bond lengths for a surface atom at each reaction coordinate of the fully relaxed $(1\bar{1}0)$ gamma surface in the $[11\bar{2}]$ direction. The coordinate labels the step in the gamma surface.

A possible explanation for the variation for the magnetic moment of an inner atom of the simulation cell as shown Figure 34 can be obtained by considering the local density of states at each reaction coordinate in Figure 36. Moving along the curve, it can be seen that the up and down densities of state overlap more and more, resulting in the magnetism dropping off. The change in the local densities of states is much less pronounced between coordinates 2 and 6, which is consistent with the flat section in the blue curve of Figure 34.

5.5.2 $(11\bar{2})$ *Gamma Surface*

The shape of the $(11\bar{2})$ gamma surface along the $[1\bar{1}0]$ direction is symmetric as indicated in Figure 37 and of a shape consistent with calculations carried out using both DFT [91] and the Mendelev empirical



(a) Reaction coordinate 0 (b) Reaction coodinate 1 (c) Reaction coordinate 2

(d) Reaction coordinate 3 (e) Reaction coordinate 4 (f) Reaction coordinate 5

(g) Reaction coordinate 6 (h) Reaction coordinate 7 (i) Reaction coordinate 8

Figure 36: Local density of states at each reaction coordinate of the fully relaxed $(1\bar{1}0)$ gamma surface in the $[11\bar{2}]$ direction for an inner atom of the simulation cell with spin S2. The red curve is the up spin density of states and the black curve the down spin density of states.

potential [64]. An interesting feature is the degree to which relaxation affects the magnitude of the energy across the gamma surface. This is made abundantly clear by Figure 37, and it can be seen that the maximum energy of the unrelaxed surface is roughly 15 times that of the relaxed surface which corresponds to significant movement of the atoms upon relaxation.

In order to view the shape of the relaxed gamma surface this is plotted separately in Figure 38. This has a broad peak shape and is again symmetric.

The situation in the [111] direction is slightly more complicated as illustrated in Figure 39. A certain amount of asymmetry is detected in

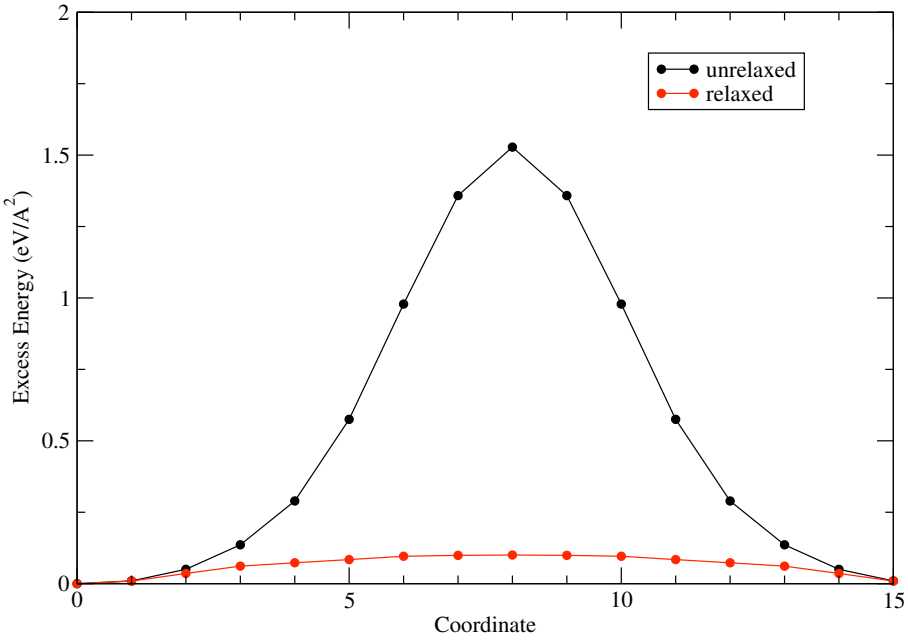


Figure 37: $(11\bar{2})$ gamma surface in the $[1\bar{1}0]$ direction

both the unrelaxed and relaxed cases, although it is more pronounced in the former. This is again consistent with published work on gamma surfaces conducted using both DFT [91] and the Mendelev empirial potential [64]. As with the $(1\bar{1}0)$ gamma surface in the [111] direction, at the central position along the [111] direction the unrelaxed and relaxed energies are very similar. This could just be a feature of the tight binding model used.

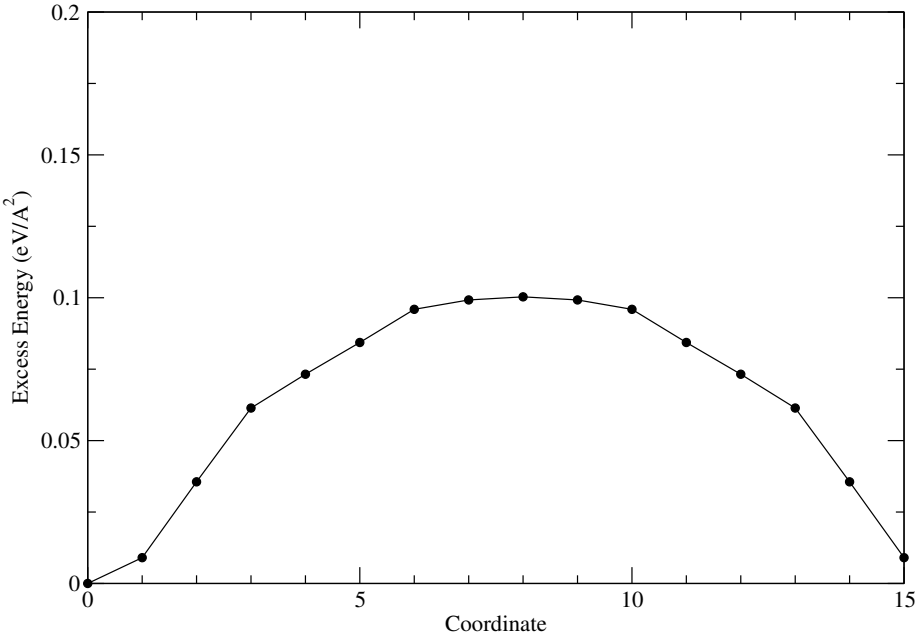


Figure 38: $(11\bar{2})$ gamma surface in the $[1\bar{1}0]$ direction relaxed

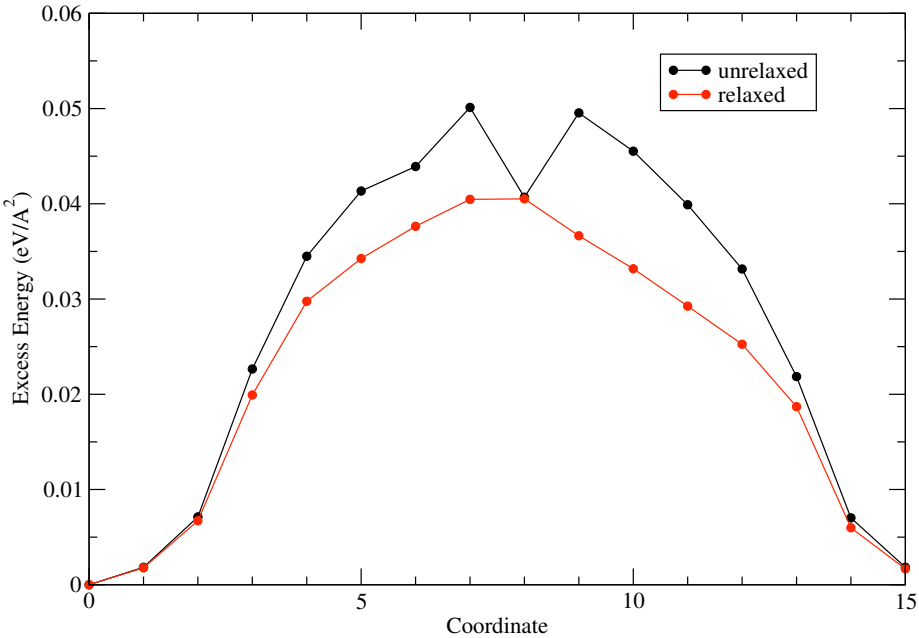


Figure 39: $(11\bar{2})$ gamma surface [111] direction

5.6 THE SIMULATION GEOMETRY

Cluster or periodic boundaries?

The first decision to make is between two clear approaches to the geometrical set up of dislocation calculations. The first is to employ periodic boundary conditions in all directions. The essence of this approach is to create multiple images of the dislocation [91, 15]. For geometrical reasons it is necessary to have two dislocations with opposite Burgers vectors in a dislocation dipole. These dipoles are then repeated periodically. The most evident advantage of this is that there are no surfaces of the dislocation to be concerned about. However, this is set against the need to work out the interactions between the dislocations in the dipole as well as the interactions between their periodic images. It is also necessary to have twice as many atoms than are going to be studied as there are two dislocations in the dipole. Extensions to this method involve using quadrupolar arrays in order to reduce image effects [80].

The second approach, and the one favoured in this work, is to use the cluster method [98, 96]. In this a dislocation is inserted down the centre of an isolated infinitely long cylinder of bulk bcc iron. The cell is periodic in one dimension along the dislocation direction. This introduces the problem of surface effects, where the atoms on the surface of the cylinder are not in the same environment as the atoms closer to the centre (i.e. the ones at surface have a lower coordination number). To avoid this problem a buffer zone of atoms around the central cylinder in introduced. All atoms in the configuration are initially displaced according to the isotropic elastic solution. The atoms in the buffer zone are then frozen throughout any calculations. Using a cluster set-up avoids the need to consider any dislocation interactions and therefore theoretically allows simulations with a larger number of atoms to be conducted. This means more complicated situations involving dislocations can be considered, for example, their interaction with both impurity atoms and other dislocations.

Which cross sectional shape to use?

To begin with, a dislocation is placed in a block of bulk bcc iron at the centre of a triangle whose points are columns of atoms in the [111] direction. To create a dislocation whisker along the [111] direction, the centre of the dislocation is treated as the centre of a cylinder cut out from an infinite solid. The need to freeze a ring of outer atoms, means that we require an outer and inner radius. The inner radius marks the boundary between the relaxed and frozen atoms whilst at the outer radius there is a free surface (Figure 40) . As the cross sectional area of the whisker is increased the difference between the inner and outer radii is kept fixed. Due to the geometry of the bcc structure it is not possible to achieve an exactly circular cross sectional area and there is a rough edge to the shape. The first attempts at simulation use these configurations and with most calculations we find that they work reasonably well, i.e surface effects are not significant. However, we find that with certain shapes it is not possible to achieve spin and charge convergence.

In later calculations the condition that all edges of the shapes are smooth was imposed, resulting in a hexagonal type shape. The position of the centre of the dislocation means that three of the edges will be one length and the other three another. For consistency, three layers of frozen atoms are used in all calculations as illustrated in Figure 41.

It is possible to isolate the energy of the relaxed atoms only. If the orbitals of the atoms in the relaxed region are labelled α or α' and those in the frozen region β or β' then their total energy can be written as eqn 5.3.

$$U = U_A + U_B \tag{5.3}$$

where U_A is the energy of the relaxed region and U_B that of the frozen region. An arbitrary decision to assign half of the cross over energy to the frozen atoms and half to the relaxed atoms is taken. The expressions for the energies of the individual atoms are hence given as

$$U_A = \rho_{\alpha\alpha'} H_{\alpha'\alpha} + \rho_{\alpha\beta} H_{\beta\alpha} + \frac{1}{2}\sum_{a a'} \phi_{aa} + \frac{1}{2}\sum_{ab} \phi_{ab} \tag{5.4}$$

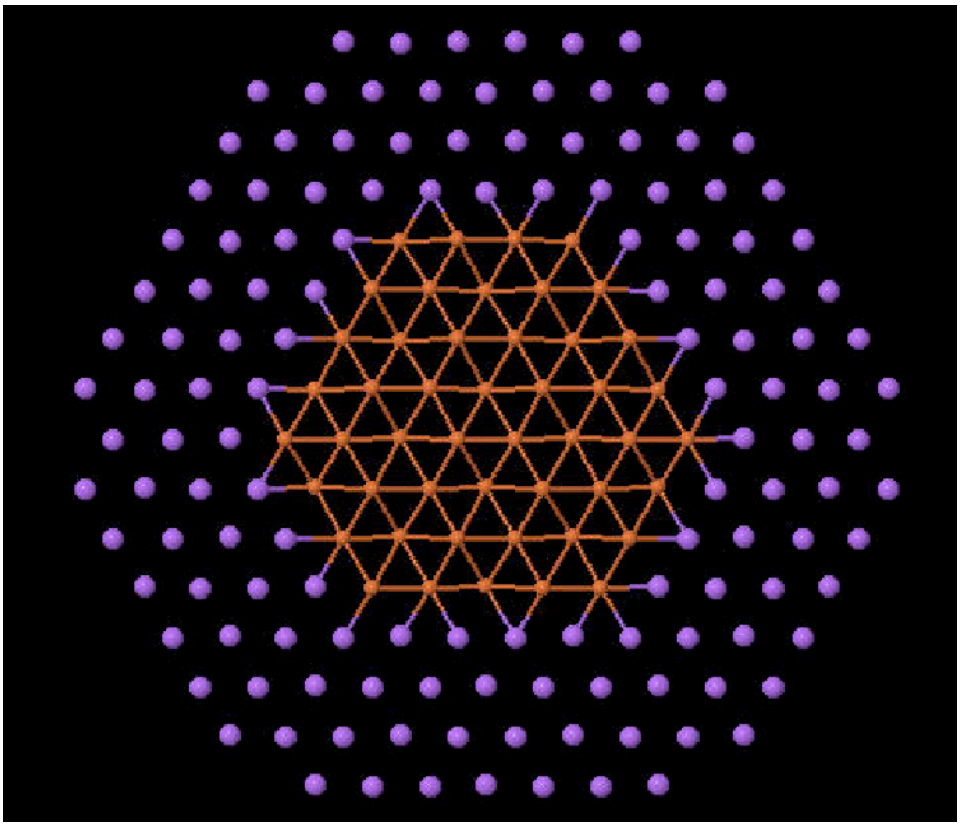


Figure 40: Cylindrical shape whisker. Orange atoms are allowed to relax, purple atoms are frozen.

$$U_B = \rho_{\beta\beta'} H_{\beta'\beta} + \rho_{\beta\alpha} H_{\alpha\beta} + \frac{1}{2} \sum_{bb'} \phi_{bb} + \frac{1}{2} \sum_{ab} \phi_{ab} \tag{5.5}$$

However, from Figures 42 and 43 it can be seen there is no noticeable improvement in the relationship between formation energy and whisker radius (which should vary linearly with the logarithm of the radius) and the difference in hard-easy core formation energy with whisker radius (which should level off with radius) when the energy of the unfrozen atoms only is considered, instead of all the atoms in the simulation. Therefore it was decided that the total energy will always be considered and calculations will be carried out with hexagonal shapes, with 3 layers of frozen atoms.

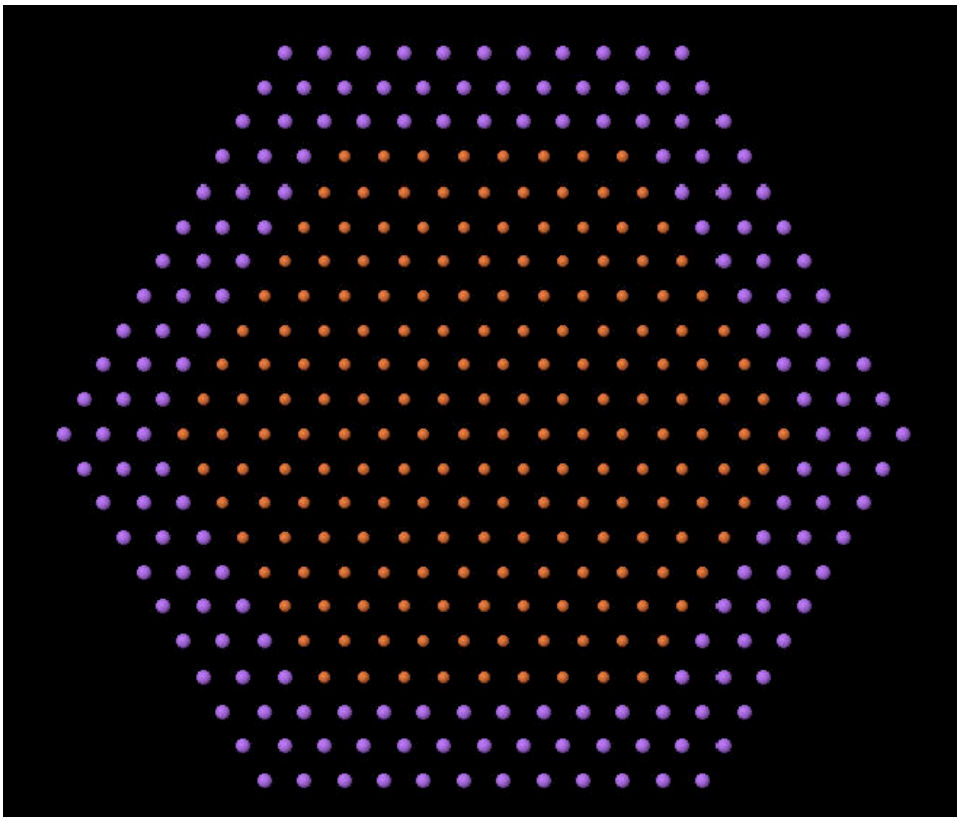


Figure 41: Hexagonal Cross sectional whisker. Orange atoms are allowed to relax, purple atoms are frozen.

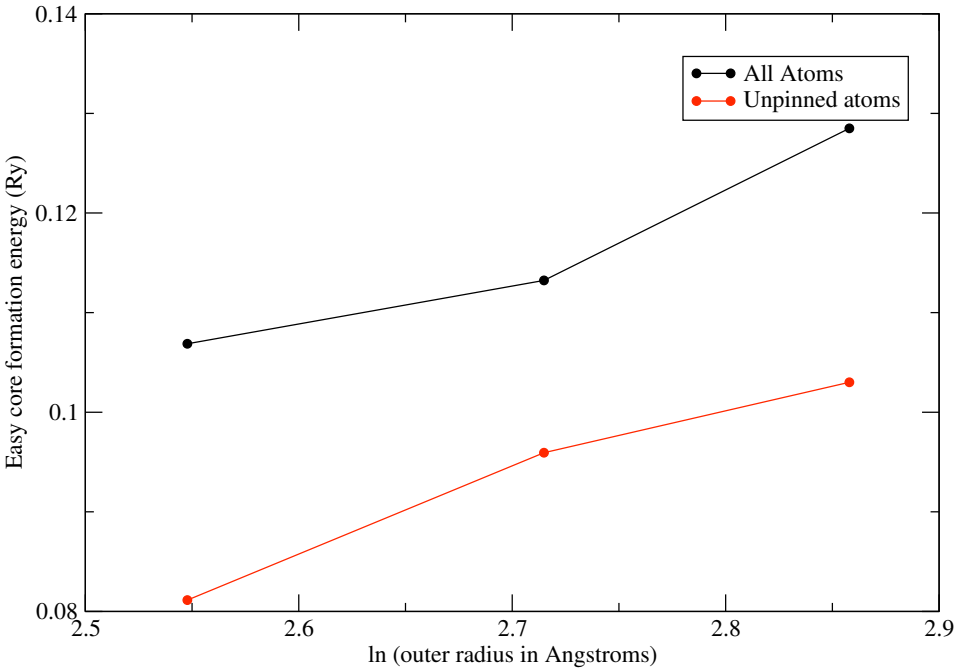


Figure 42: Variation in easy core formation energy with outer radius (defined as half the distance between two diagonally opposite corners)

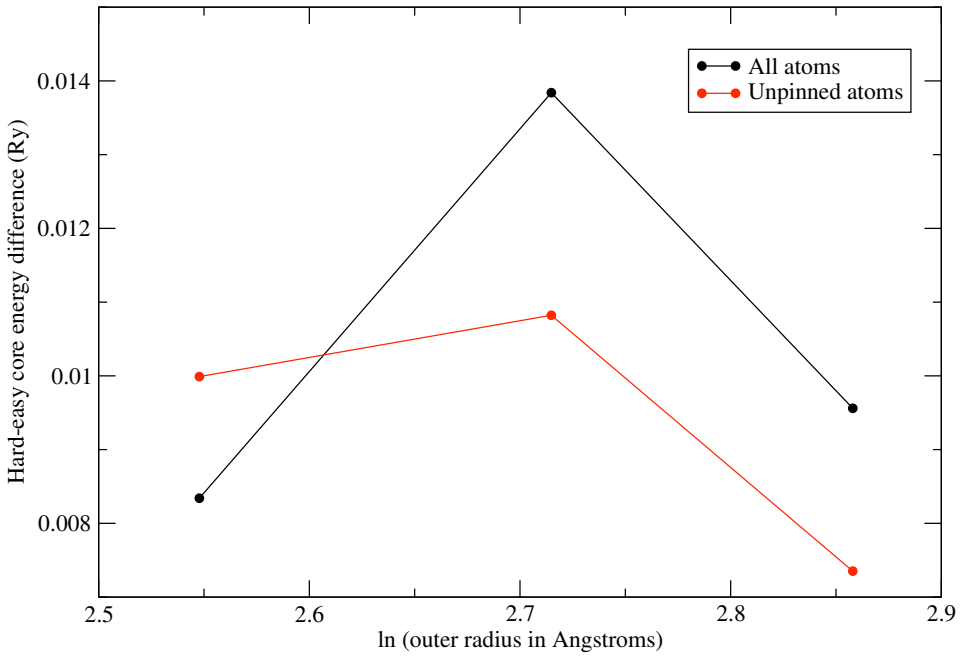


Figure 43: Variation in hard-easy core formation energy difference with outer radius (defined as half the distance between two diagonally opposite corners)

5.7 1-D ELECTROSTATICS

The whisker-like nature of the simulations of screw dislocations means that one dimensional electrostatics are very pertinent. This effect is included by using a cuboid simulation cell in which the radius of the cylinder is small in comparison to the lengths of the x and y axis. This cell can then either be repeated in the z direction only and treated using 1-D electrostatics or repeated in all directions, creating a 'toothbrush' like array in which the 'bristles' are the screw dislocation whiskers. These are far enough apart so that they do not interact with each other. The two approaches are compared later in the section.

It is first necessary to understand the scheme for including 1-D electrostatics. There is a unit cell of atoms indexed with I, where atom I has position $\vec{r}_I$ and net charge q_I. To illustrate the point a 1-D infinite array of Gaussian charges q_J and its interaction with a test charge q_I is considered. This is depicted in Figure 44.

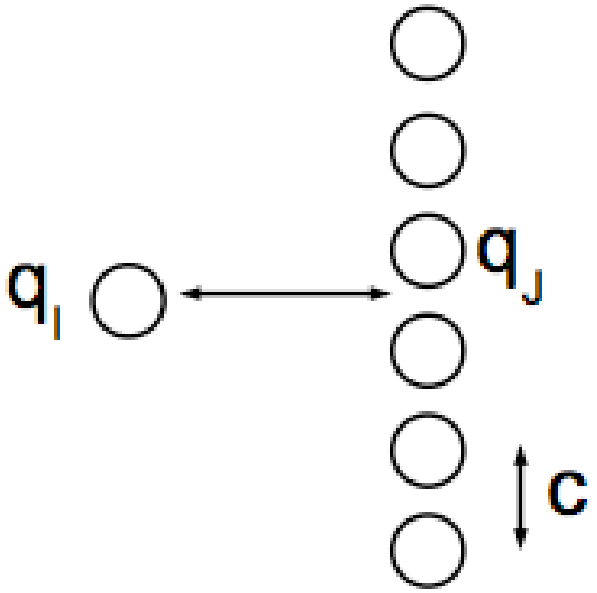


Figure 44: 1-D Boundary conditions

The form of the Gaussian charge density for an atom on site I at a position $\vec{r}_I$ is

$$\rho_I(\vec{r}) = q_I e \left(\frac{\alpha_I}{\pi}\right)^{3/2} \exp\left(-\alpha_I[\vec{r} - \vec{r}_I]^2\right) \tag{5.6}$$

The cells are repeated periodically along the z direction with period c. The net Coulomb interaction between the charge at site I and the charge at site J and all its periodic images is then

$$q_I q_J \gamma_{IJ} = \lim_{N\to\infty} \sum_{m=-N}^{N} \int d\vec{r}\, d\vec{r}\,' \frac{\rho_I(\vec{r})\rho_J(\vec{r}\,' - m\hat{e}_z)}{4\pi\epsilon_0 |\vec{r} - \vec{r}\,'|} \tag{5.7}$$

$$= q_I q_J \lim_{N\to\infty} \sum_{m=-N}^{N} g_{IJ}(\vec{r}_I - \vec{r}_J - mc\hat{e}_z) \tag{5.8}$$

where

$$g_{IJ}(\vec{R}) = \frac{e^2}{4\pi\epsilon_0}\left(\frac{\sqrt{\alpha_I\alpha_J}}{\pi}\right)^3 \int d\vec{r}\, d\vec{r}\,' \frac{\exp\left(-\alpha_I r^2\right)\exp\left(-\alpha_j[\vec{r}\,' + \vec{R}]^2\right)}{|\vec{r} - \vec{r}\,'|} \tag{5.9}$$

and $\vec{R} = \vec{r}_J - \vec{r}_I + mc\hat{e}_z$. Now, this sum diverges as $N \to \infty$. However, $\sum_J \gamma_{IJ}$ does not, provided the cells are charge neutral overall, and this is the quantity required. Initially it is fine to work with γ_{IJ} and consider the divergences later. The easiest way to perform the integral in Eq. 5.9 is to use Fourier transforms. The Gaussian functions are replaced by their Fourier transforms to give:

$$g_{IJ}(\vec{R}) = \frac{e^2}{4\pi\epsilon_0 R}\mathrm{erf}(\sqrt{\alpha_{IJ}}R) \tag{5.10}$$

This result, however, cannot be summed (see Eq. 5.8), so a different strategy is needed. The details of achieving this are arduous and so left out for simplicity. We use 1-D Ewald sums and follow a procedure similar to Smith [81].

Several tests using both Newton-Raphson and Pulay mixing with 1-D and 3-D boundary conditions are conducted. In some cases it was possible to reach self consistency and others it was not. The possible permutations and their success are presented in Table 11.

	On-site(3-D)	On-site (1-D)	Inter-site (3-D)	Inter-site (1-D)
Newton	Yes	Yes	No	No
Pulay	Yes	Yes	Yes	Yes

Table 11: Success of both Newton and Pulay Mixing with different boundary conditions. Yes = successful calculation

The tests were conducted on a particular cylindrical whisker containing 42 atoms. Using Pulay mixing convergence is achieved with both types of boundary conditions for the inter-site charge description. However, it is found that the 1-D boundary conditions are more efficient, as convergence is achieved in approximately a third of the number of self consistency (scf) loops required with 3-D boundary conditions. As well as this obvious technical advantage in favour of 1-D boundary conditions there is absolutely no ambiguity in the physical description of the interaction of charges in the whiskers of the simulations.

Newton Mixing is unfortunately not able to achieve convergence for the inter-site description for the cylindrical wire set-up as the energy escalates rapidly with each scf loop. The reason for this is not immediately obvious and to diagnose the cause of this a particular atom is selected and the change in its spin and charge charted as the calculation proceeds. The results of this are plotted in Figure 45.

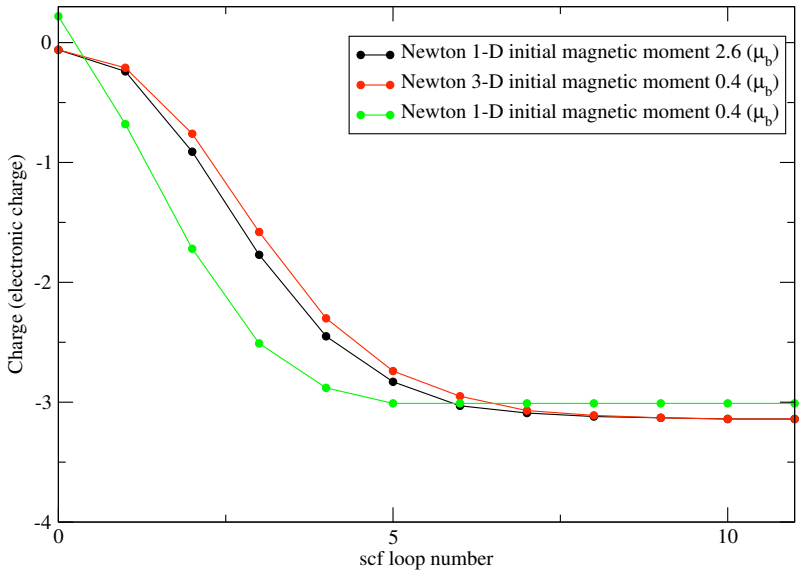


Figure 45: Change in the charge on an atom in a cylindrical whisker as inter-site calculation progresses

Both types of boundary conditions and two different values of initial spin are considered. In all cases the spin rapidly approaches zero and the size of the charge becomes very large and then levels off to a large negative value. Figures 45 and 46 comparing 1-D and 3-D boundary conditions show that there is a minimal difference in how the spin and charge on a particular atom progress with scf loops. Comparing two

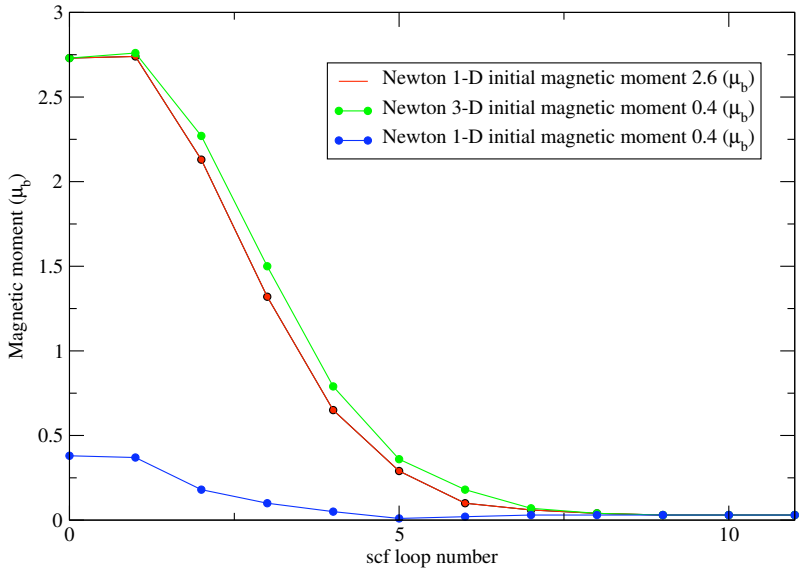


Figure 46: Change in magnetic moment on an atom in a cylindrical whisker as inter-site calculation progresses

different initial values of magnetic moment (2.6 and $0.4\mu_b$) using 1-D boundary conditions only, it can be seen that the general trends for spin and charge are the same and they reach their final value at roughly the same number of scf loops (i.e. at the same point in the calculation).

Further light is shed on why the 1-D Newton mixing has problems with convergence by considering the density of states. An 84 atom cylindrical wire is chosen for the simulations and the calculation is allowed to run until two scf loops have been completed. Two atoms are then selected; one close to the centre and one close to the surface of the wire and the density of states is projected on to each of these atoms. There is a marked difference between the shapes of the densities of states between the on-site and inter-site cases as illustrated in Figures 47 and 48. For the inter-site case, the surface and central atoms have very different densities of states indicating that using Newton mixing and 1-D boundary conditions does not give a realistic solution using the configuration in question.

Results from simulations carried out using the full inter-site description of charge are compared with the results of those using only the on-site terms. It is found that the energies and charges are very similar to each other. This is possibly because the parameters of the model employed [59] are all intended for calculations assuming local charge

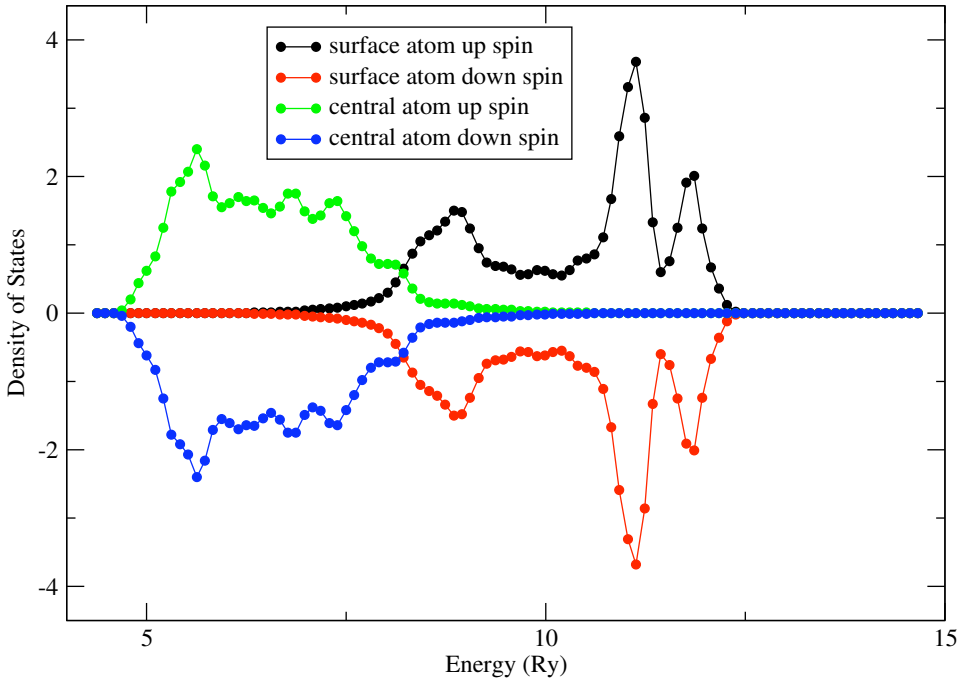


Figure 47: Inter-site local densites of states using 1-D boundary conditions

neutrality. It is also possible that on-site calculations might be more realistic as we have thrown away the s and p orbitals which move freely. These would screen the charge in the d orbitals, making it more like the on-site only calculations. Therefore in the majority of the simulations on dislocations a simpler on-site description of the charge is used.

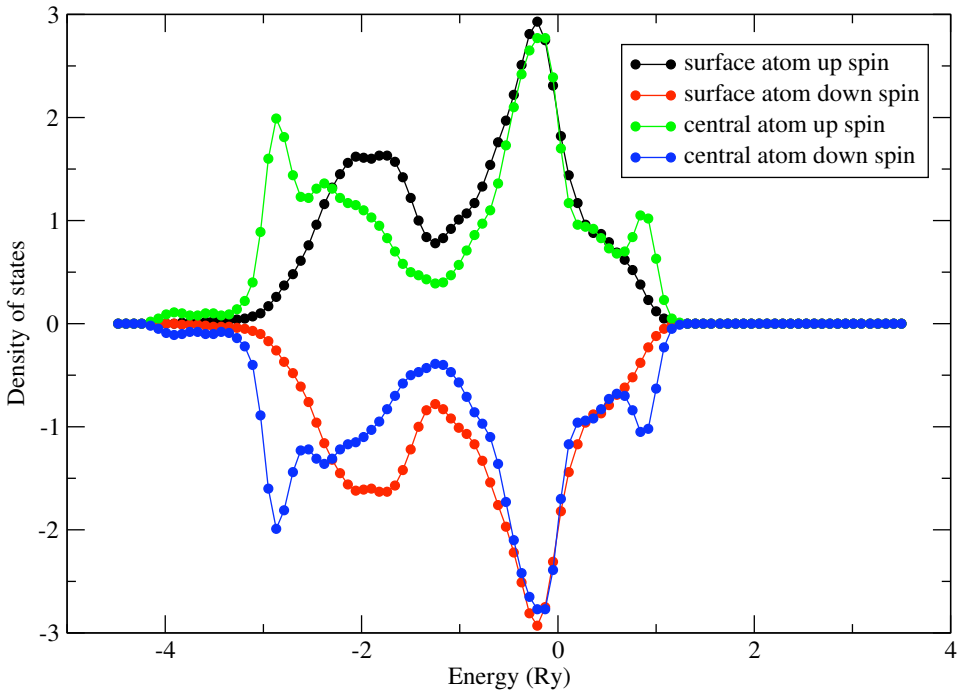


Figure 48: On-site local densities of states using 1-D boundary conditions

5.8 VARIATION IN SPINS AROUND DISLOCATION CENTRE

The spins around an easy core dislocation have three fold symmetry as can be seen from Figure 49, which matches the geometric symmetry. The difference in magnetic moments between the bulk and easy core dislocation for the coloured atoms are included in Table 12. Enough atoms were considered to make the pattern obvious. The results are similar to those recently obtained using bond order potentials [67]. There are small variations in the different values of spin as the distances between the atoms are similar to the perfect bcc structure, unlike for point defects which have much bigger variations.

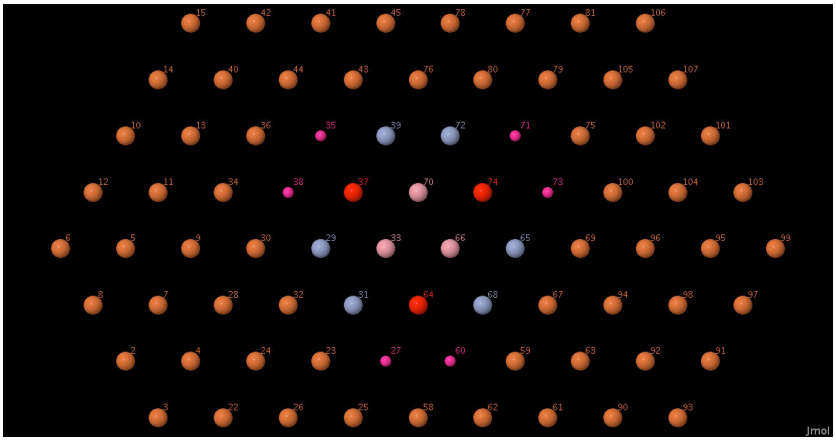


Figure 49: Spins around an easy core dislocation. Excluding orange atoms, all other atoms with the same colour have the same magnetic moment. Centre of dislocation in centre of triangle of pale pink atoms.

Colour of atom	Magnetic moment difference (bulk - dislocation) (μ_b)
pale pink	-0.03052
red	+0.08396
bright pink	-0.02194
blue	-0.09662

Table 12: Change in magnetic moments from bulk values around an easy core dislocation

5.9 VARIATION IN FORMATION ENERGY OF EASY CORE DISLOCATION WITH RADIUS OF WIRE

As hexagonal shapes are used it is first necessary to define what is meant by the radius. This is chosen to be half the distance between two atoms at diagonally opposite vertices of the shape. Using this definition an approximately linear relationship between the easy core formation energy and the logarithm of the wire radius results, in agreement with both DFT and empirical potential calculations [89, 17]. The formation energy is defined as the difference in energy between a bulk wire and the same whisker with a dislocation inserted through its centre. Both the unrelaxed and relaxed formation energies are plotted and as expected there is little difference in the two curves, except perhaps that the unrelaxed line is steeper. For completeness the resulting graph

using the cylindrical shapes is included. This is much higher in energy than that of the hexagonal shape and is also slightly more noisy. This linear relationship does not continue indefinitely with smaller and smaller radius and there is a cut-off radius as the limits of elasticity theory have been reached.

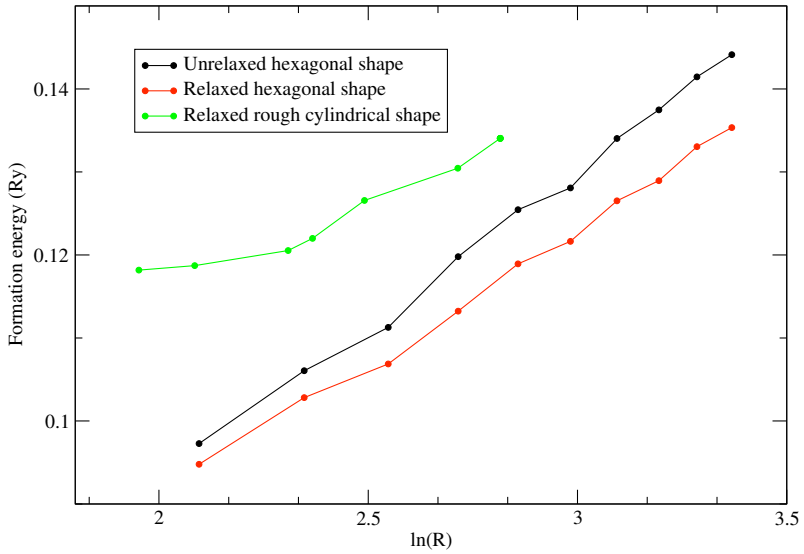


Figure 50: Variation in easy core dislocation energy with ln of outer radius (R) of hexagonal whisker

5.10 CONVERGENCE OF HARD AND EASY CORE ENERGIES

The hard core dislocation is consistently higher in energy than the easy core. The variation in the difference between hard core and easy core energy with dislocation radius is of interest. The expectation is that after a certain radius convergence is achieved and the value of the energy difference levels off [17]. This general trend is observed; however, there is a large amount of noise as can be observed in Figure 51. Ideally it would have been preferable to go to higher radii to observe true convergence but this was not possible due to constraints on computational time. A comparison is included between the cases where outer shell atoms were frozen and where they were not frozen. There does not seem to be an appreciable decrease in the noise present.

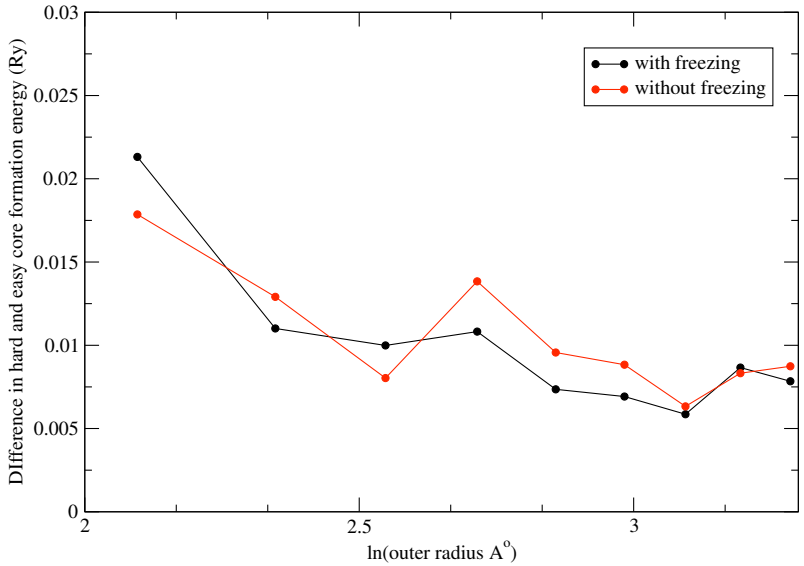


Figure 51: Hard/easy core energy convergence with ln of outer radius (R) of whisker

5.11 CONCLUSIONS

This chapter has been primarily concerned with static straight dislocations using the tight binding method. The correct geometry and the one-dimensional boundary conditions necessary for successful calculations have been established. The resulting non-degenerate easy core structure is consistent with existing density functional theory calculations and the relationship between the whisker radius and easy core formation energy and hard/easy core energy difference are as expected from similar studies using DFT.

6

PEIERLS BARRIERS

6.1 PEIERLS BARRIER CALCULATIONS

6.1.1 *What is a Peierls Barrier?*

Having established an efficient method for dealing with static dislocations in whiskers, the next logical step is to consider the motion of dislocations and the Peierls barrier, defined as the energy barrier along a path taken by a dislocation between two identical positions in a crystal. The dislocation itself moves as a straight line, and at each specified point along the path atomic relaxation is implemented to optimise the geometry. In this work the shape of the resulting energy profile is obtained using the 'generalised drag method'. Two adjacent easy core positions are used as the end points of the dislocation path and the dislocation passes through an intermediate hard core position along the $[11\bar{2}]$ direction. The position of the straight dislocation is simply moved along the line connecting the end points. Originally, the term Peierls barrier referred to the maximum stress along the stress versus displacement characteristic, as this indicates the applied stress required to initiate plastic flow.

In reality the motion of the $\frac{1}{2}\langle 111\rangle$ screw dislocation occurs via double kink propagation [21, 89, 92]; whereby a portion of the dislocation along the [111] direction moves from its stable easy core position to another position. The two edges of the kink then move away from each other dragging the entire dislocation over as described in Chapter 2. The Peierls barrier provides some insight into this process as it maps out the change in energy of a straight dislocation from its most stable position to the higher energy hard core position and gives an idea of the energy barrier the kink has to overcome.

6.1.2 *Initial Setup and the drag method*

The position vectors of the initial and final equivalent dislocation configurations are known. Linear interpolation is used to define the reaction coordinate λ, and to initialise the configurations along the pathway in 6.1.

$$\vec{R}(\lambda) = \vec{R_I} + \lambda\left(\vec{R_F} - \vec{R_I}\right) \tag{6.1}$$

$\vec{R_I}$ and $\vec{R_F}$ are the position vectors of the initial and final configurations of the dislocation. λ, the reaction coordinate, is a measure of how far the dislocation has been dragged between the initial and final configurations and $\vec{R}(\lambda)$ is the position vector corresponding to this. If the value of $\vec{R}(\lambda)$ is known then the value of λ can be determined as can be seen by rearranging eqn 6.1 to give the expression in eqn 6.2.

$$\lambda = \Lambda\left(\vec{R}\right) = \frac{\left(\vec{R} - \vec{R_I}\right).\left(\vec{R_F} - \vec{R_I}\right)}{\left(\vec{R_F} - \vec{R_I}\right).\left(\vec{R_F} - \vec{R_I}\right)} \tag{6.2}$$

At each value of λ it is necessary to obtain the minimum total energy E for each configuration of the atoms whilst keeping the reaction coordinate fixed. To this end Lagrange multiplier α is introduced and the quantity Ω is minimised.

$$\Omega = E + \alpha\left[\Lambda\left(\vec{R}\right) - \lambda\right] \tag{6.3}$$

Differentiating the expression for Ω with respect to the position vector $\vec{R}$, gives eqn 6.4

$$\begin{aligned} \frac{\partial\Omega}{\partial\vec{R}} &= \frac{\partial E}{\partial\vec{R}} + \alpha\frac{\partial\Lambda}{\partial\vec{R}} \\ &= -\vec{F} + \alpha\frac{\left(\vec{R_F} - \vec{R_I}\right)}{\left|\left(\vec{R_F} - \vec{R_I}\right)\right|^2} \end{aligned} \tag{6.4}$$

where the derivative of the total energy is related to $\vec{F}$, a Hellman-Feynman force. It should be noted that this expression had to be slightly modified in practice in the implementation in the PLATO code [54] to prevent translation of the atoms. It is now possible to set eqn 6.4 to zero

and rearrange to find the value of α, resulting in the expression in eqn 6.5.

$$\alpha = \left(\vec{R_F} - \vec{R_I}\right).\vec{F} \tag{6.5}$$

Substituting back into the expression for $\frac{\partial \Omega}{\partial \vec{R}}$ in eqn 6.4 gives

$$\frac{\partial \Omega}{\partial \vec{R}} = -\vec{F} + \frac{\left(\vec{R_F} - \vec{R_I}\right).\vec{F}\left(\vec{R_F} - \vec{R_I}\right)}{\left(\vec{R_F} - \vec{R_I}\right)^2} \tag{6.6}$$

The minimisation is carried out at a fixed value of λ, and so it is necessary that the changes in the positions of the atoms do not change the reaction coordinate. Some insight can be gained into how this condition is fulfilled by defining the normalised vector $\hat{\Delta}$.

$$\hat{\Delta} = \frac{\vec{\Delta}}{|\Delta|} = \frac{\left(\vec{R_F} - \vec{R_I}\right)}{\left[\left(\vec{R_F} - \vec{R_I}\right).\left(\vec{R_F} - \vec{R_I}\right)\right]^{0.5}} \tag{6.7}$$

Using the definition in eqn 6.7, the expression for the derivative of the energy is given in eqn 6.8.

$$-\frac{\partial \Omega}{\partial \vec{R}} = \vec{F} - (\hat{\Delta}.\vec{F})\hat{\Delta} \tag{6.8}$$

It can see from Eq. 6.8 that the component of force parallel to the reaction coordinate is subtracted from the total force, ensuring that the reaction coordinate remains fixed. In the following calculations a relative reaction coordinate is plotted; that is the initial and final reaction coordinates are labelled 0 and 1 respectively.

Which shape wire to use?

The aim is to obtain the Peierls barrier for a $\frac{1}{2}\langle 111\rangle$ screw dislocation in bcc ferromagnetic iron. It is now necessary to establish the shape of the whisker in which the dislocation will be inserted and the simulations carried out. The imperative constraint that must be fulfilled is that the initial and final positions of the path the dislocation takes give the same total energy for the whole whisker. Two possible shapes are considered. The first is an elliptical shape with completely smooth sides in which the positions of its focal points provide the starting and

end points for the dislocation. These focal points are at the centre of two triangles whose direction of spiral is the same. The Burgers vector is then chosen so that these two sites are both easy core in character. The equilibrium position of the dislocation should be in the geometric centre of the three columns of atoms in the [111] direction, which form the vertices of a triangle [93, 26]. The distance between the focal points is chosen such that there is only one triangle in between them. Naturally the atoms in this triangle spiral in the opposite direction resulting in a hard core dislocation site. The vertices of the easy core triangles are labelled by atoms (26, 28, 53) and (53, 55, 57) and those of the hard core triangle by (28, 53, 55) in Figure 52. By dragging the dislocation from one easy core site to another through a hard core site, it is expected that the shape of the Peierls barrier will be correctly established. The initial and final dislocation positions are energetically exactly equivalent as they are both geometrically equivalent. This can be observed by considering the symmetry of the wire. The ellipse configuration used contains 85 atoms in total.

The second shape considered is an exactly hexagonal shape as illustrated in Figure 53 and is made up of 169 atoms. The vertices of the easy core triangles are labelled by atoms (57, 61, 103) and (99, 103, 107) and those of the hard core triangle by (57, 99, 103). The difference in the resultant Peierls barriers between this configuration and the ellipse is of interest in indicating if the elongated nature of the ellipse has an unforeseen effect on their shape.

6.1.3 *The Shape of the Peierls Barrier*

The expected behavior of the shape of the Peierls barrier is that its minima correspond to the initial and final positions of the dislocation with a maximum at the central hard core position as produced by BOP (bond order potentials)[67] and DFT studies [91][80, 52]. Initial calculations of the Peierls barrier, however, indicate that the minima are shifted slightly away from the centre of the easy core triangle towards the hard core site for both the hexagonal and elliptical shapes. In the hexagonal case this is slightly closer to the hard core site than for the

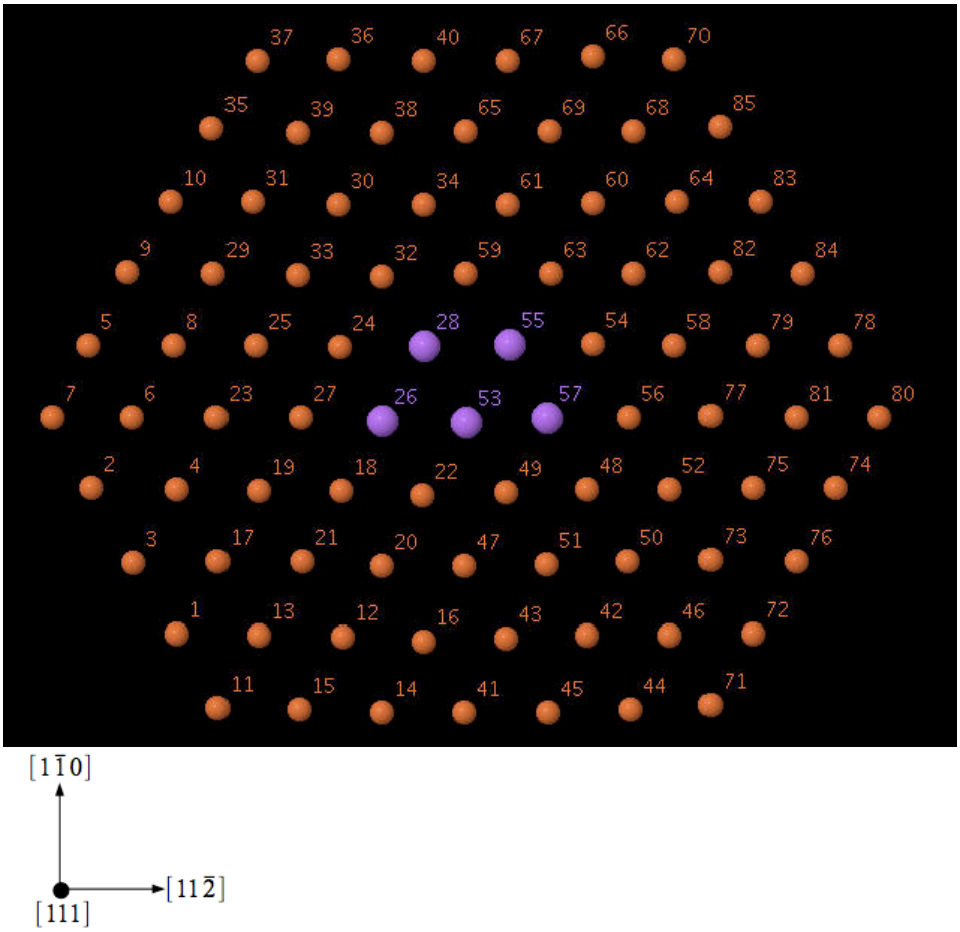


Figure 52: Jmol picture showing ellipse configuration. Dislocation is dragged between the centres of the left and right most purple triangles.

elliptical case. For both shapes there is a sharp peak in the energy when the dislocation is in the hard core position which has the highest energy as expected and there is an extremely shallow downward slope from the initial position of the dislocation, until just before the energy peaks. This is illustrated in Figure 54, the energy is given in meV/b, where b is the magnitude of the Burgers vector. One possible explanation of these is that the end points are not sufficiently minimised. This is discussed further in the section on minimised barriers.

6.1.4 *DD plots - a closer look at what is happening*

A closer look at what is happening using differential displacement plots (DD plots) can provide some insight into what is happening. Figures 55 to 59, which show the motion of a dislocation in an elliptical whisker, are labeled from 0 to 4, corresponding to the 5 points sampled

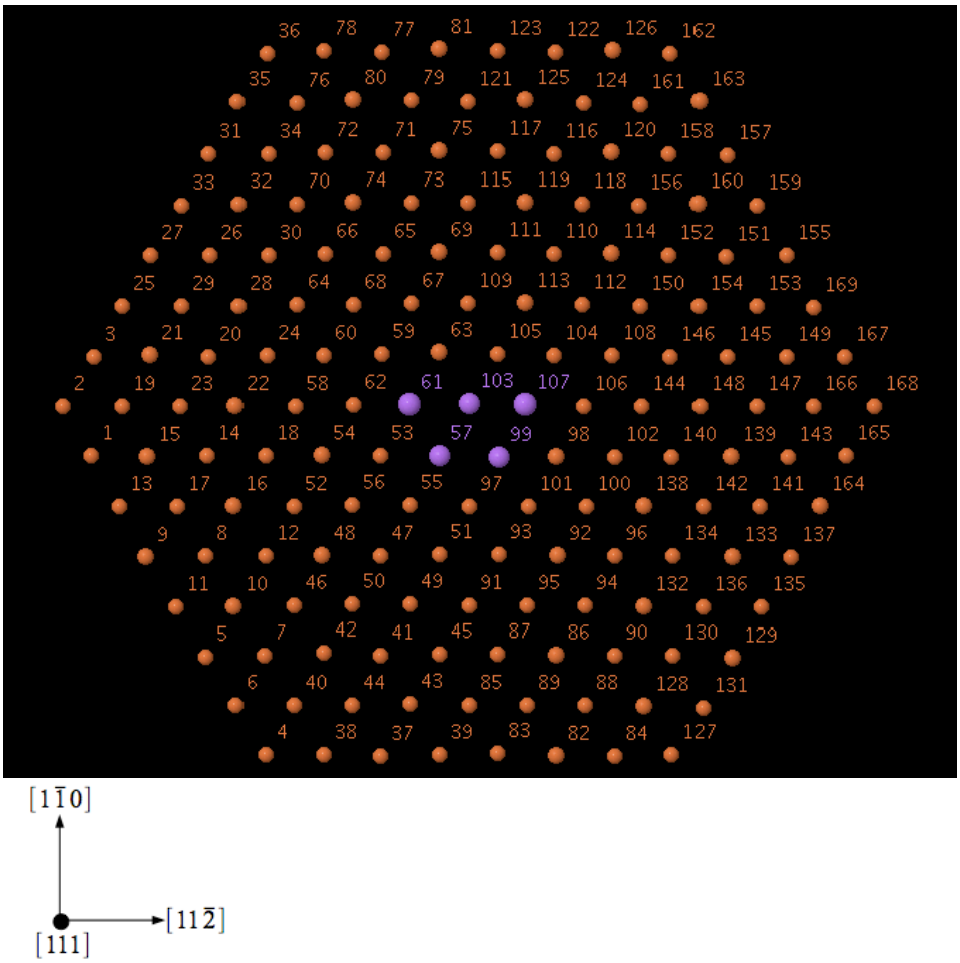


Figure 53: Jmol picture showing hexagon configuration. Dislocation is dragged between the centres of the left and right most purple triangles

between the easy- and hard-core positions. Configuration 0 is one endpoint and configuration 4 is the central hard core position. From Figure 55 (configuration 0), the dislocation has been placed in the exact centre of a triangle, which is expected to be the minimum energy configuration but turns out not to be. The direction of spiral indicates whether the dislocation is easy or hard core in character. There is very little difference between Figures 56, 57 and 58. All that can be seen is that the position of the easy core dislocation moves further to the left. In Figure 59 the dislocation is hard core in character, although it can be seen by the unequal length of the arrows forming the triangle around it that its centre does not correspond exactly to the centre of the triangle.

The differences in energy between configurations 1,2 and 3 are so tiny they are not noticeable in the DD plots for them. One possible

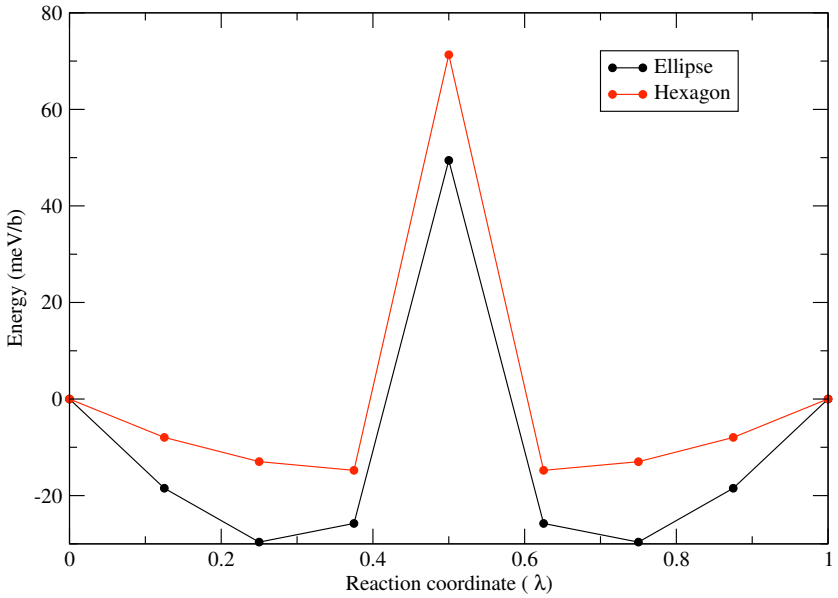


Figure 54: Comparison of the Peierls barrier of elliptical and hexagonal shapes

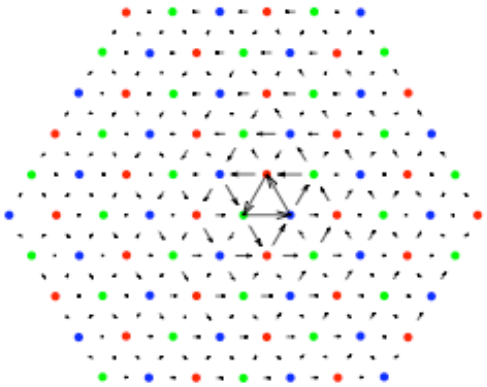


Figure 55: DD plot of configuration 0 for original ellipse Peierls barrier

explanation of the behaviour of the energies is discussed in the section on minimised barriers, in which configurations 1 and 2 are minimised further. The resulting energies are lower than all other previous values indicating that incomplete minimisation may have occurred.

Difficulties in obtaining the expected profile of the Peierls barrier could also be attributed to cancellations between the various contributions to the energy, for example the pair potential and band energy. Close to the endpoints the pair potential seems to dominate slightly as can be seen from Figure 60. This form of the pair potential could play a role in the curious profile of the Peierls barrier close to the endpoints. The potential becomes slightly negative at a distance of about $4.9a_o$.

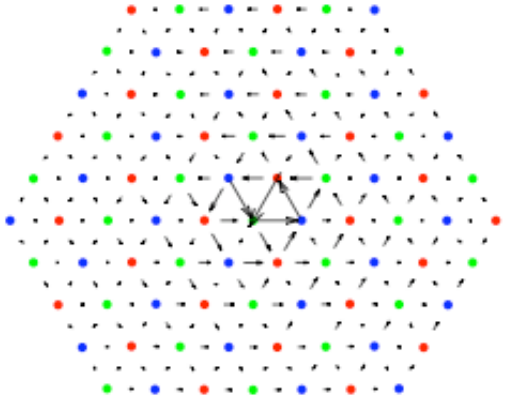


Figure 56: DD plot of configuration 1 for original ellipse Peierls barrier

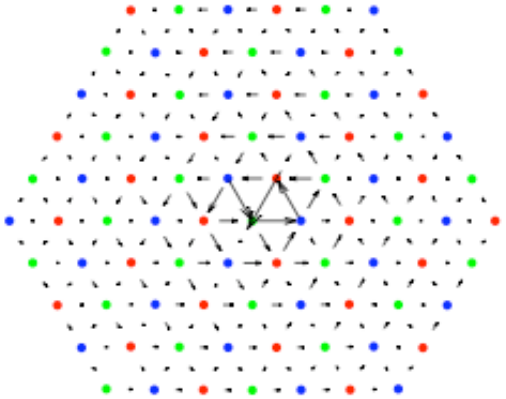


Figure 57: DD plot of configuration 2 for original ellipse Peierls barrier

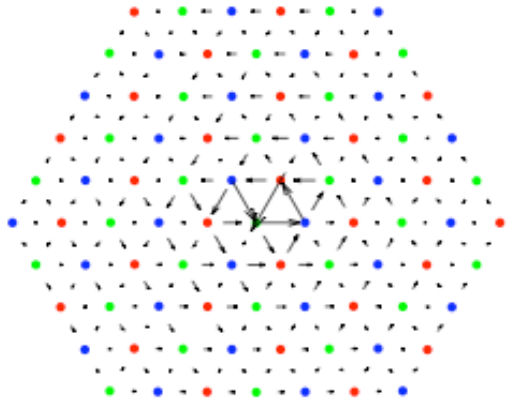


Figure 58: DD plot of configuration 3 for original ellipse Peierls barrier

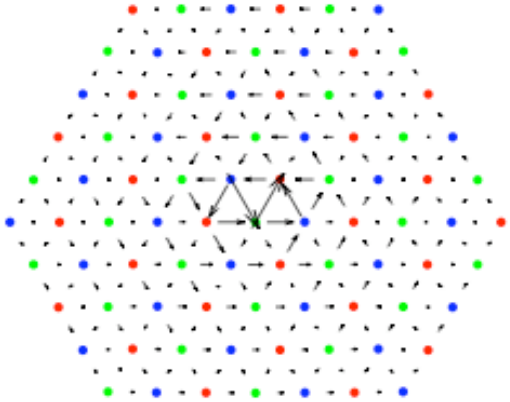


Figure 59: DD plot of configuration 4 for original ellipse Peierls barrier

Particular bond lengths fall into this region and this could have an effect as detailed in Table 13. Between the configurations 0 and 1, and 7 and 8 the bond length of the bond between atoms 57 and 55 on one end of the barrier and 26 and 28 on the other end jumps into the range where the pair potential is negative. These bonds are important as indicated by the purple atoms in Figure 52 and form the sides of the triangles between which the dislocation moves.

Configuration Number	Bond	Bond Length (a_0)
0	57-55	4.80923
1	57-55	4.97299
2	57-55	4.94937
3	57-55	4.93836
4	26-53, 53-57	4.91134 4.91128
5	26-28	4.93870
6	26-28	4.94937
7	26-28	4.97275
8	26-28	4.80923

Table 13: Bond lengths in the negative region of the pair potential for all possible configurations of the original Peierls barrier from the elliptical shape

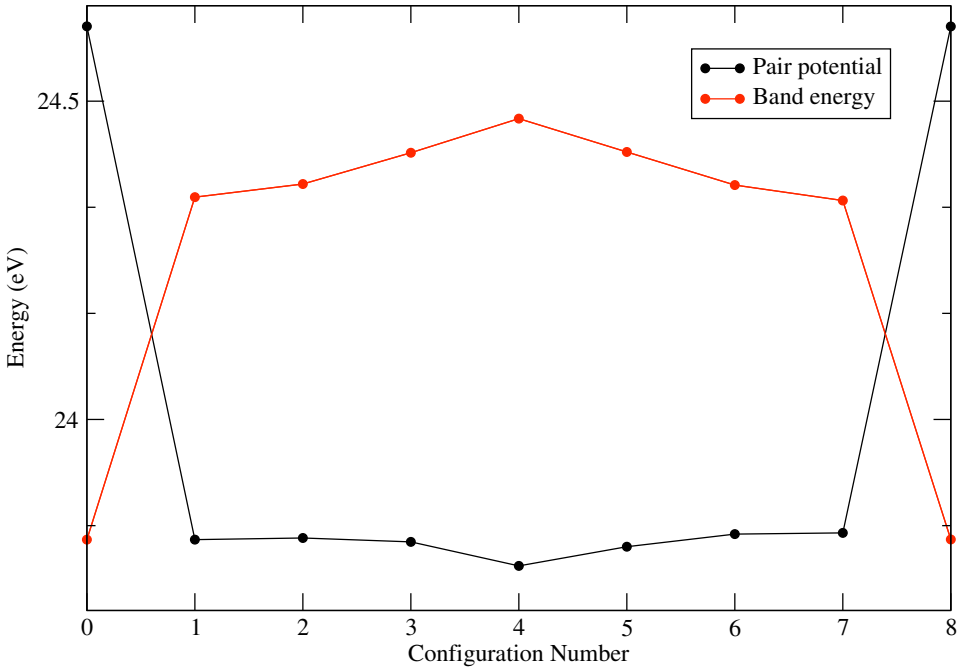


Figure 60: Contributions to the shape of the Peierls barrier for the original elliptical shape

6.2 PARAMETERS THAT CAN BE VARIED

6.2.1 *The effect of freezing*

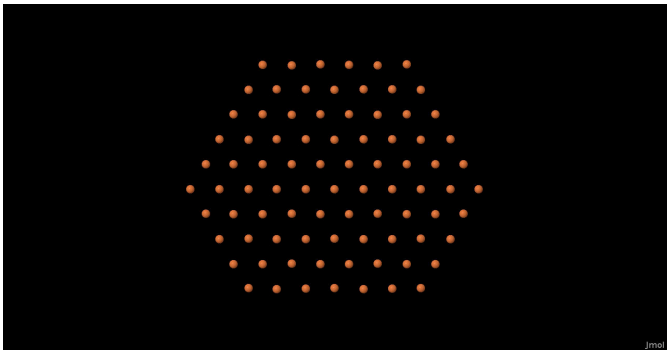
Having displaced all atoms according to the elastic solution, the effect of freezing different numbers of layers of atoms is considered. It is of paramount importance that the atoms are frozen such that the initial and final dislocation sites give the same total energy as in earlier calculations conducted without freezing. This can be achieved by ensuring that all previous symmetries are still imposed on the new arrangement. The case of the ellipse structure with 85 atoms is chosen to understand this effect.

The single outer layer of atoms is frozen first followed by the outer two layers and the Peierls barrier is recalculated in both cases as illustrate in Figure 61. Interestingly the shape of the barrier does not change; however, there are relative energy differences in the barriers between each initial set up as shown in Figure 62. This is most apparent between the completely unrestricted case and the two cases where a certain degree of freezing is implemented. This seems a reasonable outcome

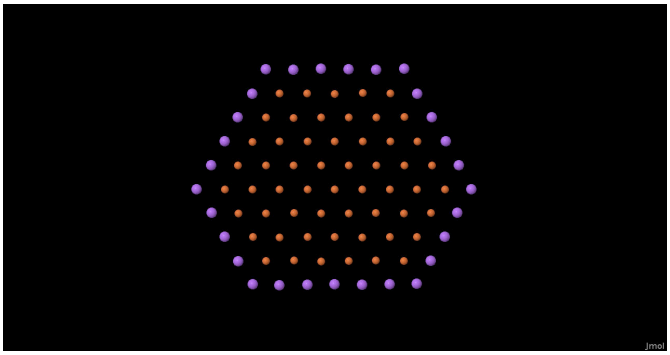
as the outer atoms of the completely free ellipse can move and reach the lowest energy solution. As the profile of the Peierls barrier is not altered by the degree of freezing it is considered simpler for further calculations to be conducted allowing all the atoms to relax in the [111] direction.

6.2.2 *The effect of k point sampling*

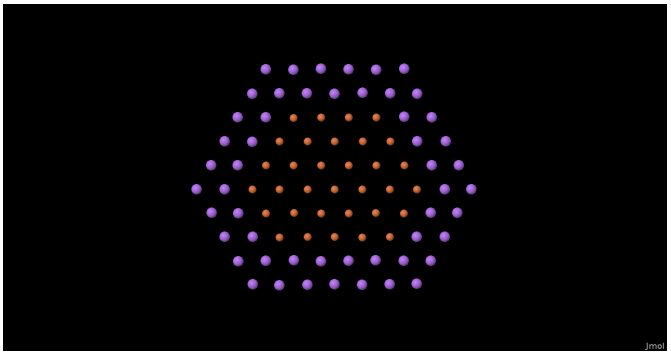
As with freezing different numbers of layers of atoms, the effect of k point sampling on the Peierls barrier does not seem to alter its shape but only to shift its energy. The number of k points are defined in the [111] direction only as free boundaries are used in the other two directions. It can be seen from Figure 63, that increasing the number of k points, reduces the energy of the Peierls barrier.



(a) No freezing



(b) One frozen layer



(c) Two frozen layers

Figure 61: Method of freezing atoms in ellipse calculations

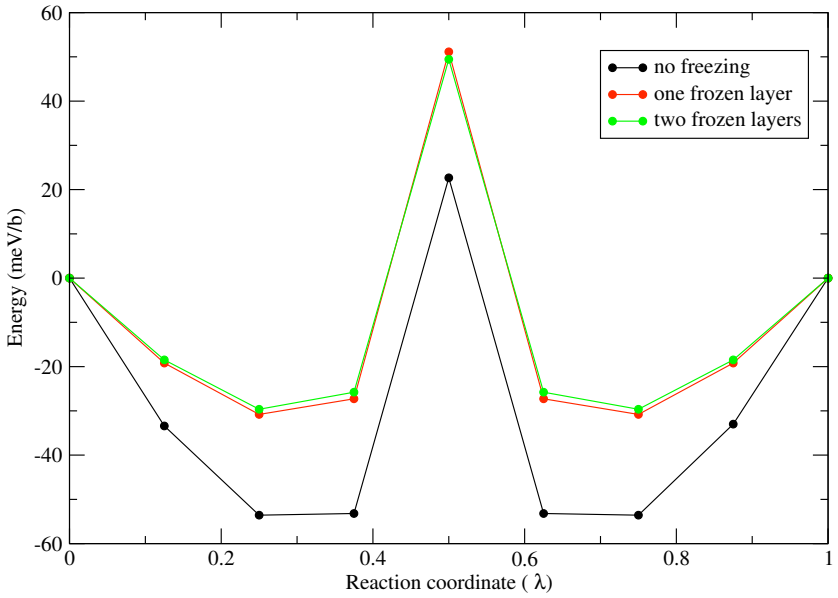


Figure 62: Effect of Freezing for ellipse calculations

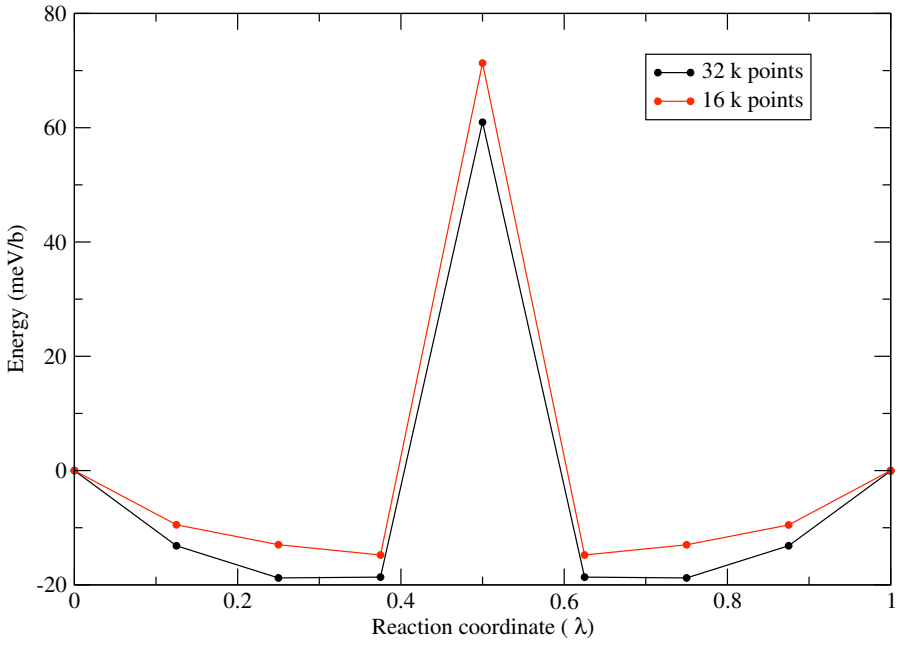


Figure 63: Peierls barrier for 169 atom hexagon with both 16 and 32 k points in the $< 111 >$ direction.

6.3 MINIMISED BARRIERS

The original barriers in Figure 54 make it clear that the minima do not occur at the initial and final points as specified. The difference in energy between the first few reaction coordinates before the hard core peak is very small, so the misplaced minima could possibly be attributed to incomplete minimisation. To investigate this, the final coordinates of the configurations corresponding to the minima were allowed to relax further. There is a small reduction in energy indicating that the original energy minimisation was not completely efficient. To rectify this the newly minimised configurations were used as the start and end points of the dislocation path, and the Peierls barrier recalculated. The result confirms that the minima are now at the endpoints; however, a new and unexpected feature emerges. There is a small local minimum at the central reaction coordinate. This is the case for both elliptical and hexagonal whisker shapes as indicated in Figures 52 and 53. An understanding of what is happening can be gained by looking at the DD plots between the initial and central configuration.

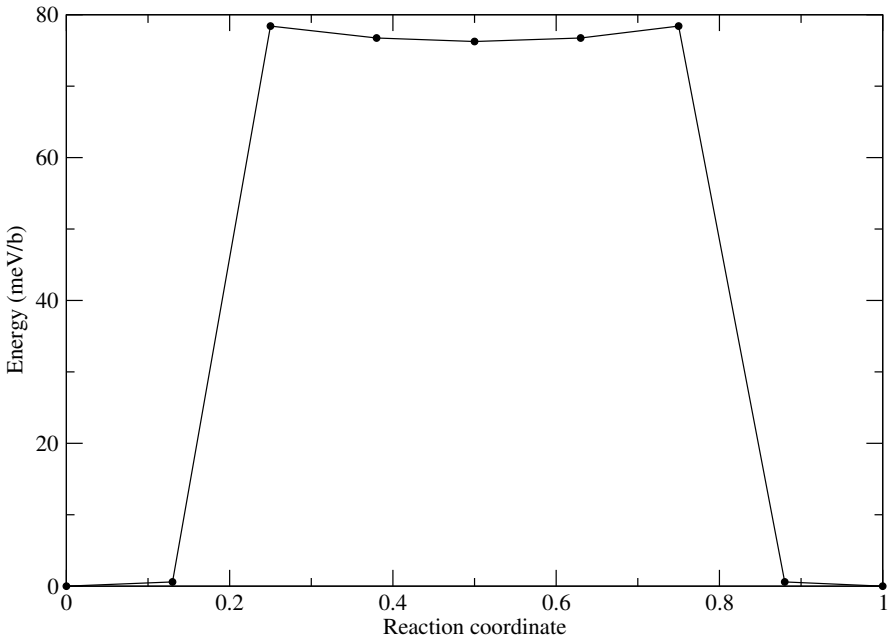


Figure 64: Minimised barrier using ellipse shape

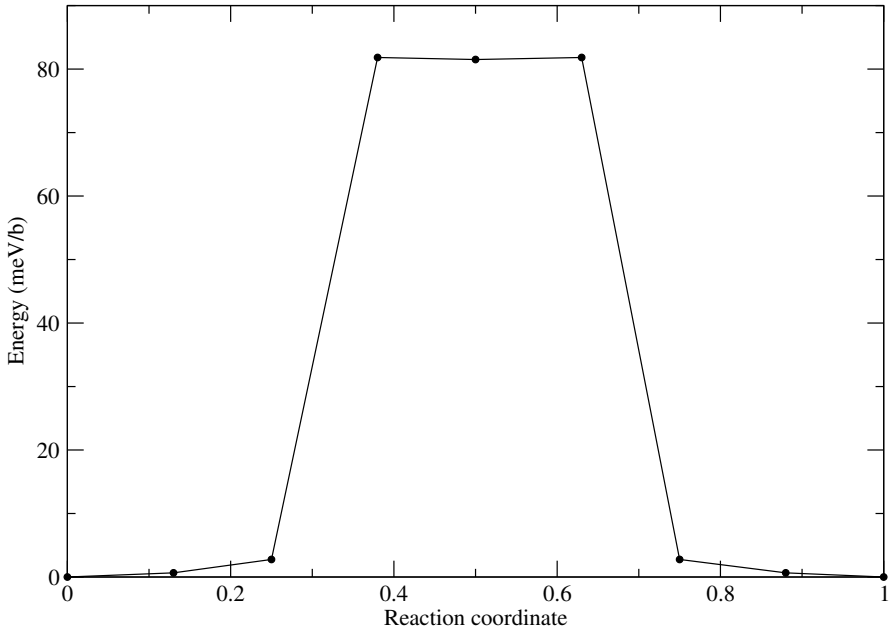


Figure 65: Minimised barrier using hexagon shape

6.3.1 *A closer look at what is happening*

Differential displacement plots are constructed from the atomic configuration corresponding to reaction coordinates of the minimised Peierls barrier for the 85 atom ellipse displayed in Figure 65; the dislocation is dragged from right to left. The initial configuration is labelled 0 and the central configuration 4. The three arrows around the position of the dislocation in the starting configuration of the DD plot (Figure 66) are almost of the same length as expected for the easy core case. The fact that they are not exactly equal, points to the position of the dislocation in the minimised atomic configurations having drifted from the exact centre of the triangle of atoms surrounding it. The first two configurations (0 and 1) indicated in Figures 66 and 67 are very similar and easy core in nature. There is then a pronounced change between configuration 1 and 2 (Figures 67 and 68) as the dislocation moves closer to the central hard core position. There is very little difference between configurations 2,3 and 4 (Figures 68, 69 and 70). This all fits in well with the shape of the Peierls barrer in Figure 65.

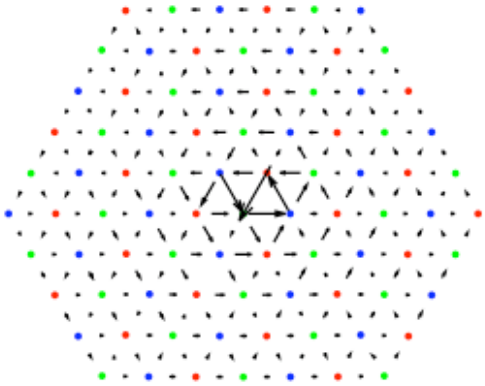


Figure 66: DD plot of configuration 0

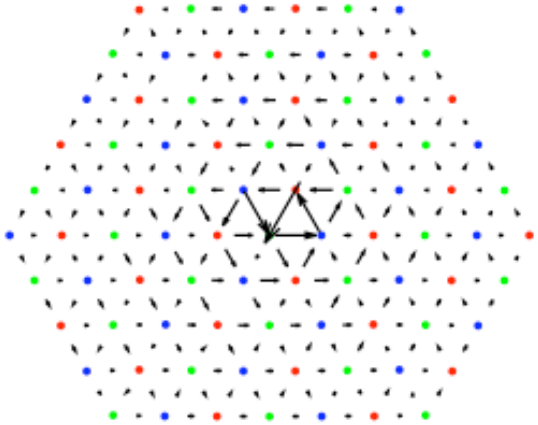


Figure 67: DD plot of configuration 1

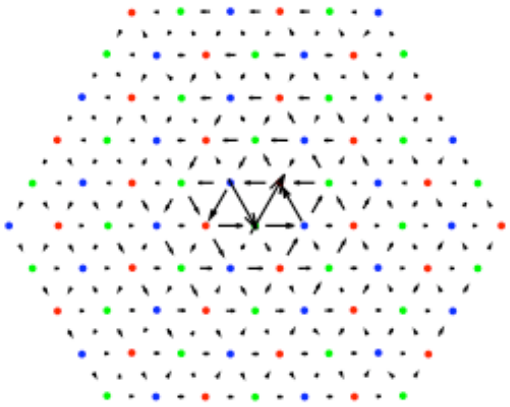


Figure 68: DD plot of configuration 2

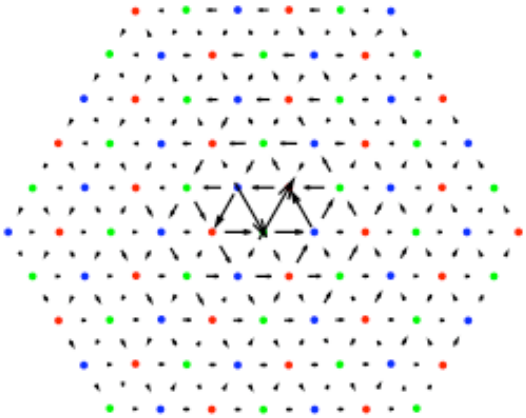


Figure 69: DD plot of configuration 3

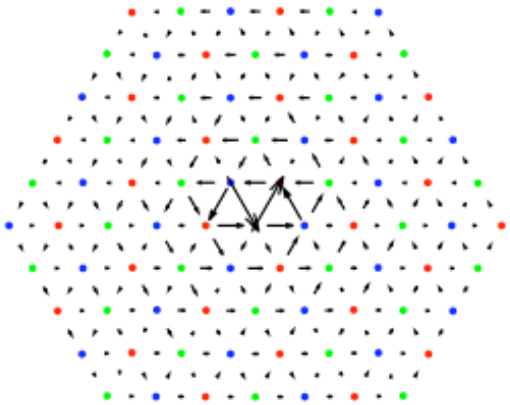


Figure 70: DD plot of configuration 4

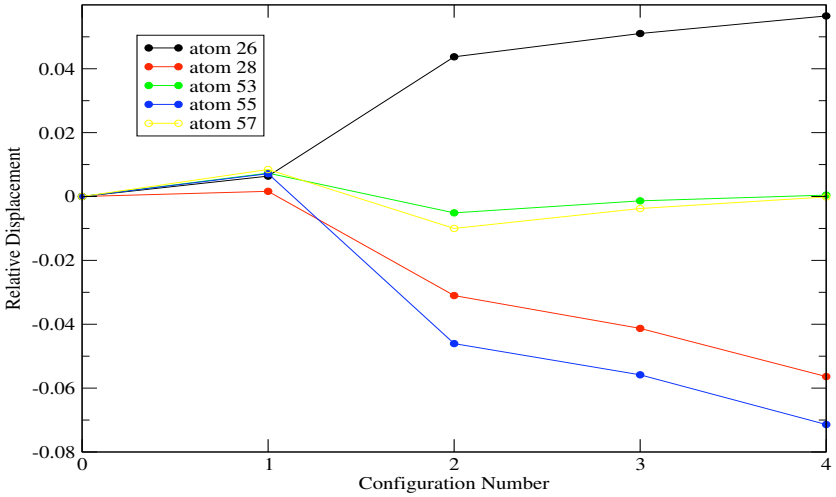
6.4 MOTION OF THE FIVE SIGNIFICANT ATOMS

It is necessary to obtain an understanding of the motion of atoms as the dislocation is dragged from the initial to the final point of the calculation. In order to achieve this the displacement in the $x\,[11\bar{2}]$, $y\,[1\bar{1}0]$ and $z\,[111]$ directions relative to the initial configuration is considered. Only the 5 significant atoms in the dislocation arrangement are considered (these are the purple atoms in Figures 52 and 53). Each of the five atoms in the hexagon and ellipse correspond to each other. Details of the atom numbers which correspond are given in Table 14.

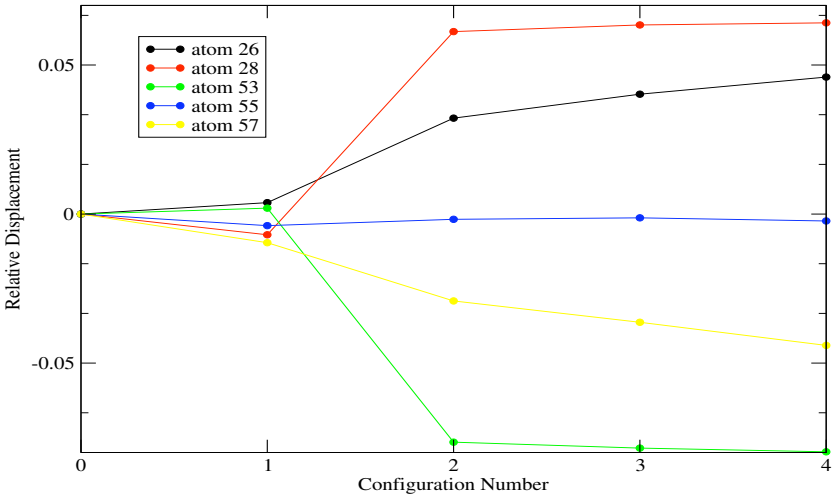
Ellipse	26	28	53	55	57
Hexagon	61	57	103	99	107

Table 14: Equivalent atoms in hexagonal and elliptical shapes

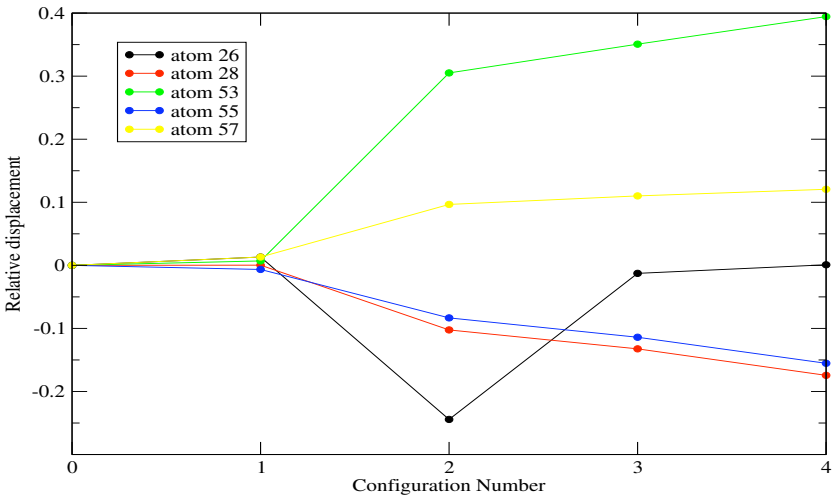
A prominent feature of the minimised Peierls barrier plots is the sudden jump in energy between configurations 1 and 2 in Figure 52 and 2 and 3 in Figure 64. This displays itself in a large change in the relative displacement for the atoms in the x,y and z directions at the corresponding configurations as illustrated in Figures 71 and 72. It can be seen that the corresponding atoms in each shape behave in a similar way in terms of their relative displacement. They are given the same colours to make this feature clearer.



(a) x direction

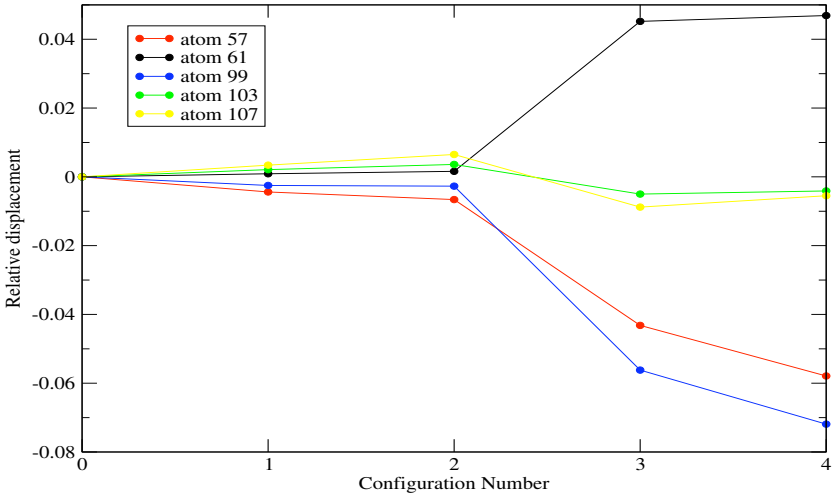


(b) y direction

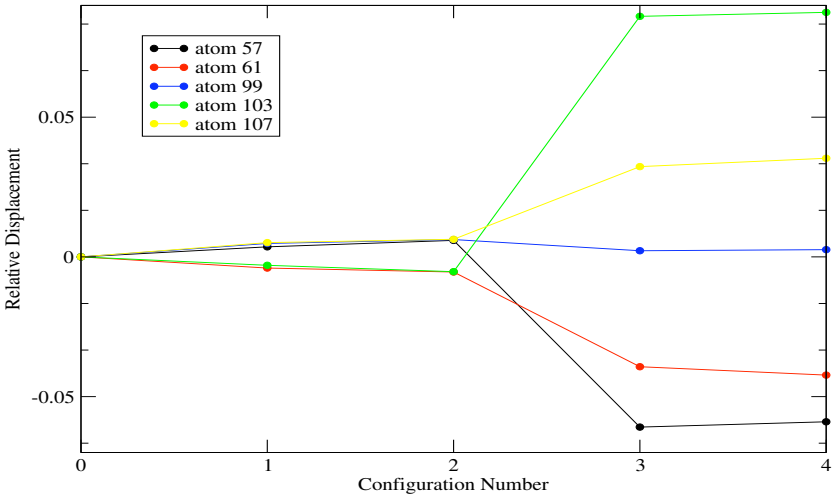


(c) z direction

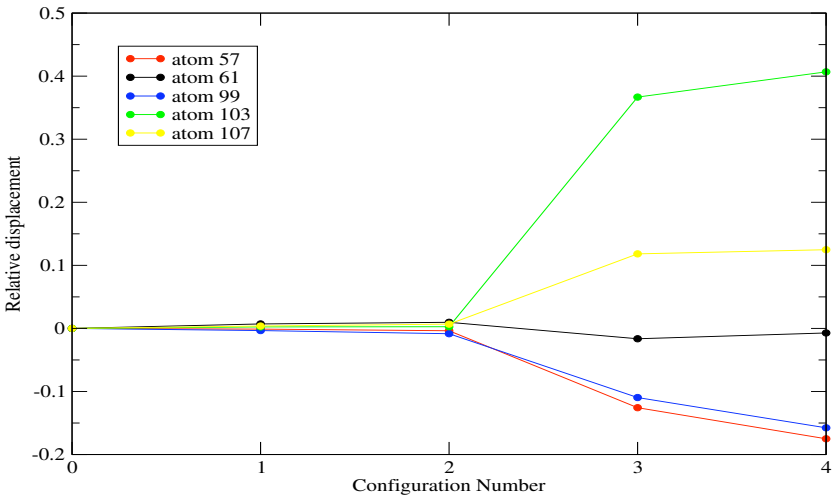
Figure 71: Relative change in displacement of atoms of each atomic configuration along the Peierls barrier for the ellipse shape whose endpoints have been minimised.



(a) x direction



(b) y direction



(c) z direction

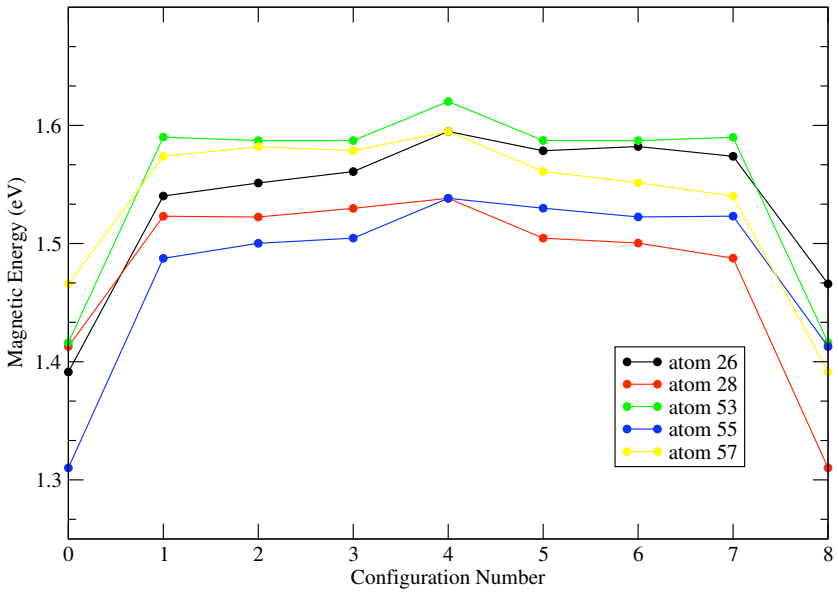
Figure 72: Relative change in displacement of atoms of each atomic configuration along the Peierls barrier for the hexagon shape whose endpoint have been minimised.

6.5 MAGNETIC CONTRIBUTION TO THE ENERGY OF THE FIVE SIGNIFICANT ATOMS

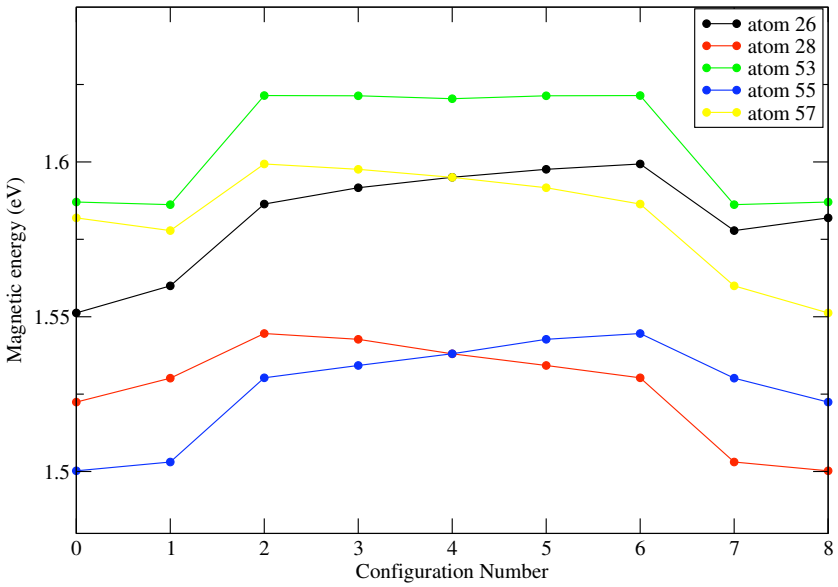
Whilst the DD plots and relative displacement diagrams provide useful clues about the abrupt changes in energy in the Peierls barriers in Figures 64 and 65, a more quantitive approach is needed to understand what is going on. It is found that the magnetic contribution to the energy is very significant (given by the expression $\frac{1}{4} I_s m_I^2$, where I_s is the Stoner parameter and m_I the magnetic moment on an atom I as detailed in Chapter 3). This energy is plotted for the five significant atoms in both shapes for both the original Peierls barriers in Figure 54 and the Peierls barriers with minimised endpoints in Figures 64 and 65.

Looking at the magnetic energy contributions in Figures 73a and 74a, it can be seen that the sharp peak corresponds to the similar peaks in the Peierls barriers of the ellipse and hexagon in Figure 54. However, the rest of the magnetic energy contribution does not share any salient features with the corresponding Peierls barriers. The explanation for the Peierls barrier where the endpoints have been minimised is quite different. The shapes of the magnetic contributions strongly mirrors the Peierls barriers in Figures 52 and 53 for both the ellipse and the hexagon. There are corresponding sharp increases in energy between the minima and the level plateau in the Peierls barrier and they occur at the same configurations. Also summing the change in magnetic energy for the five significant atoms gives a value comparable to the height of the Peierls barrier, confirming that it is the magnetic behaviour of those five atoms which creates the distinctive barrier shape.

Similar atoms in the ellipse and hexagon are given the same colour in the graphs. From this it can be seen that the central atoms (53 in the ellipse shape and 103 in the hexagon case) behave independently, but that the other four atoms form pairs which mirror each other around the central hard core configuration. This is reassuring as the energy of the two endpoints is equivalent and symmetry is preserved.

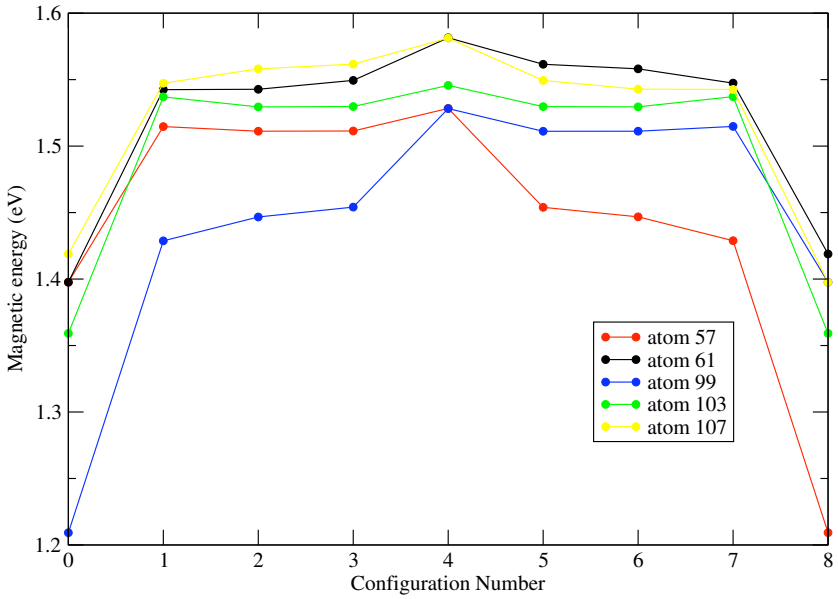


(a) Change in magnetic contribution to the energy of each atomic configuration along the Peierls barrier for the ellipse shape whose endpoints have not been minimised.

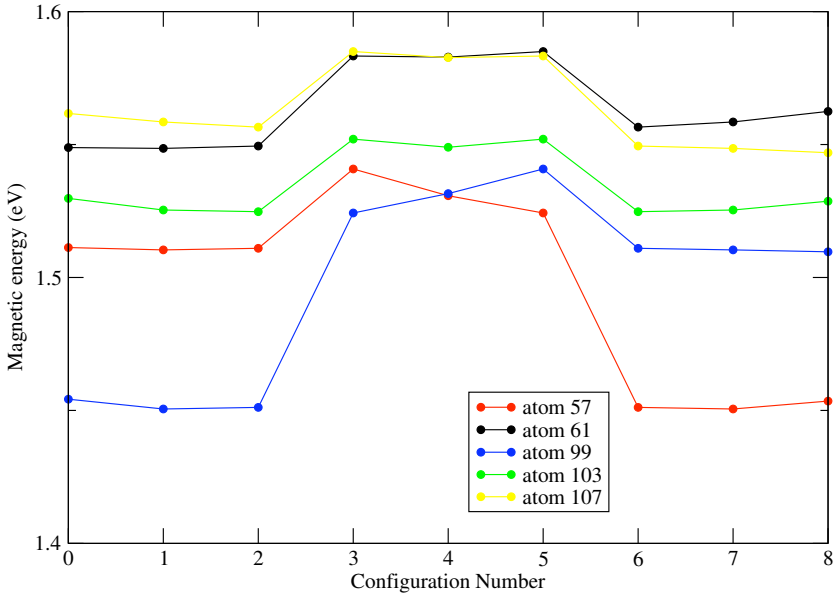


(b) Change in magnetic contribution to the energy of each atomic configuration along the Peierls barrier for the ellipse shape whose endpoints have been minimised.

Figure 73: Ellipse - Magnetic contribution to the energy of the five significant atoms



(a) Change in magnetic contribution to the energy of each atomic configuration along the Peierls barrier for the hexagon shape whose endpoints have not been minimised.



(b) Change in magnetic contribution to the energy of each atomic configuration along the Peierls barrier for the hexagon shape whose endpoints have been minimised.

Figure 74: Hexagon - Magnetic contribution to the energy of the five significant atoms

6.6 WHAT IS THE STRUCTURE?

From the discussion of the Peierls barrier calculations it is clear that there is some ambiguity in the nature of the structure which we need to investigate. One possibility is that the centre of a dislocation has moved to the middle of a bond instead of the initial position in the centre of a triangle in the [111] direction. The atoms in a structure with the dislocation at the position indicated in Figure 75 are allowed to relax in the z direction and the final energy noted. A comparison of the energy of this structure with the energies of the other initial configurations tried is included in Table 15. Its energy is higher than the energy of the minimum of the Peierls barrier, ruling out the possibility of the dislocation centre having moved to the middle of a bond. The resulting DD plot in Figure 75 can be compared to that in Figure 66 (the zero configuration of the Peierls barrier with minimised endpoints) and it can be seen that they are similar but not identical.

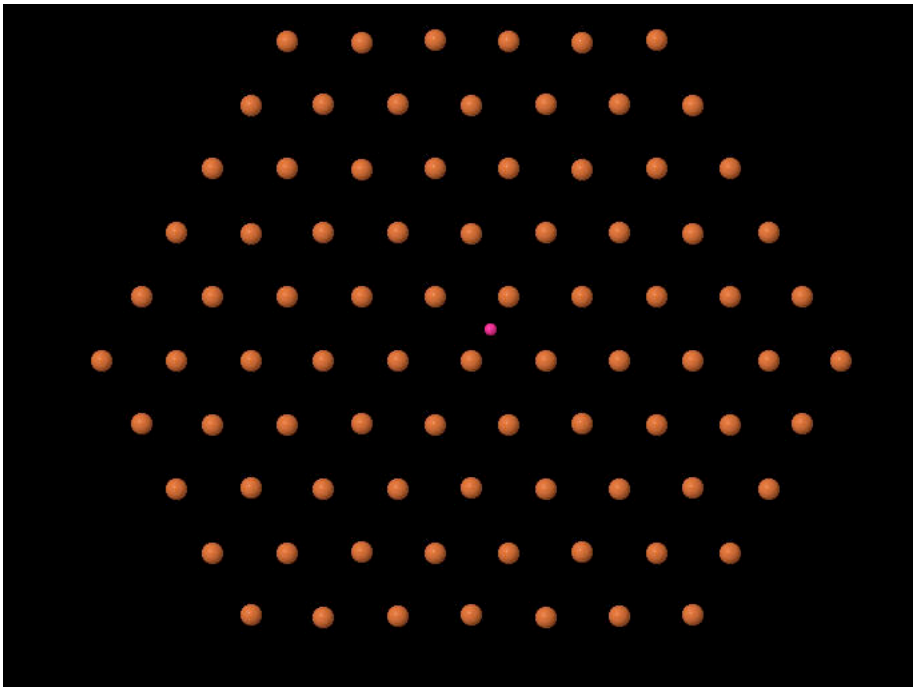


Figure 75: Dislocation in the middle of a bond. The bright pink dot marks the centre of a bond.

Low energy points	Energy (Ry)
Minimum of initial Peierls barrier	-26.7205633
Minimised endpoint of new Peierls barrier	-26.7206259
Mid-point of bond	-26.7149576

Table 15: Comparison of relaxed energies for ellipse with dislocations at different initial positions

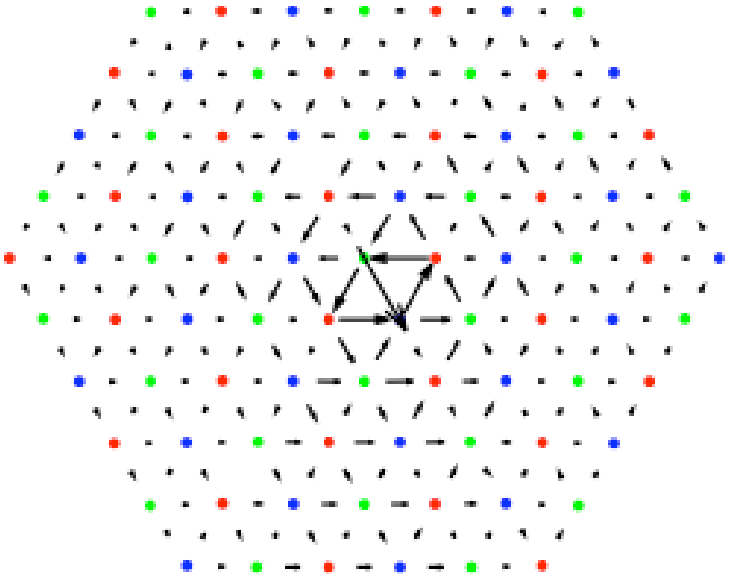


Figure 76: DD plot for dislocation placed initially in the middle of a bond

6.7 CALCULATION OF PEIERLS BARRIER USING NEW PARAMETERS

The tight binding parameters employed to calculate the original barriers were the best available at the time. Very recent parameters published by Madsen et al. [61] were used as a final check on the shape of the barrier. It was found that using these new parameters and the same charge and spin self consistency method the shape of the barrier is more similar to that predicted by DFT and that the energy minima are now at the endpoints of the dislocation drag as expected. The height of the barrier is $32meV/b$ which is in good agreement with the DFT study by Ventelon and Willaime [91]. In these calculations the cell vector was also relaxed in the z direction only. It was found that this had minimal effect on the shape of the barrier as illustrated in Figure 77. The peculiarities concerning the barrier calculated with the earlier parameters discussed

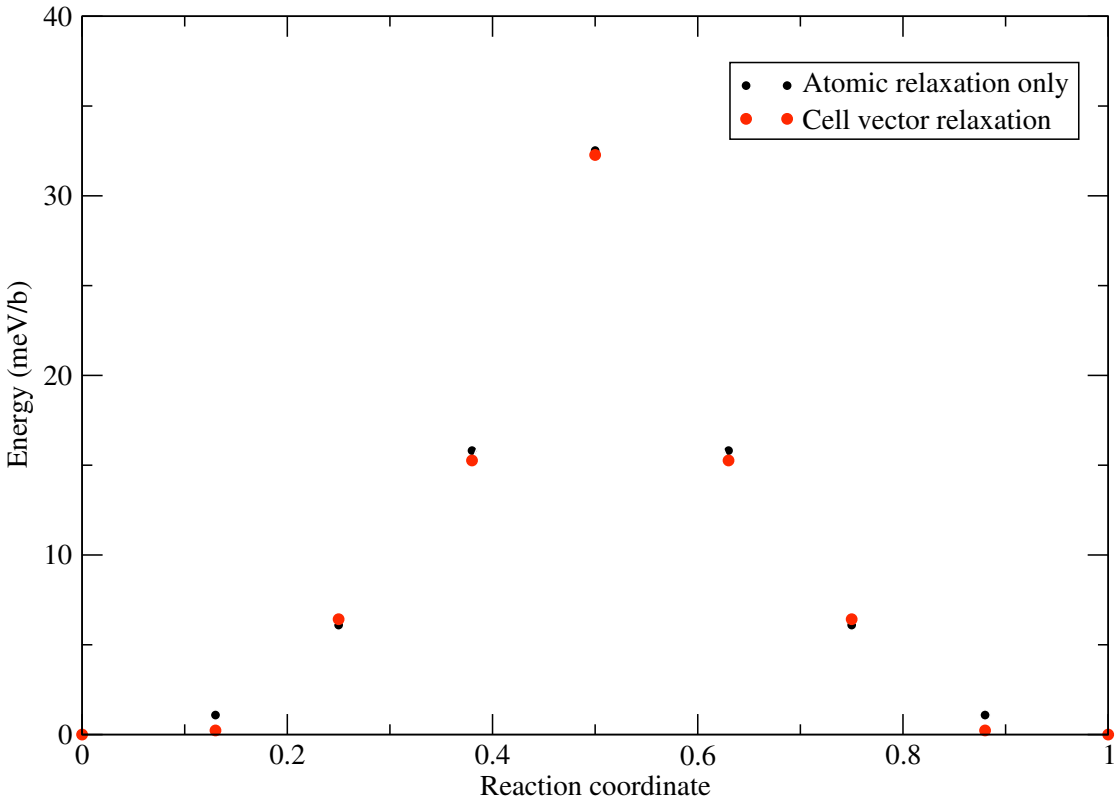


Figure 77: Peierls barrier using latest tight binding parameters.

above have been confirmed to be an artifact of the model and not an inherent feature of tight binding.

6.8 CONCLUSIONS

The initial calculation of Peierls barriers resulted in profiles where the energies of the endpoints at the focal points of the easy core triangles were not the minima. Taking the minimum of these original barriers and minimising them further provides new endpoints for the Peierls barrier calculation. The shape of this new barrier is unexpected, having a level plateau with a small local minimum in the central hard core positions where a maximum is expected. Explaining this shape is a complex task, but through a combination of looking at the changes in the positions of the atoms and the magnetic energy a coherent picture begins to emerge. The contribution of the five main atoms to the magnetic energy of each configuration is perhaps the most significant because of the similarity between this relationship and the shape of the Peierls barriers for both the hexagon and ellipse shapes used in the calculations. Results

obtained using elliptical and hexagonal whisker shapes are very similar, differing only in small details of the Peierls barrier shape. The unusual shapes of the barriers are not a feature of tight binding, but an artifact of the model itself, possibly due to the form of the pair potential as discussed previously.

7

CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

The major contribution of this work was the method for determining atomic spin and chage self-consistently using a form of the Harris-Foulkes functional extended to include both charge and spin explicitly. This method was developed for use in self-consistent tight-binding calculations, in which the interaction of charges on different atomic sites is expressed by Coulomb's law and a Hubbard U is used to model on-site charge interactions. Magnetism is included via the Stoner model of ferromagnetism. The computational task of achieving self-consistency is reduced to finding the stationary points of this novel functional, a task placed on a sound mathematical footing via a generalisation of the Newton-Raphson minimisation procedure. If exploited properly, tight binding is an efficient and accurate method of studying the spin structure of many atom systems without the computational constraints placed on the number of atoms that can be studied using DFT.

The tight binding model from which the parameters are obtained was originally built on the principle of local charge neutrality (LCN). Therefore a large value of the Hubbard U is selected to recreate this. Reducing the value of the Hubbard U makes little difference to the vacancy formation energy and whilst the charges on the atoms increase the magnitude of these charges is very small. This could be interpreted as a demonstration that the LCN assumption is a good one, or, alternatively, that the tight-binding model used forces the atomic charges to be small. It would be instructive if the method for achieving charge and spin self consistency via the extended Harris-Foulkes functional were applied to a model which does not assume that the atoms are charge neutral.

Tests conducted on simple systems during the construction of this method elucidate the complicated spin surface of iron. It is found that,

if Newton-Raphson mixing is used for spin then, even for simple bulk bcc iron, an appropriate choice of initial spin is required to prevent the simulation reaching the wrong solution. For each stage of spin mixing a constraint has to be placed on the maximum allowed change in the spin if convergence is to be attained. For some structures, such as fcc iron at certain lattice parameters, the spin simply oscillates as there is a 'dip' or valley in the energy surface as a function of spin. The solution to this problem is to use Newton-Raphson mixing for charge only and to use Pulay mixing for spin. This combination of mixing schemes is found to be efficient in almost all situations, requiring only a small number of loops to reach convergence. Pulay mixing and Newton-Raphson mixing are related via the linear response matrix, which is calculated exactly in the Newton-Raphson scheme but approximated in the Pulay approach. It should be possible to further improve the intersite charge-mixing convergence by employing a degree of preconditioning. This is an idea for the future. Improvements were made to the implementation of Pulay mixing, so that every time charge and spin were mixed together, the previous spin was remembered and spin mixing started again from this point.

The first extensions to the bulk systems considered involved point defects, in particular the vacancy and the three dumbbell interstitials. It was found that while the ordering and magnitude of the formation energies of the interstitials were in reasonable agreement with density functional theory, the spin structure of the $< 110 >$ and $< 111 >$ dumbbell interstitials was not as expected. The spins on the two dumbbell atoms were found to be lower than the bulk spin, as in DFT, and far from the interstitial atoms a bulk structure was again established. The peculiarity in the case of the $< 110 >$ and $< 111 >$ interstitials was a sudden dip of the spins at a distance from the interstitial when it seemed bulk character had already been established. The $< 100 >$ interstitial behaves more or less as expected, although the magnitude of the spin on the interstitial atoms is larger than in DFT. This is consistent with the observation that the displacements of the atoms surrounding the interstitial are greater than in DFT. It would be interesting to feed

these final structures obtained from tight binding into DFT and to find the energies and spins produced.

The limitations placed on the work thus far are related to computational requirements of the method presented in Chapter 3 and the number of atoms that can be considered as a result. Currently the scaling goes as the fourth power of the number of atoms. The ideal solution would be to develop a linear scaling technique for the Newton-Raphson scheme, enabling the time for each loop to be reduced whilst still achieving charge and spin self consistency in very few loops.

The defects already considered could be investigated more fully if more atoms could be considered and the time taken for each loop reduced. For example, in the case of straight dislocations investigated in chapter 5, it would be possible to conduct cell relaxation in the z direction as well as the existing atomic relaxation. Results with atomic relaxation only produce a non-degenerate core structure in agreement with DFT and a roughly linear relationship between the natural logarithm of the radius and the easy core formation energy as predicted by elasticity theory. However, the convergence of hard-easy core energy difference with radius is quite noisy and it would be interesting to see if this relationship was improved by conducting cell relaxation.

In order to deal with the one dimensional nature of the screw dislocation simulation cell it was necessary to adapt the mixing scheme using 1-D Ewald sums. However, it was still not possible to implement inter-site charge mixing in this method as described in Chapter 5. Future effort is definitely required to meet this challenge and create a working scheme for 1-D inter-site calculations.

Work on the motion of dislocations has so far been limited to the calculation of Peierls barriers. The tight-binding parameters used for the repulsive potential produce a barrier shape that is inconsistent with DFT. In the future other tight-binding models could be tried and variations in the shape of the barrier observed. Being able to conduct calculations in shorter times would be useful in this task. As well as straight dislocations the formation of both single and double static kinks could also be investigated as these are of interest to the process of

dislocation motion via double kink nucleation. Screw dislocations have received the majority of the scrutiny of the possible defects in fusion power plants as they are the least mobile and have the greatest effect on plasticity. Edge dislocations have been more neglected and it might be worth looking at these using the methods developed in this work using tight binding.

Defects until this point have been considered singly, and the natural extension to this would be to consider interactions between defects. The simplest examples that spring to mind are di-vacancies and di-interstitials. Following this, clusters of defects or vacancy and intersitial loops could then be studied. The interaction of straight dislocations with impurity atoms such as chromium would be particularly pertinent to the use of ferritic/martensitic steels in fusion power plants. Work has been conducted in this area employing both DFT and empirical potentials but not tight binding. The ellipse set up designed to cope with Peierls barrier calculations could be developed so that the focal points of the ellipse were further apart rather than in two adjacent triangles. A single dislocation could then be placed in the centre of each of these triangles simultaneously and their interaction investigated. The different cases that could be looked into are easy-easy, hard-easy and hard-hard core dislocations.

A potentially challenging issue to the correct functioning of a hydrogen fusion power plant relates to the phenomena of grain boundaries and their effect on plasticity. Grain boundaries are important for several reasons. At low temperatures they impede dislocation motion; the smaller the size of the individual grains the greater this effect and the stronger the material becomes (Hall-Petch Effect [8]). At high temperatures recrystallisation and grain growth occurs. In this process dislocations and other defects are 'swallowed' by the grains as they expand. Finally, embrittlement can also occur because of segregation of impurities (such as helium) at grain boundaries; this is obviously undesirable in the case of load bearing components in a power plant. This multi-faceted problem could be tackled using an improved and more

developed version of the tight binding convergency scheme discussed in this research.

The entire project has been extremely technically challenging due in the most part to the spin structure of iron and its complexities. Having developed a working scheme, it was possible to consider certain point defects and the $\frac{1}{2}$ [111] screw dislocation and gain important information about their geometric and magnetic structures. Tight binding has been shown to be a useful tool in considering magnetic materials containing extended defects. With the further developments suggested above, the efficiency of tight-binding calculations of extended defects could be improved substantially.

BIBLIOGRAPHY

[1] Iter website. URL http://www.iter.org.

[2] Kiel website. URL tf.uni-kiel.de.

[3] GJ Ackland and R Thetford. *Philos Mag A*, 56(1):15–30, 1987.

[4] GJ Ackland, DJ Bacon, AF Calder, and T Harry. *Philos Mag A*, 75 (3):713–732, 1997.

[5] GJ Ackland, MI Mendelev, DJ Srolovitz, S Han, and AV Barashev. *J Phys-Condens Mat*, 16(27):S2629–S2642, 2004.

[6] JF Annett. *Comp Mater Sci*, 4(1):23–42, 1995.

[7] K Arakawa, K Ono, M Isshiki, K Mimura, M Uchikoshi, and H Mori. *Science*, 318(5852):956–959, 2007.

[8] MF Ashby and DRH Jones. Elsevier, 2011.

[9] AV Barashev, YN Osetsky, and DJ Bacon. *Philos Mag A*, 80(11): 2709–2720, 2000.

[10] CS Becquart, C Domain, A Legris, and JC Van Duysen. *J Nucl Mater*, 280(1):73–85, 2000.

[11] J Boutard, S Dudarev, and E Diegele. *Materials Issues for Generation IV Systems, Proceedings of the NATO Advanced Study Institute on Materials Issues for Generation IV Systems: Status, Open Questions and Challenges, Cargese, Corsica, 2007, Springer, Berlin*, 2008.

[12] J-L Boutard, A Alamo, R Lindau, and M Rieth. *Cr Phys*, 9(3-4): 287–302, 2008.

[13] S Boys. *Proceedings of the Royal Society of London. Series AMathematical and Physical Sciences*, 200(1063):542–554, Feb 1950.

[14] D Brunner and J Diehl. *Phys Status Solidi A*, 124(2):455–464, 1991.

[15] W Cai, VV Bulatov, JP Chang, J Li, and S Yip. *Philos Mag*, 83(5): 539–567, 2003.

[16] J Chaussidon, M Fivel, and D Rodney. *Acta Materialia*, 54(13): 3407–3416, 2006.

[17] E Clouet, L Ventelon, and F Willaime. *Physical review letters*, 102 (5):055502, 2009.

[18] RE Cohen and S Mukherjee. *Phys Earth Planet In*, 143:445–453, 2004.

[19] PM Derlet, D Nguyen-Manh, and S. L Dudarev. *Phys Rev B*, 76 (5):054107, 2007.

[20] C Domain and C Becquart. *Phys Rev B*, 65:024103, 2001.

[21] C Domain and G Monnet. *Physical review letters*, 95(21):215506, 2005.

[22] S. L Dudarev, J. L Boutard, R Laesser, M. J Caturla, P. M Derlet, M Fivel, C. C Fu, M. Y Lavrentiev, L Malerba, M Mrovec, D Nguyen-Manh, K Nordlund, M Perlado, R Schaeublin, H Van Swygenhoven, D Terentyev, J Wallenius, D Weygand, and F Willaime. *J Nucl Mater*, 386:1–7, 2009.

[23] SL Dudarev. *Philos Mag*, 83(31-34):3577–3597, 2003.

[24] SL Dudarev and PM Derlet. *J Phys-Condens Mat*, 17(44):7097–7118, 2005.

[25] SL Dudarev, R Bullough, and PM Derlet. *Physical review letters*, 100(13):135503, 2008.

[26] MS Duesbery and V Vitek. *Acta Mater*, 46(5):1481–1492, 1998.

[27] MS Duesbery, B Joos, and DJ Michel. *Phys Rev B*, 43(6):5143–5146, 1991.

[28] J.D. Eshelby, W.T. Read, and W. Shockley. *Acta Mater*, 1(3):251 – 259, 1953.

[29] D Farkas, CG Schon, MSF DeLima, and H Goldenstein. *Acta Mater*, 44(1):409–419, 1996.

[30] R Feynman. *Physical Review*, 56:340–343, 1939.

[31] MW Finnis and JE Sinclair. *Philos Mag A*, 50(1):45–55, 1984.

[32] SP Fitzgerald and S Aubry. *J Phys.:Condens Mat*, 22(29):295403, 2010.

[33] SP Fitzgerald and Z Yao. *Phil Mag Lett*, 89(9):581–588, 2009.

[34] W Foulkes and R Haydock. *Phys Rev B*, 39:12520–12536, 1989.

[35] T Frauenheim, G Seifert, M Elstner, Z Hajnal, G Jungnickel, D Porezag, S Suhai, and R Scholz. *Phys Status Solidi B*, 217 (1):41–62, 2000.

[36] SL Frederiksen and KW Jacobsen. *Philos Mag*, 83(3):365–375, 2003.

[37] C-C Fu, F Willaime, and P Ordejón. *Physical review letters*, 92: 177503, 2004.

[38] CC Fu, J Dalla Torre, F Willaime, JL Bocquet, and A Barbu. *Nat Mater*, 4(1):68–74, 2005.

[39] M. R Gilbert, S Queyreau, and J Marian. *Phys Rev B*, 84(17):174103, 2011.

[40] CM Goringge, DR Bowler, and E Hernandez. *Rep Prog Phys*, 60 (12):1447–1512, 1997.

[41] GY Guo and HH Wang. *Chinese J Phys*, 38(5):949–961, 2000.

[42] AA Harms, KF Schoepf, GH Miley, and DR Kingdon. World Scientific, Singapore, 2002.

[43] J Harris. *Phys Rev B*, 31(4):1770–1779, 1985.

[44] T Harry and DJ Bacon. *Acta Mater*, 50(1):195–208, 2002.

[45] HC Herper, E Hoffmann, and P Entel. *Phys Rev B*, 60(6):3839–3848, 1999.

[46] JP Hirth and J Lothe. *Theory of dislocations*. Krieger Publishing, Jan 1982.

[47] AP Horsfield. *Phys Rev B*, 56(11):6594–6602, 1997.

[48] AP Horsfield and AM Bratkovsky. *Phys Rev B*, 53(23):15381–15384, 1996.

[49] AP Horsfield and AM Bratkovsky. *J Phys.: Condens Mat*, 12(2): R1–R24, 2000.

[50] AP Horsfield, DR Bowler, AM Bratkovsky, M Fearn, P Goodwin, S Goedecker, C Goringe, D Miller, R Harris, and D Nguyen-Manh. Oxon code, 2001.

[51] J Hubbard. *Proceedings of the Royal Society of London. Series AMathematical and Physical Sciences*, 276(1365):238–257, 1963.

[52] S Ismail-Beigi and TA Arias. *Physical review letters*, 84(7):1499–1502, 2000.

[53] DE Jiang and EA Carter. *Phys Rev B*, 67(21):214103, 2003.

[54] SD Kenny and AP Horsfield. *Comput Phys Commun*, 180(12): 2616–2621, 2009.

[55] SD Kenny, AP Horsfield, and H Fujitani. *Phys Rev B*, 62(8):4899–4905, 2000.

[56] M Kiritani. *J Nucl Mater*, 216:220–264, 1994.

[57] W Kohn and LJ Sham. *Phys. Rev*, 140:A1133, 1965.

[58] A Koyama, A Hishinuma, DS Gelles, RL Kleuh, W Dietz, and K Ehrlich. *J. Nucl. Mater*, 138(5):233–237, 1996.

[59] GQ Liu, D Nguyen-Manh, BG Liu, and DG Pettifor. *Phys Rev B*, 71(17):174115, 2005.

[60] P-W Ma, WC Liu, CH Woo, and SL Dudarev. *Journal of Applied Physics*, 101(7):073908, 2007.

[61] GKH Madsen, EJ McEniry, and R Drautz. *Phys Rev B*, 83(18): 184119, 2011.

[62] J Marian, BD Wirth, A Caro, B Sadigh, GR Odette, JM Perlado, and TD de la Rubia. *Phys Rev B*, 65(14):144102, 2002.

[63] J Marian, W Cai, and VV Bulatov. *Nat Mater*, 3(3):158–163, 2004.

[64] MI Mendelev, S Han, DJ Srolovitz, GJ Ackland, DY Sun, and M Asta. *Philos Mag*, 83(35):3977–3994, 2003.

[65] M Mrovec, D Nguyen-Manh, DG Pettifor, and V Vitek. *Phys Rev B*, 69(9):094115, 2004.

[66] M Mrovec, R Groeger, A. G Bailey, D Nguyen-Manh, C Elsaesser, and V Vitek. *Phys Rev B*, 75(10):104119, 2007.

[67] M Mrovec, D Nguyen-Manh, C Elsaesser, and P Gumbsch. *Physical review letters*, 106(24):246402, 2011.

[68] AHW Ngan and M Wen. *Comp Mater Sci*, 23(1-4):139–145, 2002.

[69] D Nguyen-Manh, AP Horsfield, and SL Dudarev. *Phys Rev B*, 73: 020101, Jan 2006.

[70] D Nguyen-Manh, M Mrovec, and SP Fitzgerald. *Mater Trans*, 49 (11):2497–2506, 2008.

[71] D Nguyen-Manh, M. Yu Lavrentiev, and S. L Dudarev. *J Nucl Mater*, 386:60–63, 2009.

[72] P Olsson, IA Abrikosov, L Vitos, and J Wallenius. *J Nucl Mater*, 321(1):84–90, 2003.

[73] YN Osetsky, DJ Bacon, and A Serra. *Phil Mag Lett*, 79(5):273–282, 1999.

[74] AT Paxton and MW Finnis. *Phys Rev B*, 77(2):024428, 2008.

[75] KL Petal. *Rev.Mod.Phys*, 39:395, 1967.

[76] DG Pettifor. *Physical review letters*, 63(22):2480–2483, 1989.

[77] P Pulay. *Chem Phys Lett*, 73(2):393–398, 1980.

[78] KF Riley, MP Hobson, and SJ Bence. *Mathematical Methods for Physics and Engineering*. Cambridge University Press, 2006.

[79] HE Schaefer, K Maier, M Weller, D Herlach, A Seeger, and J Diehl. *Scripta Metall Mater*, 11(9):803–809, 1977.

[80] DE Segall, A Strachan, WA Goddard, S Ismail-Beigi, and TA Arias. *Phys Rev B*, 68(1):014104, 2003.

[81] ER Smith. *J Chem Phys*, 128(17):174104, 2008.

[82] PK Soin, AP Horsfield, and D Nguyen-Manh. *Comput Phys Communications*, 182(6):1350–1360, 2011.

[83] N Soneda and TD de la Rubia. *Philos Mag A*, 78(5):995–1019, 1998.

[84] R Soulairol, C-C Fu, and C Barreteau. *J Phys-Condens Mat*, 22(29): 295502, 2010.

[85] L Stixrude, R Cohen, and D Singh. *Phys Rev B*, 1994.

[86] E Stoner. *Proceedings of the Royal Society of London. Series A*, 1938.

[87] AP Sutton, MW Finnis, DG Pettifor, and Y Ohta. *J Phys C Solid State*, 21(1):35–66, 1988.

[88] S Takaki, J Fuss, H Kugler, U Dedek, and H Schultz. *Radiat Eff Defect S*, 79(1-4):87–122, 1983.

[89] DA Terentyev and L Malerba. *Comp Mater Sci*, 43(4):855–866, 2008.

[90] A Vehanen, P Hautojarvi, J Johansson, J Ylikauppila, and P Moser. *Phys Rev B*, 25(2):762–780, 1982.

[91] L Ventelon and F Willaime. *J Comput-Aided Mater*, 14:85–94, 2007.

[92] L Ventelon, F Willaime, and P Leyronnas. *J Nucl Mater*, 386-88: 26–29, 2009.

[93] V Vitek. *Cryst Latt Def Amorp*, 5(1):1–34, 1974.

[94] GF Wang, A Strachan, T Cagin, and WA Goddard. *Phys Rev B*, 68 (22):224101, 2003.

[95] M Wen and AHW Ngan. *Acta Mater*, 48(17):4255–4265, 2000.

[96] C Woodward. *Mat Sci Eng A-Struct*, 400:59–67, 2005.

[97] C Woodward and SI Rao. *Physical review letters*, 88(21):216402, 2002.

[98] W Xu and JA Moriarty. *Phys Rev B*, 54(10):6941–6951, 1996.

[99] W Xu and JA Moriarty. *Comp Mater Sci*, 9(3-4):348–356, 1998.

[100] Z Yao, S Xu, ML Jenkins, and MA Kirk. *J Electron Microsc*, 57(3): 91–94, 2008.