Theory of Isolated Dopants in GaAs (110) Surfaces

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Frederick Tilley
Department of Physics and Astronomy
University of Leicester

September 2015
Theory of Isolated Dopants in GaAs (110) Surfaces

by

Frederick Tilley

Abstract

In this thesis we perform a range of highly accurate density functional theory (DFT) calculations for a GaAs (110) slab containing almost all of the near-surface single atomic dopants from groups III, IV and V of the periodic table. We look in detail at the relaxed geometry and local density of states of the doped surface, and using the theory of Tersoff and Hamann we generate STM images of the different dopant systems. Where possible we compare to experimental results obtaining excellent qualitative and quantitative agreement, with bond lengths and shifts in STM contrast agreeing to within 0.03 Å and 0.09 Å respectively.

We are able to show very clear trends in both the relaxed positions and STM image contrasts for the range of dopants. These trends are determined by the covalent radius of the dopants. Dopants with larger radii relax out of the surface and ones with smaller radii relax into the surface, and these relaxations cause the different contrasts in the STM images. These trends fit very well with existing results for nitrogen and silicon doped systems, and also allow us to fill in the gaps for those systems that have not been as thoroughly investigated. Our analysis applies equally across the three groups of dopants from the periodic table covering isovalent, donor and acceptor cases.

By developing a geometrical model based on the covalent radii of the dopants and host atoms, we show how the covalent radius determines the geometry of the surface, which in turn determines the contrast seen in the STM images. Using this model we are able to explain and predict the relaxation and STM images for all the dopants in this work to a high degree of accuracy without relying on DFT simulations.
Contents

List of Figures xvi
List of Tables xx
Acknowledgements xxi
List of Publications xxii

1 Introduction 1
  1.1 Introduction ................................................. 1
  1.2 Scanning Tunnelling Microscopy ............................ 4
  1.3 Synopsis ..................................................... 9

2 Theory 12
  2.1 Introduction .................................................. 12
  2.2 Density Functional Theory Calculations ................. 13
    2.2.1 Hohenberg-Kohn DFT Theory .......................... 13
    2.2.2 N-Representable Density ............................... 16
    2.2.3 Kohn-Sham Ansatz .................................... 17
    2.2.4 LDA .................................................... 18
    2.2.5 Self-Consistency ...................................... 20
    2.2.6 Accuracy of the Functionals .......................... 21
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3</td>
<td>Pseudopotentials</td>
<td>23</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Norm-Conserving Pseudopotentials</td>
<td>25</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Pseudopotential Smoothness</td>
<td>26</td>
</tr>
<tr>
<td>2.4</td>
<td>ABINIT</td>
<td>27</td>
</tr>
<tr>
<td>2.5</td>
<td>STM and STS</td>
<td>28</td>
</tr>
<tr>
<td>2.6</td>
<td>k-Point Grid</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>Geometrical Method for Estimating Impurity Positions</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>Clean GaAs Simulation</td>
<td>35</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Investigation of Pseudopotentials: Bulk GaAs</td>
<td>36</td>
</tr>
<tr>
<td>3.3</td>
<td>Clean GaAs (110)</td>
<td>42</td>
</tr>
<tr>
<td>4</td>
<td>Isovalent Dopants in GaAs</td>
<td>46</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Simulations</td>
<td>48</td>
</tr>
<tr>
<td>4.3</td>
<td>Significance of Geometry</td>
<td>53</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Covalent Radius</td>
<td>55</td>
</tr>
<tr>
<td>4.3.2</td>
<td>N:GaAs</td>
<td>64</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Second Layer Features</td>
<td>67</td>
</tr>
<tr>
<td>4.4</td>
<td>Group III and V Isoelectronic Impurities</td>
<td>69</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Group III STM Images</td>
<td>70</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Group V STM Images</td>
<td>73</td>
</tr>
<tr>
<td>4.5</td>
<td>Bi in InP</td>
<td>78</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>83</td>
</tr>
<tr>
<td>5</td>
<td>Group IV Donors on the (110) Surface of GaAs</td>
<td>84</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>84</td>
</tr>
</tbody>
</table>
5.2 Calculation ................................................. 85
  5.2.1 $E_{\text{cut}}$ and $k$-point Convergence ................. 86
  5.2.2 System Size Convergence ............................ 90
5.3 Relaxed Positions ........................................ 94
  5.3.1 Geometrical Model ................................... 103
5.4 Dopant Density of States ................................. 105
  5.4.1 LDOS Above the Surface ............................. 108
  5.4.2 STM Images ........................................ 112
5.5 Summary .................................................. 113

6 Group IV Acceptors and Group IV Charged Impurities 116
  6.1 Introduction ........................................... 116
  6.2 Group IV Acceptors .................................... 117
    6.2.1 Relaxations ....................................... 119
    6.2.2 Local Density of States .......................... 123
  6.3 Group IV Charged States ............................... 126
    6.3.1 Relaxations ....................................... 128
    6.3.2 Local Density of States .......................... 133
  6.4 Summary ................................................ 143

7 Conclusions .............................................. 146

A Example ABINIT Input File ................................. 151

References ............................................... 154
List of Figures

2.1 Single arrows represent steps made by solving the Schrödinger equation, the double line arrows represent the Hohenburg-Kohn theory that relates the ground state density to the external potential and the double headed arrow is the Kohn-Sham ansatz that links a many-body interacting density to a non-interacting one. Reproduced from [1].

2.2 Flowchart showing the steps taken to obtain a self-consistent Kohn-Sham density.

2.3 Schematic representation of a pseudopotential and pseudo-wavefunction (dot-dashed and dashed lines) compared to their all-electron counterparts (solid lines) reproduced from Payne et al. 1992 [2]. $r_c$ is the core radius beyond which the potentials and wavefunctions for both methods agree.

2.4 Representation of the trilateration process used to find the impurity location (schematic). A, B and C mark the positions of the three nearest neighbour atoms after translation to the $xy$ plane, circles are great circles of spheres centred on the nearest neighbour atoms with radii equal to the dopant-neighbour bond length, and the dopant position, I, is at one of the intersection points of these three spheres which is found by solving equations 2.49.
3.1 GaAs band structure comparison between FHI pseudopotentials (solid lines) computed with ABINIT and the PW-MT-FHI pseudopotentials (empty circles) computed with QUANTUM ESPRESSO. For clarity only every third value was plotted for the PW-MT-FHI data. 

3.2 Total energy for the simulation of bulk GaAs for each of the pseudopotential types investigated. All systems are run with a $6 \times 6 \times 6$ k-point grid. Top: total energy convergence of pseudopotentials which have the $d$-orbitals in the core. Bottom: total energy convergence of pseudopotentials which have $d$-orbitals in the valence states. The black arrows indicate which axis to read the data values from for each pseudopotential.

3.3 Total energy convergence for the simulation of bulk GaAs for each of the pseudopotential types investigated. In each case the energy at each point is compared to a run performed at one step above the highest $E_{cut}$ shown in the plot. All systems are run with a $6 \times 6 \times 6$ k-point grid. Inset: data shown on log scales to show final levels of convergence. The horizontal line indicates the level at which we take the total energy to be converged. Solid lines correspond to pseudopotentials which have the $d$-orbitals in the core, dotted lines correspond to pseudopotentials which have $d$-orbitals in the valence states.

3.4 Comparison of the conduction band effective mass, band gap (calculated at $k=(0,0,0)$), and lattice constant values for different GaAs pseudopotentials. The lattice constant was calculated first and then used in the calculations for the other two parameters. In each case the horizontal line represents the experimental value of the associated parameter. The use of semi-core gallium pseudopotentials is indicated with the use of D. The band gap value of the PSPNC pseudopotential (marked with an *) is larger than the experimental as the pseudopotential incorrectly produces an indirect band gap for GaAs.
3.5 Side view (left), and a top down view (right), of the clean (110) gallium arsenide surface. Black circles represent arsenic atoms and white are gallium atoms. The buckling angle, \( \alpha \), is shown as well as the perpendicular displacement between the surface atoms \( \Delta_{1,\perp} \). The \( x, y \) and \( z \) labelling convention for the axes displayed here will be the standard throughout this work. The dimensions of the cell are \( x = 11.73 \, \text{Å} \), \( y = 16.59 \, \text{Å} \) and \( z = 25.42 \, \text{Å} \).

3.6 The total energy as a function of \( E_{\text{cut}} \) for a clean 14 atom (110) slab, for four different \( k \)-point configurations. The inset shows the fractional difference in the total energy at each step compared to the energy of the system run at the maximum value of \( E_{\text{cut}} = 18 \) Hartree.

4.1 \( 3 \times 3 \) super-cell. Left: side view of the cell showing the surface buckling angle, \( \alpha \). Right: plan view of the (110) surface. Black dots represent arsenic atoms and white dots represent gallium atoms. The red dot shows the isovalent group III impurity position and the blue dot shows the isovalent group V impurity position. The dimensions of the cell are \( x = 11.73 \, \text{Å} \), \( y = 16.59 \, \text{Å} \) and \( z = 25.42 \, \text{Å} \).

4.2 The relative error in the total energy of a doped GaAs slab as a function of \( E_{\text{cut}} \). The total energy is compared to a simulation with an \( E_{\text{cut}} \) of 18 Hartree. The total energies for B, Tl, N and Bi doped systems at 18 Hartree are -14.89, -15.33, -14.96 and -15.51 keV respectively.

4.3 The relative error in total energy as a function of \( k \)-point grid for four impurities from groups III and V. The total energy is compared to a simulation with an \( k \)-grid of \( 5 \times 5 \times 1 \).
4.4 The relative error in total energy as a function of $E_{\text{cut}}$ for the three different nitrogen positions in the slab. The total energy in each case is compared to a simulation with an $E_{\text{cut}}$ of 18 Hartree.  

4.5 Total energy per atom as a function of the number of atoms for systems with a single nitrogen impurity. The points are the results of the simulation and the line is the results of a least squares fitting of the data using the functional form given in equation 4.2.  

4.6 Comparison of the LDOS for different simulations of a nitrogen impurity in the surface layer. (a) is our standard system consisting of a $3 \times 3$ super-cell with a $E_{\text{cut}}$ of 15 Hartree and a $2 \times 2 \times 1$ k-point grid and the rest have had one of the variables increased. (b) is for a system with an $E_{\text{cut}}$ of 18 Hartree, (c) is for a system with a $3 \times 4$ super-cell and (d) is for a system with a $3 \times 3 \times 1$ k-point grid.  

4.7 Side-on view of the final relaxed positions of (a) the group III impurities substituting gallium and (b) the group V impurities substituting arsenic, as well as their first nearest neighbours. Impurity atoms are represented by filled shapes and the nearest neighbours atoms are represented by the corresponding empty shapes. Black lines have been added as a guide for the eye.  

4.8 Graphical representation of equation 4.3 showing the difference in position as a function of bond angle.  

4.9 Figure showing how the error in position varies depending on the bond angle. Three bond angles are shown, $0^\circ$ (triangles), $90^\circ$ (squares) and $180^\circ$ (circles). The original bond configuration is given by the solid blue lines and filled shapes, with the effect of increasing the bond lengths by 3.8% being shown by the red dashed line and empty shapes.
4.10 Comparison of constant height filled state STM images for 1st (a) 2nd (b) and 3rd (c) layer nitrogen impurities in the (110) surface of GaAs along with a diagram of the surface relaxation in each case ((d), (e) and (f)). In (a), (b) and (c) the position of the nitrogen has been marked with a white cross. In (d), (e) and (f) the gallium and arsenic atoms are shown along with the nitrogen impurity (black pentagon) and the positions of the GaAs atoms for a clean surface (empty circles). The dotted line marks the surface position. The STM images were generated at 4 Å above the surface. The dimensions of the cell are $x = 11.73$ Å, $y = 16.59$ Å and $z = 25.42$ Å.

4.11 Constant current height scan of the filled states for a nitrogen doped GaAs (110) surface taken from [3]. A, B and C are features due to nitrogen impurities in the first, second and third atomic layers.

4.12 Theoretical and experimental STM data for a 2nd layer nitrogen feature. (a) experimental STM data taken from figure 4.11 (the image has been flipped horizontally to line up with the geometry of the simulation), (b) repeat of the STM data shown in figure 4.10(b), (c) line scans of experimental data showing the relative peak heights with the dotted line being taken from (a) at $x \approx 1.1$ and the solid line taken from (a) at $x \approx 0.4$, (d) line scan of simulation data taken from (b) at $x = 0.37$, and (e) diagram of the DFT surface relaxation in the (110) plane. In (e) the dots are scaled depending on their distance away from the surface. The dimensions of the cell used in the DFT simulation are $x = 11.73$ Å, $y = 16.59$ Å and $z = 25.42$ Å.
4.13 Empty state calculated STM images of Group III impurities on the 1st layer of the GaAs (110) surface: boron, aluminium, gallium, indium and thallium. We see that the smallest atom, boron, leaves a dark region in the image due to it moving into the surface. The impurities bigger than gallium relax outwards and so give brighter spots. The similarity between the aluminium and gallium images comes from them having almost identical bond lengths and thus relaxations.

4.14 Simulated filled state STM images of Group III impurities on the 1st layer of the GaAs (110) surface: boron, aluminium, gallium, indium and thallium. For the boron system the nearest neighbour arsenic atoms relax towards each other due to the small size of the impurity. The larger indium and thallium impurities push the arsenic atoms away introducing a space near the impurity location. The relaxed position of the impurity in each simulation has been marked with a white cross.

4.15 Experimental STM images of group III dopants on the (110) surface of GaAs: indium empty state, (a); indium filled state (b); aluminium empty state, (c). Images (a) and (b) were reproduced from [4] and are of the same surface region with a doping density of In$_{0.04}$Ga$_{0.96}$As. In (a) the atomic contrast is primarily from the Ga atoms and in (b) the contrast is primarily from the As atoms. Image (c) was reproduced from [5]. The regions of dark contrast are caused by the aluminium atoms, and the areas of light contrast are from the Ga atoms. The doping concentration was Al$_{0.3}$Ga$_{0.7}$As.

4.16 Simulated filled state STM images of Group V impurities on the 1st layer of the GaAs (110) surface: nitrogen, phosphorus, arsenic, antimony and bismuth. The increasing size of the impurities directly relates to the change in the intensity of the image over the impurity site.
4.17 Simulated empty state STM images of Group V impurities on the 1st layer of the GaAs (110) surface: nitrogen, phosphorus, arsenic, antimony and bismuth. The impurity location in each case is marked with a cross.  

4.18 Experimental STM images of group V dopants on the (110) surface of GaAs: phosphorus filled state, (a); antimony filled state (b); antimony empty state, (c). Image (a) was reproduced from [6], the background contrast is from the arsenic atoms, and the dark regions bisected by the lines a and b are due to the phosphorus atoms. In images (b) and (c) the background contrast is from the arsenic and gallium atoms respectively, with the bright contrast due to the antimony atoms. The antimony images were reproduced from [7].

4.19 Simulated dI/dV curves taken over clean gallium and arsenic sites (top), a single bismuth impurity site (middle) and a nitrogen impurity (bottom).

4.20 Simulated dI/dV curves taken over a clean gallium site (blue dots), the nitrogen site (black dashes) and one of the nitrogen’s neighbouring gallium sites (green solid).

4.21 Comparison of constant height filled state STM images for 1st (a) 2nd (b) and 3rd (c) layer bismuth impurities in the (110) surface of InP along with a diagram of the surface relaxation in each case ((d), (e) and (f)). In (a), (b) and (c) the position of the bismuth has been marked with a cross. In (d), (e) and (f) the indium and phosphorus atoms are shown along with the bismuth impurity (black pentagon) and the positions of the InP atoms for a clean surface (empty circles). The dotted line marks the surface position. The dimensions of the cell are $x = 12.37 \, \text{Å}$, $y = 17.46 \, \text{Å}$ and $z = 26.75 \, \text{Å}$.

4.22 Constant current height scan of the filled states for a bismuth doped InP (110) surface [8]. A, B and C are features due to bismuth impurities in the first, second and third atomic layers.
5.1 The fractional change in total energy as a function of $E_{\text{cut}}$ for three different sized Si doped super-cells. The fractional difference in total energy is measured against a calculation with $E_{\text{cut}}^{\text{max}} = 18$ Hartree. The total energies for the 56, 126 and 168 atom super-cells at 18 Hartree are -6.65, -14.91 and -19.86 keV respectively. Notice the resemblance to figure 3.6.

5.2 Comparison of the convergence of the total energy of two different Si doped systems with respect to $k$-point grid. The convergence is measured as the fractional difference between the total energy at the current $k$-point grid compared to the total energy calculated using a $5 \times 5 \times 1$ $k$-grid, $\frac{E_{\text{total}}(k) - E_{\text{total}}(5 \times 5 \times 1)}{E_{\text{total}}(5 \times 5 \times 1)}$. Lines are shown as a guide for the eye.

5.3 Energy of the silicon band gap state for three different sized super-cells. Left: convergence of the eigenvalue of the state as a function of $E_{\text{cut}}$. The arrows indicate which axis to read for the datasets. The conduction band edge was at -0.5 eV. Right: change in energy of the band gap state as a function of $E_{\text{cut}}$, $|E_{\text{bgs}}(E_{\text{cut}}) - E_{\text{bgs}}(E_{\text{cut}} - \Delta)|$ where $\Delta$ is the 1 Hartree step size between $E_{\text{cut}}$ values.

5.4 Total energy per atom as a function of the number of atoms. The points are the results of the simulation and the lines are the results of a least squares fitting of the data using the functional form given in equation 4.2.

5.5 Si doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. (a) Varying the super-cell length in the $x$ direction while keeping the $y$ length fixed at 2, (b) same as (a) but with the $y$ length fixed at 3, (c) varying super-cell length in the $y$ direction with the $x$ length fixed at 2 and (d) same as (c) but with the $x$ length fixed at 3. In each case the dimensions of the super-cell are given. The horizontal dashed line marks the energy level in the largest, 210 atom, super-cell, which we take to be the converged energy.
5.6 C doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. See figure 5.5 for details. ........................................ 95

5.7 Ge doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. See figure 5.5 for details. ........................................ 96

5.8 Sn doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. See figure 5.5 for details. ........................................ 97

5.9 Comparison of the final relaxed positions of the carbon, silicon, germanium and tin donors as well as their first nearest neighbours. Donor atoms are represented by filled shapes and their nearest neighbour arsenic atoms are represented by the corresponding empty shapes. The initial substitutional gallium site is also marked for comparison. The bonds in the case of the silicon dopant have been shown for reference. The left hand plot is a slice taken through the slab in the (100) plane and the plot on the right is a plan view of the (110) surface. The inset in the left hand plot shows a close-up of the surface nearest neighbour positions and the inset in the right hand plot shows a close-up of the dopant positions. ........................................... 98

5.10 Comparison of the final relaxed positions of the group III (left) and group IV (right) dopants along with their first nearest neighbours. Dopant atoms are represented by filled shapes and their nearest neighbour atoms are represented by the corresponding empty shapes. The dotted lines show the directional trend of the relaxations in the group. The final relaxed positions of the group III dopants lie along a line at roughly 4° to the surface normal while the relaxed positions of the group IV donors lie along a line at roughly 15° to the surface normal. ................................................ 100
5.11 Dangling bond model of the relaxation (schematic). The small dots show the number of electrons in each dangling bond. (a) shows the GaAs (110) surface relaxation as the gallium (white) tries to flatten its bonds and the arsenic (black) tries to make them more tetrahedral. (b) shows how the dangling half filled bond of the silicon donor tries to form a tetrahedral configuration and (c) shows the sp$^2$+p hybridisation of the much smaller carbon donor. . . . . 102

5.12 LDOS plots taken in the (110) plane through the dopant location for (a) carbon, (b) silicon, (c) germanium and (d) tin in the 3×3 super-cell. The black dots show the dopant location, green triangles are the bonded arsenic atoms and green squares are the second nearest neighbour gallium atoms. (e)-(h) are slices in the (¯110) plane. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 106

5.13 Comparison of LDOS at the silicon location for (a) standard calculation, (b) HGH pseudopotential, (c) $E_{cut}=17$ Hartree, (d) 3×3×1 k-point grid and (e) 210 atom super-cell. Unless stated otherwise all other variables match the standard calculation. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 107

5.14 Line scan of the dopant LDOS along $z$, centred on the dopant site for five different simulations. Inset, data in the range 10 to 22Å shown on a log scale to show the vacuum tail of the density. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 108

5.15 LDOS images taken at a height of 4Å above the surface. Circles are the dopant atoms and triangles are the nearest neighbour arsenic atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 110

5.16 Line scan of the dopant LDOS along $z$, centred on the dopant site. Inset, data in the range 10 to 18Å shown on a log scale to show the convergence of the vacuum tail of the density. The surface position was taken to be at 12 Å and all STM scans were taken at 16Å. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 111
5.17 Filled state constant height STM images of the Group IV impurities on the 1st layer of the GaAs (110) surface. These images were generated at a height of at 4 Å above the surface. .......................................................... 112
5.18 Empty state constant height STM images of the Group IV impurities on the 1st layer of the GaAs (110) surface. These images were generated at a height of at 4 Å above the surface. .......................................................... 114

6.1 Convergence of the total energy with respect to $E_{cut}$ for the silicon acceptor in a 126 atom super-cell and a $2 \times 2 \times 1$ k-grid. .......................................................... 118
6.2 Comparison of the acceptor state of a silicon atom on the surface of GaAs. The left panel shows a simulation using a $3 \times 3$ super-cell and the right panel shows a simulation using a $3 \times 4$ super-cell. The position of the silicon atom is marked with a black dot. The energy difference between these two states is 5.17 meV. .......................................................... 119
6.3 Comparison of the final relaxed positions of the carbon, silicon, germanium and tin acceptors as well as the gallium atoms bonded to them. Dopant atoms are represented by filled shapes and the gallium atoms are represented by the corresponding empty shapes. The initial substitutional arsenic site is marked as a reference. .......................................................... 120
6.4 (a)-(d) are LDOS plots taken in the $xy$ surface plane through the acceptor location for carbon, silicon, germanium and tin. (e)-(h) are slices in the $yz$ plane. The black dots show the acceptor location, the green triangles are the bonded gallium atoms and the green squares are the second nearest neighbour arsenic atoms. .......................................................... 124
6.5 STM plots for (a) carbon, (b) silicon, (c) germanium and (d) tin acceptors taken at a height of 4Å above the surface. .......................................................... 125
6.6 Line scan of the acceptor dopant LDOS along $z$, centred on the dopant site. Insert, data in the range 11 to 18 Å shown on a log scale. 127

6.7 Convergence of the total energy versus $E_{\text{cut}}$ for each of the charged configurations of the silicon impurity. All simulations were run for a 126 atom super-cell with a $2 \times 2 \times 1$ k-grid. 128

6.8 Comparison of the empty state LDOS for the negatively charged silicon donor in (a) a 126 atom cell and (b) a 168 atom cell. The top panels show the surface plane and the bottom panels shows a side on view. The silicon position is marked with a black dot. 129

6.9 Comparison of dopant positions for negative (left), neutral (middle) and positive (right) charge states on the donor site. Dopant atoms are represented by filled shapes with the neighbouring arsenic atoms being represented by the corresponding empty shapes. The initial substitutional gallium site is also marked as a reference. 130

6.10 Comparison of dopant positions for negative (left), neutral (middle) and positive (right) acceptor charge states. Dopant atoms are filled shapes and the neighbouring gallium atoms are represented by the corresponding empty shapes. The initial substitutional arsenic site is marked as a reference. 131

6.11 Comparison of LDOS for all four elements in the case of negatively charged donors. (a) to (d) are slices through the dopants in the $xy$ plane and (e)-(h) are slices in the $yz$ plane. Going from left to right is carbon, silicon, germanium and tin. 135

6.12 Comparison of LDOS for all four elements in the case of positively charged donors. (a) to (d) are slices through the dopants in the $xy$ plane and (e)-(h) are slices in the $yz$ plane. Going from left to right is carbon, silicon, germanium and tin. 136
6.13 Comparison of LDOS for all four elements in the case of negatively charged acceptors. (a) to (d) are slices through the dopants in the \( xy \) plane and (e)-(h) are slices in the \( yz \) plane. Going from left to right is carbon, silicon, germanium and tin. 138

6.14 Comparison of LDOS for all four elements in the case of positively charged acceptors. (a) to (d) are slices through the dopants in the \( xy \) plane and (e)-(h) are slices in the \( yz \) plane. Going from left to right is carbon, silicon, germanium and tin. 139

6.15 LDOS images for the negatively charged donors taken 4 Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour arsenic atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin. 140

6.16 LDOS images for the positively charged donors taken 4 Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour arsenic atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin. 141

6.17 LDOS images for the negatively charged acceptors taken 4 Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour gallium atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin. 142

6.18 LDOS images for the positively charged acceptors taken 4 Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour gallium atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin. 143
List of Tables

4.1 Average dopant bond lengths for all group III and V atoms taken from the DFT simulations compared with bond lengths calculated from the covalent radii in Ref. [9]. .......................................................... 56

4.2 Comparison of the group V dopant positions and their three nearest neighbours between the DFT simulation and the trilateration calculation. In each case the first line is the dopant position and the following three lines are the nearest neighbour positions. For the trilateration positions the three nearest neighbour positions are the ones used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions. 59

4.3 Comparison between the DFT simulation and the trilateration calculation for the group III dopant and nearest neighbour positions. For the DFT positions the first line is the dopant position and the following three lines are the nearest neighbour positions. The DFT nearest neighbour positions are used as input for the trilateration calculations in this case. The difference is measured as the Euclidean distance between the two positions. ................................. 61
List of Figures

4.4 Comparison of the position of the bismuth dopant (first line) and its three nearest neighbours (following three lines) between the DFT simulation and the trilateration calculation. For the trilateration positions the three nearest neighbour positions are clean InP slab positions used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions. 79

5.1 Experimental covalent radii for group IV atoms along with the substituted gallium atom. All values are taken from [9]. 90

5.2 Data from the least squares fitting of the total energy data to the functional form given in equation 4.2. The units of $\alpha$ and $\alpha_{\text{donor}}$ are Hartrees per atom. 92

5.3 Average dopant bond lengths for all group IV donors taken from the DFT simulations compared with bond lengths calculated from the covalent radii in Ref. [9]. 103

5.4 Comparison of the group V dopant positions and their three nearest neighbours between the DFT simulation and the trilateration calculation. In each case the first line is the dopant position and the following three lines are the nearest neighbour positions. For the trilateration positions the three nearest neighbour positions are the ones used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions. 104

6.1 Average dopant bond lengths for all group IV acceptors taken from the DFT simulations compared with bond lengths calculated from the covalent radii in Ref. [9]. 121
6.2 Comparison of the group IV acceptor positions and their three nearest neighbours between the DFT simulation and the trilateration calculation. In each case the first line is the dopant position and the following three lines are the nearest neighbour positions. For the trilateration positions the three nearest neighbour positions are the ones used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions. .......................................................... 122

6.3 Absolute displacement from the neutral dopant position given in Å. ........ 132

6.4 Comparison of silicon surface and sub-surface bond lengths data for negatively, neutral and positively charged cases, as well as their height above the sub-surface layer. All values are in Å. .................. 133
Acknowledgements

I would like to thank my supervisors Dr Mervyn Roy and Professor Peter Maksym for their guidance and infinite patience, my wife Claire for her constant support and my fellow PhD students Peter Hague, John Coxon and Steph Jinks for their encouragement and reliable coffee breaks.
List of Publications


Chapter 1

Introduction

1.1 Introduction

The physics of single dopants on semiconductors is a rapidly growing field in condensed matter physics, and the range of possible semiconductor-dopant combinations potentially covers a wide area of possible applications. Gallium arsenide (GaAs) is a popular semiconductor for doping, however most of the existing work has focused on a few systems of interest, such as the silicon donor state [10, 11, 12, 13] or the magnetic properties of manganese [14, 15, 16]. In contrast to existing work, which has tended to focus on specific dopants in specific host materials, we present a systematic study of 14 dopants, comprising almost the entirety of groups III, IV and V of the periodic table.

In this thesis we perform a range of highly accurate density functional theory (DFT) calculations of a GaAs (110) slab containing different isolated near-surface single atomic impurities. We are able to show very clear trends in the final relaxed positions of a wide range of impurities on the GaAs (110) surface, with the relaxed dopant positions being determined by the dopants covalent radius. We also show trends in the contrast of scanning tunnelling microscopy (STM) images of the
surfaces, which are also determined by the dopant’s covalent radius. The relaxed position and STM contrast trends presented in this work fit very well with existing experimental and theoretical analysis for the systems that have been studied extensively, but also fill in the gaps for those systems that have not received the same level of investigation. We show that the analysis presented in this work applies equally across three groups of impurities from the periodic table, groups III, IV and V, covering isovalent, donor and acceptor cases.

The study of a wide range of near-surface impurities is of great importance to the continuing advancement of semiconductor devices. Because of the fixed scale of atoms, the presence of impurities will become a larger and larger concern as device miniaturisation continues. In 1990 a single impurity atom would have made up around $1 \times 10^{-10}\%$ by volume of a transistor structure on a computer chip; in 2015 a single atom takes up around $1 \times 10^{-4}\%$ of the structure by volume, and as the devices continue to shrink, this percentage can only grow. Not only will an understanding of the physics of devices where dopants make up a increasing percentage of the structure be vital, but so too will the understanding of the changing environment that miniaturisation will bring. As devices shrink, a larger fraction of their structure becomes surface and so impurities that were previously in the bulk will more frequently find themselves in a very different geometrical environment with potentially very different effects on the overall properties of the device. The addition of a surface to a structure causes relaxation and reconstruction that all alter the physical environment of any donor in the region, and it is these effects that will have to be understood as device miniaturisation continues.

Solotronic devices [17] represent the pinnacle of the semiconductor device miniaturisation. These are devices whose behaviour and operation is determined by single atomic impurities. There are a wide range of potential uses of solotronic
Chapter 1

devices, such as light sources [18], transistors [19] and quantum computers [20] and an understanding of the physics of impurities is key to advancing this technology. For solotronic devices the tailored use of a specific dopant is vital to obtaining the desired operational behaviour, and so a detailed knowledge of the behaviour of different dopants is crucial. As mentioned previously however, existing work is often focused on specific systems rather than systematic studies of a wider section of dopants.

A method that has been used extensively to study surface and near surface dopants is cross-sectional scanning tunnelling microscopy (XSTM). This technique provides very detailed images of the doped surface, however it is difficult to disentangle the contributions to the image from physical and electronic effects (see section 1.2). In this thesis we look in detail at simulated STM images of the different dopants and provide a simple explanation for the contrasts seen in all the systems studied in this work.

We begin with DFT simulations of each of the ten isovalent dopants from groups III and V of the periodic table on the GaAs (110) surface. We then extend our simulations to the four dopants from group IV, first as donors then as acceptors before finally looking at the more complicated systems of the charged cases for the group IV donor and acceptors. For each system we look in detail at the relaxation of the dopant and calculate STM images of the surface. We develop a simple geometrical model to explain the results of the DFT simulations and find that the covalent radii of the dopants determines the relaxed geometry of the system, which in turn determines the STM contrast of the surface. Using our geometrical model and theoretical values of the dopant covalent radii we are able to interpret the STM images and relaxation, as well as obtain accurate estimations, to within an average of 0.19 Å, of the final relaxed positions of the dopants without any input from the DFT simulation.
1.2 Scanning Tunnelling Microscopy

A major breakthrough in the study and observation of semiconductors’ surfaces and impurities came in 1981 with the invention of the scanning tunnelling microscope (STM). This device created by Gerd Binnig and Heinrich Rohrer won them the Nobel Prize in Physics in 1986 and has become an integral part of the field. Before STM the main experimental techniques available worked on a macroscopic level, giving you information about the bulk properties of a material and generally washing out any small scale effects caused by dopants or impurities. Scanning tunnelling microscopy allowed for atomic scale measurements and resolution, giving very detailed information on an impurity’s electrostatic and structural properties.

A scanning tunnelling microscope is a purely quantum device that works due to electron tunnelling. The device consists of a metallic tip that is brought close to an atomic surface in a vacuum. The vacuum is a potential barrier between the tip and surface that would forbid classical conduction of electrons, however electrons can quantum mechanically tunnel through. The electron states in the surface and the tip both decay outwards into the vacuum, and these states can overlap allowing the electrons to tunnel between the tip and the surface. The theory behind the STM was published by Tersoff and Hamann in 1983. They found that the tunnelling current was proportional to the local density of states of the surface, which is dependant on both the applied voltage and the tip position, and depends exponentially on the distance between the tip and the surface [21]. This allowed the results of simulations to be directly compared to STM experiments. If the tip is moved across the surface the resulting changes in the measured tunnelling current can be used to build up an image of the surface structure. By applying a bias voltage to the surface or the tip the tunnelling can be controlled. With
a positive sample bias the conduction band states are imaged as electron tunnel from the tip into the empty states of the sample. When using a negative sample bias the valence band states are imaged as electrons now tunnel from the filled states of the sample to the tip.

A scanning tunnelling microscope can also be used to perform Scanning tunnelling spectroscopy (STS) measurements. For these the tip is kept fixed at a specific location above the surface, and the tunnelling current is measured as a function of the applied voltage. The slope of the I-V curve obtained by these measurements gives the electron density at the position of the tip, the local density of states.

The XSTM image of a GaAs (110) surface with impurities was measured in 1985 by Feenstra and Fein [22]. They imaged both n-type GaAs doped with silicon and p-type GaAs which was doped with zinc. From their images they were able to get a relatively good agreement of the surface structure compared to the accepted values. They found a lattice constant of 6.0 Å and a $1 \times 1$ surface periodicity of $6.5 \, \text{Å} \times 5.0 \, \text{Å}$, compared to the accepted values of $5.65 \, \text{Å}$ and $5.65 \, \text{Å} \times 4.00 \, \text{Å}$ respectively. In the p-type GaAs images they discovered defects, and due to the alignment of the clean atomic rows in the rest of the image they deduced that these were point defects rather than lattice dislocations. These defects were only seen in the p-type samples but they were not confident enough to attribute the source of the defects to the zinc doping.

In 1993 Feenstra et al extended their work on the GaAs (110) surface to study bulk impurities with STM, specifically arsenic antisite defects [23]. The difficulty of studying bulk impurities at a surface comes in part from any reconstruction that occurs when the bulk system is cleaved to form a surface. This is mitigated slightly by using the (110) surface of GaAs which undergoes no reconstruction. They used STS to identify the band gap and the impurity states within the gap,
and then used STM to image these states. The position in the band gap of the impurity state pointed to it being an arsenic antisite defect which they then imaged using STM, showing a structure that was consistent with the antisite arsenic.

In 1993 Wang et al published theoretical STM images for a silicon doped GaAs (110) surface [12]. They used the Tersoff Hamann theory in conjunction with a DFT pseudopotential calculation to compare simulated STM images of the clean (110) surface and a surface with a single silicon dopant. By looking at a range of applied voltages they saw the localised donor state of the silicon, its effect on the empty gallium states in the conduction band and its effect on the filled valence band states of the arsenic atoms. For their valence band imaging they used a -1.9 eV bias voltage and saw an image relatively unchanged from the clean surface apart from two regions of increased intensity on the neighbouring arsenic atom, which was attributed to charge leaking from the extra electron on the silicon donor. For the empty state conduction band imaging they used a +1.9 eV bias voltage and saw the usual gallium sub lattice with a large dark region over the silicon site. Because of the extra electron, the silicon is not equivalent to the surrounding gallium sites and is seen as having a weaker tunnelling probability. They also calculated an image for a small voltage window around the top of the valence band to specifically pick out the localised silicon donor state. This, they showed, was located on the silicon site with a small amount of density leaking over to the neighbouring arsenic atoms.

A year later Zheng et al performed STM experiments on silicon doped GaAs [13] inspired by Wang's work. By performing STM measurements on the (110) surface of a doped GaAs sample they were able to image silicon donors in the first five layers of the surface. For the surface layer silicon dopants they observed localised features spreading to only a couple of nearest neighbour atoms. These features are
caused by the dangling bond of the silicon which introduces a gap state predicted by Wang et al. [12]. For the four layers of sub-surface dopants they found a delocalised feature that spread around 25 Å from the dopant site, consistent with the Coulomb potential of silicon. They found the surface concentration was proportional to the bulk doping concentration of silicon, and that the count level on each layer was consistent, showing that the surface imaging was representative of the bulk doping.

Duan, Peressi and Baroni calculated XSTM images of nitrogen and indium substitutional impurities on the (110) surface of GaAs [24]. They were interested in characterising different impurity positions and impurity types based on STM images, and showed that different situations could be differentiated by using a combination of positive and negative bias images. They explained the features in the STM images using the relaxed positions of the nitrogen and indium atoms, clearly showing the different STM contrasts produced by different combinations of impurities. In chapter 4 of this thesis we give a detailed explanation of the source of the relaxation and how it leads to the different contrasts seen, not only for nitrogen and indium but for all group III and V isovalent impurities.

In 2011 Yi, Ma and Rohlfing performed LDA DFT calculations of a silicon donor on the GaAs (110) surface with specific interest in the DX center behaviour of the charged donor state [11]. They show that, as in previous work, the DX centre in bulk GaAs is metastable, and then go on to show that the DX center is unstable on the (110) surface for both the neutral and positive charge states, whereas the negative DX center is stable on the surface. For each of the surface cases positive and negative voltage STM images are calculated, which agree with experimental results. During the analysis they present the geometry and charge density distribution for the positive, negative and neutral silicon donors, however their analysis of the cause of these features does not go far beyond the Coulomb
interaction between the silicon and its nearest arsenic atom. In chapters 5 and 6 of this thesis we present our explanation for both the silicon relaxation and charge density, and compare in detail to the results of Yi, Ma and Rohlfing.

Although the use of XSTM to investigate impurities is now well established there are still limitations that can lead to uncertainties as to the interpretation of the results. One of the main limitations comes in disentangling geometrical versus electronic causes of features in the images. STM measures an energy projected part of the charge density of the surface, and so different heights or contrasts in the surface images are usually attributed to electronic structure differences such as surface density of states, band bending or changing band gap. However, because the positions of the atoms dictate the charge density and bond directions, the geometry of the system plays a large role in the appearance of the STM images [25].

Another limiting factor on the understanding of the STM results is that the images only really give you an interpretation of what is happening at the outermost surface layer. All the details from different depths or energies in the surface are “compressed” to a single number at each position on the surface. Without a detailed knowledge of the underlying structure, or by performing more detailed experimentation such as scanning tunnelling spectroscopy (STS), it becomes almost impossible to know whether to attribute STM features to the surface or to some underlying perturbation in the crystal. In chapter 3 we present experimental and theoretical results for STM surface features which are caused by sub-surface dopants.
1.3 Synopsis

In this thesis we present detailed DFT calculations of dopants on the GaAs (110) surface and develop a new geometrical model for understanding the physics of a wide range of dopants. We find that the covalent radii of the dopants determines the relaxed geometry of the system, with dopants that have larger covalent radii relaxing out of the surface and dopants with smaller covalent radii relaxing into the surface. It is this relaxation that in turn determines the STM contrast of the surface, with the dopants that have relaxed further out producing a bright contrast in the STM images over the dopant site and the dopants that relax in producing a dark contrast. The range of electronic behaviour provided by the isovalent, donor and acceptor dopants allows us to test our calculations and geometrical model in a variety of situations to establish their robustness, and the scope of dopants studied provides us with a number of cases that have been studied experimentally to compare to. We find excellent agreement when applying this model to well established cases that have been studied both theoretically and experimentally, such as isolated nitrogen and silicon dopants on the GaAs (110) surface. For those systems where there is no experimental data available we make detailed predictions to guide future experiments.

In chapter 2 we present the theory behind density functional theory and pseudopotentials, both of which are heavily used throughout this work, as well as the theory behind scanning tunnelling microscopy which is used to generate the STM images of our systems that can be compared to the results of experiments. In this chapter we also present a mathematical technique called trilateration, that we use in subsequent chapters to obtain accurate estimates of the relaxed positions of all the donors studied in this work.

Chapter 3 presents the basic ideas and simulations for our work that the
subsequent chapters are built upon. This chapter contains our analysis of the available pseudopotentials as well as the details of the construction of the GaAs (110) slab to which all the impurities are added. We also outline our standard set of convergence checks for the clean GaAs surface. These checks are then carried through to subsequent chapters as the different impurities are studied.

In chapter 4 we present our simulations and analysis of group III and group V isovalent impurities on the GaAs surface, and compare these to the results of recent STM experiments. We look in detail at the case of nitrogen in the first three surface layers and find excellent agreement between our geometrical interpretation and the DFT and experimental results, explaining positions of the features in the contrast that agree to within 2% of a lattice constant. We also check our model by simulating STM images for a single bismuth dopant on an InP (110) surface and again find excellent agreement with experimental results, indicating that our interpretation may extend beyond the GaAs (110) system.

Chapter 5 contains the work on group IV dopants on the GaAs (110) surface. We extend the analysis developed for the isovalent case to four donors from group IV. For these donor systems the localised donor state is of particular interest and was the main focus for both the convergence of the system and the subsequent analysis of the local density of states (LDOS) and STM images. By approaching the donor systems from our geometrical point of view, not only do we understand the behaviour of the relaxation and STM contrast, as in the isovalent case, but we are also able to provide an explanation for the spatial distributions of the localised donor states.

In chapter 6 we expand on the previous group IV work by looking at the group IV dopants as acceptor impurities as well as the charged cases for the group IV donors and acceptors. The addition or subtraction of an electron for the charge states caused a change in the bond lengths of the dopant, however our geometrical
explanation still provided good qualitative descriptions of the relaxations of the systems as well as of the local density of states. This implies that it is primarily the covalent radius that controls the relaxation, not the charge.

Finally chapter 7 contains the summary and conclusions for the work presented in this thesis and contains some suggestions for future work.
Chapter 2

Theory

2.1 Introduction

In this chapter we present the theory behind the different techniques used throughout this work in the simulation of doped surfaces. The bulk of the chapter is concerned with density functional theory (DFT) and the local density approximation (LDA) which are used for all the simulations in this work. We also present the details of the pseudopotential approximation which simplifies the DFT calculations, allowing us to run more complicated calculations than would be possible using all-electron calculations. In this thesis, where possible, we compare our calculations to the results of STM experiments. In this chapter we therefore outline the Tersoff-Hamann theory which is used to calculate STM images from the results of our DFT simulations. We end this chapter by presenting a new geometrical model we have developed for estimating the relaxed positions of impurities on the surface. This model is used throughout the work, with the details of the calculations given here.

In section 2.2 we outline the theory behind the DFT calculations, and also discuss the accuracy of DFT. In 2.3 we discuss the theory behind
pseudopotentials. In section 2.4 we provide details about the ABINIT software used to perform the DFT calculations and in section 2.5 we present the theory behind the scanning tunnelling microscope (STM) and scanning tunnelling spectroscopy (STS) techniques used to calculate images of the doped surfaces. Finally in section 2.6 we present a new geometrical method for estimating impurity positions.

2.2 Density Functional Theory Calculations

Density functional theory (DFT) is a quantum mechanical framework for calculating the properties of many-body systems of atoms and molecules. The theory is built on the idea that the properties of these many-body systems can be found through functionals of the ground state electron density. The theory has its beginning in the 1960’s with the work of Hohenberg and Kohn, which was expanded upon by Kohn and Sham in 1965 and Levy in 1979. In the 1990’s improvements in pseudopotentials and local density approximations gave DFT a resurgence, which was coupled with an increase in computing power meaning complicated systems, previously thought too computationally difficult to be modelled, could be simulated with relative ease and accuracy.

2.2.1 Hohenberg-Kohn DFT Theory

In their seminal 1964 paper, Hohenberg and Kohn showed that for a system of interacting particles in an external potential, \( V_{ext} \), there exists a ground state density that is uniquely determined by \( V_{ext} \) [26]. The starting point for their *reductio ad absurdum* proof is the assumption that the same ground state density can arise from two different external potentials that differ by more than just a
constant. They define two systems,

\[ \hat{H}' \Psi' = E' \Psi' \]  \hspace{1cm} (2.1)

\[ \hat{H} \Psi = E \Psi, \]  \hspace{1cm} (2.2)

where \( \hat{H} \) is the Hamiltonian operator and \( E \) is the energy. \( \hat{H}' \) and \( \hat{H} \) differ only in the form of the external potential,

\[ \hat{H}' = \hat{T} + \hat{U} + \hat{V}'_{\text{ext}} \]  \hspace{1cm} (2.3)

\[ \hat{H} = \hat{T} + \hat{U} + \hat{V}_{\text{ext}}. \]  \hspace{1cm} (2.4)

Here \( \hat{T} \) is the kinetic energy, \( \hat{U} \) is the internal energy of the system and \( \hat{V}'_{\text{ext}} \neq \hat{V}_{\text{ext}} + \text{constant} \). Since these are two different Hamiltonians there will be a different ground state wavefunction that satisfies each of equations 2.1 and 2.2. The variational principle states that the ground state wavefunction of a system will minimise the energy of that system, i.e.

\[ E = \langle \Psi | \hat{H} | \Psi \rangle < \langle \Psi' | \hat{H} | \Psi' \rangle. \]  \hspace{1cm} (2.5)

Expanding out the far right hand side, using the fact that the Hamiltonians only differ in the external potential, gives

\[ E < \langle \Psi' | \hat{H}' - \hat{V}'_{\text{ext}} + \hat{V}_{\text{ext}} | \Psi' \rangle \]  \hspace{1cm} (2.6)

\[ < E' + \int [V_{\text{ext}}(r) - V'_{\text{ext}}(r)] n(r) d^3r, \]  \hspace{1cm} (2.7)

where \( n(r) \) is the ground state density and \( \langle \Psi | \hat{V}_{\text{ext}} | \Psi \rangle = \int V_{\text{ext}} n(r) d^3r \). The same reasoning can be followed through but with the primed and unprimed quantities swapped giving,

\[ E' < E + \int [V'_{\text{ext}}(r) - V_{\text{ext}}(r)] n(r) d^3r, \]  \hspace{1cm} (2.8)

where again \( n(r) \) is the ground state density, the same one as in equation 2.7. Combining equations 2.7 and 2.8 gives the result

\[ E' + E < E + E'. \]  \hspace{1cm} (2.9)
This result shows that the initial assumption must be incorrect as it leads to a contradiction, which proves that $V_{ext}$ must be a unique functional of $n(r)$. Since $V_{ext}$ fixes the form of $\hat{H}$ then the full many particle ground state is also a functional of the density.

The second theory put forward by Hohenberg and Kohn states that the full many particle ground state energy is a functional of the density. The exact ground state density of the system will be the one that minimises the energy functional of the density. The Hamiltonian operator for a set of $N$ interacting particles can be written as

$$\hat{H} = \hat{T} + \hat{V}_{ext} + \hat{U} + E_{II}$$  \hspace{1cm} (2.10)

where $E_{II}$ is classical interaction between the nuclei, $\hat{T}$ is the kinetic energy, $\hat{V}_{ext}$ is the potential energy due to an external field and $\hat{U}$ is the internal energy from electron-electron interactions;

$$\hat{T} = \sum_{i}^{N} \left(-\frac{1}{2}\nabla_{i}^{2}\right)$$, \hspace{1cm} (2.11)

$$\hat{V}_{ext} = \sum_{i}^{N} V(r_{i})$$, \hspace{1cm} (2.12)

$$\hat{U} = \frac{1}{2} \sum_{i \neq j}^{N} \frac{1}{|r_{i} - r_{j}|}$$, \hspace{1cm} (2.13)

in atomic units. Because the kinetic and electron-electron interaction terms of the Hamiltonian are determined by $n(r)$, Hohenberg and Kohn combined them into a single functional,

$$\hat{F}_{HK}[n(r)] = \langle \Psi | \hat{T}[n(r)] + \hat{U}[n(r)] | \Psi \rangle$$ \hspace{1cm} (2.14)

which allows the total energy of the system to be written as

$$E_{HK}[n(r)] = \hat{F}_{HK}[n(r)] + \int V_{ext}(r)n(r)d^{3}r + E_{II}$$ \hspace{1cm} (2.15)
$F_{HK}[n(r)]$ is a universal functional as the calculation of the kinetic and internal energies is the same for all systems, $V_{ext}$ on the other hand is system dependent. The energy of the system will be minimised for the ground state wavefunction, $\Psi$. Any other wavefunction, $\Psi'$, will lead to a higher energy system,

$$E[n(r)] = \langle \Psi | \hat{H} | \Psi \rangle > \langle \Psi' | \hat{H} | \Psi' \rangle.$$  \hspace{1cm} (2.16)

Thus the energy of the system will be lower for the correct ground state density, $n(r) = \int \Psi^* \Psi(\mathbf{r}) d^3 \mathbf{r}$, than for any other choice of density.

These two theorems, stating the external potential and the energy of the system in terms of functionals of the density form the groundwork for DFT.

### 2.2.2 N-Representable Density

One of the issues with the Hohenburg-Kohn formalism is that it is restricted to only dealing with non degenerate systems as it is not possible to prove, using the Hohenburg-Kohn theorems, that the density is uniquely determined by the external potential if there is more than one wavefunction corresponding to the ground state. The solution to this problem was found by Mel Levy and required the use of an $N$-representable density [27]. $N$-representable densities are all the densities that can be constructed from $N$ fermionic wavefunctions, which now allows for the possibility of degenerate ground state systems. Levy defined a new universal part for the energy functional to replace $F_{HK}[n(r)]$, defined as

$$F_{LL}[n(r)] = \min_{\Psi_i \rightarrow n(r)} \langle \Psi_i | \hat{T} + \hat{U} | \Psi_i \rangle,$$  \hspace{1cm} (2.17)

where a constrained minimisation is performed over all wavefunctions, $\Psi_i$, that give the same density, $n(r)$. The energy functional of the system can now be written as

$$E_{LL}[n(r)] = F_{LL}[n(r)] + \int V_{ext}(\mathbf{r}) n(\mathbf{r}) d^3 \mathbf{r} + E_H.$$  \hspace{1cm} (2.18)
At the minimum of the total energy we have $F_{LL} = F_{HK}$. This gives a more meaningful definition of the Hohenburg-Kohn functional (equation 2.15). The functional is now given as the minimum of the kinetic and potential energies for all wavefunctions with the chosen density $n(r)$, and is no longer limited to non-degenerate systems but instead is defined for any density for a system of $N$ electrons. It is possible to construct a wavefunction, $\Psi_N$, that can describe any density, $n(r)$, for a system of $N$ electrons; any finite non-negative differentiable function can give an $N$-representable density [28]. However the densities generated by an external potential required by the HK functions are in general not known.

### 2.2.3 Kohn-Sham Ansatz

Although the Hohenburg-Kohn theories provide the link between the ground state density and the external potential they do not provide any information on what the form of the ground state density should be. The only way to find it would be to solve the Schrödinger equation for the full many-body interacting system. A potential solution to this was proposed by Kohn and Sham and involves replacing the full many-body interacting ground state density with a system of fictitious non-interacting particles that give the same ground state density [29]. This leads to a system of independent particle equations which can be easily solved with numerical techniques.

In the Kohn-Sham formalism the independent particle kinetic energy is defined as

$$T_s[n(r)] = -\frac{1}{2} \sum_{i=1}^{N} |\nabla \psi_i|^2,$$

and the classical Coulomb interaction caused by the electron density self
interacting, called the Hartree energy, is defined as

$$E_H[n(r)] = \frac{1}{2} \int \frac{n(r)n(r')}{|r - r'|} d^3r d^3r'. \quad (2.20)$$

The energy functional can now be written as

$$E_{KS}[n(r)] = T_s[n(r)] + E_H[n(r)] + E_{II} + \int V_{ext}(r)n(r)dr + E_{XC}[n(r)] \quad (2.21)$$

where $E_{XC}[n(r)]$ is the exchange-correlation functional which incorporates all of the complicated many-body interaction effects. If the HK energy functional (equation 2.15) is compared to the KS functional we can write

$$E_{XC}[n(r)] = F_{HK}[n(r)] - T_s[n(r)] - E_H[n(r)], \quad (2.22)$$

and expanding out $F_{HK}[n(r)]$ using equation 2.14 gives

$$E_{XC}[n(r)] = (T[n(r)] - T_s[n(r)]) + (U[n(r)] - E_H[n(r)]). \quad (2.23)$$

This shows that the exchange-correlation functional is the difference of the kinetic and internal energies of the true interacting system from the fictitious non-interacting system.

The relationship between the Kohn-Sham theories and the Hohenberg-Kohn theories is depicted in figure 2.1. Single arrows represent the link between properties found by solving the Schrödinger equation, the double line arrows show the completion of the loop made by the Hohenburg-Kohn theory that relates the ground state density to the external potential and the double headed arrow in the middle shows how the Kohn-Sham theory links the many-body interacting system to the non-interacting system through the ground state density.

### 2.2.4 LDA

As well as presenting the non-interacting density model, the paper by Kohn and Sham also proposed the local density approximation (LDA). This LDA assumes
\begin{equation}
V_{\text{ext}}(r) \overset{HK}{\leftrightarrow} n_0(r) \overset{\text{ansatz}}{\leftrightarrow} n_0'(r) \overset{HK}{\leftrightarrow} V_{KS}(r)
\end{equation}
\begin{align*}
\downarrow & \quad \uparrow & \quad \uparrow & \quad \downarrow \\
\Psi_i(r) & \rightarrow \Psi_0(r) & \Psi_{i=1,N_s}(r) & \leftarrow \Psi_i(r)
\end{align*}

Figure 2.1: Single arrows represent steps made by solving the Schrödinger equation, the double line arrows represent the Hohenburg-Kohn theory that relates the ground state density to the external potential and the double headed arrow is the Kohn-Sham ansatz that links a many-body interacting density to a non-interacting one. Reproduced from [1].

that the exchange-correlation energy per electron at a position \( r \), \( \epsilon_{XC}(r) \), is equal to the exchange-correlation energy of a uniform electron gas with a density \( n(r) \),

\begin{equation}
E_{XC}[n(r)] = \int \epsilon_{XC}[n(r)] n(r) d^3r.
\end{equation}

From this the exchange-correlation potential can be found as

\begin{equation}
V_{XC}(r) = \frac{\delta E_{XC}[n(r)]}{\delta n(r)} = \epsilon_{XC}[n(r)] + n(r) \frac{\partial \epsilon_{XC}[n(r)]}{\partial n(r)}.
\end{equation}

The exchange part of the potential is known exactly [30]

\begin{equation}
\epsilon_X[n(r)] = -\frac{3}{4} \left( \frac{3}{\pi} \right)^{\frac{1}{3}} n(r)^{\frac{1}{3}},
\end{equation}

however for the correlation part exact analytic expressions are only known in the high and low density limit [31]

\begin{align*}
\epsilon_C &= c_0 \ln(r_s) - c_1 + c_2 r_s \ln(r_s) - c_3 r_s & \text{for high density} \\
\epsilon_C &= -\frac{d_0}{r_s} + \frac{d_1}{r_s^2} & \text{for low density}
\end{align*}

where \( \frac{4}{3} \pi r_s^3 = n(r)^{-1} \). For any density between these two limits the correlation can be estimated using Monte Carlo simulations [32] and there now exist many different forms of correlation potentials for use in LDA calculations, such as Vosko-Wilk-Nusair (VWN) [33], Perdew-Zunger (PZ81) [34], Cole-Perdew (CP)
and the Perdew-Wang (PW92) functional [31] which is the potential used by our choice of pseudopotential. One of the main competitors to LDA is the generalised gradient approximation (GGA). For GGA the gradient of the density is also taken into account,

\[ E_{XC}^{GGA}[n(r)] = \int \epsilon_{XC}^{GGA}[n(r), \nabla n(r)]n(r)d^3r. \] (2.28)

This method can be more accurate than LDA in some cases, however unlike LDA there is no unique way of constructing the functionals. Neither LDA or GGA are able to accurately reproduce bond lengths, with LDA underestimating and GGA overestimating, however the LDA bond lengths are generally closer while the bond length errors of GGA are generally larger and less consistent [36]. Since we are looking in detail at the relaxation geometries and bond lengths of the different dopants we chose to use LDA over GGA.

### 2.2.5 Self-Consistency

The Schrödinger equation for the collection of independent single particle wavefunctions used to approximate the interacting system is solved self-consistently using the following procedure.

First an initial guess is made for the ground state density, \( n_{in}(r) \). Using this density an effective Kohn-Sham potential is built,

\[ V_{KS}(r) = V_H[n_{in}(r)] + V_{XC}[n_{in}(r)] + V_{ext}(r) + V_{II}(r), \] (2.29)

where \( V_{II} \) is the potential from the nuclei, \( V_{XC} \) is given by equation 2.25 and

\[ V_{Hartree}[n(r)] = \int \frac{n(r')}{|r-r'|}d^3r'. \] (2.30)

Using this potential, a set of independent particle Schrödinger equations are constructed and solved for the ground state wavefunctions and eigenvalues,

\[ \left[ -\frac{1}{2} \nabla^2 + V_{KS}(r) \right] \psi_i(r) = \varepsilon_i \psi_i(r). \] (2.31)
Estimate ground state density, $n(r)_{in}$

Calculate potential

Solve the Kohn-Sham equations

Calculate the ground state density, $n(r)_{out}$

Mix $n(r)_{in}$ and $n(r)_{out}$ to form a new $n(r)_{in}$

Does $n(r)_{in} = n(r)_{out}$?

No

Yes

Output results

Figure 2.2: Flowchart showing the steps taken to obtain a self-consistent Kohn-Sham density.

From these ground state wavefunctions the ground state density of this system can be calculated

$$n_{out}(r) = \sum_{i}^{N} |\psi_{i}(r)|^2,$$

which is then checked against $n_{in}(r)$. If the two densities are consistent then the procedure is finished, if not then a new input density is constructed usually by mixing the old $n_{in}(r)$ and the new $n_{out}(r)$ and the process is started again. A flowchart outlining these steps is shown in figure 2.2.

### 2.2.6 Accuracy of the Functionals

Despite its enormous success there are a number of issues that arise from DFT calculations. As a ground state calculation, no real reliability can be attributed to the energies of any of the conduction band states and in general there is no reason to expect that any of the single particle eigenvalues of a DFT calculation will equate to the true energy states of the system being modelled [37]. The one exception to this is that the highest occupied state in a finite system has
a physical meaning: its energy is the negative of the ionisation potential and controls the decay of the density into the vacuum [38, 39]. In practice however, the results from DFT calculations have proven to be so accurate that it is safe to compare them to physical systems [40].

Another area of concern with DFT is its tendency to underestimate the band gap of semiconductors. There are a number of methods of adjusting the band energies to correct for the underestimation of the band gap, however their validity and appropriateness is a matter of strong debate. Common correction methods include LDA+U [41, 42], which adds a Hubbard-like potential to the density functional to take into account large, highly localised effects that can arise from $d$ or $f$ electrons, and the GW approximation that takes into account the self-energy [43]. Recent work by Lany & Zunger [44] investigated adjustments to the band energies for systems including impurities and concluded that the band gap and the band edge energies are more affected by the failures of LDA than the gap impurity state, and as such, only adjustment to band edge energies are required to correct the calculation. They argue that the impurity levels themselves can be taken without alteration. It is still a matter of much discussion and work on it has been done as recently as 2013 by P. A. Schultz and A. H. Edwards [45]. By investigating defects in bulk silicon and gallium arsenide they show that a DFT LDA calculation can provide very accurate energies for defect levels completely independently of band gap, with systems that have Kohn-Sham gaps ranging from 0.1 to 1.1 eV giving consistent results for defect levels.

The primary concern in our investigation is the character of the impurity state and how it affects the STM images. Because of this no corrections to the eigenvalues were implemented.
2.3 Pseudopotentials

In a DFT calculation a basis set of functions is chosen to represent the wavefunctions of the non-interacting pseudo-electron in the system, with a common choice being a plane wave basis set. The number of plane waves included in the basis set is determined by a kinetic energy cut-off value, $E_{\text{cut}}$, which determines a maximum k-vector in reciprocal space defined by

$$E_{\text{cut}} > \frac{(k + G)^2}{2},$$  \hspace{1cm} (2.33)

where $G$ is a reciprocal space lattice vector. In order to describe features in real space with plane waves, a minimum wavelength is needed which is smaller than the smallest feature of the system which translates to a large maximum wave vector in reciprocal space. This means the larger the value of $E_{\text{cut}}$, the more the basis set will be able to accurately describe the details of the physical system. Atoms have very highly localised and highly oscillatory features in their core regions which require a very large number of plane waves to be described accurately. This large number of plane waves greatly increases the calculation time and thus limits what can reasonably be simulated. The solution to this problem comes in the form of pseudopotentials.

The idea behind pseudopotentials is that the core electrons of an atom do not play a large role in the bonding of the atom and so will be largely unaffected by the chemical or crystal environment, leaving them unchanged from their usual atomic configuration. It is mainly the effect of the outer shell valence electrons that dominates the behaviour of materials. The core electron states can be described by atomic wavefunctions and their effect on the valence electrons can be approximated through the use of a pseudopotential. The form of this pseudopotential is such that the corresponding pseudo-wavefunctions have no small scale structure that would require a large number of basis states to describe
and which accurately reproduce all the features of the core and valence electrons outside of a certain core radius $r_c$. This both reduces the number of electrons in the calculations, and allows for a lower value of $E_{cut}$ now that the core region is smooth. A schematic representation of a pseudopotential is shown in figure 2.3.

The earliest pseudopotentials were developed by Philips and Kleinman in 1959 [46]. Early pseudopotential construction was empirical, they were fit to experimental data and were initially local pseudopotentials. Local pseudopotentials are ones for which all the different electrons, regardless of whether they are in $s,p,d$ orbitals etc., see the same potential that only depends on $r$. It was not until the work of Chelikowsky and Cohen in 1976 that the first empirical non-local pseudopotentials were developed [47]. For non-local pseudopotentials the $s,p$ and $d$ electrons each see a different potential.
2.3.1 Norm-Conserving Pseudopotentials

A desired feature of pseudopotentials is transferability; transferable pseudopotentials developed for use in one system will perform equally well in a different system. Norm-conserving pseudopotentials are important as they are highly transferable. Satisfying the norm-conservation requirement allows the pseudo-core region to replicate the scattering of the true core region which allows the pseudopotential to behave consistently in different environments. Hamann, Schlüter and Chang proposed the first norm-conserving pseudopotentials in 1979 [48], with the requirements for a norm-conserving pseudopotential being:

- Exact and pseudo eigenvalues agree for a chosen atomic configuration.
- Exact and pseudo wavefunctions agree past a certain core radius, \( r_c \).
- The exact and pseudo integrals of the charge densities from 0 to \( r \) agree for each valence state for \( r > r_c \). Gauss’ law then guarantees that the electrostatic potential outside of \( r_c \) agrees for the exact and pseudo cases.
- The logarithmic derivatives of the exact and pseudo wavefunctions, and their first energy derivatives agree for \( r > r_c \). This minimises the error in the pseudo scattering when compared to the real core scattering.

The first two conditions are met by comparing the exact all-electron system,

\[
\left[ -\frac{1}{2} \left( \frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} \right) + V^{AE}(r) \right] \phi^{AE}_l(r) = \epsilon^{AE}_l \phi^{AE}_l(r), \tag{2.34}
\]

with the pseudo system

\[
\left[ -\frac{1}{2} \left( \frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} \right) + V^{PP}_l(r) \right] \phi^{PP}_l(r) = \epsilon^{PP}_l \phi^{PP}_l(r), \tag{2.35}
\]

and ensuring that

\[
\epsilon^{AE}_l = \epsilon^{PP}_l, \tag{2.36}
\]
\[ \phi_{l}^{AE}(r) = \phi_{l}^{PP}(r) \text{ for } r > r_{lc}. \] (2.37)

Here \( \phi_{l}^{AE}(r) \) and \( \epsilon_{l}^{AE} \) are the all-electron wavefunctions and corresponding eigenvalues for a given orbital angular momentum, \( l \), and \( \phi_{l}^{PP}(r) \) and \( \epsilon_{l}^{PP} \) are the pseudo wavefunctions and eigenvalues. This \( l \) dependence makes this a non-local method. The last two norm-conservation conditions are achieved by making sure the relationship

\[ -2\pi \left( r\phi_{l}^{PP}(r)\right)^{2} \frac{d}{dr} \left( \frac{d}{dr} \ln \phi_{l}^{PP}(r) \right) \bigg|_{r_{c}} = 4\pi \int_{0}^{r_{c}} |\phi_{l}^{PP}(r)|^{2} r^{2} dr, \] (2.38)

is met.

### 2.3.2 Pseudopotential Smoothness

In 1991 Troullier and Martin presented a method of generating smooth norm-conserving pseudopotentials [49]. A smooth pseudopotential is one that gives a rapid convergence of the total energy of a system with an increase in the plane wave basis set. They started with a form of pseudopotential defined by Kerker [50]

\[ V_{scr, l}^{PP}(r) = \begin{cases} V^{AE}(r) & \text{if } r \geq r_{lc} \\ \epsilon_{l} + \frac{l+1}{2} \frac{p'(r)}{2} + \frac{p''(r)+p'(r)}{2} & \text{if } r \leq r_{lc}. \end{cases} \] (2.39)

The smoothness condition proposed by Troullier and Martin consists of defining the \( p(r) \) as a sixth order polynomial in \( r^{2} \)

\[ p(r) = c_{0} + c_{2} r^{2} + c_{4} r^{4} + c_{6} r^{6} + c_{8} r^{8} + c_{10} r^{10} + c_{12} r^{12}, \] (2.40)

where seven \( c \) coefficients are solved for by using the conditions of norm-conservation, zero curvature at the origin and the continuity of the pseudo-wavefunction and its four derivatives at \( r_{lc} \).
2.4 ABINIT

This research used the ALICE High Performance Computing Facility at the University of Leicester using the ABINIT software package. ABINIT is able to calculate the total energy, charge density and electronic structure of atomic systems using Density Functional Theory (DFT), pseudopotentials and a planewave or wavelet basis set.

ABINIT uses a standard routine for calculating the ground state density of a system self-consistently. From an input file that contains the dimensions of the unit cell and the positions of all the atoms in the system, ABINIT initially calculates a starting density based on the pseudopotentials and exchange correlation functions chosen. From this starting density and the pseudopotential information it calculates the Hartree and exchange correlation potentials. It then solves the Kohn-Sham single particle equations self-consistently using a diagonalisation technique that gives an output density. This density is then compared with the initial starting density and if it is not consistent a new input density is formed through Pulay mixing the previous ones [51]. The more self-consistent iterations the calculation performs the more accurate the result. In reality the number of self-consistent iterations or tolerance on the change in total energy must be capped to stop the calculation running for a prohibitive amount of time. For all our calculations we have a cut-off on the total number of self-consistent cycles of 100, and a tolerance on the change in total energy of 0.1 meV between two successive iterations. In practice the energy tolerance is reached long before the limit of self-consistent cycles is reached.

ABINIT is also able to optimise and relax the geometry of a system based on the forces and stresses as well as being able to perform molecular dynamics simulations. For the relaxations we use the Broyden-Fletcher-Goldfarb-Shanno
(BFGS) routines, one of the standard relaxation procedures available in ABINT. This relaxation scheme was chosen as it does not require the use of temperatures, friction coefficients, viscous damping or zero masses. This relaxation technique takes into account the forces on the atoms as well as the second derivatives of the energy with respect to nuclear coordinates, $\frac{\partial^2 E}{\partial x_i \partial x_j}$. As with the normal self-consistent calculations, tolerances are placed on the relaxation to stop the calculation at a desired level. The same total energy tolerance and number of self-consistent cycles are used, however in addition there is a tolerance on the maximum force present in a relaxed system which is given in each chapter.

### 2.5 STM and STS

As mentioned in the previous chapter this work is concerned with STM images of the dopants in the (110) surface. By looking at these images we can compare the behaviour of the different dopant states as well as compare to experimental results for the dopants that have been investigated using a scanning tunnelling microscope. To generate the STM images from our DFT simulations we need to know how the scanning tunnelling microscope works.

In 1983 Tersoff and Hermann presented the first quantitative theory of the scanning tunnelling microscope [21]. By modelling the STM tip as a sphere, with a simple asymptotic spherical s-type wavefunction, they were able to calculate the tunnelling matrix elements between the tip and the surface. This was then used via first order perturbation theory to calculate the tunnelling current, $I$, of an STM tip,

$$I = 32\frac{\pi^3}{\hbar^2 k^4} e^2 V \phi^2 D_t(E_F) R^2 e^{2kR} \rho(r_0; E_F), \quad (2.41)$$

where $D_t$ is the density of states per unit volume of the STM tip, $R$ and $r_0$ are the radius of curvature and the centre of curvature of the tip respectively, $V$ is the
applied voltage, $E_F$ and $\phi$ are the Fermi level and work function of the surface sample, and $k$ is the inverse decay length for those surface states in a vacuum, $k = \sqrt{\frac{2m\phi}{\hbar^2}}$. $\rho(r_0; E_F)$ is an energy windowed density of states of the sample, at the tip position $r_0$,

$$\rho(r_0; E_F) \equiv \sum_{\nu} |\psi_{\nu}(r_0)|^2 \delta(E_{\nu} - E_F),$$  \hspace{1cm} (2.42)

where $\psi_{\nu}$ are the states of the sample with energy $E_{\nu}$. This is called the local density of states (LDOS). Approximating the results of an STM experiment using the LDOS from a pseudopotential calculation is now a common procedure, see for examples Refs. [14, 24, 11, 52].

Throughout the analysis of the different dopants studied in this work we repeatedly generate STM images of the surface. This aids in both comparing our models to experimental results as well as providing a visual measure of the physical changes at the surface due to the presence of the dopants. The STM images in this work were calculated using the LDOS of the system following the Tersoff-Hamman model [53]. We image over all states in an energy window determined by the band gap edges and the tip voltage, $V_{bias}$,

$$I(V_{bias}) \propto \int_{E_{LU/HO}}^{E_{LU/HO}+eV_{bias}} \rho(r, E) dE$$  \hspace{1cm} (2.43)

where $E_{LU/HO}$ is the energy of the lowest unoccupied or the highest occupied orbital and $V_{bias}$ is the applied bias voltage. For filled state imaging we use $E_{HO}$, which is the energy of the valence band edge, and for empty state imaging we use $E_{LU}$, which is the energy of the conduction band edge. $V_{bias}$ is +1 eV for empty state imaging and -1 eV for filled state imaging. By fixing one edge of the energy window to the band edges we obtain a fixed background intensity across impurity calculations with differing band gap states. In all our calculations we took the tip position, $r$, to be 4 Å above the unperturbed surface arsenic positions, which represents an average tip-sample separation [52, 54, 55, 56].
In chapter 4 we also look at scanning tunnelling spectroscopy (STS) data for the surface dopants. The most common type of characteristic measured using STS is a $dI/dV$ curve. We generate a $dI/dV$ curve by calculating $I$ for a series of 0.1 eV voltage windows. This allows the determination of the positions of the band gap edges as well as the size of the band gap. Because it is performed at a fixed point in space it is possible to obtain detailed local density of states information for individual atoms on the surface, seeing how the conduction and valence band states for different atoms vary. The STS calculations were performed 4 Å above the surface directly above either a gallium, arsenic or dopant site.

### 2.6 k-Point Grid

Many of the properties of a system, such as the total energy or density, can be calculated by integrating over all k-points in the first Brillouin zone,

$$\langle A \rangle = \frac{1}{V_{BZ}} \int_{BZ} A(k) d^3k. \quad (2.44)$$

In order to evaluate this integral computationally we use a weighted sum over certain k-points,

$$\frac{1}{N_k} \sum_k w_k A(k), \quad (2.45)$$

where the symmetry and size of the real space cell can be used to reduce the number of k-points that are needed. Monkhorst and Pack built on earlier work by Chadi and Cohen [57] to develop a method for finding these special points in the Brillouin zone that lead to efficient approximations to the full integration [58]. They gave a procedure for generating a uniform grid of k-points than can be used in the finite Brillouin zone sampling in equation 2.45. The $N_1 \times N_2 \times N_3$ grid is given by,

$$k_{n_1,n_2,n_3} = \sum_{i=1}^{3} \frac{2n_i - N_i - 1}{2N_i} b_i \quad n_i = 1, 2, 3, ..., N_i, \quad (2.46)$$
where
\[ b_1 = \frac{2\pi}{V_{cell}} a_2 \times a_3, \quad b_2 = \frac{2\pi}{V_{cell}} a_3 \times a_1, \quad b_3 = \frac{2\pi}{V_{cell}} a_1 \times a_2, \] (2.47)

and \( a_1, a_2 \) and \( a_3 \) are the real space lattice vectors.

For surface simulations, such as the ones in this thesis, the \( z \) component of the real space cell is large enough that in \( k \)-space only one \( k \)-point is needed for \( N_3 \). Once the number of \( k \)-points for each of the three dimensions is given ABINIT will generate a Monkhorst-Pack grid based on the system.

### 2.7 Geometrical Method for Estimating Impurity Positions

In chapters 4 and 5 we show that the geometry of the relaxed surface and the character of the XSTM positions depends primarily on the effective size or covalent radius \([9]\) of the impurity atoms, and the bond length of the impurity to its nearest neighbour. In this section we define the calculation used to predict the relaxed dopant positions given the bond length information.

To find the positions of the dopants a mathematical technique known as trilateration was employed \([59]\). This is the same principle as the global positioning system (GPS), using three known positions and three known lengths to find an unknown position. All that is needed to find the dopant position, \( I \), are the positions of the three nearest neighbour atoms the dopant will be bonded to, \( A, B \) and \( C \), and the three lengths of these nearest neighbour bonds, \( R_{AI}, R_{BI} \) and \( R_{CI} \) as shown in figure 2.4. We will begin by presenting the equations used to calculate the dopant position for a specific case where the equations are simplified. For this special case the positions of the three nearest neighbour atoms must all lie in the \( xy \) plane with atom \( A \) at the origin and atom \( B \) along the \( x-\)
Figure 2.4: Representation of the trilateration process used to find the impurity location (schematic). $A, B$ and $C$ mark the positions of the three nearest neighbour atoms after translation to the $xy$ plane, circles are great circles of spheres centred on the nearest neighbour atoms with radii equal to the dopant-neighbour bond length, and the dopant position, $I$, is at one of the intersection points of these three spheres which is found by solving equations 2.49.

axis. We will then present the more general case. The bond lengths $R_{AI}$, $R_{BI}$ and $R_{CI}$ are related to the covalent radii of the atoms, $r_A$, $r_B$, $r_C$ and $r_I$, via

\[
R_{AI} = r_A + r_I, \quad (2.48a)
\]
\[
R_{BI} = r_B + r_I, \quad (2.48b)
\]
\[
R_{CI} = r_C + r_I. \quad (2.48c)
\]

The position of the dopant, $(I_x, I_y, I_z) \equiv I$, is then given by

\[
I_x = \frac{R_{AI}^2 - R_{BI}^2 + B_x^2}{2B_x}, \quad (2.49a)
\]
\[
I_y = \frac{R_{AI}^2 - R_{CI}^2 + C_x^2 + C_y^2}{2C_y} - \frac{C_x}{C_y} I_x, \quad (2.49b)
\]
\[
I_z = \pm \sqrt{R_{AI}^2 - I_x^2 - I_y^2}. \quad (2.49c)
\]

Equations 2.49 have two solutions: a positive $I_z$ solution above the $xy$ plane and a negative $I_z$ solution below the $xy$ plane. In all cases we take the positive solution as the negative solution would lead to the dopant atom intersecting with the sub-surface atoms.

We can generalise equations 2.49 to work with a system of atoms in a general
coordinate system. If the three initial neighbour locations are given by

\[ A = (A_x, A_y, A_z), \quad (2.50a) \]
\[ B = (B_x, B_y, B_z), \quad (2.50b) \]
\[ C = (C_x, C_y, C_z), \quad (2.50c) \]

then the first step is to translate them all so that one of them, in this case \( A \), is located at the origin. We denote the translated positions with primes,

\[ A' = (0, 0, 0), \quad (2.51a) \]
\[ B' = (B_x - A_x, B_y - A_y, B_z - A_z), \quad (2.51b) \]
\[ C' = (C_x - A_x, C_y - A_y, C_z - A_z). \quad (2.51c) \]

Next the positions are rotated about the origin such that they all lie in the \( xy \) plane and one of them, in our case \( B \), lies along the \( x \)-axis. The rotation matrix to be applied to all three positions is given by

\[
R = \begin{pmatrix}
\cos(\gamma) \cos(\theta) \cos(\phi) & -\cos(\phi) \sin(\gamma) - \sin(\gamma) \cos(\theta) & \cos(\gamma) \sin(\phi) \\
\sin(\gamma) \cos(\theta) + \cos(\gamma) \cos(\phi) & \cos(\phi) \sin(\gamma) - \cos(\gamma) \cos(\theta) & -\cos(\gamma) \sin(\phi) \\
-\sin(\phi) \cos(\theta) & \sin(\phi) \sin(\theta) & \cos(\phi)
\end{pmatrix}
\quad (2.52)
\]

where the three angles are

\[ \theta = \arctan \left( -\frac{n_y}{n_x} \right), \quad (2.53a) \]
\[ \phi = \arctan \left( \frac{n_y \sin(\theta) - n_z \cos(\theta)}{n_z} \right), \quad (2.53b) \]
\[ \gamma = \arctan \left( \frac{B'_x \sin(\theta) + B'_y \cos(\theta)}{B'_z \cos(\phi) - B'_y \cos(\phi) \sin(\theta) + B'_z \sin(\phi)} \right). \quad (2.53c) \]

Here \( n \) is the vector normal to the plane that passes through the three translated points and is defined as, \( n = B' \times C' \). The \( \theta \) and \( \phi \) angles rotate the normal of the plane so that it lies along the \( z \)-axis, putting all three points in the \( xy \) plane, and the angle \( \gamma \) is the rotation around the \( z \)-axis that puts \( B' \) on the \( x \)-axis.
The translated and rotated position of the dopant, \( \mathbf{I}^{\prime R} \), can then be found from equations 2.49, with the components of \( \mathbf{A}, \mathbf{B} \) and \( \mathbf{C} \) being replaced with the new translated and rotated \( \mathbf{A}^{\prime R}, \mathbf{B}^{\prime R} \) and \( \mathbf{C}^{\prime R} \), where \( \prime R \) indicates that the position has undergone the translation and rotation outlined above. The position of the dopant in the original coordinate system of the super cell can then be found by first applying the inverse of 2.52,

\[
R^{-1} = \begin{pmatrix}
\cos(\theta) \cos(\phi) \cos(\gamma) - \sin(\gamma) \sin(\theta) & \cos(\phi) \cos(\theta) \sin(\gamma) + \cos(\gamma) \sin(\theta) & -\cos(\theta) \sin(\phi) \\
-\sin(\theta) \cos(\phi) \cos(\gamma) - \sin(\gamma) \cos(\theta) & -\cos(\phi) \sin(\theta) \sin(\gamma) - \cos(\gamma) \cos(\theta) & \sin(\theta) \sin(\phi) \\
\sin(\phi) \cos(\gamma) & \sin(\phi) \sin(\gamma) & \cos(\phi)
\end{pmatrix}
\]

(2.54)

and then by applying a translation of \( \mathbf{A} \).
Chapter 3

Clean GaAs Simulation

3.1 Introduction

The focus of this work is on the modelling and understanding of impurity atoms on the GaAs (110) surface, so an accurate simulation of the impurities is needed. To that end we require not only a detailed knowledge of the structure of a clean GaAs (110) surface but also details on how the structure of the surface changes once the impurity atom is added. In this chapter we present our construction of a clean GaAs (110) slab.

An accurate simulation of the GaAs (110) surface is the foundation that each section of this work rests upon, and needs to be rigorously checked before the addition of each of the different impurity types. By performing careful checks on the accuracy of our host lattice, confidence is ensured in later work when the impurity atoms are introduced to the system. While simulating a GaAs surface is not new it is still vital in ensuring an accurate groundwork for our later calculations, as well as for testing our calculations against well known results. In this chapter we present our construction of a clean GaAs (110) slab including the analysis and comparison between currently available pseudopotentials. We also
present the convergence checks on the total energy of the system with respect to the cut-off energy, $E_{\text{cut}}$, and Monkhorst-Pack grid density, $k$-grid, that were performed. These systems and checks then provide the background for later chapters.

In section 3.2 we present our analysis of the possible pseudopotentials for the clean GaAs system and our choice of the pseudopotentials that are then used throughout the rest of the work. In section 3.3 we outline the set-up of the GaAs slab and go through the detailed convergence checks that will be used for all future impurity calculations.

### 3.2 Investigation of Pseudopotentials: Bulk GaAs

There is a large number of available pre-generated pseudopotentials for the atoms investigated in this work, some of which can give very different results for the same systems [60]. Initial work involved systematically testing the available pseudopotentials in a simulation of bulk gallium arsenide. Two separate programs were used for these simulations, ABINIT [61, 62] and QUANTUM ESPRESSO [63].

From the ABINIT pseudopotential libraries there were three types of norm-conserving pseudopotential available for gallium, arsenic and the different impurities considered in this work: Fritz-Haber-Institute (FHI) [64], Hartwigsen-Goedecker-Hutter (HGH) [65] and Troullier-Martins (PSPNC) [66], with both FHI and HGH providing a semi-core variant for gallium that has the $d$-orbitals included in the valence states. From the QUANTUM ESPRESSO pseudopotential library we tested the PW-MT,FHI and PBE-
Figure 3.1: GaAs band structure comparison between FHI pseudopotentials (solid lines) computed with ABINIT and the PW-MT_FHI pseudopotentials (empty circles) computed with QUANTUM ESPRESSO. For clarity only every third value was plotted for the PW-MT_FHI data.

MT_FHI pseudopotentials, with both having semi-core versions available for gallium. The FHI pseudopotentials from the ABINIT libraries correspond to the PW-MT_FHI pseudopotentials in the QUANTUM ESPRESSO libraries, allowing for a direct comparison between the two programs. A comparison of the band structure of bulk GaAs generated by these equivalent pseudopotentials in each program can be seen in figure 3.1, showing the high level of agreement between the two calculations, with the average disagreement between eigenvalues being less than 1.3%.

To investigate the pseudopotentials, the lattice constant, band gap, and effective mass values for a simulation of bulk GaAs were compared. These tests began with a convergence check for each pseudopotential of the total energy of
Figure 3.2: Total energy for the simulation of bulk GaAs for each of the pseudopotential types investigated. All systems are run with a $6 \times 6 \times 6$ $k$-point grid. Top: total energy convergence of pseudopotentials which have the $d$-orbitals in the core. Bottom: total energy convergence of pseudopotentials which have $d$-orbitals in the valence states. The black arrows indicate which axis to read the data values from for each pseudopotential.

We measure the convergence as the fractional difference between the current run and a run performed at a maximum value of $E_{\text{cut}}$:

$$\frac{\left| E_{\text{total}}(E_{\text{cut}}^{\text{max}}) - E_{\text{total}}(E_{\text{cut}}) \right|}{E_{\text{total}}(E_{\text{cut}}^{\text{max}})}.$$  \hspace{1cm} (3.1)

Each data set in figure 3.3 has its own fixed step length in $E_{\text{cut}}$ and for each data set $E_{\text{cut}}^{\text{max}}$ is the highest $E_{\text{cut}}$ shown in figure 3.2. This method of measuring the level of convergence through the fractional difference will be the standard for all future convergence tests in this work. We took the systems to be converged when the total energy was within 1 part in $10^4$ of $E_{\text{cut}}^{\text{max}}$. 


Figure 3.3: Total energy convergence for the simulation of bulk GaAs for each of the pseudopotential types investigated. In each case the energy at each point is compared to a run performed at one step above the highest $E_{\text{cut}}$ shown in the plot. All systems are run with a $6 \times 6 \times 6$ \textit{k}-point grid. Inset: data shown on log scales to show final levels of convergence. The horizontal line indicates the level at which we take the total energy to be converged. Solid lines correspond to pseudopotentials which have the $d$-orbitals in the core, dotted lines correspond to pseudopotentials which have $d$-orbitals in the valence states.
Once the bulk GaAs system had been converged with respect to $E_{\text{cut}}$, the values of the lattice constant, band gap, and effective mass were found, and are shown in figure 3.4. The theoretical lattice constant was found first by minimising the total energy with respect to the lattice constant. The value of the lattice constant which minimised the energy was then used in the calculations of the band gap and effective mass. The conduction band effective mass was found by numerically differentiating $E(k)$ around $k = 0$ using the centred differencing method,

$$m^\ast = \frac{\hbar^2}{m_e} \frac{(\Delta k)^2}{E(0 + \Delta k) - 2E(0) - E(0 - \Delta k)}.$$  

Here $E(0)$ is the energy in joules at the conduction band minimum and $E(0 \pm \Delta k)$ are the energies either side of this at a step size $\Delta k$, which for the data in figure 3.4 was $2.27 \times 10^8 \text{ m}^{-1}$, which gave an effective mass that was converged to better than 3% when compared to a step length 10 times smaller.

The FHI-D, HGH-D, PBE-MT_FHI-D and PW-MT_FHI-D pseudopotentials which include the $d$-electrons in the valence were found to give better theoretical values for the lattice constant, however they also greatly underestimated the band gap and effective mass. The inclusion of $d$-orbitals in the pseudopotential also increased the computational effort as can be seen in figure 3.3, with a much higher value of $E_{\text{cut}}$, around four to five times larger for the pseudopotentials here, required to reach the same levels of total energy convergence. They also add more electrons to the calculations which increases the amount of computing time needed. Sufficiently converged calculations using semi-core pseudopotentials would be prohibitively time consuming when it came to simulating the much larger systems required for modelling an isolated impurity, and so were ruled out.

After the initial investigation, the ABINIT FHI pseudopotentials were chosen, specifically the gallium pseudopotentials that place the $d$-orbitals in the core rather than the valence states. They provide the best matches to the experimental
Figure 3.4: Comparison of the conduction band effective mass, band gap (calculated at \( \mathbf{k} = (0,0,0) \)), and lattice constant values for different GaAs pseudopotentials. The lattice constant was calculated first and then used in the calculations for the other two parameters. In each case the horizontal line represents the experimental value of the associated parameter. The use of semi-core gallium pseudopotentials is indicated with the use of D. The band gap value of the PSPNC pseudopotential (marked with an *) is larger than the experimental as the pseudopotential incorrectly produces an indirect band gap for GaAs.
values for band gap and effective mass, whilst still giving a reasonable value for
the lattice constant. The impurity pseudopotentials were also taken from the
FHI library to eliminate any effects that may arise from mixing different types
of pseudopotential. The effect of the choice of pseudopotential on the calculated
STM images is tested in chapters 4 and 5, showing minimal differences in the
dopant state LDOS due to any variation in the choice of pseudopotential.

3.3 Clean GaAs (110)

The clean (110) surface was set up using a cell that consisted of a standard
configuration of 7 atomic layers (14 atoms) and 6 layers of vacuum [47, 67, 68],
which has been shown in other simulations of dopants on the GaAs (110) surface
to provide well converged results [11, 24]. The (110) surface of GaAs undergoes
a relaxation but does not reconstruct and so retains its $1 \times 1$ periodicity. The
relaxation of the surface was performed using a standard Broyden-Fletcher-
Goldfarb-Shanno minimization scheme implemented by ABINIT that takes into
account the energy of the system as well as the forces [69]. The conditions imposed
on the relaxation were that the difference in total energy between two successive
self-consistent runs was less than 0.1 meV and that the maximum force allowed
between atoms was $5 \times 10^{-5}$ Hartrees per Bohr as suggested by ABINIT [70].
Figure 3.5(a) shows a side view of the (110) surface and indicates the buckling
angle of the surface, $\alpha$, and the vertical displacement of the surface atoms while
3.5(b) shows a top down view of the surface. The $x, y$ and $z$ labelling convention
for the axes displayed in this figure will be the standard throughout this work.
Due to the symmetry of the system the $x$ positions of all the atoms could be
fixed, along with the $x$ and $z$ positions of the atoms that made up the central $z$
layer of the cell.
As is well known, the result of the relaxation is that the surface arsenic atoms move outwards from their initial positions and the surface gallium atoms move inwards. The buckling angle created by this movement, $\alpha$, was found to be $30.37^\circ$ and the perpendicular displacement between the surface atoms, $\Delta_{1,\perp}$, was $0.67 \, \text{Å}$. These values agree very well with the results of experiment and other pseudopotential calculations [71, 67]. There is also vertical movement of the gallium and arsenic atoms in the second and third layers, however this is much smaller than at the surface, with the resulting $z$ displacement from the unrelaxed positions having decreased to only $0.1 \, \text{Å}$ in the second layer and $0.05 \, \text{Å}$ in the third.

The (110) surface gives rise to surface states, which in the bulk position configuration lie in the band gap. Experimentally the surface states are not found in the band gap, and once we allow our system to relax we see the gap states of the bulk configuration move down into the valence bands and up into the conduction bands, giving the empty band gap seen experimentally. The relaxation does,
however, produce the well known reduction in the the size of the gap [11, 72], by a value of 0.37 eV.

The convergence of the total energy with respect to both $E_{\text{cut}}$ and the $k$-point grid was checked for the slab calculation. The convergence with respect to $E_{\text{cut}}$ was checked by looking at the behaviour of the total energy for $E_{\text{cut}}$ values ranging from 2 to 18 Hartree (54.4 to 489.8 eV), and was repeated for four different $k$-point grids; $2 \times 2 \times 1$, $3 \times 3 \times 1$, $4 \times 4 \times 1$ and $5 \times 5 \times 1$. In each case the three numbers represent the number of $k$-points in each dimension in the Monkhorst-Pack grid generated by ABINIT. Only a single $k$-point is required in $z$ due to the large real space size of the cell in that dimension. The results of the convergence tests can be seen in figure 3.6.

At 18 Hartree the total energy becomes very well converged, having reached a total energy value that is within 0.4% of that of the 18 Hartree case. This level of convergence improves to 0.01% when $E_{\text{cut}}$ becomes 15 Hartree. This convergence pattern was found in all the $k$-point tests with only small variations being seen. The $2 \times 2 \times 1$ grid gave a convergence in the total energy of 0.01% at each value of $E_{\text{cut}}$ when compared to the $5 \times 5 \times 1$ grid. Going further to the $3 \times 3 \times 1$ grid gave a subsequent convergence in the total energy of only 0.001% at each value of $E_{\text{cut}}$ when compared to the $5 \times 5 \times 1$ case. The convergence of the system is much more sensitive to changes in $E_{\text{cut}}$ than in the $k$-point grid. For our converged calculation that will be used in later chapters we use a $2 \times 2 \times 1$ $k$-point grid and an $E_{\text{cut}}$ of 15 Hartree (408.2 eV) in most cases but 30 Hartree (816.3 eV) for carbon (see section 5.2.1).
Figure 3.6: The total energy as a function of $E_{\text{cut}}$ for a clean 14 atom (110) slab, for four different k-point configurations. The inset shows the fractional difference in the total energy at each step compared to the energy of the system run at the maximum value of $E_{\text{cut}} = 18$ Hartree.
Chapter 4

Isovalent Dopants in GaAs

4.1 Introduction

Single atomic impurities have a wide range of potential uses in semiconductor devices and so a detailed understanding of the physics of impurities is key to advancing this technology. One method of investigating the behaviour of these single atomic impurities is through XSTM. In XSTM, samples are cleaved in situ to expose the cross-section of the sample. This cleavage allows for observations of bulk structures such as interfaces or impurities at an exposed surface. Although this technique is widely used the interpretation of the results, as mentioned in chapter 1, is still an open issue [73]. In this chapter we present a simple analysis of XSTM images for isovalent impurities on the GaAs (110) surface based on DFT calculations and our geometrical method that was outlined in chapter 2. This model provides insight into disentangling the geometrical and electronic effects of features seen in STM images.

Isovalent impurities are ideal for investigating and understanding XSTM physics due to their lack of complicated charge effects. Recently XSTM has been used to investigate single isovalent nitrogen impurities in and below the surface
layer of a cleaved (110) GaAs surface [3]. The behaviour of the surface and sub-
surface nitrogen impurities has been investigated qualitatively [24] using density
functional theory (DFT) simulations, however there has been no quantitative
analysis or detailed comparison to the experimental results. Nor has there been
any detailed analysis of other isovalent impurities from groups III and V.

Here we compare the experimental XSTM images with \textit{ab initio} DFT
simulations for a nitrogen doped GaAs (110) surface. We use the geometrical
model outlined in chapter 2 to help understand the physics of the impurity system
and show that experimentally observed STM features can be explained primarily
by the effective size and position of the impurity atom. We then extend the
analysis to other isovalent impurities in the GaAs (110) surface layer from groups
III and V. We find that the predictions of our model match very well with the
full DFT relaxation of the structure, with disagreements in the bond lengths of
less than 4.2%. We also find that the calculated STM images obtained for the
group III and V impurities match with both the predictions of our geometrical
model and available experimental results.

In section 4.2 we outline the DFT calculations used and in section 4.3 we
go through the convergence checks performed on the different isovalently doped
systems. In section 4.4 we show the experimental and calculated STM images
of nitrogen impurities and present our analysis of the results, with specific focus
on a quantitative comparison of STM image features for a second layer nitrogen
impurity. In section 4.5 we extend our analysis to the other group III and V
isovalent impurities and compare to experimental results where available. In
section 4.6 we look at the behaviour of an isovalent impurity in a different host
semiconductor and finally in section 4.7 we present a summary of this chapter.
Figure 4.1: $3 \times 3$ super-cell. Left: side view of the cell showing the surface buckling angle, $\alpha$. Right: plan view of the (110) surface. Black dots represent arsenic atoms and white dots represent gallium atoms. The red dot shows the isovalent group III impurity position and the blue dot shows the isovalent group V impurity position. The dimensions of the cell are $x = 11.73 \, \text{Å}$, $y = 16.59 \, \text{Å}$ and $z = 25.42 \, \text{Å}$.

### 4.2 Simulations

The clean GaAs (110) surface obtained in section 3.3 was used to construct a 126 atom super-cell with a $3 \times 3$ surface periodicity for all impurity simulations. This super-cell is shown in figure 4.1 indicating the placement of the group III and V impurities. For the isovalent systems the group III impurities substitute a gallium atom and the group V impurities substitute an arsenic atom. After the addition of an impurity at a substitutional location the system was allowed to fully relax. Although the tolerance on the forces for the clean slab was $0.5 \times 10^{-4}$ Hartrees per Bohr the larger, more computationally intensive, doped systems were run with a slightly larger tolerance on the forces of $1 \times 10^{-4}$ Hartrees per Bohr. When checked against two other relaxations performed with the higher tolerance the positions were found to change by less than 1%.

For each of the impurities investigated we performed detailed convergence checks with respect to $E_{\text{cut}}$ and the number of $\mathbf{k}$-points. In figures 4.2 and 4.3 we show the fractional change in total energy for the largest and smallest impurities.
Chapter 4

Figure 4.2: The relative error in the total energy of a doped GaAs slab as a function of $E_{cut}$. The total energy is compared to a simulation with an $E_{cut}$ of 18 Hartree. The total energies for B, Tl, N and Bi doped systems at 18 Hartree are -14.89, -15.33, -14.96 and -15.51 keV respectively.

We see the same monotonic convergence as a function of $E_{cut}$ as we saw in the case of the total energy in figure 3.6, achieving a level of convergence of better than 1 part in 2000 for an $E_{cut}$ of 15 Hartree. The $k$-grid convergence shown in figure 4.3 does not show the same monotonic convergence as figure 4.2. The important feature for the $k$-grid convergence is that the changes in the total energy are tiny, roughly four orders of magnitude smaller than any changes due to $E_{cut}$.

For the nitrogen system we will also be looking at the behaviour of sub-surface impurities, and so the convergence of the sub-surface systems was also checked. In figure 4.4 we show the convergence of the total energy with respect to $E_{cut}$ for systems with a nitrogen impurity in the first, second and third surface layers. The
Figure 4.3: The relative error in total energy as a function of \( k \)-point grid for four impurities from groups III and V. The total energy is compared to a simulation with an \( k \)-grid of \( 5 \times 5 \times 1 \).
Figure 4.4: The relative error in total energy as a function of $E_{\text{cut}}$ for the three different nitrogen positions in the slab. The total energy in each case is compared to a simulation with an $E_{\text{cut}}$ of 18 Hartree.

The same convergence is observed independent of the position of the nitrogen impurity below the surface. A similar check against $k$-grid was performed showing the same level of convergence as in figure 4.3.

As well as the checks on $E_{\text{cut}}$ and the $k$-point grid, detailed checks were also performed to determine the convergence of the system with respect to the super-cell size. The total energy of the system, $E$, is proportional to the volume of the super-cell and thus the number of atoms, $N$,

$$E(N) = \alpha N + (\alpha_{\text{dopant}} - \alpha_{\text{Ga}}), \quad (4.1)$$

where $\alpha$ is the energy per atom for a clean semi-infinite slab, $\alpha_{\text{Ga}}$ is the energy per atom of the substituted gallium and $\alpha_{\text{dopant}}$ is the contribution to the total energy from the dopant. The total energy per atom of the system should converge for large $N$,

$$E_{\text{atom}}(N) = \alpha + \frac{\alpha_{\text{dopant}} - 59.9}{N} \quad \text{where } N > 0, \quad (4.2)$$

where -59.9 eV per atom is the energy per gallium atom [74]. We use least squares
fitting to determine the coefficients in equation 4.2 giving $\alpha = -117.94$ eV and $\alpha_{\text{donor}} = -94.2$ eV. The value of $-117.94$ eV/atom is the energy per atom of the clean (110) slab and is expected in the large $N$ regime [75, 71, 74]. In figure 4.5 we show the convergence of the total energy per atom as a function of system size for a system with a single nitrogen impurity.

As well as checking the total energy per atom of the system we also made sure the simulated STM images were converged. These checks were performed in the same manner laid out in our previous work [76], where we checked our STM image against calculations performed with an increased $E_{\text{cut}}$, an increased number of $k$-points, a larger super-cell or a different pseudopotential. Figure 4.6 shows a series of simulated STM images for a 1st layer nitrogen impurity indicating the convergence of the image. Figure 4.6(a) is a simulation with the parameters used for what we will call the standard system; a $3 \times 3$ super-cell with an $E_{\text{cut}}$ of 15 Hartree and a $2 \times 2 \times 1$ $k$-point grid. Parts (b), (c) and (d) all have one of the parameters increased; (b) shows a system with an $E_{\text{cut}}$ of 18 Hartree, (c) shows a system with a $3 \times 4$ super-cell and (d) shows a system with a $3 \times 3 \times 1$ $k$-point grid. We found that there was no discernible change in the images when

Figure 4.5: Total energy per atom as a function of the number of atoms for systems with a single nitrogen impurity. The points are the results of the simulation and the line is the results of a least squares fitting of the data using the functional form given in equation 4.2.
increasing either $E_{\text{cut}}$, the system size or the number of points in the $k$-grid. For example, going from a $3\times3$ to a $3\times4$ super-cell changes the peak positions by less than 0.04 Å and the peak intensities by less than 1.5%.

### 4.3 Significance of Geometry

A necessary step in modelling a system with an impurity is a full relaxation of the atomic positions, as the presence of the impurity will likely change the geometry of the surrounding lattice. We find that the final relaxed positions of the dopants change systematically with the position of the dopant in the column of the periodic table. In section 4.4 we show that the change in relaxed position is directly linked to the distribution of the STM contrast. Because of this link between the STM contrast and the relaxed positions it is important to understand the physics of the surface relaxation.

In figure 4.7(a) and (b) we show the relaxed positions of all the group III and V atoms in the outer layer of GaAs. For group III, boron is the dopant with the lowest atomic number and has relaxed furthest into the surface; this is the same
Figure 4.7: Side-on view of the final relaxed positions of (a) the group III impurities substituting gallium and (b) the group V impurities substituting arsenic, as well as their first nearest neighbours. Impurity atoms are represented by filled shapes and the nearest neighbours atoms are represented by the corresponding empty shapes. Black lines have been added as a guide for the eye.
for nitrogen in group V. Thallium and bismuth have the largest atomic numbers in each group and these are the dopants that have relaxed furthest out of the surface in each case. For the group III dopants a planar bonding configuration is preferred with the average bond angle in the $yz$ plane being 168.7°, while for group V a more tetrahedral configuration is seen with an average bond angle in the $yz$ plane being 92.8°. The general form of these configurations, either planar or tetrahedral, is determined by the number of valence electrons, with the different dopant sizes providing the spread of positions seen in both cases.

### 4.3.1 Covalent Radius

The covalent bond length between two atoms is known to depend on the positions of the atoms in the periodic table. The covalent bond length between atoms $A$ and $B$, $R_{AB}$, can be calculated as a sum of their individual covalent radii [9], $r_A$ and $r_B$. The covalent radius provides a measure of the effective size of an atom and generally increases with atomic number within a group of the periodic table, and generally decreases with atomic number across the groups. For an atom of element $E$ the covalent radius is usually calculated as half the homo-nuclear, E-E, bond length, or by using a bond to a reference atom or molecule of known covalent radius. The covalent radii in Ref. [9] are calculated using E-E, E-Hydrogen and E-CH$_3$ bond data. The length of a covalent bond between any pair of atoms can be estimated from the covalent radii in Ref. [9].

The inclusion of covalent radii in density functional theory is not new. Work was done back in the 80’s to try and use DFT to calculate the covalent radius of atoms. In 1986 Balbás et al., building on previous work, used DFT calculations to estimate the covalent radius for 32 atoms by calculating the radius at which the chemical potential is equal to the electrostatic potential in a free atom [77].
The values they got from this method were reasonably accurate, disagreeing on average by about 4% with the experimentally derived values we use from Ref [9]. More recently Castleton et al. have interpreted energy levels for hydrogen on III-V surfaces based on the difference between the covalent radii of the host atoms [78], as the difference in covalent radii is related to the ionicity of the material. Other sources have used the values in Ref. [9] to compare to the bond lengths from their simulations [79, 80, 81], however to our knowledge no extensive predictions or interpretations based on covalent radii like the ones we present here have been performed.

In table 4.1 we present a comparison of the average bond lengths from the DFT relaxations for dopants in the surface layer, with predicted bond lengths based on the sum of the covalent radius of the dopants and their nearest neighbours. From the theoretical covalent radii we expect to see that for each group in the periodic table impurities with larger atomic number, Z, will relax out of the surface as they will generally have a larger effective size than the substituted atom, while

<table>
<thead>
<tr>
<th>Bond Lengths (Å)</th>
<th>B</th>
<th>Al</th>
<th>Ga</th>
<th>In</th>
<th>Tl</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT</td>
<td>2.05</td>
<td>2.40</td>
<td>2.37</td>
<td>2.55</td>
<td>2.54</td>
</tr>
<tr>
<td>Sum of covalent radii</td>
<td>2.06</td>
<td>2.47</td>
<td>2.45</td>
<td>2.63</td>
<td>2.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>P</th>
<th>As</th>
<th>Sb</th>
<th>Bi</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT</td>
<td>1.89</td>
<td>2.29</td>
<td>2.38</td>
<td>2.56</td>
<td>2.67</td>
</tr>
<tr>
<td>Sum of covalent radii</td>
<td>1.95</td>
<td>2.35</td>
<td>2.45</td>
<td>2.64</td>
<td>2.75</td>
</tr>
</tbody>
</table>

Table 4.1: Average dopant bond lengths for all group III and V atoms taken from the DFT simulations compared with bond lengths calculated from the covalent radii in Ref. [9].
atoms with lower atomic numbers are expected to relax into the surface as they will generally have a smaller effective size. This trend of relaxing outwards for larger $Z$ only holds in each group of the periodic table and not across the elements in general. In each group an increase in $Z$ gives an increase in the number of electron shells as the principal and azimuthal quantum numbers increase. This has the effect of increasing the covalent radius (see each row in table 4.1). Along each row of the periodic table on the other hand, an increase in $Z$ gives an increase in the atomic number while the number of electron shells remains constant. This has the effect of decreasing the covalent radius as the nuclear charge increases (see each column in table 4.1).

In table 4.1 we see an excellent agreement between the DFT bond lengths and the bond lengths calculated as a sum of the covalent radii, with the largest error being a 4.2% mismatch in the case of thallium. For nitrogen the sub-surface relaxation was also investigated (see section 4.3.2), and we again found excellent agreement between the predicted and simulated bond lengths, with the second and third layer nitrogen dopants having average DFT bond lengths of 2.00 Å and 2.02 Å respectively compared to the sum of the theoretical covalent radii of 1.95 Å. A potential source of the error between the theoretical and DFT bond lengths comes from the pseudopotentials. For the clean GaAs slab our lattice constant is 2.2% smaller than the experimental value and so we expect to see this very slight disagreement in the final bond lengths.

In figure 4.7 dopants of low atomic number tend to relax into the surface while dopants of high atomic number tend to relax outwards. This is consistent with the tendency for the covalent radii within a group to increase with atomic number, suggesting that geometrical effects, particularly the effective dopant size, are important in these isovalent systems. The covalent radius data also explains the relaxed position of aluminium, which is slightly above the gallium site. This is
due to the aluminium atom having a slightly larger covalent radius than gallium despite having a lower atomic number. This indicates that attributing the trend in relaxed position to atomic number is only an approximation, and the correct driving force behind the relaxation is the effective size of the dopants.

In order to test the idea that the relaxed positions of the dopants depend on their effective size, we compute the surface impurity position using the trilateration technique outlined in chapter 2. In table 4.2 we present the position data for the group V dopants and their three nearest neighbours taken from the DFT relaxation and the trilateration procedure. For the trilateration procedure the three neighbours were fixed in their clean slab positions and the dopant position was calculated using the sum of the covalent radii as outlined in chapter 2. In table 4.2 we see that the sub-surface nearest neighbours show only small deviations from the clean slab positions, less than 0.08 Å in all cases, with the surface neighbours showing a generally larger displacement of between 0.05 and 0.15 Å. For nitrogen, its small covalent radius produces larger shifts in the nearest neighbours during the DFT relaxation, so fixed nearest neighbour positions cannot be used. Instead the surface nearest neighbours are shifted slightly while the sub-surface neighbour is kept fixed, in order to obtain a solution to the trilateration equations. Our interpretation that the geometry is important in the relaxation is supported by the good agreement between the trilateration positions and the DFT relaxation, with the largest difference in position being only 0.3 Å for the tin dopant. This good agreement is obtained without any input from the DFT relaxation.

In the case of group III the dopants prefer a planar bonding configuration. In contrast to the tetrahedral preference of the group V atoms, this causes a greater movement in the nearest neighbours. Because of this we cannot set up the trilateration calculation in a way that is independent of the DFT relaxation. We
<table>
<thead>
<tr>
<th></th>
<th>DFT positions</th>
<th>Trilateration positions</th>
<th>Difference (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>5.867 9.708</td>
<td>5.865 9.752</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>4.185 10.535</td>
<td>4.100 10.580</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>7.548 10.534</td>
<td>7.631 10.580</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>5.867 8.408</td>
<td>5.865 8.337</td>
<td>0.07</td>
</tr>
<tr>
<td>P</td>
<td>5.865 9.552</td>
<td>5.865 9.521</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>3.949 10.646</td>
<td>3.910 10.675</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>7.781 10.646</td>
<td>7.821 10.675</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>5.865 8.341</td>
<td>5.865 8.337</td>
<td>0.01</td>
</tr>
<tr>
<td>As</td>
<td>5.865 9.520</td>
<td>5.865 9.418</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>3.910 10.675</td>
<td>3.910 10.675</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>7.821 10.675</td>
<td>7.821 10.675</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>5.865 8.337</td>
<td>5.865 8.337</td>
<td>0.00</td>
</tr>
<tr>
<td>Sb</td>
<td>5.865 9.470</td>
<td>5.865 9.223</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>3.837 10.712</td>
<td>3.910 10.675</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>7.894 10.712</td>
<td>7.821 10.675</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>5.865 8.317</td>
<td>5.865 8.337</td>
<td>0.03</td>
</tr>
<tr>
<td>Bi</td>
<td>5.865 9.267</td>
<td>5.865 9.159</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>3.871 10.654</td>
<td>3.910 10.675</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>7.860 10.654</td>
<td>7.821 10.675</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>5.865 8.306</td>
<td>5.865 8.337</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 4.2: Comparison of the group V dopant positions and their three nearest neighbours between the DFT simulation and the trilateration calculation. In each case the first line is the dopant position and the following three lines are the nearest neighbour positions. For the trilateration positions the three nearest neighbour positions are the ones used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions.
can however test our geometrical interpretation by using the nearest neighbours’ positions from the DFT relaxation in the trilateration procedure. In table 4.3 we present the DFT position data for the group III dopants and their nearest neighbours along with the trilateration predictions using the DFT nearest neighbours.

We see that the difference in the predicted dopant position and the DFT position is larger than in the case of group V, even when using the DFT positions. This is due to the small difference between the DFT bond length, which comes from the pseudopotentials, and the sum of the covalent radii. These differences are not identical for all bond lengths, with the sub-surface bond mismatch on average 1.1% and the surface bond mismatch on average 3.8%. The bond length difference has a greater influence on the predicted position for the planar geometry than in the case of the tetrahedral geometry of the group V dopants. The relationship between the difference in position, \( D \), and the bond angle, \( \theta \), for a 3.8% bond length mismatch is given by

\[
D = R_{DFT} \left( \sqrt{1.038^2 - \cos^2 \left( \frac{180 - \theta}{2} \right)} - \sin \left( \frac{180 - \theta}{2} \right) \right), \tag{4.3}
\]

where \( R_{DFT} \) is the bond length. A plot of this relationship between \( \theta = 0^\circ \) and \( \theta = 180^\circ \) is shown in figure 4.8. This expression has a maximum at \( \theta = 180^\circ \), which gives a difference in position of 0.278\( R_{DFT} \), and a minimum at \( \theta = 0^\circ \), which gives a difference in position of 0.038\( R_{DFT} \). For figure 4.8 we see that the function of \( \theta \) on the right of Eq. 4.3 stays quite flat far away from the minimum at \( \theta = 0^\circ \) and only starts to increase significantly when \( \theta \sim 115^\circ \). In figure 4.9 we present a representation of the differences in positions for three bond angles, 0°, 90° and 180°. The original bond configurations are given in blue, all with bond lengths of 1, and the red dashed lines and empty shapes show the results of increasing the bond lengths by 3.8%. The discrepancy between the effect on the
### Table 4.3: Comparison between the DFT simulation and the trilateration calculation for the group III dopant and nearest neighbour positions. For the DFT positions the first line is the dopant position and the following three lines are the nearest neighbour positions. The DFT nearest neighbour positions are used as input for the trilateration calculations in this case. The difference is measured as the Euclidean distance between the two positions.

<table>
<thead>
<tr>
<th></th>
<th>DFT positions</th>
<th>Trilateration positions</th>
<th>Difference (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>3.910 5.131 11.168</td>
<td>3.910 5.250 11.342</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>3.910 6.760 9.940</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.233 4.158 11.830</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.588 4.158 11.830</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al</td>
<td>3.910 5.131 11.318</td>
<td>3.910 5.410 11.696</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>3.910 6.956 9.770</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.925 3.970 11.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.895 3.970 11.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In</td>
<td>3.910 5.105 11.408</td>
<td>3.910 5.358 11.745</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>3.910 7.018 9.706</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.783 3.843 12.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.038 3.843 12.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tl</td>
<td>3.910 5.097 11.522</td>
<td>3.910 5.414 11.915</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>3.910 6.970 9.770</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.777 3.849 12.069</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.043 3.849 12.069</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.8: Graphical representation of equation 4.3 showing the difference in position as a function of bond angle.

Figure 4.9: Figure showing how the error in position varies depending on the bond angle. Three bond angles are shown, $0^\circ$ (triangles), $90^\circ$ (squares) and $180^\circ$ (circles). The original bond configuration is given by the solid blue lines and filled shapes, with the effect of increasing the bond lengths by 3.8% being shown by the red dashed line and empty shapes.

$0^\circ$ and $90^\circ$ cases when compared to the $180^\circ$ case is clearly seen.

The group III impurities have a planar bond configuration, with an average bond angle of $168.7^\circ$. This causes the displacement between the DFT and trilateration positions to be larger than for groups IV (see chapter 5) or V. The geometrical interpretation still applies to group III, with the discrepancies in the dopant position being a simple consequence of the planar geometry enhancing the estimation errors. Once the nearest neighbour shifts have been accounted for the remaining discrepancy can be explained as a consequence of the planar geometry
enhancing the effect of the bond length errors. To confirm this we perform the same trilateration procedure as before but now with both the DFT bond lengths and positions to obtain the following differences in position: $D_B = 0.01 \, \text{Å}$, $D_{Al} = 0.01 \, \text{Å}$, $D_{Ga} = 0.04 \, \text{Å}$, $D_{In} = 0.02 \, \text{Å}$ and $D_{Th} = 0.03 \, \text{Å}$, showing the accuracy of the trilateration method once the errors have been reduced.

The geometrical interpretation applies to both group III and V surface layer dopants. This method works well in the case of surface impurities in tetrahedral lattices as the impurity has three points of contact with the surface, and less well for planar geometries, which enhance the effect of any bond length errors. For sub-surface impurities the relaxation becomes more complicated, involving four bonds, and can no longer be easily predicted with this method.

By using the covalent radii model we obtain bond lengths with an error of less than 4.2% and positions with an average deviation of only 0.19 Å when compared with the full DFT relaxation for all dopants and nearest neighbours. When this method was used to seed the DFT relaxation for all the surface dopants in this work we found the relaxation was more computationally efficient. For example, in the case of bismuth the number of iterations required to minimise the forces was reduced from 63 to 40.

It is the final position of the different impurities that determines the character of the STM image. We expect all impurities with a covalent radius smaller than the substituted atom to produce dark regions in the STM image as they move into the surface and impurities with a larger covalent radius than the substituted atom to produce bright regions as they move out. This is true in both the filled state imaging of group V impurities and the empty state imaging of group III impurities (see section 4.4).
Figure 4.10: Comparison of constant height filled state STM images for 1st (a) 2nd (b) and 3rd (c) layer nitrogen impurities in the (110) surface of GaAs along with a diagram of the surface relaxation in each case (d), (e) and (f)). In (a), (b) and (c) the position of the nitrogen has been marked with a white cross. In (d), (e) and (f) the gallium and arsenic atoms are shown along with the nitrogen impurity (black pentagon) and the positions of the GaAs atoms for a clean surface (empty circles). The dotted line marks the surface position. The STM images were generated at 4 Å above the surface. The dimensions of the cell are $x = 11.73 \text{ Å}$, $y = 16.59 \text{ Å}$ and $z = 25.42 \text{ Å}$.

4.3.2 N:GaAs

LDA DFT simulations were performed for nitrogen impurities in the first three surface layers of GaAs (110). Figures 4.10(a), (b) and (c) show the constant height STM images calculated from our DFT calculations for the 1st, 2nd and 3rd layer nitrogen respectively. Parts (d), (e) and (f) show a side-on view of the relaxed geometry. We have indicated the position of the clean surface as well as the clean surface atomic positions for comparison. Figure 4.11 shows an experimental filled state XSTM image. Three typical nitrogen related features can be seen on the (110) surface originating from nitrogen atoms in the first (A), second (B) and third (C) layers.

For the 1st layer impurity in figure 4.10 we see a dark region in the STM image above the nitrogen location (figure 4.10(a)). This dark region is caused by the
Figure 4.11: Constant current height scan of the filled states for a nitrogen doped GaAs (110) surface taken from [3]. A, B and C are features due to nitrogen impurities in the first, second and third atomic layers.
reduced size of the nitrogen compared to the arsenic atoms which results in it relaxing into the surface (figure 4.10(d)). This change of 0.69 Å in the position of the impurity in the z direction causes a shift in the tail of the wavefunction. The impurity wavefunction decays exponentially away from the surface, and so a small change of position in the z direction can lead to a very rapid reduction in the LDOS above the nitrogen when compared to the surrounding arsenic atoms. This matches feature A in the experimental data (figure 4.11). In both 4.10(a) and 4.10(d) we see very little displacement away from the clean positions in the (110) plane for the surrounding arsenic atoms which leaves the rest of the STM image unperturbed.

For the 2nd layer nitrogen we get a region of reduced intensity over the four arsenic surface sites closest to the nitrogen position (figure 4.10(b)). There is very little change in the initial nitrogen position, with most of the effect of the impurity being the shrinking of the bonds to the surrounding atoms (figure 4.10(e)). This reduction of the N-Ga bond lengths is felt more by the gallium atoms closest to the surface; the ones further into the slab are surrounded by more bulk structure and so are more resistant to movement. The relaxation of these surface gallium atoms causes four surface arsenic atoms to move mainly parallel to the surface plane, with a slight displacement into the surface (figure 4.10(e)). The relaxation of the arsenic atoms into the surface is significantly less in this case than for the 1st layer nitrogen, with the average shift in the z direction being only 0.12 Å. This has the effect on the STM image of causing only slight reductions in the LDOS located above those four atoms. This feature, spread over multiple arsenic sites, corresponds to feature B in figure 4.11. The agreement between the theory and experiment for this case is discussed further in section 4.3.3.

The 3rd layer nitrogen produces a very similar feature to the 1st layer case: a relatively unperturbed surface with a dark region above the nitrogen site (figure
4.10(c)). There is an even smaller change in position of the nitrogen away from the initial substitutional position than in the 2\textsuperscript{nd} layer case, and again we see that most of the response to the nitrogen is a shrinking of the bonds to its nearest neighbours. As with the 2\textsuperscript{nd} layer case this has the largest effect on the gallium atom closest to the surface (figure 4.10(f)). The relaxation of this gallium atom close to the surface pulls down the single surface arsenic atom located above it resulting in an STM image that is very similar to figure 4.10(a). The relaxation of the surface arsenic atom in the $z$ direction in this case is 0.23 Å which is smaller than in the 1\textsuperscript{st} layer simulation. This relaxation is still large enough to cause a general dark region around the atom but, because the shift is much smaller than the 0.69 Å value for the 1\textsuperscript{st} layer case, the intensity in the dark region is almost three times higher than in the 1\textsuperscript{st} layer case. This feature caused by the 3\textsuperscript{rd} layer nitrogen corresponds to feature C in figure 4.11.

### 4.3.3 Second Layer Features

A feature of interest in the 2\textsuperscript{nd} layer case is the in-plane shift of the four surface arsenic atoms, as well as slight differences in their peak intensities. In figure 4.12 we compare the data for the second layer feature from both experiment, 4.12(a), and theory, 4.12(b). Here the cell dimensions have been scaled and are given in terms of lattice constant $a_0$. From line scans of the experimental data in figure 4.12(c) we measure an average shift in the $y$ direction of the positions of the peaks marked at A of 8.3±0.8% of the clean arsenic row separation. In figure 4.12(d) we show a line scan of the data taken from our simulations where we measure a shift in the peaks at A in the $y$ direction of 10% of the clean arsenic row separation. The peaks at B are smaller than the peaks at A and are measured experimentally to have an average shift of 2.6±0.7% while the simulation gives a
Figure 4.12: Theoretical and experimental STM data for a 2nd layer nitrogen feature. (a) experimental STM data taken from figure 4.11 (the image has been flipped horizontally to line up with the geometry of the simulation), (b) repeat of the STM data shown in figure 4.10(b), (c) line scans of experimental data showing the relative peak heights with the dotted line being taken from (a) at $x \approx 1.1$ and the solid line taken from (a) at $x \approx 0.4$, (d) line scan of simulation data taken from (b) at $x = 0.37$, and (e) diagram of the DFT surface relaxation in the (110) plane. In (e) the dots are scaled depending on their distance away from the surface. The dimensions of the cell used in the DFT simulation are $x = 11.73 \, \text{Å}$, $y = 16.59 \, \text{Å}$ and $z = 25.42 \, \text{Å}$. 
shift of 2.5%. In the experimental data there is a slight asymmetry between the
two shifts at A as well as an asymmetry of the two shifts at B (figure 4.12(c)),
while in the simulation both shifts at A are symmetric and so are those at B. This
is thought to be due to the inhomogeneity of the surrounding area compared to
the perfectly periodic system in the simulation.

The difference in both shift and intensity of the two peaks at A and the two at
B can be explained by the geometry of the surface and position of the nitrogen
impurity. In figure 4.10(e) we see that the nitrogen’s neighbouring gallium atom
that is closer to the surface relaxes more than the other neighbouring atoms.
This gallium is only fixed to the surface through three bonds while the other
neighbours have four. This allows it to move more in response to the addition of
the nitrogen impurity which in turn leads to a large relaxation of the two surface
arsenic atoms bonded to it. The difference in intensities comes from the fact
that the pair of arsenic atoms at A are being pulled parallel to the surface due
to the position of their neighbouring gallium, while the two arsenic atoms at B
are mainly being pulled down into the surface. These two effects are seen in the
image as the two peaks at A having a larger in-plane shift in position while the
peaks at B have a smaller shift but a larger reduction in intensity.

4.4 Group III and V Isoelectronic Impurities

The behaviour of the nitrogen dopant in the first three surface layers of GaAs can
be understood as a simple result of the difference in size between the nitrogen
and the substituted arsenic atom. Predictions can be made from our geometrical
model on how other isovalent impurities will behave in GaAs and what would be
seen in the resulting STM images. As explained previously, our model predicts
that any impurities that are smaller than the atoms they substitute will cause
Figure 4.13: Empty state calculated STM images of Group III impurities on the 1st layer of the GaAs (110) surface: boron, aluminium, gallium, indium and thallium. We see that the smallest atom, boron, leaves a dark region in the image due to it moving into the surface. The impurities bigger than gallium relax outwards and so give brighter spots. The similarity between the aluminium and gallium images comes from them having almost identical bond lengths and thus relaxations.

darker regions in the STM images, whereas impurities that are larger than the substituted atom will cause brighter regions. To test this a range of impurities were simulated on the first layer of a GaAs (110) slab. The impurities investigated were those from groups III; boron, aluminium, indium, thallium, and those from group V; phosphorus, antimony and bismuth, which when combined with the results of nitrogen as well as the clean GaAs surface, provide a full picture of both groups of isovalent impurities.

4.4.1 Group III STM Images

Figure 4.13 shows the series of empty state STM images for the group III atoms on the (110) surface. The increasing size of the impurities down the group has the boron impurity below the surface level while the indium and thallium atoms, which are of greater effective size, are above the surface. The increased size of aluminium puts it slightly further out of the surface than gallium, which shows up
Figure 4.14: Simulated filled state STM images of Group III impurities on the 1st layer of the GaAs (110) surface: boron, aluminium, gallium, indium and thallium. For the boron system the nearest neighbour arsenic atoms relax towards each other due to the small size of the impurity. The larger indium and thallium impurities push the arsenic atoms away introducing a space near the impurity location. The relaxed position of the impurity in each simulation has been marked with a white cross.

in the STM image as an increase of about 14% in the LDOS over the aluminium site.

Because the group III impurities are further into the surface than the surrounding arsenic atoms we see a sizeable movement in the relaxation of these surface arsenic atoms. This also has the result of altering the simulated filled state STM images which are shown in figure 4.14. Here we are seeing the movement of the arsenic states due to the relaxation of the surface around the impurity. In figure 4.14(a) the small size of the boron impurity has pulled the two neighbouring surface arsenic atoms towards each other into the surface. This has the effect of merging the features above these two atoms as well as reducing their intensities by about 20%. For the larger impurities such as indium and thallium (figure 4.14(d) and (e)) the neighbouring arsenic atoms get pushed away from each other, widening the space between the peaks in the image.

Most of the isovalent dopants that we discuss in this chapter have been studied experimentally by cross-sectional STM. For group III, the elements Al and In
have been frequently observed in GaAs by XSTM due to the popularity of the GaAs/AlGaAs and InAs/GaAs multilayer structures. However, although B-doped [82] and Tl-doped GaAs [83] have been grown, these materials are still hardly studied and have not been explored by STM as far as we know.

For indium, typically a bright contrast is observed in empty state imaging conditions [84, 4], and for filled state imaging at low indium concentrations there appears to be no indium contribution to the image [4]. This is shown in figure 4.15(a) and (b). The features marked A in figure 4.15(a) are surface indium atoms seen using empty state STM imaging. The corresponding location in figure 4.15(b), the filled state STM image, show no contribution from the indium atoms. This agrees with our simulations which also show a bright contrast in the empty state imaging and no indium contrast in the filled state images. Some experimental results for indium in GaAs disagree with this, and show bright spots

![Figure 4.15](image_url)

Figure 4.15: Experimental STM images of group III dopants on the (110) surface of GaAs: indium empty state, (a); indium filled state (b); aluminium empty state, (c). Images (a) and (b) were reproduced from [4] and are of the same surface region with a doping density of In$_{0.04}$Ga$_{0.96}$As. In (a) the atomic contrast is primarily from the Ga atoms and in (b) the contrast is primarily from the As atoms. Image (c) was reproduced from [5]. The regions of dark contrast are caused by the aluminium atoms, and the areas of light contrast are from the Ga atoms. The doping concentration was Al$_{0.3}$Ga$_{0.7}$As.
in the filled state imaging [85, 86], however these results typically contain very high concentrations of indium, or are systems with very high lattice strain such as quantum dots or rings.

In empty state imaging conditions a dark contrast is typically seen for Al in Al$_{1-x}$Ga$_x$As as seen in figure 4.15(c) and Refs. [87, 5], but this is not reproduced in the simulations. For the majority of Group III dopants the large difference in size between the dopant and the Ga that it substitutes determines the physics and hence the contrast in the STM images. However, as Al is very similar to Ga in size, more subtle electronic effects may also play a role in determining the contrast. In this case the experiments in Refs. [87, 5] may not be directly comparable to our calculations; in the experiments where the concentration fraction is very high ($x > 0.3$) we would expect much larger electronic interactions between Al atoms than are present in our simulation where $x = 0.016$.

### 4.4.2 Group V STM Images

Figure 4.16 shows the series of filled state STM images generated for the five group V atoms on the (110) surface. As in figure 4.13, the increasing size of the impurities down the group moves the relaxed position from below the surface in the case of nitrogen, all the way up through the surface layer ending with bismuth at the highest position. This is seen in the calculated STM images as a change in the intensity of the image over the dopant site. The trend of increasing covalent radii down group V does not have an exception like aluminium in group III, but the overall form of the trend is consistent. In all cases the surrounding STM image is left unperturbed as the relaxation due to the impurity does not spread to other arsenic atoms on the surface.

The empty state images for the group V impurities are shown in figure 4.17,
Figure 4.16: Simulated filled state STM images of Group V impurities on the 1st layer of the GaAs (110) surface: nitrogen, phosphorus, arsenic, antimony and bismuth. The increasing size of the impurities directly relates to the change in the intensity of the image over the impurity site.

and do not show quite the same variation with impurity type as the group III filled states images do (figure 4.14). For the phosphorus and arsenic systems there is no noticeable change in the image due to the minimal relaxation. For the nitrogen case the neighbouring gallium atoms have relaxed inwards with each peak having moved away from the clean position by around $0.44 \text{"Å}$. This results in a redistribution of their empty states in the calculated STM image as seen in figure 4.17. This relaxation also causes the gallium atoms at the edges of the super-cell being forced up slightly as the ones in the middle are pulled down. For antimony and bismuth, the impurity atoms are far enough out of the surface that their associated empty states have not decayed away in strength as much as the surrounding gallium states, and so have a sizeable presence in the image above the impurity location.

All group V isoelectronic dopants that have been studied theoretically in this chapter have been observed experimentally by filled state XSTM imaging, as well as empty state imaging for both N and Sb. In all cases the calculated contrast corresponds nicely with the experimentally observed contrast.
Figure 4.17: Simulated empty state STM images of Group V impurities on the 1st layer of the GaAs (110) surface: nitrogen, phosphorus, arsenic, antimony and bismuth. The impurity location in each case is marked with a cross.

X-STM images show that N doped GaAs [3, 4] gives rise to a dark contrast on top of the N-atom, with Ref. [4] also showing weak bright contrast for empty state images of nitrogen, which agree qualitatively with the results of the simulation shown in figure 4.16. The experimentally observed filled state contrast on P atoms in P doped GaAs [6] shows the same dark contrast predicted by the simulations as shown in figure 4.18(a). In the case of Sb doped GaAs, the XSTM experiments show the predicted bright contrast on Sb atoms for both filled and empty state imaging [7] which is shown in figure 4.18(b) and (c), as is the case for filled state imaging of Bi doped III/V material [8].

To disentangle the various effects that cause the features seen in figure 4.17 we calculated an STS dI/dV curve for sites over the nitrogen and bismuth locations along with clean surface gallium and arsenic sites. These STS results can be seen in figure 4.19.

For the clean GaAs surface we find that the valance bands are primarily dominated by the filled arsenic states, with only a weak contribution from the filled gallium states. For the conduction band we see the reverse; now the gallium empty states dominate over the weaker arsenic empty states. In the case of
bismuth, its proximity to the tip increases the intensity of both its filled and empty states in comparison to the gallium and arsenic ones. In the valence bands the bismuth filled states have a larger magnitude than the arsenic states leading to the bright spot in figure 4.16. In the conduction band the bismuth states also dominate the data. Here the proximity to the tip of the bismuth has increased the intensity of the empty states to the point where they are now stronger than the gallium signals.

For the nitrogen case we see that in the valence band the filled states are diminished in comparison to the arsenic and gallium states. This is due to the relaxation of the nitrogen away from the surface, moving it further away from the tip. The conduction band is more complicated, showing a fluctuating nitrogen signal that sometimes dominates over both the arsenic and gallium states. In figure 4.20 we compare STS data calculated over the nitrogen impurity location,

![Experimental STM images of group V dopants on the (110) surface of GaAs: phosphorus filled state, (a); antimony filled state (b); antimony empty state, (c). Image (a) was reproduced from [6], the background contrast is from the arsenic atoms, and the dark regions bisected by the lines a and b are due to the phosphorus atoms. In images (b) and (c) the background contrast is from the arsenic and gallium atoms respectively, with the bright contrast due to the antimony atoms. The antimony images were reproduced from [7].](image-url)
Figure 4.19: Simulated dI/dV curves taken over clean gallium and arsenic sites (top), a single bismuth impurity site (middle) and a nitrogen impurity (bottom).
over the location of one of the nitrogen’s nearest neighbour galliums and over a clean surface gallium site. The signal seen over the nitrogen matches very closely with that of its neighbouring gallium atoms. This suggests that the nitrogen signal in figure 4.19 is actually the gallium states “spilling” over the nitrogen site due to the surface relaxation.

We can now explain the remaining empty state STM images in figure 4.17. For the nitrogen case the surrounding gallium atoms have relaxed inwards, which is seen as a redistribution of their empty states in the calculated STM image. For antimony and bismuth the impurity atoms are far enough out of the surface that their associated empty states have not decayed away in strength as much as the surrounding gallium states, and so have a sizeable presence in the image.

### 4.5 Bi in InP

In order to see if this behaviour was unique to isovalent impurities in GaAs a similar III-V doped system was simulated that had also been recently looked at experimentally [8]. The system involved isovalent bismuth impurities on an indium phosphide (110) surface. Bismuth has a larger covalent radius than the
phosphorus atom it substitutes and so an area of increased intensity is expected, similar to the case of bismuth in GaAs in figure 4.16. As with the case of nitrogen, bismuth was simulated in the first three layers of a 3×3 InP slab. For the InP slab FHI pseudopotentials that had been used for the indium and phosphorus impurities were employed. The clean slab was initially relaxed giving a lattice constant of 5.82 Å, a buckling angle, $\alpha$, of 28.4° and a perpendicular displacement between the surface atoms, $\Delta_{1,\perp}$, of 0.67 Å. These values agree very well with the results of experiment and other pseudopotential calculations [88, 89, 90], matching the average buckling angle to within 2.4% and the average perpendicular surface displacement to within 0.2%.

In table 4.4 we present the DFT and trilateration data for the first layer bismuth and its nearest neighbours. The nearest neighbour positions used in the trilateration calculation were taken from the clean InP slab positions. We find the differences in position are consistent with the group V dopants in GaAs, and again find the sub-surface nearest neighbour shows the smallest change in position. From the DFT calculation we get an InP bond length of 2.52 Å, which is only 0.4% different from the InP bond length calculated as the sum of the indium

<table>
<thead>
<tr>
<th></th>
<th>DFT positions</th>
<th>Trilateration positions</th>
<th>Difference (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bi</td>
<td>6.173 9.711 13.120</td>
<td>6.173 9.598 13.215</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>4.050 11.253 12.113</td>
<td>4.115 11.254 11.948</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>8.296 11.253 12.113</td>
<td>8.231 11.254 11.948</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>6.173 8.788 10.430</td>
<td>6.173 8.782 10.401</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 4.4: Comparison of the position of the bismuth dopant (first line) and its three nearest neighbours (following three lines) between the DFT simulation and the trilateration calculation. For the trilateration positions the three nearest neighbour positions are clean InP slab positions used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions.
and phosphorus covalent radii, 2.53 Å. We also see very good agreement between the DFT and covalent radii for the In-Bi bond length. The two bond lengths differ by only 3.8%, with the DFT relaxation giving an average In-Bi bond length of 2.82 Å, and the sum of the indium and bismuth covalent radii giving a bond length of 2.93 Å.

In Fig. 4.21 we show the calculated constant height STM images along with diagrams of the relaxed surface for a single bismuth impurity in the first three layers of the InP (110) surface. These images are all taken at a scan height of 4 Å above the clean surface (z = 16.6 Å). For comparison we present an experimental filled state XSTM image in figure 4.22. Three typical bismuth features can be seen on the (110) surface originating from bismuth atoms in the first (A), second (B) and third (C) layers.

In 4.21(a) we see a bright region over the bismuth site. Due to its larger
covalent radius the bismuth dopant has relaxed to a position further out of the surface than its initial substitutional site, which can be seen in 4.21(d). Being closer to the scan height causes the bright spot over the bismuth site in the same manner as in figure 4.16. There is very little geometrical change in the (110) plane for the surrounding phosphorus atoms which leaves the rest of the STM image unperturbed. This corresponds to the feature marked A in figure 4.22.

In 4.21(b) we see a very slight increase in the intensity of the LDOS over the phosphorus atoms surrounding the bismuth. Looking at 4.21(e) we see very little movement of the bismuth away from its initial position with most of the geometry alteration coming from the shrinking of the bond lengths to its nearest neighbours, as was seen in the nitrogen case (4.10(e)). Nitrogen in GaAs causes a 20% change in bond length compared to the arsenic it substitutes whereas bismuth causes a 15% change compared to the phosphorus it substitutes. This smaller change in bond length causes the features seen in the bismuth STM images
to be less pronounced than for nitrogen. The relaxation of the bismuth’s nearest
neighbours causes four surface phosphorus atoms to relax slightly parallel to the
surface plane which can be seen in 4.21(e). The lateral motion is significantly less
than was seen in figure 4.10(e) and so the effect on the STM image is seen as only
a slight increase in the intensity located above those four atoms. Here the average
change in $z$ of the four surface phosphorus atoms is 0.04 Å while for nitrogen it
was on average 0.13 Å. The same feature in figure 4.22 is at B, and although we
do see an increased contrast over the 4 surface atoms it is much more pronounced
than in figure 4.21. This discrepancy may come from the second layer feature
being more sensitive to the applied voltage, with the experimental images being
measured with a different bias voltage than the theoretical ones.

In figure 4.21(c) we show the results for a bismuth impurity in the third layer.
Here the STM image is similar to (a) only with a reduction in intensity over the
bismuth site. Looking at the relaxation in (f) we see only a small displacement of
the bismuth from its initial substitutional position with most of the effects of the
bismuth being seen in the shrinking of the bond lengths to its nearest neighbours.
We see that this has the largest effect on the indium atom closest to the surface.
The displacement of this indium atom pushes up the single surface phosphorus
atom located above it resulting in an STM image that is very similar to figure
4.21(a). This, again, agrees with the features in the experimental image, with C
showing an area of increased contrast over the impurity site, but with a smaller
extent than the first layer feature at A.

Figures 4.21 shows very clear parallels with figure 4.10. Whereas before the
smaller nitrogen impurity caused inwards relaxation of the surface, the larger
bismuth causes outwards relaxations. In both cases the effects are reduced as the
impurity moves further into the slab, and show varying symmetries depending
on which layer the impurity is in. For the second layer cases (figures 4.21(b) and
4.10(b)) the size of the effect on the surface is larger in the nitrogen case where there is a larger change in atom size.

4.6 Summary

In this chapter we have shown that the behaviour and appearance of nitrogen features on the GaAs (110) surface can be qualitatively explained through a simple analysis of the geometry of the system, without having to refer to any other effects. This behaviour has been verified through LDA DFT simulations and compared to experimental results. We find good qualitative agreement between the calculated and experimental STM images and in cases where atomic shifts can be measured we find excellent quantitative agreement. Extending this analysis we find that impurity size and system geometry are the main influence on the STM images for all the isovalent impurities studied. All atoms from groups III and V behave in a predictable manner based on their sizes, which gives rise to clear trends in the relaxations and STM images. We have shown that the effect of the group V (III) impurities is also visible through empty (filled) state imaging. By looking at the behaviour of bismuth impurities in InP we find similar agreement between the geometrical model and DFT calculations, indicating that our model is not limited to GaAs.
Chapter 5

Group IV Donors on the (110) Surface of GaAs

5.1 Introduction

Group IV dopants make ideal donors for solotronic devices. The extra electron of the donor introduces interactions with the host crystal that are not present in the case of isovalent impurities. The extra states introduced by donors or acceptors have been used for a variety of applications such as single photon production from nitrogen acceptors in ZnSe [91] to storing data on phosphorus donors in silicon [92]. The isovalent impurities studied in chapter 4 were an ideal test case for our geometrical model, and in this chapter we extend our analysis of dopants on the GaAs (110) surface to donors from group IV.

The donor electron introduces a number of important charge effects when compared to the isovalent case. This is first seen in the relaxation of the donor, which is now dependant on its charge state. In the previous chapter we applied our simple geometric model to isovalent impurities in GaAs to understand the relaxed positions and STM images of those impurities. In this chapter we apply
the model to group IV donors and find that, while the extra effects of the donor charge play a role in the relaxation, our model still gives accurate predictions of the bond lengths and of the final relaxed positions.

The donor also introduces localised filled states in the band gap. In a bulk environment the localised state is hydrogenic and extends over a radius of \( \sim 100 \ \text{Å} \). It was thought that as a donor moved towards a surface there would be a reduction in the binding energy of this hydrogenic state, increasing the effective radius. However, experimentally it was found that the binding energy of the donor state increased towards the surface, decreasing the extent of the state [93]. This reduction in the extent of the state makes it possible to simulate the system using DFT.

In section 5.2 we outline the DFT calculations used, as well as the various convergence checks that were employed to ensure an accurate calculation. In section 5.4 we present the calculated LDOS images of the group IV band gap states and present our analysis of the results. Finally in section 5.5 we present a summary of this chapter.

5.2 Calculation

The clean GaAs (110) surface from chapter 3 was used as the starting point for the investigation of the dopants. For the carbon, silicon, germanium and tin donors studied here we used pseudopotentials from the same FHI library that was used in previous chapters. The only compatible lead pseudopotentials included the \( d \)-electrons in the valence electrons which would have required much more time to fully converge and may have introduced discrepancies when comparing to the results for the other dopants. For these reasons lead was omitted from the investigation.
The donor state at the surface has a much smaller extent than it has in the bulk, with the probability density decaying to less than 10% of its peak value out at the first nearest neighbour atoms (see figure 5.12). However, despite the localised nature of the surface state care must be taken in constructing an appropriately sized system. To make sure an accurate model of an isolated donor is obtained a rigorous convergence test with respect to system size was performed, in tandem with the usual convergence checks with respect to the value of $E_{cut}$ and $k$-points. As in the clean system and isovalent doped systems in chapters 3 and 4 the system was allowed to fully relax after the addition of the donors with a tolerance on the forces of $1 \times 10^{-4}$ hartrees per bohr.

5.2.1 $E_{cut}$ and k-point Convergence

The convergence of the total energy of the different super-cell systems was initially measured against $E_{cut}$ and the number of points in the $k$-point grid. This test was performed for all group IV dopants in a super-cell with the standard 7 layers of GaAs and 6 vacuum layers. The results for silicon in three super-cells with different $x$ and $y$ dimensions are shown in figure 5.1. The total energy of the system was found to be converged to 1 part in $10^4$ for an $E_{cut}$ of 15 Hartree when compared to a system with an $E_{cut}$ of 18 Hartree. The curves for the three super-cells show that the convergence is almost completely independent of the physical size of the system. This is expected as $E_{cut}$ can be used as a length scale for measuring details in real space. As mentioned in chapter 2, $E_{cut}$ is related to $k$ via

$$E_{cut} > \frac{|k + G|^2}{2},$$

so we can define a real space length scale, $\lambda$ as

$$\lambda = \frac{2\pi}{|k|} \propto \frac{1}{\sqrt{E_{cut}}}. \quad (5.2)$$
Figure 5.1: The fractional change in total energy as a function of $E_{\text{cut}}$ for three different sized Si doped super-cells. The fractional difference in total energy is measured against a calculation with $E_{\text{cut}}^{\text{max}}=18$ Hartree. The total energies for the 56, 126 and 168 atom super-cells at 18 Hartree are -6.65, -14.91 and -19.86 keV respectively. Notice the resemblance to figure 3.6.

Although the systems in figure 5.1 differ in physical extent, the scale of the smallest features they contain remains constant leading to the consistent set of convergence plots as a function of $E_{\text{cut}}$. This is also why the convergence for the doped system is similar to that of the clean system seen in figure 3.6. As the pseudopotentials have removed all the inner core details, only the broad features are left and so the dopant pseudopotential and features should require roughly the same $E_{\text{cut}}$ as the gallium or arsenic atoms, especially considering their similar size.

In figure 5.2 we present the fractional changes in $E_{\text{total}}$ as a function of the number of points in the $\mathbf{k}$-point grid for two different super-cells. For each $\mathbf{k}$-point configuration the total energy is compared to a system with a $5 \times 5 \times 1 \mathbf{k}$-grid. The total energy of both dopant systems was found to be converged to a fractional difference of 1 part in $10^6$ for the $2 \times 2 \times 1 \mathbf{k}$-grid. Because of this high level of
Figure 5.2: Comparison of the convergence of the total energy of two different Si doped systems with respect to $k$-point grid. The convergence is measured as the fractional difference between the total energy at the current $k$-point grid compared to the total energy calculated using a $5 \times 5 \times 1$ $k$-grid, $\frac{E_{\text{total}}(k) - E_{\text{total}}(5 \times 5 \times 1)}{E_{\text{total}}(5 \times 5 \times 1)}$. Lines are shown as a guide for the eye.
Figure 5.3: Energy of the silicon band gap state for three different sized super-cells. Left: convergence of the eigenvalue of the state as a function of $E_{\text{cut}}$. The arrows indicate which axis to read for the datasets. The conduction band edge was at -0.5 eV. Right: change in energy of the band gap state as a function of $E_{\text{cut}}$, $|E_{\text{bgs}}(E_{\text{cut}}) - E_{\text{bgs}}(E_{\text{cut}} - \Delta)|$ where $\Delta$ is the 1 Hartree step size between $E_{\text{cut}}$ values.

For all the subsequent calculations for silicon, germanium and tin the value of...
$E_{\text{cut}}$ was chosen to be 15 Hartree. For carbon the band gap state energy had converged to less than 2 meV for an $E_{\text{cut}}$ of 18 Hartree however when it came to simulating constant height STM images at 4 Å above the surface we found that the STM contrast was highly sensitive to the exponentially decaying tail of the carbon donor state into the vacuum. A value of 30 Hartree was needed to produce a converged STM image and so it is this value that is used in all subsequent carbon calculations.

### 5.2.2 System Size Convergence

Once the individual calculations were converged with respect to the $k$-point grid and $E_{\text{cut}}$ the convergence of the energy of the dopant state with system size was investigated for each dopant. As well as the initial $2 \times 2$ cell, eleven other super-cells were investigated, $2 \times 3$ and $3 \times 2$ (84 atoms), $2 \times 4$ and $4 \times 2$ (112 atoms), $3 \times 3$ (126 atoms), $4 \times 3$ and $3 \times 4$ (168 atoms), $5 \times 2$ and $2 \times 5$ (140 atoms) and $5 \times 3$ and $3 \times 5$ (210 atoms). The values for the covalent radii of the group IV donors are given in table 5.1, along with the covalent radius of gallium for comparison. After the addition of the dopant all super-cells up to 126 atoms were allowed to relax. Any super-cells larger than 126 atoms were constructed from the relaxed positions of the 126 atom super-cell with the additional atoms added in their clean surface positions. This was done as the outer atoms of the 126 atom super-cell had moved

<table>
<thead>
<tr>
<th>Covalent Radii (Å)</th>
<th>C</th>
<th>Si</th>
<th>Ge</th>
<th>Sn</th>
<th>Ga</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.75</td>
<td>1.16</td>
<td>1.21</td>
<td>1.40</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Table 5.1: Experimental covalent radii for group IV atoms along with the substituted gallium atom. All values are taken from [9].
Initially we checked the convergence of the total energy per atom of the system with respect to super-cell size which is shown in figure 5.4. As a point of reference the total energy per atom in the $3 \times 3$ (126 atom) cells for silicon, germanium and tin are all within 0.3 eV of the infinite size limit and the carbon case is within 0.7 eV. In the previous chapter the total energy per atom of the system is given in equation 4.2 as $E_{\text{atom}}(N) = \alpha + (\alpha_{\text{dopant}} - 59.9)N^{-1}$ where $N > 0$. As before we use least squares fitting to determine the coefficients in the equation which are given in table 5.2.

For all four impurities the value of $\alpha$ was calculated to be within $1 \times 10^{-3}$ Hartrees of the expected -117.94 eV/atom for a clean (110) slab in the large $N$ regime [74]. The values of $\alpha_{\text{dopant}}$ were found to be in good agreement with the calculated total energy per atom of carbon, -155.06 eV per atom [94]; silicon, from -106.98 eV per atom [95] to -108.03 eV per atom [96]; and germanium, -106.78

---

**Figure 5.4:** Total energy per atom as a function of the number of atoms. The points are the results of the simulation and the lines are the results of a least squares fitting of the data using the functional form given in equation 4.2.

Table 5.2: Data from the least squares fitting of the total energy data to the functional form given in equation 4.2. The units of $\alpha$ and $\alpha_{\text{donor}}$ are Hartrees per atom.

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>Si</th>
<th>Ge</th>
<th>Sn</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>-117.94</td>
<td>-117.94</td>
<td>-117.94</td>
<td>-117.94</td>
</tr>
<tr>
<td>$\alpha_{\text{donor}}$</td>
<td>-151</td>
<td>-106</td>
<td>-106</td>
<td>-95</td>
</tr>
<tr>
<td>Residual sum of squares</td>
<td>2.1E-6</td>
<td>6.0E-7</td>
<td>1.3E-6</td>
<td>8.2E-6</td>
</tr>
</tbody>
</table>

eV per atom [95].

After the total energy of the system had been converged the investigation turned to the dopant state in more detail, specifically the convergence of the band gap state eigenvalue. For each super-cell the energy of the band gap state introduced by the donor atom was plotted as a function of system size. Figure 5.5 shows the variation of the energy of the silicon dopant state as the size of the super-cell is changed. Each section of the figures has one of the lateral dimensions of the super-cell kept constant as the other is varied. (a) and (b) are for super-cells where the length in the $y$ direction has been fixed at 2 and 3 primitive unit cells respectively while the length in the $x$ direction is increased, and (c) and (d) are for super-cells that have the lengths in the $x$ direction fixed at 2 and 3 primitive cells respectively while the length in the $y$ direction is increased. The surface periodicity for each data point is given in the plots.

The variation in the energy of the dopant state is much more pronounced for changes in the length of the super-cell in the $x$ direction compared to changes in the $y$ direction. This is essentially because the length of the primitive cell in the $y$ direction is $\sqrt{2}$ times longer than in the $x$ direction. Another factor thought to contribute to the asymmetry in the convergence is that the (110) surface geometry has zig-zag chains in the plane of the surface in the $x$ direction which can lead to greater defect-defect interactions than are possible in the $y$ direction [97].
Figure 5.5: Si doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. (a) Varying the super-cell length in the $x$ direction while keeping the $y$ length fixed at 2, (b) same as (a) but with the $y$ length fixed at 3, (c) varying super-cell length in the $y$ direction with the $x$ length fixed at 2 and (d) same as (c) but with the $x$ length fixed at 3. In each case the dimensions of the super-cell are given. The horizontal dashed line marks the energy level in the largest, 210 atom, super-cell, which we take to be the converged energy.
Figure 5.5(a) and (b) generally show larger oscillations than those in (c) and (d), with the biggest energy change coming from a change in the $x$ length of the super-cell. This is true for all donors (figures 5.6, 5.7 and 5.8). Figure 5.5(c) shows that the super-cells with a width of 2 primitive cells in the $x$ direction produce consistently low energies regardless of their length in the $y$ direction. In this case the dopants are only 7.82 Å away from their closest periodic images and so a sizeable dopant-dopant interaction is expected. Figure 5.5(d) shows the least amount of variation with super-cell dimension. Once the super-cell gets above $3 \times 3$ there is very little variation in the energy of the silicon band gap state. This, again, is the case for all the donors investigated here.

In figure 5.6 we see that the variation in the energy of the carbon gap state shows a much bigger dependence on the length in the $x$ direction when compared to the other group IV elements, especially in the case of 5.6(b), but the overall qualitative trends remain the same.

The variation of the energy of the germanium dopant state can be seen in figure 5.7 and for tin in figure 5.8. Both of these follow almost identical trends to silicon. Based on the results for all four dopant types we used the $3 \times 3$ cell for all subsequent calculations. This system had the energy of the gap state converged to better than 10 meV when compared to the $3 \times 5$ for all donors and was small enough to keep the calculation time manageable, down to a couple of hours.

## 5.3 Relaxed Positions

Figure 5.9 shows the side view and top down view of the final relaxed positions of all the group IV donors, along with their first nearest neighbours and the gallium substitutional site. Carbon relaxes into the surface, with a relaxed position roughly 0.2 Å further in than the substitutional gallium site. The bond lengths
Figure 5.6: C doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. See figure 5.5 for details.
Figure 5.7: Ge doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. See figure 5.5 for details.
Figure 5.8: Sn doped GaAs: variation in the energy of the localised state in the gap with respect to super-cell size. See figure 5.5 for details.
Figure 5.9: Comparison of the final relaxed positions of the carbon, silicon, germanium and tin donors as well as their first nearest neighbours. Donor atoms are represented by filled shapes and their nearest neighbour arsenic atoms are represented by the corresponding empty shapes. The initial substitutional gallium site is also marked for comparison. The bonds in the case of the silicon dopant have been shown for reference. The left hand plot is a slice taken through the slab in the (100) plane and the plot on the right is a plan view of the (110) surface. The inset in the left hand plot shows a close-up of the surface nearest neighbour positions and the inset in the right hand plot shows a close-up of the dopant positions.
Chapter 5

99

to the nearest neighbour arsenic atoms on the surface are 1.99 Å and the bond lengths to the arsenic in the next layer down are 1.94 Å. This increases the local surface buckling angle to 37.77° from the clean surface value of 30.37°, and also has the effect of pulling in the nearest neighbour arsenic atoms. This relaxation into the surface is primarily due to the much smaller effective size of carbon when compared to the gallium atom it substitutes.

Because of the extra electron present on the donor the silicon relaxes to a position further out of the surface than the gallium substitutional location despite being an atom with a smaller effective size. This position is also further out than was predicted by the geometrical model (see section 5.3.1). The local buckling angle is reduced to 15.59°, and changes the in-plane nearest neighbour bond lengths to 2.36 Å and the length of the nearest neighbour bond into the surface to 2.34 Å. The two nearest neighbour arsenic atoms in the surface plane experience the largest change in position. These atoms have only three bonds holding them in place, one of which is to the donor atom. The second layer arsenic atom on the other hand is bonded to four other atoms of which two are further into the bulk. This constrains the relaxation of the subsurface atoms, leading to their reduced movement during the relaxations. This data obtained for the relaxation of silicon matches very well with the results of similar DFT calculations [11] and a quantitative discussion of the data is presented in chapter 6.

Like silicon, the germanium and tin donors also move out from their initial substitutional location. Germanium and silicon in particular are closer in radius to each other than any of the other group IV elements, with respective covalent radii of 1.16 and 1.21 Å and hence show very similar behaviour. Germanium moves out of the surface, decreasing the local buckling angle to 11.14° and increasing the bond lengths with its neighbours to 2.44 Å and 2.42 Å. Tin is the group IV dopant with the largest covalent radius investigated and ends up
Figure 5.10: Comparison of the final relaxed positions of the group III (left) and group IV (right) dopants along with their first nearest neighbours. Dopant atoms are represented by filled shapes and their nearest neighbour atoms are represented by the corresponding empty shapes. The dotted lines show the directional trend of the relaxations in the group. The final relaxed positions of the group III dopants lie along a line at roughly $4^\circ$ to the surface normal while the relaxed positions of the group IV donors lie along a line at roughly $15^\circ$ to the surface normal.

furthest out of the surface, decreasing the local buckling angle to $5.84^\circ$, and increasing the nearest neighbour bond lengths to $2.58 \, \text{Å}$.

For the group III isovalent dopants on the gallium site the relaxation was primarily in the $z$ direction, either straight into the surface for dopants smaller than gallium or straight up out of the surface for bigger ones. These relaxed positions are shown again in figure 5.10 along with the relaxed positions of the group IV donors. For group IV we see the final relaxed positions of the dopants
lie on a common line, however this line is tilted in the $y$ direction, away from the neighbouring surface arsenic atoms (figure 5.10 right). The final relaxed positions of the group IV donors lie along a line at roughly $15^\circ$ to the surface normal. This rotation of the relaxation direction is caused by the donor electron. The donor electron interacts with both the bulk and the surrounding arsenic atoms providing an additional effect during the relaxation that was not present in the case of the isovalent dopants.

A way of looking at how the presence of the donor electron effects the relaxation is to consider a dangling bond model of the surface. The gallium atom (or group III isovalent dopants) has only three valence electrons and forms only three bonds at the surface. To minimise the total energy these three bonds all want to be in the same plane separated by $120^\circ$ [98]. The surface arsenic atom (or the group V dopants) on the other hand has five valence electrons and three bonds, leaving two of the electrons to form a full dangling bond. To minimise the energy the four bonds try to obtain a tetrahedral or trigonal pyramidal configuration. The competition between these two effects produces the buckling angle of the clean GaAs surface.

When the dopant is a group IV donor, rather than a group III isovalent dopant, the atom at the gallium site now has a half filled bond which will have a different geometrical configuration that minimises the total energy. In figure 5.11 we illustrate the dangling bond model for the clean GaAs (110) surface, a silicon donor and a carbon donor. For the silicon, germanium and tin donors the preferred geometry is trigonal pyramidal, similar to the arsenic site, and so the atom moves out to reduce the bond angles. This leads to a relaxed position further out of the surface despite the effective size of the silicon donor being smaller than the gallium it substitutes. For the carbon atom the greatly reduced effective size and localised nature of the donor electron cause an inwards relaxation. The
three bonds to the neighbouring arsenic atoms seem to form a sp\(^2\) trigonal planar geometry, with the remaining donor electron in a p configuration at right angles to the trigonal plane (see figure 5.12(e)). This resembles the configuration for the carbon bonding geometry in a C\(_2\)H\(_4\) ethylene molecule [99].

Sabish, Krüger and Pollmann identified three main effects that go into the relaxation of the GaAs (110) surface [100]. The three effects are: the electronegativity of the two different surface atoms, the different bond hybridisations and geometries of surface atoms, and the charge density distribution along the surface bonds. They identify the charge density distribution as being the least important effect in GaAs as the charge density is still located away from the surface anions and thus still covalent. This effect only becomes dominant in systems with ionic bonding. The other two effects they outline match with the description we provided here. In order to minimise the total energy in the presence of the donor electron the final relaxed position of the donor is as far away as possible from the rest of the bulk structure, with the bond

![Figure 5.11: Dangling bond model of the relaxation (schematic). The small dots show the number of electrons in each dangling bond. (a) shows the GaAs (110) surface relaxation as the gallium (white) tries to flatten its bonds and the arsenic (black) tries to make them more tetrahedral. (b) shows how the dangling half filled bond of the silicon donor tries to form a tetrahedral configuration and (c) shows the sp\(^2\)+p hybridisation of the much smaller carbon donor.](image)
angles aligned to obtain a more energetically favourable bonding configuration.

5.3.1 Geometrical Model

As with the group III and V dopants we can use the trilateration calculation from chapter 2 to see if the relaxed position of the donors is also governed by their effective size. In Table 5.3 we present a comparison of the average bond lengths from the DFT relaxations for donors on the GaAs (110) surface, with predicted bond lengths based on the sum of the covalent radius of the donors and their nearest neighbours. We see very good agreement between the DFT values and the sum of the covalent radii, with the average difference being only 0.62% and the largest being 1.12% in the case of tin. We also see the expected trend of increasing covalent radii with increasing atomic number. Using these values, along with the clean surface positions of the nearest neighbours we can calculate the predicted relaxed position using trilateration. These results are presented in table 5.4 along with the DFT positions.

In table 5.4 we see that the sub-surface nearest neighbours show only small deviations from the clean slab positions as was the case for the group V dopants. Here the shift of the sub-surface neighbours is less than 0.05 Å in all cases, with the surface neighbours showing a generally larger displacement of between 0.01

<table>
<thead>
<tr>
<th>Bond Lengths (Å)</th>
<th>C</th>
<th>Si</th>
<th>Ge</th>
<th>Sn</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT</td>
<td>1.96</td>
<td>2.35</td>
<td>2.43</td>
<td>2.58</td>
</tr>
<tr>
<td>Sum of covalent radii</td>
<td>1.96</td>
<td>2.37</td>
<td>2.42</td>
<td>2.61</td>
</tr>
</tbody>
</table>

Table 5.3: Average dopant bond lengths for all group IV donors taken from the DFT simulations compared with bond lengths calculated from the covalent radii in Ref. [9].
Table 5.4: Comparison of the group V dopant positions and their three nearest neighbours between the DFT simulation and the trilateration calculation. In each case the first line is the dopant position and the following three lines are the nearest neighbour positions. For the trilateration positions the three nearest neighbour positions are the ones used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions.
and 0.12 Å. For carbon, the nearest neighbour positions had to be shifted away from their clean surface positions in order to find a solution to the trilateration equations, similar to nitrogen in chapter 4. They were moved towards the DFT relaxed positions, while the sub-surface neighbour was fixed in its clean surface position, until a solution to the trilateration equations was obtained. The good agreement here between the DFT relaxed positions and the trilateration positions indicates that the interpretation we had for the isovalent dopants, that the geometry is important in the relaxation, also applies to the donors. It is worth noting again that this good agreement is obtained without any input from the DFT relaxation in all cases apart from carbon.

5.4 Dopant Density of States

For donors a key area of interest is the localised band gap state that they introduce to the system, and so we looked at the behaviour of these states in detail. Here we present constant height LDOS plots for each of the donors’ band gap state. All the LDOS images of the band gap state presented in this chapter were calculated at the Γ point of the Brillouin zone ($k=(0,0,0)$). The Γ point was chosen as it provides the largest contribution to an experimental STM image [21] and in the limit of an isolated dopant is equivalent to all other $k$-points.

The plots of the LDOS at the dopant location for each donor show how the local environment affects the state. Figure 5.12 shows slices of the LDOS data in the $xy$ and $yz$ planes through the location of the dopant, along with the positions of the donor (blue circles), the three arsenic atoms immediately bonded to the donor (green triangles), and the sub-surface second nearest neighbour gallium atoms (green squares).

Going down the group from carbon to tin there is a clear redistribution of
Figure 5.12: LDOS plots taken in the (110) plane through the dopant location for (a) carbon, (b) silicon, (c) germanium and (d) tin in the 3×3 super-cell. The black dots show the dopant location, green triangles are the bonded arsenic atoms and green squares are the second nearest neighbour gallium atoms. (e)-(h) are slices in the (110) plane.
the density. The region of high density that is located on the slab side of the carbon atom is rapidly reduced, with the density in the vacuum reducing at a slower rate with increasing effective size. Silicon, germanium and tin also display an increase in density along the nearest neighbour bond directions. This redistribution towards the arsenic atoms is primarily due to the reduction of the buckling angle as the dopant atom moves outwards and into line with the surface atoms. The silicon LDOS at the donor height matches well with the results of other DFT calculations [11] (see Chapter 6).

In order to check the robustness of the LDOS images a number of extra calculations were performed. We investigated the effect on the silicon band gap state LDOS of changing $E_{\text{cut}}$, the $k$-point grid, the super-cell size, and the pseudopotential. $E_{\text{cut}}$ was increased to 17 Hartree, the $k$-point grid was changed to a $3\times3\times1$ grid, and the number of atoms in the system was increased to 210. The different pseudopotentials used were taken from the Hartwigs-Goedecker-Hutter (HGH) library on the ABINIT website.

Figure 5.13 shows the silicon LDOS for the different simulations. Panel (a) shows the silicon gap state LDOS from figure 5.12 for the $3\times3$ cell. The simulation with the HGH pseudopotentials is shown in (b), (c) is the calculation with the

![Figure 5.13: Comparison of LDOS at the silicon location for (a) standard calculation, (b) HGH pseudopotential, (c) $E_{\text{cut}}=17$ Hartree, (d) $3\times3\times1$ k-point grid and (e) 210 atom super-cell. Unless stated otherwise all other variables match the standard calculation.](image-url)
increased value of $E_{\text{cut}}$, (d) is the increased k-point calculation and (e) is for the larger 210 atom super-cell. In figure 5.14 we have taken a line scan of the LDOS through the slab at the location of the silicon. Here we can see the quantitative agreement between the five different simulations, with the largest change in peak heights being less than 3%. The inset in figure 5.14 shows the decay of the density into the vacuum on a log scale, with the fluctuations at around 18 Å indicating the level at which we start to see noise in our calculation coming from our choice of $E_{\text{cut}}$.

5.4.1 LDOS Above the Surface

The LDOS of the band gap state was also investigated above the surface plane. LDOS images were generated at a height of 4 Å above the (110) surface, which is taken as the position of the outermost arsenic atoms. As well as LDOS images in the (110) plane, line-scans were made in the z direction to observe the decay

Figure 5.14: Line scan of the dopant LDOS along z, centred on the dopant site for five different simulations. Inset, data in the range 10 to 22Å shown on a log scale to show the vacuum tail of the density.
Chapter 5

109

of the LDOS into the vacuum.

Figure 5.15 shows the LDOS images taken at a constant height of 4 Å above the surface for each dopant. For silicon, germanium and tin the features that were located around the surface arsenic sites are now completely gone and the image is dominated by a single peak that covers the dopant site. In this case the trend among the impurities is an increase in both maximum height and width of the peak going from silicon to tin.

In order to explain the features seen in the LDOS plots above the surface, the density at the $x$ and $y$ coordinate of the donor has been plotted in the $z$ direction in figure 5.16. The inset shows the decay of the LDOS into the vacuum on a log scale indicating the level of convergence of the impurity state wavefunction. The low noise in the vacuum tail of the LDOS gives an indication of the good level of convergence achieved in the system, with the carbon line scan in particular showing the necessity of the 30 Hartree cut-off energy to obtain accurate results at 4 Å above the surface.

In figure 5.16 a trend is seen for the peak position of the LDOS to move away from the dopant location with increasing dopant effective size. This is coupled with a decrease in the height of the peak and a loss of the symmetry perpendicular to the surface. These changes in the LDOS are caused by the change in the local geometry of the dopant. The carbon dopant is furthest into the surface and its LDOS is highly confined around the atom in symmetrical peaks into and out of the surface. As the impurities move out of the surface the LDOS distribution changes; the density now favours the vacuum side of the dopant and spreads out away from the dopant site, this is coupled with a decrease in the height of the peak (see figure 5.12e-h). Because of this spreading, the LDOS of tin, which had a low peak close to the dopant, has a larger LDOS than the other impurities out in the vacuum.
Figure 5.15: LDOS images taken at a height of 4Å above the surface. Circles are the dopant atoms and triangles are the nearest neighbour arsenic atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin.
Figure 5.16: Line scan of the dopant LDOS along $z$, centred on the dopant site. Inset, data in the range 10 to 18 Å shown on a log scale to show the convergence of the vacuum tail of the density. The surface position was taken to be at 12 Å and all STM scans were taken at 16 Å.
Figure 5.17: Filled state constant height STM images of the Group IV impurities on the 1st layer of the GaAs (110) surface. These images were generated at a height of at $4 \, \text{Å}$ above the surface.

5.4.2 STM Images

In figure 5.17 we present simulated filled state STM images for the four group IV impurities. As in the previous chapter these images are generated using the Tersoff-Hamman model given by equation 2.43. For filled state imaging $E_{\text{HO}}$ was the energy of the highest occupied molecular orbital which corresponds to the donor band gap state and $V_{\text{bias}}$ was $-1 \, \text{eV}$, and for empty state imaging $E_{\text{LU}}$ was the energy of the lowest unoccupied molecular orbital and $V_{\text{bias}}$ was $+1 \, \text{eV}$. As before, the scan height for our STM images was $4 \, \text{Å}$ above the unperturbed surface arsenic positions ($z = 16 \, \text{Å}$).

The images show the same pattern as Fig. 5.15, now with the filled arsenic states present. As the effective size of the impurity increases the atoms relax to a higher position out of the surface as was seen in figure 5.9. This is seen in the STM image as an increasing area of LDOS over the impurity site, as the localised filled donor state is moved closer to the tip position. For the case of carbon, the impurity has relaxed far enough into the surface to cause a sizeable change in
position of its two nearest neighbours on the surface. These atoms are pulled
down into the surface by the small carbon impurity, reducing their contribution
to the STM image.

For silicon and germanium, their upwards relaxation combined with their donor
electron causes a region of increased LDOS to be seen between the arsenic
sites. The silicon STM image shows very good qualitative agreement to both
the experimental results [12] and other theoretical simulations [72]. Germanium
relaxes to a slightly higher position out of the surface than silicon, which causes
a slightly higher peak over the donor site. For tin, which relaxes to a greater
height than both the silicon and germanium, we see an even greater peak in the
LDOS above the donor location.

In figure 5.18 we present simulated empty state STM images for the four group
IV impurities. Here we see similar behaviour to that seen in figure 5.17. The
carbon impurity has relaxed far enough into the surface that it appears as a dark
empty region in the STM, with the relaxation causing a slight change in the
density around the surrounding gallium atoms. For the other impurities there is
again a trend of the contrast in the STM image increasing with the effective size
as the larger impurities have a final relaxed position that is closer to the scan
height.

\section{5.5 Summary}

In this chapter we have shown that the behaviour of the donors on the GaAs
(110) surface is heavily influenced by their charge and effective size. We then
calculated simulated STM images for four group IV dopants and determined
group dependent trends in the character of the localised band gap state and the
STM images.
Figure 5.18: Empty state constant height STM images of the Group IV impurities on the 1st layer of the GaAs (110) surface. These images were generated at a height of at 4 Å above the surface.

The variation in the dopant state energy with system size was investigated for a number of different super-cell sizes and for four different impurities. This variation was consistent among all impurities, having converged to better than 6 meV for a 3×3 super-cell with a kinetic energy cut-off of 30 Hartree for the carbon system and 15 Hartree for all other donors, using a 2×2×1 k-point grid. For this size system the total energy per atom of the system was converged to within 10 meV of the \( N \to \infty \) value of -117.94 eV/atom.

In the case of all the donors larger than carbon a reduction in the buckling angle between the dopant and its arsenic neighbours in the surface layer was seen, coupled with an increase in the bond lengths to all three nearest neighbours. For carbon there was an increase in the buckling angle and a decrease in the bond lengths as it moved into the surface. For the four donors investigated there is a clear trend of reduced buckling angle and increasing nearest neighbour bonds for increasing effective size. This is partly to do with the effective size of the dopant; going from carbon to tin the effective size increases, pushing the surrounding
atoms away and moving out of the surface. This is not the full picture however, as the silicon dopant is smaller than gallium. The other cause of this outwards relaxation is the extra electron present on all these donors which pushes the dopant out of the surface in a similar way to the arsenic atoms during the clean (110) relaxation.

The local density of states for the dopant band gap state showed no real variation with super-cell size provided the $x$ dimension is greater than two. At the impurity location there is a reduction in the peak intensity of the LDOS with increasing impurity effective size but with a roughly constant physical extent of the state, extending out to the first nearest neighbours for all four impurities. At a scan height of 4Å above the surface the intensity trend is now reversed, the peak intensity now increases with the effective size of the impurity, with carbon having the smallest peak and tin the biggest. A change in the width of the peak is also seen with the peak broadening going from carbon to tin. STM images of the surface displayed the same trends as in chapter 4, with the larger impurities producing areas of bright contrast due to their final relaxed position being closer to the scan height, while the impurities with smaller effective sizes produce areas of darker contrast due to their final relaxed positions being further into the surface.
Chapter 6

Group IV Acceptors and Group IV Charged Impurities

6.1 Introduction

Group IV dopants in GaAs are not limited to donors. Through the substitution of an arsenic atom each of the group IV atoms can act as an acceptor. This is achieved experimentally by either controlling the temperature and doping concentration or by controlling the concentrations of gallium and arsenic during growth [101]. For group IV doping an excess of arsenic during the crystal growth will lead to gallium vacancies which will be populated with group IV donors, while an excess of gallium atoms will lead to arsenic vacancies which will lead to group IV acceptors. While work has been done with silicon as a donor at and near the (110) surface there does not seem to have been as much interest in its acceptor state or the rest of group IV outside of bulk investigations [37, 102, 103].

In this chapter we follow the same procedure we outlined for investigating the surface donors but apply it to the investigation of surface acceptors. After reaching a converged calculation set-up for all of the acceptors we analyse
the relaxation and local density of states of the acceptor atom and provide an explanation of the dopant behaviour based on its effective size and charge configuration.

We also look at different charged cases of both the donors and acceptors in this chapter. When performing an STM experiment on a surface the voltage on the STM tip can induce charged states in the dopant atom. This effect has been used by Garleff et al. to control the charge state of a silicon donor on the GaAs (110) surface [104, 10]. Different charged states can be obtained depending on both the location and applied voltage of the STM tip. The electrons that tunnel into or out of the localised dopant state cause a change in the local potential of the system and the valence configuration of the dopant which gives rise to both a change in the geometry of the surface around the location of the dopant and a change in the appearance of the LDOS of the dopant. The charged states investigated in this chapter provide an interesting comparison to the group III and V dopants as in some of the charged configurations the group IV dopants will be isovalent.

The layout of this chapter will be as follows: in section 6.2 we outline the calculation of the acceptor systems and look at the relaxation and LDOS images obtained, with comparisons to the donor behaviour. In section 6.3 we look at the charged cases of the dopants in both the donor and acceptor configurations and in section 6.5 we give a summary of this chapter and our conclusions.

### 6.2 Group IV Acceptors

The acceptor states are generated through the substitution of an arsenic atom, leaving the final system with one less electron. After the addition of the acceptor the systems were fully relaxed and then systematically checked in the same manner as the donors to ensure a converged calculation. In figure 6.1 we show
Figure 6.1: Convergence of the total energy with respect to $E_{\text{cut}}$ for the silicon acceptor in a 126 atom super-cell and a $2 \times 2 \times 1$ k-grid.

the convergence of the total energy with respect to $E_{\text{cut}}$ for a 126 atom silicon acceptor system with a $2 \times 2 \times 1$ k-point grid. The convergence is almost identical to that of the donors and isovalent impurity systems, showing a change of only a few parts in 1000 for a system with an $E_{\text{cut}}$ of 15 Hartree. For the rest of the group IV acceptors we see the same level of convergence as in the donor systems, and we used the same values of $E_{\text{cut}}$ as in the donor cases, 30 Hartree for carbon and 15 Hartree for silicon, germanium and tin. The convergence with respect to k-grid was also found to be the same as in the donor cases, with all converged systems in this chapter being run with the same $2 \times 2 \times 1$ grid. This same level of convergence for both the donor and acceptor cases is expected as there have been no new features introduced to the system that would require a higher number of plane waves to accurately describe, and the periodicity and scale of the system has not changed so the same k-grid performs equally well.

Convergence checks on the localised acceptor state with respect to system size were also carried out and a $3 \times 3$ super-cell was found to once again give
Figure 6.2: Comparison of the acceptor state of a silicon atom on the surface of GaAs. The left panel shows a simulation using a $3 \times 3$ super-cell and the right panel shows a simulation using a $3 \times 4$ super-cell. The position of the silicon atom is marked with a black dot. The energy difference between these two states is 5.17 meV.

excellent levels of convergence. Figure 6.2 shows the comparison between the silicon acceptor state in a $3 \times 3$ 126 atom super-cell and a $3 \times 4$ 168 atom super-cell, where the difference in the energy of the localised acceptor state is only 5.17 meV.

### 6.2.1 Relaxations

The relaxed positions of the dopants were calculated using a full DFT relaxation, with all the same conditions and tolerances that were used in the previous calculations. In figure 6.3 we present the final relaxed positions of all the acceptor systems along with the first nearest neighbour positions. As in the case of the donors we find that carbon ends up furthest into the surface, silicon and
germanium have similar final relaxed positions, near to the substitutional site, and tin is the furthest out of the surface. The acceptors have one electron not used in bonding and so prefer a tetrahedral configuration like arsenic. Despite having the same covalent radius as arsenic the germanium atom relaxes inwards, reducing the buckling angle compared to the clean GaAs surface. This is due to the half filled dangling bond of the germanium atom causing the bonds to have a less tetrahedral configuration than for the arsenic, shifting slightly towards a more planar configuration. For arsenic the bond angle in the $xy$ plane is $89^\circ$, while for germanium it is $84^\circ$. We also see the same slight displacements in the nearest neighbour atoms, with the sub-surface atoms moving the least due to their full set of four bonds fixing them in place and the surface neighbours moving the most due to them only having two non-dopant bonds.

As with the group III and V dopants we can use the trilateration calculation.

![Figure 6.3](image)

Figure 6.3: Comparison of the final relaxed positions of the carbon, silicon, germanium and tin acceptors as well as the gallium atoms bonded to them. Dopant atoms are represented by filled shapes and the gallium atoms are represented by the corresponding empty shapes. The initial substitutional arsenic site is marked as a reference.
to see if the relaxed position of the acceptors is also governed by their effective size. In Table 6.1 we present a comparison of the average bond lengths from the DFT relaxations for acceptors on the GaAs (110) surface, with predicted bond lengths based on the sum of the covalent radius of the dopant and their nearest neighbours. As before the carbon neighbours had to be moved slightly from their clean slab positions to obtain a solution to the trilateration equations. Here we see slightly worse agreement between the DFT values and the sum of the covalent radii than for the donor cases. The average difference is 3.02%, whereas before it was 0.62%, with the largest being 3.33% for carbon. This could be caused by the gallium pseudopotential having a slightly less accurate covalent radius compared to the arsenic one. The gallium covalent radius being 0.07 Å too small would account for the 3% difference. Using the covalent radii values, along with the clean surface positions of the nearest neighbours we can calculate the predicted relaxed position using trilateration. These results are presented in table 6.2 along with the DFT positions. The larger bond length error for the group IV acceptors is evident in the poorer estimate of the relaxed acceptor positions when compared to the donor cases.

<table>
<thead>
<tr>
<th>Bond Lengths (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
</tr>
<tr>
<td>DFT</td>
</tr>
<tr>
<td>Sum of covalent radii</td>
</tr>
</tbody>
</table>

Table 6.1: Average dopant bond lengths for all group IV acceptors taken from the DFT simulations compared with bond lengths calculated from the covalent radii in Ref. [9].
<table>
<thead>
<tr>
<th>Element</th>
<th>DFT positions</th>
<th>Trilateration positions</th>
<th>Difference (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1.956 4.309 11.179</td>
<td>1.955 4.275 11.170</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>0.190 5.030 11.234</td>
<td>0.200 5.050 9.823</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>3.721 5.030 11.234</td>
<td>3.710 5.050 11.768</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>1.955 2.880 9.868</td>
<td>1.955 2.807 9.853</td>
<td>0.07</td>
</tr>
<tr>
<td>Si</td>
<td>1.955 4.225 11.792</td>
<td>1.955 3.938 11.970</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>−0.104 5.176 11.346</td>
<td>0.000 5.145 11.276</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>4.015 5.176 11.346</td>
<td>3.910 5.145 11.276</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>1.955 2.825 9.885</td>
<td>1.955 2.807 9.853</td>
<td>0.04</td>
</tr>
<tr>
<td>Ge</td>
<td>1.955 4.236 11.845</td>
<td>1.955 3.888 12.052</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>−0.126 5.182 11.354</td>
<td>0.000 5.145 11.276</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>4.037 5.182 11.354</td>
<td>3.910 5.145 11.276</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>1.955 2.831 9.888</td>
<td>1.955 2.807 9.853</td>
<td>0.04</td>
</tr>
<tr>
<td>Sn</td>
<td>1.957 4.236 12.160</td>
<td>1.955 3.718 12.331</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>−0.207 5.226 11.401</td>
<td>0.000 5.145 11.276</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>4.122 5.226 11.401</td>
<td>3.910 5.145 11.276</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>1.956 2.858 9.933</td>
<td>1.955 2.807 9.853</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 6.2: Comparison of the group IV acceptor positions and their three nearest neighbours between the DFT simulation and the trilateration calculation. In each case the first line is the dopant position and the following three lines are the nearest neighbour positions. For the trilateration positions the three nearest neighbour positions are the ones used as input to the trilateration procedure. The difference is measured as the Euclidean distance between the two positions.
6.2.2 Local Density of States

As before, the LDOS for each of the acceptors were investigated, both in the plane of the dopant as well as at a height of 4 Å to see how the different acceptor states behave. Figure 6.4 shows \( xy \) and \( yz \) slices of the LDOS taken through the location of the acceptor.

Like in the case of the donors, the main peak of the density is located on the vacuum side of the dopant. Going from silicon to tin we see the same reduction of the peak intensity and spreading of the density that was present for the donor cases. Where the donors had regions of high density along the directions of the surface bonds however we now see a region of high density between the surface bonds. The reason for this difference comes from the relative position of the dopant and thus the LDOS slices. In the donor case silicon, germanium and tin all move out of the surface enough to move into the surface plane allowing the \( xy \) slice to cut through the donor-arsenic surface bonds. In the acceptor case however the dopant location is a lot higher and so the region of density located along the bonds to the surface gallium atoms is missed out by the plot. The carbon atom shows the same roughly symmetrical lobe structure as in the donor case that resembles the \( sp^2+p \) hybridisation of ethylene, however here the carbon did not relax far enough into the surface to bring all three of its neighbouring bonds into the same plane.

In figure 6.5 we have taken a slice through the LDOS at a height of 4 Å above the surface. The silicon, germanium and tin states show a similar LDOS distribution to the donor cases, however the acceptors have larger intensities due to their increased proximity to the scan height. This also leads to the density spreading over a larger area. Carbon now shows an area of density over the acceptor site that was not present in the donor images due to the depth of the carbon donor
Figure 6.4: (a)-(d) are LDOS plots taken in the $xy$ surface plane through the acceptor location for carbon, silicon, germanium and tin. (e)-(h) are slices in the $yz$ plane. The black dots show the acceptor location, the green triangles are the bonded gallium atoms and the green squares are the second nearest neighbour arsenic atoms.
Figure 6.5: STM plots for (a) carbon, (b) silicon, (c) germanium and (d) tin acceptors taken at a height of 4Å above the surface.
below the surface. Although the carbon acceptor is closer to the scan position than in the donor case it is still below the neighbouring atoms, and its density is localised enough around the acceptor position that the signal at this height is very weak.

In figure 6.6 we show a line scan along $z$ through the $x$ and $y$ location of the dopant, showing the distribution of the density and on the log scale insert how the density decays into the vacuum. A very similar distribution is seen to the case of the donors, with the LDOS having a large peak at a slightly higher $z$ position than the dopant location and then decaying into the vacuum. The graph is very similar to figure 5.16 with a few main differences. The maximum height of the LDOS around the carbon dopant is reduced and seems to fall away faster than in the donor case showing a larger difference in intensity between the carbon LDOS and the silicon and germanium LDOS.

### 6.3 Group IV Charged States

As stated in the introduction STM experiments can induce charged states in a dopant atom. The addition or removal of an electron will not only change how the dopants interact with the surface from a charge point of view, but also what bond geometry will be favourable under the new valence configuration. For each of the four dopants from group IV both the positive and negative charge configuration will be investigated at the donor and acceptor sites by adding or removing an electron from the simulation.

The new systems were once again converged with respect to $E_{\text{cut}}$, $k$-grid and system size. The convergence of the total energy of the system is unchanged by the addition or removal of an electron, as can be seen in figure 6.7.

For $k$-grid and super-cell size we also saw the same levels of convergence as with
Figure 6.6: Line scan of the acceptor dopant LDOS along $z$, centred on the dopant site. Insert, data in the range 11 to 18Å shown on a log scale.
previous calculations so the covered system for the charged calculations was the same 126 atom super-cell with a $2 \times 2 \times 1$ $k$-point grid and an $E_{\text{cut}}$ of 15 Hartree for silicon, germanium and tin and 30 Hartree for carbon. In figure 6.8 we show a comparison of the localised state associated with a negatively charged silicon donor in both the converged $3 \times 3$ super-cell, (a), and a larger $4 \times 3$ super-cell, (b). Here the difference in energy between these two states is less than 3 meV.

### 6.3.1 Relaxations

The charged systems underwent a full DFT relaxation in the same manner as all previous simulations. In figure 6.9 we show the relaxations for the negatively and positively charged group IV dopants on the gallium donor site, along with the neutral donor relaxations for comparison.

For the negatively charged case the addition of an electron causes the dopant valence configuration to match that of the arsenic atoms, and an outwards

---

**Figure 6.7**: Convergence of the total energy versus $E_{\text{cut}}$ for each of the charged configurations of the silicon impurity. All simulations were run for a 126 atom super-cell with a $2 \times 2 \times 1$ $k$-grid.
Figure 6.8: Comparison of the empty state LDOS for the negatively charged silicon donor in (a) a 126 atom cell and (b) a 168 atom cell. The top panels show the surface plane and the bottom panels show a side view. The silicon position is marked with a black dot.
Figure 6.9: Comparison of dopant positions for negative (left), neutral (middle) and positive (right) charge states on the donor site. Dopant atoms are represented by filled shapes with the neighbouring arsenic atoms being represented by the corresponding empty shapes. The initial substitutional gallium site is also marked as a reference.

relaxation of the dopant is clearly seen. Carbon still relaxes inwards due to its effective size, but its final relaxed position is slightly further out of the surface than in the neutral case. For the other three dopants both the repulsion of the filled dangling bond and preferred tetrahedral geometry cause a large movement which has the effect of reducing the buckling angle of the surface as both surface atoms now have the same electronic configuration. The overall vicinity of the final positions of the dopants are based on the arsenic-like valence electron configuration, with the exact positions varying depending on their effective size.

For the positively charged case there is a trend in the positions of the dopant towards the site of the original substituted gallium atom. In this situation the donors now have the same valence charge configuration as gallium, but with a different effective size depending on the atom. In this case the preferable bonding geometry seems to match the trigonal planar geometry of the substituted gallium atom. The difference in effective size of the donors limits how close they can get to the original gallium position, and leads to the spread in positions seen. The silicon and germanium atoms, having a comparable covalent radius to gallium,
end up closest to the gallium site whereas the smaller covalent radius of the carbon and the larger covalent radius of the tin atoms cause them to end up at positions further into and out of the surface respectively.

In figure 6.10 we show the results of the charged cases for dopants in the acceptor positions, along with the neutral acceptor positions. For the negatively charged case, like in the case of the donor, the dopant has an arsenic-like electron configuration and so the dopants move out towards the initial arsenic site. The extra electron that is not present in the neutral case causes them to be repelled further away from the bulk, and rotates the bonds to minimise energy. The rough line along which all the positions lie now passes closer to the substituted arsenic position than in the neutral case, as the acceptors now have the same valence configuration as arsenic.

When the acceptor state is positively charged we see similar behaviour as in the case of the negative donors. In this situation however the valence configuration of both surface atoms is gallium-like, so the dopants move into the surface, reducing the buckling angle and becoming level with the surface gallium positions. Now
that the dangling bond is empty and they have only three valence electrons the
dopants try to obtain the same trigonal planar bonding configuration as their
neighbouring gallium atoms, but are ultimately limited by their effective size and
the similar behaviour in the neighbouring gallium.

Table 6.3 shows the absolute change in position of the various charged cases
from their neutral location. In all four charged configurations carbon shows the
smallest amount of movement away from its initial neutral positions. In this case
the small effective size of carbon dominates the relaxation, with the charge effects
only causing minimal perturbations away from the neutral position. We see there
is also a trend for the acceptors to move more than the donors. This is mainly
caused by the different surrounding structures in the donor and acceptor cases.
The donors are generally below the arsenic atoms further into the surface, while
the acceptor positions are higher up and so less constrained by the presence of
their nearest neighbours.

We can compare our results with recent work by Yi, Ma and Rohlfing [11],
who investigated the behaviour of silicon on the GaAs (110) surface with specific
analysis of the charge distribution of positive, negative and neutral silicon on the
gallium site. They provide data for the final relaxed positions and LDOS for the
silicon donor in all three charged cases which we can compare to the results of our
systems. In table 6.4 we show our data and the data from Ref. [11] for the bond

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>Si</th>
<th>Ge</th>
<th>Sn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Acceptor</td>
<td>0.104</td>
<td>0.442</td>
<td>0.470</td>
<td>0.614</td>
</tr>
<tr>
<td>Negative Acceptor</td>
<td>0.218</td>
<td>0.497</td>
<td>0.536</td>
<td>0.534</td>
</tr>
<tr>
<td>Positive Donor</td>
<td>0.026</td>
<td>0.420</td>
<td>0.481</td>
<td>0.550</td>
</tr>
<tr>
<td>Negative Donor</td>
<td>0.079</td>
<td>0.477</td>
<td>0.466</td>
<td>0.468</td>
</tr>
</tbody>
</table>

Table 6.3: Absolute displacement from the neutral dopant position given in Å.
Surface bonds 2.45 2.46 2.36 2.37 2.31 2.36 
Sub-surface bonds 2.41 2.42 2.34 2.36 2.30 2.31 
Height above 2\textsuperscript{nd} layer 2.13 2.20 1.82 1.78 1.46 1.49 

Table 6.4: Comparison of silicon surface and sub-surface bond lengths data for negatively, neutral and positively charged cases, as well as their height above the sub-surface layer. All values are in \textgreek{A}.

lengths and height above the sub-surface layer for the different silicon charged states. The data shows very good agreement, with an average difference between bond lengths being less than 0.03 \textgreek{A}. When comparing to their plots of the charge distribution we also find very good qualitative agreement. We see the same key features in our LDOS plots as are presented in their work. Both our work and theirs agree that going from the neutral to negative case gives a reduction in the density in the region between the surface neighbours which we have attributed to the change in geometry of the system with respect to the $xy$ plane of the scan. We also agree with their results, showing a large change in the distribution of the density when going from the neutral to the positively charged case, which we attribute to a change in the bond geometry.

### 6.3.2 Local Density of States

As with the case of the neutral dopants the LDOS for each of the charged systems was investigated. In all cases the density imaged is that of the localised dopant state whether it is occupied or unoccupied.

For the negatively charged cases shown in figure 6.11 the states imaged are the fully occupied local dopant states. As in the neutral donor case the carbon
density shows a highly localised set of lobes perpendicular to the neighbouring bonds. Other than carbon the donor atoms have relaxed out of the surface to positions in line with the neighbouring arsenic atoms. The \(yz\) slices of the LDOS generally show a high case of asymmetry with a high density peak away from the bulk similar to the neutral donor case. Silicon, germanium and tin all show a similar density distribution as the neutral tin case in figure 5.12, they are all now roughly in line with the surface arsenic atoms and the density shows up along the surface bonds. The extra electron present in the now filled dangling bond causes the intensities seen in figure 6.11 to be generally larger than in figure 5.12.

The LDOS for the positive donors is shown in figure 6.12, here we image the localised empty dangling bond state of the donor. Here all the dopants show a very similar density distribution; a symmetrical lobe-like distribution perpendicular to the plane of the nearest neighbour bonds. They all resemble the carbon distributions seen in the neutral donor case. The three filled nearest neighbour bonds have flattened into the trigonal planar configuration leaving the remaining empty state that is imaged here resembling a \(p\) orbital. The lobes spread out further with the increasing effective size of the dopant and, other than carbon, show roughly consistent intensities.

The filled negative acceptor states are shown in figure 6.13. Here we see the most prominent asymmetry of the localised carbon lobes in the \(yz\) plane of all the carbon cases. The vacuum side of the density has a larger size and a larger peak intensity than the slab side. The addition of the extra electron distorts the previously preferred \(sp^2 + p\) -like bond hybridization geometry that has been present in all carbon cases up until now, towards a more tetrahedral configuration that has been displayed by the other group IV dopants. The rest of the dopants continue the trend of the peak of the LDOS decreasing in intensity and increasing spatial distribution in the \(yz\) plane as the effective size of the dopant increases.
Figure 6.11: Comparison of LDOS for all four elements in the case of negatively charged donors. (a) to (d) are slices through the dopants in the $xy$ plane and (e)-(h) are slices in the $yz$ plane. Going from left to right is carbon, silicon, germanium and tin.
Figure 6.12: Comparison of LDOS for all four elements in the case of positively charged donors. (a) to (d) are slices through the dopants in the \( xy \) plane and (e)-(h) are slices in the \( yz \) plane. Going from left to right is carbon, silicon, germanium and tin.
As with the neutral acceptors the $xy$ slice is far enough above the surface atoms that there are only faint details present in the density around the neighbouring atoms.

The positively charged acceptor states are shown in figure 6.14, where we have imaged the localised empty dangling bond states associated with the dopants. Here the density distribution resembles that of the positive donor cases (figure 6.12) with all four of the acceptors showing the roughly symmetrical distribution into and out of the surface. The deficit of negative charge causes the acceptors to move into the surface, and the three remaining bonds try to flatten into the same plane. As the effective size of the dopants limits how far they can relax into the surface there is slight preference towards the vacuum side for the density seen in the $yz$ slices.

We also took slices of the data at a height of 4 Å above the surface which are shown in figures 6.15, 6.16, 6.17 and 6.18. The negative donor states in figure 6.15 are very qualitatively similar to those of the neutral case (figure 5.15). The main differences are that for the donors other than carbon the density shows a more circular distribution, whereas before it was a more oval shape. This change in shape is likely due to the height of the donor above the surface, which puts less constraints on the density distribution. In the neutral case tin was already starting to display a more circular distribution and now that silicon, germanium and tin are all above the neutral tin position their circular distribution is more apparent. The density seen here is also of a larger intensity, due to the closer proximity of the donor to the scan height.

For the positively charged donors in figure 6.16 the dopant positions are further into the surface than for any of the other charged or neutral systems and are well below the height of their nearest neighbours in the surface layer. This depth causes a change in the appearance of the density and we no longer see just a large
Figure 6.13: Comparison of LDOS for all four elements in the case of negatively charged acceptors. (a) to (d) are slices through the dopants in the $xy$ plane and (e)-(h) are slices in the $yz$ plane. Going from left to right is carbon, silicon, germanium and tin.
Figure 6.14: Comparison of LDOS for all four elements in the case of positively charged acceptors. (a) to (d) are slices through the dopants in the \(xy\) plane and (e)-(h) are slices in the \(yz\) plane. Going from left to right is carbon, silicon, germanium and tin.
Figure 6.15: LDOS images for the negatively charged donors taken 4Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour arsenic atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin.
peak over the location of the dopant. As the effective size of the dopant increases the density shifts from the neighbouring surface atoms to the dopant position. For carbon the density is almost entirely located over the two neighbouring surface atoms, and for tin it is almost completely localised above the dopant site. Silicon and gallium display a density distribution over both the dopant and neighbour positions, with the peaks in the silicon density being above the neighbour sites and the peak in the germanium case being over the dopant site. This behaviour is due to both the increasing spread of the density into the vacuum as the effective size of the dopant increases, and the gradual movement of the dopant out of the surface. Here the shape of the peak in the density for tin is clearly being influenced by the surrounding atoms.

For the negatively charged acceptor states the dopant is closer to the scan
Figure 6.17: LDOS images for the negatively charged acceptors taken 4 Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour gallium atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin.

height than for any other system and thus have the densities with the largest intensity. Looking at the $yz$ scan of the density in figure 6.13 we see that the density shows the most pronounced rotations towards the left side of the acceptor as the dopant increases in effective size. This is seen as a distortion of the circular density distribution in the slices taken 4 Å above the surface, with carbon showing the least distortion and tin showing the most.

For the positively charged acceptor states we see the usual trend of increasing spread and intensity of the density as the dopant atom increases in effective size, as well as a slight distortion of the circular distribution. The distortion occurs around the positions of the two nearest neighbour atoms on the surface ($y \approx 5$ Å in figure 6.14). A similar effect is seen in figure 6.15 at $y \approx 4$ Å. In both cases the dopants are roughly level with their surface neighbours which is causing a
Figure 6.18: LDOS images for the positively charged acceptors taken 4 Å above the surface. Black circles represent the dopant and the green triangles represent the nearest neighbour gallium atoms. (a) carbon, (b) silicon, (c) germanium and (d) tin.

noticeable effect in the density distribution.

6.4 Summary

The group IV dopants were investigated as acceptors through substitution of an arsenic atom. During relaxation of the system a general inwards movement was seen from the initial substitutional location due to the local deficit of charge. The one exception to this was tin, which ended up at a position further out from the surface. This can be attributed to the size of the atom, which physically prevents the tin from relaxing further into the surface. The LDOS of the localised states around the position of the acceptor showed a peak on the vacuum side of the acceptor as in the case of the donors. When a slice was taken at a height of 4 Å
the behaviour seen in the density followed the same trend as in the case of the donors - an increase in both the size and intensity of the peak density located above the dopant site as the dopants increase in effective size.

We then turned our attention to the charged states of the group IV dopants, for both the donor and acceptor configurations. The final relaxed positions of the different dopant types under different charge configurations could be understood by a comparison to the behaviour of gallium and arsenic on the clean surface. The positive donors behaved like the gallium atoms, moving towards the original substitutional site. Likewise the negative acceptors took on arsenic characteristics and moved towards the initial arsenic location. For the negative donor and positive acceptor cases the dopant matched the valence configuration of its surface neighbours. The relaxation minimised the surface buckling angle as much as the effective size of the dopant would allow, drawing the dopant level with its surface neighbour.

Looking at the LDOS for all of the different charge states, along with those of the neutral cases (figures 5.12 and 6.4), showed consistent trends in the behaviour of the group IV dopants. This behaviour is explained by considering the effective size, electrostatic nature, and the preferential bond geometry of each dopant in terms of energy minimisation. Dopants with a dangling, unbonded electron try to relax out of the surface, if their effective size permits it, to minimise the Coulomb interaction of the electron and the bulk. Dopants with more than one unbonded electron will relax further out as the size of this effect is increased. As well as moving out they will also try to alter the directions of their bonds to minimise energy, predominately preferring a tetrahedral configuration. Dopants with empty dangling bonds tend to relax inwards, as far as their effective size will allow. Without the extra electron the group IV dopants only have three bonds and the preferential bonding geometry to minimise energy now resembles an $sp^2$
planar geometry.
Chapter 7

Conclusions

In this thesis we have presented a systematic investigation of 12 different single atomic impurities on the (110) surface of GaAs. By performing accurate LDA DFT simulations and comparing to available experimental results we have shown that the behaviour of each of the impurities on the surface can be qualitatively explained primarily through a simple analysis of the geometry of the system and the effective size of the impurity. This simple analysis based on the geometry allows us to obtain very accurate initial estimates of the relaxations of all of the dopants from groups III, IV and V. It also allowed us to make qualitative predictions about the contrast of STM images for a wide range of different dopants, as well as providing an explanation of why we see what we do in both the theoretical and experimental STM images. We obtain excellent qualitative and quantitative agreement for both well-established systems, such as silicon and nitrogen in the GaAs (110) surface, and systems that have not received the same levels of experimental or theoretical investigation.

In chapter 4 we simulated nitrogen dopants in the first three layers of the (110) surface, showing that the behaviour of the relaxation, and the resulting STM images of the surface, could be understood simply as a result of the effective
size of the nitrogen atom given by its covalent radius. We obtained qualitative agreement between the calculated and experimental STM images and where the shifts in the relaxed positions of the surface atoms could be measured we found excellent quantitative agreement. Extending the investigation, all the isovalent dopants from groups III and V were found to behave in a predictable way based on their effective sizes, and we presented clear trends in both the final relaxed positions and the contrast of the STM images. We developed a geometrical model to explain why we see the relaxation we do from the DFT simulations, which produced very accurate estimates of the relaxed positions of the group V donors, to within 0.14 Å, however its estimates for the group III donors were more sensitive to the small disagreements between theoretical and DFT bond lengths. When these discrepancies were taken into account however the model was able to produce accurate estimates, on par with those for group V. We briefly turned our attention to bismuth impurities on the InP (110) surface and found similar agreement between the geometrical model, DFT calculations and experimental results, indicating that our model is not limited to GaAs.

In chapter 5 we extended the analysis of the isovalent dopants to group IV donors. These represent a more complicated picture than the isovalent dopants due to their excess electron, however our geometrical and size based analysis still provided very accurate predictions of both the relaxed positions and the STM images. The same trend in the relaxed positions and STM images was seen as in the case of the isovalent dopants, with the larger donors relaxing out of the surface producing bright contrast and the smaller donors relaxing inwards producing dark contrast. The more tetrahedral bond configuration of the donor led to more accurate estimates by the geometrical model than for group III dopants, with the average difference in dopant position being only 0.13 Å.

The donors introduce a band gap state into the band structure and this was
also investigated. The variation in the donor state energy with respect to system size was investigated for a number of different super-cell sizes for each of the donors, and the LDOS of the dopant band gap state was calculated. In each case the LDOS was clearly linked to the geometry, with the carbon donor state tightly confined around the donor atom in two symmetrical lobes at right angles to the planar bonding, while the LDOS around the tin dopant was much more diffuse, spreading out towards the vacuum side in the direction of the tin relaxation.

In chapter 6 we continued our investigation of the group IV dopants, but now acting as acceptors on the arsenic site. The relaxation of the acceptors showed the same expected trend of the final positions being determined by the effective size and charge of the dopants. The group IV acceptors generally relaxed into the surface, with all but tin ending up at a position lower than the initial substitutional location. This difference from the isovalent cases, like with the donor relaxation, was attributed to the difference in valence electron configuration which causes the acceptor relaxation to prefer a different energy minimising geometry. The tendency for the acceptors to favour a more planar geometry led to the estimates from the geometrical model being slightly worse than for the donors, with the estimates on average being within 0.35 Å of the DFT relaxed positions. Looking at the LDOS of the acceptor states we found they showed similar distributions to their donor counterparts, which was again determined by the geometry and predicted by our model.

We then investigated the charged states of the group IV dopants, for both the donor and acceptor configurations. The relaxed positions of the different dopant types under different charge configurations could be understood based on the effective size of the dopants but also by a comparison to the behaviour of gallium and arsenic on the clean (110) surface. The positive donors behaved like the gallium atoms, with relaxed positions around the original substitutional
site, while the negative acceptors relaxed towards the initial arsenic location. For the negative donor and positive acceptor cases the dopants had the same valence electron configuration as their surface neighbours and the final relaxed positions were such that the dopants drew level at the surface to minimise the buckling angle with their neighbours.

For all the dopants studied in this work the STM images have shown consistent trends in the contrast, with impurities with larger effective sizes than the substituted atoms producing areas of bright contrast and impurities with smaller effective sizes than the substituted atom giving rise to areas of dark contrast. This behaviour is explained by considering the effective size, electrostatic nature, and the preferential bond geometry of each dopant in terms of energy minimisation. The isovalent dopants from groups III and V have the same valence electron configuration as the substituted atoms, and so their relaxed positions are predominantly determined by their effective size. For the group IV donors and acceptors the impurities now have a different valence electron configuration to the host gallium or arsenic. The relaxed positions are now determined by both the bonding geometry and effective size of the dopants. The relaxed positions of the donors and acceptors follow the same ordering regardless of charge, with the dopants that have smaller covalent radii being further into the surface and the ones with larger covalent radii being further out. The effect of the charge is simply to shift the overall positions of the dopants to minimise the total energy of the system, while maintaining their order relative to each other.

By showing how the resulting STM images are primarily determined by the geometry of the system we have provided some insight into some of the issues mentioned at the beginning of this thesis with regards to disentangling the geometrical and electronic effects on the of features seen in STM images.

Preliminary work was done on modelling donors from group VI through
the substitution of an arsenic atom and acceptors from group II through the substitution of a gallium atom. With more time we may have been able to fully investigate groups II and VI. We found that the group II and VI doped systems were much harder to converge with respect to system size. For the largest group II systems there were still differences in the dopant state ranging from 50 to 200 meV while for group VI we started to see the convergence trends that were present for group IV, however they occurred at system sizes that were too difficult for us to investigate fully due to the time constraints of the PhD. With more time the calculations for these systems could be fully developed to see if the models we have for explaining the physics of the relaxation and density distribution extend more generally to other species of dopant.

As well as looking at other types of dopant, future work could expand upon the simulations we performed for Bi on the InP (110) surface. By expanding to all of the group V atoms in the InP (110) surface, as well as to dopants in other III-V surfaces we would gain some insight into whether our geometrical interpretation applies to surfaces in general where the dopant has three covalent bonds.

Another possible area for future work is in calculating formation energies. Experimental results have shown that silicon displays a bistable configuration in the GaAs (110) surface, which can be switched through the application of an appropriately charged STM tip [10]. It should be possible to extend the investigation of the charged states in chapter 5 to all the dopants studied in this work, then, through the calculation of their formation energies, we could obtain insight into whether the same bistable configuration exists for other dopants and whether the energy of the switching barrier can be overcome with an STM for controllability. This would require more accurate simulations of the dopant states due to the presence of the STM tip, with the use of time dependent DFT as well as including tip-induced band bending [13, 105].
Appendix A

Example ABINIT Input File

Shown here is the .in input file for ABINIT for the calculation of a 4x3 super-cell with a single silicon donor in the surface layer.

```
ndtset  1
chsymbreak  0
chkprim  0
nsppol  2
mkmem  0

#Definition of the planewave basis set
ecut  15.0
kptopt  1
nshiftk  2
shiftk  0.5 0.0 0.0
       0.0 0.5 0.0
ngkpt  2 2 1
prtden  1
toldfe  3.67d-6  #Hartrees (0.1meV)
```
# Definition of the 4x3 unit cell
acell *5.53 angstrom
rprim sqrt(8) 0.0 0.0
    0.0 3.0 0.0
    0.0 0.0 sqrt(169/8)

# Definition of the atom types
ntypat 3
znucl  31 33 14

# Definition of the atoms
natom 168
typat  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
       1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
       1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
       1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
       2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
       2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
       2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
       2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

# Atomic positions
xangst
*omitted to save space *

# Definition of the SCF procedure
nstep 100

diemac 10.0

diemix 0.33


[8] To be published XSTM experiments performed in the group at the Eindhoven University of Technology.


References


References


References


