Theoretical Modelling and Characterisation of Concerted Tip and Surface Effects in NC-AFM

Bamidele, Joseph Temitope Oladimeji

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to:
• Share: to copy, distribute and transmit the work

Under the following conditions:
• Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
• Non Commercial: You may not use this work for commercial purposes.
• No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Title: Theoretical Modelling and Characterisation of Concerted Tip and Surface Effects in NC-AFM

Author: Joseph Bamidele

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENSE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. http://creativecommons.org/licenses/by-nc-nd/3.0/

You are free to:

- Share: to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Theoretical Modelling and Characterisation of Concerted Tip and Surface Effects in NC-AFM

PhD Thesis

Joseph T. Bamidele

May 7, 2013

A thesis submitted to King’s College London for the degree of Doctor of Philosophy in the Department of Physics, School of Natural and Mathematical Sciences.
Declaring of Authorship

I, Joseph Bamidele, declare that this thesis titled, “Theoretical Modelling and Characterisation of Concerted Tip and Surface Effects in NC-AFM” is the result of work undertaken between October 2009 and July 2012 under the supervision of Prof. Lev N. Kantorovich. I confirm that:

- This work was done wholly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from work of others, the source is always given. With the exception of such quotations, the thesis is entirely my own work.
- Where any work included in this thesis has been done in collaboration with another group or other groups, my own contribution has been clearly indicated.

Signature:

Date:
Abstract

Rising interest in functionalising molecules and surface and building nano-devices, that is nanotechnology, has influenced demand for techniques that image molecules and surfaces with atomic precision. One such technique that has answered this call is Non-Contact Atomic Force Microscopy (NC-AFM), which has a very successful history of directly probing many atomic properties on a wide variety of systems. Interpretation of measurements, however, is highly dependent on knowledge of the atomic probe tip, which is generally elusive and is the most problematic issue surrounding the field. Another such issue is in understanding the atomic processes behind measurements, which to a large extent, also depends on knowledge of the atomic tip termination. Since its development, theorists have been attempting to fill in this knowledge gap, making significant progress, but a substantial gap still remains. This thesis therefore addresses these two issues from a theoretical viewpoint.

After describing the theoretical model used to simulate NC-AFM behaviour, it is used in conjunction with various theoretical techniques to study three important effects on NC-AFM measurements contributing to these issues: different tip terminations, chemical and structural; the relative stabilities of the surfaces and instabilities within the surface; surfaces defects.

The systems studied were the $p(2 \times 1)$ and $c(6 \times 1)$ reconstructions of the oxidised copper (110) surface and the buckled dimer $c(4 \times 2)$ reconstruction of the silicon (001) surface. The stabilities of the two Cu(110):O surfaces were studied thermodynamically, where conventional theoretical methods were found to unexpectedly fail to accurately describe them, and corrections for this were studied. Using this knowledge, NC-AFM simulations were performed using two different chemical tip terminations, and were compared to experimental results, where tip identification was found to be possible on the $c(6 \times 1)$ surface and the $p(2 \times 1)$ surface if a common ad-atom defect is present. Instabilities in the third surface were studied using two different structural tip terminations, where dissipation signals were found to probe the tip-surface interactions in a specific way, due predominantly to soft modes in the surface. The role of defects in NC-AFM manipulations of this surface were then studied, which was found to be crucial to successful manipulation of the surface. As a result of these studies, a tip identification protocol whose successful implementation is novel as well as a novel atomic switching manipulation protocol have been demonstrated.

This research demonstrates the important role of theory in addressing some of the most problematic issues limiting unambiguous interpretation of NC-AFM measurements. The studies have elucidated the atomic processes behind atomic NC-AFM dissipation signals when probing a surface with soft modes and have increase our ability to characterise different atomic tips.
Dedication

This thesis is dedicated to my mother and sister, who have been my pillar of strength throughout my life. They, along with Nora and the late Phil Merritt, auntie Ester, other family members and all my friends, have encouraged and supported me throughout these past three years. It is because of these people that I have been able to achieve this much and reach greater heights.

For the many great memories I have gained while studying for my PhD and for motivating me during my most difficult periods, I thank my fellow PhD students (many of whom have already finished): Dr. Anna Kostouki, Sabrina Beniken, Neil Haria, Dr. Manuela Mura, Kwasi Kwakwa, Dr. Arnold de Lavallaz, Giovanni Peralta, Spyros Sypsas, Dr. Ariadni Vergou and many others.

For access to books, journals and counsel I give my heartfelt thanks to the entire King’s College London physics department, in particular to Julia Kilpatrick and Prof. Jean Alexandre for access to large quantities of beer during our annual ale/beer tasting evenings.

Thanks to all these people I have mentioned (and my sincere apologies to those I should have but have forgotten to include), I have reached this stage in my life, successfully completing this thesis and along with it my PhD. Of course none of this would have been possible without the guidance of a particularly great scientist, my supervisor Prof. Lev N. Kantorovich.
Contents

Abstract
Dedication
Contents
List of Figures
List of Tables
Acknowledgements
Publications and Workshops
Conventions

I Introduction

1 Introduction
   1.1 Historical Overview
   1.2 Scanning Probe Microscopy (SPM)
      1.2.1 The Scanning Tunnelling Microscope (STM)
      1.2.2 The Atomic Force Microscopy (AFM)
   1.3 Non-Contact Atomic Force Microscopy
      1.3.1 Issues with the NC-AFM
   1.4 Aim
   1.5 Overview of Study

2 The Non-Contact Atomic Force Microscope
   2.1 Experimental Setup of the NC-AFM
   2.2 Operation of the NC-AFM
   2.3 Macroscopic Forces
      2.3.1 Macroscopic van der Waals Interactions
   2.4 Microscopic Forces
      2.4.1 Microscopic van der Waals Interactions
3.14.1 Frequency Shift Curves ............................................. 87
3.14.2 Force Curves .......................................................... 89
3.14.3 Dissipation Curves .................................................... 89
3.15 Summary ................................................................. 90

III Results ................................................................. 91

4 Thermodynamic Stabilities of Cu(110):O Surfaces ....................... 92
  4.1 Motivation .............................................................. 92
  4.2 The Cu(110):O Surface ................................................ 93
    4.2.1 Introduction ..................................................... 93
    4.2.2 Previous Studies ................................................ 93
  4.3 General Calculation Details ......................................... 94
    4.3.1 Theoretical Model .............................................. 96
    4.3.2 Additional Variations of the Cu(110):O Surface Reconstructions ...................................................... 97
  4.4 Gibbs Free Energy Differences (GFE) ................................ 98
    4.4.1 PBE and PW ..................................................... 98
    4.4.2 Higher Precision Method: $E_{PBE} + F_{vib}$ ..................... 100
  4.5 The Critical Role of Non-Local Interactions ......................... 101
    4.5.1 Calculation Details .............................................. 101
    4.5.2 Higher Precision Method: PBE-D2 ............................... 102
    4.5.3 Analysis .......................................................... 103
    4.5.4 Summary .......................................................... 107
  4.6 The Effect of Using Non-Local Density Functionals .................... 108
    4.6.1 Introduction ..................................................... 108
    4.6.2 Calculation Details .............................................. 108
    4.6.3 Cu Bulk Properties ............................................. 109
    4.6.4 Higher Precision Method: vdW-DF ............................... 110
    4.6.5 Analysis .......................................................... 111
    4.6.6 Summary .......................................................... 115
  4.7 Application ........................................................... 115
  4.8 Conclusion .......................................................... 116

5 Chemical NC-AFM Probe Tip Fingerprinting on the Oxidised Cu(110) Surface .......................................................... 118
  5.1 Motivation .............................................................. 118
  5.2 Determining the Tip Apices .......................................... 118
    5.2.1 Introduction ..................................................... 118
    5.2.2 Experimental Setup ............................................. 119
    5.2.3 Theoretical Model .............................................. 119
    5.2.4 Imaging the Cu(110):O Surface ................................ 122
    5.2.5 Determining the Tip Types ..................................... 127
    5.2.6 Summary .......................................................... 130
  5.3 Contrast Inversion .................................................... 131
5.3.1 Introduction .................................................. 131
5.3.2 Contrast Inversion when Scanning the c(6 × 2) Surface .............. 131
5.3.3 Contrast Inversion when Scanning the p(2 × 1) Surface .............. 134
5.3.4 Summary ..................................................... 136

5.4 Conclusion ..................................................... 136

6 NC-AFM Manipulation of Oxidised Cu(110) Surfaces ...................... 139
6.1 Motivation ...................................................... 139
6.2 Introduction ................................................... 139
6.3 Theoretical Model ............................................. 140
6.4 Creating the Letter ‘X’ ......................................... 141
6.5 Conventional Methods vs. vdW Corrected Methods ...................... 141
6.6 Diffusion of Super Cu Atom Across an Added Row ...................... 144
6.6.1 Without a Tip ............................................... 144
6.6.2 In the Presence of the O-Terminated Tip .......................... 145
6.6.3 In the Presence of the Cu-Terminated Tip .......................... 146
6.6.4 Tip Assisted Transition - ‘Carrying’ the Super Cu Atom ............ 147
6.7 Summary ....................................................... 150
6.8 Conclusion ..................................................... 150

7 Tip and Surface Effects in NC-AFM Imaging of Si(001) .................. 152
7.1 Motivation ...................................................... 152
7.2 Introduction ................................................... 152
7.2.1 Previous Dissipation Studies .................................. 152
7.2.2 The Buckled Dimer Si(001)-c(4 × 2) Surface ..................... 153
7.3 Theoretical Model ............................................. 153
7.4 Experimental Setup ............................................ 155
7.5 The H3 Tip ...................................................... 156
7.6 The Dimer Tip .................................................. 160
7.7 Comparison of Tip Models ....................................... 163
7.7.1 Tip Deformation ............................................. 164
7.7.2 Hysteresis .................................................. 165
7.8 Preliminary Experiments ......................................... 167
7.9 Summary ....................................................... 171
7.10 Conclusion .................................................... 172

8 The Role of Defects in Controllable NC-AFM Manipulation on Si(001) 175
8.1 Motivation ...................................................... 175
8.2 Introduction ................................................... 175
8.3 Experimental Setup ............................................ 176
8.4 Theoretical Model ............................................. 176
8.5 Toggling the Dimers ............................................ 177
8.6 Forces Required to Flip the Dimers ................................ 179
8.6.1 Variations of the Forces ..................................... 179
<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.7</td>
</tr>
<tr>
<td>8.8</td>
</tr>
<tr>
<td>8.9</td>
</tr>
<tr>
<td>8.10</td>
</tr>
<tr>
<td>IV</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>V</td>
</tr>
</tbody>
</table>

### 8.7 Three-in-a-Row Structure

- 8.8 The Pristine Surface: Tip Induced Variations in Potential Energy Surface
  - 8.8.1 The Stability of the Three-in-a-Row Structure
  - 8.8.2 Correlated Dimer Flips

### 8.9 The Defective Surface: Defect Induced Variations in Potential Energy Surface

- 8.9.1 Non-Local Variations in Surface Stability when Scanning a Defective Surface
- 8.9.2 Two Dimer Vacancy (2DV)
- 8.9.3 Emulating Sub-Surface Defects
- 8.9.4 Strained Surface
- 8.9.5 The Role of Defects in Phason Pair Formation
- 8.9.6 Summary

### 8.10 Conclusion

### IV Summary & Conclusion

### 9 Summary

### 10 Conclusions

### V Bibliography

References
List of Figures

1.1 A STM image of two complete Si 7 × 7 unit cells. .......................... 23
1.2 “The Beginning” - A STM image of Xenon atoms arranged on the Nickel (110) surface. ............................................................. 24
1.3 STM and NC-AFM images of Pentacene. ....................................... 26

2.1 Photograph and schematic of the NC-AFM cantilever-sample holder. ........ 32
2.2 Photograph and schematic of the home-built NC-AFM chamber. ............. 33
2.3 Electron micrographs of NC-AFM cantilever and tip. ............................ 33
2.4 Schematic of the NC-AFM experimental setup. ..................................... 34

3.1 Schematic illustrating the cantilever and sample surface system. ................. 43
3.2 Schematic illustrating the partitioning of the tip into macroscopic and microscopic regions. ................................................................. 45
3.3 Schematics illustrating a point-charge above a single-layer surface and a sphere above a semi-infinite surface. ...................................... 47
3.4 Schematic representation of the self-consistent cycles as obtained by the Hohenberg-Kohn theorems and the Kohn-Sham equations. ......................... 63
3.5 Representation of the three dimensional periodic boundary conditions. ........ 73
3.6 Schematic of the nano-asperity and sample surface. ............................... 75
3.7 Schematic illustrating the key features of the minimum energy path as described by the nudged elastic band method. ............................... 84
3.8 An illustration of the origin of frequency shifts and a schematic of a frequency shift plot. ............................................................... 87
3.9 Schematic of the tip model used to calculate theoretical spectroscopy (∆f-z ) curves. 88

4.1 Possible PBE-relaxed geometries of the Cu(110):O surface. .................... 94
4.2 One possible geometry of the (1 × 1) reconstructed oxidised Cu(110) surface. .......................... 97
4.3 Two possible geometries of the (2 × 1) reconstructed oxidised Cu(110) surface. . 97
4.4 Two possible geometries of the (2 × 2) reconstructed oxidised Cu(110) surface. 97
4.5 Three possible geometries of the (3 × 1) reconstructed oxidised Cu(110) surface. 97
4.6 Five possible geometries of the (6 × 2) reconstructed oxidised Cu(110) surface. . 97
4.7 Gibbs free energy plots for p(2 × 1)-like geometries of Cu(110):O system calculated using PBE. ........................................................ 98
<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.8</td>
<td>Gibbs free energy plots for the Cu(110):O system calculated using PBE.</td>
<td>99</td>
</tr>
<tr>
<td>4.9</td>
<td>Gibbs free energy plots for the Cu(110):O surface calculated using PW.</td>
<td>100</td>
</tr>
<tr>
<td>4.10</td>
<td>Gibbs free energy plots for the Cu(110):O system calculated using PBE-D2.</td>
<td>102</td>
</tr>
<tr>
<td>4.11</td>
<td>Electronic density plots of the AS-c(6×2) and AS-(6×2)₁ structures calculated using PBE.</td>
<td>104</td>
</tr>
<tr>
<td>4.12</td>
<td>Iso-surfaces of the electronic density of the AS-c(6×2) and AS-(6×2)₁ structures calculated using PBE.</td>
<td>104</td>
</tr>
<tr>
<td>4.13</td>
<td>Gibbs free energy plots for the Cu(110):O system calculated using PBE-rescaled.</td>
<td>105</td>
</tr>
<tr>
<td>4.14</td>
<td>Plot of energy differences between the AS-c(6×2) and AS-(6×2)₁ structures as a function of the Cu bulk lattice constant.</td>
<td>107</td>
</tr>
<tr>
<td>4.15</td>
<td>Gibbs free energy plots for the Cu(110):O system calculated using vDW-DF1.</td>
<td>110</td>
</tr>
<tr>
<td>4.16</td>
<td>Gibbs free energy plots for the Cu(110):O system calculated using vDW-DF1 and vDW-DF2, both employing smaller lattice constants than they predict.</td>
<td>112</td>
</tr>
<tr>
<td>4.17</td>
<td>Electronic density plots of the AS-(6×2) and AS-(6×2)₁ structures calculated using PBE-vDW-DF1.</td>
<td>113</td>
</tr>
<tr>
<td>4.18</td>
<td>Iso-surfaces of the electronic density of the AS-c(6×2) and AS-(6×2)₁ structures calculated using PBE and revPBE-vDW-DF1.</td>
<td>114</td>
</tr>
<tr>
<td>4.19</td>
<td>Minimum energy profiles for a super Cu atom diffusing along a p(2×1) row.</td>
<td>115</td>
</tr>
<tr>
<td>5.1</td>
<td>Ball models of the p(2×1) and c(6×2) phases.</td>
<td>119</td>
</tr>
<tr>
<td>5.2</td>
<td>Cu- and O-terminated tip models.</td>
<td>120</td>
</tr>
<tr>
<td>5.3</td>
<td>Schematic of the cell model used in our theoretical calculations.</td>
<td>121</td>
</tr>
<tr>
<td>5.4</td>
<td>Experimental NC-AFM images of the c(6×2) surface.</td>
<td>123</td>
</tr>
<tr>
<td>5.5</td>
<td>Experimental NC-AFM topography images of the p(2×1)/c(6×2) interface and island.</td>
<td>124</td>
</tr>
<tr>
<td>5.6</td>
<td>Experimental NC-AFM topography images of the p(2×1)/c(6×2) interface, highlighting a reconstruction initiation region.</td>
<td>125</td>
</tr>
<tr>
<td>5.7</td>
<td>Experimental NC-AFM topography images of a type I → type II contrast transition and Cu deposition.</td>
<td>126</td>
</tr>
<tr>
<td>5.8</td>
<td>Experimental NC-AFM topography scan line and corresponding image of the Cu(110)-c(6×2) surface taken with the Si cantilever.</td>
<td>127</td>
</tr>
<tr>
<td>5.9</td>
<td>Experimental and theoretical topography scan lines of the Cu(110)-c(6×2) surface.</td>
<td>128</td>
</tr>
<tr>
<td>5.10</td>
<td>Experimental and theoretical spectroscopy curves of the Cu(110)-c(6×2) surface.</td>
<td>129</td>
</tr>
<tr>
<td>5.11</td>
<td>Experimental and theoretical two-dimensional NC-AFM constant frequency shift map of the topography of the Cu(110)-c(6×2) surface.</td>
<td>132</td>
</tr>
<tr>
<td>5.12</td>
<td>Experimental NC-AFM topography image, scan line and theoretical two-dimensional constant frequency shift map of the topography of the Cu(110)-p(2×1) surface.</td>
<td>135</td>
</tr>
<tr>
<td>6.1</td>
<td>NC-AFM topography scans of the p(2×1) surface being manipulated.</td>
<td>141</td>
</tr>
<tr>
<td>6.2</td>
<td>Model illustrating Cu deposition on the p(2×1) surface and corresponding NC-AFM topography images.</td>
<td>142</td>
</tr>
<tr>
<td>6.3</td>
<td>Geometry models of the possible ad-atom reconstructions of the p(2×1) surface.</td>
<td>143</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Gibbs free energy difference plots for the possible ad-atom reconstructions of the $p(2 \times 1)$ surface calculated using PBE.</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>Gibbs free energy difference plots for the possible ad-atom reconstructions of the $p(2 \times 1)$ surface calculated using PBE-D2 and PBE-vdW-DF1.</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>Schematic depicting how the lateral tip position is represented.</td>
<td></td>
</tr>
<tr>
<td>6.7</td>
<td>Model and corresponding minimum energy profile for the transition of a super Cu atom across an added row.</td>
<td></td>
</tr>
<tr>
<td>6.8</td>
<td>The minimum energy profile for the transition of the super Cu atom to evolve across the alternating Cu-O rows for different O-terminated tip-surface separations.</td>
<td></td>
</tr>
<tr>
<td>6.9</td>
<td>The minimum energy profile for the transition of the super Cu atom to evolve across the alternating Cu-O rows for different Cu-terminated tip-surface separations.</td>
<td></td>
</tr>
<tr>
<td>6.10</td>
<td>Illustrations of the transition of the super Cu atom across the Cu-O row by tip-induced barrier reduction and tip assistance.</td>
<td></td>
</tr>
<tr>
<td>6.11</td>
<td>Model of the relaxed geometry of the Cu-terminated tip in close proximity to the surface.</td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>Si H3 and dimer tip models.</td>
<td></td>
</tr>
<tr>
<td>7.2</td>
<td>Theoretical approach and retraction $F$-$z$ curves for different H3 tip-surface separations.</td>
<td></td>
</tr>
<tr>
<td>7.3</td>
<td>Theoretical dissipation signals, ball model of surface and dissipation scan lines using the H3 tip.</td>
<td></td>
</tr>
<tr>
<td>7.4</td>
<td>Theoretical $F$-$z$ curves for seven lateral positions of the H3 tip.</td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>Theoretical two-dimensional constant frequency shift map of the topography and dissipation using the H3 tip.</td>
<td></td>
</tr>
<tr>
<td>7.6</td>
<td>Theoretical approach and retraction $F$-$z$ curves for different dimer tip-surface separations.</td>
<td></td>
</tr>
<tr>
<td>7.7</td>
<td>Theoretically calculated dissipation signals using the H3 and dimer tips.</td>
<td></td>
</tr>
<tr>
<td>7.8</td>
<td>Theoretical approach and retraction $F$-$z$ curves for different dimer tip-surface separations, which result in tip deformation.</td>
<td></td>
</tr>
<tr>
<td>7.9</td>
<td>Vertical displacements of the tip apex and surface dimer from their ideal positions as a function of H3 tip-surface separation.</td>
<td></td>
</tr>
<tr>
<td>7.10</td>
<td>Vertical displacements of the tip apex and surface dimer from their ideal positions as a function of dimer tip-surface separation.</td>
<td></td>
</tr>
<tr>
<td>7.11</td>
<td>NC-AFM topography and corresponding dissipation images and scan lines of the Si(001) surface taken at 5 K.</td>
<td></td>
</tr>
<tr>
<td>7.12</td>
<td>Non-sequential NC-AFM topography and dissipation images of different regions of the Si(001) surface.</td>
<td></td>
</tr>
<tr>
<td>7.13</td>
<td>Topography image, simultaneously measured dissipation image and corresponding scan lines taken at $\Delta f = -8$ Hz.</td>
<td></td>
</tr>
<tr>
<td>7.14</td>
<td>Topography image, simultaneously measured dissipation image and corresponding scan lines taken at $\Delta f = -18$ Hz.</td>
<td></td>
</tr>
<tr>
<td>7.15</td>
<td>Topography image, simultaneously measured dissipation image and corresponding scan lines taken at $\Delta f = -25$ Hz.</td>
<td></td>
</tr>
</tbody>
</table>
7.16 Topography image, simultaneously measured dissipation image and corresponding scan lines taken at $\Delta f = -28$ Hz.  

8.1 Experimental NC-AFM $\Delta f$ vs. $z$ curves for toggling dimers on the Si(001) surface.  

8.2 A sequence of NC-AFM topography images of the Si(001) surface, depicting dimers that are toggled sequentially along a row.  

8.3 Experimental and theoretical NC-AFM $F$-$z$ and corresponding experimental dissipation curves for toggling dimers on the Si(001) surface.  

8.4 Theoretical $F$-$z$ curves for flipping a dimer on the Si(001) surface using the dimer tip.  

8.5 Relaxed geometries of the Si(001) surface just before manipulation (the $c(4 \times 2)$ surface) and after manipulation (the three-in-a-row configuration) using the H3/Si(111) tip.  

8.6 Minimum energy profiles for the system to evolve from the pristine $c(4 \times 2)$ surface to a surface containing a phason pair via the three-in-a-row structure and via simultaneous flipping of two dimers.  

8.7 Minimum energy profiles for the system to evolve from the pristine $c(4 \times 2)$ surface to a surface containing a phason pair.  

8.8 Evolution of the kinetic energy of different atoms of the surface and the tip apex during an NVE ensemble molecular dynamics simulation.  

8.9 Low temperature NC-AFM images of the manipulation of phason pairs due to scanning.  

8.10 The minimum energy profile for the system to be manipulated from the $c(4 \times 2)$ surface with a two dimer vacancy to a surface containing a phason pair via the three-in-a-row structure.  

8.11 Schematics of the systems used in simulating a strain along the dimer row.  

8.12 The minimum energy profiles for the defective system to be manipulated from the $c(4 \times 2)$ surface to a surface containing a phason pair via the three-in-a-row structure.
List of Tables

4.1 Calculated Cu bulk lattice constant and cohesive energy as determined by different theoretical methods. .................................................. 95
4.2 Comparison of inter-layer separations for the Cu(110) surface slab. ...................... 96
4.3 Calculated Gibbs free energy differences (meV) of surfaces relative to the AS-c(6 × 2) phase as determined by different theoretical methods. ......................... 100
4.4 Calculated Gibbs free energy difference (meV) of surfaces relative to the AS-c(6 × 2) phase as determined by PBE and PBE-D2 methods. ......................... 103
4.5 Calculated Gibbs free energy difference (meV) relative to the AS-(3 × 1) phase as determined by different theoretical methods. ......................... 111
Acknowledgements

First and foremost I would like to thank my supervisor Prof. Lev N. Kantorovich as none of the research contained within this thesis would have been possible without his vision and guidance. He has been a continuous source of great advice, encouragement and inspiration. For the individual projects I would like to thank all of my collaborators in: Japan - Prof. Yasuhiro Sugawara and his team at Osaka University; Slovakia - Prof. Ivan Štich and his team at the Slovak Academy of Sciences; Nottingham, U.K. - Prof. Philip Moriarty and his team at the University of Nottingham. In particular I would like to thank Dr. Yoshitaka Naitoh (Osaka), Dr. Robert Turanský (Slovakia), Samuel Jarvis and Dr. Adam Sweetman (Nottingham). All of these people have contributed significantly to the success of my research.

In addition to the excellent teams I’ve worked with, I would also like to thank my second supervisor Dr. Carla Molteni for her constructive comments and suggestions. For warmly welcoming me and aiding my development in presenting at conferences, I thank the non-contact atomic force microscopy (NC-AFM) community.

For financial assistance, I thank the UK Engineering and Physical Sciences Research Council (EPSRC) funding body for sponsoring me through the studentship grant, providing me with the opportunity to work on these exciting projects, and the High Performance Computing (HPC) Materials Chemistry Consortium (MCC) for additional funding and access to the U.K. HPCx and HECToR HPC facilities. For assistance with King’s College London’s HPC resources, I am especially grateful to the cluster manager Dr. Alessio Comisso whose help has truly been an invaluable aid throughout my PhD. I also thank Dr. Colin Bannister for access to the University of Nottingham HPC facility. For general support and access to academic and intellectual resources I thank the entire Physics department at King’s College London.

I would also like to thank Prof. Alison Mainwood, who was the first person to suggest I work towards a PhD after completing my Masters project with her. It is because of her initial push, which eventually led me to start my PhD with Lev, that I have managed to reach this point.
Publications and Workshops

List of Publications

  The author participated in discussions to determine the theoretical approaches used, performed a large portion of the theoretical calculations, and wrote part of the letter. The author also implemented the NEB algorithm into the DFT code used in this letter.

  The author suggested several of the theoretical approaches used, performed some of the theoretical calculations, and wrote part of the article. The author also implemented the NEB algorithm into the DFT code used in this article.

  The author implemented the NEB algorithm into the DFT code used in this article.

  The author performed most of the theoretical calculations presented here and wrote the article.

The author performed half of the theoretical calculations presented here and co-wrote the article.


  The author performed most of the theoretical calculations presented here and co-wrote the article.


  The author performed half of the theoretical calculations presented here and co-wrote the article.


  The author performed most of the theoretical calculations presented here and co-wrote the manuscript.


  The author performed all of the theoretical calculations presented here and co-wrote the manuscript.

- **J. Bamidele, L. Kantorovich**, “Kinetics of Tri-Methyl-TPP”, *In Preparation 2012*

  The author performed all of the theoretical calculations presented here.

### Conferences, Meetings and Courses Attended

#### International Conferences

- Poster presented at the 13th International Non-Contact Atomic Force Microscopy (NC-AFM) Conference, “*ab initio* Modelling of Anomalously Large Dissipation During NC-AFM Imaging of Si(001) Surface”, Kanazawa, Japan, 2010


• Oral presentation at 14th International Non-Contact Atomic Force Microscopy (NC-AFM) Conference, “NC-AFM Tip & Surface Species Identification on Oxidized Cu(110)”, Lindau, Germany, 2011


**Other Oral Presentations**

• King’s College London postgraduate research seminar, “Dimer Flipping on Si(110)”, 2010

• King’s College London postgraduate research seminar, “Reference System for Scanning Probe Tip Fingerprinting”, 2011

• Second prize award for best presentation at the Thomas Young Centre (TYC) student day, “Surface Scanning Probe Tip Fingerprinting”, 2011

• Materials Chemistry Consortium (MCC) meeting, “Surface Scanning Probe Tip Fingerprinting”, 2011

**Other Poster Presentations**

• Poster presented at Physics By the Lake, “*ab initio* Modelling of Anomalously Large Dissipation During NC-AFM Imaging of Si(001) Surface”, Lake District, U.K., 2010

• Poster presented at the International Review of the Thomas Young Centre (TYC), “Modelling Dissipation on Si(001) under NC-AFM”, 2010

• Poster presented at the Thomas Young Centre Thermodynamics and Kinetics of Surfaces and Interfaces from Simulation workshop, “How Important is vdW in Stabilizing Crystal Surfaces? A Cu:O Study”, 2011
• Second prize award for best poster at King’s College London NMS Graduate Day poster session, “Huge and Complex Dissipation Signals from Small and Simple NC-AFM Scans”, 2011

Other Workshops and Scientific Courses

• Fortran 95 course by Numerical Algorithms Group (NAG), Oxford, U.K., 2010

• Physics by the Lake, Lake district, U.K., 2010

• Thomas Young Centre (TYC) Master class by Prof. Richard M. Martin, “Theory and Quantitative Methods for Electrons in Materials: Concepts, Methods and Prospects”, 2011

• Thomas Young Centre (TYC) Master class by Dr. Jerry Tersoff, 2011

• Thomas Young Centre (TYC) Master class by Prof. Lev N. Kantorovitch, “Non-Equilibrium Green’s Functions”, 2011

• Parallel Programming with MPI course by Numerical Algorithms Group (NAG), London, U.K., 2011

Conventions

The following conventions for units are typically used throughout this thesis:

- **Time Scale:** Femtoseconds (fs), where 1 fs = $10^{-15}$ s.
- **Length Scale:** Nanometres (nm) or Ångströms (Å), where 10 Å = 1 nm = $10^{-10}$ m.
- **Energy Scale:** ElectronVolts (eV), where 1 eV = $1.602178 \times 10^{-19}$ J.
- **Force Scale:** NanoNewtons (nN) or electronVolts per Angstrom, where 1 eV/Å = 1.602178 nN.
- **Surface Nomenclatures:** Surfaces cut from the bulk are specified by the Miller Index, i.e. a (001) surface refers to a plane cut from a bulk normal to the z-axis in Cartesian co-ordinates. If the surface reconstructs after being allowed to relax, then each reconstruction is specified by their extension of the primitive unit cell of the surface, i.e. a 3×2 reconstruction is six times the size of the surface unit cell.
- **Natural Units:** $\hbar = m_e = e = 4\pi\epsilon_0 = 1$.
- **Font:** *Italic font* in text indicates noteworthy names and phrases, or to emphasise particular points, otherwise it is used to represent mathematical symbols and equations, while **bold font** is used to represent vectors.
- **Notation:** The acronym NC-AFM is used interchangeably for non-contact atomic force microscope and non-contact atomic force microscopy.
- **Citations:** If there are only one or two authors, both names are given. If there are more than two authors then the first authors name is given followed by *et al.* This convention is used for all citations given in the main body of the text, figures and tables, the names of all the authors are given in the bibliography.
Part I

Introduction
Chapter 1

Introduction

1.1 Historical Overview

“There’s plenty of room at the bottom” is both the name of the famous talk given by Richard P. Feynman from which the quote at the top of the page is taken, arguably the root of nanotechnology, and a perfect description of how there is much to learn and do in the area of direct manipulation of individual atoms. The field of nanotechnology was built up from this foundation, led by the famous engineer who popularised molecular nanotechnology, Eric Drexler. Due to work of Drexler and other scientists, this question posed in Feynman’s talk was in fact answered within Feynman’s lifetime.

“The revolutionary Feynman vision launched the global nanotechnology race.”

K. Eric Drexler
Engineer (1955 - present)

Modern day experiments routinely consist of functionalising molecules such that they will detect and bond to others, assembling by themselves, this is called self-assembly. These molecules can be tailored to form specific two-dimensional structures, such as nano-porous structures to be used as templates on which to confine functionalised guest molecules [1], an application of which would be in data storage. Their structures, along with many other properties, can now also be studied and even imaged at the atomic/molecular level. This is largely because 22 years after that famous talk by Feynman, the first of two important developments in this area of research occurred, which began to turn nanotechnology into the broad field it is today. This was the development of sophisticated Scanning Probe Microscopy (SPM) techniques.
1.2 Scanning Probe Microscopy (SPM)

1.2.1 The Scanning Tunnelling Microscope (STM)

“It always seems impossible until it’s done.”

Nelson R. Mandela

Civil Rights Leader and Politician (1918 - Present)

In 1981, Gerd K. Binnig and Heinrich Rohrer invented the **Scanning Tunnelling Microscope** (STM) [2], which allowed for the visualisation and identification of individual atoms on a surface, through measuring the tunnelling current between a surface and an atomically sharp tip. The first example of this is shown in **Fig. 1.1**, where Binnig imaged the atomic structure of Si (7×7) surface. However, they were not the first to observe individual atoms directly. This trophy belongs to the field ion microscope [3], which was used by Erwin W. Müller in 1955 [4] to image tungsten atoms on a metal tip and by Gert Ehrlich in 1966 [5, 6] to image atoms and perform quantitative calculations.

The importance of the invention of the STM was its capacity to “directly” image surfaces atom by atom, their geometries (or rather to directly probe their density of states, as was consequently understood) and and its widespread applicability, not shared by other techniques of the time [7]. This was an ability not held by any other popular imaging techniques such as X-ray
1.2 Scanning Probe Microscopy (SPM)

Photoelectron Spectroscopy (XPS), (Scanning) Transmission Electron Microscopy ((S)TEM) and Low Energy Electron Diffraction (LEED).

This technique has since repeatedly and reproducibly achieved atomic contrast when imaging atoms [8] and molecules [9] on surfaces. It has also allowed for nano-manipulations to be performed [8, 10], a feat first performed by Donald M. Eigler and Erhard K. Schweizer when they arranged individual Xenon atoms on a Nickel surface and formed the letters “IBM” in reference to their laboratory, as shown in Fig. 1.2.

![The Beginning - A STM image of Xenon atoms arranged on the Nickel (110) surface, the first atomic manipulation experiment performed by D. M. Eigler and E. K. Schweizer and published in 1990 [8].](image)

**Figure 1.2:** “The Beginning” - A STM image of Xenon atoms arranged on the Nickel (110) surface, the first atomic manipulation experiment performed by D. M. Eigler and E. K. Schweizer and published in 1990 [8].

1.2.2 The Atomic Force Microscopy (AFM)

For the first time, physicists were able to image atoms, and not only that, but were also able to manipulate them. However, despite STM being a very powerful tool for studying and characterising the electronic [9], structural [11, 12], vibrational [13] and transport properties [14, 15] on a single-molecule scale, it is based on the phenomenon of quantum tunnelling, and at low T is therefore limited only to conducting surfaces and thin semiconducting or insulating films on conducting substrates. Due to this, the STM fails to meet the need of being able to also atomically manipulate purely ionic or semiconducting surfaces (i.e. all types of surfaces). However, this issue was addressed in 1986 with the invention of the *Atomic Force Microscope* (AFM) by Gerd K. Binnig, Calvin F. Quate and Christoph Gerber [16], which measures (indirectly) the forces between a surface and an atomically sharp tip. In order to obtain atomic contrast, the tip and sample must be sufficiently close enough to detect changes in forces above different atomic positions, so it therefore measures the chemical (short-ranged) forces.
1.2 Scanning Probe Microscopy (SPM)

1.2.2.1 Types of AFM

There are two main modes for AFM operation: Static and Dynamic (D-AFM\(^1\)). Static mode is commonly referred to as Contact (C-AFM\(^2\)), while dynamic mode can be separated further into two groups: Tapping/Intermittent Contact (T-AFM\(^3\)/IC-AFM) and Non-Contact (NC-AFM\(^4\)).

When operated in contact mode the tip scans across the sample surface while maintaining a constant tip deflection by adjusting the separation of the cantilever and sample. These adjustments provide information of the surface topography. This mode is very successful, applicable on many types of surfaces, but is, however, a form of destructive imaging as the tip-surface separation is such that the tip essentially scratches the surface during the scan. As this happens both the surface and tip are modified as a result. Therefore, while the vertical resolution is very good, \(\approx 0.1\) nm, the lateral resolution is much less, \(\approx 1\) nm. For this reason C-AFM usually does not provide true atomic resolution as 1 nm can span across several atoms.

Tapping mode is generally used in ambient conditions, where the samples usually develop a layer of liquid meniscus. In this mode the tip penetrates this layer in order to be sufficiently close enough to detect the short-range forces, intermittently tapping softly, but not damaging, the sample, which could quite possibly be aqueous. Although, it is a far less destructive technique than contact mode and offers the same vertical resolution, the lateral resolution suffers as a result, \(\approx 5\) nm.

Non-contact mode is when this probing tip is oscillated at a high enough frequency such that the maximum restoring force of the oscillations is equal to or greater than the strongest attractive force that can be felt between the tip and sample surface. The result of this is that the tip never comes into contact with the surface. In practice however, often the tip still does accidentally crash into the surface. Nonetheless, because of the site-specific nature of the tip response (due to the criterion that the restoring force must be equal to or greater than the tip-surface interaction), this mode greatly improves the lateral resolution, achieving atomic resolution, \(> 0.1\) nm.

Furthermore, within these three modes, the range of different flavours of modern AFMs is now very diverse \([17]\), each designed to measure specific properties. These include: Phase Detection Microscopy (PDM), Piezoelectric Force Microscopy (PFM), Force Modulation Microscopy (FMM), Current-Sensing AFM (CS-AFM), Kelvin Probe Force Microscopy (KPFM), Electrostatic Force Microscopy (EFM), Magnetic Force Microscopy (MFM), Magnetic Resonance Force Microscopy (MRFM), Scanning Voltage Microscopy (SVM)/Nanopotentiometry, Friction Force Microscopy

\(^{1}\) Also written as DM-AFM (dynamic mode), DFM-AFM (dynamic force mode), DAFM and dAFM.

\(^{2}\) Also written as CM-AFM (contact-mode AFM), c-AFM and cAFM.

\(^{3}\) Also written as TM-AFM (tapping-mode) and t-AFM.

\(^{4}\) Also written as NCAFM, nc-AFM and ncAFM.
1.3 Non-Contact Atomic Force Microscopy

(FFM)/Lateral Force Microscopy (LFM), Contact-Resonance AFM (CR-AFM), and Conductive AFM (CAFM\(^1\)).

### 1.3 Non-Contact Atomic Force Microscopy

Using these tools, we have arrived at a stage where atomic and molecular imaging and manipulation are common practice and the physics behind nano-machine (self-assemblers) is becoming much better understood. In spite of all the benefits of these techniques, there are some experimental elements that can not be controlled, such as the atomic structure of the tip termination. This, along with other factors, means that SPM measurements are not always easy to interpret and what you see may not exactly be what you have! It is for this reason that our group has chosen to focus on one of these techniques, NC-AFM, to study some of these issues in the hope of allowing further progress in this field.

Since its development, the NC-AFM has proven an excellent tool for performing similar tasks to the STM, atomic [19, 20, 21] and chemical resolution [22], and nano-manipulation [19, 23], by measuring forces between the tip and surface [24, 25]. An example of an NC-AFM image is shown in Fig. 1.3, where Pentacene molecules can clearly be resolved\(^2\). In addition, NC-AFM, unlike STM, can be performed on conducting, semiconducting and insulating surfaces, allowing it

---

\(^1\)Caution as this is also written as C-AFM and is not be confused with contact mode.

\(^2\)Fig. 1.3 is a special example in that the image is taken with a CO adsorbed tip which has been found to produce specially high-resolution NC-AFM images.
1.3 Non-Contact Atomic Force Microscopy

to be the tool of choice for probing surface structures and building nano-devices on all types of surfaces.

As with all experimental techniques, there is also the need for an in-depth theoretical understanding, and since its invention, there have been several advances in the development of theoretical methods for analysing and interpreting experimental results [26, 27], as well as tools for theoretical simulations [28, 29], creating the theoretical backbone for NC-AFM. Although much of the theory is generally well understood, there are still issues with interpreting images. This is because instead of directly measuring distances between the tip and surface, it is the change in the frequency of the tip oscillations (or change in amplitude) that is measured, which is directly related to the force between the tip and surface (which in turn is related to the tip-surface separation). Therefore, first and foremost, detailed knowledge of both the tip and sample surface must be known in order to characterise the forces acting within them. This will allow theory to propose the model(s) of the surface which result(s) in NC-AFM measurements that are comparable to experiments.

“If we knew what it was we were doing, it would not be called research, would it?”

ALBERT EINSTEIN
Physicist (1879 - 1955)

1.3.1 Issues with the NC-AFM

1.3.1.1 Surfaces

Starting with the sample, although bulk materials are often difficult to model, surfaces are comparatively more complicated physical structures that need to be understood\(^1\). Sharing their borders with both the bulk material and the outside environment, surface atoms may interact with: bulk atoms; other surface atoms; atoms/molecules in their surrounding environment. This can often lead to several different, but sometimes similar, reconstructions when cut from the bulk. If a well described surface is known then there is still the practical issue of producing and maintaining a “clean” surface, as they are usually quite easily contaminated with other atomic species (impurities). These impurities can sometimes greatly affect results\(^2\). However, modern NC-AFM can be operated in \textit{ultra-high vacuum} (UHV) conditions, which greatly reduces the problem of contamination, but does not eliminate it.

“God made the bulk; surfaces were invented by the devil.”

WOLFGANG E. PAULI
Physicist (1900 - 1958)

\(^1\)The bulk is often simpler to model than surfaces due to their three-dimensional periodicity, except at their boundaries, where surface effects complicate the physics.

\(^2\)Impurities, apart from local-features, may cause long-ranged interactions, and hence long-ranged distortion of the crystal lattice around them.
1.3.1.2 Probe Tip

Assuming contamination is not an issue, the atomic structure of the tip is usually unknown before, during and after experiments, as the tip itself is performing the imaging and therefore cannot be imaged itself. Also, the tip often changes during scanning or manipulation, due to tip crashes, sometimes multiple times. This means we can usually only have just a very general description of the tip, but there are instances when this is just not enough and the exact nature must be known [30]. Although the principle behind atomically probing surfaces with a NC-AFM is beautifully simple (use a sharp tip to measure the force between it and the surface), the implementation is certainly not.

“All art is at once surface and symbol. Those who go beneath the surface do so at their own peril.”

Oscar F. O. W. Wilde
Writer and Poet (1854 - 1900)

1.4 Aim

The main motivation for this study into the *theoretical modelling and characterisation of concerted tip and surface affects in NC-AFM* was primarily to investigate the effects of different tip terminations, surface instabilities and defects on NC-AFM imaging and atomic manipulation of different systems. The secondary objective was to develop a protocol for on-the-fly tip termination identification/enforcement.

1.5 Overview of Study

The first two chapters of this thesis detail all the methodologies used, starting from a discussion of the experimental setup of the NC-AFM in Chapter 2, moving on to the theoretical modelling of it in Chapter 3. Within the modelling, the theoretical knowledge required for the numerical simulations is described, introducing density functional theory (DFT) and explaining how it gives access to the energy of the system and the forces acting within it, via an approximate solution to the many-body Schrödinger equation. For the simulations, a section is dedicated to the plane-wave and localised basis-set methods implemented in the *Vienna Ab initio Simulation Package* (VASP) and *Spanish Initiative for Electronic Simulation with Thousands of Atoms* (SIESTA) codes, respectively, used to perform our calculations. Discussions on the construction of the systems and convergence parameters used in the calculations are presented. In addition, methods that were required to be manually implemented, such as the nudged elastic band (NEB) method in SIESTA, and tools designed specifically to analyse the systems, are thoroughly explained.
Chapter 4 introduces the first set of systems studied, the oxidised Cu(110) insulator surfaces, (henceforth referred to as Cu(110):O) and describes our theoretical understanding of the thermodynamic stabilities of its stable ground-state reconstructions. The research is an in-depth study to not only obtain the most suitable parameters for later NC-AFM modelling of the systems, but also to probe and understand even the subtle properties of the surfaces, such as how changes in the parameters describing the system can critically distort results. This leads on to a discussion on the role of long-ranged dispersion forces in the relative stabilities of the surfaces, an effect not before reported on a solely insulator surface.

Chapter 5 then follows this up by a collaborative experimental and theoretical NC-AFM study of the two ground-state reconstructions of Cu(110):O surface: p(2 × 1) and c(6 × 2), and a discussion on the nature of the tip-surface forces for different tip terminations. The experimental reproducibility is also discussed, which leads on to the formulation of a protocol for identifying and discriminating between three different atomic tip terminations, the only three terminations likely for the systems (Si, Cu and O). This is achieved on the c(6 × 2) surface, by essentially scanning and crashing the tip into the surface, then scanning again. Furthermore, the protocol also allows for the enforcement of one of two tip terminations out of the possible three.

In Chapter 6 the research on the Cu(110):O surfaces is concluded by a study of NC-AFM nano-manipulation performed on these surfaces. The manipulations of these surfaces are performed experimentally and the mechanisms are studied theoretically, both with different tip terminations.

Looking now at tip and surface effects for a different surface, a thorough study of dissipation on the Si(001) - c(4 × 2) surface is presented in Chapter 7. This chapter focuses on how differently structured tips affect mechanically induced NC-AFM dimer flipping (manipulation), up to the contact regime, and how this in turn affects NC-AFM topography and dissipation imaging of that surface. Through this, a complex architecture behind the experimental feedback signals (measurements) as a function of tip type and positioning is uncovered.

Lastly, with the knowledge of how the tip structure affects imaging and manipulations on the Si(001) - c(4 × 2) surface gained in Chapter 7, controlled reversible manipulations of the surface is performed and studied experimentally followed by thorough theoretical analysis in Chapter 8. In this study, the effects of surface dopants, as well as surface defects, on controllable nano-manipulation is investigated and the mechanism behind controllable manipulations is uncovered and detailed.

All the results are then summarised in Chapter 9 and the study as a whole is concluded with a discussed in Chapter 10, interpreting each of the results within the context of how both the tip and surface affect NC-AFM results separately and simultaneously. Some suggestions are
then made on the future application of this research and on possible further research to address outstanding issues.
Chapter 2

The Non-Contact Atomic Force Microscope

As already implied in the introduction, operating the AFM in non-contact mode gets around the issue of the tip regularly crashing into the surface. It is operated under the principle that the effective\(^1\) cantilever stiffness, \(k_{\text{eff}}\), must be equal to or greater than the gradient of the maximum tip-surface force felt, \(F_{ts}\), over the whole range of tip-surface separations, where \(z\) is the tip-surface separation. In other words, \(k_{\text{eff}} \geq \max |\frac{\partial F_{ts}}{\partial z}|\), where \(k_{\text{eff}}\) depends on the cantilever spring constant, \(k\), and \(F_{ts}\) depends on both the macroscopic and microscopic forces. However, rather than simply using a very stiff cantilever, \(k_{\text{eff}}\) must also still be comparable to the spring constant corresponding to only the chemical forces so that measurements of the cantilever deflection are above the experimental noise [26]. This criterion is met by oscillating the cantilever about an equilibrium tip-surface separation, \(z_0\). This is because by using an appropriate intrinsic cantilever stiffness, \(k\), and oscillation amplitude, \(A_0\), the restoring force of the tip can be made greater than the maximum tip-surface force, i.e. \(A_0 k_{\text{eff}} \geq |F_{ts}(z)|\), where \(z = (z_0 - A_0)\). In other words, by oscillating the cantilever, the elastic force of the oscillations can be made to be such that it is always greater than the maximum force of attraction between the cantilever tip and the surface. The result of this is the prevention of the tip jumping into contact with the surface. Since the imaging signal is the cantilever deflection, both the frequency and amplitude of the oscillations can be measured, and also means that either the frequency or amplitude can be modulated. When the frequency is modulated, NC-AFM is referred to as Frequency-Modulated AFM (FM-AFM), and when the amplitude is modulated it is known as Amplitude-Modulated AFM (AM-AFM). Throughout this thesis experimental measurements and theoretical simulations are performing using FM-AFM.

\(^1\)When interacting with the surface \(k_{\text{eff}}\) depends on both the intrinsic stiffness, \(k\), and the contribution due to the tip-surface interaction, \(\frac{\partial F_{ts}}{\partial z}\).
2.1 Experimental Setup of the NC-AFM

In order to put any theoretical methodologies into context, a description of the experimental setup of the NC-AFM must first be given. This can then be followed by a detailed description of how it can be theoretically modelled and how each of the relevant properties can be calculated theoretically, as is done in Chapter 3.

![Figure 2.1: Photograph (a) and schematic (b) of the NC-AFM cantilever-sample holder used to perform the experiments discussed in Chapter 5 and Chapter 6. Both images were taken by and used with the permission of Prof. Y. Sugawara’s group at Osaka University.](image)

In a crude description, the NC-AFM consists of a probe tip on an oscillating cantilever attached to a sensor that measures the oscillation frequency, which changes upon tip interaction with the sample, all suspended on springs in order to negate external vibrations. The actual size of the NC-AFM can vary, but the core equipment, the cantilever holder shown in Fig. 2.1, is only a few centimetres wide. However, that being said, the housing of the holder can be much bigger (≈1 m tall), because the sample may need to be scanned under specific conditions. Such is the case for all the experiments studies in this thesis, which were performed in ultra-high vacuum (UHV) conditions to minimise/control contaminants. Therefore, for all the experiments performed in our research, the holder was placed inside a UHV chamber. An example of this is shown in Fig. 2.2. Furthermore, all of the experiments required the samples to be cooled to very low temperatures in order to “freeze” the surfaces in a particular phase and/or to reduce the mobility of surface atoms, so that transitions (changes) may be studied. The unwanted mobility of surface atoms, or of the tip position relative to the surface, is referred to as thermal noise/drift, however, with development of atom tracking methods [26], drift is less of an issue.
2.1 Experimental Setup of the NC-AFM

In all the experiments presented here, silicon cantilevers were used, which before entering the main chamber, indicated in Fig. 2.2(a), were cleaned *in-situ* in the preparation chamber, indicated in Fig. 2.2(b). This was usually done either by electron-beam heating or Ar\(^+\) bombardment, to remove any contaminants and oxide layers that may be present due to exposure to the atmosphere, as shown schematically in Fig. 2.3(a). The result of which is the production of a clean tip, Fig. 2.3(b), which can then be coated if necessary, as is the case in Fig. 2.3(c).

In FM-AFM, the cantilever tip is oscillated at the resonance frequency, \(f_0\), for the given tip-surface interaction field, by a piezoelectric actuator, maintaining a user defined constant amplitude
2.1 Experimental Setup of the NC-AFM

A schematic of the experimental setup of the NC-AFM is shown in Fig. 2.4. The amplitude and oscillation frequency are measured by an optical sensor, a laser beam bouncing off the back of the cantilever onto a photo-diode detector. This signal is then fed into the Phase Locked Loop (PLL), then into two loops: the Automatic Gain Control (AGC) loop, and the loop containing the Automatic Distance Control (ADC) block. The PLL outputs the frequency shift $\Delta f$, caused by the tip-surface interactions changing the frequency from its initial value, $f_0$, i.e. $\Delta f = f - f_0$, where $f$ is the current oscillation frequency. The AGC loop maintains the amplitude of the cantilever oscillations, i.e. it keeps $A_0$ constant at some preset value. It does this by adjusting the strength of the excitation signal, $A_{exc}$, that is delivered to the piezoelectric actuator. In order to decouple the equations for the frequency and oscillation damping, the cantilever is oscillated at resonance, a task achieved by the Phase Shifter, which shifts the phase of the signal by -90 °. Hence, $A_{exc}$ can also give information about the surface as the changes in the excitation signal are directly related to the oscillation damping caused by the tip-surface interactions. The ADC loop maintains a user defined $\Delta f$ by way of the ADC block, which adjusts the position of the sample, moving it closer to or further away from the tip, so as to maintain a constant average tip-surface separation, $z_0$. In AM-AFM experiments $z_0$ is adjusted until the amplitude equals a preset value, $A_0$. This change in tip-surface separation is therefore used to image the topography of the surface. The two processes discussed here, restoring the amplitude and frequency shift, are, however, not instantaneous and
2.2 Operation of the NC-AFM

some relaxation time, typically within a few oscillations, is required to reach their preset values after any tip-surface separation change.

As mentioned, two important properties can be measured: changes in height and excitation signal. The map of the change in height with the respect to the lateral tip-surface positions gives the topography of the surface and that of the excitation signal gives the damping. Since damping in the system is due to energy being dissipated throughout the tip-surface system, further references to damping will be given as dissipation. Alternatively, the NC-AFM can be operated in constant height mode, where $z_0$ is fixed and the change in the frequency shift is measured. In the images produced, there is usually an associated scale to describe how the topography of the surface is represented.

2.2 Operation of the NC-AFM

As implied, the topography map produced consists of the changes in $z_0$ values required to maintain a constant $\Delta f$, which may change for each lateral position. In fact each pixel (lateral data point) consists of an average of several hundred tip oscillations. Initially the cantilever is positioned far away from the surface (large $z_0$), such that it is free to oscillate at the user defined $A_0$. Whilst oscillating, $z_0$ is then reduced, then as the tip enters the attractive region of the interaction, the oscillation frequency reduces. $z_0$ is then adjusted until the preset $\Delta f_0$ is obtained. After the ADC adjusts $z_0$ each time to maintain constant $\Delta f$, the tip is moved in the lateral direction; the tip performs this operations at every specified point in the $x$ direction. The tip then moves in the $y$ direction, and scans back across the $x$ directions. This procedure is repeated for all specified lateral tip positions. Hence, in terms of this example, when the tip rasters, there is a fast direction, the $x$ direction (typically at 10-50 nm s$^{-1}$), and a slow direction, the $y$ direction. The raster speed depends on the experimentalist’s preference to some degree, but most importantly depends on the tip type and stiffness, where tip type refers to whether a simple cantilever or a tuning fork is used, while the stiffness is material dependent. By using an appropriate $k$ and $A_0$, the tip is able to maintain oscillations in the attractive regime of the tip-surface interaction and thus never (or rarely, as is the case practically) crashes into the surface. Hence, during imaging, the instantaneous tip-surface separation, $z$, which is given by $z = z_0 - A_0$, is typically of the order of inter-atomic bond lengths, 2 - 5 Å.
2.3 Macroscopic Forces

2.3.1 Macroscopic van der Waals Interactions

Due to the close proximity of the tip and sample (within several Ångströms of each other), there will be a large van der Waals (vdW) contribution to tip force. However, this contribution is also dependent on the large sizes of the tip and sample (several microns), and is hence relatively insensitive to the comparatively small changes in the lateral tip position (a few Å, many orders of magnitude smaller). It can therefore be modelled simply as a background contribution to force, independent of lateral position. The vdW interaction describes the pairwise inter-atomic electrostatic attraction between particles in the system; it could be due to:

- polar atoms (dipole-dipole), where permanent dipoles interact and the interaction strength is dependent on the orientation of the dipoles, which is known as the Keesom interaction;
- polar atoms (dipole-induced dipole), where a permanent dipole induces a dipole moment in a neighbour and the two dipoles interact, which is known as the Debye interaction;
- neutral atoms (instantaneous dipole-instantaneous dipole), where spontaneously and instantaneously fluctuating dipoles interact with each other, which is known as the London dispersion interaction (sometimes referred to as the London interaction or dispersion interaction) \[33\]. This is usually the largest contributor to the vdW interaction.

Since the macroscopic portion of the tip and surface contain hundreds of billions of atoms, it is computationally prohibitive to treat all of these interactions quantum mechanically (for reasons outlined in Section 3.8.2.3). Therefore an approximation is made for the background contribution in all calculations, where the macroscopic portion of the tip is modelled by a sphere, and important microscopic contributions, such as a cluster terminating the tip (a few hundred atoms or less) are considered quantum mechanically.

Other macroscopic forces may contribute but tend to be significant only when the tip and sample have different charges. Such contributions are: electrostatic forces, magnetic forces, image and capacitance forces. However, for the systems discussed in this thesis, the tip and sample do not have different charges as, although there is an electrical connection between the tip and sample, the system is electrically grounded at 0 V bias. Thus these contributions are not present and will not be discussed further, but the author refers the reader to the book by Adam Foster and Werner A. Hofer \[27\] for further reading.
2.4 Microscopic Forces

2.4.1 Microscopic van der Waals Interactions

Although vdW interactions are the weakest atomic interactions, they can sometimes play a critical role, e.g. when studying the interaction of molecules with a surface. Therefore, there is a need for them to be accounted for quantum mechanically. This can be achieved by using one of several methods, two of which are described in Section 3.8.2.3. This is only done in cases where the vdW interactions are deemed crucial and the number of atoms involved is small enough to allow the calculations to be completed in a reasonable amount of time.

2.4.2 Chemical Forces

The chemical forces are generally what give rise to the atomic resolution in NC-AFM measurements, and they are governed entirely by electronic interactions of the tip-surface system. Therefore the nature of these forces vary depending on the material of both the tip and the surface; the inter-atomic forces within both the tip and surface and between them may be due to:

- ionic bonding, where electrons are taken or removed, resulting in differently charged ions which have long-ranged electrostatic interactions;
- metallic bonding, where electrons are delocalised and move throughout the whole material, effectively being shared by all the metallic atoms;
- covalent bonding, where electrons are shared between two or a small group of atoms.

Generally ionic and covalent bonds are very strong bonds. Where molecules are imaged (adsorbed onto the sample surface), additional contributions to the tip forces can come from the interactions between the tip and molecule, which may be due to:

- vdW interactions between molecules, molecules and the surface, and between molecules and the tip;
- hydrogen bonding, where electronegative atomic species (such as Oxygen) covalently bonded to Hydrogen atoms causing an elongation of the electron distribution on the Hydrogen atoms, which polarises the Hydrogen atoms, allowing them to interact with negative dipoles;
- covalent interaction between the molecule and tip.

It is worth noting here that there is often confusion as to whether the dispersion force is a subset of the vdW force or if they are the same thing. However for clarification, all three interactions contribute to the vdW force, but generally the largest contributor is the London force.
2.5 Summary

In order to model the macroscopic forces, approximations need to be made, and can be done without changing the physical interpretation of the results. This allows for a simpler treatment of them, mostly due to the fact that the macroscopic part of the system can quite accurately be considered as not quantum mechanical, but classical. However, because microscopic forces are due to the electronic interactions, which are entirely quantum mechanical, the methodology behind the modelling of them is certainly not so simple.

The systems discussed in this thesis predominantly bond only covalently, but some involve a metallic substrate. In any case, the methods described in the next chapter describe how to model all three types of chemical interactions, as well as the macroscopic and microscopic vdW interactions.
Part II

Theory
Chapter 3

Theoretical Methodology

This chapter outlines all the important theoretical methods used for analysis, including the modeling of the NC-AFM, the calculation of forces, energies and transitions paths. All of the theory in this section is derived from first principles, not based on any empirical (experimental) results, as such, the theory is referred to as ab initio, i.e. from first principles.

3.1 Classical Mechanical Description of the System

“Nothing exists except atoms and empty space; everything else is opinion.”

DEMOCRITUS
Greek Philosopher (460-370 B.C.)

In order to properly describe how to model an NC-AFM experiment at the atomic level, from the general motion of the cantilever (macroscopic forces) to the tip-surface interactions (microscopic forces), we must first have a description of atomic interactions. To this end, it is best to start from the very general and basic law governing the motion of any mechanical system (quantum or otherwise), the Hamilton principle, otherwise known as the principle of least action.

3.1.1 The Lagrangian

To completely define the mechanical state of any system, only the co-ordinates ($x$) and velocities ($\dot{x}$) at any given time ($t$) of all objects constituting the system need be known [34]. These three quantities alone describe entirely where the system is and the nature of its motion. In fact, if $x$ is known for any $t$, then $\dot{x}$ is simply derived from this, and the trajectory for $x(t)$ is such that the condition of least action is met. The function that contains all of these quantities, $L(x, \dot{x}, t)$, is called the Lagrangian of the system.

---

1The exception to this is in the description of vdW where material dependent terms appear in the macroscopic approximation of the tip and in the Grimme method discussed in Section 3.8.2.3.1.
3.1 Classical Mechanical Description of the System

3.1.2 Action

The trajectory \( x(t) \) corresponding to the least action will therefore be such that the action, as defined when the system goes from \( x_1 \) at time \( t_1 \) to \( x_2 \) at time \( t_2 \), given by

\[
S = \int_{t_1}^{t_2} L(x, \dot{x}, t) \, dt,
\]

is minimised. The necessary mathematical condition for a function to have a minimum is that the variation of the function is zero, namely that

\[
\delta S = \delta \int_{t_1}^{t_2} L(x, \dot{x}, t) \, dt = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} \right) \, dt = 0,
\]

where \( \delta S \) is a small variation of \( S \). Setting \( x = x(t) \) as the function for which \( S \) has its minimum, implies that any variation of that function, i.e. \( x = x(t) + \delta x(t) \), changes \( S \) to the first order. This also requires that for \( t = t_1 \) and \( t = t_2 \),

\[
\delta x(t_1) = \delta x(t_2) = 0,
\]

i.e. that the ends of the tried trajectories are fixed.

3.1.3 Lagrange’s Equations

Hence, for a system with \( n \) degrees of freedom, by substituting \( \delta \dot{x} = d\delta x/dt \), integrating by part, and using Eq. 3.3, Eq. 3.2 can be written as \( n \) sets of differential equations, known as Lagrange’s equations\(^1\), given by

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) - \frac{\partial L}{\partial x_i} = 0 \quad (i = 1, 2, ..., n).
\]

If the Lagrangian for a system is known, then Lagrange’s equations (Eq. 3.4) give the relations between co-ordinates, velocities and accelerations, known as the system’s equations of motion.

To generate the form of the \( L \), a frame of reference must be chosen for which observations of the mechanical system can be made. An example of such a frame is the inertial frame, in which space and time are homogeneous, the former also being isotropic. Incidentally, choosing this frame as the reference frame leads to several convenient simplifications of Eq. 3.4. These are that in the absence of external fields, the homogeneity of space and time implies that \( L \) has explicit dependences on neither the radial vector, nor time, and due to the isotropy of space in the inertial frame, \( L \) also shows no explicit dependency on the direction of the velocity. The first

\(^1\)These are also known as Euler’s equations or Euler-Lagrange equations.
of these simplifications implies that $L$ depends only on velocity, which when applied to Eq. 3.4 leads to the result that the velocity is constant. The second simplification implies that $L$ depends only on the magnitude of the velocity, i.e. $L = L(|\dot{x}|^2)$. Together these are known as the law of inertia.

These results are proven simplest in the inertial frame but are also true for any other frame of reference. Also, after expanding the Lagrangian in terms of the velocities of the particles relative to each other for a system consisting of $N$ non-interacting particles, $L$ is found to be additive and of the form

$$L = \sum_{i=1}^{n} \frac{1}{2} m_i \dot{x}_i^2,$$

(3.5)

where $n$ is the total number of degrees of freedom$^1$ as in Eq. 3.4, and $m_i$ is known as the mass of the $i^{th}$ particle.

### 3.1.4 Newton’s Equations

Now considering a system of $N$ interacting particles in a closed system (where particles interact with other particles within that system and no others), Eq. 3.5 becomes

$$L = \sum_{i=1}^{n} \frac{1}{2} m_i \dot{x}_i^2 - U(x_1, x_2, ..., x_n),$$

(3.6)

The Lagrangian in Eq. 3.6 now holds the two principle physical quantities that determine the motion of any system. The first term, the non-interacting term, is called the kinetic energy of the system, $T$, while the second term, describing the interactions within the system, is known as the potential energy of the system, $U$. To work out the equations of motions of the system, Eq. 3.6 is substituted into Eq. 3.4, which gives

$$\frac{\partial L}{\partial \dot{x}_i} = m_i \ddot{x}_i = -\frac{\partial U}{\partial x_i}.$$

(3.7)

Written in this form, these equations of motion are referred to as Newton’s equations, where the term on the right is called the force, $F_i$, and depends on the derivative of the potential energy or the second time derivative of the distance. This also corresponds to Newton’s second law of motion.

### 3.2 Modelling NC-AFM Tip Dynamics

There is enough knowledge now to be able to model the macroscopic cantilever dynamics, relating the force of the cantilever oscillations to the displacement of the cantilever tip. A full

$^1$Each particle has $3N$ degrees of freedom as both $x$ and $\dot{x}$ are 3-dimensional.
3.2 Modelling NC-AFM Tip Dynamics

A computational model exists for the NC-AFM called the virtual NC-AFM or vNC-AFM. However, since this research requires only knowledge of the atomic processes occurring in the junction of the tip-surface interactions, a simpler model, detailed in this chapter, is instead used. For more information on the vNC-AFM, which can model the real-time dynamical response of the system, the author refers the reader to the PhD thesis of Jérôme Polesel [35], the paper by J. Polesel-Maris and S. Gauthier [36], or the PhD thesis of Thomas Trevethan [37].

![Diagram of cantilever and sample system](image.png)

**Figure 3.1:** Schematic illustrating the cantilever and sample surface system. Image is an enlargement of the tip-surface system show in Fig. 2.4.

In general, to model the dynamics of an oscillating cantilever, illustrated in Fig. 3.1, a solution to the three dimensional harmonic oscillator must be obtained. However, as explained in the previous chapter, the cantilever tip raster across the sample surface, and each lateral position imaged consists of hundreds of vertical oscillations. Therefore, assuming that the oscillation frequency is very high in comparison to the raster speed (which experimentally is the case), the problem of a cantilever tip interacting with the sample surface is greatly reduced; the problem becomes a driven, one-dimensional (considering the tip-surface separation) harmonic oscillator acting within an external field, where only vertical oscillations are sufficient to consider. Thus, the cantilever equation of motion is a second order differential equation written as

\[ m\dddot{z} + m\omega_0^2 \ddot{z} + kz = F_{ts}(z) + F_d\cos(\omega t), \]  

where \( z \), as always, is the instantaneous tip-surface separation and \( m \) the effective mass of the cantilever, given as \( m = k/w_0^2 \). \( \omega_0 \) and \( Q \) are the natural frequency and Q-factor, respectively, which are both material dependent. \( \omega_0 \) is given as \( \omega_0 = 2\pi f_0 \), where \( f_0 \) is the natural frequency (resonant frequency of the free cantilever). Finally, \( F_d \) is the amplitude of the driving force of the cantilever, which forces the cantilever to oscillate at the frequency \( \omega \). This can be rewritten more clearly as

\[ m\dddot{z} + m\omega_0^2 \ddot{z} + [kz - F_{ts}(z)] = F_d\cos(\omega t), \]  

(3.9)
3.2 Modelling NC-AFM Tip Dynamics

where the tip-surface force, $F_{ts}$, can now clearly be seen to also determine the motion of the cantilever. The solution to this problem contains two parts: the non-linear complimentary solution, $z_c$, which is transient since the cantilever is subject to a friction force (the second term, containing $Q$), and the particular (steady-state) solution, $z_p$, due to the driven oscillations. Assuming enough relaxation time, the dynamics are governed solely by the steady-state solution.

Taking into account both the sinusoidal driving force and external potential within which the tip is acting (the tip-surface interactions), the general form of the steady-state solution can be given as

$$z_p(t) = \alpha \sin(\omega t) + \beta \cos(\omega t)$$

$$= A_p \left[ \sin(\omega t) \sin(\vartheta) + \cos(\omega t) \cos(\vartheta) \right]$$

$$= A_p \cos(\omega t + \vartheta),$$

where $\alpha$ and $\beta$ are constants of the trial solution, and $A_p$ is the amplitude, given by $A_p = \sqrt{\alpha^2 + \beta^2}$, while $\vartheta$ is the oscillation phase shift, given by $\sin(\vartheta) = \alpha/A_p$ or $\cos(\vartheta) = \beta/A_p$. In fact, there are more terms (using a Fourier expansion). However, taking only the leading term is most cases sufficient as the tip oscillation is predominantly sinusoidal. This is reflected in Eq. 3.10. Typically during NC-AFM experiments, a phase shift of $-\pi/2$ ($-90^\circ$) is introduced to the driving/excitation signal to achieve oscillations at the resonant frequency of the tip, and applying this shift to $z_p$ gives $z_p = A_p \sin(\omega t)$. By substituting this into Eq. 3.9, the following is obtained

$$\frac{F_{ts}(z)}{m} + \frac{F_d}{m} \cos(\omega t) = -\omega^2 A_p \sin(\omega t) + \frac{\omega_0}{Q} \omega A_p \cos(\omega t) + \omega_0^2 A_p \cos(\omega t)$$

$$= A_p \left[ (\omega_0^2 - \omega^2) \sin(\omega t) + \frac{\omega_0}{Q} \omega \cos(\omega t) \right].$$

$A_p$ can now be seen to be the effective amplitude of the oscillations which has been changed from $F_d$ due to the tip-surface forces, $F_{ts}$. Next, using $z_0$, with its usual meaning as the average tip-surface separation (c.f. Fig. 3.1), the instantaneous tip height, $z(t)$, can be described as $z(t) = z_0 + z_p(t)$, where $z_p(t)$ is defined as in Eq. 3.10. Now, using this new expression for $z(t)$, multiplying Eq. 3.11 by $\sin(\omega t)$ and performing a time-integral on both sides gives

$$\omega_0^2 - \omega^2 = \frac{\omega}{m \pi A_p} \int_0^{2\pi/\omega} F_{ts}(z_0 + z_p(t)) \sin(\omega t) dt,$$  

(3.12)

which, by changing the variable of integration, is simplified to give the oscillation frequency $\omega$ as

$$\omega = \sqrt{\omega_0^2 - \frac{1}{m \pi A_p} \int_0^{2\pi} F_{ts}(z_0 + A_p \sin(\theta)) \sin(\theta) d\theta}.$$  

(3.13)
Similarly, by instead multiplying Eq. 3.11 by $\cos(\omega t)$ before integrating, an expression for the effective amplitude can also be determined. This analysis and resulting equations are valid when anharmonic effects can be neglected, i.e. when $F_{ts}$ is much smaller than the restoring force of the cantilever.

As a final remark on modelling the tip dynamics, by solving the same mechanical problem, but for the tip far away from the surface, i.e. where $F_{ts} = 0$ (for free oscillations), the resonant frequency can be found by calculating the maximum amplitude, when $\frac{dA_p}{d\omega} = 0$, and in fact, this turns out to be given by $\omega_d$, where, using $l = 1/(2Q)$, $\omega_d = \omega_0 \sqrt{1 - l^2}$.

### 3.3 Tip-Surface Interactions

It is clear from Eq. 3.13 that to accurately model how measurements are taken with the NC-AFM, the force due to the tip-surface interaction, must be known, i.e. $F_{ts}$ must also be accurately modelled.

![Figure 3.2: Schematic illustrating the partitioning of the tip into macroscopic and microscopic regions.](image)

To model the interaction between an NC-AFM tip and systems discussed in this thesis, the macroscopic and microscopic portions of the cantilever tip are separated. The former is treated classically while the latter is treated quantum mechanically. This can be done because the atomic resolution is obtained due to chemical forces, which are very short-ranged and only act within $\approx 0.5$ nm.

Thus, only the cluster of atoms terminating the tip, the nano-asperity, need to be calculated very accurately. The separation of these regions and the nano-asperity are illustrated in Fig. 3.2. In this way, microscopic forces (chemical, which include microscopic vdW), responsible for atomic resolution, can be isolated from the larger tip-surface system (the macroscopic portion). These can then be evaluated to a much higher precision, as is necessary.

As discussed in the previous chapter, the macroscopic regions interact via vdW interactions and the microscopic via chemical interactions, contributing energies $U_{vdW}$ and $U_{\mu}$, respectively.
3.4 Modelling Macroscopic Forces

The macroscopic interactions depend only on the instantaneous tip-surface separation, \( z \), and are sensitive only to changes in the separation that are many orders of magnitude larger than inter-atomic distances (see Section 3.23). Hence, \( U_{vdW} \) is generally insensitive to the lateral motion of the tip, whilst microscopic forces are of course very sensitive to this. Since the interactions have been partitioned, there should also be a term that couples the interaction regions together, however, for a large enough microscopic tip model, the coupling of the two interaction regions can safely be neglected. That being said, the relative positions of the macroscopic and microscopic tips need to be considered, and this is done with an offset parameter, illustrated in the right image of Fig. 3.2. This offset, has physical meaning. If it is negative then the microscopic tip sits inside the macroscopic model, and in this case there will be double counting of the vdW interaction due to the overlapping part; however, for small values, the effect of this must be negligible. Conversely, if the offset is positive then the two portions of the tip have a gap, which would imply a larger microscopic tip is present. In this case, a part of the macroscopic interaction due to the atoms in the gap is missing, however, for small values\(^1\), due to these atoms being further away from the surface than the atoms of the microscopic tip, this contribution can also be neglected. Hence, the tip-surface force can be calculated simply as the sum of the macroscopic and microscopic forces, i.e.

\[
F_{ts}(x, y, z) = F_{vdW}(z) + F_{\mu}(x, y, z).
\]  

(3.14)

3.4 Modelling Macroscopic Forces

As mentioned, there is only one macroscopic interaction relevant for this research, vdW, and it is treated in a similar manner to the treatment proposed by Jacob Israelachvili [38], starting from a simple model, then building up the physical attributes of a more realistic system.

Initially a point-charge above a single-layer surface is considered, depicted in Fig. 3.3(a). The charge will induce dipoles in surface atoms, whose interaction gives the London dispersion energy\(^2\),

\[
U_{vdW}(r) = -\frac{C_6}{r^6},
\]  

(3.15)

where \( C_6 \) is a system dependent constant\(^3\) and \( r \) is the inter-atomic separation. Since the macroscopic tip is very far from the surface and assuming a very large number of particles, \( N \), i.e. \( N \to \infty \) (a realistic assumption for tip sizes \( > \mu m \)), the continuum approximation can be used, describing the full interaction as an integral over the surface, \( S \). Therefore, writing this interaction

\(^1\)i.e. within a few Ångströms since the vdW contribution to the force scales as \( 1/z^2 \) (c.f. Eq. 3.23).

\(^2\)Considering the first term in the expansion only.

\(^3\)Although often referred to as the London constant, it also contains the Keesom and Debye constants, cf. Section 2.3.
3.4 Modelling Macroscopic Forces

![Schematics illustrating a point-charge at distance \( r_0 \) above a single-layer surface (a) and a sphere at distance \( r \) above a semi-infinite surface (b).](image)

Figure 3.3: Schematics illustrating a point-charge at distance \( r_0 \) above a single-layer surface (a) and a sphere at distance \( r \) above a semi-infinite surface (b).

Energy in terms of a point-charge, positioned at \( r_0 \), separated from the surface by a distance \( z_0 \), gives

\[
U_{vdW}(z_0) = -\sum_i^N \frac{C_6}{|r_0 - r_i|^6} = -\int_S \frac{C_6}{|r_0 - r|^6} \rho_S \, d\mathbf{r} = -A \int_S \frac{d\mathbf{r}}{|r_0 - r|^6},
\]

where \( \rho_S \) is the surface density (assumed uniform over the volume of a given material) and \( A \) is a constant that has absorbed all of the \( C_6 \) coefficients. Treating the problem now in cylindrical co-ordinates\(^2\) transforms \( r_0(0,0,z_0) \) and \( r_i(x_i,y_i,z_i) \) into \( r_0(0,0,z_0) \) and \( r_i(\eta_i, \theta, z_i) \). Therefore the distance between the point-charge and surface atoms, \( l_i \), is given by \( l_i = |r_0 - r_i| = \sqrt{\eta_i^2 + z_0^2} \).

Then using the Jacobian associated with this change of co-ordinates simplifies the integral, giving

\[
U_{vdW}(z_0) = -A \int_0^\infty d\eta \int_0^{2\pi} \frac{\eta}{16} d\theta = -2\pi A \int_0^\infty \frac{\eta}{(\eta^2 + z_0^2)^{3/2}} d\eta = -\frac{\pi A}{2z_0^4},
\]

where \( z_0 \) is now the \( z \)-separation of the point-charge and the surface. Extending this simple model now to a semi-infinite surface, i.e. to a point-charge above a three-dimensional surface, requires only integration from the surface to the bottom of the material. The thickness of the slab can be thought of as infinite due to its length scales relative to that of the point particle, giving the result that

\[
U_{vdW}(z_0) = -\frac{\pi A}{2} \int_{z_0}^\infty \frac{1}{z_0^3} dz_0 = -\frac{\pi A}{6z_0^4}.
\]

\(^1\)The \( z_0 \) used here is not to be confused with the average tip-surface separation, in this description of how to model the vdW forces only, it is the instantaneous charge-surface separation.

\(^2\)\( x = \eta \cos(\theta), y = \eta \sin(\theta) \) and \( z = z \).
Two commonly used macroscopic tip models are the spherical and sphere-terminated conical tip models. The latter is said to be more accurate than the former as it more accurately represents a sharp tip (in the macroscopic sense). In the conical tip model the cone is place on top of the sphere/hemisphere and hence is further away from the surface. Although the results may be different, the treatment of both are very similar, however, for this research only the simpler spherical tip model was used as the cantilever tips used in the experiments typically had relatively large termination radii ($\geq 10$ nm)\(^1\). Thus, finally, the description of the interaction energy between the surface of a sphere and slab surface, where any point on the outside of the sphere is separated from the slab surface by \(z_S\), as in Fig. 3.3(b), is obtained by integrating Eq. 3.18 over the volume of the sphere, \(V\), i.e.

\[
U_{vdW}(z_0) = -\frac{\pi A}{6} \int_V \frac{\rho_{Sph}}{z_S^3} dV,
\]

where \(z_0\) is the \(z\)-separation of the base of the sphere and the surface, and \(\rho_{Sph}\) is the density of the sphere. Then, by writing all of the variables in terms of \(z_0\) and the sphere radius \(R\), and by using the same methodology that was used to obtain Eq. 3.17 (but with several more lines of algebra), the solution is found to be

\[
U_{vdW}(z_0) = -\frac{H}{6} \left[ \ln \left( \frac{z_0}{z_0 + 2R} \right) + \frac{2R(z_0 + R)}{z_0(z_0 + 2R)} \right],
\]

where \(H = \pi^2 A \rho_{Sph} = \pi^2 C_6 \rho_S \rho_{Sph}\) and is known as the Hamaker constant [39], which is entirely material dependent. Finally the force, \(F_{vdW}(z_0)\), for such interactions can be now obtained as the derivative of the interaction energy (refer to Eq. 3.7), and is given by

\[
F_{vdW}(z_0) = -\nabla_{z_0} U_{vdW}(z_0) = \frac{2HR^3}{3z_0^2(z_0 + 2R)^2}.
\]

Although this is the full form of the force, it is often not used as a simplification is usually applied after considering the relative sizes of the tip apex radius and its distance from the surface, i.e. utilising that typically \(R \gg z_0\). This simplification reduces both Eq. 3.20 and Eq. 3.21, respectively, to

\[
U_{vdW}(z_0) = -\frac{H}{6} \left[ \ln \frac{z_0}{2R} + \frac{R}{z_0} \right],
\]

and

\[
F_{vdW}(z_0) = \frac{HR}{6z_0^2}.
\]

Due to the inverse square relationship between this force and the tip-surface separation, at large separations it is negligible, however at small enough separations this macroscopic force can in fact dominate all other forces. It can also been seen that there is a dramatic change in the inverse

\(^1\)Since the distance of the conical portion of the tip to the surface will be large in comparison to the variation of the tip-surface distance the change would be small $< 10\%$ of the background interaction.
power of the vdW force calculated for two atoms (Eq. 3.15) to that of a sphere and a surface, (Eq. 3.23) from decaying as the 7th power of the distance to a much slower decay of just the 2nd power.

### 3.5 Classical Mechanics → Quantum Mechanics

In NC-AFM experiments, it is the microscopic (atomic) tip-surface interactions that give rise to atomic resolution, thus their chemistry needs to be modelled. It is therefore important to make clear to the reader that all chemistry is determined solely by electronic interactions. This means that regardless of whether the system consists of small clusters, bulk materials or semi-infinite surfaces, to model any of their properties, a correct description of the electrons in the material is required. This necessitates the use of quantum mechanics. For additional information outside of this thesis on describing electronic interactions, the author refers the reader to the thesis of Manuela Mura [40], the book by Richard Martin [41], entitled “Electronic Structure”, and the book of Lev Kantorovich [42], entitled “Quantum Theory of the Solid State: An Introduction”.

In order to now describe the microscopic forces, the classical discussion will now be concluded, leading into the quantum mechanical framework that allow these forces to be calculated.

#### 3.5.1 Conservation Laws

Now that the form for the Lagrangian is known (Eq. 3.6), several crucial physical laws can be derived, for which the starting point is the total time differential of the $L$ for a closed system,

$$\frac{dL}{dt} = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}_i} \dot{x}_i + \sum_{i=1}^{n} \frac{\partial L}{\partial x_i} \ddot{x}_i + \sum_{i=1}^{n} \frac{\partial L}{\partial t}.$$  \hspace{1cm} (3.24)

Conveniently, applying the homogeneity of time to this immediately eliminates the third term in Eq. 3.24. Using this results and then replacing the $\partial L/\partial x_i$ terms with the expression obtained by rearranging Eq. 3.4 gives that

$$\frac{dL}{dt} = \sum_{i=1}^{n} \frac{d}{dt} \left( \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} \right),$$  \hspace{1cm} (3.25)

or in other words, that

$$\frac{d}{dt} \left( \sum_{i=1}^{n} \dot{x}_i \frac{\partial L}{\partial x_i} - L \right) = 0$$  \hspace{1cm} (3.26)

therefore the terms within the bracket must be a constant, i.e. that

$$\sum_{i=1}^{n} \dot{x}_i \frac{\partial L}{\partial x_i} - L \equiv E,$$  \hspace{1cm} (3.27)
where the constant $E$ for the closed system is known as the energy of the system. Using Eq. 3.6, that $L = T(x, \dot{x}) - U(x)$, where $T$ is a quadratic function and substituting this into Eq. 3.27 gives

$$\sum_{i=1}^{n} \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} = \sum_{i=1}^{n} \dot{x}_i \frac{\partial T}{\partial \dot{x}_i} = 2T,$$

(3.28)

which reveals the more conventional form for the energy,

$$E = T(x, \dot{x}) + U(x)$$

(3.29)

Now applying the homogeneity of space means that $L$ does not change upon translation of all particles in the system by the same vector, or more specifically for a system of $N$ atoms,

$$\delta L = L(x + \delta x) - L(x) = 0.$$  

(3.30)

Since $\delta x$ is arbitrary, this is equivalent to writing

$$\sum_{j=1}^{N} \frac{\partial L}{\partial x_j} = 0,$$

(3.31)

i.e. the sum of all forces within the system is zero. Substituting this into Eq. 3.6 gives

$$\sum_{j=1}^{N} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_j} \right) = \frac{d}{dt} \sum_{j=1}^{N} \frac{\partial L}{\partial \dot{x}_j} = 0,$$

(3.32)

meaning that

$$\sum_{j=1}^{N} \frac{\partial L}{\partial \dot{x}_j} \equiv P$$

(3.33)

is a constant of the motion, where $P$ is called the momentum of the system. However, to obtain the conventional form of the momentum, the Lagrangian from Eq. 3.6 is substituted into this equation and the differentiation is performed, giving

$$P = \sum_{j=1}^{N} m_j \dot{x}_j.$$  

(3.34)

The momentum, $p_j$, of the $j$th particle in the system, is therefore given by

$$p_j = m_j \dot{x}_j.$$  

(3.35)

Thus, the kinetic energy ($T_j$) of the $j$-th particle can also be written as

$$T_j = \sum_{i=1}^{3} \frac{p_i^2}{2m}.$$  

(3.36)
3.6 Quantum Mechanical Description of the System

Here, the author highlights the notation used to the reader, that the index \( i \) has been used for summations over \( n \) degrees of freedoms, and \( j \) for summations over \( N \) particles to describe momentum.

Similarly to how the energy, force and momentum for a closed system are defined, by considering the isotropy of space, the angular momentum \( M \) is obtained\(^1\), written as

\[
M \equiv \sum_{j=1}^{N} x_j \times p_j.
\]  

(3.37)

Rather than simply defining quantities and playing around with calculus, real physical implications can also be drawn out from these equations. For example, Eq. 3.31 proves that the sum of all forces acting on particles within a closed system must be equal to zero, which is a crucial principle for later calculations. This can also be thought of in terms of Newton’s third law of motion, that for every action, there is an equal and opposite reaction.

### 3.5.2 The Hamiltonian

The function described in Eq. 3.27, which gives the energy of a closed system, is in fact the function of most relevance in quantum mechanics. It is known as the Hamiltonian of the system,

\[
H \equiv \sum_{i=1}^{n} \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} - L,
\]  

(3.38)

and is the function that gives the total energy of a closed system. In the following text the Hamiltonian is often written as an operator \( \hat{H} \). While this operator is often referred to as the Hamiltonian in the text, it is also called, more correctly, the Hamilton operator. This operator acts on the function describing the particles in a system and its eigenvalue is the total energy of that system.

3.6 Quantum Mechanical Description of the System

"Everything we call real is made of things that cannot be regarded as real."

**Niels H. D. Bohr**

*Physicist (1885 - 1962)*

The Hamiltonian has so far been derived quite generally, and where terms were explicitly given, they were for describing a classical mechanical system. However, the fact that electrons can also diffract implies that they can be described simultaneously in terms of both particles and waves, so called wave-particle duality, first suggested by de Broglie. This, along with the very small

\(^1\)L is normally used but to more clearly distinguish between the Lagrangian and angular momentum the more archaic \( M \) representation has been used.
length scales for electrons, means that their motion must be described quantum mechanically. Furthermore, since the electronic motions considered in this thesis are all assumed to be non-relativistic, the non-relativistic quantum mechanical form of the Hamiltonian must be used.

### 3.6.1 The Hamiltonian of a Quantum Mechanical System

Given that $\hat{H}$ must act on the function that describes the system, $\Psi$, this must now also be described. Rather than explicitly tracking each particle in the system, the function $\Psi$ is instead used such that its modulus squared describes the probability for all electrons to have particular co-ordinates, and is known as the *wave-function*. $\Psi$, the wave-function, completely describes the probability of the particles existing in particular quantum mechanical state of a physical system at any time. If $\Psi$ is known, then the probability of finding a particle in any particular state at any time $\alpha(t)$ is given by $|\Psi(\alpha)|^2$. In other words, if the wave-function of a system is known for any instance in time, then all the properties of that system in that instance as well as the behaviour of that system in all future instances is known, i.e. $\Psi$ at any time, $t$, must be determined by $\Psi$ at that time and by the principle of superposition, it has been found that this relation must be linear and its general form is given by

$$\hat{H}\Psi = \imath\hbar\frac{\partial\Psi}{\partial t}. \quad (3.39)$$

For more details, the author refers the reader to Chapter 2 of the book “Quantum Mechanics” by Landau and Lifshitz [43]. The Hamiltonian, for a closed system, or system in a constant external field, cannot explicitly contain time, and since it commutes with itself, its expectation value must be conserved. From the results in Section 3.5.1, it can be deduced that a Hamiltonian whose expectation value is conserved gives the average energy of the system, which is invariant in time (is a stationary solution). Thus, a particular state, $i$, of the system described by the wave-function, $\Psi_i$, that gives a definite value for the average energy, $E_i$, is a stationary state of the system, and transforms Eq. 3.39 into an eigenvalue equation, which is of course satisfied. This equation is given by

$$\imath\hbar\frac{\partial\Psi_i}{\partial t} = \hat{H}\Psi_i = E_i\Psi_i, \quad (3.40)$$

where $E_i$ is the eigenvalue of the energy. Naturally, the full wave-function $\Psi$ of a closed system or one under a constant external field gives the total energy of the system, i.e. $\hat{H}\Psi = E\Psi$. Using Dirac notation\(^1\) and a normalised wave-function, this is also written as

$$\hat{H}|\Psi\rangle = E|\Psi\rangle. \quad (3.41)$$

\(^1\)Also known as bra-ket notation, where $\langle\Psi|\Phi\rangle = \int_V \Phi(r_1, r_2, \cdots, r_{N_e})\Psi^\ast(r_1, r_2, \cdots, r_{N_e})d^3r_1d^3r_2\cdots d^3r_{N_e}$. 

52
Thus, the energy is a functional of the wave-function (indicated by the square brackets),

\[ E[\Psi] = \langle \Psi | \hat{H} | \Psi \rangle. \tag{3.42} \]

The energy calculated from \( \Psi \) will not be the minimum energy of the system (lowest energy/ground-state energy - henceforth referred to as ground-state energy), \( E_0 \), unless \( \Psi \) is the ground-state wave-function \( \Psi_0 \). In other words,

\[ E[\Psi] \geq E_0, \tag{3.43} \]

where \( E_0 = E[\Psi_0] \). This is the variational theorem.

### 3.6.2 The Schrödinger Equation

The Hamiltonian for a single quantum mechanical particle existing within an external field (such as the field generated by the interaction with other particles in the system) is given by

\[ \hat{H} = \frac{\hat{p}^2}{2m} + \hat{U} = -\frac{\hbar^2}{2m} \nabla^2 + \hat{U}. \tag{3.44} \]

Substituting Eq. 3.44 into Eq. 3.40 gives the wave-equation for a single particle moving in an external field,

\[ i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + U \Psi. \tag{3.45} \]

This equation, for a particle in a quantum state not changing with time, i.e. when \( i\hbar \partial \Psi/\partial t = E \Psi \), is therefore given as

\[ -\frac{\hbar^2}{2m} \nabla^2 \Psi + U \Psi = E \Psi. \tag{3.46} \]

These equations, Eq. 3.45 and Eq. 3.46, are known as the time-dependent Schrödinger equation and time-independent Schrödinger equation, respectively, where the first terms are the kinetic energies and the second the potential energies.

### 3.6.3 The Many-Body Schrödinger equation

Since, it is not the focus of this thesis to discuss the evolution in time of any quantum mechanical systems, only the time-independent Schrödinger equation will be discussed henceforth. Furthermore, the systems discussed in this thesis consist of many atoms, most of which each have several electrons. For this reason, instead of using the one-body Schrödinger equation, derived from the single particle Hamiltonian, the many-body Schrödinger equation is used.

The total wave-function, to accurately describe the system, must be a function of every nuclear (atomic core) and electronic position, as well as electronic spin, denoted by indices \( n, e, \) and \( \sigma \), respectively, and capital letters for atoms, lower case for electrons, i.e. \( \Psi = \Psi(R_n, r_e, \sigma_e) \). The electronic spin, \( \sigma_e \), can either be ‘up’ or ‘down’, corresponding to states orthogonal to each
other. Therefore, for an \( N \)-body system (consisting of \( N_n \) nuclei and \( N_e \) electrons), the total wave-function has \( 3N_n + 4N_e \) degrees of freedom, three spatial co-ordinates and spin (for electrons).

Writing out the many-body Schrödinger equation compactly gives

\[
\hat{H} \Psi = \left\{ \hat{T}_n + \hat{U}_{nn} + \hat{T}_e + \hat{U}_{ne} + \hat{U}_{ee} \right\} \Psi = E \Psi, \tag{3.47}
\]

where the \( \hat{T} \) terms are the kinetic energy operators, the \( \hat{U} \) terms are the Coulomb energy operators (positive if repulsion, negative if attraction) and the subscripts indicate which particles the operators act on. Writing out these terms explicitly, for an \( N_n \) atom, \( N_e \) electron system, gives:

\[
\hat{T}_n = -\sum_{I=1}^{N_n} \frac{\hbar^2}{2M_I} \nabla_I^2, \tag{3.48}
\]

\[
\hat{U}_{nn} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{l=1}^{N_n} \sum_{j=1}^{N_n} \frac{Z_I Z_J e^2}{|R_I - R_J|} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{I=1}^{N_n} \sum_{J>I} Z_I Z_J e^2 |R_I - R_J|, \tag{3.49}
\]

\[
\hat{T}_e = -\frac{\hbar^2}{2m_e} \sum_{i=1}^{N_e} \nabla_i^2, \tag{3.50}
\]

\[
\hat{U}_{ee} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \frac{e^2}{|r_i - r_j|} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N_e} \sum_{j>i} e^2 |r_i - r_j|, \tag{3.51}
\]

\[
\hat{U}_{ne} = -\frac{1}{4\pi\epsilon_0} \sum_{I=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_I e^2}{|R_I - r_i|}. \tag{3.52}
\]

The 1/2 pre-factor has been used in the \( \hat{U}_{nn} \) and \( \hat{U}_{ee} \) terms to discount for any double counting. The ionic and electronic positions are represented by \( \mathbf{R} \) and \( \mathbf{r} \) respectively; their masses are given by \( M \) and \( m \), respectively. All other symbols have their usual meanings. In order to simplify Eq. 3.47, for use in the remainder of this chapter, standard atomic (natural) units have been adopted, simplifying this expression to

\[
\hat{H} = -\frac{1}{2} \sum_{l=1}^{N_n} \frac{1}{M_l} \nabla_I^2 + \sum_{l=1}^{N_n} \sum_{j>l} \frac{Z_I Z_J}{|R_I - R_J|} - \frac{1}{2} \sum_{i=1}^{N_e} \nabla_i^2 + \sum_{i=1}^{N_e} \sum_{j>i} \frac{1}{|r_i - r_j|} - \sum_{l=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_I}{|R_I - r_i|}. \tag{3.53}
\]

Although the time-independent many-body Schrödinger equation has been simplified, in its exact form, an analytical solution can only be found for a very simple system, such as for a hydrogen atom. This is because, in addition to the fact that there are \( 3N_n + 4N_e \) degrees of freedom, the nuclear motion is coupled to those of the electrons. Therefore, approximations are required in order to be able to solve Eq. 3.47.
3.6 Quantum Mechanical Description of the System

3.6.4 The Born-Oppenheimer/Adiabatic approximation

To help reduce the complexity of the many-body Schrödinger equation, the Born-Oppenheimer approximation is used. In 1927, in the book by Max Born and Julius R. Oppenheimer [44], it was suggested that because nuclear masses were much larger than electronic masses, the nuclei move about two orders of magnitude slower than electrons, i.e. $\dot{R}_n \ll \dot{r}_e$, implying that electrons may be considered to obtain their equilibrium configurations “instantaneously” after any nuclear motion. This also implies a spatial localisation of the nuclear wave-function. This part of the Born-Oppenheimer approximation originates from the Adiabatic theorem.

Now, within the adiabatic approximation, which assumes the electrons instantaneously adapt their wave-function according to changes in nuclear positions, the electronic system will remain in the ground-state, and nuclear co-ordinates simply appear as parameters in the electronic wave-function; nuclei may be treated classically as moving in a potential generated by the electrons, where the nuclear kinetic energy acts as a perturbation, if required. Thus, due to the disparity between relative nuclear and electronic motion, the nuclei are considered frozen for each electronic calculation; electrons instantaneously adjust to the potential created by the nuclei. This allows for the wave-function to be written as

$$\Psi(R_n, r_e, \sigma_e) = \chi(R_n) \Phi(r_e, \sigma_e; R_n),$$  \hspace{1cm} (3.54)

where the wave-function $\chi_n(R_n)$ is a function only of the nuclear positions\(^1\), and $\Phi(r_e, \sigma_e; R_n)$ depends on both the electronic degrees of freedom, including the spin, and parametrically on nuclear positions.

This effectively decouples the electronic and ionic motions, finally allowing for a redefinition of the Schrödinger equation,

$$\hat{H}\Psi = \{\hat{H}_n + \hat{H}_e\} \chi \Phi,$$  \hspace{1cm} (3.55)

where $\{\hat{T}_n + \hat{U}_{nn}\}$ can be considered the nuclear Hamiltonian, $\hat{H}_n$, and $\{\hat{T}_e + \hat{U}_{ee}\}$ describes the electronic motion, that is coupled to the nuclei by $\hat{U}_{ne}$, collectively known as the electronic Hamiltonian, $\hat{H}_e$.

Hence, the electronic Hamiltonian operator ($\hat{H}_e$), which is the term that actually gives the properties of the system for a given nuclear configuration, gives

$$\hat{H}_e \Phi(r_e, \sigma_e; R_n) = \epsilon \Phi(r_e, \sigma_e; R_n),$$  \hspace{1cm} (3.56)

where

$$\hat{H}_e(r_e; R_n) = \hat{T}_e(r_e) + \hat{U}_{ee}(r_e) + \hat{U}_{ne}(r_e; R_n).$$  \hspace{1cm} (3.57)

\(^1\)It is this term therefore, that holds the translational, vibrational and rotational information.
Thus, the electronic energy functional, given by
\[ E_e[\Phi] = T_e[\Phi] + U_{ee}[\Phi] + U_{ne}[\Phi], \] (3.58)
contains three terms: the kinetic energy \( T_e \), the electron-electron interaction energy \( U_{ee} \), and the energy due to the external interaction \( U_{ne} \). From these last two equations, the parametric dependence of the electronic energy on the nuclear configuration can be seen in the \( \hat{U}_{ne} \) term.

### 3.7 Solving the Time-Independent Schrödinger Equation for Electronic Systems

#### 3.7.1 Hartree-Fock Method

Early attempts to solve the full \( N_e \)-body problem in Eq. 3.56 involved reducing it to \( N_e \) one-body equations by way of the Hartree method. This is based on the idea that if the exact wave-function for an \( N_e \)-electron system is not known, then an approximate wave-function, \( \Phi_H \) may be used, given as
\[ \Phi_H(r) = \prod_{i=1}^{N_e} \phi_i(r_i, \sigma_i), \] (3.59)
where \( \phi_i(r_i, \sigma_i) \) are electronic states (also known as single-electron or molecular orbitals). However, this approximate wave-function does not lead to a natural solution of Eq. 3.56 (the ground-state energy) but rather a higher energy solution. Nevertheless, by using the variational principle, the ground-state wave-function (and therefore, ground-state energy) may be determined to a reasonable precision, by varying the wave-function about the minimum of the energy. While this is a very powerful method, it neglects the fermionic nature of the electrons and their spin.

The spin contribution is such an issue because of the Pauli’s exclusion principle, which is closely related to the indistinguishability of electrons. The former states that no two fermions (which electrons are) can simultaneously exist in the same quantum state, which results in a non-local exchange contribution to the energy, known as the exchange energy. The effect this exchange contribution has on same-spin electron is the appearance of “exchange holes” around electrons, due to same-spin electrons avoiding each other. This implicit anti-symmetrisation of the wave-function causes a spatial separation of electrons, resulting in a reduction of the likelihood of finding two same-spin electrons in close proximity to each other. Although neglected in the original Hartree method, it is accounted for in the more popular Hartree-Fock (HF) method, given by the Slater determinant
\[ \Phi(r, \sigma) = \frac{1}{\sqrt{N_e!}} \det |\phi_1(r_1, \sigma_1)\phi_2(r_2, \sigma_2)\cdots\phi_{N_e}(r_{N_e}, \sigma_{N_e})|. \] (3.60)
3.7 Solving the Time-Independent Schrödinger Equation for Electronic Systems

Using the HF method, the ground-state orbitals \( \phi_i \) can be determined by applying the variational theorem (cf. Eq. 3.43) using \( \phi_i \) to calculate the energy under the constraint that the orbitals are orthonormal (i.e. that \( \langle \phi_i | \phi_j \rangle = \delta_{ij} \)). Doing this now accounts for the fermionic nature of electrons. This methods leads to a set of equations known as the Hartree-Fock equations, which must be solved self-consistently. The HF equations describe non-interacting electrons under the influence of the electronic Coulomb potential and non-local exchange potential, which together are referred to as the mean field. Hence the HF method is also known as the self-consistent field (SCF) method.

A contribution to the energy, however, not accounted for in the HF method comes from an inability of this method to properly describe the correlation of electrons of either spin. In fact the difference between the exact energy, and the energy calculated using the HF method is known as the correlation energy.

Although the HF method is still very powerful and informative, the total energy of a system calculated using the HF method will differ from the true value in a non-trivial way due to the inability of the method to account for the correlation energy.

3.7.2 Hohenberg-Kohn Theorem

The basis of density functional theory (DFT) is that instead of using electronic positions to define the wave-function, the electron density, \( n \), is used, which is a measure of the probability finding single electron of spin \( \beta \) existing within radius \( r \) from an atom or molecule\(^1\). The density of a finite system should be such that it vanishes as \( r \) goes to infinity, i.e. \( n(r \to 0) = 0 \). It should also be such that the integration of \( n \) over all space\(^2\) gives the total number of electrons, i.e. \( \int n(r) \, dr = N_e \). It therefore has the distinct advantage of reducing the \( 4N_e \) variables in the many-body problem to simply 4 variables (consisting of position and spin), while still containing many (but not all) of the physical information about the system\(^3\). This is made possible due to the Hohenberg-Kohn theorems, proposed by Pierre C. Hohenberg and Walter Kohn in 1964 [45], made for an electronic system acting within an external potential, in this case \( \nu_{ex} \).

**Theorem I**

The first theorem gives:

---

\(^1\) \( n(r) \, d^3r \) gives the number of electrons expected to be found within the small volume enclosed by \( d^3r \).

\(^2\) The volume integral over all \( r \) is \( \int_V \, dr \), but henceforth is shall simply be written as \( \int \, dr \).

\(^3\) Properties of the system such as energies and forces, can be determined from the electronic density, and magnetisation from the spin density. However, there are some properties that are beyond the ability of DFT to determine, a few examples of which are the full kinetic energy, electron current, and the exact interaction energy.
This theorem is proven by contradiction that the external potential and the ground-state wave-function are determined uniquely by the ground-state density. Hence the energy of the system has a one-to-one correspondence with the ground-state density and the wave-function. Furthermore, for an isolated molecular or crystalline system, under no other influences, the external potential acting on the electrons, $v_{ext}$, which determines $U_{ext}$, is due solely to the external nuclear potential acting on those electrons, $v_{ne}$, which determines $U_{ne}$, i.e. $v_{ext} = v_{ne}$. To summarise, this theorem allows the total energy of the $N_e$-electron system to be determined by the single electron density $n(r, \sigma)$,}

$$ n(r, \sigma) = \langle \Phi(r, \sigma) | \hat{n}(r, \sigma) | \Phi(r, \sigma) \rangle$$  \hspace{1cm} (3.61)

using

$$ \hat{n}(r, \sigma) = \sum_{i=1}^{N_e} \delta(r - r_i) \delta_{\sigma, \sigma_i},$$  \hspace{1cm} (3.62)

where $\delta(r - r_i)$ is the density of electron $i$ at position $r$. Eq. 3.61 is the form resulting from a normalised wave-function, and henceforth, for convention, $n(r, \sigma)$ will simply be written as $n(r)$\textsuperscript{1}.

Thus, the total electronic energy, as described by Theorem I, can be written as a functional of the ground-state density, given by

$$ E_e[n(r)] = F[n(r)] + \int n(r)v_{ne}(r)dr,$$  \hspace{1cm} (3.63)

where $F[n(r)]$ is an unknown universal functional of the density and the second term is the energy, $U_{ne}$, due to the external field, $v_{ne}$, i.e. the nuclear potential energy.

\textbf{Theorem II}

The second of the Hohenberg-Kohn theorems states [41] that:

\begin{quote}
A universal energy functional can be defined in terms of the density and is valid for any external field. Furthermore, the density that minimises this functional is the ground-state density, and the corresponding energy is the ground-state energy.
\end{quote}

\textsuperscript{1}The spin inclusive forms of some of the equations may differ slightly, however they will not be included in this text, this is for simplicity as the author wishes only to convey the methodologies.
The reasoning behind this is that, from Theorem I, it has been shown that the density uniquely determines the external potential and thus the energy, which in turn determines $\Phi(r)$ by way of the variational method. Therefore, $\Phi(r)$ must be a functional of $n(r)$, even though from Eq. 3.61 it is known that $n(r)$ is determined from $\Phi(r)$. Thus, by minimising the energy functional using the variational method, the ground-state density can be determined, i.e. that

$$E_e[n(r)] = \min_{\Phi(r)\rightarrow n_0} \langle \Phi(r)|\hat{H}_e|\Phi(r)\rangle,$$

(3.64)

where the density of the minimised energy functional is the ground-state density, $n_0(r)$. This theorem establishes a variational principle for determining the ground-state density.

### 3.8 Density Functional Theory

The following fundamental statement can be made based on the two Hohenberg-Kohn theorems: “In order to determine the ground-state energy of the system, the energy functional can be minimised using the Lagrange multiplier, $\mu$, enforcing the constraint that the number of electrons is conserved (that $\int n(r)dr = N_e$)”. Using Eq. 3.63, this gives

$$\delta \left[ F[n(r)] + \int n(r)v_ne(r)dr - \mu \left( \int n(r)dr - N_e \right) \right] = 0,$$

(3.65)

where $\mu$ is known as the electronic chemical potential. A slightly neater representation of this, after performing functional differentiation, is the equivalent Euler-Lagrange equation for Eq. 3.65, which is

$$\frac{\delta F}{\delta n(r)} + v_{ne}(r) = \mu.$$

(3.66)

Since Theorem II states that the universal energy functional $F[n(r)]$ does not depend on the external potential, the second term in Eq. 3.65 and Eq. 3.66 is trivial. Thus, Eq. 3.65 gives that if $F[n(r)]$ is known, then the exact ground-state density and energy can be obtained.

As one might infer from the text in Theorems I and II, DFT is generally restricted to studies of ground-state properties. However, extensions can be made to excited state, but since this is outside the scope of this research, treatment of excited states will not be discussed. Nonetheless, the HK theorems show that, unlike the Hartree and Hartree-Fock methods, there exists a, now popular, numerical methodology for solving Eq. 3.56, DFT, that reduces the number of degrees of freedom in Eq. 3.56 considerably. For further material on DFT beyond what this thesis offers, the author refers the reader to the book of Kieron Burke [46] in addition to those of Martin [41] and Kantorovich [42].
3.8 Density Functional Theory

3.8.1 Kohn-Sham Equations

As previously stated the universal functional is unknown, however, By comparing Eq. 3.63 with Eq. 3.58, it can be seen that the universal energy functional consists of the kinetic and electron-electron interaction energy terms, i.e.

\[ F[n(r)] = T_e[n(r)] + U_{ee}[n(r)]. \]  (3.67)

The analytical forms of both \( T_e \) and \( U_{ee} \) are not known as the exact description of the contributions to the kinetic energy from electron-electron interactions in terms of the density is not known. To approximate \( F[n(r)] \), Kohn and Lu J. Sham suggested [47] that the fully-interacting \( N_e \)-electron system (electrons with electrons, and electrons with nuclei), the real system, could be exactly mapped onto a fictitious non-interacting \( N_e \)-electron system (the non-interacting electron gas model), where electrons in state \( i \), described by one-electron orbitals, \( \phi_i \), move within an effective single-electron potential, known as the Kohn-Sham potential, \( \nu_{KS} \). The crucial point here is that the ground-state density of the non-interacting system is exactly the same as that of the fully-interacting system.

Now, unlike for the fully-interacting system, the kinetic energy of the non-interacting system that gives the same density as the interacting system is known, given as

\[ T_{KS}[n(r)] = -\frac{1}{2} \sum_{i}^{N_e} \langle \phi_i(r) | \nabla^2 | \phi_i(r) \rangle. \]  (3.68)

Hence, for \( N_e \)-electron non-interacting system, the energies of the one-electron orbitals, denoted by \( \epsilon_i \) is given by

\[ \left\{ -\frac{\nabla^2}{2} + \nu_{KS} \right\} \phi_i^{KS}(r) = \epsilon_i \phi_i^{KS}(r). \]  (3.69)

These equations (Eq. 3.69) are known as the Kohn-Sham equations and orbitals which satisfy them are known as the Kohn-Sham orbitals, \( \phi_i^{KS} \). Furthermore, as \( v_{ee} = 0 \) for the non-interacting system, and so \( U_{ee} = 0 \), Eq. 3.66 for the non-interacting system is

\[ \frac{\delta T_{KS}}{\delta n(r)} + \nu_{KS}(r) = \mu. \]  (3.70)

Since Eq. 3.68 requires knowledge of the one-electron orbitals, it makes sense to also construct the density in terms of these one-electron orbitals. Doing so gives the density as

\[ n(r) = \sum_i |\phi_i(r)|^2, \]  (3.71)

where \( i \) runs over all electrons. Since the ground-state density and energy is to be determined, electrons are not likely to occupy the higher energy states, it is therefore worthwhile rewriting the
density (Eq. 3.71) in terms of the lowest \( N \)-occupied electronic states \( (N_{\text{occ}}) \) that contribute the most to the ground-state density. This gives \( n(r) \) as

\[
n(r) = \sum_{i} N_{\text{occ}} |\phi_i(r)|^2, \tag{3.72}
\]

where the density is the sum over all the occupied orbitals. However, since it is not known before any calculation whether a state is occupied or not, it is more convenient to write Eq. 3.72 as

\[
n(r) = \sum_{i} w_i |\phi_i(r)|^2, \tag{3.73}
\]

where \( w_i \) is a weighting prefactor that is either 1 if occupied or 0 if not (a step function), hence could be given by the Heavyside or Fermi-Dirac function, and \( N_w \) is the number of orbitals used to construct the density. Eq. 3.73 is the same as Eq. 3.72 if \( w_i = 1 \) for \( i \leq N_{\text{occ}} \) and \( w_i = 0 \) for \( i > N_{\text{occ}} \).

Now, \( F[n(r)] \) for the fully-interacting system can be defined in terms of the kinetic energy of the non-interacting system and the electronic Coulomb repulsion energy (the Hartree term), \( T_{KS} \) and \( U_H \), respectively, i.e.

\[
F[n(r)] = T_{KS}[n(r)] + \frac{1}{2} \int n(r) v_H[n(r')] dr + U_{xc}[n(r)], \tag{3.74}
\]

where the second term is the Coulomb energy, \( U_H \), and \( v_H \) is the Coulomb potential. However, this definition of \( F[n(r)] \) produces an additional third term, \( U_{xc} \), known as the exchange-correlation energy, which accounts for the exchange and correlation energies discussed in Section 3.7.1, as well as for the missing part of the kinetic energy, due to electron-electron interactions, which needs to be accounted for as the kinetic energy of the interacting electron gas has been replaced by that of the non-interacting gas. Writing \( F[n(r)] \) this way allows the Euler-Lagrange equations given in Eq. 3.66 and Eq. 3.70 to be related by

\[
\mu = \delta F \delta n(r) + v_{ne} = \delta T_{KS} \delta n(r) + v_{KS}, \tag{3.75}
\]

which is only consistent if

\[
v_{KS} = v_{ne} + \delta U_H \delta n(r) + \delta U_{xc} \delta n(r). \tag{3.76}
\]

This can be written more succinctly as

\[
v_{KS} = v_{ne} + v_H + v_{xc}, \tag{3.77}
\]
3.8 Density Functional Theory

where \( v_H \) is given as

\[
v_H[n(r)] = \frac{\delta U_H}{\delta n(r)} = \int \frac{n(r')}{|r-r'|} \, dr',
\]

and \( v_{xc} \) is the exchange-correlation potential, given as

\[
v_{xc}[n(r)] = \frac{\delta U_{xc}}{\delta n(r)}.
\]

Using the form of the density described in Eq. 3.73, the one-electron orbitals used to construct the density can be determined by varying \( F[n(r)] \) under the constraint that the one-electron orbitals are orthonormal, i.e. Eq. 3.65 becomes

\[
\delta \left[ F[n(r)] + \int n(r) v_{nc}(r) \, dr - \sum_{i,j} \epsilon_{ij} \left( \langle \phi_i | \phi_j \rangle - \delta_{ij} \right) \right] = 0,
\]

where \( \epsilon_{ij} \) are the corresponding Lagrange multipliers. Finally, after determining \( v_{xc} \), which is exactly unknown but is frequently approximation using the generalised gradient approximation (GGA) (see Section 3.8.2), the total electronic energy of the fully-interacting system, \( E_e \), as defined by Eq. 3.63, can now be written in terms of the one-electron energies of the non-interacting system, \( \epsilon_i \), where \( T_{KS} \) is common to both systems. This is given as

\[
E_e = \sum_{i}^{N_u} \epsilon_i - \frac{1}{2} \int n(r) v_H \, dr - \left( \int n(r) v_{xc} \, dr - U_{xc} \right),
\]

where the second term is simply needed to avoid double counting of the Coulomb energy contribution, as is the third term (the two terms within the brackets), related to the exchange-correlation energy.

In summary, the Hohenberg-Kohn theorems reduce the number of degrees of freedom that the many-body problem is usually associated with, while still being a many-body problem. Due to the relationships of the Kohn-Sham equations and the electronic density, the Kohn-Sham equations must be solved iteratively (self-consistently). Eq. 3.73 uses the Kohn-Sham orbitals to construct the density. The density is then used to determine the Kohn-Sham (effective) potential. The Kohn-Sham potential is then used to calculate the eigenvalues of the Kohn-Sham equations which are used to calculate the Kohn-Sham orbitals. Due to the Kohn-Sham ansatz, the many-body problem of interacting particles, is directly mapped onto many single-body problems of non-interacting particles. An illustration of this is given in Fig. 3.4. These self-consistent calculations are terminated only when the specified precision has been met.

As mentioned before, the exact form of the KE expressed via density for the full-interacting system is not known but it is known for the non-interacting system, this is an advantage of using

\(^1\) The way in which this is achieved can be found in any good DFT books (c.f. Refs. [41, 42]).
the Kohn-Sham equations. Furthermore, the Kohn-Sham equations and Hatree-Fock equations are both one-electron equations which must be solved self-consistently, the difference being that the Kohn-Sham equations also take into account electron correlation. The price, however, of this advantage is the appearance of the exchange-correlation term, whose exact analytical form is unknown. It should also be noted here that the true electronic wave-function is not the Slater determinant of the Kohn-Sham orbitals, it cannot be determined from the density which is constructed from the Kohn-Sham orbitals.

### 3.8.2 Exchange and Correlation Functionals

The Hohenberg-Kohn theorems and Kohn-Sham equations, although allowing for the simplification of the fully-interacting many-body Hamiltonian, do not simplify the complexities arising from the electron-electron interactions. Rather, they move these complexities into one functional, the exchange-correlational term, $U_{xc}$. Therefore, although Kohn-Sham theory is exact within the Born-Oppenheimer approximation, its practical application is not, as the exact form of $U_{xc}$ is unknown, and is instead approximated. The implicit form of $U_{xc}$, however, from Kohn-Sham theory is $U_{xc}[n(r)] = T_{e}[n(r)] - T_{KS}[n(r)] + U_{ee}[n(r)] - U_{H}[n(r)]$, where $U_{ee}$ is the exact energy due to electron-electron interactions, and $T_{e}$, as opposite to $T_{KS}$ is the exact kinetic energy of the fully interacting electron system.

For this reason, $U_{xc}$ is usually where the largest source of error comes\(^1\). It is therefore no surprise that this is an active field of research in physics, and several researchers are attempting

---

\(^1\)Errors from other approximation may have larger contributions, cf. Section 3.8.3.3 or Section 3.8.4.
3.8 Density Functional Theory

to improve the accuracy of \( U_{xc}[n(r)] \) or are trying to find the exact functional, to reach the top of “Jacob’s Ladder” [48], where each metaphorical rung is a step closer to the exact functional (the top rung).

3.8.2.1 Local Density Approximation

The simplest of these approximations (the bottom rung) is the local (spin) density approximation, L(S)DA, of the form

\[
U_{xc}^{LDA}[n(r)] = U_{x}^{LDA}[n(r)] + U_{c}^{LDA}[n(r)] = \int n(r)\nu_{xc}^{LDA}[n(r)]d^3r,
\]

where \( \nu_{xc}^{LDA} \) is the exchange-correlation potential of an electron in a uniform electron gas, for the given density, \( n \). This assumes an electron gas with a slowly varying density, and reproduces the exact known energy of a uniform electron gas. It is a local functional and it does not take into account any large fluctuations of the density across the system. This approximation generally over-simplifies the electron-electron interaction, so although it speeds up calculation, the cost is the accuracy of the calculations. Surprisingly, however, LDA yields good results for metal systems. This is because metal systems largely consist of delocalised electrons, very similar to the electron system the approximation was made from. Furthermore, when applied to van der Waals (vdW) interacting systems, LDA has also been known to yield surprisingly reasonable results. This was later found out to be for the wrong reasons [49], a case of accidental over-binding, found to occur within the LDA for most systems.

3.8.2.2 Generalised Gradient Approximation

The next rung up is the generalised gradient approximation (GGA), of the form

\[
U_{xc}^{GGA}[n(r)] = \int n(r)\nu_{xc}^{LDA}[n(r),\nabla n(r)]d^3r.
\]

This is a semi-local method that differs from LDA in that there is also a dependence on the gradient of the electronic density. This gradient term means that the behaviour of the density around point \( r \) is also taken into account, solving some (but not all) of the inherent problems surrounding the LDA treatment (the problems arising from the locality of the \( U_{xc} \) term).

Several slightly different flavours of the GGA functional exist, differing in how they treat the exchange and correlation effects, such as Perdew-Wang 91 (PW91) [50, 51], Perdew-Burke-Ernzerhoft (PBE), [52] and revised Perdew-Burke-Ernzerhoft (revPBE) [53], and many others.
3.8.2.3 Accounting for van der Waals Interactions

Since, LDA and PBE both completely lack the mechanism to account for the highly non-local vdW interaction, several methods have been developed, two of which will now be discussed.

3.8.2.3.1 Semi-Empirical “C₆ Corrections”/Grimme Method

It is well known that the inter-atomic vdW contribution to the energy of two neutral atoms is of the form

\[ U_{\text{vdW}} \propto \frac{C_6}{r^6}, \]  

(3.84)

where \( C_6 \) is a species dependent constant and \( R \) is the inter-atomic distance (cf. Section 3.4), and several researchers have attempted to implement this within DFT as semi-empirical methods, most noticeably Grimme [54, 55], which will be discussed, and Tkatchenko and Scheffler [56].

The method developed by Grimme, DFT-D2 [54], is numerically very efficient, costing not much more than vanilla (ordinary) DFT methods. This is because the calculation of \( U_{\text{vdW}} \) is not self-consistent, and is in fact calculated empirically; for a given geometry, it calculates the relatively simple sum

\[
U_{\text{vdW}} = -s_6 \sum_{i>j}^N C_6^{ij} \frac{1}{r_{ij}^6} \frac{1}{1 + e^{-d\left(\frac{r_{ij}}{R_{ij}} - 1\right)}},
\]  

(3.85)

where \( C_6^{ij} = \sqrt{C_6^i C_6^j} \) and \( s_6 \) are system dependent empirical constants, and \( r_{ij} = r_i + r_j \) is the inter-atomic distance between atom \( i \) and \( j \). This vdW contributions is cut-off at the radius \( R_{ij} \), but to ensure a smooth transition from the short to long-ranged description, it is introduced with scaling factors \( d \) and \( s_R \) which act as smoothing parameters. As the name suggests, the Grimme correction can be implemented with different functionals, such as PBE or some hybrid-GGA functionals, and gives the total energy as \( E_e = E_{\text{DFT}} + U_{\text{vdW}} \). Whilst the method is quite accurate for some systems and is computationally cheap, it has the serious drawback of no dependence on the electronic density. In addition to this, due to its material dependance, it is quite inflexible and not easily transferable. However, some of these issues have been addressed in DFT-D3 [55], in which a co-ordination number has been introduced, which holds information about the chemical environment.

3.8.2.3.2 Self-Consistent van der Waals Density Functional

Since the scheme suggested by Kohn and Sham, Section 3.8.1, is exact in nature and does not imply a particular form of \( U_{\text{xc}} \), vdW interactions can also be accounted for self-consistently. This

---

1Considering the first term in the expansion only.
3.8 Density Functional Theory

is precisely what has been done by Dion et al. [57] with their approximate non-local correlation functional, $U_c$, originally called vdW-DF (now known as vdW-DF1).

Using the adiabatic-connection fluctuation-dissipation theorem (ACFDT), $U_c$ may be split into two parts, the short-range and long-range (non-local), referred to as “sr” and “nl”, respectively, in equations. This gives the correlation energy as $U_c = U_{c;sr} + U_{c;nl}$.

The former, since it describes the short-range interactions, may be reasonably described by local or semi-local methods (LDA or GGA), typically LDA. However, to describe the exchange, it was suggested that the revPBE functional be used for the short-range description since it apparently shows weak-binding [57].

This non-local contribution to the energy is given by

$$U_{c;nl} = \frac{1}{2} \int \int n(r) \varphi(r, r') n(r') d^3r d^3r',$$

where $\varphi(r, r')$ is a kernel function as defined in [57], dependent on the gradient of the electron density, accounting for semi-local contributions. Hence,

$$U_{xc} = U_{xc;GGA} + U_{c;nl} + U_{c;LDA} - U_{c;GGA},$$

where $U_{xc;GGA}$ is $U_{xc}$ calculated using GGA, $U_{c;GGA}$ and $U_{c;LDA}$ are the correlation energies as calculated by GGA and LDA, respectively. This implementation effectively replaces the $U_{c;GGA}$ with the sum of the non-local, $U_{c;nl}$, and short-range, $U_{c;LDA}$, contributions. An improvement on vdW-DF1 [58], known as vdW-DF2 reduces the dependence of the functional on the uniform electron gas model, where rPW86 (PW86R) [59] was suggested for the short-range description.

The obvious advantage of this method is the inherent dependence on the electron distribution. Additionally, it can be implemented self-consistently, and also contributes to the calculated forces. The drawback of the vdW-DF1/2 however, is that the 6-dimensional integrals need to be calculated which, if calculated self-consistently, is computationally very expensive, often prohibitively so for large systems. However, schemes do exist that allow such calculations to be performed within a reasonable amount on time. When necessary in this research, the implementation by Andris Gulans was used [60]. Despite the obvious benefit of having these methods, the accuracy of them, namely what should be the description of the short-range interaction, has been found to not yet be so clear cut [61, 62].

In this thesis the PW91, PBE, revised PBE and revised PW86 exchange-correlation functionals have been used, and in order to account for vdW interactions either the PBE-D2 method or vdW-DF1/2 functionals have been used. Generally, however, unless otherwise stated, the PBE functional was used.
3.8.3 Solving the Kohn-Sham Equations: Bloch Theorem

Since this thesis will mainly be studying crystalline systems, this section will now discuss the methodology behind modelling such systems. For calculations of a periodic structure, such as infinite crystals, the periodic nature of the effective potential $v_{KS}$ is used.

Based on the fact that the $v_{KS}$ is generated by the periodically repeating nuclei, it is useful to introduce the translation vector that separates the equivalent nuclei. This is given by

$$R = \sum_{i=1}^{3} n_{i} a_{i},$$

(3.88)

where $n_{i}$ is the integer number of times the lattice vector, $a_{i}$, is repeated in the $i^{th}$ direction. It therefore has the property that $v_{KS}(r) = v_{KS}(r + R)$. Then assuming a large number, $N$, of periodically repeating primitive cells, of volume $\Omega_{\text{cell}}$, such that the system is within a large cell volume $\Omega = N \Omega_{\text{cell}}$, the electrons can be modelled as being acted upon by a periodically repeating potential, for which the electron density must also have the same periodicity.

If $\hat{R}$ is the translation operator, then the property of translational invariance (see Section 3.5.1), implies that $\hat{R} \cdot \hat{H}_{e} = \hat{H}_{e} \cdot \hat{R}$. The Bloch theorem [63] then states that the eigenvalue for $\hat{R}$ is $e^{ik \cdot R}$, for a normalised wave-function, where $k$ is the wave-vector. Thus, for a function $u_{k}(r)$, with the same periodicity as $v_{KS}(r)$, the Kohn-Sham orbitals can be constructed as

$$\phi_{k}(r) = e^{ik \cdot R} u_{k}(r),$$

(3.89)

satisfying the orthogonality condition

$$\int_{\Omega_{\text{cell}}} \phi_{k}(r)^{\ast} \phi_{k'}(r) d^{3}r = \delta_{kk'}. $$

(3.90)

With the Kohn-Sham orbitals constructed this way, the equation $\phi_{k}(r + R) = e^{ik \cdot R} \phi_{k}(r)$ holds. Eq. 3.89 is known as the Bloch-wave expansion (also known as a Bloch-state).

The energy eigenvalue in the Kohn-Sham equation, Eq. 3.69, is therefore given by

$$\hat{H}_{e} \phi_{k}(r) = \epsilon_{k} \phi_{k}(r).$$

(3.91)

The plot of the energies in Eq. 3.91 with respect to different $k$ ($\epsilon_{k}$ in $k$-space/reciprocal-space) for a given material gives the band-structure for that material, which is usually plotted along the directions of Brillouin zone (BZ - the primitive cell in reciprocal-space).
3.8 Density Functional Theory

3.8.3.1 Plane-Waves as Implemented in VASP

VASP stands for the Vienna Ab initio Simulation Package (VASP)\(^1\) [64, 65, 66, 67], and is a DFT code that expands the one-electron wave-function in Eq. 3.89 in terms of plane-waves, and the total energy is minimised with respect to the expansion coefficients \(c_k\). Now whilst the particular methods outlined in this section may in fact not necessarily be the most efficient for calculations (and hence may not be the specific implementation in VASP), they do give a clear description of the logic and underlying theory behind the algorithms used in plane-wave calculations, commenting also on limitations.

Working in \(k\)-space, using \(N_G\) reciprocal-lattice vectors, \(G_j\), the periodic function \(u_k(r)\) is expanded into the Fourier series,

\[
u_k(r) = \frac{1}{\sqrt{\Omega}} \sum_{j=1}^{NG} c_k(G_j)e^{iG_j \cdot r}, \tag{3.92}\]

where

\[
c_k(G) = \frac{1}{\sqrt{\Omega}} \int_{\Omega_{cell}} u_k(r)e^{-iG \cdot r}d^3r. \tag{3.93}\]

This uses that \(G = 3 \sum_{i=1}^{3} m_i b_i\), \(\tag{3.94}\)

where, \(m_i\) is the integer number of times the basis vectors \(b_i\) is repeated in the \(i^{th}\) direction of the reciprocal-lattice, such that \(a_i \cdot b_j = 2\pi \delta_{ij}\).

This gives that the Bloch-wave expansion of the wave-functions used in Eq. 3.91 in terms of plane-wave basis sets is

\[
\phi_{i,k}(r) = \frac{1}{\sqrt{\Omega}} \sum_{j=1}^{NG} c_{i,k}(G_j)e^{i(k+G_j) \cdot r}, \tag{3.95}\]

where the normalised basis function \(\frac{1}{\sqrt{\Omega}} e^{i(k+G_j) \cdot r}\) is called a plane-wave, hence this is known as the plane-wave expansion. Very often, as it is in the VASP code, the coefficients \(c_k(G)\) are usually calculated using Fast Fourier transformations (FFT) which also simplifies the calculation of the wave-function. For more information on this the author refers the reader to the VASP manual [68].

Thus, \(v_{KS}\) can be written as

\[
v_{KS}(r) = \sum_{j=1}^{NG} v_{KS}(G_j)e^{iG_j \cdot r}, \tag{3.96}\]

\(^1\)Specifically version 4.6 was used for all VASP calculations performed.
where

\[ u_{KS}(G) = \frac{1}{\sqrt{V_{\text{cell}}}} \int_{\Omega_{\text{cell}}} u_{KS}(r)e^{-iG \cdot r} \, d^3r. \]  

(3.97)

### 3.8.3.2 Localised Basis-Sets as Implemented in SIESTA

Alternatively, however, \( \phi_k(r) \) may be expanded in terms of a superposition of localised orbitals, centred on and expanded around each atom.

This is the method implemented in the Spanish Initiative for Electronic Simulation with Thousands of Atoms, (SIESTA)\(^1\). This DFT code expands the wave-function in Eq. 3.89 in terms of localised basis-sets, using, unlike VASP, a linear combination of atomic orbitals (LCAO).

Each of these is presented as the product of a radial function, \( \Theta_{Inl} \), and a spherical harmonic, \( Y_{lm} \), where \( I \) denotes the atom number and \( nlm \) have their usual meanings as atomic quantum numbers. Localised orbitals, \( \vartheta(r) \), also known as a localised basis-set, are each of the form

\[ \vartheta_{Ilnm}(r) = \Theta_{Iln}(r_I)Y_{lm}(\hat{r}_I), \]  

(3.98)

where \( r_I = r - R_I \) and \( \hat{r}_I = \frac{r - R_I}{|r - R_I|} \). Hence, the wave-function is determined as an expansion of these basis sets, constructed in a similar way as Eq. 3.95, i.e.

\[ \phi_i(r) = \sum_\lambda c_{i,\lambda}\vartheta_{\lambda}(r), \]  

(3.99)

where \( \lambda = Ilnm \). The Bloch-wave expansion of the wave-functions is thus

\[ \phi_{i,k}(r) = \sum_\lambda c_{i,k,\lambda}\vartheta_{\lambda}(r)e^{ikR_\lambda}. \]  

(3.100)

When using localised basis sets, the chemical accuracy depends (not solely) on the quality of the basis sets used e.g. how spatially localised the orbitals are or the number of orbitals per atom.

### 3.8.3.3 Energy Cutoff

Now, by using the wave-function expansion in terms of plane-waves and solving Eq. 3.91 for an infinite crystal, the problem of describing an infinite number of electrons has been exactly mapped into the reciprocal-space problem of using an infinite number of reciprocal-lattice vectors within the first BZ to describe the wave-function. Hence, Eq. 3.91 becomes

\[ \sum_{j=1}^{N_{G_i}} \left\{ \frac{|k + G_i|^2}{2}\delta_{kj} + v_{KS}(G_i - G_j) \right\} c_k(G_j) = \epsilon(k)c_k(G_i). \]  

(3.101)

\(^1\)Specifically, all calculations performed using SIESTA were done with version 3.0.1.
3.8 Density Functional Theory

Usual treatment of this problem involves diagonalisation of the Hamiltonian matrix in the curly brackets. Since, obviously, an infinite representation of the system is not possible in practice, the size of the Hamiltonian matrix (the number of plane-waves used) is determined by the energy cutoff (called ENCUT in VASP), \( E_{\text{cut}} \), where \( E_{\text{cut}} = \frac{G_{\text{cut}}^2}{2} \), where \( G_{\text{cut}} > |k + G| \). All plane-waves with kinetic energy less than \( E_{\text{cut}} \) are included.

The energy cutoff therefore acts as a limiter on the chemical accuracy and representability of the calculation.

3.8.3.4 Reciprocal-Space Description of the Density

3.8.3.4.1 \( k \)-point Sampling

Another limiter is the \( k \)-point\(^1\) sampling approximation, therefore, a brief discussion will be given on how the \( k \)-point sampling may influence results.

In the representations given so far, the electrons have been considered within an infinite number of reciprocal-lattice vectors, \( G_j \). Since the orbitals are not periodic but depend on a periodic potential (i.e. they satisfy the Bloch theorem), both the reciprocal vectors and the vectors sampling the BZ are required to represent them. The latter only need to be correctly sampled at all the irreducible \( k \)-points within the first BZ, which by way of symmetry operations, can be extended over the whole reciprocal-space (\( k \)-space). This is achieved by a finite mesh in the BZ; specifically the Monkhorst-Pack scheme [70] has been used to do this throughout this thesis, which in essence provides a means to evaluate the integral

\[
n(r) = \frac{1}{\Omega_{\text{BZ}}} \int_{\text{BZ}} n_k(r) d^3k,
\]

where \( \Omega_{\text{BZ}} = \frac{(2\pi)^2}{\Omega_{\text{cell}}} \). In this equation, \( n_k(r) \) is given by Eq. 3.73, where only orbitals associated with the particular \( k \)-point specified are used. In SIESTA however, when the Monkhorst-Pack scheme is not used, a parameter known as the \( k \)-grid cutoff is used, which determines the fineness of the \( k \)-grid used [71].

3.8.3.4.2 Partial Occupancies

As discussed before, the construction of the charge density used to evaluate Eq. 3.102 is summed over states, using a weight to describe the occupancy of each state. For materials that have completely (or nearly completely) filled band-structures, such as insulators and semi-conductors, the convergence of these calculations are quicker and can be done accurately using a relatively

\(^1\)Often \( k \)-points are referred to as \( k \)-points, however only the former is used in this thesis.
small number of $k$-points. This is achieved using the equation

$$
\epsilon_k = \sum_i N \frac{1}{\Omega_{BZ}} \int_{BZ} \epsilon_i, k \Theta(\epsilon_i, k - \epsilon_F) d^3 k,
$$

(3.103)

where $\Theta(\epsilon_i, k - \epsilon_F)$ is the Dirac step function associated with the Fermi energy $\epsilon_F$. In practice, for computational efficiency, this is actually calculated over a special finite set of $k$-points and is evaluated as

$$
\epsilon_k = \sum_i N_k \sum_k w_k \epsilon_i, k \Theta(\epsilon_i, k - \epsilon_F),
$$

(3.104)

where $N_k$ is the $N^{th}$ special $k$-point and $w_k$ is the weighting factor associated with $k$-point $k$. However, this equation converges progressively slowly as the number of $k$-points increases. Therefore, for metals, for which partial occupancies are used, the convergence is much slower, making the calculation very computationally expensive. This is because metals have delocalised electrons and unoccupied states near the Fermi level, the highest occupied state, meaning that they are likely to have overlapping bands and many partially-filled bands. Due to this, there is a necessity to better reproduce these states near the Fermi level, which have a non-trivial shape. This can be done by using more $k$-points. It is for this reason that the weight of the density at any $k$-point is usually approximated by a smoothing function, which in VASP is, by default, obtained by the method of Methfessel and Paxton [72], involving the expansion of the step function in a complete orthonormal set of functions. However, other methods do exist.

Lastly, since construction of $n_k(r)$ is done in reciprocal-space, as the real-space cell size increases, fewer $k$-points are needed to get an accurate description of the density. This is good as the fewer the number of $k$-points needed to describe the system, the quicker the calculations, because summations are required for each $k$-point, and each calculation is computationally quite expensive. However, caution is required, as a poor $k$-point sampling will lead to a poor description of the density, which in turn will lead to poor accuracy of the calculations.

### 3.8.4 Pseudo-Potentials

A further approximation is used in order to increase the accessibility of the solutions to the Kohn-Sham equations, this is the pseudo-potential (PP) approximation. This approximation is based on the frozen core approximation, in which the effect of the chemical environment of the atom only weakly affects the core region (region around the nucleus) and hence may be treated as a first-order perturbation of the interaction with the more weakly bound electrons. Following this logic, the electrons can be split into two groups: the strongly bound (and therefore, chemically inert) core electrons, and the more weakly bound valence electrons (usually involved in the chemical bonding between atoms).
The need for this to be done stems from the fact that the Coulomb potential term in the Hamiltonian, Eq. 3.52, which gives \( v_{ne} \propto \frac{Z_n}{|R_n - r_e|} \), has a discontinuity at \(|R_n - r_e| = 0\). The core region itself is strongly attractive, however, due to the orthogonality condition, the wave-functions describing the valence electrons oscillate rapidly in the core region. There is also a repulsive contribution from the localised core electrons. These oscillations effectively act as a repulsive potential which balance out the Coulombic attractive potential, \( v_{ne} \). Representing these oscillating orbitals and the strongly localised core electrons would require many basis sets. This is what the pseudo-potential, \( v_{pp} \), imitates, by introducing a smooth repulsive potential in the core region \[73\]. Therefore, at small distances, smaller that the cutoff distance, i.e. \(|R_n - r_e| \leq r_{cut}\), the potential acting on the electrons is replaced with a fitted potential \[1\]. These method, typically combined with the frozen core approximation effectively eliminates them (the core electrons) from the electronic calculations. However, outside of this cutoff \(|R_n - r_e| > r_{cut}\), \( v_{pp} \) accurately describes the real potential, but since overlap between the core and valence wave-functions is neglected, the accuracy generally improves as \(|R_n - r_e| \to \infty\). Thus allowing \( \hat{v}_{pp} \) to be written as \( \hat{v}_{pp} = \hat{v}_{ne} + \hat{v}_{sc} + \hat{v}_{val} \), where \( \hat{v}_{sc} \) is the screening effect due to the core electrons, and \( \hat{v}_{val} \) is the electronic-valence electron interaction potential. When used with Eq. 3.91, the corresponding eigenfunctions become pseudo-eigenfunctions, valid only in the region outside atomic cores. Elsewhere the wave-function is continuous, without nodes, which allows for much quicker basis-set convergence, necessarily reducing the computational cost. This means that a limitation on the accuracy of any calculation is the choice of pseudo-potential.

Since electrons in the core region play little to no role in chemical interactions, the physical and chemical properties studies by DFT are not greatly affected by their absence. Thus, the correct use of pseudo-potentials may not only greatly reduces the computational cost, but may also do so without greatly reducing its chemical accuracy. However, this depends on there being no significant modification to the core electrons within a calculation\[2\]. Three such pseudo-potentials are norm-conserving \[74\], ultra-soft \[75\] and the projector augmented-wave \[76, 77\] pseudo-potentials, abbreviated to NC-PP, US-PP and PAW-PP, respectively. Normally the pseudo-potentials act only on valence electrons, but PAW-PP acts on all. These method essentially differ by how they treat the wave-functions around the cutoff, whether they strictly (NC-PP) or more loosely (US-PP) conserve the norm of each pseudo-potential with respect to the all-electron wave-function. However NC-PP and US-PP do not allow for the restoration of the all-electron wave-function, while the PAW method does. It achieves this by decomposing the all-electron wave-function into three parts: one which is exact inside the core region, smoothly tending to zero outside (partial

---

1. This depends on the method implement, i.e. this is true for norm-conserving pseudo-potentials but not for projector augmented-wave pseudo-potentials.
2. An example, to name but one, where modification does occur is in x-ray absorption.
all-electron wave-function); one which is smooth every inside and exact outside the core region (partial pseudo-wave-function); and one consisting of a set of projector functions. While other pseudo-potentials exists, the SIESTA calculations described in this thesis employed NC-PP (used in this thesis for describing semiconductor materials) and VASP calculations employed US-PP and PAW-PP (used for describing metallic materials).

3.9 Modelling Surfaces

Figure 3.5: Representation of the three dimensional periodic boundary conditions as applied to the $p(2 \times 1)$ phase of the oxidised Cu(110) surface, with a sufficiently thick slab. Supercell is indicated by the thick red lines. This also illustrates the significant vacuum gap between the uppermost atoms of the bottom structure and the lowermost atoms of the top structure.

Although the theory described by the Bloch theorem in Section 3.8.3 so far has been for systems periodic in three dimensions, i.e. a crystal, for the majority of the thesis, surfaces are discussed. For this reason, it is very convenient, and accurate, to describe a surface as a semi-infinite crystal (infinite in two dimensions), a semi-periodic structure. Due to the robustness of the Bloch theorem, the implementation of periodic boundary conditions (PBC) is not solely limited to the study of periodic systems. In fact, non-periodic systems can be modelled as periodic systems with vacuums gaps (separations) between each repeating unit, large enough that the images (repeating units) can be said to be non-interacting. Such a unit is called the supercell
3.9 Modelling Surfaces

but will hence also be referred to as an image. In this manner, a system of interest can be modelled, such as isolated atoms, molecules, clusters, mono-layers or semi-infinite surfaces, as well as crystals. In the same fashion, single defects in crystals and at surfaces can also be modelled, by using a large enough unit cell that the defects in the images do not interact with each other. An illustration of a model of a surface is shown in Fig. 3.5, where the unit cell has been highlighted. In this figure, the cell can be seen to be periodic in the lateral co-ordinates, but has a vacuum gap separating the top layer of the unit cell from the bottom layer of the image above it.

If the images undesirably interact strongly, then the calculations will produce unphysical results, therefore, generally, the larger the vacuum, the more accurate the calculation. However, for a system represented by plane-waves, a large vacuum gap results in a finer grid in the reciprocal-space and hence more plane-waves are needed, and this necessity for more plane-wave makes the calculation more computationally expensive. When localised basis-sets are used, as the orbitals are centred on the atoms, there is no such problem.

Furthermore, when modelling a surface, even the number of layers required (thickness of the slab) can be optimised. Certainly, a slab with lateral periodicity and semi-infinite in the third direction would yield the most accurate results, however, this is often not possible in practice. Hence, a sufficiently thick slab is used, thick enough to accurately represent the bulk, surface and bulk-surface interface, but thin enough to yield the best possible speed for the calculation.

The convergence of a surface, made up of \( N_l \) layers, with respect to the number of layers used to construct it, can be determined using the idea that convergence is reached when the energy of adding another layer, is equivalent to the energy of the layer if it were in the bulk, \( E_l \). Another way of saying this, is that this convergence is reached when the change in energy \( (\Delta E) \) of the system upon adding another layer tends to zero, i.e. \( \Delta E \to 0 \) as \( N_l \to \infty \), where

\[
\Delta E = E(N_l) - N_l E_l - \left[ E(N_l + 1) - (N_l + 1) E_l \right]
\]

For single species slabs, where a layer can be modelled simply as a chain of that atomic species, the term \( E_l \) is simply the bulk energy (energy of that atomic species in the bulk).

However, to determine the number of layers to use in a slab on which atoms or molecules have been adsorbed (resulting in a reconstructions of the surface), \( \Delta E \) is given by

\[
\Delta E = E_X(N_l) - E_s(N_l) - \left[ E_X(N_l + 1) - E_s(N_l + 1) \right],
\]

where \( E_X \) is the energy of the surface with non-native species adsorbed and \( E_s \) is the energy of the native surface. \( \Delta E \) here is the change in reconstruction energy. Typically when studying surfaces with adsorbed atoms or molecules, both methods are used in determining the appropriate number
of layer. Doing so sometimes lead to very different results. In such cases the author uses the result from the latter method.

![Diagram of Region I: Frozen, Region II: Free to Relax, Region III: Free to Relax, Region IV: Frozen]

Figure 3.6: Schematic of the nano-asperity and sample surface. Regions are indicated where the atoms are free to relax or are frozen during geometry optimisation calculations.

Since the slab thickness affects the cell size, as well as the physical representability of the system, such tests are, of course, performed during the systems initial construction. Additionally, to reduce the number of the degrees of freedom, and more importantly, to accurately describe the bulk material for a surface, the bottom layers of the slab are typically “frozen/fixed” (they are not allowed to relax). The freezing of certain layers in the surface model is illustrated in Fig. 3.6, where atoms that are allowed to relax are those in the region where most interactions take place, the most reactive regions, which correspond to regions II and III in the image. Those that are frozen corresponds to the less reactive regions, I and IV, and are considered to only contribute negligibly to the interactions.

The way in which these regions are frozen is generally done by:

1. relaxing all atoms in the bulk material,

2. creating the slab (and tip if present) system from the relaxed bulk geometry (increase the vacuum gap in the vertical direction),

3. not allowing the lowest $N_f$ layers of atoms (and highest layers in tip if present) to relax for all subsequent calculations, starting from their relaxed bulk geometry.

By doing this and using a sufficient number of frozen layers, the bulk material (the semi-infinite) surface can realistically be modelled in a manner computationally cheaper than by using many layers. The more layers that are frozen, the cheaper the calculation, however, keeping more layers
frozen means that less layers are used to accurately represent the reactive part of the surface (the portion of the surface that is free to relax). Conversely, the fewer the number of frozen layers, the more accurate the surface is described as more atoms are free to relax into their global minimum positions, but the less accurate the semi-infinite bulk is described as less atoms will likely occupy their ideal bulk positions. So the number of frozen layers is also a parameter that can affect the chemical accuracy of the calculation and needs to be optimised for each system.

3.10 Modelling Microscopic Forces

Summarising what has been achieved so far, it has been determined that in order to relax a geometry (to find the minimum energy structure), the energy of the system needs to be minimised with respect to the atomic positions, i.e. the force acting on each atom should be zero. It has been shown in Eq. 3.7 that the forces acting on the atoms depend only on their potential environment, which is related to the total energy of the system. Therefore, by having access to the total energy, the force can be quantum mechanically determined.

So as to not forget why it is necessary to use these complex methods (such as DFT), the remainder of this chapter describes how the simplifications/solutions offered by these methods are utilised in order to perform theoretical calculations of NC-AFM experiments.

In short, for a given atomic arrangement, the electron density can be calculated, which is then used to calculate the sum of the nuclear and electronic energies, $E_n$ and $E_e$, respectively, i.e.

$$E = E_n + E_e.$$  \hfill(3.107)

Having knowledge of the energy of the system gives access to the force, $F_{\mu_i}$, acting on atom $i$, as defined by Eq. 3.7. Since these calculations are static with respect to the atoms, there is no contribution from $T_n$, thus the force is given as

$$F_{\mu_i} = -\frac{\partial E}{\partial R_i}.$$  \hfill(3.108)

Futhermore, by using the Hellmann-Feynman theorem \cite{78, 79}, this force can be calculated from Eq. 3.46 using just the electronic density ($n$) for a given nuclear geometry\footnote{Strictly speaking this is true for plane-wave method. For method employing non-local basis sets, like SIESTA, contributions arising from the need to differentiate of the basis set with respect to the atomic position also need to be considered, these are called Pulay forces.}, $(\mathbf{R})$. When the sum of all forces on atoms and the individual forces on all atoms allowed to relax is close enough to zero (or a precision defined by the force convergence flag in DFT packages), the system is considered to be relaxed and has reached its potential energy minimum/equilibrium geometry (hopefully not a local but a global minimum). This is so as to meet the requirement given in Eq. 3.31.
Calculations performed this way, minimising the energy for a given geometry, are referred to as static equilibrium calculations.

For modelling NC-AFM experiments it is crucial to know the tip forces, \( F_\mu \) (cf. Section 3.3), for this reason the definition of it is now given.

Since the force depends on the interaction potential, for a tip above a surface, the forces acting on the tip atoms depend on both the tip and surfaces atoms. As described in Section 3.9, typically some layers of atoms are frozen, this means that the sum of the forces acting on these atoms will be non-zero, as they are not allowed to relax into their equilibrium positions. This is illustrated in Fig. 3.6, where such layers correspond to regions I and IV.

The tip-surface interactions change only with respect to the relative tip-surface separation, \( z \), so naturally the force is determined from the derivative with respect to \( z \). Hence, for a microscopic system the microscopic tip force is given by

\[
F_\mu = - \left. \frac{\partial E(z_f, z_r)}{\partial z} \right|_{eq} = - \left[ \frac{\partial E}{\partial z_f} \frac{\partial z_f}{\partial z} + \frac{\partial E}{\partial z_r} \frac{\partial z_r}{\partial z} \right]_{eq},
\]

where frozen atoms are denoted by \( z_f \) and atoms allowed to relax by \( z_r \). It can easily be shown that the geometries of the frozen surface atoms are independent of \( z \), however, their interactions with the tip and consequential contributions to the tip force are accounted for. Furthermore, atoms that are allowed to relax within the tip will relax into their equilibrium positions; i.e. \( \frac{\partial E}{\partial z_r} \bigg|_{eq} = 0 \). Conversely, the frozen atoms linearly depend on \( z \), leading to the results that \( \frac{\partial z_f}{\partial z} = 1 \). Thus the microscopic forces acting on the tip can be simplified to

\[
F_\mu = - \left. \frac{\partial E}{\partial z_f} \right|_{eq},
\]

which is simply the sum of the vertical forces acting on only the frozen atoms within the tip. For a description of the tip forces for when the system is not in its potential minimum, the author refers the reader to the the paper [80] by and PhD thesis [37] of Thomas Trevethan.

### 3.11 Additional Contributions to the Energy

#### 3.11.1 Lattice Vibrations - Phonons

The DFT calculated energy is not exactly the free energy of the system, as \( E^{DFT} \) does not contain the contributions to the energy due to the vibrational modes of the crystal, the origin of which will become clear in the following discussion. A crystal consists of atoms vibrating about their lattice sites, for this reason, a crystalline system is well described as a periodic arrangement of vibrating atoms. For an optimised crystal structure, with geometry \( \mathbf{R} \), the energy is minimised,
3.11 Additional Contributions to the Energy

and can then be expanded to the second order about $R$, ignoring higher order terms. Each element in the second order term contains

$$\Phi_{I,I'} = \frac{\partial^2 E(R)}{\partial R_I \partial R_{I'}},$$

(3.111)

where $R_I$ and $R_{I'}$ are the positions of atom $I$ and $I'$, respectively. These elements form a matrix $\Phi$, the *Force constant matrix* (also called the *Hessian*). Often, these terms are instead contained within $D$, the Dynamical matrix, where

$$D_{I,I'} = \frac{1}{\sqrt{M_I M_{I'}}} \Phi_{I,I'},$$

(3.112)

and $M_I$ and $M_{I'}$ are the masses of atoms $I$ and $I'$, respectively. This is the harmonic approximation, within which small displacements cause oscillations about $R$. Thus, to a good approximation, the system can be modelled as a crystal lattice of harmonic oscillators, the quantum solution of which is known to give a non-zero energy related to the vibration of these atoms,

$$E_{\text{vib}} = \sum_{k,j} \hbar \omega_{k,j} (n_{k,j} + \frac{1}{2}),$$

(3.113)

where $k_j$ is the wave-vector on the $j$th-band of the vibration with frequency $\omega_{k,j}$ for the state $n_{k,j}$. From this equation it can be seen that states are equally separated by $\hbar \omega_{k,j}$, where the first-state/ground-state energy, $E_{\text{vib}}^0$, is given by

$$E_{\text{vib}}^0 = \frac{1}{2} \sum_{k,j} \hbar \omega_{k,j},$$

(3.114)

known as the *zero-point energy*. The dependence on different $k$ is the result from atomic displacement eigenvectors for a periodic crystal obeying the Bloch theorem (see Section 3.8.3). Hence, the frequencies themselves are found to be related to $k$ due to the relation

$$\det \left| D(k) - \omega_{k,j}^2 \right| = 0,$$

(3.115)

This relation between $\omega_{k,j}$ and $k$ is known as the dispersion relation.

For a crystal lattice this vibrational contribution to the energy depends on collective vibrations, i.e. it relates to groups of atomic vibrations, known as *phonons*. Furthermore, if the crystal consists of more than one atomic species or differently spaced atoms then there will be two types of phonons, acoustic and optical. Acoustic phonons refer to coherent vibrations (vibrations in the same direction), while optical phonons refer to vibrations of atoms in different directions, and each phonon profile is known as a branch.

Since the average of all atomic vibrations about the equilibrium geometry is zero by symmetry within the harmonic approximation [42], the approximation is applied to the free energy
3.11 Additional Contributions to the Energy

Instead of the potential energy, which effectively takes into account some anharmonic effects, therefore, the Hamiltonian of the relaxed system of energy \( E_0 \), with atoms in their equilibrium positions, \( \mathbf{R}_0 \), is given by

\[
\hat{H} = \hat{H}_0(\mathbf{R}_0) + \sum_{k,j} \hbar \omega_{kj} \langle \hat{n}_{kj} + \frac{1}{2} \rangle,
\]

where \( j \) refers to the phonon branch, and the summation is performed over \( N_k \) wave-vectors and \( N_j \) branches. The operator \( \hat{n} \), which changes the state of the system, is the number operator\(^{1}\). This approximation is known as the quasi-harmonic approximation, which thus allows the total energy, \( E \), of the system to be given by \( E = E_{\text{DFT}} + E_{\text{vib}} \), or explicitly by

\[
E = E_{\text{DFT}}(\mathbf{R}_0) + \sum_{k,j} \left[ \frac{1}{2} \hbar \omega_{kj} + \frac{1}{\beta} \ln(1 - e^{-\beta \hbar \omega_{kj}}) \right],
\]

where the second term is the zero-point energy, \( \beta \) is the inverse temperature, equal to \( \frac{1}{k_B T} \). The last two terms in this equation make up \( E_{\text{vib}} \).

For more information on phonons and the quasi-harmonic approximation, outside of the discussion provided in this thesis, the author refers the reader to Chapter 4 of the book by Lev Kantorovich \[42\].

Throughout this thesis, contributions to the energy due to phonons were calculated, using TETR \[81\], a code written by Lev Kantorovich. The calculations were performed employing the frozen phonon method, which builds up the force constant matrix by displacing atoms and calculating the forces on those displaced atoms.

3.11.2 Gibbs Free Energy (GFE)

As already mentioned, the aim of the Schrödinger equation, all the approximations made and methods developed to solve it, is to calculate the total energy of the system. However, the energy calculated by Eq. 3.107, refers to the a quantity known as the internal energy of the system, and is made up of the kinetic and potential energies of the particles in a system. This energy ignores changes due to temperature, pressure and volume. In reality, such changes are in fact important. For instance, in systems consisting of a gas-surface interface, a contribution to the energy must be included to account for volume changes, \( V \), in the system, when gas is released or absorbed by the surface. For a system at constant pressure, \( P \), the term \( PV \) accounts for the work done on the system to create volume. Additionally, a particular state of a system at temperature, \( T \), can likely be achieved in several ways. As an example, for a single atomic

\(^{1}\)The number operator \( \hat{n} \) discussed here is not to be confused with the density operator introduced in Section 3.7.2.
species system, the ordering of atoms \( \langle R_1, R_2 \rangle \) may equally be \( \langle R_2, R_1 \rangle \). The measure of this multiplicity is known as the entropy, \( S \), of the system, from the second law of thermodynamics, which describes how nature tends towards disorder in a system. For this reason, a system at temperature \( T \) accumulates entropy. Furthermore, as the system does this and there is a more even distribution of the energy of the system, that energy becomes less able to do work, in order words, \( TS \) energy is removed from the system.

Now using this information, the free energy of the system at pressure \( P \) and temperature \( T \), can be defined as

\[
G = U - TS + PV,
\]

and the quantity of energy \( G \) is called the **Gibbs free energy** (GFE) of the system. In this equation, \( U \) is the internal energy of the system, given by the DFT calculated energy of the system, \( E_{DFT} \), i.e. \( U = E_{DFT} \).

In application now to surfaces, an expression for the Gibbs free energy difference\(^1\), \( \Delta G \), must now be given. For a metal surface of chemical species \( X \), that has adsorbed or released atoms \( y_i \), from diatomic gases in environment around the surface, where \( i \) denotes the gas, \( \Delta G \) has been derived by Karsten Reuter, Catherine Stampfl, and Matthias Scheffler [82, 83, 84, 85]. It has been shown that for a metal surface in thermodynamic equilibrium with a gaseous environment, the chemical potentials of the surface metal atoms can be given by the bulk metal chemical potential. This can then be used in the expression for \( \Delta G \), which is given as the difference in the GFE of unadsorbed surface (known also as the bare/clean surface), and the adsorbed surfaces, \( G_{surf}^X \) and \( G_{surf}^{y_i} \), respectively. Terms are also included to account for the cost of losing or gaining \( N \) atoms in the transition between the two phases. Thus, \( \Delta G \) is written as

\[
\Delta G = G_{surf}^{y_i} - G_{surf}^X - \Delta N_X \mu_X - \sum_i \Delta N_{y_i} \mu_{y_i},
\]

where \( \Delta N_X \) and \( \Delta N_{y_i} \) are the differences in the number of metal and gas atoms, respectively, between the bare and adsorbed surface, with chemical potentials \( \mu_X \) and \( \mu_{y_i} \), respectively.

Simplifying this to a single gaseous species, as is studied in this thesis, the temperature \( T \) and pressure \( P \) dependence of the chemical potential of the gas is given by

\[
\mu_y(T, P_Y) = \frac{1}{2} \left[ E_{DFT}^Y + \tilde{\mu}_y(T, P_0^Y) + k_B T \ln \left( \frac{P_Y}{P_0^Y} \right) \right],
\]

where \( E_Y \) is the total DFT calculated energy of the diatomic molecular gas, \( Y \). \( \tilde{\mu}_y \) consists of the rotational and vibrational contributions to the energy of \( Y \) in addition to its ideal gas entropy using the reference pressure for \( Y \), \( P_0^Y = 1 \) atm. The \( 1/2 \) prefactor is because it is only single

\(^1\)The Gibbs free energy difference is often referred to simply as the Gibbs free energy in this thesis.
atoms from the diatomic gas molecule that adsorb at a specific site, or are released from the surface, not the molecule as a whole. \( \mu \) can in fact be quite easily calculated from thermodynamic tables [82, 86] through the equation

\[
\tilde{\mu}(T, P_Y) = \frac{1}{2} \left[ H(T, p_Y^0) - H(0K, p_Y^0) - T[S(T, p_Y^0) - S(0K, p_Y^0)] \right],
\]

(3.121)

where \( H \) and \( S \) are the enthalpy and entropy of the system, respectively, calculated both for the given temperature and \( T = 0K \). \( \mu_x \) is taken as the bulk energy.

### 3.12 Nudged Elastic Band (NEB) Method

For information about the kinetics and energetics of some of the systems covered in the thesis, information about the lowest energy pathway from some initial stable state to a final stable state is required. Physically, and naturally of course, this corresponds to the minimum energy pathway (MEP) or energy profile along the global potential energy surface (PES). The MEP is usually defined using a reaction co-ordinate, this is a co-ordinate that describes the state/progress of the reaction\(^1\). The \textit{Nudged Elastic Band} (NEB) method [87] is one such commonly used method for determining the MEP, and therefore energy barrier, of a state transition, be it a chemical reaction or manipulation. The energy barrier, important for estimating transition rates within the harmonic transition state theory (hTST), is defined as the difference in energy between the initial state energy and the maximum along the MEP (known as the saddle point). For more information on hTST, which is important for studying the kinetics of a system, the author refers the reader to the paper by Gregory Mills [88] and references therein.

#### 3.12.1 NEB Theory

NEB is based on the chain-of-states method, in which a chain of intermediate states (replicas/images) is generated between the fixed end points, the initial and final states. If the initial image is given by \( R_0 \) and the final by \( R_N \), then there are \( N - 1 \) intermediate images. The intermediate images are constructed such that they describe the transition from \( R_0 \) to \( R_N \), an example of which is simply a linear interpolation between the initial and final configurations.

#### 3.12.1.1 Plain Elastic Band (PEB)

In the chain-of-state method there are \( N + 1 \) images, initial and final states included \((R_0 \cdots R_N)\). These, collectively referred to as a band, are connected to their neighbours by \( N \)

\(^1\text{This is typically the position(s) of an atom or a group of atoms.}\)
springs, with spring constant \( k \), and are used to map out the MEP by constructing the object function \[ S(R_1, \ldots, R_N) = \sum_{i=1}^{N-1} E(R_i) + \sum_{i=1}^{N} \frac{k}{2} (R_i - R_{i-1})^2, \] (3.122)
and minimising it with respect to the images \( R_1, \ldots, R_N \). The inclusion of the springs introduces the second term, the spring force term, \( F^{spr} \), meaning that the method maps the chain-of-states (the band) onto an elastic band stretched between two fixed ends points. This method, as described by Eq. 3.122, is known as the plain elastic band (PEB) method.

Although it is a very powerful concept, the PEB method suffers from two important problems:

- the elastic band tends to cut corners, deviating from the MEP,
- the images tend to slide down towards the end points.

The first of these problems is due to the stiffness of spring forces acting within the images in the region where the the MEP is curved. In short, this is due to the presence of the *spring force perpendicular to the band*. The second is due to the atomic forces (henceforth referred to as real forces) acting within the images (e.g. the forces on atoms calculated by DFT given, by Eq. 3.108), \( F^{real} \), acting along the band which results in non-equidistant spacings between the images. In short, the second problem is due to the presence of the *real force parallel to the band*, and due to this the region of the saddle point has the lowest resolution.

### 3.12.1.2 NEB

The problems highlighted in the PEB method can be resolved by using a band force \( F^{NEB} \) [87, 89, 90], made up of the correct projections of \( F^{real} \) and \( F^{spr} \), where

\[
F_i^{NEB} = F_i^{real,|\bot|} + F_i^{spr,|\parallel|},
\] (3.123)

In this reformulation\(^1\) only the perpendicular components of the real force (acting within the images) and the parallel components of the spring force (acting along the band, between images) are kept. This is given, respectively, by

\[
F_i^{real,|\bot|} = F_i^{real} - F_i^{real} \cdot \hat{\tau}_i \hat{\tau}_i
\] (3.124)

and

\[
F_i^{spr,|\parallel|} = k(|R_{i+1} - R_i| - |R_i - R_{i-1}|) \hat{\tau}_i,
\] (3.125)

\(^1\)The improved tangent method has been employed whenever the NEB method [90] has been used through this research.
where $\hat{\tau}_i$ is the normalised tangent (parallel vector), $\hat{\tau}_i = \frac{\tau_i}{|\tau_i|}$, to the path of the band at image $i$, determined by only the energy of image $i$ and its immediate neighbours. It is given by

$$\tau_i = \begin{cases} \tau_i^+ & \text{if } E_{i+1} > E_i > E_{i-1}, \\ \tau_i^- & \text{if } E_{i+1} < E_i < E_{i-1}, \end{cases}$$

(3.126)

where $E_i = E(R_i)$ is the energy (in this case DFT calculated energy) of image $i$, and

$$\tau_i^+ = R_{i+1} - R_i, \quad \text{and} \quad \tau_i^- = R_i - R_{i-1}.$$

(3.127)

This is a good approximation when a sufficiently large number of images are used in the calculation. However, in order to allow for a smooth transition to and from the saddle point (to avoid convergence issues), in the event that its neighbouring images are either both higher or both lower in energy than the saddle point, the tangent is taken as a weighted average of the two vector tangents of the neighbouring images. In this case, if image $i$ is at a maximum ($E^\text{max}_i$) or minimum ($E^\text{min}_i$), i.e. $E_{i+1} > E_i < E_{i-1}$ or $E_{i+1} < E_i > E_{i-1}$, respectively, the tangent used is then:

$$\tau_i = \begin{cases} \tau_i^+ \Delta E_{i}^{\text{max}} + \tau_i^- \Delta E_{i}^{\text{min}} & \text{if } E_{i+1} > E_{i-1}, \\ \tau_i^+ \Delta E_{i}^{\text{min}} + \tau_i^- \Delta E_{i}^{\text{max}} & \text{if } E_{i+1} < E_{i-1}, \end{cases}$$

(3.128)

where

$$\Delta E_i^{\text{max}} = \max(|E_{i+1} - E_i|, |E_{i-1} - E_i|),$$

(3.129)

and

$$\Delta E_i^{\text{min}} = \min(|E_{i+1} - E_i|, |E_{i-1} - E_i|).$$

(3.130)

Now, as long as a sufficient number of intermediate images are used for the calculation, this method allows for the band to converge towards the MEP with respect to the force, i.e. $|\mathbf{F}| < F_{\text{max}}$, where $F_{\text{max}}$ is a specified maximum force tolerance. It is due to the introduction of these projected force terms that the intermediate images are adjusted (nudged) towards the MEP, hence the name nudged elastic band method.

The resulting MEP following a single reaction co-ordinate is then represented as illustrated in Fig. 3.7. In this example, for the system to evolve from its initial state to the intermediate state, it must overcome an energy barrier of $\Delta E_1$. However, to go from the intermediate state back to its initial state the system only needs to overcome an energy barrier of $\Delta E_2$, which is smaller than $\Delta E_1$. Alternatively, to evolve from the intermediate state to its final state, it must overcome an energy barrier of $\Delta E_3$. If either $\Delta E_2$ or $\Delta E_3$ is very small (or the same order as atomic vibrations), then the intermediate state will be stable only for a short time, i.e. the system will quickly evolve to either the initial or final state, depending on which barrier is smaller, and in this case the intermediate state is known as meta-stable. For the system to go from the final state
to the intermediate state it must overcome an energy barrier of $\Delta E_4$. It should be clear from this
description, that by calculating a single band (MEP) for a reaction/process, the profiles for both
the forward and reverse reactions/process will be known.

![Energy Profile Diagram]

**Figure 3.7:** Schematic illustrating the key features of the minimum energy path as described by the
nudged elastic band method.

### 3.12.2 Implementation

For an efficient NEB calculation it is of added benefit to not only use a sufficient number
of intermediate images, but to also ensure that an odd number of images is used. An even number
of images would likely interpolate across the saddle point, however by having an odd number of
images (in addition to an appropriate spring constant), it will be more likely that an image either
exactly be on or very close to the true MEP saddle point, which in turn ensures that the saddle
point will have the maximum resolution, improving the accuracy of the calculated energy barrier.

In order to use NEB, the energies and forces (which are simply the energy gradients)
must be calculated at each image, and these calculations are performed using DFT in the manner
already outlined in this chapter. For most of the NEB calculations performed with VASP,
the standard built-in implementation of NEB was used. However, for some of the calculations
performed with VASP, and for all calculations performed with SIESTA, which does not have an
built-in NEB routine, an NEB code written by the author was instead used. The code, called
BEINEB, standing for **B**arrier **E**stimation **I**mplementing the **N**udged **E**lastic **B**and method, was
designed to minimise a user-specified band. It is able to use any DFT code as a black box, re-
quiring only the energies and forces acting on atoms for images specified by BEINEB, which it
calculates via a minimisation algorithm. At the time of writing this thesis, BEINEB has been implemented for VASP and SIESTA, and the NEB calculations were performed using the Velocity Verlet minimisation algorithm (cf. Section 3.13.1.1) which was found to perform sufficiently well.

Additionally, as an initial band guess (the initial input band to be minimised), where possible, linear interpolations were used, calculated by IISAI, another code written by the author, standing for Intermediate Image Selector and Interpolator. For NEB calculations performed with BEINEB, the MEP is plotted as a function of image number, as the images, after a band relaxation by BEINEB, are approximately equidistant if \( k \) is the same for all springs [90]. The same may be true of the calculations performed with the built-in VASP algorithm, but to be safe, for VASP calculations, the MEP is plotted as a function of an atomic co-ordinate that suitably describes the transition.

3.13 Unconstrained Global Minimisation Methods

In order to find stable nuclear configurations, the global (or local) minimum(s) of the PES must be found, where the corresponding configuration(s) is/are the stable structure(s). Since, typically the starting geometry is a non-stable geometry, minimisation calculations need to be performed. This is also the case in NEB calculations, where the band is required to be minimised. Methods that search for the minimum are also known as, and commonly referred to, as relaxation or optimisation methods. If information about the final geometry of the system is available, then that information can be fed into the minimisation method, and such methods are called constrained minimisations. In the event that no information is available, then an unconstrained minimisation is required. Furthermore, some of these methods can be used as a way in which to include time in the calculations, for studies of how systems evolve in time, such methods are known as molecular dynamics (MD) methods and are based on Newtonian dynamics. As some of these methods are used for this research, a brief description of them will now be given in this section.

3.13.1 Molecular Dynamics

Based on Newton’s equations of motion, these methods integrate Eq. 3.7, using the current potential environment that the atoms are acting within at time \( t \). From this, the velocities of atoms are derived and hence, their future positions, at time \( t + \Delta t \), can also be derived from Eq. 3.7.
3.13 Unconstrained Global Minimisation Methods

3.13.1.1 Velocity Verlet Method

A popular molecular dynamics method, and one used in this research, is called Velocity Verlet\(^1\). In this method, the accelerations of the atoms, \(a\), are obtained from the derivative of the energy \(E\), by substituting \(\text{Eq. 3.108}\) into \(\text{Eq. 3.7}\). The velocities, \(v\), and positions, \(R\), are then calculated by the following equations in order:

1. \(R(t + \Delta t) = R(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2\).
2. \(a(t + \Delta t) = -\frac{1}{M_n} \frac{\partial E[R(t+\Delta t)]}{\partial R(t+\Delta t)}\), where \(M_n\) is the nuclear mass.
3. \(v(t + \frac{\Delta t}{2}) = v(t) + \frac{1}{2}a(t)\Delta t\).
4. \(v(t + \Delta t) = v(t + \frac{\Delta t}{2}) + \frac{1}{2}a(t + \Delta t)\Delta t\).

This particular implementation is known as the half-step algorithm, and allows \(v(t + \Delta t)\) to be calculated from \(a(t + \Delta t)\) via this ‘half-time step’. However, this assumes that \(a(t + \Delta t)\) depends only on \(R(t + \Delta t)\) and not \(v(t + \Delta t)\). It is worth noting that is it very important to choose the time-step, \(\Delta t\), wisely, typically an order of magnitude less than the characteristic time of phonon vibrations (\(\approx 10\) fs). Choosing a value too large will lead to integration errors, and too small a value will likely cause rounding errors, and could render the calculation too expensive.

When implemented in BEINEB, only the component of the velocity that is parallel to the force, and along the same direction, is kept\(^2\). This allows for performing efficient minimisation of the energy by means of MD type simulations. Thus, the velocities used, \(v^{used}\) are

\[
v^{used} = \begin{cases} (v \cdot F)\hat{F} & \text{if } v \cdot \hat{F} > 0, \\ 0 & \text{else.} \end{cases} \tag{3.131}
\]

Using this definition of the velocities, if the system goes beyond the minimum, the velocities are immediately reduced to zero. Conversely, the velocities will increase if the force continually acts in the same direction.

3.13.2 Other Minimisation Methods

Instead of the Velocity Verlet method, a quasi-Newton method is used for the NEB calculations performed with the built-in VASP implementation [68]. In this method an approximation to the second order derivative (analogous to \(a\) in the Velocity Verlet method) is calculated from the inverse Hessian matrix, which is built up from the derivatives of the force, the second order derivatives of the potential.

\(^1\)A variation of Velocity Verlet algorithm (cf. \text{Eq. 3.131}) is used in the NEB calculations performed with SIESTA (cf. \text{Section 8}).

\(^2\)By doing this however, the prerequisite for an accurate calculation is to have a good initial approximation of the reaction path.
3.14 Comparing Theory to Experiments

There also exists completely non-Newtonian minimisation methods, and one such popular method is the *Conjugate Gradient method*. This method searches for the vector (of atomic position) for which the force is minimised, where the search direction is determined as the conjugate to the force.

The Velocity Verlet method was used in BEINEB due to it being quick to implementation and producing sufficiently quick and accurate results.

3.14 Comparing Theory to Experiments

Now that the framework has been laid down for modelling the dynamics of the NC-AFM, inclusive of the tip-surface interactions, a brief description is now given of how to directly compare experimental results to theoretical predictions, namely through \( \Delta f \), \( F \) (force) and dissipation vs. \( z \) plots. The author points out here that for the remainder of the thesis \( z \) refers only to the average tip-surface separation.

3.14.1 Frequency Shift Curves

![Figure 3.8](image)

*Figure 3.8:* (a) Schematic illustrating the origin of frequency shifts, where the *black* and *red* curves represent the \( A-f \) plot for the non-interacting and interacting tip, respectively. (b) An illustration of how the topography using a user defined \( \Delta f \) in constant \( \Delta f \) FM-AFM is obtained, where the different curves correspond to different lateral tip positions.

As mentioned, the tip is oscillated at the resonant frequency, where its amplitude is at its maximum. However, when the tip interacts with the surface, the position of the peak amplitude shifts, as illustrated in *Fig. 3.8*(a). This relative shift is known as the frequency shift, \( \Delta f \), and is the quantity that is measured experimentally (one of the jobs of the PLL shown in *Fig. 2.4*). By using *Eq. 3.13* and an appropriate scaling factor, \( \Omega \), \( \Delta f \) is typically given in terms of \( \Omega \), where \( \Omega = \omega/\omega_0 \), giving that

\[
\Delta f = 2\pi\omega_0(\Delta\Omega - 1).
\] (3.132)
Furthermore, due to microscopic forces, $\Delta f$ changes depending on the lateral position of the tip above the surface, i.e. depending on which atom the tip interacts most with. Hence, in constant frequency shift and amplitude FM-AFM (called constant frequency shift FM-AFM), the topography of the surface is determined from the changes in the average tip-surface separations ($\Delta z_0$) which corresponds to maintaining the preset $\Delta f$ for different lateral positions. An illustration of this is given in Fig. 3.8(b), where $\Delta z_0 = z_2 - z_1$. Similarly, constant height and amplitude FM-AFM (called constant height FM-AFM), which maintains a preset $\Delta z_0$, gives the $\Delta f$ map of the surface. Conversely, in AM-AFM, topography is instead determined using a preset $\Delta A$ (Fig. 3.8(a)).

![Diagram of tip model](image)

**Figure 3.9:** Schematic of the tip model used to calculate theoretical spectroscopy ($\Delta f-z$) curves, consisting of a macroscopic sphere terminated with a microscopic (atomistic) model. The portion of the microscopic tip that sits inside the macroscopic sphere is defined by the offset.

Experimentally, $\Delta f-z$ is measured directly, however, theoretically, it must be modelled via Eq. 3.13 and Eq. 3.132. This equation depends on the tip-surface interactions, namely $F_{ts}$, which can now be calculated as all the components of Eq. 3.14 can be determined from the macroscopic tip model (via Eq. 3.23) and DFT (via Eq. 3.110). To do this in our studies, we used a spherical tip model, as shown in Fig. 3.9, where a large sphere was assumed for the macroscopic portion of the tip, terminated by the microscopic tip considered atomistically in our DFT calculations.
The overall force acting on the whole tip is evaluated from a combination of the macroscopic vDW forces due to the large spherical part and the microscopic forces on the atomistic part calculated by DFT.

### 3.14.2 Force Curves

Whilst the $\Delta f$ curves give information about the topography of the surface, they do not directly give information about the inter-atomic processes occurring due to the tip-surface interactions. It is therefore necessary to be able to plot a force-distance curve, $F$-$z$. Although it is possible to plot the full $F$-$z$ curves both theoretically and experimentally, for direct measurement of the short-ranged/atomic forces, it is necessary to extract the long-ranged forces from the experimental data [91, 92]. This is also necessary for better comparison against theory, as it eliminates the influence the macroscopic spherical tip model has on results. Theoretically, these full $F$-$z$ curves can easily be calculated directly from the macroscopic tip model (see Eq. 3.23) and from the microscopic forces predicted by DFT calculations via Eq. 3.110, using only the latter for plots of the short-ranged forces. Typically, however, only the short-range $F$-$z$ curves are plotted. In all instances throughout this thesis, the $F$-$z$ curves were obtained for the lowest atom of the tip apex initially separated from some pre-defined reference co-ordinate of the surface before relaxation by a distance $z_i$, and all subsequent values of $z$ correspond to the decrease or increase in tip-sample separation prior to relaxation. Experimentally this is much more challenging, and must be extracted from the frequency shifts, typically done using the Sader-Jarvis inversion algorithm [93]. This algorithm gives the full $F$-$z$ curve while the short-ranged forces can be determined by measuring $\Delta f$-$z$ spectra at several different lateral positions and extracting the common long-range part of the curves, or in pseudo-constant height mode on different surface terraces close to an atomic step edge as is done in [19].

### 3.14.3 Dissipation Curves

Damping in NC-AFM corresponds to energy dissipation in the cantilever-sample system due to non-conservative tip-surface relaxation processes\(^1\). This phenomenon has been most effectively explained by the adhesion hysteresis mechanism, whereby atomic relaxations in the junction (i.e. both in the tip and sample) on approach and retraction differ [94, 95, 96, 97, 98, 99], leading to the tip and surface atoms following two different trajectories and hence to the tip force hysteresis (two different $F$-$z$ curves). Although the debate about the mechanism is still somewhat open [26, 96, 97, 100, 101]. Using this mechanism, this energy, $E$, per oscillation may indeed be

\(^1\)This could otherwise be interpreted as the amount of additional energy required for the oscillating cantilever to maintain a constant amplitude at resonance, where changes are due to the tip-surface interactions.
calculated [102], and is given as

\[ E = \oint F(z)dz = \int_{z_{\text{min}}}^{z_{\text{max}}} (F_\uparrow - F_\downarrow) \, dz, \quad (3.133) \]

where \( F_\downarrow \) and \( F_\uparrow \) are the tip-surface force upon tip approach to and retraction from the surface, respectively. Since it is the short-ranged forces that are predominantly behind the changes in \( F_{\text{ts}} \), this energy can be extracted theoretically directly from DFT calculated forces. However, although this dissipation energy may also be measured experimentally, it is averaged over several hundred tip oscillation cycles and direct comparison with theory per oscillation is generally not possible. That being said, there is still a benefit in comparing the two, as the general trends they show should not differ.

3.15 Summary

In this chapter, all of the relevant methodologies and theories have been described, along with their relevance, applications and limitations. During the preparatory calculations for each system, the issues of the limitations were addressed; the real-space energy cutoff (affecting \( E_{\text{cut}} \), the number of plane-wave/localised basis-sets), \( k \)-point sampling, exchange-correlation functional, vacuum gap, system size and macroscopic-microscopic tip offset were all converged in order to achieve the most accurate results, as quickly as possible. Hence, all results presented in this thesis were obtained from calculations performed using converged parameters and the convergence of the parameters are not discussed further, with the exception of Chapter 4. Furthermore, it is also been shown explicitly how to compare theoretical predictions to experimental measurements.

In order to model experiments and make theoretical predictions, this author has developed various codes, namely to model transition paths and to calculation \( \Delta f-z \) and dissipation-\( z \) curves.
Part III

Results
Chapter 4

Thermodynamic Stabilities of Cu(110):O Surfaces

4.1 Motivation

In an attempt to characterise/identify and control the atomic termination of an NC-AFM tip (the species terminating the tip asperity), the oxidised Cu(110) insulator surface has been chosen. This surface, henceforth referred to as Cu(110):O, was chosen since in order to manipulate the terminating species of the tip apex, there must be more than one atomic species present and the Cu(110):O surface is a very stable example of such a surface. If only tip enforcement was important then by simply choosing a monoatomic sample and crashing the tip into the sample, the termination could be identified with near certainty. However, what this thesis researches is a method for simultaneously identifying both the tip and the surface, and for controllably enforcing specific tip terminations, which can then be used for controlled manipulations. Cu(110):O is ideal for this since it is an example of the simplest stable non-monoatomic surface, a binary oxide surface.

Although this was the main aim for this research on Cu(110):O, before any NC-AFM study of the surface can be performed, first the surface geometries must be described and identified, and the correct model for the surface constructed, which is the focus of this chapter.

In addition to the specific motivation behind this study, copper surfaces are important in their own right as industrial catalysts [103, 104, 105] for example, and surface oxidation is believed to play crucial role [106] in their catalytic activity.
4.2 The Cu(110):O Surface

4.2.1 Introduction

When oxygen adsorbs onto the clean Cu(110) surface, at high oxygen exposures (2000 L), it forms a semiconductor/metal interface [107], consisting of \( p(2 \times 1) \) and \( c(6 \times 2) \) patches. At low oxygen exposures it forms a \( p(2 \times 1) \) phase, while at higher exposures a \( AS-c(6 \times 2) \) phase [108]. The \( p(2 \times 1) \) phase, shown in Fig. 4.1(a), has been studied extensively both experimentally and theoretically (see Ref. [109] and references therein), and its “added row” (AR) structure is generally accepted, where AR refers to its alternating row/missing row conformation [108, 110] in the (110) direction, and where alternating Cu and O atoms lie along the (001) direction. The \( p(2 \times 1) \) phase, shown in Fig. 4.1(a), has been studied considerably experimentally, where its structure has been confirmed by STM [111], NC-AFM [108] and LEED [112], but much less theoretically [109]. It has a “double added row” structure, different from the single AR structure of the \( p(2 \times 1) \) phase. Another key feature of the AS-\( c(6 \times 2) \) surface is the presence of additional copper atoms in every second row, known as “super” copper atoms [108, 111], which are bound between two oxygen atoms, referred to as the “added strand” (AS) [109]). As a result, the oxygen atoms bonded to them are raised slightly out of plane and are buckled in towards the super Cu atoms, acquiring the names “high” [108] or “buckled” [111] oxygen atoms. These high oxygen atoms are bonded to the super Cu atoms to form “super” CuO\(_2\) molecules (or units) on the surface. The AS-\( c(6 \times 2) \) structure represents tiling the surface with Cu\(_5\)O\(_4\) templates resulting in a AS-\( c(6 \times 2) \) arrangement of the “super” CuO\(_2\) units. Other variations of this building principle are depicted in Fig. 4.1. For example, the AS-(6 \( \times \) 2)-(1 structure, Fig. 4.1(e), uses tiling with Cu\(_9\)O\(_8\) units resulting in a AS-(6 \( \times \) 2) arrangement with a single super CuO\(_2\) unit. Note that AS/AR notation is used to distinguish between similar reconstructions.

4.2.2 Previous Studies

The \( p(2 \times 1) \) phase is the well known ground state of the Cu(110):O surface at lower oxygen exposures, and appears to be well described by standard DFT methods. At the same time, however, although the AS-\( c(6 \times 2) \) reconstruction has been shown experimentally to be the ground state at higher oxygen exposures, previous theoretical thermodynamic calculations [109] found this reconstruction as thermodynamically favourable from several other geometrically similar surface reconstructions by only a marginal amount (\( \approx 25 \) meV per surface unit cell). If these DFT calculations [109] were indeed correct, then one would expect to observe these other reconstructions alongside the AS-\( c(6 \times 2) \) phase, but only the latter has been experimentally reported at higher oxygen exposures, before the bulk Cu\(_2\)O oxide forms.
4.3 General Calculation Details

All calculations discussed in this chapter, unless otherwise stated, were performed with VASP, and PAW-PP using 11 and 6 valence electrons on every copper and oxygen atom, respectively. Electronic exchange and correlation were initially described using the PBE. For calculations of the bulk Cu, a face centred crystal (fcc), with a space group number 225, \textit{Fm}\text{-}\bar{3}\textit{m}, a similar set of parameters as in S. Y. Liem \textit{et al.} [113], was used. The Brillouin zone was sampled using the Monkhorst-Pack method with a \((30 \times 30 \times 30)\) k-point set, giving a dense k-point grid of \(\approx 0.03\) Å\(^{-1}\) in each direction. Partial occupancies were handled employing the broadening approach suggested by Methfessel and Paxton with a width of 0.2 eV. A plane-wave cutoff of 1000 eV was used in the bulk calculations and for all subsequent calculations, a plane wave cut-off of 500 eV was used and the k-point sampling was scaled from that of the bulk. Typically all structures were fully relaxed with a threshold of \(10^{-6}\) eV for the energy convergence and \(10^{-2}\) eV/Å for forces.

Calculations for the bulk using these parameter gave a lattice constant, \(a_0\), and cohesive energy, \(E_{coh}\), of 3.64 Å and 3.50 eV, respectively, shown in Table 4.1. These are typical values for GGA calculations [114], and is in good agreement with experimental values of 3.61 Å and 3.49 eV [115, 116], respectively, also comparing remarkably well against previous DFT studies employing the GGA approximation [62, 109, 113].

When performing calculations to determining the optimal number of layers to use, oscillation...
### 4.3 General Calculation Details

Table 4.1: Calculated Cu bulk lattice constant \(a_0\) and cohesive energy \(E_{\text{Coh}}\) as determined by different theoretical methods. Two sets of experimental values are given, the second of which is corrected for zero-point energy effects (indicated by ZPEC). The notation PP denotes pseudo-potential, SR the short-range description used (exchange), and LR the long-range description used (correlation). PBE/PBE (PBE-D2 rescaled) refers to a PBE calculation rescaled using the PBE-D2 lattice constant, this is later referred to as PBE-rescaled.

<table>
<thead>
<tr>
<th>PP</th>
<th>SR / LR</th>
<th>(a_0) (±0.02 Å)</th>
<th>(E_{\text{Coh}}) (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAW</td>
<td>PBE / PBE</td>
<td>3.64</td>
<td>3.50</td>
</tr>
<tr>
<td>US</td>
<td>PBE / PBE</td>
<td>3.64</td>
<td>3.52</td>
</tr>
<tr>
<td>PAW</td>
<td>revPBE / revPBE</td>
<td>3.68</td>
<td>3.11</td>
</tr>
<tr>
<td>US</td>
<td>revPBE / revPBE</td>
<td>3.68</td>
<td>3.07</td>
</tr>
<tr>
<td>PAW</td>
<td>PW86 / PW86</td>
<td>3.68</td>
<td>5.16</td>
</tr>
<tr>
<td>US</td>
<td>PW86 / PW86</td>
<td>3.68</td>
<td>3.34</td>
</tr>
<tr>
<td>PAW</td>
<td>PW86R / PW86R</td>
<td>3.68</td>
<td>3.19</td>
</tr>
<tr>
<td>US</td>
<td>PW86R / PW86R</td>
<td>3.68</td>
<td>3.27</td>
</tr>
<tr>
<td>PAW</td>
<td>PW91 / PW91</td>
<td>3.64</td>
<td>3.54</td>
</tr>
<tr>
<td>US</td>
<td>PW91 / PW91</td>
<td>3.64</td>
<td>3.55</td>
</tr>
<tr>
<td>PAW</td>
<td>C09x / C09x</td>
<td>3.56</td>
<td>4.24</td>
</tr>
<tr>
<td>US</td>
<td>C09x / C09x</td>
<td>3.56</td>
<td>4.16</td>
</tr>
<tr>
<td>PAW</td>
<td>PBE / PBE-D2</td>
<td>3.573</td>
<td>3.91</td>
</tr>
<tr>
<td>PAW</td>
<td>PBE / vdW-DF1</td>
<td>3.674</td>
<td>3.38</td>
</tr>
<tr>
<td>US</td>
<td>PBE / vdW-DF1</td>
<td>3.68</td>
<td>3.46</td>
</tr>
<tr>
<td>PAW</td>
<td>PBE / vdW-DF2</td>
<td>3.68</td>
<td>3.14</td>
</tr>
<tr>
<td>US</td>
<td>PBE / vdW-DF2</td>
<td>3.68</td>
<td>3.18</td>
</tr>
<tr>
<td>PAW</td>
<td>revPBE / vdW-DF1</td>
<td>3.712</td>
<td>2.99</td>
</tr>
<tr>
<td>US</td>
<td>revPBE / vdW-DF1</td>
<td>3.72</td>
<td>3.00</td>
</tr>
<tr>
<td>PAW</td>
<td>revPBE / vdW-DF2</td>
<td>3.72</td>
<td>-</td>
</tr>
<tr>
<td>US</td>
<td>revPBE / vdW-DF2</td>
<td>3.72</td>
<td>-</td>
</tr>
<tr>
<td>PAW</td>
<td>PW86 / vdW-DF1</td>
<td>3.72</td>
<td>3.17</td>
</tr>
<tr>
<td>US</td>
<td>PW86 / vdW-DF1</td>
<td>3.68</td>
<td>3.29</td>
</tr>
<tr>
<td>PAW</td>
<td>PW86 / vdW-DF2</td>
<td>3.76</td>
<td>2.92</td>
</tr>
<tr>
<td>US</td>
<td>PW86 / vdW-DF2</td>
<td>3.76</td>
<td>3.01</td>
</tr>
<tr>
<td>PAW</td>
<td>PW86R / vdW-DF1</td>
<td>3.72</td>
<td>3.11</td>
</tr>
<tr>
<td>US</td>
<td>PW86R / vdW-DF1</td>
<td>3.72</td>
<td>3.24</td>
</tr>
<tr>
<td>PAW</td>
<td>PW86R / vdW-DF2</td>
<td>3.76</td>
<td>2.86</td>
</tr>
<tr>
<td>US</td>
<td>PW86R / vdW-DF2</td>
<td>3.72</td>
<td>2.97</td>
</tr>
<tr>
<td>PAW</td>
<td>PW91 / vdW-DF1</td>
<td>3.56</td>
<td>4.55</td>
</tr>
<tr>
<td>US</td>
<td>PW91 / vdW-DF1</td>
<td>3.56</td>
<td>4.55</td>
</tr>
<tr>
<td>PAW</td>
<td>PW91 / vdW-DF2</td>
<td>3.56</td>
<td>4.26</td>
</tr>
<tr>
<td>US</td>
<td>PW91 / vdW-DF2</td>
<td>3.60</td>
<td>4.25</td>
</tr>
<tr>
<td>PAW</td>
<td>C09x / vdW-DF1</td>
<td>3.56</td>
<td>4.07</td>
</tr>
<tr>
<td>US</td>
<td>C09x / vdW-DF1</td>
<td>3.56</td>
<td>4.06</td>
</tr>
<tr>
<td>PAW</td>
<td>C09x / vdW-DF2</td>
<td>3.60</td>
<td>3.81</td>
</tr>
<tr>
<td>US</td>
<td>C09x / vdW-DF2</td>
<td>3.60</td>
<td>3.79</td>
</tr>
<tr>
<td>PAW</td>
<td>PBE / PBE (rescaled to PBE-D2 (a_0))</td>
<td>3.573</td>
<td>3.49</td>
</tr>
<tr>
<td>PAW</td>
<td>PBE / vdW-DF1 (rescaled to 3.582 Å)</td>
<td>3.582</td>
<td>3.35</td>
</tr>
<tr>
<td>US</td>
<td>PW86R / vdW-DF2 (rescaled to 3.553 Å)</td>
<td>3.582</td>
<td>2.86</td>
</tr>
<tr>
<td>Exp.</td>
<td>Literature</td>
<td>3.61 [115]</td>
<td>3.49 [116]</td>
</tr>
</tbody>
</table>
4.3 General Calculation Details

Oscillations of contractions and expansions of the inter-layer separations were observed, similar to those discussed in Ref. [113], summarised in Table 4.2 for \(N_l\) layers. Additionally, oscillations in the surface energy as the slab size increased were observed, similar to those discussed in the study by J. Harl and G. Kresse [117], reminiscent of quantum beating [118]. However, whilst being an interesting observation, since the number of layers will remain unchanged when comparing different surfaces, it has been assumed that does not affect the study\(^1\). Therefore, upon comparing these results with experimental results, and by using the expression for the change in energy of a system with adsorbed atoms upon addition of layer, (cf. Eq. 3.106), a five layer slab with a 7 Å vacuum gap was found to be suitable. Hence, to construct the oxidised slabs using the optimised parameters, an additional, adsorbed, layer was placed on top. The bare Cu(110) slab was represented using a \((6 \times 2)\) unit cell with five atomic layers, and all oxidised surfaces had an additional layer on which adsorbed oxygens form the reconstructed surface with additional Cu atoms. All calculations of the various surface reconstructions were performed using the \((6 \times 2)\) cell to reduce finite size effects, enabling unambiguous comparison. The bottom two layers were kept frozen in the calculated bulk positions so as to emulate the bulk material, and all other atoms were allowed to relax. Both the cell size and the number of layers used were checked for convergence and the calculated energies were found to be accurate to within \(<12\) meV.

Table 4.2: Comparison of inter-layer separations for the Cu(110) surface slab as a percentage of the separation in the bulk, calculated using PBE. \(N_l\) denotes the number of layer in the slab, and \(d_{ij}\) refers to the inter-layer separation of the \(i^{th}\) and \(j^{th}\) layer. S. Y. Liem et al. results refer to a \(k\)-point sampling of \(16 \times 16 \times 16\). Positive values indicate expansion and negative values indicate contractions.

<table>
<thead>
<tr>
<th>(N_l)</th>
<th>(d_{12})</th>
<th>(d_{23})</th>
<th>(d_{34})</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-9.52</td>
<td>3.21</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-10.26</td>
<td>4.63</td>
<td>-1.35</td>
</tr>
<tr>
<td>9</td>
<td>-9.90</td>
<td>3.62</td>
<td>-0.28</td>
</tr>
<tr>
<td>S. Y. Liem et al. [113]</td>
<td>-10.8</td>
<td>5.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Exp. [113]</td>
<td>-3.0 to -10.0</td>
<td>0.0 to 8.0</td>
<td>-</td>
</tr>
</tbody>
</table>

4.3.1 Theoretical Model

As an initial step in our understanding of the nature of the ambiguity, models for the surface reconstructions used by X. Duan et al. [109] were revisited. At variance with the previous calculations based on localised basis sets, we used the plane wave VASP code, setup as described in Section 4.3.

The experimental conditions of high pressure and temperature of the oxygen gas environment being in thermal equilibrium with the copper surface are explicitly taken into account in

\(^1\)This was indirectly found to be the case later on from calculations to determine the potential energy surface, which were performed using different sized slabs.
4.3 General Calculation Details

*ab initio* surface thermodynamics (see Section 3.11.2). This is done using the chemical potential of oxygen, $\Delta \mu_O$ (defined relative to half the $O_2$ energy), which is a function of the molecular oxygen gas partial pressure, $p_{O_2}$, and temperature, $T$. Thus, for the same slab size, the most stable surface structure has the lowest Gibbs free energy (GFE) difference, $\Delta G$, where

$$
\Delta G^{\text{slab}}_{O/Cu} = G^{\text{slab}}_{O/Cu} - G^{\text{slab}}_{Cu} - \Delta N_{Cu} \mu_{Cu} - \Delta N_O \Delta \mu_O \quad (4.1)
$$

for the given range of $\Delta \mu_O$ ($p_{O_2}$ and/or $T$), relative to the bare surface.

Since the $p(2 \times 1)$ phase, Fig. 4.1(a), is known to be the ground state reconstruction at lower oxygen exposures, we use the bare surface and the $p(2 \times 1)$ surface as reference systems, and expect the AS-c(6 × 2) surface to be the next thermodynamically favorable surface at higher oxygen exposures.

### 4.3.2 Additional Variations of the Cu(110):O Surface Reconstructions

**Figure 4.2:** One possible geometry of the (1 × 1) reconstructed oxidised Cu(110) surface, colour coded as in Fig. 4.1.

**Figure 4.3:** Two possible geometries of the (2 × 1) reconstructed oxidised Cu(110) surface, colour coded as in Fig. 4.1.

**Figure 4.4:** Two possible geometries of the (2 × 2) reconstructed oxidised Cu(110) surface, colour coded as in Fig. 4.1.

**Figure 4.5:** Three possible geometries of the (3 × 1) reconstructed oxidised Cu(110) surface. Purple balls represent oxygen atoms, that lie lower that the other surface oxygen and other atoms are colour coded as in Fig. 4.1.

**Figure 4.6:** Five possible geometries of the (6 × 2) reconstructed oxidised Cu(110) surface. Purple balls represent oxygen atoms, that lie lower that the other surface oxygen and other atoms are colour coded as in Fig. 4.1.

Initially, over 30 variations of the AS reconstruction of the Cu(110):O surface, which might serve as possible candidates at higher oxygen exposure, were considered. The representative geometries of the different variations of the (1 × 1) - Fig. 4.2, (2 × 1) - Fig. 4.3, (2 × 2) -
4.4 Gibbs Free Energy Differences (GFE)

Fig. 4.4, (3 × 1) - Fig. 4.5 and (6 × 2) - Fig. 4.6 reconstructions are shown, and of those, seven geometries were studied in more detail, all corresponding to the lowest energy structures, shown in Fig. 4.1(b-h).

4.4 Gibbs Free Energy Differences (GFE)

4.4.1 PBE and PW

![Gibbs Free Energy Plot](image_url)

**Figure 4.7:** Gibbs free energy plots for $p(2 \times 1)$-like geometries of Cu(110):O system calculated using PBE without considering vibrations. The brown (horizontal) and blue lines correspond to the clean Cu(110) and $p(2 \times 1)$ surfaces, respectively. The black lines correspond to four other $p(2 \times 1)$ geometries with different arrangements of super Cu atoms.

Before studying the AS-c(6 × 2) phase, to both test our setup and verify previous findings, we show in Fig. 4.7 the calculated GFE plot for the $p(2 \times 1)$ and $p(2 \times 1)$-like geometries with different arrangements of super Cu atoms (one of which is shown on the left in Fig. 4.3). Here the $p(2 \times 1)$ surface (blue line) clearly lies lowest in energy. We also studied variants of the AR structures, (1 × 1), (2 × 1) and (3 × 1) phases, considering all plausible variations in which super Cu atoms are also present. In line with previous findings [113], we found that the (2 × 1) phase is indeed the ground-state structure at low oxygen exposures.

Having established that our setup models the $p(2 \times 1)$ phase correctly, we studied the AS-c(6 × 2) phase. Considering the same cell size, we modelled the seven lowest energy reconstructions, shown in Fig. 4.1(b-h), using 8 oxygen atoms per unit cell, differing only in the number of super Cu atoms per unit and their locations; hence their gradients on the GFE difference plot vs. $\Delta \mu_O$ are the same.

Comparison of our PBE calculations, shown in Fig. 4.8(a), with those in the previous study by X. Duan et al. [109] immediately shows excellent agreement, that for low values of $\Delta \mu_O < -1.97$ eV the clean surface is favourable, for $-1.97 < \Delta \mu_O < 1.38$ eV the $p(2 \times 1)$ is favoured, and then for $-1.38 < \Delta \mu_O < -1.24$ eV the AS-c(6 × 2) surface is favoured, until the bulk-oxide forms
Figure 4.8: Gibbs free energy plots for the Cu(110):O system calculated using PBE without considering vibrational contributions to the energy (a) and including vibrations (b). The brown (horizontal), blue and red lines correspond to the clean Cu(110) surface, p(2 x 1) and AS-c(6 x 2) phases, respectively. The black lines correspond to the other six of the seven AS-c(6 x 2)-like geometries shown in Fig. 4.1. The thick green line indicates the systems that are thermodynamically favourable for the given range of ∆µ_O; the relevant regions are also shaded in colour and labelled to highlight this.

(-1.24 eV corresponds to the calculated ∆µ_O value in the bulk-oxide, Cu_2O, assuming the same value of µ_Cu).

It can be seen here that although the AS-c(6 x 2) surface is indeed calculated as the most favourable in the appropriate window of ∆µ_O, there are several other structures very near in energy, see Table 4.3. In fact, the AS-(6 x 2)_1 structure in Fig. 4.1(e) differs only by 27 meV, which is well beyond the precision of DFT.

For the purpose of comparing with the previous studies done by J. Harl and G. Kresse [117] and S. Y. Liem et al. [113], calculations both with and without vibrations were also performed using the PW91 density functional, in the same way as in the PBE calculation, but using ultra-soft pseudo-potentials instead. These plots are shown in Fig. 4.9. Upon comparison of PBE and PW calculations, we found that the calculated order of structures in terms of their stability depends crucially on the density functional used: for instance, when a PW functional was used, shown in Fig. 4.9(a), the AS-(6 x 2)_1 was instead found to be more favourable by 20 meV without the vibrational contribution; this is in contrast with the PBE calculation that found the AS-c(6 x 2) phase as the most stable. Note, however, that these energy differences are very small, so whilst these results should be considered far from being conclusive, they do strongly indicate the failure of the PW functional at describing these systems, and the sensitivity of these predictions on the functionals used.

Also note in passing that, in contrast to the previous study [109], our calculations have shown that the structure in Fig. 4.1(f) is in fact separated from the lowest energy structure by a considerable gap, even though this was one of the surfaces previously found to be very close in
Delta G to the others; this means that at least some of the ambiguity found in [109] can indeed be attributed to the localised basis set used there. Nevertheless, the fundamental issue still remains that there are several surfaces thermodynamically indistinguishable from the AS-c(6 \times 2) surface as obtained by the conventional DFT methods.

Table 4.3: Calculated Gibbs free energy differences (meV) of surfaces relative to the AS-c(6 \times 2) phase as determined by different theoretical methods. The asterisks indicate the lowest energy structures for each method of calculation.

<table>
<thead>
<tr>
<th>Structure</th>
<th>(\Delta G)</th>
<th>(\Delta G + \Delta F^{\text{vib}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-c(6 \times 2)</td>
<td>0*</td>
<td>0*</td>
</tr>
<tr>
<td>AR-(3 \times 1)</td>
<td>62</td>
<td>214</td>
</tr>
<tr>
<td>AS-(3 \times 1)</td>
<td>1171</td>
<td>1086</td>
</tr>
<tr>
<td>AS-(6 \times 2)</td>
<td>27</td>
<td>148</td>
</tr>
<tr>
<td>AS-(6 \times 2)_1</td>
<td>127</td>
<td>19</td>
</tr>
</tbody>
</table>

4.4.2 Higher Precision Method: \(E_{\text{PBE}} + F^{\text{vib}}\)

As a next step we have included the contributions to the energy due to vibrations, \(F^{\text{vib}}\), using the frozen phonon method within the quasi-harmonic approximation [42, 81] (cf. Section 3.11.1). To calculate \(F^{\text{vib}}\), vibrational modes for each competing surface were calculated. In our calculations, the lowest three layers of the slabs, common to all surfaces and separated from the oxidised layer by two atomic layers, were frozen. For validation of cell size, \(F^{\text{vib}}\) for a single system was compared to that of a larger cell size, and negligible error was seen (<2%). The plots of \(\Delta G + F^{\text{vib}}\) is given in Fig. 4.8(b), and Fig. 4.9(b), summarised in Table 4.3.

As can be seen in Fig. 4.8(a) and Table 4.3, there is some improvement in the relative stability of the AS-c(6 \times 2) surface in the PBE case, bring the theoretical predications closer...
4.5 The Critical Role of Non-Local Interactions

The author reminds the reader that this study of Cu(110):O surface was conducted with the aim of deepening our understanding of the surface and providing a valid theoretical model (correct calculation parameters). To that end, although inclusion of vibrations has been shown to introduce added stability to the AS-c(6 × 2) surface relative to the other surface, the reason behind the poor description of the system (relative to the other similar systems) is still unknown, i.e. why adding vibrations appears to add stability to the AS-c(6 × 2) surface for the PBE density functional, but not for the PW. Therefore, given the semiconductor/metal nature of the Cu(110):O surface with super CuO$_2$ units formed on it, we could ask ourselves whether the lack of dispersion interaction (non-local correlation) in conventional DF methods is the reason for the failure of DFT in describing this surface. If this is indeed the case, then that could provide an explanation of the failure of $\Delta G_{PW} + F_{vib}$ as compared to $\Delta G_{PBE} + F_{vib}$; that it may be due to their different treatments of the gradient of the exchange-correlation density functional. In order to describe non-local interactions we initially used the PBE-D2 method (cf. Section 3.8.2.3.1), in which the PBE functional is employed with the appropriate DFT-D2 correction.

4.5.1 Calculation Details

Calculations accounting for non-local interactions were performed using the DFT-D2 correction (PBE-D2) [54]. These were performed initially by Andris Gulēns at Aalto University using VASP 5.2 and then subsequently by the author using an implementation in VASP 4.6 developed by the group of Prof. Ivan Štich. This method predicts a decrease of the Cu lattice constant from 3.64 Å (PBE) to 3.57 Å (PBE-D2), still agreeing with the experimentally calculated lattice constant to within 2%, however, the cohesive energy calculated using the PBE-D2 method comes out with a larger error as compared to the PBE prediction, shown in Table 4.1. The $\Delta G$ results are shown in Fig. 4.10, also summarised in Table 4.4. Recently V. G. Ruiz et al. [119] have shown
that for some systems, such as organic molecules physisorbed on metal surfaces, many-body collective response (screening) of the substrate electrons may be needed in determination of dispersion coefficients and vdW radii in DFT-D type of vdW theory, an approach going effectively beyond the pairwise description. Although such an approach is deemed not necessary for description of the Cu(110):O surface due to formation of a semiconductor/metal interface upon oxidation, DFT-D2 method uses empirically determined $C_6$ dispersion coefficients and hence, these might change considerably upon oxygen adsorption. For that reason our collaborators, Prof. Ivan Štich’s group, have explicitly checked the distribution of $C_6$ coefficients on the surface and in the slab using a procedure proposed by A. Tkatchenko and M. Scheffler for estimating $C_6$ and vdW radius by a Hirshfeld partitioning of the electron density [56]. For example, for the AS-$c(6 \times 2)$ phase only small variations of up to $\approx 6\%$ for Cu atoms in different environments and much smaller variations for oxygen atoms have been noticed. These variations are much smaller than differences between dispersion coefficients of free metallic atoms determined by different methods available.

4.5.2 Higher Precision Method: PBE-D2

![Figure 4.10: Gibbs free energy plots for the Cu(110):O system calculated using PBE-D2 without considering vibrational contributions to the energy (a) and including vibrations (b). The colour coding is the same as in Fig. 4.8.](image)

As one can see from Fig. 4.10(a) and Table 4.4, this method shows both that the $p(2 \times 1)$ phase is the most favourable for a reasonable range of $\mu_O$ values, although this range of $\mu_O$ is significantly narrower than that calculated by the ordinary PBE, and that the AS-$c(6 \times 2)$ surface is the most favourable at larger values of $\mu_O$. Note, however, that surprisingly the separation between the AS-$c(6 \times 2)$ structure and the next one in energy, AS-$c(6 \times 2)_1$, was found to increase by more than a factor of 4 relative to the PBE result, reaching 113 meV, without the need to include vibrations. Including now the $F_{vib}$ term, as shown in Fig. 4.10(b) and Table 4.4, adds yet more stability to the AS-$c(6 \times 2)$ surface, separating it significantly further from its
nearest competitor by 0.14 eV, which is 0.1 eV more stable as compared to the corresponding PBE result.

**Table 4.4:** Calculated Gibbs free energy difference $\Delta G$ (meV) of surfaces relative to the AS-$c(6 \times 2)$ phase without/with vibrational contribution $F^{\text{vib}}$. $E_{\text{DFT}}$ ($E_{\text{vdW}}$) denote PBE (vdW) contributions to the “adsorption” energy per super Cu atom relative to the “bare” AR-(3 $\times$ 1) structure. $n$ is the number of super Cu atoms per $6 \times 2$ cell.

<table>
<thead>
<tr>
<th>Structure</th>
<th>n</th>
<th>$\Delta G$</th>
<th>$E_{\text{DFT}}$</th>
<th>$E_{\text{DFT}}$</th>
<th>$E_{\text{vdW}}$</th>
<th>$E_{\text{DFT}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-$c(6 \times 2)$</td>
<td>2</td>
<td>0&quot;</td>
<td>0&quot;</td>
<td>0&quot;</td>
<td>-31</td>
<td>-205</td>
</tr>
<tr>
<td>AR-(3 $\times$ 1)</td>
<td>0</td>
<td>62/214</td>
<td>582/572</td>
<td>353/621</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AS-(6 $\times$ 2)</td>
<td>4</td>
<td>1171/1086</td>
<td>882/786</td>
<td>1055/1013</td>
<td>277</td>
<td>154</td>
</tr>
<tr>
<td>AS-(6 $\times$ 2)$_1$</td>
<td>1</td>
<td>27/148</td>
<td>113/251</td>
<td>128/300</td>
<td>-35</td>
<td>-265</td>
</tr>
</tbody>
</table>

### 4.5.3 Analysis

#### 4.5.3.1 Electronic Density

In an attempt to see more clearly why accounting for non-local interactions correctly results in the favourability of the AS-$c(6 \times 2)$ phase, we also studied the electronic density of the surfaces. We try to identify if there are any localised pockets of electron density. Since in conventional techniques (such as PBE) non-local interaction between these “pockets” is not properly accounted for, one would expect to find such regions of the density in the AS-$c(6 \times 2)$ phase specifically, but not in the others. This might then shed light on why the vdW based methods provide an additional stability for this particular phase as compared with the others.

Here we compare the AS-$c(6 \times 2)$ phase with its closest competing structure AS-(6 $\times$ 2)$_1$.

The electronic density plots for these two surfaces, calculated using PBE, are shown in **Fig. 4.11**. We also show the iso-surfaces of the electronic densities of the two structures in **Fig. 4.12**.

Comparing and analysing the results presented in **Fig. 4.11** and **Fig. 4.12**, the densities of the two surfaces, one can indeed find “pockets” of the electron density on both of them, localised in equivalent regions. Furthermore, the cuts through the “high O - super Cu - high O” chain of atoms on both surfaces show that the density distributions are almost identical. The AS-(6 $\times$ 2)$_1$ surface has some missing super Cu atoms with respect to the AS-$c(6 \times 2)$ surface, however, the relaxed O atoms around those positions are predicted to move in and up as if there were super Cu atoms present. Due to this effect, the density cuts through the regions where the super Cu atoms are missing (the AS-(6 $\times$ 2)$_1$ surface) shown in Figs. 4.11 and 4.12 demonstrate that the density closely resembles the one of the AS-$c(6 \times 2)$ surface which has super Cu atoms present. Note that the effect of the oxygen atoms displacing towards the positions they would occupy if the super Cu atom was present is seen in each of the geometries relaxed by different functionals. For conventional methods the displacement is small, the smallest for PBE (as measured from the
4.5 The Critical Role of Non-Local Interactions

Figure 4.11: Electronic density plots of the AS-c(6 × 2) (left panels) and AS-(6 × 2)\textsubscript{1} (right panels) structures calculated using PBE. Top panels: the cut is made perpendicular to the surface and passing through the “high O - super Cu - high O” chain of atoms. Bottom panels: the cut is made along the perpendicular direction to that used in the left panels and goes along the -O-Cu-O- row of atoms for the AS-(6 × 2) surface and across the missing super Cu site on the AS-(6 × 2)\textsubscript{1} surface. Grey and red balls denote copper and oxygen atoms respectively (as in Fig. 4.1), and the translucent blue regions represent the electronic density distribution.

Figure 4.12: Top and side views of the iso-surfaces of the electronic density (blue regions) of the AS-c(6 × 2) (a) and AS-(6 × 2)\textsubscript{1} (b) structures calculated using PBE. Grey and red balls denote copper and oxygen atoms respectively (as in Fig. 4.1), and the translucent blue regions represent the electronic density distribution.

ideal position), however, when accounting for non-local interactions (PBE-D2) these displacements significantly increase.

From this it can be concluded that the electron density distribution does not explain why
the AS-c(6 × 2) phases becomes more favourable than the AS-(6 × 2)_1 upon consideration of vdw.

4.5.3.2 Adsorption of Super Cu: vdw Contribution

The energy balance between the chemical (PBE) and physical (vdW) interactions in different structures can also be compared by considering each phase as being obtained by “adsorbing” additional super Cu atoms (with the “gas phase” energies $\mu_{\text{Cu}}$) on the “bare” AR-(3 × 1) phase, Fig. 4.1(c). In the 7th and 8th columns of Table 4.4 the relative (with respect to the “bare” phase) chemical (PBE) and vdw contributions to the total energies of the three structures are given, calculated per super Cu atom. The AS-(6 × 2)_1 structure which is the nearest contender to the AS-c(6 × 2) phase, is seen to lie much lower in energy, as compared to the AR-(3 × 1) phase, due to considerable PBE and vdw contributions upon the addition of a single super Cu atom. Adding two atoms (yielding the AS-c(6 × 2) phase) results in larger reduction of the chemical and considerable reduction of the vdw energies per added atom; overall this phase becomes more favourable. Adding four super Cu atoms to construct the AS-(3 × 1) structure results in significant chemical repulsion between CuO$_2$ units, while the vdw energy (per added atom) does not change significantly from the AS-c(6 × 2) contribution. Identical analysis for the PBE-only calculation shown in the preceding column of Table 4.4 shows a significantly smaller difference between the two phases, AS-c(6 × 2) and AS-(6 × 2)_1, and hence explains the failure of the theory to identify the experimentally observed phase at high oxygen exposures. This analysis indicates a delicate balance of the chemical and physical interactions in this system.

4.5.3.3 Dependence of $\Delta G$ on the $a_0$
Looking now only on the chemical (PBE) contribution of the PBE-D2 calculation (the 7th column in Table 4.4) shows that the GFE difference between the nearest competitors AS-c(6 × 2) and AS-(6 × 2)_1 already amounts to substantial 145 meV. Since the only difference between this and the PBE-only (the preceding column) calculation was in the geometry of the surfaces, which must be significantly affected by different bulk lattice constants, 3.64 Å in PBE and 3.57 Å in PBE-D2, one may think that contraction of the bulk lattice constant when accounting for the non-local interaction must be the decisive factor. Indeed, when we performed PBE calculations on the four contenders listed in Table 4.3 with the reduced lattice constant of 3.58 Å, referred to as PBE-rescaled, the plot of which is shown in Fig. 4.13, we found an energy gap of 128 meV between the AS-c(6 × 2) and its nearest competitor AS-(6 × 2)_1, which further increases to 300 meV after including the vibrational contribution. Note that in these calculations no buckling of the CuO₂ units was found which means that this effect, which is only observed using the DFT-D2 method, seems to be of secondary importance.

Additionally, from the PBE-rescaled calculations, it can be seen that, indeed, both the p(2 × 1) and AS-c(6 × 2) phases are found to be favourable, with the p(2 × 1) phase favourable for a large range of µO values. We also found similar displacements of the high O atoms where no super Cu atoms were present (as discussed in Section 4.5.3.1). We found displacements of ≈0.3 Å, much larger than in the PBE calculations and slightly less than calculated by PBE-D2. These calculations demonstrate that by reducing the lattice constant a₀ to the value predicted by PBE-D2, which roughly accounts for the non-local interaction, one can achieve correct ordering of different phases by using conventional PBE density functional based calculations, which, consequentially, are less expensive.

To further investigate the role the reduction in the lattice constant( a₀) has in stabilising the AS-c(6 × 2) phase, we compared again the AS-c(6 × 2) and AS-(6×2)_1 structures. Since both of them have the same number of O atoms, their GFEs are parallel to each other, hence a direct comparison can be made for any value of µO. In Fig. 4.14 we show the difference in ∆G of the two structures calculated with the PBE method as a function of a₀. It was found that the energy gap between the two structures increases as a₀ decreases (Fig. 4.14(a)): although the PBE calculated lattice constant gives the minimum bulk energy (Fig. 4.14(b)), decreasing a₀ results in larger changes in energy of the AS-c(6 × 2) structure as compared to the AS-(6×2)_1 structure (Fig. 4.14(c)), leading to the increase of the energy gap between them at a smaller a₀ values. Since the attractive non-local interaction results in a shrinking of a₀, as we observed with the PBE-D2 method (see Table 4.1), this indeed leads the AS-c(6 × 2) structure to become more energetically favourable when the non-local interaction is accounted for.
The Critical Role of Non-Local Interactions

4.5.4 Summary

To summarise, we have explicitly reported on failure of conventional DFT methods (PBE and PW) to conclusively determine the AS-c(6 × 2) phase as the ground-state reconstruction of the Cu(110) surface at higher oxygen exposures. We demonstrated that considering vibrations generally strengthens its relative stability. However, the stabilities are very sensitive to the choice of the exchange-correlation functional, and hence, conventional methods are still unable to conclusively determine the most stable structure. Surprisingly, we found that non-local interactions play a crucial role in unambiguously isolating the AS-c(6 × 2) phase as the most thermodynamically stable at high oxygen exposures.

Typically, non-local interactions, namely van der Waals (vdW), are routinely used in the description of sparse matter [120], in studying of interactions of closed-shell molecules and clusters with surfaces [121, 122] and among themselves [123, 124], in systems of biological relevance [120], etc... However, contrary to these examples, surfaces are prototypical systems with higher reactivity featuring “strong” chemical bonds, and for them non-local interactions are expected to play only

Figure 4.14: (a) Plot of the Gibbs free energy (GFE) difference between AS-c(6 × 2) and AS-(6× 2)_1 structures (solid line with asterisks) as a function of the Cu bulk lattice constant, \(a_0\). (b, c) Bulk and slab total energies, respectively, as functions of \(a_0\), relative to the PBE calculated \(a_0 = 3.64\) Å value. All calculations were performed using the PBE method. Blue squares and red circles indicate where PBE (3.64 Å) and (PBE-rescaled, 3.58 Å) lattice constants, respectively, are on the plot.
4.6 The Effect of Using Non-Local Density Functionals

a minor, if any, role. Therefore, it is very interesting that we have now found that there are surface systems that are stabilised by a delicate balance between short-range chemical and longer-range vdW (physical) interactions. The semiconductor/metal interface formed on the Cu(110) surface upon oxidation stabilises the super CuO\(_2\) units by a delicate combination of interactions. The latter interaction needs to be added to the standard DFT in order to provide the necessary reduction of the bulk lattice constant and correctly capture both effects.

Furthermore, we believe that the effect we observe may not be limited to only the Cu(110):O surface. Stabilities of other surfaces, such as the oxidised Ag(110) [125], Ni(110) [110], and Au(110) [126], which have missing row reconstructions, with the possible presence of super metal atoms, are also expected to be affected by the non-local correlation effects, which may need to be included.

4.6 The Effect of Using Non-Local Density Functionals

4.6.1 Introduction

In a study [127], detailed above, the issue of the accuracy of conventional DFT methods in studying the thermodynamic stabilities of Cu(110):O (oxidised Cu(110)) surfaces was investigated. In this study it was shown that conventional DFT methods fail to accurately describe the AS-c(6 × 2) surface, that only when accounting for non-local interactions does theory agree with experimental results.

Our aim therefore was to study these surfaces in greater detail, with self-consistent methods for treating the non-local interactions instead of the semi-empirical method used in Ref. [127]. In the previous study the Grimme DFT-D2 method [54] was used to describe non-local interactions, however, this method does not affect the electronic density. Therefore, the self-consistent van der Waals density functional (cf. Section 3.8.2.3.2) [57, 58], vdW-DF1/2, was used in this study, drawing comparisons with the PBE and PBE-D2 results discussed in [127]. In this way the effects the choice of the long-range (LR) descriptions have on results were also investigated. A similar investigation was also performed on the effects of the choice of the description of short-range (SR) interactions, by changing the description of the SR interactions. Furthermore, this study also investigated whether or not the main contribution to the accuracy originates from the recalculation of the lattice constant, taking into account non-local interactions, as is suggested in Ref. [127].

4.6.2 Calculation Details

The calculations were performed in a very similar manner to Section 4.5. We used the plane wave VASP code, employing the GGA-PBE density functionals. Calculations accounting for
4.6 The Effect of Using Non-Local Density Functionals

Non-local interactions were performed using vdW-DF1 and vdW-DF2 density functionals [57, 58]. All calculations performed using PBE and vdW-DF1 functionals, were performed using PAW-PP, while US-PP were used with vdW-DF2 functionals, both describing every copper and oxygen atoms with 11 and 6 valence electrons, respectively. For the bulk copper, the parameters outlined here yielded the FCC unit cell lattice constants and cohesive energies shown in Table 4.1.

4.6.3 Cu Bulk Properties

The effects the choice of the SR and LR descriptions of the system has on results is determined by studying the effects on the bulk properties, i.e. lattice constant and cohesive energy. For this part of the study, a wide selection of functionals have been employed: LDA [128], PBE, revised PBE (revPBE), PW86 [129], PW86R, PW91 and C09x [130]. Furthermore, the choice of pseudo-potential for each of these methods has also been varied, using PAW and US. The results from this large selection of functional mixings are given in Table 4.1. This is a good place to note that, for the actual study of the surfaces, the methods implemented were limited to the PBE, PW, PBE-D2, PBE-vdW-DF1, revPBE-vdW-DF1 and PW86R-vdW-DF2 density functionals.

Comparing the results with previous DFT studies employing the PBE and PW density functionals for LR and SR interactions [109, 113], as well as the PW86R-vdW-DF2 results with Ref. [62], shows excellent agreement. However, the table as a whole also highlights the wide variety of results that can be obtained depending on the method implemented. Generally GGA methods suffer from under-binding, yielding a larger bulk lattice constant, this is generally consistent among the results but the cohesive energy tends to vary widely. Moreover, for the vdW methods (PBE-D2 and vdW-DF1/2), one would always expect the lattice constant to contract due to the additional attractive contribution that the inclusion of vdW interactions should bring. However, a large majority of these results show an increase in the lattice constant by up to 5% and give up to a ±18% error in the cohesive energy, as compared to experimental values. This problem likely results from the fact that these method were not initially designed for solids, but for layered materials and molecules; these method been applied to solids (bulk materials) only recently (c.f. Ref. [62]). Therefore, caution is advised when choosing how to describe LR interactions (correlation) in solids.

Although it is clear that none of the methods in Table 4.1 are ideal to use, keeping in line with many past studies, the vdW-DF1 was used to study the Cu(110):O surface in more detail, in particular the electronic distribution on the AS-c(6 × 2) surface and its thermodynamically indistinguishable competitors.

\(^1\)Nomenclature: PBE-vdW-DF1 refers to using the PBE functional to describe electron exchange, while using the vdW-DF1 functional to describe non-local interactions (correlation). revPBE-vdW-DF1 and PW86R-vdW-DF2 have similar descriptions.
4.6.4 Higher Precision Method: vdW-DF

The Gibbs free energy plots for the Cu(110):O system calculated using (a) PBE, (b) PBE-vdW-DF1 and (c) revPBE-vdW-DF1, all without considering vibrational contributions of the slabs. The colour coding is the same as Fig. 4.8. Note that the CuO$_2$ bulk was not calculated using the vdW-DF1 functionals and therefore those regions are not included in images (b) and (c).

The Gibbs free energy difference plots and contributions to the energy due to vibrations were calculated as before. These calculations were fully self-consistent and we were able to perform complete geometry relaxations, possible thanks to an efficient implementation of the vdW-DF1 method [131].

The Gibbs free energy plots for O/Cu(110) without considering vibrations are shown in Fig. 4.15 and Table 4.5 gives the summary of results with and without vibrational contributions$^1$. Table 4.5 considers both types of systems, with the predicted larger lattice constant, and with shorter lattice constants (similar to that predicted by PBE-D2).

Upon studying Table 4.5, one can see that the vdW-DF1 (predicted) calculations, which predicted a larger lattice constant, even without considering the vibrational contribution to the GFE, immediately predict that the AS-c(6 × 2) reconstruction is highly favourable. The vdW-

$^1$The CuO$_2$ bulk has not been calculated here because at the time of writing this thesis our key focus was the stability of the AS-c(6 × 2) phase alone and not CuO$_2$. 

110
4.6 The Effect of Using Non-Local Density Functionals

DF1 calculations employing the larger lattice constant predict (without vibrations) that the surface is more than twice as stable as the AS-\(c(6 \times 2),1\) surface in the PBE-D2 calculations without vibrations, more stable than even the PBE-D2 calculations with vibrations.

It is worth noting, that for all methods except when employing the PBE density functional (Fig. 4.15(a)), the range of \(\Delta \mu_O\) for which \(p(2 \times 1)\) is favourable is significantly narrowed. This is in contrast to the findings in Ref. [108], which shows that the \(p(2 \times 1)\) should be most favourable for a very wide range of \(\Delta \mu_O\). This observation once again highlights the fact that the methods for treating non-local interactions in surfaces are not yet ideal and have to be used with caution. Nonetheless, it is clear that both the semi-empirical D2 method, depending only on nearest neighbour distances, and the self-consistent method give the same qualitative results. Thus, it can safely be said that the ambiguity in the theoretically calculated ground state structure at high oxygen exposures, made using conventional DFT techniques, can be resolved by accounting for the non-local interactions.

Table 4.5: Calculated Gibbs free energy difference (meV) relative to the AS-(3 \(\times\) 1) phase as determined by different theoretical methods. Asterisk indicates the lowest energy structure for each method of calculation.

<table>
<thead>
<tr>
<th>Structure</th>
<th>PBE (predicted)</th>
<th>vDW-DF1 (reduced (a_0))</th>
<th>vDW-DF2 (reduced (a_0))</th>
<th>PBE-rescaled (PBE-D2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-(c(6 \times 2))</td>
<td>-62*</td>
<td>-582*</td>
<td>-1166*</td>
<td>-353*</td>
</tr>
<tr>
<td>AS-(c(6 \times 2))</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AR-(3 (\times) 1)</td>
<td>1109</td>
<td>300</td>
<td>563</td>
<td>702</td>
</tr>
<tr>
<td>AS-(6 (\times) 2)</td>
<td>-35</td>
<td>-469</td>
<td>-296</td>
<td>-225</td>
</tr>
<tr>
<td>AS-(6 (\times) 2),1</td>
<td>-214*</td>
<td>-572*</td>
<td>-1268*</td>
<td>-621*</td>
</tr>
<tr>
<td>AR-(3 (\times) 1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AS-(3 (\times) 1)</td>
<td>872</td>
<td>196</td>
<td>-407</td>
<td>392</td>
</tr>
<tr>
<td>AS-(6 (\times) 2),1</td>
<td>-66</td>
<td>-321</td>
<td>-640</td>
<td>-321</td>
</tr>
</tbody>
</table>

4.6.5 Analysis

4.6.5.1 Dependence of \(\Delta G\) on the Method Employed

The calculations were also performed with the vdw-DF1 and vdw-DF2 functionals employing a shorter lattice constant, 3.58 Å and 3.55 Å, respectively, shown in Fig. 4.16. It can be seen from Fig. 4.16 that the vdw-DF1/2 calculations with a reduced lattice constant both still find the AS-\(c(6 \times 2)\) phase as much more favourable, in agreement with the calculations performed using the extended (predicted) lattice constant. Calculations with PBE-vdw-DF1 using the reduced lattice constant determined the AS-\(c(6 \times 2)\) phase to be more stable than the AS-(6 \(\times\) 2),1 phase by a further 0.12 eV as compared to that of the extended lattice constant. Furthermore,
4.6 The Effect of Using Non-Local Density Functionals

Figure 4.16: Gibbs free energy plots for the Cu(110):O system calculated both without considering vibrations, shown in the top panels (a,b), and while accounting for vibrational contributions of the slabs, shown in the bottom panels (c,d). The calculations were performed using PBE-vdW-DF1 (a,c) and PW86R-vdW-DF2 (b,d), both employing smaller lattice constants than they predict, 3.58 and 3.55, respectively (see Table 4.1 for predicted values). The colour coding is the same as Fig. 4.8.

the calculation with PW86R-vdW-DF2 shows a further increase in stability of 0.28 eV, as compared to the PBE-vdW-DF1 employing the reduced lattice constant, unambiguously identifying the AS-c(6 × 2) phase as the ground-state. Again the results of the stabilities improve greatly by including vibrations. However, it appears that performing these calculations with a reduced lattice constant significantly underestimates the stability of the p(2 × 1) surface such that it is often never favourable\(^1\), which contradicts experiment. In short, these calculations predict even more stability (twice as stable as the predictions made with the larger lattice constant). However, due to the computational cost of performing the calculations with the vdW-DF1/2 functionals, the contributions to the energy due to vibrations were not considered extensively, although preliminary calculations with PBE-vdW-DF1 employing the shorter lattice constant showed that the nearest neighbour separation increases to over 0.6 eV.

As already mentioned, the vdW-DF1/2 functionals tend to predict an expansion of \(a_0\) in contradiction to what is expected. This issue has been highlighted in recent studies [62, 132], and

\(^{1}\)This is not the case for the PW86R-vdW-DF2 calculation that also accounts for vibrations, in which the p(2 \(\times\) 1) phase is favourable for a narrow range of \(\mu_O\).
4.6 The Effect of Using Non-Local Density Functionals

while it has yet to be fully addressed, J. Klimeš et al. [62] has recently developed the optB88-vdW and optB86b-vdW functionals, which appear to give extremely accurate predictions for the Cu bulk lattice constant, and the optB88-vdW functional also accurately predicts the cohesive energy.

An interesting observation can be made from Table 4.5, that although the PBE-vdW-DF1 and revPBE-vdW-DF1 methods calculate a larger lattice constant, they both still find the AS-c(6 × 2) surface as most favourable by a large gap, where an ordinary PBE calculation using the same larger lattice constant would predict a different surface as favourable (cf. Fig. 4.14). This suggests that the non-local effects are not only important for determining the lattice constant, but also in describing the surfaces themselves, which could indicates that this result may not in fact be dependent on the lattice constant used (for realistic values of \( a_0 \)).

4.6.5.2 Electronic Density

![Figure 4.17](image)

Figure 4.17: Electronic density plots of the AS-(6×2) and AS-(6×2)_1 structures calculated using PBE-vdW-DF1. The cuts are made along the same directions through the super Cu (for the AS-(6×2) surface) and the missing super Cu (AS-(6×2)_1 surface) position as in Fig. 4.11, and the same colour code is used.

In an attempt to understand the origin of this added stability upon inclusion of non-local interactions, it is useful to probe the electronic distributions on the competing surfaces. Hence, a comparison has been made between the AS-c(6 × 2) phase and its closest competing structure AS-(6×2)_1. The electronic density plots for these two surfaces, calculated using the PBE and
revPBE-vdW-DF1 methods, are shown in Fig 4.11 and Fig 4.17, respectively. The iso-surfaces of the electronic densities of the two structures are also shown in Fig. 4.18.

Figure 4.18: Top and side views of the iso-surfaces of the electronic density, as calculated by PBE (a,b) and revPBE-vdW-DF1 (c, d) methods, of the AS-c(6×2) (a, c) and AS-(6×2)_1 (b, d) structures. Grey and red balls denote copper and oxygen atoms respectively (as in Fig. 4.1), and the translucent blue regions represent the electronic density distribution.

Analysing the charge density in fact reveals “pockets” of electron density between added -Cu-O-Cu- rows in both AS-c(6×2) and AS-(6×2)_1 structures, albeit to a lesser extent in the case of the AS-(6×2)_1 surface as calculated using PBE (Fig 4.18(b)). Our expectation was that existence of such pockets should contribute to additional non-local interactions; these would be lacking in the PBE based calculation, but accounted for in the vdW-DF1 method. However, since these density pockets were found in both systems, for both PBE and PBE-vdW-DF1 calculations, this finding still does not explain why the AS-c(6×2) structure has a lower GFE as compared with its competitors when calculated including non-local interactions. Detailed analysis of the density of states for both systems using both PBE and PBE-vdW-DF1 also reveals no discernible differences that would explain why the AS-c(6×2) is significantly more stable when calculated accounting for non-local interactions.
4.6.6 Summary

To summarise, we have explicitly shown that inclusion of non-local effects is indeed crucial to unambiguously determining the AS-c(6 × 2) phase as the ground-state reconstruction of the Cu(110) surface at higher oxygen exposures. Our analyses reveal that the relative stability/ordering of the AS-c(6 × 2) surface is very sensitive to the bulk lattice constant $a_0$, used: the vdW based methods result in a slight reduction of bulk lattice constant, $a_0$, which appears to be the main reason for increasing the energy gap between the AS-c(6 × 2) phase and its competitors. Furthermore, while a major contributor to its added stability is in the accurate description of the lattice constant, there is very likely other non-local interactions occurring on/within the surfaces themselves.

4.7 Application

Figure 4.19: Minimum energy profiles relative to the initial state of the system, for a super Cu atom diffusing along a $p(2 \times 1)$ row (Y-axis), as indicated by atomistic models (i-iii), calculated using the NEB method, employing various electron exchange-correlation methods, as indicated in the graph legend. Here silver, light grey, dark grey, pink and red balls denote bulk Cu, Cu in ARs, super Cu, O in ARs and high O atoms, respectively, as in Fig. 4.1.
As evidence of the impact these effects may have in the context of general surface science, we also present results of our diffusion calculations, using the VASP built-in NEB algorithm, that were made for a single super Cu atom on the $p(2 \times 1)$ surface moving along the alternating rows of copper and oxygen atoms by one lattice spacing (from state (i) to (iii) in Fig. 4.19). The resulting band is shown in the top part of Fig. 4.19. Using a conventional PBE method (black band) an artificial minimum was found corresponding to an experimentally unobserved geometry shown in Fig. 4.19(ii), with the super Cu atom bonded to two surface copper atoms from two adjacent added rows, surrounded by four (rather than two) surface oxygen atoms. This intermediate structure is separated from both the initial (i) and final (iii) states by a barrier of $\approx 0.55$ eV. However, when recalculating the band, now accounting for non-local interactions with the vdW-DF1 method with the reduced lattice constant (green band), the intermediate point becomes only a shallow dip; also the energy barriers are increased by $\approx 0.2$ eV with respect to the PBE calculated band. The band predicted by revPBE-vdW-DF1 using the reduced lattice constant (red band) has a similar barrier to that calculated by PBE-vdW-DF1 with the reduced lattice constant, however, the intermediate state lies lower in energy, but is still much less stable as compared to the PBE calculation. The shift of the position of the intermediate geometry on the graph is due to change in lattice constant between the different methods. Surprisingly, results reproduced by the PBE-D2 (blue band) and PBE-rescaled (purple band) calculations, also raise the energy of this artificial minimum (although much more modestly) and also significantly reduces the dip, but slightly increases the barriers. In short, all the methods which, in some way, correct for non-local interactions predict the intermediate state (ii) as less stable than the initial and final image, and reduce the barrier to go to and from this state, which the conventional PBE calculation failed to correctly predict. Despite the fact that these methods reduce the depth of the dip, the remaining dip is probably still unphysical. This indicates that although these methods appear to head in the correction direction (in terms of making the correct surface more energetically favourable) they are still far from perfect.

4.8 Conclusion

Since these types of results have not been reported before, this leads to the assumption that DFT fails in this way for a small class of surfaces, possibly only reconstructed surfaces that differ very subtly. In the systems considered here, the surfaces differed only by the number and arrangement of one adsorbed atom (henceforth referred to as ad-atom), of the same atom species as the bulk (super Cu atoms), i.e. there were 1-4 super Cu present, and in the case of the diffusion, the difference was in whether the super Cu atom was bonded to surface oxygens, or other surface copper atoms. These observations have lead us to suggest a computationally cheap
4.8 Conclusion

*ab initio* method of very crudely taking account of some of the non-local interactions for such cases, by employing the vdW calculated bulk lattice constant but performing the calculations using conventional DFT methods (such as the PBE-rescaled calculations that were mentioned here), which we believe may give a much better qualitative description of the system. At the same time, we found that care is needed in applying DFT techniques with vdW corrections as results for the stability of the \( p(2 \times 1) \) phase, which is known to be the most stable at lower oxygen exposures, vary, and it does not always come out as favourable as it should, depending on the method used. In addition, some vdW methods predict larger lattice constants than the GGA based methods, opposite to what one would expect when non-local interactions are accounted for. However, even cases where the vdW methods predict larger lattice constants, they still predict the correct \( c(6 \times 2) \) structure, AS-\( c(6 \times 2) \), as the most favourable. This indicates that some non-local correlation effects must play a role in stabilising the surface in addition to the lattice constant. We believe that this study will stimulate further work in identifying other systems where the non-local interaction, quite unexpectedly, plays an essential role.

Lastly, the surface effects discussed here may play crucial roles in theoretically modelling NC-AFM experiments on the AS-\( c(6 \times 2) \) surface, henceforth referred to as \( c(6 \times 2) \) since it is now surely known to be the correct reconstruction of the surface. Where these effects are important and where they are not will be discussed further in the next chapter.
Chapter 5

Chemical NC-AFM Probe Tip Fingerprinting on the Oxidised Cu(110) Surface

5.1 Motivation

By using the work presented in Chapter 4, a working theoretical description of the Cu(110):O surface was obtained. This was then used to accurately model NC-AFM experiments on the surface, which was in fact our original aim.

5.2 Determining the Tip Apices

5.2.1 Introduction

We present here a NC-AFM study of the oxygen-terminated Cu(110) surface, with $c(6 \times 2)$ islands on the $p(2 \times 1)$ reconstructed surface, the models of which are shown in Fig. 5.1(a) and Fig. 5.1(b), respectively. As we shall propose below, this binary oxidised surface, on which several STM experiments have already been performed [133, 134, 135, 136, 137, 138], may help in solving one key outstanding problem of the SPM, namely identification during experiment the chemical tip identity. Interpretation of experimental NC-AFM images typically requires considerable theoretical input, which is impossible without the clear chemical and structural model for the tip apex. However, whilst it is possible to image tips at the mesoscopic level with e.g. transmission electron microscopy [139, 140], this only helps in evaluating the macroscopic component of the force responsible for the overall tip force magnitude, but not for the lateral resolution; the structure of the actual tip asperity determining the lateral resolution remains an experimentally unknown parameter that can never be known a priori. In many cases this makes determining surface species
and interpreting atomic structures very problematic. Even with the aid of theoretical analysis the tip can only be determined \textit{a posteriori}, if at all, but in real experiments it is most beneficial to know the asperity before and \textit{during} scanning.

![Figure 5.1:](image)

\textbf{Figure 5.1:} (a) Ball model of the \(p(2 \times 1)\) phase (top and side views), where light red, grey and light grey balls depict low O, added Cu, and bulk Cu atoms, respectively. (b) Ball model of the \(c(6 \times 2)\) phase, where additionally dark grey and red balls depict the super Cu and high O atoms.

### 5.2.2 Experimental Setup

All experiments were carried out by a home-built NC-AFM operating under ultra-high vacuum conditions at 78K by our collaborators in Osaka, Prof Yasuhiro Sugawara’s group. The frequency-modulation technique [141] was used to detect the tip-surface force interaction. Commercial silicon cantilevers were oscillated keeping the oscillation amplitude constant at 6.7 nm. The cantilevers typically had a \(Q\)-factor of 150000, spring constant of 40 N/m and resonance frequencies of 150 kHz. Before experiments, Cu(110):O surface was prepared. First, Cu(110) surface was cleaned by repeated Ar ion sputtering and annealed. Then Cu(110) surface was exposed to oxygen. The tip apex was cleaned by Ar ion sputtering \textit{in situ} and was found to have a radius of curvature less than 10 nm. Both the cantilever and the sample were always electrically grounded.

### 5.2.3 Theoretical Model

In order to verify the experimental conclusions, the nature of the interactions between Cu- and O-terminated tips and the \(c(6 \times 2)\) surface were studied theoretically. All theoretical calculations were performed using DFT as implemented in the VASP code. We employed a plane wave basis set and PAW-PP, considering 11 and 6 valence electrons on every copper and oxygen atoms, respectively. Plane waves were chosen for this calculation since the use of the computationally more efficient localised basis sets led to a sizeable, difficult to correct, tip-surface separation dependent basis set superposition error qualitatively modifying the force-distance curves. Electronic
5.2 Determining the Tip Apices

exchange and correlation were described using the GGA as employed by the PBE density functional. In calculations of the bulk Cu, the PBE calculated lattice constant was used, the Brillouin zone was sampled using the Monkhorst-Pack method with a dense \((30 \times 30 \times 30)\) \(k\)-points grid and fractional occupancies were handled employing the broadening approach suggested by Methfessel and Paxton with a width of \(\sigma = 0.2\) eV. For all subsequent calculations, the \(k\)-point sampling was scaled from that of the bulk and a plane wave cut-off of 500 eV was used. Typically forces on atoms were converged to 0.01 eV/Å.

![Cu-Terminated tip and O-Terminated tip models](image)

**Figure 5.2:** Cu- and O-terminated tip models, colour coded as in Fig. 5.1.

Since experimentally the tip is crashed into a Cu-O cluster several times in order to form tip type I or II, as will be explained in more detail below, we assume that a Cu-O cluster terminates the Si cantilever. Therefore, for our tip models, we considered a microscopic Cu-O cluster terminating a spherical macroscopic Si cantilever. To model the O-terminated tip cluster we used a relatively stiff Cu tetrahedron employing an O atom asperity, in which the lowest four atoms were allowed to relax upon interaction with the surface. The same model was used for the Cu-terminated tip with a Cu atom asperity instead. Both tips, shown in Fig. 5.2, consisted of a 10 atom tetrahedral structure with either Cu or O atoms at their apex. Due to our cell size, material, DFT code used, and limitations on available resources, a laborious search through tip-phase space was computationally prohibitive. Instead preliminary calculations were performed on a collection of tip models and the two tips we discuss were found to be the most structurally stable tips. Other tips consisted of the apex bonded to two or four other atoms, however these either involved using much larger tip models or were simply unstable. Given everything, the 10 atom tetrahedral tip model was the best choice. Of course we expect that a much larger tip model would be required for better quantitative agreement. Calculations with the O-terminated and Cu-terminated tips were performed by the author and Dr. Robert Turanský in the group of Prof. Ivan Štich, respectively.

To model the surface a 6-layer slab was used with 78 atoms in the unit cell; the bottom two layers were kept fixed to simulate the missing bulk; the simulation cell is shown in Fig. 5.3. This cell was chosen to provide adequate separation of neighbouring tip structures. The lattice vectors shown in Fig. 5.3 for the slab were chosen so as to maximise the nearest neighbour distance.
5.2 Determining the Tip Apices

Figure 5.3: Schematic of the cell model used in our theoretical calculations. Lattice vectors $a_1 = 8.53$ Å, $a_2 = 14.55$ Å and the angle between them $\alpha = 64.76^\circ$. Plan view of the surface is on the left and side view on the right. Light grey, grey and dark grey balls correspond to bulk, added and super Cu atoms. Light red and red balls correspond to low and high O atoms as in Fig. 5.1.

between tip images for the given number of atoms in the cell. The nearest neighbour distance between tip images chosen was 4.1 Å, which was found to be very reasonable and introduces little finite size effects due to periodic boundary conditions. This was tested against a cell twice as large (extended in the lattice direction $a_1$) with a tip nearest neighbour distance of 10.1 Å, and was shown to have negligible differences in the vertical forces acting on the tip atoms of less than 0.02 eV/Å.

From our previous study on the thermodynamics of the Cu(110):O in Chapter 4, we know that the vdW interaction may play a role in determining the tip-surface interaction, and could potentially affect the orderings of the force above different surface sites. However, we also showed that this effect seems only to be apparent when treating similar but different surface, whereas for this study, we are studying the same surface, simply imaged with different tips. Thus, the vdW interaction is not expected to change the relative ordering of forces over different positions on the surface. For the sake of completeness, preliminary calculations were performed to check this
by comparing PBE against PBE-D2 calculations for different tip-surface separations and lateral positions. The results of this test showed that, although the magnitude of the PBE-D2 calculated forces were approximately twice as large as the PBE forces, the relative orderings above the three most important sites of the $c(6 \times 2)$ reconstructed surface (the super Cu atom, high O atom and hollow site) remained unaltered.

To realistically reproduce the experimental scan lines and spectroscopy curves theoretically, contributions due to macroscopic van der Waals (vdW) interaction between the macroscopic tip and the surface must be included as described in Section 3.14.1. In our calculations the macroscopic vdW contributions were found to be small, this was due to a rather small radius of the sphere used of 25 Å and an offset of +3.2 Å, indicating that for more quantitatively accurate calculations a larger microscopic (atomistic) tip model should be used (as we expected). Note that the value of $H = 1.16$ eV for the Hamaker constant (corresponding to silicon) has been used in our calculations of the vdW contribution to the force. We optimised the tip radius and the offset distance to achieve the best agreement with the experimentally measured scan lines shown in Fig. 5.9, and experimental values were used for all other parameters.

5.2.4 Imaging the Cu(110):O Surface

In the experiments performed by out collaborators, initially a pure silicon tip was prepared and used to scan the surface, producing the image shown in Fig. 5.4(a). In this image both Cu and O atoms are imaged: Cu atoms as bright protrusions and O as dark depressions. We note that the continuous NC-AFM imaging with a pure Si tip termination is extremely difficult as the tip is easily contaminated by surface species (Cu and O atoms). It was noticed by taking a large area image of the oxidised Cu(110) surface shown in Fig. 5.4(b), that there are several Cu-O clusters distributed on the surface. These clusters (indicated by the arrow in Fig. 5.4(b)) are located mainly at domain boundaries, kink and/or step edge sites, appearing as an aggregation of Cu and O atoms after the surface cleaning process. To enforce a particular tip termination, the following procedure was used. The tip was intentionally crashed into one such Cu-O cluster as schematically shown in Fig. 5.4(c). Tip crashes reproducibly led to two different types of images: I and II. Images of both types taken over the same scanning region, consisting of a $c(6 \times 2)$ edge on the $p(2 \times 1)$ surface are shown in Fig. 5.4(d, e), respectively. These clearly differ from the image taken with the pure Si tip, Fig. 5.4(a), showing two different arrangements of atoms with the same super-cell symmetry. In type I image, a single bright spot is seen at the expected positions of the super Cu atoms as highlighted in the schematic of the $c(6 \times 2)$ surface below image (d). Conversely, in type II image, two close spots are imaged instead, located around the
5.2 Determining the Tip Apices

![Image of Si tip, CuO cluster, Cu(110)-O surface, and types I and II.](image)

**Figure 5.4:** Experimental NC-AFM measurements of (a) the pure $c(6 \times 2)$ surface with the Si tip before contamination with the surface species, and (b) a large area of the oxidised Cu(110) surface containing a Cu-O cluster (indicated by an arrow) used to modify the tip apex. The experimental procedure is depicted by a cartoon (c), illustrating the Si tip crashing into a Cu-O cluster on the surface, thereby resulting in tips of either types I or II. Experimental NC-AFM topography images of the pure $c(6 \times 2)$ surface are shown, performed with tip type I (d) and tip type II (e), with the schematic of the top layer of the $c(6 \times 2)$ surface overlaid in each case. At the bottom of images (a, d, e) the corresponding schematic of the $c(6 \times 2)$ surface is given to highlight the imaged atoms.

same position as in type I and hence must be attributed to the two high O atoms which surround each super Cu atom, as highlighted in the schematic below image (e).

To further study these two different tip terminations, a large area of the oxidised Cu(110) surface, shown in Fig. 5.5, was scanned with both tip types. In the left part of the two images one can see a $p(2 \times 1)$ reconstructed surface. It contains equidistant (001) rows of Cu and O atoms characteristic for this phase. One can see that images of this particular phase look the same with both tip types. In the right part of the two images one can see a part of a big island with characteristic $c(6 \times 2)$ reconstruction, while an additional layer of atoms on top of the island can be seen in the far bottom-right corner of the two images; note that we do not know precisely the structure of this additional layer. One can clearly see the characteristic signatures of the tips type I and II in the images, (a) and (b), respectively, of the $c(6 \times 2)$ reconstructed island, namely,
5.2 Determining the Tip Apices

Figure 5.5: Experimental NC-AFM topography images of the same area of the oxidised Cu(110) surface performed using tip type I (a) and II (b). An elongated island of (most likely) Cu and O atoms on top of the Cu(110)-c(6×2) surface is clearly visible in the bottom right corner of the images. A much clearer indication of this restructuring of the Cu-O rows can be seen in Fig. 5.6 in the circled areas as well as at the edges of the c(6×2) island as indicated by the white arrows. Additionally, here the registry of the bright features imaged by both tip types can be seen more clearly, indicated by the black arrows. For the type I tip, the large bright protrusion on the c(6×2) island is out of registry with the protrusions seen on the p(2×1) surface, as shown in Fig. 5.6(a). This is in line with the deduction that the type I tip images only Cu atoms, strongly indicating that the large bright protrusion must be super Cu atoms. It then follows that the protrusions on the p(2×1) surface are the Cu atoms which run along the Cu-O (001) rows. The super Cu atoms appear out of registry because they are bonded to two oxygen atoms each, which are not imaged as bright by this tip. Therefore, the super Cu atoms are out of registry with the Cu atoms in the Cu-O (001) rows and are in registry with the O atoms, which are not imaged. This supports our findings. Conversely, in Fig. 5.6(b) the bright protrusions appear in registry with the bright spots on the p(2×1) surface. This can be explained by the same logic, assuming
that the type II tip images only O atoms, which lie within the Cu-O (001) rows, the high O atoms would naturally appear in registry with the \( p(2 \times 1) \) rows. This shows that each tip indeed only images one species: the type I tip images only Cu atoms and the type II tip images only O atoms. Thus, when scanning the \( p(2 \times 1) \) surface which is made up of equidistant rows of alternating equidistant Cu and O atoms, with both tip types, the topographies appear identical.

**Figure 5.6:** Experimental NC-AFM topography images of the same area of the oxidised Cu(110) surface performed using tip type I (a) and II (b). The circled area in both images shows initiation of the reconstruction of the Cu-O (001) rows from the missing row to the double row/missing row structure. A reduced distance between the Cu-O (001) rows can also be seen at the edges of the \( c(6 \times 2) \) island (indicated by white arrows). Black arrows indicate the registry of \( p(2 \times 1) \) and \( c(6 \times 2) \) features.

To summarise, the two types of images should be attributed to two different tip terminations (I and II) obtained from the pure Si tip when crashing it into a Cu-O cluster. This conclusion is further supported by observations made during scanning, when the tip changes between the two types. A typical example is shown in Fig. 5.7(a). One can see that initially super Cu atoms were imaged (type I), then two tip changes occurred during the same scan resulting in a change of contrast in which high O atoms were imaged (type II). We are sure that the atom appeared due to a deposition from the tip rather than diffusion across the surface, as the barriers, from our preliminary diffusion calculations (see Section 4.7), for super Cu atoms to diffuse are quite high (>0.6 eV), energies which are not freely accessible in the 78 K environment the experiments were performed in. Clearly, as Cu and O are the most abundant species on the surface, these results are highly reproducible (similar results obtained in separate experiments), and since usually no impurities are initially imaged near the location of where the tip is crashed into the Cu-O cluster, one has to assume that tips type I and II correspond to the tips terminated with either Cu or O atoms.
5.2 Determining the Tip Apices

It is essential to point out that these results were gathered from different sets of experiments. Furthermore, there is often no change in contrast even when the tip deposits an atom onto the surface, see Fig. 5.7(b). Here, a type I tip was used as only Cu atoms are imaged. Then undoubtedly a Cu atom, as indicated by arrows, was deposited to the surface from the tip (i.e. a structural change must have occurred in the tip). However, we see that the contrast did not change after the deposition, meaning that the tip apex had also not been chemically altered, i.e. it remained type I (stable before, during and after the manipulation). Therefore, it is not the full structure, but specifically the chemical species terminating the tip that is essential to the contrast, as these results were found to be highly reproducible, with little change in contrast between experiments. Hence, we conclude that the images showing types I and II were obtained with two different classes of tips, terminated either with Cu or O, although in each particular instance the actual structures of the tips having the same termination might have been different.

These conclusions are corroborated by scan lines, Fig. 5.8 and Fig. 5.9, taken along the (1T0) direction for the Si, type I and type II tips. It can clearly be seen than all three scan lines differ significantly from each other. The Si tip images super Cu and high O atoms as bright and dark, respectively. This differs from the tip types I and II, which only image either one or the other species. Type I tip images the super Cu atoms as bright, and the high O atoms appear darker than the hollow sites features, while for the same region, the type II tip images the high O only as bright and interacts significantly less with both the super Cu and hollow site Cu atoms,
5.2 Determining the Tip Apices

imaging them both as darker features.

![Graph and images showing tip height and distance](image)

**Figure 5.8:** (a) Experimental NC-AFM topography scan line of the Cu(110)-c(6 × 2) surface taken with the Si cantilever before it is dipped into surface. The corresponding topography image is shown in (b), where the white line indicates where the scan line was taken. The scan line goes across three missing rows, over the super Cu, high O and hollow sites as shown in the schematic beneath (a). The same colour code is used as in Fig. 5.1. Results were obtained using $\Delta f = -18.8$ Hz.

5.2.5 Determining the Tip Types

Using our theoretical model, we have simulated the topography scan lines along the same direction as in experiment, passing through the O-Cu-O sequence of atoms, using DFT. Comparison of the theoretical and experimental results in Fig. 5.9(a, d) strongly suggests that the type I tip is the O-terminated tip, imaging super Cu atoms as shown in Fig. 5.9(c), and the type II tip is Cu-terminated, imaging high O atoms as shown in Fig. 5.9(f). Indeed, our theoretical analysis shows without doubt that before and after the deposition, shown in Fig. 5.7(b), the tip must have been O-terminated and the deposited Cu atom originated from further along the tip and not from the apex as one might naively assume.

The key feature observed so far was the ability to experimentally modify the tip apex reproducibly and reversibly by dipping it into the Cu-O clusters abundantly present on the surface to produce two distinctly different NC-AFM images of the c(6 × 2) surface. Although this method does seem to ensure only two possible chemical terminations of the tip (either Cu or O), the actual tip structure as well as chemical identity of other species present above the apex atom cannot be controlled by this procedure and hence, one may expect, this may affect the measured NC-AFM image. Surprisingly however, we find that although occasionally there are indeed secondary features present in our images (where the rows running in the (001) direction can be observed), the main characteristics of the images (either bright spots around super-Cu atoms or two bright spots above high-O atoms) remain clearly identifiable. In Fig. 5.10(d) we show images which feature
5.2 Determining the Tip Apices

Figure 5.9: Experimental (black solid curves) and theoretical (red dashed curves) topography scan lines of the Cu(110)-c(6 × 2) surface. (a) Scan lines with experimental tip type I and the theoretical O-terminated tip; (b) the corresponding tip model and (c) the experimental topography image. (d) scan lines with experimental tip type II and the theoretical Cu-terminated tip; (e) the corresponding tip model and (f) the experimental topography image. All scan lines go across three missing rows, over the super Cu, high O and hollow sites as shown in the schematic beneath (d). The same colour code is used as in Fig. 5.1. Experimental results for Type I and II tips were obtained using $\Delta f$ values of -19.5 Hz and and -24.4 Hz, respectively. Theoretical results for O- and Cu-terminated tips were obtained using $\Delta f$ values of -22 Hz and -24 Hz, respectively.

To illustrate, additionally, that the primary features are not affected by the actual tip structure and determined solely by the chemical identity of the tip apex atom, we compare in Fig. 5.10 experimental and simulated spectroscopy, $\Delta f(z)$, curves over the most important surface sites: super Cu atom, high O atom, and the hollow site. In spite of the fact that our theoretical models
5.2 Determining the Tip Apices

Figure 5.10: (a) Experimental (left panels) and (b) theoretical (right panels) spectroscopy curves of the Cu(110)-c(6×2) surface over a super Cu atom (red lines), high O atom (blue lines), and a hollow site (cyan lines) for Type I apex. The theoretical tip model is shown in panel (c). Panel (d) shows the corresponding experimental image with spectroscopic sites indicated by crosses. Note that bright spots around super-Cu atoms have oval-like shapes. Panels (e-h) depict the same information for Type II apex.

are unlikely to be exactly the same as the actual experimentally used tip apices, the gross features of both experimental and simulated contrasts are the same and hence not affected by the possible secondary structural tip differences, i.e. Type I (O-terminated) apex yields super Cu atoms as brightest spots and Type II (Cu-terminated) apex yields high O sites as brightest. We reiterate here that preliminary results indicated a larger tip model is required for softer contact. That is to say, a larger microscopic tip model, better describing the short-range tip-surface interactions by allowing more atoms in the tip to relax, would allow more tip atoms to reach their equilibrium geometries at smaller tip-surface separations. This would yield more accurate descriptions of the tip-surface energies and forces (which by necessity would be smaller as there would be less
5.2 Determining the Tip Apices

internal stress within the tip), and thus better quantitative agreement (at smaller $\Delta f$ values). However qualitatively, the results produced by our theoretical model are in very good agreement with experimental measurements.

So, for tip-surface separations at which NC-AFM imaging is normally done ($\approx 0.2 - 0.3$ nm), our theoretical results agree well with experiment: 1) both Cu- and O-terminated tips image only one of the two surface species each; 2) Cu-terminated tips image a different surface species than O-terminated tips. Moreover, theoretical calculations helped interpretation of the experimental images and provided unambiguous characterisation of the chemical nature of the tip apex. Our results have also indicated that care should be taken when interpreting the outcome of manipulation experiments, as a usually made assumption that atoms deposited on the surface from the tip must have come from the tip asperity may not necessarily be true.

5.2.6 Summary

To summarise, three unique image signatures are found when experimentally imaging the $c(6 \times 2)$ oxidised Cu(110) surface, where, depending on whether the tip apex is terminated with Si, Cu or O, one of three types of contrast is obtained:

i. Si - both surface species are imaged with super Cu appearing as protrusions and high O as holes;

ii. Type I - only super Cu are imaged as bright,

iii. Type II - only high O are imaged as bright.

A Si termination is distinctly different and therefore clearly identifiable from the type I and II terminations. The contrasts obtained by the type I and II tip terminations show different atoms, and this surface allows for switching between the two tip terminations with high reproducibility. These experimental findings are supported by theoretical simulations, which were able to identify the type I and type II tips as being O- and Cu-terminated, respectively.

We find in our NC-AFM study of the structure of the partially oxidised Cu(110) surface that it is solely the chemical identity at the tip apex that is essential to the contrast for this system, the actual structure of the apex is of the secondary importance. Whilst this idea in itself may not be new, we have also found that the contrast for this system can be changed fairly easily with a reasonable degree of control, i.e. one can keep indenting the tip into the surface until the required signature is acquired. Control here is in reference to the primary features of the tip (the end atom), but not necessarily the secondary ones, and this, as we have found, may be enough in many cases. In short, we have found that this system can be used for “fingerprinting” and enforcement of the chemical identity of the tip asperity both before as well as at any stage of
5.3 Contrast Inversion

5.3.1 Introduction

In the NC-AFM study of the Cu(110):O-c(6×2) surface by S. Kinoshita et al. [108], when imaging the surface, the super Cu atoms were always observed as brightest, however, different sub-features were also observed. They report on three contrasts which appear to have the same primary feature, that the super Cu atoms are imaged brightest, however, the secondary features in each contrast appear to differ in their imaging of the alternating Cu and O atoms in the Cu-O rows running along the (001) direction. Thus, in each of these contrasts only the type I tip (O-terminated tip, as we now understand from our current study) was probing the surface, and in our experiments we also observed the ARs running along the (001) direction, although these rows can be resolved much more clearly in the images obtained by S. Kinoshita et al. [108]. Since we have deduced that an O- or Cu-terminated tip must be present to observe Cu or O atoms, respectively, the contrasts observed in that paper suggest that perhaps both O and Cu atoms may have be near the tip apex, or there is some sort of contrast change at different tip-surface separations.

Having successfully studied different types of contrast in NC-AFM measurements of the c(6×2) surface, and having identified the tip terminating chemical species by theoretically analysing experimental results, we theoretically studied some of the more subtle details that can be observed in the contrast, like these secondary features. All experimental data discussed here was obtain by our collaborators, the group of Prof. Sugawara in Osaka.

5.3.2 Contrast Inversion when Scanning the c(6 × 2) Surface

To study the contrast, both experimental and theoretical topography graphs at constant \(\Delta f\) were obtained for a range of \(\Delta f\) values, the results of which are presented in Fig. 5.11. Although it is clear that the tip model used is crude as only the lowest four atoms are allowed to relax, we observe remarkable qualitative agreement between the experimental and theoretical results for the two sets of images (O-terminated and Cu-terminated). In (a) and (b) at small |\(\Delta f\)| values, the “high O - super Cu - high O” chains and the hollow site are imaged as protrusions (bright), with the super Cu atoms appearing as a maximum, while between them are depressions (dark). Then, as |\(\Delta f\)| increases, the hollow site is imaged as darker than the chain. In short, both experimental and theoretical results indicate that for all \(\Delta f\) values the super Cu is imaged as brightest with the O-terminated (type I) tip.
While our theoretical model describes the primary features observed accurately, there are differences from experimental observations. As $|\Delta f|$ increases, our model predicts that the super
5.3 Contrast Inversion

Cu atoms will appear much brighter than the high O atoms, and for larger $|\Delta f|$ values, the high O atoms are predicted to appear darker than both the super Cu atoms and the hollow sites. However, experimentally the type I tip initially images the “high O - super Cu - high O” chain as bright with the super Cu atoms appearing as brightest, then at larger $|\Delta f|$ values, the tip continues to image the chain as brighter that the rest of the surface, but instead of the super Cu appearing clearly as brightest, the high O atoms appear almost as bright. To summarise, for the O-terminated (type I) tip, at smaller $|\Delta f|$ values the tip images Cu atoms as brightest, but as the tip-surface distance reduces both the super Cu and high O atoms are imaged as bright. This result may also give a valid explanation of the different contrasts observed for the type I tip in Fig. 5.9(c), where the super Cu atoms appear circular, and in Fig. 5.10(d), where they appear elliptical and some sub-surface features can also be observed. Similarly, they also help in explaining the three similar but different contrasts by S. Kinoshita et al. [108], as at different $\Delta f$ values either just the Cu atoms appear as brightest at smaller $|\Delta f|$ values, or both the Cu atoms and the O can be imaged at larger $|\Delta f|$ values.

Analysing now (c) and (d), although our model predicts a $\Delta f$ scale for our theoretically calculated Cu-terminated tip $\Delta f$-z curves for the (Fig. 5.10) that is four times larger than is experimentally measured, we observed even better qualitative agreement in Fig. 5.11 than with the O-terminated tip, although there are some subtle differences, particularly at larger $\Delta f$ values (closer to the surface). At small $|\Delta f|$ values, the Cu-terminated tip images the “high O - super Cu - high O” chain as bright, with the super Cu atom as brightest. As the tip-surface interaction increases, the super Cu and high O atoms in the chain are imaged as equally bright. When there is strong tip-surface interaction, the super Cu atoms are imaged as protrusions, while the high O atoms are imaged clearly as bright features. Differences between theory and experiment arise in the imaging of the hollow site, where experimentally they are initially imaged as darker features than the “high O - super Cu - high O” chain, then at large $|\Delta f|$ values they are imaged as features that are equally as dark as super atoms. However, in the theoretically calculated profiles, the hollow sites are initially imaged darker than the chain, but for large $|\Delta f|$ values they are predicted to be imaged much darker than the super Cu atoms. In short, for the Cu-terminated (type II) tip, there is a contrast inversion when imaging the c(6 × 2) surface at difference $\Delta f$ values. Since this is experimentally measured and predicted by theory, this is another strong indication that the tip model is indeed correct, i.e. that the type II tip is indeed the Cu-terminated tip. This helps to explain why, when imaging O atoms with the Cu-terminated tip, either bright circular (Fig. 5.10(h)) or elliptical/oblong (Fig. 5.9(f)) features are also observed where O atoms are expected. Of course, this assumes the validity of our tip model, which is likely to be only an idealised tip model, so the imaging of these features may still very well be due to secondary atoms in the tip. Furthermore, when referring back to the $\Delta f$-z curve Fig. 5.10(f), a crossover can be
5.3 Contrast Inversion

observed in the super Cu and high O curves both theoretically and experimentally, which is the origin of the topography contrast inversion.

It is clear from this last result that interpretation of NC-AFM imaging on the c(6 × 2) surface should be made with caution, as one of two cases can arise, either:

i. the Cu are clearly imaged at small |Δf| values, where they appear as circular, but then at larger |Δf| values both the Cu and O atoms are imaged, appearing elliptical, which corresponds to the O-terminated tip;

ii. the Cu atoms are imaged at small |Δf| values, where they appear as circular, but then at larger |Δf| values the O atoms appear as bright and the Cu atoms appear a dark, which corresponds to the Cu-terminated tip.

This may make tip identification from experiments difficult, however, the distinction between the two can easily be made by measuring at large |Δf|: if only O atoms are imaged, then clearly the tip must be Cu-terminated and if either the “high O - super Cu - high O” chains or only the super Cu atoms are images, then the tip must be O-terminated. However, other alternatives may exist, where the tip apex, for example, may consist of both Cu and O atoms, a Cu atom may be near the apex of a O-terminated tip or vice versa. As such, in the event of an ambiguous contrast, the best method to determine the chemical termination may in fact be to scan the surface at a relatively large |Δf| values, as all ambiguity is eliminated.

5.3.3 Contrast Inversion when Scanning the p(2 × 1) Surface

We already know that the periodicities of Cu and O atoms in the row running along the (001) direction of the p(2 × 1) surface are identical, therefore identification of the chemical species terminating the tip is impossible, however, since most of the experimental measurements were taken on a c(6 × 2) island on the p(2 × 1) surface, it is worthwhile studying whether the same contrast inversion can be observed on the p(2 × 1) surface. This will also act as a check of the validity of our model, since to accurately simulate a scan line on the c(6 × 2) surface, our model must also accurately simulate a scan line on the p(2 × 1) surface. To that end, we have also performed similar theoretical simulation of a scan line on the p(2 × 1) surface, taken along the row of alternating Cu and O atoms in the (001) direction, and have compared our results to an equivalent experimental scan line.

Fig. 5.12(a-c) show the results for the Cu-terminated tip scanning the row, while Fig. 5.12(d-f) show results for the O-terminated tip. In (a) and (d), topography images of the same region of the p(2 × 1) surface with the type I and II tips, respectively, reveal identical topographies, as expected. Similarly, the scan lines shown in (b) and (e) are also identical however, they show
that for both tips, only one species is imaged, measured with a periodicity of ≈0.31 nm, while theoretical calculations predict a periodicity of ≈0.30 nm, in good agreement. We can also observe that the Cu-terminated tip does indeed produce a contrast inversion as $|\Delta f|$ increases: initially imaging the O atoms as darkest followed by Cu atoms at small $|\Delta f|$ values; imaging both Cu and O atoms as equally bright at larger values; imaging O atoms as darkest at the largest $|\Delta f|$ values.

The O-terminated tip, only ever images the Cu atoms as bright, predicting that the O atoms will be imaged ever darker as $|\Delta f|$ increases. While, based on our study of the scan line on the $c(6 \times 2)$ surface (Fig. 5.11) with the O-terminated tip, we expect that these results will be in good agreement with experiments at small $|\Delta f|$ values, we also may expect that for large $|\Delta f|$ both the O and Cu atoms would be imaged as almost equally bright, with the Cu atoms imaged as slightly brighter.

Although we expect a subtle difference between theory and experiments using this simple model, for typical imaging $\Delta f$ values, we observed good agreement between theory and experiments, as both predict that only either Cu or O atoms will be imaged, and that the imaged species
will be different for the Cu- and O-terminated tips. This shows that our model is indeed valid for
typical imaging $\Delta f$ values, and also shows that experimental identification of the tip impossible
on the $p(2 \times 1)$.

5.3.4 Summary

To summarise, we performed calculations to determine the dependence of the topography
of the $c(6 \times 2)$ and $p(2 \times 1)$ surfaces in the tip-surface interaction ($\Delta f$). When imaging the $c(6 
\times 2)$ at different $\Delta f$ values, we observed that different contrasts can be obtained: the tips can
image either Cu atoms, O atoms or the “high O - super Cu - high O” chains as brightest. We then
used these results to explain the different contrasts obtained when imaging the Cu atoms, in both
recent experimental measurements and in those previously reported by S. Kinoshita et al. [108]
circular and elliptical shapes as well as sub-surface features), as being due to measurements taken
by the O-terminated tip at different tip-surface interaction regimes, and also show that in each
case, the Cu atoms are still imaged as brightest. We then went on to show why identification by
scanning the clean $p(2 \times 1)$ surface is not possible.

5.4 Conclusion

In conclusion, we have performed a joint experimental/theorical study of the oxygen-
terminated Cu(110) surface. NC-AFM experiments identified patches of two coexisting phases,
$p(2 \times 1)$ and $c(6 \times 2)$, and the latter was found to exhibit a reversal of the image contrast with
respect to the chemical, as opposed to structural, tip apex identity. Although changes in contrast
after contact with surfaces being scanned have been previously reported (for example in Ref. [143]),
it is clear that the oxidised Cu(110) surface offers some advantages over other systems studied so
far; it allows for much more control over the changes of the chemical species at the tip apex (the
tip class), and also, if required, the apex can be changed to one of the two stable terminations
by locating a Cu-O cluster on the surface and crashing the tip into the cluster. It is still not
possible to change the tip apex to the desired termination, either Cu or O termination is obtained
and there is little control over this, however, our theoretical simulations suggest without doubt
that the knowledge of the chemical nature of the tip asperity can be obtained immediately after
the dipping protocol by either scanning across a small region of a $c(6 \times 2)$ patch or performing
spectroscopy above the three main sites (hollow, high O and super Cu). Although we discuss
contrast changes, our emphasis is on neither the observation of contrast changes due to accidental
(and often frequent) tip changes during scanning, nor the variety of different contrasts, but the
control we have in obtaining these different contrasts. The experimental protocol our collabora-
tors successfully performed was shown to make it possible to fingerprint/enforce the chemical tip
identity of the tip before, during and after scanning. This, we believe, addresses one of the key problems facing the NC-AFM community, namely identification of the chemical nature of the tip asperity during the course of experiment.

Once the desired termination is obtained (NC-AFM topography image is seen), measurements or controlled manipulations of this surface can proceed with clear knowledge of the tip apex. In the case of tip change (which can easily be recognised experimentally in most cases by a sudden change of contrast), one should still be able to tell the tip apex termination. If the termination does not correspond to any of these three known signatures (e.g. a foreign contaminant is present on the apex), then one of the two terminations (Cu or O) can be enforced again by performing the crashing protocol. Surprisingly, although the concept is not new, to our knowledge the procedure proposed here has not been vigorously pursued in the past.

Alternative methods for identifying the tip have been previously proposed, such as: determining the tip polarity by scanning [144] or performing spectroscopy measurements [145] on an ionic system, performing a laborious search in the “tip phase space” to reproduce the results of NC-AFM [143] or STM/NC-AFM measurements [30], ramping the bias voltage to probe the tip chemical identity [146] and performing frequency shift vs. bias measurements to verify if the tip is metallic [147]. However these method can be either too cumbersome, limited and may not allow for stable and reversible tip enforcement. More importantly, the mentioned methods either require substantial theoretical effort, which can only be exercised after the actual experiment, or provide only qualitative information about the tip apex structure (e.g. polarity). To control the tip termination when performing NC-AFM imaging, impressive results have been achieved by functionalising the tip apex [148] to guarantee the known tip termination, however, this can frequently be unstable when performing manipulation experiments. Furthermore, while several methods exist for ensuring a single species termination, the protocol shown here allows for a reasonable amount of control over switching between two different terminations in NC-AFM experiments.

In addition, we believe that this particular surface may have even wider utilisation and appeal. Firstly, we propose that the Cu(110) surface, partially oxidised, be used as the main surface or substrate for nano-manipulation and building nano-devices, e.g. by adsorbing molecules onto either the clean or oxidised Cu(110) surface [133, 135, 149, 150, 151, 152]. This is because Cu-O clusters are commonly available on it after surface preparation [153], and so are islands of the p(2 × 1) missing row phase with some super Cu atoms in the missing rows, deforming the nearest two O atoms either side of them as in the c(6 × 2) phase [108, 111, 153]. Hence, the tip identification protocol successfully used in this study, may be performed at any stage of the experiment using the oxidised part of the surface. Secondly, significant progress has been recently reported at the TRNM V conference (2012) by P. Liljeroth et al. who, coincidentally, performed an NC-AFM experiment in which a tip was prepared on the Cu(111) surface, before transferring it to the...
actual surface of interest inside the same (cold) chamber \cite{154}. In the experiments performs by our collaborator on this surface, when imaging at larger tip-surface separations (i.e. at imaging height, not manipulation height), the tips were found to be very stable as well as being able to take highly reproducible measurements. In fact, even when performing manipulations with these tips their contrasts were mostly found to remain unchanged. Although extremely challenging, this method, if and when successful, may open new avenues in NC-AFM experiments, and then the oxidised Cu(110) surface we believe could be a good choice for a reference system for controlling the chemical nature of the tip apex when scanning other surfaces.

For the surface studied here, the key feature required for it to serve as a reference is the ability to distinguish between the atomic arrangements of the metal and oxygen atoms. Since this is the only criterion, identification of a much larger class of reference systems of this type with a super metal atom in a missing row or having other characteristic arrangements of several atomic species, which may also show variable signatures for different tip terminations, may also be possible. For instance, these could be binary oxidised metal surfaces \cite{155} such as oxidised Cu(100) \cite{109, 156}, Ag(100) \cite{156}, Ag(110) \cite{125}, Ag(111) \cite{157}, Ni(110) \cite{110, 158, 159}, and Au(110) \cite{126}, which have either a missing row reconstruction (on which there may be super metal atoms), or distinct and different arrangements of metal and oxygen atoms.

As interesting as the observation of contrast inversion at different tip-surface interaction strengths with the Cu-terminated tip is, which also helps to explain the different contrasts that may be obtained with the Cu-terminated tip (e.g. circular vs. elliptical), the result itself is not new. In fact, contrast inversions have been observed several times, particularly on the Si(111) \cite{160, 161, 162} and TiO$_2$ \cite{163, 164} surfaces, as well as other surfaces \cite{165, 166, 167}. The inversion in our case is likely due to the chemical reactivities of the Cu-terminated tip. This, combined with the relative bond strengths and equilibrium bond lengths of Cu-Cu and Cu-O bonds (as the Cu-Cu bond length is larger than for Cu-O, but the bond strength is weaker), could very possibly lead to the crossover observed in the $\Delta f$-$z$ curves for the Cu-terminated tip that we observed. This hypothesis can be tested by examining more Cu-terminated tip models (more and less chemically reactive), but is outside the scope of this study.

We hope this study will stimulate more research in the directions proposed to investigate other surfaces with similar properties, which may serve as convenient references. Also, we believe advances are required on the instrumentation side which would enable performing controlled scans at any point during experiment of a reference system. The protocol and a reference system designed for the chemical tip identification/enforcement as outlined here may well be just a first step in the effort to build a whole database for chemical tip fingerprinting. As an example of a study that can be performed using this technique, the work presented in the next chapter, about Cu manipulation on the oxidised Cu(110) surface, was largely based on this.
Chapter 6

NC-AFM Manipulation of Oxidised Cu(110) Surfaces

6.1 Motivation

Now that imaging of the Cu(110):O-\(c(6 \times 2)\) has been studied in great detail, both experimentally and theoretically, and a good understanding of topography contrasts on the surface and an adequate theoretical model has been obtained, we studied manipulation of super Cu atoms on the Cu(110):O surface.

6.2 Introduction

We present here a NC-AFM study of manipulation of super Cu atoms on the Cu(110):O-\(p(2 \times 1)\) surface. Experiments have been performed, in which super Cu atoms, initially randomly located near a \(c(6 \times 2)\) island on the \(p(2 \times 1)\) surface, were controllably manipulated to form the letter ‘X’ on the \(p(2 \times 1)\) surface. These manipulations consisted mostly of vertical manipulations, but lateral manipulations were also observed. Vertical manipulation with NC-AFM has already been studied in great detail on several different surfaces, and the mechanism is generally well understood to be either an interchange mechanism [168], where the tip exchanges an atom with the surface, or the more common combination of removal and deposition (the nano-indentation method) [169, 170, 171]. We focus only on lateral manipulations here, leaving vertical manipulation investigations for future work.

Several experimental and theoretical studies of lateral manipulations with NC-AFM have previously been performed [19, 23, 172, 173], which so far fall into one of two categories: pulling [173], in which the surface atom diffuses in the lateral direction towards the tip, and the more common tip-induced barrier reduction [19, 172], in which the barrier for the surface atom to diffuse or be
manipulated reduces due to the presence of the tip. In the latter case the atoms may move in any direction, including towards or away from the tip. In a sense the pulling mechanism can be thought of as a sub-group of tip-induced barrier reduction, as in both cases, it is the presence of the tip that allows the atoms to diffuse. Given the structure of the $p(2 \times 1)$ surface along the (001) direction (that there are no physical obstructions preventing diffusion), the mechanism for lateral manipulation along the (001) rows can likely be easily understood as tip induced. For a detailed description of a tip induced manipulation by barrier reduction, the author refers the reader to Chapter 8.

This leaves only lateral manipulation in the (110) direction, across the Cu-O rows, which we focus on because the mechanism for this type of manipulation is not so clear, as the super Cu atom must jump over a continuous row of alternating Cu and O atoms. While this manipulation may also be due to a tip-induced reduction of the energy barrier to diffuse, most of the previous studies on the mechanisms behind lateral manipulation were limited to semiconducting or ionic insulator surfaces, so it is nonetheless still worthwhile to study manipulation on this non-ionic insulator surface.

6.3 Theoretical Model

The same theoretical model that was described in Section 5.2.3 was used, differing in that to model the surface, a $(6 \times 3)$ extension of the $p(2 \times 1)$ reconstruction consisting of five layers was used, as shown in Fig. 4.19.

In all cases where a minimum energy path (MEP) or an energy barrier was calculated, NEB calculations were performed. Calculation of all the minimum energy paths (MEP), which give the energy barriers, were performed with the VASP built-in NEB algorithm. As with the previous study, calculations with the O-terminated and Cu-terminated tips were performed by the author and Dr. Robert Turanský in the group of Prof. Ivan Štich, respectively. For such calculations, the initial geometry was taken from a fully relaxed tip approach or retraction calculation, and the final geometry was approximated and relaxed, then checked. Since the NEB method is only used to calculate part of the diffusion at a time, typically either only 3 or 5 intermediate images were used to represent the band, which was considered relaxed when the total force acting on the band was less than 0.01 eV/Å. For each NEB calculation, the tip was in a fixed position, the tip position was only changed between calculations. Furthermore, the reaction co-ordinate chosen to describe the transition path is the $x$-axis lateral position of the Super Cu atom. Hence in places where the curves appear to go back on themselves, this is actually describing the position of the super Cu atom (moving back or forth). Here it is worth once again saying that the calculation of one MEP gives both the forward and reverse transitions.
Although we have tested that to calculate the force spectra on the \( c(6 \times 2) \) surface, vDW does not need to be accounted for, this was likely because the same surface was studied. However, when performing NEB calculation, difference systems are modelled (the band images), and as shown in Fig. 4.19, vDW does indeed need to be accounted for in such cases. For this study, the flavour of vDW corrected DFT we used was PBE-D2, since we used large cell sizes, which makes the vDW-DF methods computationally prohibitive.

### 6.4 Creating the Letter ‘X’

![Figure 6.1: NC-AFM topography scans of the \( p(2 \times 1) \) surface being manipulated. A random distribution of super Cu atoms, (a), is vertically and laterally manipulated, forming an intermediate arrangement (b), finally forming the letter ‘X’ from seven atoms, (c), indicated by the black lines.](image)

Our collaborators, Prof. Yasuhiro Sugawara and his group, using the same experimental set up as described in Section 5.2.2, performed NC-AFM experiments on the \( p(2 \times 1) \) surface in the vicinity of a \( c(6 \times 2) \) island, where there was an abundance of super Cu atoms. In the example given in Fig. 6.1, super Cu atoms were randomly distributed in a small region, as shown in (a). These atoms were then either manipulated away from the area or to form the letter ‘X’, as shown in intermediate image (b), in order to finally form the letter ‘X’ in (c). Although these manipulations were performed near a \( c(6 \times 2) \) island, the production of the letter ‘X’ is clearly done only on the \( p(2 \times 1) \) surface. Since in these manipulations, the Cu atoms are imaged on the \( p(2 \times 1) \) as well as the edge of the \( c(6 \times 2) \) island, the tip performing these manipulations can confidently be taken as being O-terminated, however, calculations with both tip terminations were performed to avoid ambiguity.

### 6.5 Conventional Methods vs. vDW Corrected Methods

Although we initially assumed that the atom being manipulated was a Cu atom, it is necessary to be certain of this. So, to establish exactly what the chemical species of the atom
6.5 Conventional Methods vs. vdW Corrected Methods

Figure 6.2: Model (colour coded as in Fig. 5.1) illustrating Cu deposition on the \( p(2 \times 1) \) surface, shown with respect to the Cu and O surfaces, as is imaged by the O and Cu-terminated tips, respectively. Examples of corresponding NC-AFM topography images are shown next to the schematic and the super Cu atom feature, as imaged by the respective tips, is highlighted by a blue circle and enlarged.

being manipulated is, we again carefully study the contrast of the NC-AFM images.

In Fig. 6.2, we show two typical contrasts observed with the O- and Cu-terminated tips, type I and II, respectively. As in Chapter 5, the O-terminated (type I) tip imaged a bright spot, centred between four other bright spots, and the Cu-terminated (type II) tip imaged two bright spots, in registry with the \( p(2 \times 1) \)-(001) rows. Since we know that Cu and O atoms are the most likely chemical species of the ad-atoms that produce these image features, there are only four possible combination that could give rise to the contrasts of these features, shown in Fig. 6.3.

Calculating \( \Delta G \) for these four possible variations using PBE, the results of which are plotted in Fig. 6.4, shows that structures (a) and (b) are clearly the most thermodynamically favourable out of the four, in a wide range of the O chemical potential, meaning that the imaged ad-atom is indeed Cu, i.e. super Cu and not an O adatom. At this point, upon directly comparing the (a) and (b) structures to the NC-AFM images, it can clearly be seen that only (a) matches the experimental observations in Fig. 6.2. However, the plots presented in Fig. 6.4, which were calculated using conventional DFT methods (i.e. 0 K, PBE calculation, without considering vdW or vibrations), determine structure (b) as most favourable, lower in energy than structure (a) by \( \approx 0.05 \) eV. Here it is worth re-emphasising that structure (b) has never been experimentally observed.

Since structures (a) and (b) are quite similar, differing only in the location of the super
Figure 6.3: Geometry models (colour coded as in Fig. 4.5) of the possible ad-atom reconstructions of the \( p(2 \times 1) \) surface that could give rise to the features imaged in Fig. 6.2: (a) super Cu atom located between two O atoms (the expected geometry); (b) super Cu located between two Cu atoms; (c) O atom located between two O atoms; (d) O atom located between two Cu atoms.

Figure 6.4: Gibbs free energy difference plots for the possible ad-atom reconstructions of the \( p(2 \times 1) \) surface that could give rise to the features imaged in Fig. 6.2. The plots were calculated using PBE without considering vibrations. The brown (horizontal) line corresponds to the clean/bare Cu(110), the green line corresponds to the \( p(2 \times 1) \) surface with a single super Cu atom bonded to two O atoms ((a) in Fig. 6.3), and other lines correspond to variations of the ad-atom surface, labelled as in Fig. 6.3.

Cu atom, one might think vdW interactions also need to be accounted for. For this reason \( \Delta G \) was re-calculated for the same four structures, now accounting for vdW (PBE-D2 and PBE-vdW-DF1), shown in Fig. 6.5. The results show that upon inclusion of vdW interactions, structure (b) is in fact found to be significantly less favourable (by \( \approx 0.19 \) eV as calculated by PBE-D2 and by \( \approx 0.57 \) eV as calculated by PBE-vdW-DF1) than structure (a). This once again shows that when studying structures with similar but still different geometries, vdW may also play a role in determining the ground-state (most stable) structure. This result explains the artificial minimum
6.6 Diffusion of Super Cu Atom Across an Added Row

observed in the diffusion calculation shown in Fig. 4.19. Thus, all further calculations will be performed with PBE-D2, which, although is semi-empirical, gives results qualitatively consistent with the vdW-DF results, as we have seen from Fig. 4.19.

6.6 Diffusion of Super Cu Atom Across an Added Row

![Figure 6.6: Schematic depicting how the lateral tip position is represented. The model is colour coded as in Fig. 5.1.](image)

The MEP calculations were performed for different tip positions, where all lateral positions are given with respect to the super Cu atom in the “high O - super Cu - high O” chain, as indicated in Fig. 6.6; the x-axis runs along the “high O - super Cu - high O” chain, the y-axis runs perpendicular to x, within the surface, and the z-axis runs out of the xy-plane. The initial position of the tip is (0 Å, 0 Å, 5 Å), which corresponds to the tip positioned directly above the super Cu atom, with an unrelaxed tip-surface separation of 5 Å. Unless otherwise stated the y position is always 0 Å.

6.6.1 Without a Tip

With our models fully set up and tested, we studied diffusion of the super Cu atom across the alternating Cu-O rows, which run along the (001) direction. To that end, we performed NEB calculation for the Cu atom to diffuse across the row without the presence of the tip, and the result is shown in Fig. 6.7. This MEP indicates that the super Cu atom must overcome a large energy barrier of \(\approx 2.1 \text{ eV}\), where it travels directly over an O atom. The band was also calculated for the Cu atom to travel over the mid-point, slightly laterally displaced, between a Cu and O atom in the row instead of directly above an O atom; however, the barrier for that transition
6.6 Diffusion of Super Cu Atom Across an Added Row

Figure 6.7: Model (colour coded as in Fig. 5.1) depicting the initial (i) and final states (ii) of the transition of a super Cu atom across an added row (an alternating Cu-O row running along the (001) direction), and the corresponding minimum energy profile for the transition, given relative to the energy of the initial state.

was larger as the Cu atom bonds to both the Cu and O in the row, and then breaks the bond to the Cu atom in order to reach the final state. The result of the minimisation of other energy paths studied was the same as shown in Fig. 6.7, indicating that the Cu atom must indeed travel directly over the O atom in the row, a transition with an energy barrier that at 78 K renders it very unlikely. This means that diffusion without the tip (self-diffusion) of the super Cu atoms across the row is highly unlikely to occur.

6.6.2 In the Presence of the O-Terminated Tip

The MEP was then recalculated for different lateral position of the O-terminated tip, and the results are shown in Fig. 6.8. The MEPs calculated in (a), where the tip is positioned asymmetrically with respect to both the initial and final geometries, correspond to the tip forcing the super Cu atom out of its current position and into the next (the forward transition) and “pulling” the atoms towards it (the reverse transition). While in (b), where the tip is positioned symmetrically, both the forward and reverse transition correspond to “pulling”. In Fig. 6.8, we observed that for none of the tip positions we have studied does the energy barrier reduce to a value surmountable at 78 K\(^1\). Even for very small tip-surface separations, when the tip is positioned above the super Cu atoms, (a), the energy barrier for super Cu to break its bond with O atom 1 reduces, but to travel over O atom 2, the barrier remains at \(\approx 2\ \text{eV}\). Further calculations were performed to the tip positioned above the O atom 2 at small separations from the surface (small \(z\)), and laterally displaced at \(y = 0.9\ \text{Å}\) and \(y = 1.0\ \text{Å}\). These positions were thought to allows the

---

\(^1\)By surmountable we mean that the transition rate corresponding to that temperature and energy barrier is large enough that the transition is likely to occur within a timescale on a similar order of magnitude as the experiment duration.
6.6 Diffusion of Super Cu Atom Across an Added Row

Figure 6.8: The minimum energy profile for the transition of the super Cu atom to evolve across the alternating Cu-O rows, given relative to the energy of the initial state, for different O-terminated tip-surface separations (given in legends) and lateral positions.

Cu atom to travel around O atom 2 instead of directly over it, however, the barrier to break the bond with O atom 1 remains insurmountable, >1 eV. We therefore concluded that tip-induced barrier reduction with the O-terminated tip does not appear to allow for the manipulation of the super Cu atom across the Cu-O rows.

6.6.3 In the Presence of the Cu-Terminated Tip

After Prof. Ivan Štich’s group performed similar calculations with the Cu-terminated tip, shown in Fig. 6.9, we found that, in (a), although as the tip-surface separation reduced the barrier for the super Cu atom to break its bond with O atom 1, for all tip-surface separations the barrier for the super Cu atom to travel over oxygen atom 2 remained ≈2 eV, and is thus insurmountable at 78 K. This is the same result as we found for the O-terminated tip. However, for the Cu-terminated tip positioned at $x = 2.5$ Å, (b), almost directly above O atom 2, the energy barrier reduces significantly, down to 0.8 eV (from the lowest to highest energy along the reaction path) for the tip positioned ≈3.9 Å away from the surface and displaced in the $y$-direction by 0.9 Å from O atom 2. The energy barriers here correspond to the super Cu atom breaking away from O atom 1 or 3, for the forward and reverse transitions, respectively. While the energy barriers are much lower, they are still too large for the super Cu to overcome at 78 K.

The author notes the frequent use of the term surmountable. To help quantify this, by using harmonic transition rate theory, what is considered a surmountable barrier is one that gives a calculated transition rate on the same order of magnitude as the inverse of the time the tip
6.6 Diffusion of Super Cu Atom Across an Added Row

Figure 6.9: The minimum energy profile for the transition of the super Cu atom to evolve across the alternating Cu-O row, given relative to the energy of the initial state, for different Cu-terminated tip-surface separations (given in legends) and lateral positions.

spends above the “high O - super Cu - high O” region. Using this system as an example, taking a rough estimate of the NC-AFM raster speed of 10 nm s$^{-1}$ and a “high O - super Cu - high O” chain length of $\approx 1.5$ Å (c.f. Fig. 5.9), the amount of time the tip spends in that region is $\approx 0.015$ s, requiring an energy barrier of at least $\approx 0.20$ eV to make the transition likely to occur within that time.

6.6.4 Tip Assisted Transition - ‘Carrying’ the Super Cu Atom

From Fig. 6.9 and Fig. 6.8, we deduced that the transition for the super Cu atom to travel across a Cu-O row could not occur via tip-induced barrier reduction (or the pulling mechanism), since for all tip positions studied, for both tips, a barrier $>0.8$ eV was always present, which is insurmountable at 78 K. We therefore again reviewed the barrier for the transition to occur without the tip (Fig. 6.7). We noticed that if the super Cu atom could be brought to the position directly above O atom 2 (or near that position), and the Cu atom did not remain bonded to the tip upon full tip retraction, then statistically, there was a high probability that super Cu atom could then relax in the final state (it could also obviously relax back into its initial state). Combining this with our observations in Fig. 6.8(a) and Fig. 6.9(a), where, at small tip-surface separations above the super Cu atom, the barrier for the super Cu atom to break its bond with O atom 1 significantly reduces, we constructed a different path for the transition to occur. Along this path, rather than the tip removing then depositing atoms, interchanging atoms or simply
reducing the barrier for the transition, the tip would ‘carry’ the Cu atom to the transition state, then upon retraction the system would be able to relax into the final state (manipulated state). If this mechanism is valid, then this would be a new mechanism for manipulation.

Figure 6.10: Cartoons illustrating the schematics for transition of the super Cu atom across the Cu-O row by: (a) tip-induced barrier reduction, where the tip-surface separation does not change; (b) tip assistance. Grey and red shapes denote Cu and O atoms, respectively.

The schematic depicting the mechanism for conventional lateral manipulation is shown in Fig. 6.10(a), where the presence of the tip, at small enough tip-surface separations, reduces the barrier for the atom to diffuse, making the transition statistically more likely. In this mechanism, the tip is not required to move during the manipulation. However, as discussed, this method does not appear to work for our system. Therefore, we have performed calculations for our tip-assisted mechanism, which is illustrated in Fig. 6.10(b). Here, the vertical movement of the tip is crucial,
as depicted; the tip makes a close enough approach to the surface to reduce the barrier for the tip to ‘pick up’ the Cu atom, but instead of removing it from the surface upon retraction, it shares the Cu atom with O atom $2$ ($v$), then as the tip fully retracts ($vii$), the Cu atom can either relax back into its initial state ($viii$) or into a new state (our final state), shown in ($ix$). Since, if the manipulated atom is taken away by the tip, it would be a vertical manipulation, for this mechanism to work, it would have to occur over a single tip oscillation cycle. The calculations, corresponding to Fig. 6.10(b) were performed as follows:

1. the tip was brought into close proximity to the super Cu atom, and the system was allowed to relax at each different tip-surface separation ($i$-$ii$);

2. at a tip-surface separation of $\approx 1.9$ Å, the energy barrier for the Cu atom to bond to the tip and O atom $1$ was then calculated using 5 images, since due to the Cu-O bond length, it is highly unfavourable for super Cu to bond to O atom $1$ ($iii$), i.e. there is much larger energy barrier;

3. the tip was then retracted from the surface, and at each tip-surface separation the system was relaxed ($iv$);

4. at a tip-surface separation of $\approx 2.1 - 2.3$ Å, the energy barrier (calculated using 5 images) for the Cu to break its bond to O atom $1$ and to instead bond to O atom $2$ was calculated ($v$);

5. the tip was then retracted until the Cu atom was near the transition point, when the tip-surface separation is $\approx 4.8$ Å, relaxing the system at each separation ($vi$-$vii$);

6. then lastly, the energy barrier (calculated using 7 images) for the Cu atom to relax into the final state was calculated ($ix$).

These calculations were initially performed with the O-terminated tip, and the barrier associated with step $2$ is less than 0.1 eV, which is surmountable at 78 K. Since the tip retractions simply involve geometry relaxations, they do not have any associated barriers. For step $4$ the barrier was calculated to be less than 0.15 eV, which is again surmountable at 78 K. For the final step, for the Cu atom to reach the final, manipulated, state, the energy barrier was calculated to be negligible ($< 0.01$ eV). This clearly shows that this manipulation mechanism is indeed viable, as all of the calculated barriers are less than 0.2 eV, which at 78 K is not difficult for the system to overcome.

These calculations were repeated for the Cu-terminated tip by Prof. Ivan Štich’s group, however, at small tip-surface separations, this 10 atom tetrahedral tip model used (see Fig. 5.2...
6.7 Summary

Figure 6.11: Model of the relaxed geometry of the Cu-terminated tip in close proximity to the surface, colour coded as in Fig. 5.1.

was found to be highly unstable, resulting in the tip depositing the apex atom, a Cu atom, onto the surface. The geometry of this is shown in Fig. 6.11.

6.7 Summary

To summarise, we present experimental images of manipulations of super Cu atoms on the \( p(2 \times 1) \) surface near a \( c(6 \times 2) \) island, where a random distribution of atoms are formed into the letter ‘X’. Manipulation of super Cu atoms on the \( p(2 \times 1) \) surface appear to be highly reproducible as several atoms are manipulated in order to produce the letter ‘X’. We then performed theoretical calculations of the transition of super Cu atoms across the alternating row Cu and O atoms running along the (001) direction, that form the added row structure on the \( p(2 \times 1) \) surface, lateral manipulations. Our calculations reveal that known lateral manipulation mechanisms fail to describe the transition of super Cu atoms across these rows, for both the O- and Cu-terminated tips. As a result, we then go on to propose, describe and show the validity (plausibility) of a new mechanism, that allows for the transition with small energy barriers.

6.8 Conclusion

While tip-induced barrier reduction and “pulling” mechanisms work in describing lateral manipulations of atoms on a variety of systems, and may very well also explain the diffusion of
super Cu atoms along Cu-O rows on the $p(2 \times 1)$ surface, these methods do not appear to explain how a super Cu atom would diffuse across, which requires the Cu atoms to travel over O atoms. However, the new mechanism proposed here, a tip-assisted transition, which has shown to be at least plausible from our calculations, does allow for this diffusion. Although there is currently not enough experimental data on the lateral manipulations to verify these findings. This mechanism, which involves the tip “carrying” the super Cu atom to the transition state, if shown to be the experimental mechanism, would mean a completely new mode of manipulation. If this is indeed the case, it would aid in understanding lateral manipulations of atoms in system consisting of rows, where the atom is manipulated across (over) a row. It should be noted however, that within this mechanism, there is a small chance that the manipulated atom, when it is at its highest position, directly between the tip and surface, may be taken away by the tip upon retraction. This would simply result in a vertical manipulation, removal of the atom.

We hope this research will stimulate experiments and theoretical investigations into lateral manipulation across row, where there are physical barriers (a wall of atoms) that need to be traversed over, as little work has been done on this mode of manipulation. Furthermore, we hope this will also stimulate studies to prove the applicability of this new mechanism (or disprove), and if proven correct, will also provide new insight into lateral manipulation experiments of atoms across systems consisting of rows.
Chapter 7

Tip and Surface Effects in NC-AFM Imaging of Si(001)

7.1 Motivation

Now moving on to a different surface, a series of studies probing the instabilities and low vibrational frequencies (soft modes) of the Si(001) bucked-dimer surface is presented. Initially, the role of the surface and different tips used in NC-AFM measurements of dissipation were studied, as previous experimental results had indicated that there was a complex interplay between both the tip and surface in producing the dissipation contrast. Thus, here a detailed theoretical study on the complex architecture that may be behind observed dissipation signals is presented, and we shall show that our main conclusions are broadly supported by experimental results.

7.2 Introduction

7.2.1 Previous Dissipation Studies

Damping in non-contact atomic force microscopy (NC-AFM) corresponds to energy dissipation in the cantilever-sample system (the amount of additional energy required for the oscillating cantilever to maintain a constant amplitude at resonance.) This phenomenon has been most effectively explained by the adhesion hysteresis mechanism, whereby atomic relaxations in the junction (i.e. both in the tip and sample) on approach and retraction differ [94, 95, 96, 97, 98, 99], leading to the tip and surface atoms following two different trajectories and hence to the tip force hysteresis (two different force vs. distance, $F$-$z$, curves). This is thought to be the main mechanism, although the debate about the mechanism is still somewhat open [26, 96, 100, 101, 174].

Typically dissipation of the order of 0.01–1 eV per oscillation (henceforth referred to as p/o) cycle is experimentally measured at stable imaging conditions [98, 99], but signals exceeding
7.3 Theoretical Model

1 eV p/o (and up to 10 eV p/o) have also been reported [175, 176]. Recently, theoretical studies have made tremendous contributions to our understanding of what causes dissipation and in answering the question of why and how it provides atomic contrast. However, these studies have been mostly limited to either ascertaining the dominant mechanism behind dissipation (e.g. one or a few atoms of the surface jumping up to the tip then returning to the surface upon tip retraction) or interpreting dissipation signals. Dissipation effects were also discussed in the context of manipulation with the AFM probe [28, 168, 177, 178]. At the same time, there has been little discussion on the architecture behind the individual signals that are averaged over in experimental measurements because of the tip oscillations, and the complexities related to contrast changes as a function of the tip-surface distance for systems where no atomic depositions or removals occur. Having said that, there has been a discussion on dissipation behaviour as a function of the tip-surface distance [179], presented in the context of pulling a chain of atoms from the tip, however, this was a study of a completely different type of system. Moreover, no study has previously been done at the near-contact regime alongside a discussion as to whether this can still lead to reproducible atomic contrast both in topography and dissipation.

7.2.2 The Buckled Dimer Si(001)-c(4 × 2) Surface

We study the Si(001) surface [180, 181, 182, 183, 184], chosen due to its known dissipative behaviour when scanned using NC-AFM. Although it is now generally accepted that its low temperature (< 120 K) ground state geometry is the buckled dimer c(4 × 2) phase [19, 185, 186, 187, 188, 189], the p(2 × 2) phase is also observed in local surface areas [189]. A possible channel by which dissipation is thought to occur in this surface is primarily due to the flipping of these dimers [187]. We shall see that this is generally true only for not very close approaches; more complex dissipation mechanisms involving the atoms at the tip apex as well are at work if the tip comes very close to the surface, especially in the near contact or even contact regimes. In our study we use silicon tips, effectively avoiding effects due to different atomic species. Thus, this system allows for thorough scrutiny of only the mechanical tip-surface processes that occur.

7.3 Theoretical Model

Using DFT we meticulously examine the nature of dissipation up to the near-contact regime. We focus on tracking the changes in dissipation signals as a function of the tip-surface separation and lateral position of the tip, revealing the complexities behind these changes. We go on to characterise the atomistic processes that can lead to extremely large dissipation signals (>1

---

1The c(4 × 2) phase corresponds to the anti-symmetric buckling of dimers (only on dimer atom is images per dimer) across and along adjacent rows, while the symmetric p(2 × 1) phase corresponds to symmetric buckling (i.e. both dimer atoms are imaged).
7.3 Theoretical Model

eV) in great detail. Then based on preliminary experimental results, which show sharp contrast and discontinuous intensity changes in the profiles of observed dissipation signals, we qualitatively confirm our theoretical conclusions.

To study dissipation in this system, we have performed extensive \textit{ab initio} DFT calculations carried out using the localised basis set code SIESTA, utilising periodic boundary conditions, double-zeta polarised basis set which described the valence electrons using 13 orbitals and NC-PPs with an orbital-confining cutoff radii of 15 meV (chosen after testing different values). We chose to use SIESTA here as it is sufficient for describing a semiconductor surface and performs the calculations faster than VASP, which was chosen for studies of metallic materials. We used the PBE density functional, a kgrid\_cutoff of 6 Å (SIESTA parameter, c.f. Section 3.8.3.3) to describe the bulk and a single Gamma ($k = 0$) k-point sampling to describe sample the slab. The surface was modelled by an eight-layer slab with a cell containing eight dimers arranged in two rows with four dimers in each row; the bottom four layers were frozen in the bulk geometry and unsaturated bonds at the bottom of the slab were terminated with hydrogen atoms; the four uppermost layers of the slab were allowed to relax. Preliminary comparison tests between this slab and one in which 6 layers were allowed to relax were performed and no significant differences between results were observed, hence this slab was used for the main study. Two tip models (shown in Fig. 7.1) were used: (i) a relatively stiff standard 25 atom pyramidal Si(111) cluster with a single dangling bond [188, 190], where only 4 apex atoms were allowed to relax (H3 tip), and (ii) a much more flexible dimer tip [189, 191] containing 48 atoms, where 15 atoms at its apex were allowed to relax. In all our calculations, forces on the atoms which were free to relax were converged to 0.01 eV/Å. The distance between the tip apex atom (closest to the surface) and the surface prior to relaxation was used to measure the distance between tip and surface (the tip-surface separation).

![Si H3 Tip](image1)

![Si Dimer Tip](image2)

\textbf{Figure 7.1:} Si H3 (a) and dimer (b) tip models. Turquoise and white balls depict Si and H atoms, respectively.
In order to be able to calculate the topography and dissipation spectroscopy curves for different frequency shifts when the closest tip-surface separation is variable, the following approach was undertaken. Initially the unrelaxed distance between the tip apex and the upper dimer atom beneath the tip was set to 0.7 nm. Using increments of 0.01 nm, this separation was then reduced to a minimum distance of 0.02 nm, corresponding to the very near-contact regime, and then retracted back to 0.7 nm using the same increments. Depending on the lateral tip position a separation of 0.02 nm may lead to tip-surface bonding (contact). After each change of the tip-surface separation, the system was allowed to fully relax. Such small increments were chosen to accurately capture any point where deviations occur between the approach and retraction $F$-$z$ curves. Retraction calculations were performed in a very specific way in order to capture all possible hysteresis loops. Where hysteresis curves occur can be seen upon comparison of the approach and retraction $F$-$z$ curves, therefore, the tip-surface separation where the retraction curve deviated from the approach curve (if this occurs) was noted. Then starting from the geometry corresponding to the next tip-surface separation larger than this value on the approach curve (which was at a slightly larger tip-surface separation) a retraction was started again. This ensures that upon retraction the approach curve will be followed until the next point of deviation (if any). This process of retraction was done until the tip-surface separation was 0.7 nm so that all deviations were accounted for. By performing the calculations in this way, we were able to calculate all the complete sets of hysteresis loops in the $F$-$z$ curves for a number of sites along the line across a dimer.

In order to calculate the frequency shift curves and correspondingly the scan lines, which can be compared with experiment, a contribution due to macroscopic interactions with the surface (the vdW force) was added to the DFT calculated tip force using a spherical tip model [38, 39], as explained in Section 3.14.1. Note that an average tip force over the whole oscillation cycle, not just on its approach as it is usually done, was used in calculating the frequency shift in cases when the tip force demonstrated a hysteric behaviour.

### 7.4 Experimental Setup

The experiment was performed by Prof. Sugawara’s group in Osaka, with a home-built NC-AFM operated in ultra-high vacuum (UHV) at a low-temperature [192]. The frequency-modulation technique [141] was used with a commercial silicon cantilever as a force sensor, with 40 N/m spring constant and 171 kHz resonance frequency. In order to remove the native oxide layer and any contamination on the tip, the cantilever was cleaned by Ar$^+$ ion sputtering, and hence the tip apex is likely to be terminated by dangling bonds. As a sample, an As-doped n-type Si(001) wafer was used. All measurements were performed at 5 K. Topographic images were measured from the feedback signal for the distance control between the tip and the sample at a constant
frequency shift $\Delta f$ and at a constant oscillation amplitude $A=13.5\ nm$. Dissipation images were obtained simultaneously with the topography during the scanning. Both the cantilever and the sample were always electrically grounded to 0 V.

7.5 The H3 Tip

We start by discussing the results of our theoretical calculations with the stiffer tip. A complete set of approach and retraction simulations performed as described above is shown in Fig. 7.2, for the tip positioned above a site close to the lower dimer atom (position 5 in Fig. 7.3(b)).

As shown in Fig. 7.2(a), if the tip closest approach, $d$, during oscillations is not closer than 0.26 nm, it follows the same path upon retraction producing no dissipation, as there is not deviation between the approach and retraction curves up to this value. Closer approaches to about 0.17 nm, Fig. 7.2(d), results in a different retraction path up to 0.5 nm (a hysteresis loop). Therefore for tip oscillations with $d$ anywhere between 0.17–0.26 nm, a small but fixed dissipation is observed, corresponding to a dimer flipping event, as depicted in Fig. 7.2(e,f).

It should be noted that in our static zero temperature calculations, when the tip is removed after flipping the dimer, a three-in-a-row conformation of the surface dimers remains, composed of three Si dimers flipped in the same way. This is not observed experimentally and has previously [19, 189] been explained by the fact that the three-in-a-row structure is mechanically unstable and the system eventually returns to the original conformation of alternating dimers characteristic for the $c(4 \times 2)$ reconstruction even at 5 k. This is also discussed in Chapter 8. Therefore, as soon as the dimer on a pristine Si(001) surface flips, it will flip back upon tip retraction. Although this is not captured in our static calculations, this feature affects neither the results nor their interpretation.

For $d$ values between 0.06–0.17 nm, Fig. 7.2(g), a second hysteresis loop is observed which is due to the dimer being forcefully flipped back as the tip apex pushes back down the lower dimer atom, Fig. 7.2(h,i), leading to greater dissipation. At this point the tip bonds to several surface atoms and is noticeably distorted. At near-contact with the surface, with $d$ between 0.02–0.06 nm, a third and significantly larger hysteresis loop is revealed, Fig. 7.2(j). Here the repulsion due to the lower dimer atom, now pushed back into the surface, increases, forcing it to go further into the surface, Fig. 7.2(k,l). Hence we see two dimer flipping events for a single approach. Upon retraction, the tip is completely restored and large dissipation occurs as a result of reversing all of the processes described, breaking several bonds. Interestingly, if retracted from any distance between 0.02–0.06 nm, the second hysteresis loop was found to join the third.

Note that the specific behaviour of the dissipation signal, which remains the same over some interval of tip-surface distances and changes abruptly beyond that, results in a step-like
Figure 7.2: Left panels: four pairs of theoretical approach (black) and retraction (red) $F$-$z$ curves for different tip-surface separations, corresponding to the H3 tip positioned as in (b), slight away from the lower dimer atom (position 5 in Fig. 7.3(b)). The curves correspond to different minimum tip-surface separations on the approach, before the retractions were made, of: (a) 0.26 nm, (d) 0.17 nm, (g) 0.06 nm, and (j) 0.02 nm. Right panels: for each set of curves the corresponding geometries for characteristic tip-surface separations on the approach and retraction (indicated by arrows on the left panels) are shown. Models are colour coded as in Fig. 7.1 and orange balls depict the two dimer atoms beneath the tip.

dependence of dissipation on the tip height (at zero temperature). This peculiar behaviour is due to a finite number of hysteresis loops and different critical turning points for each of them.
Figure 7.3: Theoretical results for the H3 tip: (a) Dissipation signal as a function of tip-surface separation for seven points along a dimer, as indicated in (b) using the same colour code; (c) Dissipation scan lines across a dimer for constant $\Delta f$ values ranging from -15 Hz to -44 Hz, i.e. corresponding from ‘very weak’ to ‘extremely strong’ tip-surface interaction strengths. Inset is the corresponding topographic scan lines for the same frequency shift values, employing the same colour code. The numbers in the dissipation spectroscopy (a) and all scan line images (c) corresponds to the lateral positions numbered along the dimer shown in (b). All balls in (b) represent Si atoms; red and orange balls depict the upper and lower dimer atoms, respectively.

These types of calculations were performed for seven lateral positions taken along a line passing through a dimer, as shown in Fig. 7.3(b). Full results of these calculation, for all lateral positions, are shown in Fig. 7.4. In our simulations, upon retraction of the tip, both the tip and sample nearly always regain their original geometries. The only time this was not the case was for the tip positioned above the upper dimer atom (position 3) for tip-surface separations closer than 0.03 nm, where the upper dimer atom was removed from the surface. The system however may heal itself during subsequent oscillations, but this was not investigated. This shows the tip and sample are typically not permanently deformed, indicating the results are reproducible over many oscillations, as they should be to produce clear experimental images. For some lateral positions,
Figure 7.4: Theoretical $F_z$ curves for seven lateral positions of the H3 tip which correspond to the positions in Fig. 7.3(b) along the dimer row. Approach curves are given in black, retraction curves from the smallest tip-surface separation are given in red, and all other tip-surface retraction curves are as specified in the legends.
at near-contact, the atomic forces were large enough to displace the tip-terminating atom; once
displaced, the lower dimer atom springs back up, and due to the presence of the rest of the tip,
the dimer immediately goes into its flipped conformation, i.e. in some cases up to three dimer
flipping events occur during a single oscillation. The general trend that can be deduced from
these calculations is that the tip interacts more strongly with the lower dimer atom than the
upper dimer atom at smaller tip-surface separations, forming and destroying more bonds. This
results in more dissipation measured over the lower dimer atom region for a wider range of $\Delta f$
values than over the upper dimer atom region.

Dissipation spectroscopy curves using an oscillation amplitude of 13.5 nm are shown in
Fig. 7.3(a) for the seven lateral positions, numbered in (b). It can immediately be seen, that for
the region including the lower dimer atom (points 4-7), not only does the onset of dissipation occur
at larger separations than for the other lateral positions 1-3 (corresponding the the area around
the upper dimer atom), as we would expect (due to its affinity for flipping), but the magnitude
increases in a discontinuous, step-like, manner, as more processes begin to contribute to the total
dissipation signal as was explained above.

The scan lines in Fig. 7.3(c) is an alternative method of representing the data in Fig. 7.3(a),
best suited for a comparison with experimental images, where results for the microscopic tip have
been incorporated into the macroscopic tip model as described in Section 3.14.1. This allows
both the topography and the dissipation scan lines to be plotted as functions of the tip-surface
separation. These two-dimensional maps of these plots are shown in Fig. 7.5, where the overall
trends discussed as tip-surface interactions increase can be seen more easily.

To further aid comparison with experiment (see Section 7.8) in Fig. 7.3(c) we have sim-
plified our qualitative discussion by distinguishing five important regions of tip-surface separation:
‘very weak’, ‘weak’, ‘strong’ and ‘very strong’ tip-surface interactions, as well as an ‘extremely
strong’ interaction region corresponding to the contact/near-contact regime (described in the cap-
tion). Over these regions theory predicts that for large separations (very weak interactions, dark
green curve), almost no dissipation is measured and the topography shows the ground-state surface
(Fig. 7.3(c), inset), as expected. Then for reducing separations, we observe increasing dissipation
above the lower dimer atom (turquoise curve), gradually extending over the entire lower dimer
atom region (purple, red and black curves).

7.6 The Dimer Tip

Our theoretical model based on the stiff tip predicts a very large (over 3 eV) dissipation
signal for some lateral positions if the tip makes a close approach. This must be an overestimation
due to a rather crude tip model in which only the four bottom atoms were allowed to relax. At
close approaches more atoms of the tip must be affected which should yield softer reaction of the tip during contact with the surface and, hence, smaller dissipation. There are clear experimental indications [140] that Si tips terminated with a dimer are also realistic tip models. To verify this, calculations were also performed for a single lateral position using another much larger dimer tip containing 15 bottom atoms allowed to relax. This was done in collaboration with Samuel Jarvis from Prof. Phillip Moriarty’s group in Nottingham who performed many of the calculations with the dimer tip. The results are shown in Fig. 7.6 for tip apex positioned above the upper dimer atom. Note that various tip orientations of the dimer tip with respect to the surface dimer are possible and these would affect the ability of the tip to manipulate the surface [193]; only one such position is considered here, shown in Fig. 7.6.

At large separations the tip ‘feels’ the presence of the dimer, but does not flip it, and so any retraction made from a minimum distance $d$ of 0.25 nm will not lead to any hysteresis in the $F$-$z$ curve. Reducing the separation past this, to $d = 0.07$ nm, leads to a small but significant hysteresis loop. This loop occurs due to the flipping of the surface dimer, which flips back upon
Figure 7.6: Left panels: three pairs of theoretical approach (black) and retraction (red) $F_z$ curves for different tip-surface separations, corresponding to the dimer tip apex positioned above the upper dimer atom (position 2 in Fig. 7.3(b)). The curves correspond to different minimum tip-surface separations on the approach, before the retractions were made, of: (a) 0.25 nm, (d) 0.07 nm and (g) 0.02 nm. Right panels: for each set of curves the corresponding geometries for characteristic tip-surface separations on the approach and retraction (indicated by arrows on the left panels) are shown. The colour scheme is as in Fig. 7.2.

tip retraction. In the case of the pyramidal tip, the flipping of the dimer when the tip is placed above the upper dimer atom/region was due to the very stiff apex forcing down the upper dimer atom(s). However, in the case of the dimer tip, due to the orientation of the tip, the surface dimer flips due to the upper dimer atom of the dimer tip apex bonding to the surface lower dimer atom. Then upon retraction, the surface dimer flips back to its original conformation due to the apex lower dimer atom being above the surface upper dimer atom (which was flipped into the lower
position on approach), see Fig. 7.6(e-f). The formation/destruction of the bonds between the tip dimer apex and the surface dimer are shown as two sharp kinks in the \( F-z \) curve, Fig. 7.6(d).

A second a considerably larger hysteresis loop occurs for any retraction made from \( d < 0.07 \) nm. This results in considerably more dissipation being measured, and is due to the dimer apex being split apart by the surface dimer, shown in Fig. 7.6(i). However, upon retraction of the tip, the dimer apex comes back together and the tip is fully restored. Therefore in this case dissipation occurs as a results of both the tip and surface atomic rearrangements, as well as the creation and destruction of tip-surface bonds.

### 7.7 Comparison of Tip Models

![Graph showing comparison of dissipation signals for two tips](image)

**Figure 7.7:** Theoretically calculated dissipation signals as a function of tip-surface separation for the two tips (pyramidal, black, and dimer, red), where the apex is initially laterally positioned slight away from the upper dimer atom (position 3 in Fig. 7.3(b)).

Direct comparison between the two tip models for the dissipation signal as a function of the tip height for the lateral position above the surface upper dimer atom is shown in Fig. 7.7. It can immediately be seen that at separations above 0.07 nm our larger tip model predicts a small but noticeable dissipation above the upper dimer atom whereas the small tip model predicts no dissipation at these distances, as follows from our discussion above and Fig. 7.3. At the same time, in the near contact regime \( (d < 0.07 \) nm) dissipation calculated with the dimer tip becomes very large (over 1 eV p/o), which is a result of large tip and surface deformations at small tip-surface
separations. Importantly, however, the larger and much more flexible dimer tip also predicts large but significantly smaller dissipation and softer contact (smaller tip force) than the stiffer pyramidal tip model - a result to be expected.

7.7.1 Tip Deformation

![Graph showing force vs. tip-sample separation for approach and retraction.](image)

**Figure 7.8:** Left panels: two pairs of theoretical approach (black) and retraction (red) $F$-$z$ curves for different tip-surface separations, corresponding to the dimer tip apex positioned slightly away from the lower dimer atom (position 5 in Fig. 7.3(b)), as in Fig. 7.2. The curves correspond to different minimum tip-surface separations on the approach, before the retractions were made, of: (a) 0.20 nm and (d) 0.02 nm. Right panels: for each set of curves the corresponding geometries for characteristic tip-surface separations on the approach and retraction (indicated by arrows on the left panels) are shown. The colour scheme is as in Fig. 7.1.

Although in most of our simulations both tips were completely restored upon retraction for a majority of the lateral positions, there are some positions where the dimer tip loses an atom (deposits an atom onto the surface). An example of such a case is shown in Fig. 7.8 for the dimer tip apex at lateral position 5, where upon retraction from a small tip-surface distance, one of the atoms from the dimer termination of the tip breaks its bonds to the tip and is deposited onto the surface. This is shown by a kink in the $F$-$z$ retraction curve from a tip-surface separation 0.44 nm onwards, see Fig. 7.8. However, it may be the case that the tip is in fact restored (the deposited...
atom returns to the tip) if it is “dipped” into the Si(001), but this has not been explored in our study. This occurs for the dimer tip but not for the H3 tip (which in one case picks up a surface atom), most likely due to the fact that the dimer tip has more instabilities, and more atoms were allowed to relax, i.e. due to its flexibility, whereas the H3 tip is much more rigid. Furthermore, these deformation (deposition/removal) only occur where the tip and surface have bonded. It is also important to note that the curves show only tip-surface distance relative to the tip far from the surface, where there is no deformation, the real tip-surface separations are masked. In the cases where deposition/removal occur the absolute tip-surface separations (including stretching of both the surface dimer and the tip apex) is closer to the Si-Si bond length, which corresponds to a tip-surface bond. Nonetheless, in a large majority of our simulations, the dimer tip is completely restored upon retraction.

### 7.7.2 Hysteresis

As seen from the detailed geometries along the approach-retraction cycles shown in Fig. 7.2 and Fig. 7.6 for the H3 and dimer tips, respectively, their geometries are not the same at the same height above the surface on approach and retraction for the values of $d$ smaller than some critical height. This means that, as mentioned above, part of the calculated dissipation is due to the tip-related hysteresis alone. To characterise this effect qualitatively, we extracted the vertical displacements of both tip apices and surface dimer atoms relative to their ideal positions (i.e. taken from their initial relaxed geometries when there was no tip-surface interaction), and presented them as a function of the tip-surface separation. This was done using the relaxed positions of tip and surface atoms from our full calculations along the approach-retraction cycle. We expect that since the dissipation is primarily due to the tip inducing surface dimer flipping and the creation/destruction of tip-surface bonds, the tip apex atom would experience hysteresis for the same range of values of $d$ as the surface dimer. The results are shown in Fig. 7.9 and Fig. 7.10 for the pyramidal and dimer tips position slightly away from the upper dimer atom (position 3 in Fig. 7.3(b)), respectively.

Interestingly, observations of the role the tip plays in the dissipation effects can be seen when comparing the two tips. Much larger displacements occur within the more rigid pyramidal tip-surface system than within the dimer tip-surface system. Furthermore, two hysteresis loops form within the dimer tip-surface system (Fig. 7.10(a,d)), corresponding to when the lowest apex dimer atom bonds to the uppermost surface dimer atom (first loop at a larger tip-surface separation) and when the apex upper dimer atom bonds to the surface lower dimer atom (second loop). This, along with the observation that the magnitude of the hysteresis loops for the dimer tip-surface system is clearly significantly smaller than for the pyramidal tip-system, falls well
in line with our expectations. In fact, the description of the mechanism behind the dissipation provided by this representation can be thought of as an alternative method to our $F$-$z$ curves (considering only the major contributors to the dissipation), however, we now also simultaneously

Figure 7.9: Vertical displacements of the tip apex and surface dimer from their ideal positions as a function of tip-surface separation for the pyramidal tip, where the tip is initially positioned as in Fig. 7.7.

Figure 7.10: Vertical displacements of the tip apex and surface dimer from their ideal positions as a function of tip-surface separation for the dimer tip, where the tip is initially positioned as in Fig. 7.7.

in line with our expectations. In fact, the description of the mechanism behind the dissipation provided by this representation can be thought of as an alternative method to our $F$-$z$ curves (considering only the major contributors to the dissipation), however, we now also simultaneously
get some insight into the contribution of the tip to the overall dissipation. Thus, it follows from these results that the relaxation at the tip likely also contributes significantly into the overall dissipation effect. This is especially so for the dimer tip as hysteresis in the displacements of the tip apex is very similar in magnitude to those of the surface. This is not so much the case for the pyramidal tip-surface system where the tip displacements are much less than in comparison to the surface, albeit still a rather significant portion.

Surprisingly, theoretical calculations reveal that dissipation effects are very sensitive to the specific lateral position of the tip; it is much more difficult to induce a non-conservative distortion at the surface in stiff areas (like around the upper dimer atoms studied here), whereas softer regions (around lower dimer atoms) are much more susceptible to this kind of response. At the same time, if soft tips are more prone to non-conservative distortion than rigid ones, the actual dissipation energy may in the cases of rigid tips be still very significant as these are able to substantially deform the surface underneath.

The discussion presented above is qualitative, based solely on geometry. It is much more difficult (if at all possible) to give an accurate quantitative estimate of the portion of the dissipation energy due solely to the tip and hence it is much more difficult to localise the primary source of the non-conservative interaction. In other words, while it is certainly true that dissipation occurs due to the presence of the tip, there are contributions to the energy that may be from the tip, surface or both.

7.8 Preliminary Experiments

These predictions of theory were tested against preliminary experimental results performed at 5 K and presented in Fig. 7.11. Checks were also made to ensure that any dissipation measured was not just an artefact of instrumentation error with no physical significance. These experiments and checks were performed by the group of Prof. Yasuhiro Sugawara.

We show results for four tip-surface interaction strengths, ranging from very weak to very strong. Likewise with our theoretical calculations, we analyse the NC-AFM images and scan lines taken along a dimer. When moving from left to right in Fig. 7.11, $|\Delta f|$ increases, so the average tip-surface distance decreases. In the case of very weak tip-surface interactions, when no manipulation of the surface by the tip is to be expected, shown in Fig. 7.11(a), the $c(4 \times 2)$ phase with zigzag pattern is observed in the topographic image (i, ii), with the upper dimer atom imaged higher than the lower dimer atom, corroborating that this phase is the ground state for this surface, while there is no information in the dissipation image (iv, v).

In the case of a weak tip-surface interaction, Fig. 7.11(b), the $c(4 \times 2)$ phase is clearly observed in both topographic and dissipation images, although these image patterns are different;
Figure 7.11: NC-AFM topography (i, ii) and corresponding dissipation (iv, v) images and scan lines (taken across the blue line in the images), respectively, of the Si(001) surface taken at 5 K, with a cartoon depicting the cross-section of surface on which the scan line is taken (iii). Scans of the same area were made sequentially without a change of tip apex, at frequency shift $\Delta f$ values of -10 Hz (a), -20 Hz (b), -22 Hz (c), and -30 Hz (d). Red arrows indicate where significant dissipation is measured. The directions in which the strength of the tip-surface interaction changes is schematically indicated at the bottom.

the upper dimer atom still appears higher than the lower dimer atom in the topography scan line, whereas the lower dimer atom is imaged in the dissipation signal. This means that the contrast in the dissipation image shows an inversion to that in the topography, corresponding to the tip interacting to the lower dimer atom of the surface upon its approach, resulting in the dimer
Interestingly, when the tip-surface interaction becomes stronger, Fig. 7.11(c), bright lines along dimer rows with flicker noise (due to flipping of surface dimer atoms induced by the tip-surface interaction) are observed in the topographic image [187], when an average of both dimer atoms is observed in the corresponding scan line. The flicker noise can be identified as apparent jumps in the image, bright spots appearing as misaligned with the main bright line features. In the dissipation image, however, the c(4 × 2) phase is still observed; one can also see that dissipation increases above the lower dimer atom trough, indicating that the lower dimer atom region is imaged rather than the lower dimer atom alone. This is a rather unusual behaviour, when completely different image patterns were resolved simultaneously in topographic and dissipation images: the observed contrast in the dissipation image strongly differs from that in topography.

Finally, in the case of very strong tip-surface interaction, Fig. 7.11(d), the p(2 × 1) phase is observed in both topographic and dissipation images. Most remarkably, in comparison to the other scan lines, there is another sharp increase in the signal intensity, and dissipation of ≈1.4 eV p/o is observed above both dimer atoms.

As strong evidence for the reproducibility of these experimental results, Fig. 7.12 and the proceeding sets of additional topography and dissipation images demonstrate that large dissipation
signals were observed in several experimental runs taken at different times and, most likely, with different tips. At the small frequency shift used in Fig. 7.13, only small dissipation ($\approx 0.2$ eV) that provides no contrast is produced. The same result is seen in Fig. 7.11 for what is very likely an atomically different tip (since the images were taken at different times). Notice an increase of the dissipation signal in Fig. 7.14, where large dissipation occurs ($\approx 0.8$ eV) in the region where topography is lowest (very likely the lower dimer region), indicating the location of the instability and producing an inverted image as compared to the topography. This same result is seen in the set of images in Fig. 7.11 corresponding to $\Delta f = -20$ Hz, and can be seen to be due to flipping events in that region. Also, note that the flipping event seemed to occur frequently in the bottom
7.9 Summary

part of the image. The dissipation contrast is weak in the right side of the image in Fig. 7.15. However, as in Fig. 7.14, taken at a somewhat smaller frequency shift, a rather large dissipation signal (>0.8 eV) was measured, producing an image with a completely different contrast from the topography. As with Fig. 7.13 and Fig. 7.14, the trends observed and explained in this chapter do not seem to change very much between these measurements taken at different times, even though these images were likely produced with different tips (different not only from the one used in Fig. 7.11, but also from each other). In the final reiteration of this point, emphasising the reproducibility of our finding, in Fig. 7.16 very large dissipation is noticed between 1.2 and 1.6 eV, once again in line with Fig. 7.11.

Direct comparison between experiment and theory may be misleading due to the lack of knowledge of the actual tip structure used in any given experiment and because of this one cannot expect exact agreement. Although this is indeed the case here, we do note that the dimer tip model has a promise of delivering closer agreement with experiment in the current case than the stiff pyramidal tip. For instance, experimental dissipation was observed to initially increase above the lower dimer atom by a little more than 0.1 eV, before finally increasing by a further 1.1 eV. Comparing the results obtained with the dimer tip we find not only the qualitative agreement, but also quantitative, as initially dissipation increased by 0.1 eV, before a further 1 eV increase. Also, the absolute value of the dissipation energy observed (around 1.4 eV) is much closer to the value calculated with the dimer tip (just over 1.0 eV) than with the pyramidal tip (over 3.0 eV).

7.9 Summary

In summary, we have described not only the processes, which may lead to extraordinarily large dissipation signals, but we have also detailed the complexities of their underlying architectures. We find a complex network of interactions, where increasing numbers of bond formations and destructions are observed as the tip makes its approaches and retractions. Typically the tip interacts with a lower dimer atom, causing the dimer to flip. At smaller separations the newly flipped dimer is forced back into its original conformation as repulsion between the tip and upper dimer atom increases. At this point, several bonds form between the tip and surface dimer. In some cases the tip apex is displaced out of plane with the dimer, freeing the strained atoms beneath it, allowing the dimer to flip back into its flipped orientation. These formed bonds then break upon retraction of the tip, generally leading to very large dissipation. Furthermore, we also predict discontinuous contrast and intensity changes, which are experimentally observed and therefore confirmed. These features (in most cases) are due to fully reversible processes involving a single dimer flipping multiple times. These processes, being reversible, result in multiple hysteresis loops on the $F$-$z$ curves. These hysteresis loops encapsulate different amounts of dissipation,
which are measured at different tip-surface separations, increasing as the separation decreases. This causes dissipation signal signatures for a given frequency shift to depend critically on how many of these elementary processes occur, resulting in the contrast and intensity of the measured dissipation to vary widely.

7.10 Conclusion

These experimental results confirm that dissipation does indeed initially occur over the lower dimer atoms, spreading over the entire lower dimer region at smaller tip-surface separations. In fact, as is explained above, repeating the experiment can easily verify this particular point, and in our case the results are highly reproducible (see Figs. 7.13 - 7.16). Moreover, on the theoretical side, two quite different tips have been used in our calculations, one also containing a soft mode (the dimer tip), and in both cases we have got qualitatively similar results, which broadly agree with experiment. In other cases instabilities arise due to atoms of the surface being attracted by the tip (atoms jumping on the tip and back), and these cases have previously been presented in literature (e.g. Ref. [176]). Our results show that dissipation increases in a fairly discontinuous manner (jumps in magnitude) between different $\Delta f$ values at different lateral positions. They also demonstrate that huge dissipation, well in excess of 1 eV, may also be measured, and crucially shows that the surface atomic structure and contrast patterns in topographic and dissipation images strongly depend on the tip-surface interaction, and may differ for the same frequency shift as in Fig. 7.11(c). In the theoretical topography scan lines obtained with the stiff tip, inset in Fig. 7.3(c), the general trend observed in our experiment, Fig. 7.11, is correctly predicted, apart from very close separations where, at variance with experimental observations, only one peak may be visible for the two dimer atoms. We believe this is most likely related to the crude tip model used, so a full calculation with the bigger tip would be highly desirable. To be certain of this, one would need to perform calculations using different classes of tips (larger, double dimer, larger irregular clusters etc...). Unfortunately, even the number of different classes are many and calculation of the full scan line with just one much bigger dimer tip is computationally beyond our reach right now (it would require a great deal of time and resources). However, this research has shown that dissipation measurements may provide us with valuable information about the dynamic response of the surfaces, adsorbed molecules, surface defects, etc... on the external stimuli. This can then be used, in conjunction with theory, to gain better understanding of their structures and mechanisms of energy dissipation in these systems, particularly those like the Si(001) surface which have instabilities in the surface. In addition to this, for surfaces with soft modes, manipulations can be performed by scanning the surface at small tip-surface separations,
simultaneously obtaining information about the manipulated surface from the change in frequency shift and information of the pre-manipulated surface from the dissipation signal.

This stems from the fact that, by measuring topography, which is characteristic of the equilibrium surface geometry, one can make a picture of the static surface, but by also measuring dissipation, one can obtain the dynamical response of the surface to the external probe that is the tip. Only specific surface modes respond to this, where there are instabilities, and the dissipation measurements indicate where instabilities exist on the surface - which is very different information to that given by topography itself. The Si(001) surface represents a textbook example of a genuinely unstable system which is prone to instability and hence dissipation.

It is widely accepted now in the AFM community, that dissipation can only be observed in systems which contain internal degrees of freedom characterised by soft modes. Generally, dissipation can be due to instabilities in both the surface and tip, however, if reproducible results are obtained from different experiments, which necessarily means using different tips, then the dissipation is not due to the tip (or rather mainly not due to the tip), and hence is representative of the surface itself. Thus, our results strongly indicate that dissipation is generally not mainly due to the tip.

Our calculations correspond to zero temperature; clearly, the predicted step-like behaviour is expected to be smoothed out in a room temperature experiment by thermal motion [95, 99, 168, 178, 194]. This is because thermally activated jumps occur even with a barrier present, and this stochastic behaviour would be statistically averaged over several hundred oscillations, steps in the dissipation spectroscopy curves would appear smoother [95] in comparison with those predicted by the zero temperature theory; the magnitude of the dissipation signal would also be reduced [95, 194]. Moreover, even at 5 K this smoothing out effect is to be expected as well, albeit in a lesser extend, as every measured point corresponds to averaging over hundreds of tip oscillations, and each of those oscillation cycles will be slightly different than the other. This can be modelled however, if enough calculations could be performed on the surface that a tip-surface force field (describing different tip heights and lateral positions) could be constructed. Then the vAFM could be used (as it was in the past, c.f. Ref. [178]) and theoretical images could be produced that directly mimic the experimental images, including the stochastic distortion of tip trajectories.

We also note that, although our results compare reasonably well with available experimental data for weak to strong tip-surface interaction strengths, our tip models are still relatively simple. For strong tip-surface interactions, a larger more flexible tip model would likely achieve a better description, leading to even softer contact with the surface as indicated by our dimer tip. In particular, at small tip-surface distances, dissipation effects are caused by both tip and surface, with the relative contributions depending on the stiffness (or softness) of either of them. Of course,
a real tip used in actual experiments may well be very different to any theoretical model. However, such tips, much more realistic, are computationally very demanding. Moreover, it is likely to be extremely difficult, if at all possible, to find the ‘true’ tip model for each and every experiment. At the same time we do believe that the general trends revealed in this study will remain qualitatively unchanged for a wide variety of possible tips.

The results presented here provide an extremely detailed, although qualitative, understanding of dissipation on the Si(001)-c(4 × 2) surface at low temperatures. We hope this work will stimulate further research into atomistic processes happening during NC-AFM imaging approaching the near-contact regime. The further research could include a quantitative investigation, where averaging over oscillations is also taken into account, to allow for a more direct comparison between theory and experiment. One could also research into the differences (if any) between using a more flexible tip (48 atom without a dimer termination) and a dimer tip in order to determine how much of flexibility the dimer contributes to the 48 atom tip model.
Chapter 8

The Role of Defects in Controllable NC-AFM Manipulation on Si(001)

8.1 Motivation

In 2006, a protocol was suggested by Prof. Lev Kantorovich [188] that would allow the dimers covering the Si(001)-c(4 × 2) surface to be manipulated into predetermined arrays by scanning across the surface in particular directions with the NC-AFM. Since, at this point, we have in-depth knowledge of instabilities on the Si(001) surface, which when combined with previous studies, such as the one by Y. J. Li et al. [187], have allowed us to locate the lateral and vertical tip position above the surface in order to induce flipping of surface dimers, we are able to test this protocol. After this was attempted experimentally by our collaborators, the group of Prof. Phillip Moriarty in Nottingham, we used theoretical calculations to find the mechanism behind these manipulations. This chapter gives a report of our findings, the results of which are published in Physical Review Letters [19] and Physical Review B [189].

8.2 Introduction

In this study, we introduced a new approach, “flipping a switch”, which is at the atomic level. This involves toggling one or more bistable atoms by way of forming and destroying a single covalent bond between an NC-AFM tip and the Si(001) buckled-dimer surface. Although a multitude of different atomic manipulations have recently been performed using dynamic force microscopy (DFM) [92, 168, 169, 173], the ability to controllably switch back and forth between two modes of an atomic scale switch, has until recently been a mode of control elusive to us. This
mode allows for the modification of the geometric structure through the use of mechanical energy only, also allowing for the modification of the electronic structures, both achieved without the necessity to laterally displace, completely remove, or deposit surface atoms or ad-atoms. Another motivation behind this study is to achieve a further advancement in the field of data storage; there have been logical schemes proposed for memory units based on silicon dimers [195, 196, 197]. Though long sought after, the realisation of these schemes has not been possible up until now. This is because the STM has been used to induce the switching of dimers in experiments on the Si(100) surface in the past [198, 199], and since electron tunnelling produces a significant long-range interaction [197], this method appears not to allow for controllable manipulation of the surface dimers using a tip with atomic precision. Recently, however, it was demonstrated by Y. J. Li et al. [187], in their study to determine the Si(100) ground-state reconstruction, that it is indeed possible to achieve atomic manipulation with atomic precision on this surface using the NC-AFM. This was done by altering the tip-surface interaction, which in turn modifies the Si(100) surface reconstruction observed in NC-AFM images.

8.3 Experimental Setup

Experiments were performed by Prof. Phillip Moriarty’s group using a NC-AFM/STM operated in ultra-high vacuum (UHV) under zero applied bias conditions at 5 K in the qPlus mode introduced by Giessibl [92, 200]. The Si(001) surfaces were Boron doped, and after preparation, the Si(100) surface reconstruction was checked by performing STM, using a qPlus sensor before beginning NC-AFM experiments. Further experimental details can be found in the Supplementary Material (SM) for the study by A. Sweetman et al. in Ref. [19].

Initially the tip was operated in frequency shift mode, imaging the c(4 × 2) surface, then after positioning the tip above the lower dimer atom, the tip-surface separation was reduced, measuring ∆f until a sharp jump (discontinuity) in the ∆f vs. z curve was observed (indicating a manipulation). The tip-surface separation was increased until the original ∆f was achieved, then the surface was imaged again.

8.4 Theoretical Model

The system was described using the SIESTA code in the same way as in the study of dissipation on Si(001) discussed in the previous chapter, differing in that here the slab used to describe the surface consisted of six layers with sixteen dimers arranged in two rows of eight dimers, to allow for adequate separation of the surface dimers that are manipulated and the system image. Choosing six layers for relaxation was based on previous studies where this was
found to be sufficient to describe manipulation on this surface [187, 201]. Also, both the H3 and the dimer tip were used in these calculations. For more information, the author refers the reader to the SM of Ref. [19].

To produce the $F$-z curves, the forces are calculated while the tip was moved in quasi-static steps from 8 Å, until just after the dimer was flipped, then the tip was retracted. The author also notes that the reference tip-surface separation used here is different from that used in Chapter 7, but the scale is the same.

To perform the NEB calculations, BEINEB was used, simulating each band using 17 images. This number of images was used in order to obtain a good description of the intermediate minimum (c.f. Section 8.7). Before each calculation of the bands, initial and final geometries were relaxed and intermediate geometries were obtained by linear interpolation. The bands were then relaxed until the image energies varied by less than 0.01 eV and the forces by less than 0.1 eV/Å; these tolerances were tested and found to adequately describe the PES. Then, after obtaining the first band, the tip-surface separation was reduced and all subsequent bands were initially obtained from the previously relaxed preceding bands, i.e., the relaxed bands were used as inputs for slightly smaller tip-surface separations. This calculation method ensured that evolution of the band as the tip-surface separation decreased was smooth. Near the smallest tip-surface separation (corresponding to just before the dimer flipped in our DFT calculations), typically around 3.47 Å, small decreases in tip-surface separation of 0.02 Å were used to provide a clearer picture of the changes in the tip-surface separations that cause the dimer to flip.

All MD simulations were *ab initio* and an NVE ensemble (constant number of atoms, energy and volume) was used, which is sufficient for our purposes.

All NEB calculations were performed using the code developed by the author, BEINEB, all $F$-z calculations were performed by Samuel Jarvis, as were the NEB calculations. In this study the author provided assistance in setting up the NEB calculations, analysing them and performed all of the MD calculations. The reaction coordinate used in all PES (NEB) plots in this chapter is the image number.

### 8.5 Toggling the Dimers

A typical result for toggling the dimers, where the original $c(4 \times 2)$ configuration is changed (manipulated) into a phason pair structure\(^1\) [202] is shown in Fig. 8.1(a), i.e. flipping one dimer typically creates a phason pair. After manipulation, the phason pair structure is manipulated back into the original $c(4 \times 2)$ configuration, as shown in Fig. 8.1(b). In can also be observed \(^1\)Here, phason refers to a buckled dimer phase defect, where two dimers next to each other are in the same orientation and phason pair refers to two phasons in different orientations.
The approach we have adopted involves measuring the dimer flipping event appears as a sharp jump in the amplitude, 250 pm; Phasons also form naturally on Si(100) [and the structure... pseudoconstant height mode on different...]. The variation... vs... spectrum acquired during retraction of the tip. The insets show NC-AFM images, and illustrations of the atomic relaxation was considered complete when forces... buckled configuration to a phason pair state [Fig. 3(b)]. (A movie of the induced phason motion is available online [Fig. 8.1].)

Figure 8.1: Experimental NC-AFM $\Delta f$ vs. $z$ curves for toggling dimers on the Si(001) surface, where the black and red curves describe the approach and retraction, respectively. In (a) the tip is positioned above the lower dimer atom of the $c(4 \times 2)$ surface, creating a phason pair. In (b) the tip is positioned above the lower dimer atom of one of the phasons, restoring the original $c(4 \times 2)$ surface. Cartoons at various points and schematics of the manipulation and topography images produced using a $\Delta f$ value of -9.1 Hz are given in the inset in both (a) and (b). In topography images the $\times$ indicates the tip position. Reproduced from A. Sweetman et al. [19].

From Fig. 8.1 that the $\Delta f$ threshold to manipulate the dimers depends on the initial configuration of the surface; the dimers in the phason pair configuration flip at a different $\Delta f$ value than for the $c(4 \times 2)$ surface. This indicates that the force required to flip the dimers depends on the initial state of the system. In this way, and in the absence of tunnelling electrons which were previously reported to cause the dimers to flip, it is possible to controllably manipulate a long sequence of dimers along a row. Such a sequence of manipulations is shown in Fig. 8.2.

Figure 8.2: A sequence of NC-AFM topography images of the Si(001) surface, depicting dimers that are toggled sequentially along a row, where $\times$ indicates the tip position for each manipulation. Initially starting from a phasons in (a), which are shown in white boxes, to finally producing a $c(4 \times 2)$ region, shown in (k). Reproduced from A. Sweetman et al. [19].

It should be noted here that the expected intermediate state, three dimers in the same orientation in a row (henceforth referred to as three-in-a-row), is not seen experimentally. In
other words, we do not see that a single dimer has been manipulated, but instead see that two dimers are manipulated; correlated/concerted flipping occurs.

8.6 Forces Required to Flip the Dimers

$F$-$z$ data was then produced in Nottingham using the Sader-Jarvis inversion algorithm (a technique to determined the force from the frequency shift) in order to determine the forces required to flip the dimers. These results were then compared to theoretical simulations and the results are shown in Fig. 8.3. Firstly, from this we observed excellent agreement between theory and experiments for both tip models; in particular we see even better agreement for the dimer tip. The small amount of dissipation experimentally measured provides strong evidence that the manipulation occurred over a single oscillation cycle (or at least over a very small number of oscillations), resulting in small measured dissipation values as the signal is averaged over many oscillations. Furthermore, the calculated hysteresis loops from which theoretical dissipation may be calculated for both tips almost exactly match the profile discussed in Chapter 7. Therefore, it is safe to assume that the tip used in the experiments is a relatively rigid tip. These results strongly indicate that the manipulation occurs as a result of a single bond forming between the dimer and the tip apex, and have found the force at the onset of the dimer flip (the threshold force) to be approximately 0.6 nN.

![Figure 8.3](image)

**Figure 8.3:** Experimental and theoretical NC-AFM $F$-$z$ and experimental dissipation curves for toggling dimers on the Si(001) surface. In theoretical calculation, results for both the H3/Si(111) and dimer tip were collected. In (a) the tip is positioned above the lower dimer atom of the $c(4 \times 2)$ surface, creating a phason pairs. In (b) the tip is positioned above the lower dimer atom of one of the phasons, restoring the original $c(4 \times 2)$ surface. In topography images the $\times$ indicates the tip position. Reproduced from A. Sweetman et al. [19].

8.6.1 Variations of the Forces

Although we now have a map of the forces, we have also found that the force threshold for flipping the dimers depends on the initial state of the system and can vary widely. We found
that the threshold force is different depending on the type of dimer being manipulated, i.e. the state of the surface in the region near the tip probing area. These thresholds were found to be similar to or below the threshold for the pristine c(4×2) surface. From simulated F-z curves in Fig. 8.4 we determined the threshold forces for the tip positioned above different lower dimer atoms along and next to the three-in-a-row structure. From this, we observed that the middle dimer of the three-in-a-row structure (position iii, the same dimer that would need to be flipped in order to form the three-in-a-row structure from the pristine surface), is the least stable out of them, with a threshold force of ≈0.2 nN, nearly three times smaller than for the pristine surface. The adjacent dimer (position ii) has a threshold of 0.5 nN, more than twice as large as the middle dimer, and the dimer beginning/terminating the c(4×2) configuration (position i), next to the three-in-a-row structure, has a threshold of 0.6 nN similar to that of the pristine surface. These results clearly show that the force required to flip the dimers critically depend on the configuration of the neighbouring dimers.

8.7 Three-in-a-Row Structure

Although not seen experimentally, our simulations predict that at 0 K the three-in-a-row structure is stable when the dimer is flipped by either tip, the same results obtained in our dissipation study. Similarly when flipping a dimer in configuration i in Fig. 8.4 a five-in-a-row structure is formed, also not seen experimentally. The relaxed geometries before manipulation and after (when the three-in-a-row structure is formed) for both tips positioned at the threshold
The Pristine Surface: Tip Induced Variations in Potential Energy Surface

tip-surface separation (the separation where the dimer flips) is shown in Fig. 8.5.

![Diagram](image)

**Figure 8.5**: Relaxed geometries of the Si(001) surface just before manipulation (the $c(4 \times 2)$ surface) and after manipulation (the three-in-a-row configuration) using the H3/Si(111) tip, (a) and (b), respectively, and for the dimer tip (c) and (d), respectively. Si and H atoms are represented by yellow and white balls, respectively, while blue and green balls indicate the dimers along the row being manipulated and the dimer immediately beneath the tip, respectively. Image is adapted from A. Sweetman et al. [19].

8.8 The Pristine Surface: Tip Induced Variations in Potential Energy Surface

To study the dimer flipping transition in more detail we mapped the minimum energy
path for the manipulation using the NEB method in the absence of the tip, effectively, when the tip is positioned at the threshold separation and for several intermediate tip heights. The breakdown of the minimum energy profiles (MEP or bands) is shown in Fig. 8.6, summarised in Fig. 8.7. The energy profile of the formation of a phason pair from the pristine surface for large tip-surface separations (when the tip is effectively absent) is shown in Fig. 8.6(a). Although this is the path our calculation determined as being the MEP on the PES, the energy profile for the transition from the initial (i) to final (iv) state (see Fig. 8.6(c)) can in fact follow one of two distinct paths (as shown in Fig. 8.6(b)), both of which were explored. The two paths relate to a different combination of dimer manipulations:

1. one with a single large barrier (≈330 meV), which corresponds to a correlated motion in which two dimers flip simultaneously (red circles);

2. one consisting of two separate barriers (blue triangles) of different sizes (<330 meV), corresponding to the sequential flipping of dimers.

The latter refers to the flipping of the dimer beneath the tip, first forming a three-in-a-row structure, followed by the flipping of an adjacent dimer, forming a phason pair. These two different profiles evolve through different local MEPs from i to iv, illustrated in Fig. 8.6(c). At larger tip-surface separations, it can clearly be observed that the system prefers to flip the dimers sequentially rather than simultaneously as both barriers for sequential manipulation are almost 30% smaller as compared to simultaneously flipping two adjacent dimers. This becomes even more
evident at smaller tip-surface separations, as the energy profile converges on the sequential dimer flipping pathway (grey bands in Fig. 8.7), irrespective of the initial pathway chosen (sequential or simultaneous) for the NEB optimisation.

**Figure 8.7:** (a) Minimum energy profiles relative to the initial state of the system, which evolves from the pristine c(4 × 2) surface (i) to a surface containing a phason pair (iv). The profiles are plotted as a function of image number, where geometries between images are approximately equidistant. (b) NC-AFM topography image of the Si(001) surface as 77 K with a boron defect (white box), using Δf = 40 Hz. The profile for the system containing a two dimer vacancy (2DV) is offset for clarity. Included are the energy barriers for the pristine and 2DV systems at the closest tip-surface separation, and a cartoon representing part of the 2DV cell, colour coded as in Fig. 8.6. Image is adapted from A. Sweetman et al. [19].

All bands where the dimers flip sequentially (Fig. 8.6(b) and Fig. 8.7) demonstrate several features:

1. they all have a distinct minimum in the middle which corresponds to the three-in-a-row structure (ii);

2. when the tip approaches the surface, the total energies of both the three-in-a-row and phason pair structures reduce, i.e. they become more stable;

3. when the tip comes very close to the surface, the barrier for a single dimer in the c(4 × 2) phase to flip, forming the three-in-a-row structure, reduces smoothly to zero;

4. the barrier to evolve to the phason pair configuration does not collapse at any tip-surface separation.
8.8 The Pristine Surface: Tip Induced Variations in Potential Energy Surface

The third feature described here, suggests that the flipping of the first dimer at 5 k would occur spontaneously at sufficiently small tip-surface separations, creating the three-in-a-row structure. Hence, it can be deduced, from our NEB calculations, that for any tip-surface separation, the MEP for the transition from the pristine $c(4 \times 2)$ surface to the phason pair state requires the system to evolve through the three-in-a-row configuration, i.e. by flipping the two adjacent dimers sequentially.

8.8.1 The Stability of the Three-in-a-Row Structure

The results of the NEB calculations therefore pose two further questions:

1. why do we never observe a three-in-a-row structure in experiments?
2. how is the phason pair formed?

8.8.1.1 Pristine Surface $\rightarrow$ the Three-in-a-Row Configuration

The first of these questions may be answered by a closer analysis of the energy profiles. If a three-in-a-row structure is formed at tip-surface separations of $\approx 3.47 \, \text{Å}$, then the calculated energy barrier to evolve to a phason pair structure is always unsurmountable at 5 K, changing from $\approx 60 \, \text{meV}$ at the the smallest tip-surface separation, to 110 meV at the largest one (equivalent to when the tip is not present). However, the barrier to return to a $c(4 \times 2)$ structure from the three-in-a-row structure smoothly reduces, becoming negligible as the tip is removed from the surface. However, the barrier does not disappear for the transition from the phason pair to the $c(4 \times 2)$ structure, i.e. the phason pair is a stable structure.

To verify our calculated transition from the pristine $c(4 \times 2)$ surface to the three-in-a-row configuration (and simultaneously the reverse transition) the energy profile for this transition was calculated using more stringent convergence parameters. This was then compared to that of the evolution from the pristine $c(4 \times 2)$ surface to the phason pair, which has the three-in-a-row configuration as an intermediate state. The result is shown in Fig. 8.6(a), and excellent agreement is observed between the profile of the direct evolution to the three-in-a-row configuration (black circles) and the profile for where the three-in-a-row configuration is the intermediate state (blue triangles). Both calculations determine an energy barrier of 17 meV for the system to return to the pristine $c(4 \times 2)$ surface from the three-in-a-row configuration when the tip is removed. The author notes here that this 17 meV barrier is outside the precision of our DFT calculations, and may in fact therefore be zero. Furthermore, in the 5 K molecular dynamics (MD) calculations (discussed in Section 8.8.2), the three-in-a-row structure was found to be unstable as oscillations

\[ \text{An energy barrier of 60 meV at 5 K corresponds to a transition rate of } \approx 10^{-61} \text{ (a timescale of } 10^{53} \text{ years), i.e. very unlikely to happen.} \]
of the kinetic energy of the central dimer were observed above 17 meV. This shows that if the 17 meV barrier is present for the system to evolve from the three-in-a-row structure to the $c(4 \times 2)$ surface, then it is surmountable at 5 K. This in fact occurred in one such calculation. These facts clearly indicate that if the tip during approach traps the system into the three-in-a-row structure, then upon tip retraction, the system will be unstable at 5 K and will return to the more stable $c(4 \times 2)$ surface.

8.8.1.2 The Three-in-a-Row Configuration $\rightarrow$ The Phason Pair

The second question appears to be much more challenging, particularly so as not much is known about either the three-in-a-row configuration or the phason pair outside of this study. Furthermore, in all of our calculations, we clearly observe that the barrier from the surface with a three-in-a-row configuration to evolve into a surface with a phason pair structure is unsurmountable at 5 K for any tip-surface separations.

In an attempt to understand the cause of this barrier, we re-calculated our energy profiles using more stringent convergence parameters, and after performing these calculations (which were also allowed to run for many more steps), we found that this energy barrier did not reduce. Ruling out insufficient convergence in our NEB calculations, we reviewed the optimisation procedure implemented in our NEB code, and investigated whether the process was artificially steered through a local minimum that does not lie along the MEP. This is an issue as the final MEP depends on the initial transition path model. More specifically, we tested to see if simultaneously flipping two dimers could lead to lower barriers than sequentially flipping the dimers at small tip-surface separations, which would avoid the intermediate three-in-a-row structure. To do so, we performed a constrained minimisation calculation at the threshold tip-surface separation, where the dimer beneath the tip apex and an adjacent dimer were both flipped simultaneously. However, this constrained minimisation produced a barrier similar to that of the red band in Fig. 8.6(b), significantly larger than for sequentially flipping the dimers. In short, our results strongly suggest that the MEP involves the system evolving through the three-in-a-row configuration.

These findings, that the pristine system is statistically very unlikely to evolve into the phason pair system from the three-in-a-row configuration, and the calculated very small barrier of 17 meV (or less) to return to the pristine $c(4 \times 2)$ surface from the three-in-a-row configuration as the tip retracts, can be combined to explain the shape of these potential energy landscapes. As a result, one may except that when trying to flip a dimer on the pristine $c(4 \times 2)$ surface, the system will undergo many flip-but-flip-back events if it gets trapped in the three-in-a-row configuration as the tip retracts. However, such a conclusion would seem to contradict the fact that phason pairs can be routinely created.
8.8 The Pristine Surface: Tip Induced Variations in Potential Energy Surface

8.8.2 Correlated Dimer Flips

To summarise, we found in our NEB simulations (performed by the Nottingham group) that the barrier from three-in-a-row to the phason pair structure is unsurmountable at 5 K. However, this barrier appears too high only if the system is stabilised into the three-in-a-row structure at that very low temperature. In reality, the three-in-a-row structure may serve only as an intermediate state of the system after the first dimer is flipped and may only exist a very short time. Moreover, this state is only accessible at very close tip-surface separations when the barrier from the $c(4 \times 2)$ state disappears completely (the green band in Fig. 8.7). If we now consider this energy profile in more detail, we see that when the barrier collapses and the system falls into the three-in-a-row state, $\approx 300$ meV of energy is released into the system. This energy would be sufficient to surmount the energy barrier separating the three-in-a-row and the phason pair structures which is $\approx 80$ meV at the threshold tip-surface separation, if released into the necessary region. If this is indeed the case, then the second dimer would flip as a result of the energy released by the first dimer flipping, i.e. the flipping of the second dimer would be directly correlated to the flipping of the first dimer. However, the energy may be dispersed (dissipated) into the surrounding system in which case insufficient energy would reach the adjacent dimer and this process (as described here) would not be possible.

To study this aspect of the process and verify this idea an alternative technique was required. We therefore continued our investigation by performing ab initio MD simulations. The system was initially prepared at the geometry corresponding to the image which is next to the $c(4 \times 2)$ state along the green band in Fig. 8.7 to ensure that the system will quickly relax into the three-in-a-row structure during the course of the dynamics. The tip was positioned directly above the dimer ‘down’ atom. The time step was chosen to be 1 fs and we used the same SIESTA setup as described in Section 8.4 to perform the calculations. Although no initial velocities were given to the atoms, after about 200 MD steps the dimer beneath the tip (tip dimer) flipped and the system reached the three-in-a-row structure with substantial release of the kinetic energy (KE) on dimer atoms directly underneath the tip. This is shown in Fig. 8.8, where the evolution of the KE of the atoms most relevant in the dimer flipping process (see the inset), have been plotted as a function of MD step number. Here it can be seen that the dimer (green) directly beneath the tip apex (shown black) has its largest KE after 200 steps, which corresponds to the flipping of the dimer - referred to here as the ‘tip dimer’. Much of this energy is then transferred to the tip apex (around 70 meV); however due to the higher atoms of the tip being artificially frozen, the energy that would otherwise have been transmitted throughout the tip, is instead reflected back to the surface (a limitation of our frozen tip model). Some of the energy acquired by the tip dimer immediately after it flips and also after it is reflected from the frozen part of the tip, is
dissipated throughout the surface, predominantly via the atoms connecting adjacent dimers (red and turquoise, referred to here as ‘connecting atoms’). For the pristine surface the left and right connecting atoms each gain around 50 meV in kinetic energy. However, the problem with the perfect surface is that these connecting atoms are connected to four Si atoms with no large bias to transmit this energy in a particular direction, therefore the kinetic energy transferred to the left and right dimers is very small, typically around 10 meV. This naturally follows since out of 8 available bonds (from two connecting atoms on either side of the tip dimer) only two of them are bonded to a dimer, hence a quarter of the kinetic energy of the connecting atoms is transferred to the connecting dimer. Concluding, we see that initially the injected energy causes oscillations of the dimer atoms, but then this energy is seen to transfer to the tip apex atoms and then back to the dimer atoms and eventually to the whole system.
As the simulation progresses, the atomic oscillations spread across the whole system and eventually die away. Importantly, however, we do not observe a flip of either adjacent dimer, i.e. the system stabilises into the three-in-a-row state. In other words, there is insufficient kinetic energy transfer to surmount the barrier to the phason pair state. This result was initially attributed to the symmetrical tip position which does not favour either of the two dimers adjacent to the dimer which was forced to flip. Consequently, several non-symmetrical tip positions with respect to the dimer were also tried resulting in different amounts of KE transferred to the two adjacent dimers. However, in none of the cases did the second dimer flip after the first one was flipped.

These MD calculations demonstrate without doubt that the energy gained after the first dimer is flipped cannot be transferred in sufficient amounts to the neighbouring dimer to stimulate its sequential flip; this result does not seem to depend on the actual lateral tip position either. This statement can be made as the tip used in this calculation represents an edge case where the maximum amount of KE remains within the surface and other tip models (larger or with more atoms allowed to relax) would lead to softer contact and would absorb more of the energy. Conversely, MD calculations of the system with a three-in-a-row structure, without the presence of the tip, show that the 17 meV barrier is surmountable at 5 K, and as such we conclude that the three-in-a-row structure is unstable in the absence of the tip, even at 5 K.

The overall conclusion which is drawn from the above calculations (both MD and NEB) is that the observed formation of the phason pair structure cannot be explained for a pristine surface (without defects) at any tip position. This, however, is a strong indication that surface defects may also play a role. This is in fact proven to be the case when further experiments were performed to specifically investigate whether the presence of defects affect the manipulation of the dimers.

8.9.1 Non-Local Variations in Surface Stability when Scanning a Defective Surface

In Fig. 8.9 we found local variations in the ability for dimers to flip when scanning the Si(001) surface with phason pair defects at low temperatures. Here, it can clearly be observed that the phason pairs can be manipulated from the $c(4 \times 2)$ phase to the $p(2 \times 1)$ phase during scanning without manipulating “ordinary” dimers (meaning dimers of the pristine $c(4 \times 2)$ surface). The images in (a) and (b) clearly indicate that under the same imaging conditions, the dimers in the phason pairs are more easily manipulated that other surface dimers. When the $\Delta f$ in increased,
as in (c) and (d), other dimers begin to be manipulated, which makes identification of the phason pair more difficult, however still not all dimers are manipulated. Then scanning again at further increased $\Delta f$ values, as in (e) and (f), shows that even at large $\Delta f$ values, where most dimers are manipulated, there are still some dimers which maintain their buckled conformation, indicating that the barriers to flip those dimers are significantly larger than for the other dimers.

This point about the variation in the ability to flip dimers is again demonstrated in Fig. 8.9(g), where instead of the presence of phason defects, Boron defects are present. Here, the presence of the Boron defects appears to have strongly non-local effects, reducing the barrier for dimers along the same row to flip, as the row inline with the defects (green box) appear to be

![Figure 8.9: Low temperature NC-AFM images of the manipulation of phason pairs due to scanning, where the phason pairs are changed from $c(4 \times 2)$ configuration to $p(2 \times 1)$. Scanned before manipulation at $\Delta f = -26$ Hz (a) and after at $\Delta f = -27$ Hz (b). The same surface is scanned again at $\Delta f = -28$ Hz (c), where the phason pair become difficult to distinguish from other dimers, and at $\Delta f = -29$ Hz (d), where other dimers are likely being manipulated alongside the phason pairs. At $\Delta f = -31$ Hz (e) the $p(2 \times 1)$ configuration is observed, however some dimer rows still remain in their buckled state (indicated by white arrows). Similar results can also be observed when scanned at $\Delta f = -34$ Hz (f). In (g) a different portion of the surface is scanned with a different tip at $\Delta f = -40$ Hz and here, Boron ad-dimer defects are observed (indicated by the black box). The dimers along the row of the Boron defects appear to be in the $p(2 \times 1)$ configuration (green box), while the dimers the adjacent row (red box) appear to be in the $c(4 \times 2)$ configuration. A cartoon of the dimer flipping under the influence of the tip is shown in (h). Image reproduced from A. Sweetman et al. [189].]
in the $p(2 \times 1)$ phase (i.e. the dimers are manipulated such that both dimer atoms are imaged), while in the adjacent row (red box) dimers are imaged in the $c(4 \times 2)$ phase. This effect can be very long-ranged, sometimes extending over 30 dimers in a row.

Since, in all the theoretical calculations performed thus far the surface has been ‘pristine’ (defect free), to investigate the influence defects may have on the formation of the phason pair (e.g. reducing the 80 meV energy barrier for the system to evolve from the three-in-a-row configuration to the phason pair), further NEB calculations were also performed on three types of defective surfaces:

1. with two dimers missing;
2. with substitution defects;
3. with surface strain.

### 8.9.2 Two Dimer Vacancy (2DV)

![Figure 8.10](image.png)

**Figure 8.10**: The minimum energy profile (offset for clarity) for the system to be manipulated from the $c(4 \times 2)$ surface with a two dimer vacancy (2DV) (i in Fig. 8.6(c)) to a surface containing a phason pair (iv in Fig. 8.6(c)) via the system containing the three-in-a-row structure (ii in Fig. 8.6(c)). A cartoon describing the 2DV is shown in (c), colour coded the same as in Fig. 8.4. The energy profiles are plotted as a function of image number, where geometries between images are approximately equidistant. Image is adapted from A. Sweetman et al. [19].

To study the effect of missing dimer defects on the correlated flipping of two dimers to form the phason pair, the energy profile, shown in Fig. 8.10(a), for an extended $2 \times 10$ dimer slab with two dimers removed (creating a dimer vacancy referred to as 2DV), as depicted in Fig. 8.10(b), was calculated. This was chosen as a representative example. This calculation was performed for the tip positioned at the threshold tip-surface separation for the first dimer to flip,
\section{The Defective Surface: Defect Induced Variations in Potential Energy Surface}

and the results were compared to the profile for a pristine/perfect surface, shown in Fig. 8.10(a). We found that the energy barrier for the system to evolve from a three-in-a-row configuration to the phason pair state increased by nearly 40 \%, from 80 meV to 110 meV. This makes the barrier even more unsurmountable, but also demonstrates that defects can vastly change the potential landscape, in this case increasing the barrier to evolve on to the phason pair state. It is of course possible to model other dimer defects, however, as an example of a computationally accessible defect, the 2DV defect was chosen. It may also be possible that other defects, such as boron ad-dimer related structures \cite{203} will instead reduce the barrier for the three-in-a-row to phason pair transition.

\subsection{Emulating Sub-Surface Defects}

Some of the MD simulations were performed for systems in which the masses of selected surface atoms were either increased or reduced. This approach allowed us to simulate different possible responses of the Si lattice with different substitutions to the energy injected into the neighbouring dimers due to the flip of the dimer beneath under the tip. Specifically, using the H3 tip, the masses of atoms $^1$ connected to both the atoms of the tip dimer (dimer directly beneath the tip, the first dimer to flip) and dimer adjacent to the tip dimer were modified as these atoms must be predominately responsible for the energy transfer between the two dimers. Varying amounts of energy transfer were observed depending on the particular model, however, these energies never exceeded 20 meV in our simulations (with a frozen tip), an insufficient amount of energy to flip either adjacent dimer.

\subsection{Strained Surface}

In another type of defect simulation, we varied the distance between two neighbouring dimers along the same row as shown in Fig. 8.11. This distance was either reduced or increased to simulate two types of an external strain on the surface due to an unknown remote defect. Doing this allows one to emulate the strain at the upper layers approximately; this is the simplest model, otherwise something more elaborate would need to be done. This is suitable for the purposes of this study which is to determine the relative effect of strain (that leads to compression or expansion of the system) on the PES.

\footnote{Connecting atoms; c.f. red and turquoise coloured atoms in Fig. 8.8.}

with the bottom atoms as in the bulk. To prepare such a system, we first moved all atoms to the right of a particular dimer in the cell, dimer 1 in Fig. 8.11(a) and dimer 2 in Fig. 8.11(b) by some distance $\Delta a$, where $\Delta a = a - a_0$, and correspondingly adjusted the cell lattice vector in the direction along the dimer row. Then the atoms of the surface were allowed to relax keeping
the atoms in the lower two layers of the slab (together with the capping H atoms) fixed, as usual. Using this system, MD simulations were performed in the same way as before, with the tip at the threshold position, above the ‘down’ atom of dimer 1 in Fig. 8.11 (typically at the tip-surface separation of 3.65 Å and 3.53 Å). Performing these simulations did not result in the flip of the dimer 2, however, varying energy transfers to both atoms of that dimer (dimer 2) were observed in the simulations after the tip dimer (dimer 1) flipped, depending on the amount of the shift ($\Delta a$), its direction and where the shift was applied. The largest energy transfer of up to 50 meV was obtained when the distance between dimers 2 and 3 was increased by $\Delta a = 0.3$ Å as shown in Fig. 8.11(b). We note that in these simulations all tip atoms were frozen, and as such the energy transfer to the surface is likely to be overestimated.

8.9.5 The Role of Defects in Phason Pair Formation

Our simulations predict that if a single dimer is flipped from the pristine $c(4 \times 2)$ configuration, then a stable three-in-a-row structure (never observed experimentally) is formed with no energy barrier, and that the barrier for the system to evolve on to form the phason pair is 80 meV, a barrier we have calculated to be unsurmountable at 5 K. In fact our calculations predict that if the system is manipulated into the three-in-a-row structure from the pristine $c(4 \times 2)$ surface during tip approach, then upon tip retraction, the system will simply return to its original $c(4 \times 2)$ configuration. However, both theoretical and experimental calculations have also revealed that the presence of defects can significantly affect the barrier for the system to go from the having a three-in-a-row structure to a phason pair.
Figure 8.12: The minimum energy profiles (offset for clarity) for the defective system to be manipulated from the $c(4 \times 2)$ surface (i) to a surface containing a phason pair (iv) via the system containing the three-in-a-row structure (ii). The first energy barrier for each type of calculation vanishes at the threshold tip position, while the second barrier for each calculation is: $\approx$80 meV for the pristine surface; $\approx$110 meV for the system containing two dimer vacancies (2DV); $\approx$60 meV for the compressed system; $\approx$45 meV for the expanded system. Cartoon describing how the threshold tip-surface separation is defined, and of the initial, intermediate and final states of the system are given in the insets. The energy profiles are plotted as a function of image number, where geometries between images are approximately equidistant. Image reproduced from A. Sweetman et al. [189].

Calculations of the threshold force to flip a dimer and the energy profile for the system to go from a $c(4 \times 2)$ phase to that containing a phason pair in the presence of a defect have thus far been limited only to surfaces with dimer defects, therefore we have also performed calculations to determine the energy profile for forming the phason pair near where the surface is strained. The aim of this was to simulate a long-ranged surface strain effect due to the presence of a distant defect, and was achieved by artificially introducing a lattice mismatch near the vicinity of the phason pair. The same setup as with our MD simulation emulating the same type of surface strain performed previously (see Section 8.9.4) was used, and an illustration of the lattice mismatch described is shown in Fig. 8.11. The spacing of the fixed atoms was changed by a 0.3 Å along the direction of the rows, and, after allowing the lattice to relax, NEB calculations were
performed in order for us to explore the energy barriers to form the phason pair from the $c(4 \times 2)$ surface, with particular interest in how the emulated strain affects the barrier to go from the three-in-a-row to the phason pair structure when the tip is positioned at the threshold tip-surface separation. The results of the NEB calculations are shown in Fig. 8.12.

In the calculations where surface strain was emulated (i.e. artificial contraction and expansion of the lattice spacing was introduced), shown in Fig. 8.12, a drop in the barrier for the system to evolve from containing the three-in-a-row structure to the phason pair was found. Combining these results, that the barrier for the second dimer to flip can be reduced to \(\approx 45 \text{ meV}\) due to surface strain (when the dimer is in the vicinity of surface expansion), with that from the MD simulation of the same surface, where up to \(\approx 50 \text{ meV}\) is observed to be transferred to an adjacent dimer, we have found that the creation of the phason pair is very likely a result of surface strain. Of course this does not mean that the flip will necessarily occur, we it can at least be now said that the system is in fact likely to have access to the energy necessary for it to occur. We note that although the sub-surface displacement is moderate, the actual response at the surface is relatively small \(\approx 0.1 \text{ Å}\), of the same order as (or smaller than) reported surface strains of 0.1 - 0.8 Å, induced by surface defects [26, 184]. Indeed, strains of this magnitude would also be of the same order as the noise in the measurements during FM-AFM imaging, \(\approx 0.1 \text{ Å}\) [26, 184]. Since these modest calculations reveal a reduction in the barrier, it may indeed also be possible that some defects and resulting strain fields may lower the barrier even further, but an extensive theoretical investigation of the effect that different defects has on the PES, which would require larger simulation cells, is beyond the scope of this study.

### 8.9.6 Summary

We have seen in our NEB and MD simulations of the pristine surface that the energy transferred to the surface after the tip dimer flipped is dispersed into the tip and the surface in such a way that the dimer next to the tip dimer does not receive enough energy to flip. This means that our simulations cannot explain the formation of the phason pair structure in this case. Therefore, we have to conclude that certain surface defects must play a crucial role in triggering this process.

Upon considering the possible role(s) a surface defect (or defects) should play in the formation of the phason pair structure, we note that the energy gain in the system after the tip dimer flipped (of the order of 300 meV) is to be shared between the surface and the tip. This means that much less than 300 meV will actually be available for the second dimer to induce its flipping. Hence the role of the defect(s) must be to: increase the amount of energy transferred to the dimers adjacent to the tip dimer, in which case the second dimer will have enough energy transferred to
8.10 Conclusion

it to flip immediately after the first dimer flips; considerably reduce the barrier required to flip the second dimer, in which case the second dimer will be flipped by the kinetic energy acquired by its atoms after the tip dimer has flipped; remove it altogether, in which case the transition should happen spontaneously, immediately following the flip of the first dimer.

We also stress that after the tip is retracted, the phason pair structure should remain stable. Therefore, the defect must not destabilise the phason pair such that the barrier for it to go back to the three-in-a-row structure remains unsurmountable at 5 K after the tip is fully retracted.

Thus, the role of the defects is therefore two-fold: they should not only serve to dramatically increase the affinity of the three-in-a-row structure to go into the phason pair in the presence of the tip, but at the same time, the barrier to go in the reverse direction (at large tip-surface separations) should not be reduced, i.e. the defects should not destabilise the phason pair structure. This subtle interplay of the defects most likely explains the large number of failed manipulation attempts we observed, as described in the Section 8.9.1.

From these results it can be deduced that forming the phason pair from a pristine $c(4 \times 2)$ surface is not solely the result of the energy barriers collapsing due to the proximity of the tip. This is based solely on the theoretical results, however experimental observations also support this. In fact it was observed that the ability to create/inject a phason pair into the surface critically depends on the presence of defects in the vicinity of the dimers being manipulated.

8.10 Conclusion

In the main, the results described here clearly demonstrate that it is indeed possible to routinely and reversibly manipulate dimers on the buckled-dimer Si(001) surface, to toggle the dimer bond angle via short-ranged covalent tip-surface interactions above the lower dimer atom. This represents a new reversible atomistic manipulation and switching protocol, that enables direct, in-plane, mechanical control of bistable atoms. We show clearly that the presence of defects in the vicinity of the dimers being manipulated critically determine the ability of those dimers to be controllably manipulated.

It must be stressed that due to the uncertainty in the microscopic tip structure, and limitation on the size of the surface and microscopic tip that can be modelled within a reasonable amount of time, it is not feasible to model all of the types of defects present in the real system. Therefore, the conclusions we have drawn from our theoretical calculations give, at best, only plausible interpretations of experimental results. Nonetheless our simulations provide key insight into many of the experimental observations. Our simulations provide a possible mechanism to explain the creation of phason pairs, which is an interplay of the residual energy left in the system
after manipulating the first dimer and a reduction in the barrier to flip the second dimer due to the presence of surface strain. The controllable and reversible toggling of these dimers marks a new form of atomic manipulation.
Part IV

Summary & Conclusion
Chapter 9

Summary

Under a common theme, the different studies presented in this thesis, different pieces of the same metaphorical puzzle, combine to show that the success of the NC-AFM in some systems can be attributed solely due to the presence of either defects, instabilities or both.

The investigation was carried out on an oxidised metal insulator surface and a semiconductor surface, two contrasting surfaces. It was broken into five separate studies, summarised in Chapters 4-8, three studies of the former and two of the latter, of:

1. the thermodynamic stabilities of the reconstructed oxidised Cu(110) surfaces (Cu(110):O surface), where the necessity to take great care when modelling surfaces beyond what is typically accepted as sufficient, was proven;

2. NC-AFM imaging of the Cu(110):O surface, where simultaneous tip termination and surface species identification was achieved, leading to the implementation of an on-the-fly tip identification protocol that depends on the presence of ad-atom defects;

3. NC-AFM manipulation of the Cu(110):O surface, where the protocol implemented in the 2nd study was used, the results from the 1st study were shown to significantly affect the theoretical simulations, and a new mechanism for lateral atomic manipulation was proposed;

4. NC-AFM dissipation on the Si(001)-c(4 × 2) surface, where the dissipation mechanism in the system was determined, and unexpectedly large dissipation results due to surface and NC-AFM tip instabilities were shown to be qualitatively accurate when compared to preliminary experimental data;

5. NC-AFM manipulation of the Si(001)-c(4 × 2) surface, where a new type of controllable and reversible atomic manipulation and switching protocol has been implemented, based on the surface instabilities, which also demonstrates the crucial role of defects in the manipulation.
In terms of the aim, surface stabilities were studied on both the Cu and Si surfaces, each providing different pieces of information. Surface stabilities of the Cu(110):O surfaces were studied, by way of the Gibbs free energy difference, $\Delta G$. This showed that conventional DFT models of the surface fail to accurately describe the $c(6 \times 2)$ reconstruction of the surface; the difference in $\Delta G$ between similar surfaces was found to be of the order of 30 meV in some cases. The exact nature of the poor description was not determined, but it appears to be related to non-local correlation effects, namely vdW interactions. Initially, it was thought that the necessity to account for vdW in the calculation of the bulk lattice constant, thereby obtaining a better (smaller) value for it, was the issue, as the stability of the correct surface (relative to other possible surfaces) was shown to increase as the lattice constant reduces from the PBE calculated value to the vdW calculated value. However, subsequent DFT + vdW calculations using the PBE calculated lattice constant was shown to correctly describe the surface (at least qualitatively). As a result, even though obtaining a better description of the lattice constant and performing subsequent calculations using conventional DFT methods (a protocol proposed in this research) yields a qualitatively correct description of the surface, the exact nature of the poor description was not able to be determined beyond its relation to vdW interactions. nonetheless, if not accounted for, this failure to correctly describe the surface was shown to significantly affect interpretation of NC-AFM manipulation on the surface. It was also found that this failure appears to play a significant role in theoretical calculations, both quantitatively and qualitatively when comparing similar surfaces that differ by only a few atoms on the surface.

Instabilities on the $c(4 \times 2)$ reconstruction of the Si(001) surface with two different tip terminations (the H3 and dimer tips) were studied, by way of NC-AFM probing. By plotting the $F_z$ curves, the dissipation mechanism when imaging the surface with the NC-AFM probe was shown to be adhesion-hysteresis. Analysis of the corresponding geometries showed that dissipation in this system is due to the flipping of surface dimers, and dimers within the tip, if present. Furthermore, the amount of dissipation measured also appears to vary depending on the tip apex; large dissipation with the H3 tip and smaller (but still large) dissipation with the dimer tip. These results appear to be due to the tip termination (containing soft modes or not), the orientation of the tip (in the case of the dimer tip) and the number of atoms allowed to relax in the tip, which would lead to softer contact. As the dimer apex alone allows for greater flexibility at the point of contact, it is believed that this factor dominates. That being said, the general trends between the different tips are similar. Since large dissipation was found for both tips, dissipation must be due predominantly to soft modes in the surface, i.e. surface instabilities, specifically the flipping of surface dimers. The energy from the cantilever is injected into the surface, i.e. it goes into flipping the surface dimers, thus the cantilever driving signal must increase to maintain a constant oscillation amplitude. This means that the dissipation signal serves to probe the
interactions between the tip and surface in a specific way. As a result of this tip and surface interplay, a complex network of dimer flipping events and tip-surface reconstructions at different tip-surface separations was observed, which show that the composition of the dissipation signal, before averaging (as in experimental measurements) is in fact very complex.

Using the knowledge about the Cu(110):O surface gained from the thermodynamic stability study, the effect different tip terminations and defects have on NC-AFM imaging were also studied on the Cu(110):O surface. Here, it was shown that due to the different periodic arrangements of O and Cu atoms on the c(6 × 2) surface, or more specifically due to the presence of super Cu atoms (Cu ad-atom defects), which raise two O atoms (one either side of the super Cu atom), three different tip terminations (the atomic species terminating the tip) could be identified by simply scanning the surface. It was found that Si-terminated tips image the super Cu atoms as bright protrusions and high O atoms as dark depressions, Cu-terminated tips image only O atoms, and O-terminated tips image only Cu atoms at the typically used frequency shifts. Due to the relative ease of changing the tip to one of two different tip terminations (Cu or O), out of the most likely three, this allowed for the implementation of an on-the-fly protocol for tip identification and enforcement. It was then shown that both tip and surface species identification was not possible on the more commonly utilised p(2 × 1) reconstruction of the surface. A detailed comparison between theory and experiment was also made, where the subtle features of the different tip terminations were also discussed in the context of how they affect NC-AFM imaging. This comparison shows that the signatures of the two different terminations are distinctly different, but that confusion or ambiguity can arise due to their imaging of sub-surface features, which can be cleared up by simply imaging at a different frequency shift. Using the knowledge of the tip structure that the protocol readily allows for, lateral manipulation of the super Cu atoms on the p(2 × 1) surface was studied for both the O and Cu-terminated tips. The theoretical simulations of the manipulation was shown to be very sensitive to the descriptions of similar surfaces, as was found in the thermodynamic stability study. It was found that lateral manipulation across the rows present in the p(2 × 1) surface was not possible by the typical tip-induced barrier reduction or pulling mechanisms previously proposed. This led to the proposition of a new mechanism, called here “tip-assistance”, which can be better understood as a carrying mechanism that occurs over a single oscillation cycle. This new mechanism consists of the tip making a close approach to the surface (a tip-surface separation of nearly 2 Å), where the super Cu atom bonds to both the tip and surface, then upon tip retraction, the atom continues to be bonded to both the tip and surface until the atom is located almost directly above the wall of the row. At this point the atom may: relax into the new position across the row, with a negligible barrier; relax into its initial position, with a negligible barrier; be removed by the tip, equivalent to a vertical manipulation.
The role of defects in NC-AFM manipulation was studied on Si(001)-c(4 × 2) surface, where the surface coated with buckled dimers in anti-phase with neighbours in all directions, was manipulated, during a single manipulation, into a structure containing a phason pair (a pair of neighbouring dimers buckled in the same orientation, in anti-phase with a neighbouring pair). The manipulation effectively injects a phason pair into the c(4 × 2) surface, which can be manipulated back into the original conformation. The successful, reproducible execution of reversible manipulation of the dimers is a novel switching protocol. It was both theoretically and experimentally found that the presence of defects in the proximity of the manipulated dimers (or even as far as 30 Å away from it, along the same row) affects the ability of the dimers to be manipulated. These defects result in either effectively pinning the dimers to the surface, or allowing the manipulation to occur with a greater chance of success. In our theoretical simulations of the manipulation, it was in fact shown that without the presence of defects, i.e. on a pristine surface, it is not possible for the manipulation to occur (or at the very least it could not be found). Instead the system is trapped in an intermediate state not observed experimentally. It was also found that some defects make the transition from the pristine surface to the surface containing a phason pair even less energetically accessible. However, conversely, under the influence of surface strain, likely due to the presence of certain types of defects, the transition becomes more energetically accessible.

The accuracy and efficiency of theoretical methods has jumped by leaps and bounds over the years and now theoretical simulations and analysis of NC-AFM experiments are undoubtedly indispensable for bridging the gap between experiments and accurate physical interpretation. However, this gap will always exist unless methods, proven theoretically, enabling one to accurately characterise the tip on any surface are developed. In the NC-AFM community today, knowledge of the tip remains the most problematic issue in interpretation of NC-AFM experiments as contrast in images generally can not be unambiguously interpreted without knowledge of the tip-surface interaction. This research has shown that dissipations signals are highly sensitive to the soft modes in the tip-surface region and describe the nature of the dependence. This research has also described how tip identification can be made on the Cu(110):O-c(6 × 2) surface or Cu(110):O-p(2 × 1) surface containing super Cu atoms. Hence, either of the surfaces (the latter only if it contains super Cu atoms) could be used as a benchmark to characterise the tip apex before performing important experiments. In addition to this, this research has shown that further improvement, or at the very least caution, is required in theoretical calculations of surface properties, as conventional methods may unexpectedly fail in a non-trivial way.
Chapter 10

Conclusions

_Theoretical modelling and characterisation of concerted tip and surface effects in NC-AFM_ has been the common theme connecting this research, where the main objective was to investigate how concerted (combined) surface instabilities, defects, and tip terminations affect NC-AFM imaging and manipulation. A secondary objective was to develop a protocol for on-the-fly tip termination identification and enforcement. Both of these objectives were successfully achieved by thoroughly studying various different aspects of two very different systems, with a range of different theoretical techniques backed by experiments. This study, while addressing many issues, raises further questions and elucidates areas within the fields of NC-AFM interpretation and condensed matter physics that require further research.

A prime example of this comes out of the thermodynamics stability study presented in Chapter 4, where the vdW interaction was shown to be crucial in describing a type of system previously not thought to require consideration of non-local interactions. The important steps made in this thesis was to identify that conventional DFT methods fail to accurately describe this system and can have drastic effects on results (c.f. the diffusion calculation in Section 4.7). Furthermore, the root of this issue was identified as a lack of description of the non-local interactions. However, this is just the tip of the iceberg as in this study, although a variety of techniques were used to address the initial problem, not every possible technique or method was used to described the difference phases (e.g. hybrid functionals, larger models, full vibration calculations). This raises the question of whether the use of hybrid functionals would address the problem, which was outside both the scope and our means at the time of the study. In fact all of these are very valid questions, ones worth investigating, alongside answering the still unanswered questions of whether this issue is true for other surface, if so, which ones and what is the best way to accurately address the problem (for what surfaces or types of surfaces should theoretical modellers be more wary of non-local interactions). The first two questions raised here can answered by a time-consuming
search through all the different functionals and known surfaces (either with similar physical properties to the system studies here or all types of surfaces). The last question relates to a rapidly evolving field and the answer will likely only become clear in time, as existing methods become more robust and more methods are developed that are designed to be applied to surfaces.

Chapter 5 and Chapter 6, although they are relatively niche areas, highlight and begin to probe into two interesting areas that need to be further addressed, namely a protocol for on-the-fly tip determination/enforcement and the study of lateral manipulation across rows of atoms protruding out of the surface almost as high as the atom being manipulated, a physical barrier. For the former, there needs to be, at least, a method that allow more control than is currently available for both imaging and manipulation, as for only imaging, functionalised tips has already been shown to provide amazing results, but are inherently too unstable for manipulation. This study of the oxidised Cu(110) presents a surface that allows for such a protocol. However, one may argue that it is too limiting, if it only works on this surface. That may well be the case, but the main feature that allows the protocol to work are also identified and a whole class of surfaces suggested as potential alternatives on which the same protocol may work. It is this author’s strong believe that this must be an area of focus, to prove (or disprove) theoretically and experimentally that this can be achieved on a wider variety of surfaces. This would allow for a much better union of theory and experiments as the time spent in theoretically identifying the class of tips used experimentally (let alone the specific tip) would be shortened significantly. In addition to this, from the perspective of an experimentalist, this would make interpretation of NC-AFM measurements quicker and easier to understand.

The study of dissipation in Chapter 7 is a fairly self-contained one, but it raises the issue of understanding the single oscillation atomic processes that occur when obtaining dissipation images and the information that can be obtained from them. In this instance it is related not to jumps of atoms as previously reported, but instead to the breaking and forming of bonds at difference approach and retraction tip-surface separations and the flipping of dimers. From these calculations, it was shown that topography and dissipation images taken at relatively large frequency shifts can give rise to difference contrasts, the former giving the structure of the manipulated surface and the latter giving the unperturbed structure. The further areas of research this study invokes are purely scientific curiosities, but may be found to be useful. An example of this is that although experimentally the atomic processes are averaged over many oscillations, by comparing the two images (topography and dissipation) it is possible to meaningfully speculate as to what the general atomic processes are in manipulation experiments (achieved via scanning, not spectroscopy). This is because during the manipulation simultaneous images of the unperturbed and manipulated surface can be obtained and compared, as opposed to imaging, manipulating, then imaging again, where between the steps it is possible for the surface to change (not likely but still possible).
Furthermore, it has been shown that it may be possible to crudely characterise the tip as containing soft modes or not. Hence another possible extension of this study could be to investigate if it is indeed possible to characterise them in this way when imaging semiconductor surfaces like Si with an Si tip (or a Ge surface with a Ge tip etc...). If so, then this would dramatically reduce the search through the different tip models that agree with the experimental signatures.

In the final study presented in Chapter 8 the role of defects in successful NC-AFM manipulation was studied and found to likely be the main factor in allowing for successful single dimer manipulation (which generate a phason pair). This study however investigates different defective surfaces only up to the point of determining whether they can allow for the successful manipulation, but the defective surfaces themselves are not further studied nor their roles further quantified. A very interesting continuation of this research would be to study a boron defect, as is observed experimentally, and its affects both the row it is on and neighbouring rows, as the propensity for the rows to flip was experimentally found to be affected. Such a study could then be extended to studying the effects on dimer flipping of different defects commonly present on an Si(001) surface and to similar studies on similar surfaces, i.e. the Ge surface. In addition to this, and most importantly in an ideal world, these studies should include different, larger tip models. A variety of studies, such as the ones mentioned, could elucidate the importance and generality of the role defects have in the manipulation of such surfaces.

The studies presented in this thesis have further demonstrated the crucial role theory has in our interpretation of experimental NC-AFM results in a meaningful physical context and has suggested means to tackle the most problematic area in NC-AFM interpretation, tip identification. The NC-AFM is without doubt a tool that will prove pivotal in future scientific advancements in nanotechnology. So each step towards our better physical understanding and interpretation of NC-AFM experiments is a step towards the future.

“If I have seen further, it is by standing on ye shoulders of giants.”

Sir Isaac Newton

Physicist and Mathematician (1642-1727)
Part V

Bibliography
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


