

***Ab initio* Modelling of Low-Dimensional Systems**

Trevor Peter Hardcastle

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

Institute for Materials Research
School of Process, Environmental and Materials Engineering

University of Leeds

April 2014

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Trevor Peter Hardcastle to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2014 The University of Leeds and Trevor Peter Hardcastle

Acknowledgements

This research has been carried out by a team which has included Andrew Scott, Che Seabourne, Rik Brydson, Uschi Bangert, Quentin Ramasse, Recep Zan, Ken Livi, Konstantin Novoselov, Demie Kepaptsoglou, Keith Refson and Barbara Montanari. Sincere thanks are due to Matt Probert and Phil Hasnip for sharing their expertise on electronic structure modelling, and to Stefan Wendt for providing guidance in the interpretation of scanning tunnelling microscopy data. I gratefully acknowledge the use of high performance computing facilities in this work, including the ARC1, ARC2 and N8 Polaris machines at the University of Leeds and the Hartree machine at STFC Laboratories in Daresbury. I'd like to thank the UK EPSRC for the Doctoral Training Award which funded this research, and also myself for eventually deciding not to write a gushing acknowledgements section.

Chapters 1, 2 and 6: Written solely by T. P. Hardcastle with feedback and comments from A. J. Scott and R. M. D. Brydson.

Chapter 3: Written solely by T. P. Hardcastle with feedback and comments from A. J. Scott, R. M. D. Brydson, C. R. Seabourne, U. Bangert, Q. M. Ramasse and R. Zan. STEM images from Figure 3.1 obtained by Q. M. Ramasse at SuperSTEM Laboratory, Daresbury.

Chapter 4: Written solely by T. P. Hardcastle with feedback and comments from A. J. Scott, R. M. D. Brydson, C. R. Seabourne and K. J. T. Livi. STM image from Figure 4.13 obtained by group of S. Wendt at Aarhus University and reproduced from Ref. 114.

Chapter 5: Written solely by T. P. Hardcastle with feedback and comments from A. J. Scott, R. M. D. Brydson, K. Refson, B. Montanari, C. R. Seabourne, U. Bangert, Q. M. Ramasse and D. Kepaptsoglou.

Appendices A and B: Written solely by T. P. Hardcastle with feedback and comments from A. J. Scott, R. M. D. Brydson, C. R. Seabourne and K. J. T. Livi.

List of publications

The candidate is a co-author on the following articles which were published during the period November 2010 - March 2014:

["Energy of Step Defects on the TiO₂ Rutile \(110\) Surface: an *ab initio* DFT Methodology"](#)

T. P. Hardcastle, C. R. Seabourne, K. Livi, R. M. D. Brydson and A. J. Scott. *The Journal of Physical Chemistry C* **117**, 23766 (2013)

["Mobile metal adatoms on single layer, bilayer and trilayer graphene: an *ab initio* DFT study with van der Waals corrections correlated with electron microscopy data"](#)

T. P. Hardcastle, C. R. Seabourne, R. Zan, R. M. D. Brydson, U. Bangert, Q. M. Ramasse, A. J. Scott and K. S. Novoselov. *Physical Review B* **87**, 195430 (2013)

["Atomic-Scale Surface Roughness of Rutile and Implications for Organic Molecule Adsorption"](#)

K. J. T. Livi, B. Schaffer, D. Azzolini, C. R. Seabourne, **T. P. Hardcastle**, A. J. Scott, D. Sverjensky, R. M. Hazen and R. M. D. Brydson. *Langmuir* **29**, 6876 (2013)

["*Ab initio* modelling, polarity and energetics of clean rutile surfaces in vacuum and comparison with water environment"](#)

T. P. Hardcastle, R. M. D. Brydson, K. J. T. Livi, C. R. Seabourne and A. J. Scott. *Journal of Physics: Conference Series* **371**, 012059 (2012)

List of conference presentations

Invited oral presentations:

- SuperSTEM meeting, STFC Laboratories, Daresbury, March 2013

Contributed oral presentations:

- NanoteC13, University of Surrey, August 2013
- Theory, Modelling and Computational Methods for Semiconductors (TMCS IV), University of Salford, January 2014
- The International Conference on Metallurgical Coatings and Thin Films* (ICMCTF), San Diego, April 2014. ***Graduate Student Award Finalist**

Contributed poster presentations:

- Materials Research Society Conference (MRS 2014), San Francisco, April 2014
- CASTEP workshop, Frankfurt, September 2012
- European Microscopy Congress (EMC), Manchester, September 2012
- Electron Microscopy and Analysis Group Conference (EMAG), University of Birmingham, September 2011

Abstract

In this thesis, theoretical electronic structure methods have been used to study systems of interest to materials science and engineering. The many-body problem of quantum mechanics has been reviewed, and it has been explained that the density functional theory (DFT) of Hohenberg, Kohn and Sham is a very practical approach to solving it. The core details of DFT have been explicitly laid out. Having then demonstrated an awareness of the many different powers and varied capabilities of DFT in predicting material properties, systems laying at the current frontiers in nanoelectronics (chapters 3 and 5) and theoretical surface science (chapter 4) have then been focused on. In chapter 3, the behaviour of metal adatoms on graphene substrates has been predicted using DFT. From adatom binding energy and migration energy calculations, it has been theoretically suggested that single Cr, Au and Al adatoms diffuse randomly on graphene at room temperature until they collide with edge sites or defects, where they form stable bonds. This prediction has been used to explain experimental electron microscopy data which shows that metal adatoms evaporated onto graphene by chemical vapour deposition (CVD) have only ever been observed at edge sites and defects, and never on the pristine regions. In chapter 4, a new methodology has been developed for predicting the energies of step defects on crystalline solid surfaces, and it has been applied to steps on the (110) surface of TiO₂ rutile. The limitations of currently published methods of calculating step energies have been explained in detail, and it has been demonstrated that this new method is much more reliable. The method has been used to predict the shape of a terrace island on the (110) surface of TiO₂ rutile, and the prediction has been found to compare well with published experimental electron microscopy data. In chapter 5, current progress on an ongoing project has been summarised which investigates whether there is an energetic advantage to multiple substitutional nitrogen dopants in graphene occupying the same sublattice. The results are inconclusive so far, although it has been shown so far that magnetic effects are unlikely to be playing a role. In chapter 6, the accomplishments of this thesis have been summarised and future directions suggested.

Table of Contents

Acknowledgements.....	3
List of publications	4
List of conference presentations.....	5
Abstract.....	6
Table of Contents	7
List of Tables	10
List of Figures	11
List of Abbreviations.....	14
List of Symbols	15
Preface	17
Chapter 1: Introduction.....	18
1.1 Electronic structure theory and its relevance to materials engineering	18
1.2 Aims and objectives of this thesis	19
Chapter 2: Density functional theory.....	23
2.1 The many-body problem	23
2.2 What actually is a functional anyway?.....	31
2.3 Hohenberg-Kohn theorems.....	36
2.4 The Kohn-Sham equations.....	40
2.5 Periodicity.....	47
2.6 Practical calculations: finite plane wave basis set	57
2.7 Practical calculations: pseudopotentials.....	62
2.8 Non-periodic systems under periodic boundary conditions	66
2.9 Concluding remarks	67
Chapter 3: Mobility of metal adatoms on graphene	70
3.1 Abstract.....	70
3.2 INTRODUCTION.....	71
3.2.1 Recent theoretical studies of graphene-metal systems.....	71
3.2.2 Electron microscopy studies of graphene-metal systems	74
3.2.3 Long-ranged lattice perturbations.....	76
3.2.4 Binding sites, binding energy and dipole corrections	78
3.3 METHOD.....	80

3.3.1 Basis set parameters.....	80
3.3.2 Preliminary structure optimisation	89
3.3.3 Binding energy calculations.....	95
3.3.4 Migration energy calculations.....	101
3.4 RESULTS AND DISCUSSION.....	102
3.4.1 Binding energies and bond distances.....	102
3.4.2 Mobility of metal adatoms.....	106
3.4.3 Possible effects of the electron beam.....	107
3.5 CONCLUSIONS	108
Chapter 4: Steps on the TiO₂ rutile (110) surface	110
4.1 Abstract.....	111
4.2 INTRODUCTION.....	111
4.2.1 The importance of steps on TiO ₂	111
4.2.2 Existing studies of steps on TiO ₂	112
4.2.3 Step energy calculations are difficult because of quantum size effects	114
4.2.4 Linear fitting methods.....	114
4.2.5 Methodology.....	116
4.3 METHOD.....	122
4.3.1 Basis set, exchange correlation functional, bulk reference state and vacuum thickness	122
4.3.2 Slab calculations	123
4.4 RESULTS AND ANALYSIS.....	129
4.4.1 Initial remarks.....	129
4.4.2 $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ steps.....	136
4.4.3 $\langle 1\bar{1}1 \rangle_R$ and $\langle 1\bar{1}1 \rangle_R$ steps: general comments	137
4.4.4 $\langle 1\bar{1}1 \rangle_R$ step: detailed analysis.....	138
4.4.5 $\langle 1\bar{1}1 \rangle_R$ step: detailed analysis.....	139
4.4.6 Comparison with published STM data.....	140
4.5 CONCLUSIONS	145
Chapter 5: Ordering of substitutional N dopants in graphene	147
5.1 Abstract.....	147
5.2 INTRODUCTION.....	149
5.2.1 Experimental observations and suggested theoretical investigations	149
5.3 METHOD.....	150

5.3.1. Basis set and computational considerations	150
5.3.2. Calculation details: Single N dopant.....	152
5.3.3. Calculation details: Two N dopants	153
5.4 RESULTS.....	155
5.4.1. Single N dopant: initial remarks.....	155
5.4.2. Two N dopants: there is no apparent intrinsic energetic advantage to N dopants collecting in one sublattice	158
5.5. Proposed improvements to this project	162
Chapter 6: Concluding remarks	165
List of References	170
Bibliography	178
Appendix A : Uncertainty analysis for data in chapter 4	179
A.1 Statistical uncertainty in step energies.....	179
Appendix B : Statistical fitting methods for surface energies	183
B.1 Using linear regression to evaluate surface energies is flawed .	183
B.2 A nonlinear alternative	185
B.3 Application: merits and limitations.....	190

List of Tables

Table 3.1. Recently published DFT-calculated Au adatom/graphene binding energies on the 3 high symmetry adsorption sites of single layer graphene.	72
Table 3.2. (Ad)atom binding energies metal-carbon distances associated with the fully relaxed structures.....	105
Table 3.3. Calculated migration barriers for Au, Cr and Al on the lowest energy migration pathways on pristine monolayer, bilayer and trilayer graphene.	107
Table 4.1. The convention used to define the number of layers, L, in terms of the number of TiO_2 units, m.	126
Table 4.2. The (110) terrace surface energy per unit area, step energy per unit length and step-step interaction constant on the four vicinal surface types.....	136
Table 5.1. Single N dopant: final magnetic moments, energies and forces.	156
Table 5.2 Preliminary single N dopant calculations carried out to try to identify themes which could be influencing the results. ...	163
Table A.1. The sensitivity of γ_{110}, β and q to changes in supercell energy.	182
Table B.1. The dependence on the chosen data range of the intercept of a linear fit of total slab energies against the number of atoms.	185
Table B.2. Supercell energies calculated using CASTEP divided into two branches.....	192

List of Figures

Figure 2.1. <i>The electron density field is a single-valued function of r.</i>	33
Figure 2.2. <i>The input variables of a functional F.</i>	33
Figure 2.3. <i>The difference between d_n and δn.</i>	34
Figure 2.4. <i>A regular 4×4 Monkhorst Pack k points grid defined in a 2 dimensional hexagonal lattice.</i>	59
Figure 2.5. <i>The kinetic energy cutoff circle in 2 dimensions for a reciprocal hexagonal lattice.</i>	60
Figure 2.6. <i>All-electron wavefunction and potential compared with pseudo-wavefunction and pseudopotential.</i>	65
Figure 2.7. <i>Ball-and-stick representations of periodic and non-periodic systems modelled under periodic boundary conditions.</i>	67
Figure 3.1. <i>STEM HAADF images at 60 keV showing preferential binding of metal atoms to edge defects, hydrocarbon-contaminated regions and metal clusters.</i>	74
Figure 3.2. <i>Symmetry-breaking caused by fixing atoms on the supercell perimeter without appealing to lattice symmetries.</i>	78
Figure 3.3. <i>The high symmetry adsorption sites located at the vertices of the symmetry-reduced Wigner Seitz cell boundaries.</i>	79
Figure 3.4. <i>Periodic cells used for basis set.</i>	82
Figure 3.5. <i>Monotonic convergence of total supercell energies with increasing kinetic energy cut off for graphene and metal-graphene systems.</i>	83
Figure 3.6. <i>Monotonic convergence of total supercell energies with increasing kinetic energy cut off: metal atoms.</i>	84
Figure 3.7. <i>Convergence of binding energies with kinetic energy cut off.</i>	85
Figure 3.8. <i>Non-monotonic convergence of total supercell energies with decreasing in-plane k point spacings: graphene and metal-graphene systems.</i>	86
Figure 3.9. <i>Non-monotonic convergence of total supercell energies with decreasing in-plane k point spacings for the metal atoms.</i>	87
Figure 3.10. <i>Convergence of binding energies with decreasing in-plane k point spacings.</i>	88

Figure 3.11. <i>Convergence of graphene supercell energies to determine vacuum spacing required to decouple intercellular graphene-graphene interactions along the vacuum direction.</i>	91
Figure 3.12. <i>The changes in energy between successive total supercell energy calculations of a single metal atom in cubic supercells.</i>	92
Figure 3.13. <i>The convergence of metal adatom + single layer graphene total supercell energy with vacuum thickness.</i>	93
Figure 3.14. <i>The 3 single layer supercells before geometry optimisation used for the adatom + graphene systems for (a) site A, (b) site B and (c) site H.</i>	96
Figure 3.15. <i>The multilayer input supercells for (a) site A₁, (b) site A₂ (c) site B, and (d) site H.</i>	98
Figure 3.16. <i>Example supercells used for metal binding to monolayer edges with accompanying unit cells.</i>	99
Figure 3.17. <i>Example supercells used for migration activation barrier calculations.</i>	99
Figure 3.18. <i>The calculated binding energy for metal atoms adsorbed on the pristine substrates and bound at monolayer edge sites.</i>	101
Figure 3.19. <i>Electron density images showing the difference in bonding character between adsorption and edge sites for Au...</i>	103
Figure 4.1. <i>The edge of a terrace island bound by [001]- and [1̄11]- direction step structures on the stoichiometric (110) surface of TiO₂ rutile.</i>	113
Figure 4.2. <i>Cartoon illustrations of the physical features corresponding to the terms in equation (4.1).</i>	117
Figure 4.3. <i>Ball-and stick model of a semi-infinite periodic (430) surface of TiO₂ rutile, consisting of (110) terraces and <001>_{Ti} steps.</i>	120
Figure 4.4. <i>Top-down views of the projection of the high-index supercells onto two terraces of a semi-infinite (110) surface bounded by (a) <001>_O (b) <001>_{Ti} (c) <1̄11> and (d) <1̄11>_R steps.</i>	124
Figure 4.5. <i>Side views of the set of 24 supercells used for the <001>_O step using ball and stick representation.</i>	125
Figure 4.6. <i>The physical reason for the oscillations in the energetic components plotted in Figures 4.8, 4.9 and 4.10: thickness-dependent alternating configurations of electron density.</i>	127
Figure 4.7. <i>The convergence of surface energy per unit area with slab thickness.</i>	129
Figure 4.8. <i>Layer-by-layer cross-check.</i>	131

Figure 4.9. <i>The convergence of step energy per unit length with increasing slab thickness for the four steps studied.....</i>	132
Figure 4.10. <i>The convergence of the step-step interaction constant with increasing slab thickness for the four steps studied.</i>	133
Figure 4.11 <i>The total surface energy of the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ slabs resolved into terrace, step and step-step interaction components.....</i>	134
Figure 4.12 <i>The total surface energy of the $\langle 1\bar{1}1 \rangle$ and $\langle 1\bar{1}1 \rangle_R$ slabs resolved into terrace, step and step-step interaction components.....</i>	135
Figure 4.13. <i>Experimental scanning tunnelling electron microscopy data.....</i>	141
Figure 5.1. <i>A ball-and-stick illustration of substitutional nitrogen dopants in graphene, with nitrogen indicated in blue. In this example, 5 out of 7 of the dopants lay in the same sublattice. ...</i>	148
Figure 5.2. <i>Convergence of supercell energy differences with k points sampling.....</i>	152
Figure 5.3. <i>The irreducible Wigner-Seitz cell, whose boundaries are indicated in red, of the 6x6 supercell showing all unique locations for the second substitutional N dopant.....</i>	154
Figure 5.4. <i>Spin density slices showing sublattice asymmetry for (a) LDA, (b) PBE, (c) HSE and (d) sX-LDA.....</i>	157
Figure 5.5. <i>Relative total energies of the various configurations resolved into sublattices A and B.....</i>	159
Figure 5.6. <i>Final magnetic moments of the various configurations resolved into sublattices A and B.....</i>	161
Figure B.1. <i>The flawed nature of evaluating surface energy using a linear fit of supercell energies against slab size.....</i>	184
Figure B.2. <i>Supercell slab energies plotted against increasing supercell size with a sketch of the best fit functions $\gamma^\uparrow(m)$ and $\gamma^\downarrow(m)$.....</i>	186

List of Abbreviations

BOA: Born Oppenheimer approximation

DFT: density functional theory

HK: Hohenberg-Kohn

KS: Kohn-Sham

LDA: local density approximation

GGA: generalised gradient approximation

SCF: self-consistent field

BZ: Brillouin zone

PBC: periodic boundary conditions

PBE: Perdew-Burke-Ernzerhof (parametrisation of GGA)

HAADF: high-angle annular dark field

STEM: scanning transmission electron microscopy

TS: Tkatchenko-Scheffler (dispersion correction scheme of)

LST: linear synchronous transit

QST: quadratic synchronous transit

STM: scanning tunnelling microscopy

QSE: quantum size effects

MP: Marchenko-Parshin (step-step interaction potential of)

(L)BFGS: (low memory) Broyden-Fletcher-Goldfarb-Shanno (geometry optimisation algorithm of)

EDFT: ensemble DFT minimisation

List of Symbols

Ψ	many-body wavefunction
\vec{r}_i	position coordinate of i^{th} electron
\vec{R}_I	position coordinate of I^{th} nucleus
t	time (absolute, non-relativistic)
\hat{H}	Hamiltonian operator
Z_I	charge of I^{th} nucleus
M_I	mass of I^{th} nucleus
m_e	mass of electron
\hat{T}	many-electron kinetic energy operator
\hat{V}_{ext}	many-electron external nuclear potential operator
\hat{V}_{int}	many-electron internal electron potential operator
\hat{n}	total electron density operator
$n(\vec{r})$	total electron density
$n_0(\vec{r})$	ground state total electron density
ψ	many-electron wavefunction
χ_i	individual electron state
ψ^*	conjugate many-electron wavefunction
$\varphi_i(\vec{r})$	i^{th} Kohn-Sham wavefunction
$\varphi_i^{ps}(\vec{r})$	i^{th} Kohn-Sham pseudowavefunction
$\phi_{ps}(r)$	radial component of Kohn-Sham pseudowavefunction
E	total system energy
$V_{ext}(\vec{r})$	external nuclear potential
$V_{ps}(\vec{r})$	external nuclear pseudopotential
$T(n(\vec{r}))$	many-electron kinetic energy functional
$T_S(n(\vec{r}))$	Kohn-Sham kinetic energy functional
$E_{Hartree}(n(\vec{r}))$	Hartree energy functional
$E_{XC}(n(\vec{r}))$	exchange-correlation energy functional
$\varepsilon_{XC}(n(\vec{r}))$	exchange-correlation energy density functional

$E_{ext}(n(\vec{r}))$	external nuclear potential energy functional
V_{KS}^{ps}	effective Kohn-Sham pseudopotential
E_{II}	nuclear Coulomb energy
N	total number of Kohn-Sham states/electrons
ξ_i	energy eigenvalue of i^{th} Kohn-Sham state
H_{KS}	effective Kohn-Sham Hamiltonian
\vec{R}	direct space Bravais lattice vector
\vec{G}	reciprocal space Bravais lattice vector
$\hat{T}(\vec{R})$	Bravais lattice translation operator
$u_{\vec{k}}(\vec{r})$	cell-periodic Bloch function
Ω_{BZ}	Brillouin zone volume
Ω_{cell}	Wigner-Seitz cell volume
$c_{\vec{G}}$	\vec{G} vector coefficient of Fourier mode
s^i	k-point spacing along i^{th} reciprocal lattice vector
E_{cut}	kinetic energy cutoff
k_B	Boltzmann constant
T	temperature
E_b	binding energy
γ_{hkl}	surface energy per unit area of (hkl) plane
θ_{hkl}	angle between (hkl) plane and low-index reference plane
β	step energy per unit length
d_{hkl}	step-step distance on surface (hkl) in low-index plane
q	step-step interaction constant
ε_{hkl}	total surface energy in nominal supercell on (hkl) surface
ε_t	terrace component of total surface energy
ε_s	step component of total surface energy
ε_{ss}	step-step interaction component of total surface energy

Preface

In this thesis, theoretical electronic structure methods are employed to try to explain experimental electron microscopy observations of materials of current interest to materials science and engineering. Particular emphasis is placed on designing calculation methods with a view to differentiating between artificial aspects of the calculations which have the potential to be misleading, and aspects which can be used to motivate robust and valid physical arguments of real explanatory merit. The entire thesis was written by myself.

Trevor P. Hardcastle

Institute for Materials Research, Leeds (April 2014).

Chapter 1: Introduction

1.1 Electronic structure theory and its relevance to materials engineering

Materials, and their remarkable properties, play an extremely important role in many areas of engineering and technology. Much of the advancement in these fields relies on successfully exploiting material properties in particular contexts, such as using semiconducting materials to manipulate and systematically control electric currents and thereby open the door to modern electronics. With the exception of radioactive isotopes, most of the chemical, mechanical, magnetic and electrical properties of materials can be very well explained in terms of the behaviour of their constituent electrons alone. Therefore, gaining an understanding of the behaviour of electrons in materials is one of the most pressing objectives in theoretical materials science, because such an understanding forms the foundation upon which to explain and predict material properties which can ultimately lead to new technologies. Nowadays, all credible theories which attempt to understand electrons in materials are rooted in quantum mechanics. Generally speaking, material properties which emerge directly from the behaviour of electrons can be divided into two categories: properties which emerge from electrons in the quantum mechanical *ground state*, and those which emerge from electrons in *excited states*. Ground state properties include cohesive energy, equilibrium crystal structures and the phase transitions between them, electric and magnetic susceptibilities, and mechanical properties such as elastic constants. Excited state properties include the band gap in insulators and semiconductors as determined by the energy gap between ground and excited states, the interactions of materials with electromagnetic waves, e.g. photon-induced surface plasmons in metals, and emission spectra determined by energy gaps between electron states. Physical theories which deal with the quantum mechanical states of electrons in materials are most generally referred to as theories of *electronic structure*. The motivational premise of this thesis is that the application of electronic structure methods

to materials used in engineering and technology can result in significant advancement in these fields. The specific goal of this thesis is to try to *demonstrate* that this is indeed true by applying electronic structure methods to three systems, and to then argue that the insight gained from these investigations results in significant, and, most importantly, *new* understanding of these systems. The specific aims and objectives of the thesis are now outlined.

1.2 Aims and objectives of this thesis

There are two overriding aims of this thesis. The first aim of this thesis is to demonstrate to the reader that the author has a robust understanding of the methods of electronic structure theory which have been applied in this thesis. This is obviously a necessary prerequisite because it adds credibility to the results and conclusions drawn from the investigations carried in the subsequent chapters. To achieve this aim, the many-body problem of quantum mechanics will be outlined, and an awareness of the various formalisms which have been developed to try to solve it will be demonstrated. Density functional theory^{1,2} (DFT) - the core method used on all systems studied in the thesis - will then be introduced, and the significant practical advantages of DFT over many-body methods will be explained. The Kohn Sham (KS) equations² of DFT will be derived in full with explanations of all of the mathematical steps, using Schrödinger's wave equation as the starting point. The KS equations will then be modified so that they can be applied to periodic crystalline systems by appealing to some basic crystallography, again with full commentary of the mathematical steps. It will be explained that imposing periodic boundary conditions offers significant practical advantages, and that the periodic KS equations can be solved for a vast array of both periodic *and non-periodic* systems in order to determine their properties. Finally, the basic conventions for solving the periodic KS equations numerically in a practical way using computers will then be outlined in a non-rigorous way.

The second aim of this thesis is to then apply the periodic KS equations of DFT to low-dimensional systems of current interest to surface science and

nanoelectronics, and to demonstrate that novel and significant insight emerges from these investigations which has the power to both increase scientific understanding and influence future directions of technological advancement. In order to achieve this aim, continued emphasis is placed throughout the entire thesis on carefully designing the DFT calculations and evaluating the extent to which the model outputs can be validly compared with published experimental observations of the systems of interest. Due to the wide availability of well-developed commercial DFT codes nowadays, it is very easy to perform high level calculations using DFT. However, with great power comes great responsibility. There exists a great responsibility with theoretical modellers to try to ensure that output results obtained from a DFT code can be used to reliably explain experimental observations, or even better, predict them. Theoretical calculations must be robust despite the necessary approximations, and they must be carefully designed so that they simulate the system observed in an experiment as faithfully as possible. In addition to this, an experimental observation of a system of interest may be affected by the fact that in order to observe the system, one must *interact* with it which may affect its physical state. Thus, the most pressing objective when applying established theoretical methods of electronic structure is not just to merely execute a code, but to give a full account of all of the possible discrepancies between the code output, the available experimental data and the factors which may influence it, and the true system that both of these approaches are trying to shed light on. With this in mind, the specific investigations carried out in chapters 3, 4 and 5 are now described.

In chapter 3, a theoretical investigation into the interaction between graphene and metals is conducted. For graphene to be integrated into future nanoelectronic circuits, it needs to be interfaced - contacted - with materials used in current devices, like highly conductive metals. A proper understanding of the behaviour of metal atoms on graphene substrates would be needed in order to determine the stability of such interfaces and the prospects for long-term large-scale manufacture. The specific objective of chapter 3 is to begin to assess this, starting at a basic level, by investigating the types of bonds that metals form with graphene, and to predict how single metal atoms behave on graphene. To achieve this,

periodic ground state DFT calculations, i.e. zero Kelvin calculations, will be carefully designed and applied to try to predict the behaviour of single metal adatoms on graphene. Recently obtained experimental electron microscopy images obtained by collaborators will then be presented which show metal-graphene systems *at room temperature*, and the conclusions drawn from the calculations will be compared with the experimental observations. Since the theoretical model differs from the system observed experimentally by a temperature difference of about 300 K, a careful and conservative assessment of the extent to which the calculations are representative of the experimental data is given.

The aim of chapter 4 is to develop and apply a new and general theoretical method of characterising and understanding the nature of step defects on solid crystalline surfaces. Knowing the structure of surfaces is of widespread importance. Many of the processes of interest to new technologies take place at surfaces and interfaces, such as diodes which derive their functionality from the existence of an interface, or heterogenous catalytic reactions which occur at surfaces. Specifically, a new theoretical method is designed to be used in conjunction with electronic structure codes for calculating the energies of step defects and the energies associated with the interactions between neighbouring steps. This method will be explained in full, and it will be explained why a widely published method of calculating the energies of step defects on solid surfaces can sometimes fail *dramatically* depending on which material is studied. It will be argued in detail why the method presented in this thesis is a significant improvement over this widely published method. This will be achieved by directly carrying out DFT simulations of stepped TiO₂ rutile surfaces, and then applying the flawed published method to the output data and showing that the resulting step energy depends sensitively on the arbitrary range of output data chosen to be included in the analysis. The improved method developed in this thesis will then be tested by applying it to the same DFT data and then predicting the shape of a terrace island for direct comparison with published experimental electron microscopy data showing an annealed TiO₂ rutile surface at 120 K. As in chapter 3, a careful assessment of the validity of comparing the zero-temperature DFT output data with the system observed

experimentally is given, along with an attempt to account for defects which are visible in the experimental data but which were not accounted for theoretically, the statistical uncertainty associated with step energy calculations which are extremely sensitive to the output values, and finally, any possible consequences of having neglected smaller structures like kinks.

In chapter 5, a project which is incomplete but ongoing at the time of writing is presented which focuses on substitutional nitrogen dopant atoms in graphene. In some recent experimental microscopy publications³⁻⁵ it has been noticed (although some collaborators are currently not convinced by current experimental data) that there seems to be an overall tendency for substitutional nitrogen dopants in graphene to be observed mostly in only one of the two carbon sublattices. This observation suggests the possible existence of a strong long-ranged N-N interaction causing collective sublattice binding to occur, perhaps during synthesis. Besides this being a very interesting apparent phenomenon in its own right, it also has immediate implications for the possibility of N/graphene-based nanoscale devices which may derive their functionality from particular configurations of dopants. The aim of chapter 5 is to summarise the current understanding of the outputs obtained from theoretical calculations of nitrogen-doped graphene systems, and to outline proposed improvements for further calculations. Preliminary calculations carried out so far suggest that there is *not* an energetic advantage to collective sublattice binding of substitutional nitrogen dopants; in fact, the opposite is suggested by the current results. However, there are still many reasons to suggest that the calculations so far performed may not yet be properly representative of the systems being observed experimentally, although there is no *a priori* reason for assuming that more stringently converged calculations would give different results. To conclude chapter 5, proposed refinements to the calculations are briefly outlined.

Finally, chapter 6 aims to summarise the key findings of this thesis, to explain why these findings are novel and significant, and to anticipate future directions of further enquiry. Firstly though, chapter 2 now commences with a discussion of basic electronic structure theory.

Chapter 2: Density functional theory

In this chapter, the many body problem of quantum mechanics is outlined, the Hohenberg-Kohn theorems are proved and the Kohn-Sham equations are derived. The Kohn-Sham equations are then expressed in the context of periodic solids. Practical aspects of solving the equations, such as using a finite plane-wave basis and using pseudopotentials are then discussed.

2.1 The many-body problem

Consider an isolated neutral system of N electrons and M atomic nuclei which are all interacting with each other. The mathematical object which can be used to represent them is the quantum mechanical many-body wavefunction, $\Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N, \vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M, t)$, where \vec{r}_i denotes the position vector of the i^{th} electron, \vec{R}_I denotes the position of the I^{th} nucleus, and t is time. The Schrödinger equation states that the many-body wavefunction Ψ is an eigenstate of the Hamiltonian operator \hat{H} if it satisfies

$$i\hbar \frac{d\Psi}{dt} = \hat{H}\Psi = E\Psi, \quad (2.1)$$

where E is the total energy of the electrons + nuclei system. Following the notation of Martin⁶, (followed loosely throughout this whole chapter) the Hamiltonian operator of this system can be written as:

$$\hat{H} = -\sum_i^N \frac{\nabla_i^2}{2} - \sum_{i,I}^{N,M} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_{I \neq J}^M \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} + \frac{1}{2} \sum_{i \neq j}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} - \sum_I^M \frac{\nabla_I^2}{2M_I}, \quad (2.2)$$

where Hartree atomic units have been used so that $\hbar = e = m_e = 4\pi/\epsilon_0 = 1$, M_I denotes the mass of the I^{th} nucleus, and Z_I denotes the charge of the I^{th} nucleus. Each term in equation (2.2) has a very clear physical identity. The first term is the sum of the individual kinetic energy operators of the N electrons. The second term corresponds to the interaction of each and every electron with each and every nucleus, the third term corresponds to the interactions amongst the nuclei, and the fourth term to the interactions amongst the electrons. The fifth term gives the kinetic energy operators of

the M nuclei. In the third term, the interaction of a given nucleus with itself is excluded from the summation by summing over all I, J except the cases for which $I = J$, and double-counting of the total number of nucleus-nucleus interactions is avoided by multiplying the bare sum by $1/2$. Identical reasoning applies to the electron-electron term.

The general solution in the form of the many-body wavefunction Ψ is practically impossible to find analytically in virtually all cases, so to begin to make the problem of solving equation (2.1) tractable, the Born-Oppenheimer approximation⁷ (BOA) is used. The detailed derivation is bypassed here, but the BOA neglects the motion of the nuclei in the ground state on the basis that $M_I \gg m_e$, and treats the nuclei as a static array of point charges such that the fifth term from the Hamiltonian of equation (2.2), which is proportional to $1/M_I$ and therefore comparatively small, can be neglected (or, strictly speaking, treated as a small perturbation and ignored):

$$-\sum_I^M \frac{\nabla_I^2}{2M_I} \xrightarrow{BOA} 0. \quad (2.3)$$

Under the BOA, the Hamiltonian in equation (2.2) therefore simplifies to (where N and M are also now dropped)

$$\hat{H} = -\sum_i \frac{\nabla_i^2}{2} - \sum_{i,I} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|}. \quad (2.4)$$

Equation (2.3) implies that the nuclear position vectors \vec{R}_I are fixed, so all explicit \vec{R}_I -dependence of the nuclear interaction terms can be withdrawn from the wavefunction Ψ :

$$\Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N; \vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M, t) \xrightarrow{BOA} \Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N, t). \quad (2.5)$$

and instead expressed more simply as fixed but adjustable parameters on which the energy eigenvalue $E = E(\vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M)$ and Hamiltonian $\hat{H} = \hat{H}(\vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M)$ depend. Also, it is clear that the spatial and temporal variables of Ψ can be separated by writing it as the following product of spatial and temporal components:

$$\Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N, t) = e^{\frac{-iEt}{\hbar}} \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N), \quad (2.6)$$

because the Schrödinger equation (2.1) is clearly satisfied by this expression:

$$\begin{aligned} i\hbar \frac{d}{dt} \left(e^{-\frac{iEt}{\hbar}} \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N) \right) &= E e^{-\frac{iEt}{\hbar}} \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N) \\ &= E\Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N, t) \end{aligned} \quad (2.7)$$

Equations (2.6) and (2.7) are important, because they show that the time-independent electronic wavefunction ψ can generally be extracted for *any* system, even if its explicit form is not known. The total system energy E is equal to the expectation value of the Hamiltonian operator, and E is also clearly time-independent by virtue of the variable separation shown in equation (2.6):

$$\begin{aligned} E &= \langle \Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N, t) | \hat{H}(\vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M) | \Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N, t) \rangle \\ &= \langle \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N) \left| e^{\frac{iEt}{\hbar}} e^{-\frac{iEt}{\hbar}} \hat{H}(\vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M) \right| \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N) \rangle \\ &= \langle \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N) | \hat{H}(\vec{R}_1, \dots, \vec{R}_I, \dots, \vec{R}_M) | \psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N) \rangle, \end{aligned} \quad (2.8)$$

where Dirac notation has been adopted. The key point of this discussion, made manifestly clear by the final line of equation (2.8), is that under the BOA, the only quantum-mechanical quantities on which the total system energy E now depends are the spatial coordinates of the electrons, $\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N$. The problem of a time-dependent system of $N + M$ quantum-mechanically-interacting electrons and nuclei has therefore been reduced to a time-independent problem of N quantum-mechanically interacting electrons. Note that spin coordinates have been omitted from ψ in this simple analysis to keep the notation simple, but should be regarded as implicit throughout the entire chapter.

To make the notation less verbose, the nuclear coordinate-dependence of \hat{H} and the electron coordinate-dependence of ψ are now omitted except where needed, and the following simplifying notation is adopted:

$$\hat{T} = - \sum_i \frac{\nabla_i^2}{2}, \quad (2.9)$$

$$\hat{V}_{ext} = - \sum_{i,I} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|}, \quad (2.10)$$

$$\hat{V}_{int} = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|}, \quad (2.11)$$

and

$$E_{II} = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|}, \quad (2.12)$$

so that equation (2.8) can be written explicitly but more compactly as

$$\begin{aligned} E &= \langle \psi | \hat{H} | \psi \rangle = \langle \psi | \hat{T} + \hat{V}_{ext} + E_{II} + \hat{V}_{int} | \psi \rangle \\ &= \langle \psi | \hat{T} + \hat{V}_{ext} + \hat{V}_{int} | \psi \rangle + E_{II} \end{aligned} \quad (2.13)$$

Equation (2.13) shows that the total energy E of a given system of electrons and nuclei could be known, if only ψ could be found, because the terms \hat{T} , \hat{V}_{ext} , E_{II} and \hat{V}_{int} are already known. At this level of theory, the electronic states are the only states which require a quantum mechanical description in the form of a wavefunction, and the choices of subscript in the chosen notation in (2.10) and (2.11) reflect this redirected focus on the electrons only: \hat{V}_{ext} is the Coulomb potential *external to the electrons* created by the nuclei. \hat{V}_{int} is the potential created amongst the electrons themselves, or *internally*. Note that the comparatively trivial Coulomb nucleus-nucleus term E_{II} , being the only purely classical term, is the only term which does not take the form of an operator and therefore can be known by merely specifying the nuclear coordinates \vec{R}_I without needing to solve for ψ . The terms \hat{T} and \hat{V}_{int} depend only on the electronic coordinates \vec{r}_i and are common to all possible systems: they are *universal*. That is, they always take the same form in all problems. In contrast, the nuclear-nuclear interaction energy E_{II} and the external potential operator \hat{V}_{ext} both depend on (in addition to \vec{r}_i) the nuclear coordinates \vec{R}_I which are adjustable parameters, so \hat{V}_{ext} and E_{II} are unique to each system in that they both depend on the chosen arrangement of the atomic nuclei. Once the nuclear coordinates \vec{R}_I are chosen and fixed, the terms \hat{V}_{ext} and E_{II} then make the Hamiltonian \hat{H} unique to that particular

system. In this sense, \hat{V}_{ext} and E_{II} completely *determine* the Hamiltonian \hat{H} and also the corresponding ground state energy E . It is important to note that the converse is not true: if just the ground state energy E of a system is known, then the Hamiltonian \hat{H} cannot necessarily be determined. This is obvious; there are many different arrangements of nuclei and electrons which can have the same total energy, but any one given arrangement of nuclei and electrons can only have one total energy.

So, what is known about the wavefunction $\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N)$ describing the N interacting electrons? Unfortunately, despite the simplifications introduced so far, ψ is almost always a prohibitively difficult solution to find analytically for realistic systems. One notable analytic solution is the H atom, for which ψ can be found analytically, although the H atom is of course not a many-body problem. Shortly after the development of the foundations of quantum mechanics in the late 1920's, the first serious step towards applying the theory to many-electron systems was made by Hartree⁸ in 1928. Hartree showed that Schrödinger's equation can be *approximately* solved for a system of many electrons by treating the interactions between the electrons as an external potential experienced independently by each of the electrons. In this sense, the electrons are regarded in Hartree's approach as effectively *non-interacting*. The only requirement is that the external potential experienced by each and every electron must be *self-consistent* for all electrons in the system, which, in the absence of an analytic solution, demands a numerical approach. (This technique is revisited in detail in section 2.4.) In 1930, Slater pointed out⁹ that Hartree's problem can be solved by variational minimisation of a trial wavefunction. The variational method, being a longstanding cornerstone of classical mechanics, placed Hartree's method onto more firm theoretical footing. Also, in a different paper,¹⁰ Slater showed that the wavefunction of a general system of non-interacting electrons can be written as a totally antisymmetric determinant of single-particle wavefunctions, a *Slater determinant*, where each of those constituent wavefunctions corresponds directly to each non-interacting electron. Fock¹¹ then published the first explicit example of calculations applied to many electron atoms using Hartree's approach with fully antisymmetric determinant wavefunctions as suggested by Slater. The

method used by Fock¹¹ is now generally referred to as the *Hartree-Fock method* (HF). By writing the electronic wavefunction ψ as a Slater determinant (but neglecting spin for this simple example) of non-interacting electron wavefunctions, the total wavefunction in HF theory automatically satisfies antisymmetric *exchange symmetry* required of fermions by Fermi-Dirac statistics. For example, the wavefunction representing a two-electron system can be written as

$$\psi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} (\chi_1(\vec{r}_1)\chi_2(\vec{r}_2) - \chi_1(\vec{r}_2)\chi_2(\vec{r}_1)) = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_1(\vec{r}_1) & \chi_2(\vec{r}_1) \\ \chi_1(\vec{r}_2) & \chi_2(\vec{r}_2) \end{vmatrix} \quad (2.14)$$

where the wavefunctions χ_1 and χ_2 are two single-particle states and the $\frac{1}{\sqrt{2}}$ normalises ψ . If the positions of these two states are now *exchanged*, so that $\vec{r}_2 \leftrightarrow \vec{r}_1$ (or equivalently, if the states are relabelled using $\chi_1 \leftrightarrow \chi_2$) we obtain

$$\psi(\vec{r}_2, \vec{r}_1) = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_1(\vec{r}_2) & \chi_2(\vec{r}_2) \\ \chi_1(\vec{r}_1) & \chi_2(\vec{r}_1) \end{vmatrix} = -\psi(\vec{r}_1, \vec{r}_2) \quad (2.15)$$

It is clear that the two-electron wavefunction ψ , when written in this way, satisfies the demand that it be antisymmetric under exchange of its two constituent particle wavefunctions χ_1 and χ_2 . It is also easy to see that ψ satisfies the *Pauli exclusion principle*: suppose now that the two constituent wavefunctions are exactly the same, so that $\chi_1 = \chi_2$, then we obtain $\psi = 0$. This of course means that the probability of observing the two electrons in a superposed state is zero. These ideas are easily implemented in N -electron systems by expressing ψ in the form suggested by Slater:

$$\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N; \sigma_1, \dots, \sigma_i, \dots, \sigma_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\vec{r}_1, \sigma_1) & \cdots & \chi_N(\vec{r}_1, \sigma_1) \\ \vdots & \ddots & \vdots \\ \chi_1(\vec{r}_N, \sigma_N) & \cdots & \chi_N(\vec{r}_N, \sigma_N) \end{vmatrix} \quad (2.16)$$

where $\{\sigma_i\}$ labels the spin states of each electron. Despite the merits of the HF approach, it fails to fully account for *electron correlation* which originates from the individual Coulomb interactions amongst the individual electrons. This means that it can fail to predict material properties which depend sensitively on electron correlation, such as van der Waals interactions. In order to improve the accuracy of HF theory, one must deviate from the simpler Slater determinant form of the wavefunction in equation (2.16).

Approaches of this type are generally called *post-Hartree-Fock* theories. They include Møller–Plesset perturbation theory¹² proposed in 1934, coupled-cluster approaches which use techniques from quantum field theory as originally suggested by Čížek in 1966,¹³ and configuration interaction techniques developed by Langhoff and Davidson¹⁴ in 1974 and further developed by Pople *et al.* in 1987.¹⁵ These approaches are known to be computationally very expensive, but they can be very accurate and are sometimes used in areas of computational chemistry where the number of electrons involved is quite small. However, in order to broaden the appeal of the quantum mechanical theory of electrons to larger problems in materials science, a slightly different approach is called for.

Probably the most successful and influential approximation to the quantum mechanical theory of many electrons of the 20th century is DFT, first properly formalised in 1964/1965 by Hohenberg, Kohn and Sham.^{1,2} Between them, they showed that accounting for electron exchange *and* correlation can be achieved by recasting the problem of finding $\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_N)$, which depends on $3N$ coordinates, as the alternative problem of finding the *total electron density* $n(\vec{r})$, which is a simple scalar field and only depends on 3 coordinates. The electron density of a system of N electrons is equal to the expectation value of the electron density operator $\hat{n}(\vec{r})$, where the operator is defined as

$$\hat{n}(\vec{r}) = \sum_{j=1}^N \delta(\vec{r} - \vec{r}_j). \quad (2.17)$$

The electron density $n(\vec{r})$ at a point \vec{r} is the expectation value of the electron density operator:

$$\begin{aligned} n(\vec{r}) &\stackrel{\text{def}}{=} \langle \psi | \hat{n}(\vec{r}) | \psi \rangle \\ &= \int \psi^*(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \left(\sum_{j=1}^N \delta(\vec{r} - \vec{r}_j) \right) \psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) d^3r_1 d^3r_2 \dots d^3r_N \\ &= \sum_{j=1}^N \int \psi^*(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \left(\delta(\vec{r} - \vec{r}_j) \right) \psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) d^3r_1 d^3r_2 \dots d^3r_N \end{aligned}$$

$$\begin{aligned}
 &= \int \psi^*(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) \psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) d^3r_2 d^3r_3 \dots d^3r_{N-1} d^3r_N \\
 &+ \int \psi^*(\vec{r}_1, \vec{r} \dots, \vec{r}_N) \psi(\vec{r}_1, \vec{r} \dots, \vec{r}_N) d^3r_1 d^3r_3 \dots d^3r_{N-1} d^3r_N \\
 &+ \dots \\
 &+ \int \psi^*(\vec{r}_1, \vec{r}_2 \dots, \vec{r}) \psi(\vec{r}_1, \vec{r}_2 \dots, \vec{r}) d^3r_1 d^3r_2 \dots d^3r_{N-1}
 \end{aligned} \tag{2.18}$$

In order to arrive at the last equality in equation (2.18), the property of delta functions that for some function $\Phi(\vec{r})$, $\int \Phi(\vec{r}) \delta(\vec{r} - \vec{r}') d^3r = \Phi(\vec{r}')$, was used to remove one of the integrals from each of the N terms. Each one of these N remaining terms consists of $N - 1$ integrals over all space, and each is clearly a function of \vec{r} . The first of these terms is the probability density of finding the first electron (and *only* the first electron) at point \vec{r} . The second term is the probability density of finding the second electron at point \vec{r} . Since electrons are indistinguishable by wavefunction antisymmetry, these two probabilities *must be exactly equal*. Therefore, we can simply write the sum of all of these terms as N times the first term; that is, the probability of finding *any one* of the N electrons at point \vec{r} is equal to the N times the probability density of finding the first electron - the chosen reference electron - at point \vec{r} . The electron density $n(\vec{r})$ can therefore be written as

$$n(\vec{r}) = N \int \psi^*(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) \psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) d^3r_2 d^3r_3 \dots d^3r_N. \tag{2.19}$$

The first of many specific advantages of trading the difficult electronic wavefunction ψ for the much more simple object $n(\vec{r})$ is that, like the nucleus-nucleus interaction term from the Hamiltonian, the external potential quantum mechanical operator \hat{V}_{ext} can be converted into a classical term. This can be easily shown using Coulomb's law applied to the discrete array of positive point charges representing the nuclei and the continuous and negatively-charged electron density field $n(\vec{r})$. At a general point in space, \vec{r} , the nuclei create a classical Coulomb potential given by

$$V_{ext}(\vec{r}) = - \sum_I \frac{Z_I}{|\vec{r} - \vec{R}_I|}, \tag{2.20}$$

so that the total energy associated with the Coulomb interaction of the continuous charge distribution $n(\vec{r})$ with the nuclei is found by integrating all of the infinitesimal contributions to the total energy over all infinitesimal volume elements d^3r :

$$\int V_{ext}(\vec{r})n(\vec{r})d^3r = - \int \sum_I \frac{Z_I n(\vec{r})}{|\vec{r} - \vec{R}_I|} d^3r \quad (2.21)$$

The total system energy from equation (2.13) can therefore be simplified further to

$$E = \langle \psi | \hat{T} + \hat{V}_{int} | \psi \rangle + \int V_{ext}(\vec{r})n(\vec{r})d^3r + E_{II} \quad (2.22)$$

This is indeed much more simple because the only two terms which still require operator forms are the two universal terms, \hat{T} and \hat{V}_{int} . The other two terms - those which uniquely determine the Hamiltonian - are now just classical terms. The primary simplifications to the Hamiltonian required for DFT to be formally introduced are now complete. Firstly though, a basic explanation of functionals is worthwhile, as it will aid understanding in the following sections.

2.2 What actually is a functional anyway?

Let us briefly return to absolute basics and define a function to start with, using the electron density function $n(\vec{r})$ as the example. The physical unit of $n(\vec{r})$ is $e\text{\AA}^{-3}$ in three spatial dimensions, or electronic charge units per unit volume. In p spatial dimensions (where $p = 3$ in virtually all realistic materials science applications), the function $n(x_1, x_2, \dots, x_p)$ is a rule for converting a discrete set of input quantities x_1, x_2, \dots, x_p (positions in Euclidian space) to a single output quantity. If the input variables all have their values changed by infinitesimal amounts dx_1, dx_2, \dots, dx_p , then the function output value changes by an amount given by

$$dn = \frac{\partial n}{\partial x_1} dx_1 + \frac{\partial n}{\partial x_2} dx_2 + \dots + \frac{\partial n}{\partial x_p} dx_p \quad (2.23)$$

The particular set of values which the quantities x_1, x_2, \dots, x_p can take which produces the *smallest possible* output value of n is called the global

minimum of the function $n(x_1, x_2, \dots, x_p)$. At the global minimum of any function $n(x_1, x_2, \dots, x_p)$, it is always true that

$$dn = 0 \quad (2.24)$$

In general, any point on the function at which $dn = 0$ is called a *stationary point*, which includes the global minimum, local minima, local maxima and points of inflection. At a minimum, changes dx_1, dx_2, \dots, dx_p will always result in an increase in the value of the function n . Figure 2.1 shows an example function $n(x_1, x_2)$ in 2 dimensions. By specifying the values of the two inputs x_1 and x_2 (specified by a vector \vec{r}) and thereby picking a point on the image, one (and only one) value of n is found, evaluated using the colour scale.

A *functional* of the electron density, $F(n(x_1, x_2, \dots, x_p))$, is essentially the same as a function, except in that it is a rule for converting infinitely many input quantities to a single scalar output quantity, where these input quantities are the *infinitely many different values of the continuous function* $n(x_1, x_2, \dots, x_p)$. The notation $F(n(x_1, x_2, \dots, x_p))$ does *not* mean " F evaluated using the single input quantity n , where n is the value of the function $n(x_1, x_2, \dots, x_p)$ at the point (x_1, x_2, \dots, x_p) ". Rather, it means " F evaluated using an infinite number of input quantities, where this range of input quantities is all of the possible values of the function $n(x_1, x_2, \dots, x_p)$ over all points (x_1, x_2, \dots, x_p) ". To understand this, see Figure 2.2 which shows the electron density field $n(x_1, x_2)$ from Figure 2.1 partitioned into small squares of width dx_1 and height dx_2 . (Allow the function n to be exactly zero at all points beyond the borders of the image, so that these points do not contribute anything further to the total electron density.) Consider the mean value of the electron density in each of these squares. Let the mean electron density in the square which is located in the i^{th} square along the x_1 axis and the j^{th} square along the x_2 axis be denoted $\bar{n}_{i,j}$. We could then define a function, let us call it f , whose output value has physical units of, say, energy, where this output energy depends on the set of discrete input quantities $\bar{n}_{i,j}$, and therefore write $f = f(\bar{n}_{-\infty, -\infty}, \dots, \bar{n}_{i,j}, \dots, \bar{n}_{\infty, \infty})$. By allowing the squares to become vanishingly small and infinitely dense, so that $dx_1 \rightarrow 0$ and $dx_2 \rightarrow 0$, the squares degenerate to an infinite set of distinct

points, and so the input quantities on which the function f depends become the infinitely many different values of the continuous function $n(x_1, x_2)$.

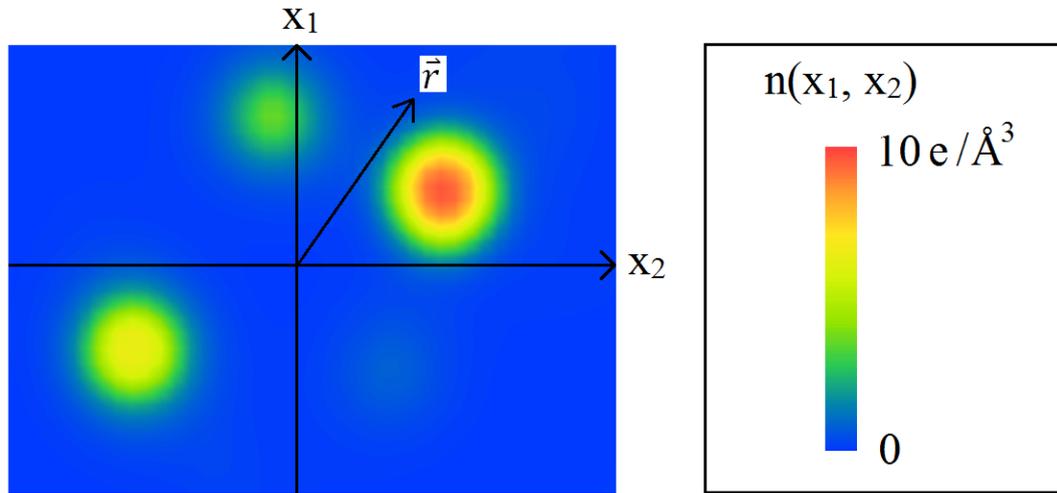


Figure 2.1. The electron density field is a single-valued function of \vec{r} .

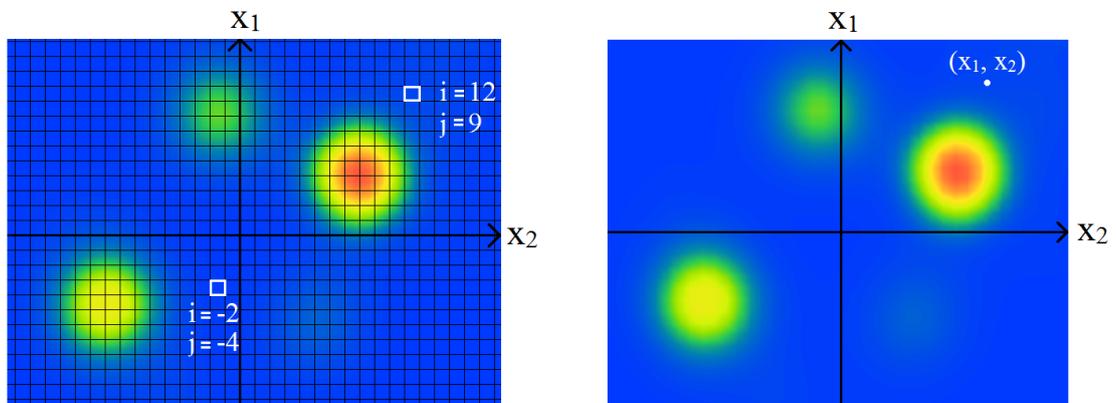


Figure 2.2. The input variables of a functional F .

They can be regarded as all possible values of the continuous function $n(x_1, x_2)$. In the left panel, the mean electron density in each square can be regarded as the discrete set of input parameters of some function f . If the size of each of the squares is allowed to vanish to make an infinitely dense grid, then each square degenerates to a point, and in this limit, the discrete input parameters to the original function f become the infinitely many values of the continuous function $n(x_1, x_2)$, and therefore become the infinitely many input variables of the functional F .

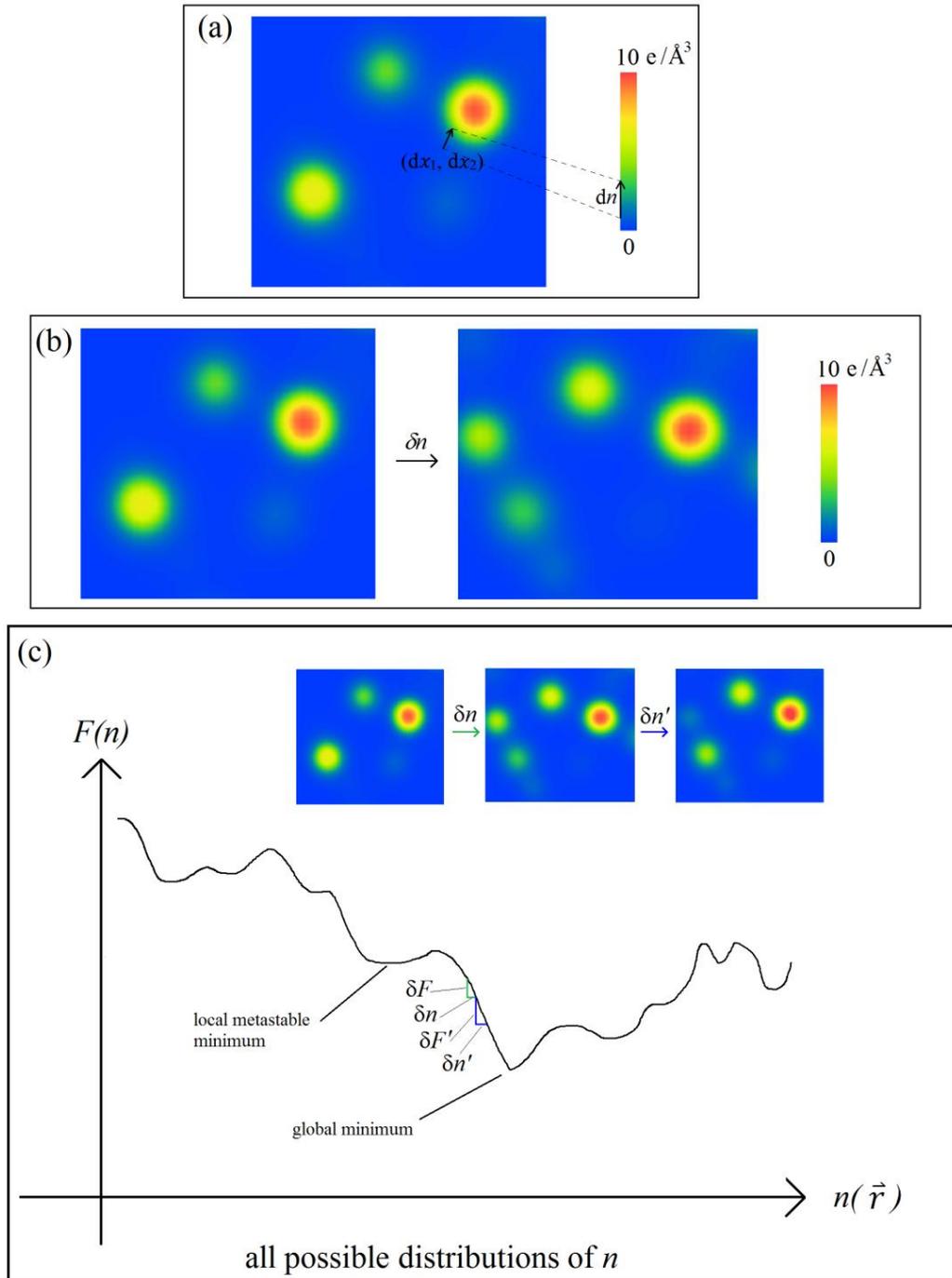


Figure 2.3. The difference between dn and δn .

(a) dn is the change in the value of the fixed function n at two close and different points in the field. (b) δn is a change in the overall shape of the field at all points in space. (c) δF is the change in the total energy in response to a change in the overall density distribution δn . F is generally a very complicated function of n in realistic problems. The curve in (c) shows that F is single-valued in n .

At this limit, the function f formally becomes a *functional* F and is denoted $F(n(x_1, x_2))$. When the function $n(x_1, x_2)$ is given as the input to the functional F , the functional F gives an energy value as its output. (This last statement is written with the benefit of hindsight. It has not actually been proved from the information so far presented in this thesis that such a functional F can be defined. It will be proved with the second Hohenberg-Kohn theorem in section 2.3.)

The reader is now reminded of the meaning of equation (2.23): Equation (2.23) shows (in the 2D case) that if infinitesimal changes dx_1 and dx_2 are made to the inputs x_1 and x_2 of the function $n(x_1, x_2)$, then a small change dn in the value of the function n is observed. If such a change is made, then of course nothing happens to the functional $F(n(x_1, x_2))$. This is because the changes dx_1 and dx_2 simply amount to moving from one point to another in the electron density distribution without actually changing the function $n(x_1, x_2)$. To observe an infinitesimal change in the *functional*, δF , the function $n(x_1, x_2)$ must itself be changed over all space. This change in the electron density is denoted $\delta n(x_1, x_2)$ and it is clearly a function because one must account for all of the points in space to fully account for changes in the overall appearance of the distribution. Assuming now that the overall appearance of the electron density has now changed by an amount δn , there is a corresponding change in the total energy functional, equal to δF . The rate at which the total energy F changes with respect to small changes in the entire distribution $\delta n(x_1, x_2)$ is called the *functional derivative of F* and is written as $\frac{\delta F}{\delta n(x_1, x_2)}$. We are now ready to define this infinitesimal change in the total energy, δF , with respect to changes in the electron density $\delta n(x_1, x_2)$:

$$\delta F = \iint \frac{\delta F}{\delta n(x_1, x_2)} \delta n(x_1, x_2) dx_1 dx_2 \quad (2.25)$$

This is the generalisation (in 2D) of the sum seen in equation (2.23) for the case of a function. Since F depends on the value of n at all points in space, the changes made to n must be integrated over all points in space. Figures 2.3 (a) and (b) show the difference of a change in the density dn in response to the changes dx_1 and dx_2 , and a change in the total density δn in response

to a change in the function n over all space. In analogy with equation (2.24), it is always true at all stationary points that

$$\delta F = 0. \quad (2.26)$$

Earlier in this section, it was stated that a functional is a rule for converting an input function (electron density) to an output value (energy). However, nothing has yet been said about what this rule actually is. To re-establish some sense of physicality and to understand why this discussion on functionals is relevant to problems of electrons and nuclei in quantum mechanics, formal DFT is now introduced in the form of the Hohenberg-Kohn theorems¹ and Kohn-Sham equations.²

2.3 Hohenberg-Kohn theorems

Let the particular case of the electron density $n(\vec{r})$ in the quantum mechanical ground state be denoted $n_0(\vec{r})$. The first Hohenberg-Kohn (HK) theorem makes the following claim:

- *The ground state electron density $n_0(\vec{r})$ uniquely determines the external potential $V_{ext}(\vec{r})$ up to a constant*

Essentially what this means is that if a ground state electron density field $n_0(\vec{r})$ is known, *there can only possibly be one corresponding arrangement of nuclei*. The converse is also true: for a known arrangement of nuclei, there is *only one possible ground state electron density $n_0(\vec{r})$* . There is therefore a one-to-one mapping between the ground state electron density $n_0(\vec{r})$ and the external potential $V_{ext}(\vec{r})$. Here's the proof.

Consider two *different* arrangements of nuclei, i.e. two different external potentials $V_{ext}^{(1)}$ and $V_{ext}^{(2)}$, each of which (along with $E_{II}^{(1)}$ and $E_{II}^{(2)}$) completely determines its respective Hamiltonian:

$$\hat{H}^{(1)} = \hat{T} + \hat{V}_{int} + \hat{V}_{ext}^{(1)} + E_{II}^{(1)} \quad (2.27)$$

$$\hat{H}^{(2)} = \hat{T} + \hat{V}_{int} + \hat{V}_{ext}^{(2)} + E_{II}^{(2)} \quad (2.28)$$

Also consider $\psi_0^{(1)}$, the ground state eigenstate of $\hat{H}^{(1)}$, and $\psi_0^{(2)}$, the ground state eigenstate of $\hat{H}^{(2)}$. Since $\psi_0^{(1)}$ and $\psi_0^{(2)}$ are eigenstates of $\hat{H}^{(1)}$ and $\hat{H}^{(2)}$

respectively, they both satisfy the Schrödinger equation, and so the corresponding ground state energies $E_0^{(1)}$ and $E_0^{(2)}$ are equal to the expectation values of their respective Hamiltonians:

$$\begin{aligned} E_0^{(1)} &= \langle \psi_0^{(1)} | \hat{H}^{(1)} | \psi_0^{(1)} \rangle \\ &= \langle \psi_0^{(1)} | \hat{T} + \hat{V}_{int} | \psi_0^{(1)} \rangle + \int V_{ext}^{(1)}(\vec{r}) n_0^{(1)}(\vec{r}) d^3r + E_{II}^{(1)} \end{aligned} \quad (2.29)$$

$$\begin{aligned} E_0^{(2)} &= \langle \psi_0^{(2)} | \hat{H}^{(2)} | \psi_0^{(2)} \rangle \\ &= \langle \psi_0^{(2)} | \hat{T} + \hat{V}_{int} | \psi_0^{(2)} \rangle + \int V_{ext}^{(2)}(\vec{r}) n_0^{(2)}(\vec{r}) d^3r + E_{II}^{(2)} \end{aligned} \quad (2.30)$$

Let us now make an assumption: let us assume that the two different potentials $V_{ext}^{(1)}(\vec{r})$ and $V_{ext}^{(2)}(\vec{r})$ give rise to the *same ground state electron density* $n_0(\vec{r})$, such that

$$n_0^{(1)}(\vec{r}) = n_0^{(2)}(\vec{r}) = n_0(\vec{r}), \quad (2.31)$$

so that equations (2.29) and (2.30) can be written as

$$E_0^{(1)} = \langle \psi_0^{(1)} | \hat{T} + \hat{V}_{int} | \psi_0^{(1)} \rangle + \int V_{ext}^{(1)}(\vec{r}) n_0(\vec{r}) d^3r + E_{II}^{(1)}, \quad (2.32)$$

and

$$E_0^{(2)} = \langle \psi_0^{(2)} | \hat{T} + \hat{V}_{int} | \psi_0^{(2)} \rangle + \int V_{ext}^{(2)}(\vec{r}) n_0(\vec{r}) d^3r + E_{II}^{(2)}. \quad (2.33)$$

$\psi_0^{(1)}$ is the ground state wavefunction of $\hat{H}^{(1)}$, so it must be true that

$$E_0^{(1)} = \langle \psi_0^{(1)} | \hat{H}^{(1)} | \psi_0^{(1)} \rangle < \langle \psi_0^{(2)} | \hat{H}^{(1)} | \psi_0^{(2)} \rangle, \quad (2.34)$$

because applying $\hat{H}^{(1)}$ to any wavefunction which is not its ground state wavefunction $\psi_0^{(1)}$ must, by definition, produce some energy eigenvalue which is greater than its ground state energy $E_0^{(1)}$. By combining equations (2.27) and (2.28), the Hamiltonian $\hat{H}^{(1)}$ can be written in terms of $\hat{H}^{(2)}$ as follows:

$$\begin{aligned} \hat{H}^{(1)} &= \hat{T} + \hat{V}_{int} + \hat{V}_{ext}^{(1)} + E_{II}^{(1)} \\ &= \hat{T} + \hat{V}_{int} + \hat{V}_{ext}^{(2)} + E_{II}^{(2)} + \left(\hat{V}_{ext}^{(1)} + E_{II}^{(1)} \right) - \left(\hat{V}_{ext}^{(2)} + E_{II}^{(2)} \right) \end{aligned}$$

$$= \hat{H}^{(2)} + \left(\hat{V}_{ext}^{(1)} + E_{II}^{(1)} \right) - \left(\hat{V}_{ext}^{(2)} + E_{II}^{(2)} \right), \quad (2.35)$$

so that inequality (2.34) can be written as

$$\begin{aligned} E_0^{(1)} &< \left\langle \psi_0^{(2)} \left| \hat{H}^{(1)} \right| \psi_0^{(2)} \right\rangle \\ &= \left\langle \psi_0^{(2)} \left| \hat{H}^{(2)} + \hat{V}_{ext}^{(1)} - \hat{V}_{ext}^{(2)} \right| \psi_0^{(2)} \right\rangle + E_{II}^{(1)} - E_{II}^{(2)} \\ &= E_0^{(2)} + \int \left(V_{ext}^{(1)}(\vec{r}) - V_{ext}^{(2)}(\vec{r}) \right) n_0(\vec{r}) d^3r + E_{II}^{(1)} - E_{II}^{(2)} \end{aligned} \quad (2.36)$$

Now, by identical reasoning with which inequality (2.34) was written, the fact that $\psi_0^{(2)}$ is the ground state wavefunction of $\hat{H}^{(2)}$, but not the ground state of $\hat{H}^{(1)}$, can be used to write

$$E_0^{(2)} = \left\langle \psi_0^{(2)} \left| \hat{H}^{(2)} \right| \psi_0^{(2)} \right\rangle < \left\langle \psi_0^{(1)} \left| \hat{H}^{(2)} \right| \psi_0^{(1)} \right\rangle, \quad (2.37)$$

which, again, by combining equations (2.27) and (2.28), results in

$$\begin{aligned} E_0^{(2)} &< \left\langle \psi_0^{(1)} \left| \hat{H}^{(2)} \right| \psi_0^{(1)} \right\rangle \\ &= \left\langle \psi_0^{(1)} \left| \hat{H}^{(1)} + \hat{V}_{ext}^{(2)} - \hat{V}_{ext}^{(1)} \right| \psi_0^{(1)} \right\rangle + E_{II}^{(2)} - E_{II}^{(1)} \\ &= E_0^{(1)} + \int \left(V_{ext}^{(2)}(\vec{r}) - V_{ext}^{(1)}(\vec{r}) \right) n_0(\vec{r}) d^3r + E_{II}^{(2)} - E_{II}^{(1)} \end{aligned} \quad (2.38)$$

Inequalities (2.36) and (2.38) can then be added to give

$$E_0^{(1)} + E_0^{(2)} < E_0^{(1)} + E_0^{(2)}, \quad (2.39)$$

which is clearly absurd. Therefore, the premise - equation (2.31) - is false! It follows that it is impossible for two distinct external potentials $V_{ext}^{(1)}$ and $V_{ext}^{(2)}$ to correspond to the same total ground state electron density $n_0(\vec{r})$. Therefore, if $n_0(\vec{r})$ is known, V_{ext} is known.

The second Hohenberg-Kohn theorem states the following:

- *For any external potential V_{ext} , the total energy of a system can be always expressed as a universal functional of the electron density, $E(n(\vec{r}))$. The ground state energy is the global minimum of the functional $E(n(\vec{r}))$, and is located at $E(n_0(\vec{r}))$.*

The functional can be written as

$$E(n(\vec{r})) = \int V_{ext}(\vec{r})n(\vec{r})d^3r + F(n(\vec{r})) + E_{II}. \quad (2.40)$$

Equation (2.40) should be compared with equation (2.22). The second Hohenberg-Kohn theorem effectively states that the two remaining quantum mechanical operator terms $\hat{T} + \hat{V}_{int}$ from equation (2.22) can *also* be expressed as functionals of the electron density $n(\vec{r})$:

$$\langle \psi | \hat{T} | \psi \rangle = T(n(\vec{r})), \quad (2.41)$$

$$\langle \psi | \hat{V}_{int} | \psi \rangle = E_{int}(n(\vec{r})), \quad (2.42)$$

where $T(n(\vec{r}))$ is the functional for the total kinetic energy of a system of N *interacting* electrons, $E_{int}(n(\vec{r}))$ gives the total energy of these electron-electron interactions, and $F(n(\vec{r})) = T(n(\vec{r})) + E_{int}(n(\vec{r}))$. Equation (2.40) therefore reads

$$E(n(\vec{r})) = T(n(\vec{r})) + E_{int}(n(\vec{r})) + \int V_{ext}(\vec{r})n(\vec{r})d^3r + E_{II}. \quad (2.43)$$

The second Hohenberg-Kohn theorem can be proved quite trivially as follows. Consider two systems: systems (a) and (b) with different ground state electron densities $n_0^{(a)}(\vec{r})$ and $n_0^{(b)}(\vec{r})$. By virtue of the first Hohenberg-Kohn theorem, it immediately follows that $n_0^{(a)}(\vec{r})$ must correspond to only one unique ground state wavefunction $\psi_0^{(a)}$, external potential $V_{ext}^{(a)}$, Hamiltonian $\hat{H}^{(a)}$ and energy $E_0^{(a)}$, and similarly that $n_0^{(b)}(\vec{r})$ must correspond to only one unique $\psi_0^{(b)}$, $V_{ext}^{(b)}$, $\hat{H}^{(b)}$ and $E_0^{(b)}$. Since $\psi_0^{(a)}$ and $\psi_0^{(b)}$ are necessarily different, it must be true that

$$\begin{aligned} E_0^{(a)} &= \langle \psi_0^{(a)} | \hat{H}^{(a)} | \psi_0^{(a)} \rangle \\ &= T(n_0^{(a)}(\vec{r})) + V_{int}(n_0^{(a)}(\vec{r})) + \int V_{ext}^{(a)}(\vec{r})n_0^{(a)}(\vec{r})d^3r + E_{II}^{(a)} \\ &< \langle \psi_0^{(b)} | \hat{H}^{(a)} | \psi_0^{(b)} \rangle \\ &= T(n_0^{(b)}(\vec{r})) + V_{int}(n_0^{(b)}(\vec{r})) + \int V_{ext}^{(a)}(\vec{r})n_0^{(b)}(\vec{r})d^3r + E_{II}^{(a)} \end{aligned} \quad (2.44)$$

by exactly the same reasoning used for inequalities (2.34) and (2.37). Note that the nuclear terms $V_{ext}^{(a)}$ and $E_{II}^{(a)}$ on both sides of inequality (2.44) are the same. Put simply, what inequality (2.44) shows is that if the electron density from system (a) were replaced with the electron density from system (b) without changing the arrangement of the nuclei, then the total energy of the resultant system would be higher. Thus, for system (a), *any* electron density which differs from its ground state density $n_0^{(a)}(\vec{r})$ results in a system energy which is greater than $E_0^{(a)}$. Therefore, it follows that $E_0^{(a)}$ is indeed the ground state energy and that finding the ground state density $n_0^{(a)}(\vec{r})$ is a sufficient condition to know the global minimum of the total energy functional E .

In conclusion, the consequence of the two Hohenberg-Kohn theorems is that every single term which contributes to the total energy takes the form of a functional of the electron density $n(\vec{r})$. The $3N$ -dimensional problem of finding the ground state electron wavefunction ψ has therefore been further reduced to the 3-dimensional problem of finding the ground state total electron density n_0 . This shows the extremely important status of the electron density in DFT.

2.4 The Kohn-Sham equations

The obvious next task is to determine the form of the functionals $T(n(\vec{r}))$ and $E_{int}(n(\vec{r}))$. Unfortunately, the exact forms are not known. The incredibly lucrative approach used by Kohn and Sham is to recast the entire problem as a non-interacting theory, building on the original idea of Hartree's⁸ idea in 1928 of converting interacting theories to non-interacting theories via an effective external potential. Kohn and Sham's thinking is best illustrated by carefully rewriting the expression for the total energy functional (2.43) and inspecting the newly written terms. This is explained now in two steps, the first of which relates to $T(n(\vec{r}))$ and the second of which relates to $E_{int}(n(\vec{r}))$.

Firstly, let the total electron density $n(\vec{r})$ be expressed explicitly in quantum mechanical form once again by expressing it as the sum of the densities of some system of N *non-interacting* wavefunctions:

$$n(\vec{r}) = \sum_{i=1}^N \langle \varphi_i(\vec{r}) | \varphi_i(\vec{r}) \rangle, \quad (2.45)$$

where the wavefunctions $\varphi_i(\vec{r})$ are *orthonormal*:

$$\langle \varphi_i(\vec{r}) | \varphi_j(\vec{r}) \rangle = \delta_{i,j} \quad (2.46)$$

where $\delta_{i,j}$ is the Kronecker delta. The wavefunctions $\varphi_i(\vec{r})$ are the *Kohn-Sham wavefunctions*. The choice of (2.45) and (2.46) is entirely deliberate and has a clear physical motivation: note that the total density here could be represented as the trace of an $N \times N$ matrix, where the term in the i^{th} row and i^{th} column is $\langle \varphi_i(\vec{r}) | \varphi_i(\vec{r}) \rangle$, but where *all off-diagonal terms are zero*; that is, there are no terms of the form $\langle \varphi_i(\vec{r}) | \varphi_j(\vec{r}) \rangle$ by orthonormality. The form of (2.45) therefore shows explicitly that the wavefunctions $\varphi_i(\vec{r})$ are non-interacting.

By virtue of the second HK theorem, the total kinetic energy of these N non-interacting Kohn-Sham states - found using the usual kinetic energy operator from equation (2.9) - can be written as a functional of the electron density:

$$-\frac{1}{2} \sum_i^N \langle \varphi_i(\vec{r}) | \nabla^2 | \varphi_i(\vec{r}) \rangle = T_S(n(\vec{r})) \quad (2.47)$$

This should be compared with $T(n(\vec{r}))$ from equation (2.41); the case of the total kinetic energy of the N -electron wavefunction ψ .

Secondly, let us define the *Hartree term*.⁸ The Hartree term is the classical expression for the total Coulomb energy arising from the interaction of each and every infinitesimal volume element of a continuous charge distribution with each and every other infinitesimal volume element of that same distribution:

$$E_{Hartree}(n(\vec{r})) = \frac{1}{2} \int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r d^3r' \quad (2.48)$$

Expressed another way, the Hartree term shows what the total electron-electron interaction energy *would* be if the electrons amounted to nothing more than a continuous charge distribution fully characterised by Coulomb's law. If this were true, then one could write $E_{int}(n(\vec{r})) = E_{Hartree}(n(\vec{r}))$. In

reality, the quantum mechanical effects of *electron exchange and correlation* are not accounted for by the Hartree term. The Hartree term is therefore a useful reference object, because it can be considered as the classical component of the total internal electronic interaction energy E_{int} , which can be subtracted to leave the purely quantum mechanical component.

From this perspective, it should now be clear why the *exchange-correlation functional*, $E_{XC}(n(\vec{r}))$, is defined as follows:

$$E_{XC}(n(\vec{r})) = (T(n(\vec{r})) - T_S(n(\vec{r}))) + (E_{int}(n(\vec{r})) - E_{Hartree}(n(\vec{r}))) \quad (2.49)$$

The exchange-correlation functional E_{XC} is the difference between the true kinetic energy of the electronic wavefunction ψ and that of all of the KS orbitals φ_i , and the difference between the interaction energy of a purely classical charge density field, and that of the true quantum-mechanical electron system, both added together. In addition to this, let us also adopt the term $E_{ext}(n(\vec{r}))$ to denote the energy associated with the Coulomb interaction between the nuclei and $n(\vec{r})$:

$$E_{ext}(n(\vec{r})) = \int V_{ext}(\vec{r})n(\vec{r})d^3r \quad (2.50)$$

The expressions from (2.49) and (2.50), can now be substituted into the expression for the total energy functional $E(n(\vec{r}))$ in equation (2.43) to get:

$$E(n(\vec{r})) = T_S(n(\vec{r})) + E_{Hartree}(n(\vec{r})) + E_{XC}(n(\vec{r})) + E_{ext}(n(\vec{r})) + E_{II}. \quad (2.51)$$

Let us now review the physical interpretation of each of the terms in equation (2.51). The first term is the total kinetic energy of the Kohn-Sham orbitals; *not* the kinetic energy of the N -electron wavefunction ψ . The second term is the classical Coulomb contribution to the electron-electron interaction energy. The third term is the sum of *i*) the quantum mechanical component of the electron-electron interaction energy and *ii*) the difference between the kinetic energies of the Kohn-Sham orbitals and that of the true N -electron system. The last two terms are unchanged from equation (2.43); they only depend on the nuclei. Besides the BOA, the accuracy of DFT as a predictive theory is influenced only by what form E_{XC} takes. In Kohn and Hohenberg's paper, for the case where the density $n(\vec{r})$ is slowly-varying with \vec{r} , the

exchange-correlation energy functional $E_{XC}(n(\vec{r}))$ was expressed as the following integral over all space:

$$E_{XC}(n(\vec{r})) \approx \int n(\vec{r})\epsilon_{XC}(n(\vec{r}))d^3r, \quad (2.52)$$

where $\epsilon_{XC}(n(\vec{r}))$ is the *exchange-correlation energy density*. Equation (2.52) is the general form of the *local density approximation* (LDA) devised by Hohenberg and Kohn. For cases where $n(\vec{r})$ has more significant gradients in space, E_{XC} can readily be expressed as a power series (of which equation (2.52) is the first term) in gradients of the density $n(\vec{r})$:

$$E_{XC}(n(\vec{r}), \nabla n(\vec{r})) = \int n(\vec{r})\epsilon_{XC}^{(1)}(n(\vec{r}))d^3r + \int |\nabla n(\vec{r})|^2 \epsilon_{XC}^{(2)}d^3r + \dots \quad (2.53)$$

Such forms of E_{XC} are generally referred to as *generalised gradient approximations* (GGA). Note the similarity of equation (2.52) - the energy of exchange and correlation - to that of equation (2.21) for the energy of the interaction between the density $n(\vec{r})$ and external nuclear Coulomb potential $V_{ext}(\vec{r})$. The form of equation (2.52) exposes a very important conceptual shift: even though the true physical nature of $\epsilon_{XC}(n(\vec{r}))$ is ultimately rooted in the quantum-mechanical aspects of the electron-electron interactions, it is regarded in KS theory as one component of an *effective external potential* experienced by the total electron density $n(\vec{r})$, and in turn, each and every one of the N KS orbitals φ_i . The energy E_{XC} can therefore be viewed on the same footing as $E_{Hartree}$ and E_{ext} . The sum of these three terms is therefore the *total potential energy* associated with the density $n(\vec{r})$, and therefore their functional derivative with respect to the total $n(\vec{r})$ is the *total potential*:

$$\begin{aligned} & \frac{\delta}{\delta n(\vec{r})} \left(E_{Hartree}(n(\vec{r})) + E_{XC}(n(\vec{r})) + E_{ext}(n(\vec{r})) \right) \\ &= \frac{\delta}{\delta n(\vec{r})} \left(\frac{1}{2} \int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r d^3r' + \int n(\vec{r})\epsilon_{XC}(n(\vec{r}))d^3r + \int V_{ext}(\vec{r})n(\vec{r})d^3r \right) \\ &= \frac{\delta}{\delta n(\vec{r})} \int \left[\frac{1}{2} \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' + \epsilon_{XC}(n(\vec{r})) + V_{ext}(\vec{r}) \right] n(\vec{r}) d^3r \\ &= \frac{\delta}{\delta n(\vec{r})} \int [V_{KS}(n(\vec{r}))]n(\vec{r}) d^3r, \end{aligned} \quad (2.54)$$

where the total integrand in square brackets is defined as the total *Kohn Sham potential*:

$$V_{KS}(n(\vec{r})) \stackrel{\text{def}}{=} \int \frac{1}{2} \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' + \varepsilon_{XC}(n(\vec{r})) + V_{ext}(\vec{r}). \quad (2.55)$$

The total energy functional can now be rewritten in the form:

$$E(n(\vec{r})) = T_S(n(\vec{r})) + \int V_{KS}(n(\vec{r}))n(\vec{r}) d^3r + E_{II} \quad (2.56)$$

Now reconsider the explicit KS wavefunction expression for $T_S(n(\vec{r}))$ in equation (2.47). By design, it is a sum of familiar kinetic energy operator expectation values. Its kinetic energy operator $-\frac{1}{2}\nabla^2$ only acts, like all normal operators in linear algebra, on the terms to the right, i.e. $|\varphi_i(\vec{r})\rangle$ but not $\langle\varphi_i(\vec{r})|$. It is therefore fairly straightforward to see that if $T_S(n(\vec{r}))$ were to be differentiated with respect to $\langle\varphi_i(\vec{r})|$, the resultant term would be of the form $-\frac{1}{2}\nabla^2|\varphi_i(\vec{r})\rangle$. This is of course a very familiar term: it is a kinetic energy operator applied to an eigenstate. This alludes to the task ahead, which is to evaluate the functional derivative of $E(n(\vec{r}))$ with respect to the KS eigenstates $\langle\varphi_i(\vec{r})|$ and attempt to determine the form of the resulting Hamiltonian.

Consider once again equation (2.45): the total density is equal to the sum of the modulus squares of the individual KS wavefunctions φ_i , i.e. $n(\vec{r}) = \int [\varphi_1\varphi_1^* + \dots + \varphi_i\varphi_i^* + \dots + \varphi_N\varphi_N^*] d^3r$. Thus, we are free to consider the functional derivative of the total energy functional $E(n(\vec{r}))$ with respect to any one of the φ_i or φ_i^* states. The functional E can therefore be written in the much more revealing form as a functional whose arguments are the N KS eigenstates and their N complex conjugates, so that we can write $E(n(\vec{r})) = E(\varphi_1, \varphi_1^*, \dots, \varphi_i, \varphi_i^*, \dots, \varphi_N, \varphi_N^*)$. By that reasoning, the same must apply for T_S and V_{KS} , so that $T_S = T_S(\varphi_1, \varphi_1^*, \dots, \varphi_i, \varphi_i^*, \dots, \varphi_N, \varphi_N^*)$ and $V_{KS} = V_{KS}(\varphi_1, \varphi_1^*, \dots, \varphi_i, \varphi_i^*, \dots, \varphi_N, \varphi_N^*)$.

Let us set the functional derivative of E with respect to the complex conjugate of one of the KS states, say, the k^{th} state, equal to zero with a view to finding the minimum:

$$\frac{\delta E}{\delta \varphi_k^*} = \frac{\delta}{\delta \varphi_k^*} \left(T_S + \int V_{KS} n(\vec{r}) d^3r + E_{II} \right) = 0 \quad (2.57)$$

Equation (2.57) alone does not fully constrain the problem: for example, setting $n(\vec{r}) = 0$ everywhere is a stationary point, but a trivial and unhelpful solution. For the general case, the constraint that the *total number of electrons* $N = \int n(\vec{r}) d^3r$ is constant, or $\delta N = 0$, needs to be enforced. This can be achieved by requiring that the variation of $\int n(\vec{r}) d^3r$ with respect to variations in the total density $n(\vec{r})$ be equal to zero:

$$\frac{\delta N}{\delta n(\vec{r})} = \frac{\delta}{\delta n(\vec{r})} \int n(\vec{r}) d^3r = 0, \quad (2.58)$$

So, the global minimum of $E(n(\vec{r}))$ can be found, avoiding trivial solutions, by imposing the constraint of fixed N by using the method of Lagrange multipliers, which is achieved by writing

$$\frac{\delta}{\delta \varphi_k^*} \left(E(n(\vec{r})) - \xi_k \left(\int n(\vec{r}) d^3r - N \right) \right) = 0 \quad (2.59)$$

and attempting to solve for $|\varphi_i(\vec{r})\rangle$, where ξ_k is the Lagrange multiplier. Note the k subscript on ξ_k . It signifies the fact that it is the Lagrange multiplier which results from minimisation with respect to φ_k^* . One could have just as well sought to minimise with respect to any one of the N complex conjugate KS states. The explicit wavefunction expressions for the kinetic energy functional T_S and total $n(\vec{r})$ - equations (2.47) and (2.45) - can be substituted into equation (2.59), so that it reads:

$$\begin{aligned} 0 &= \frac{\delta}{\delta \varphi_k^*} \left(E(n(\vec{r})) - \xi_k \left(\int n(\vec{r}) d^3r - N \right) \right) \\ &= \frac{\delta}{\delta \varphi_k^*} \left(-\frac{1}{2} \sum_i^N \langle \varphi_i(\vec{r}) | \{\nabla^2 | \varphi_i(\vec{r})\} \rangle + \int V_{KS}(n(\vec{r})) n(\vec{r}) d^3r \right. \\ &\quad \left. - \xi_k \left(\int \sum_{i=1}^N \langle \varphi_i(\vec{r}) | \varphi_i(\vec{r}) \rangle d^3r - N \right) + E_{II} \right) \end{aligned} \quad (2.60)$$

Some simplifications are in order. Firstly, note that curly brackets have been used in the kinetic term in order to show that the expression $\nabla_i^2 |\varphi_i(\vec{r})\rangle$ should be considered as a single term because ∇_i^2 acts on $|\varphi_i(\vec{r})\rangle$, whereas

$\langle \varphi_i(\vec{r}) |$ is independent of the ∇_i^2 operator. It is therefore easy to see that the derivative of the kinetic term with respect to the state φ_k^* (or in Dirac notation, $\langle \varphi_i(\vec{r}) |$) reduces down to a single term:

$$\frac{\delta}{\delta \varphi_k^*} \left(-\frac{1}{2} \sum_i^N \langle \varphi_i(\vec{r}) | \{ \nabla^2 | \varphi_i(\vec{r}) \rangle \} \right) = -\frac{1}{2} \nabla_k^2 | \varphi_k(\vec{r}) \rangle \quad (2.61)$$

where the only term which survives is the term $i = k$ by virtue of the orthonormality condition from equation (2.46). Secondly, the Lagrange multiplier term can also be simplified by appealing to orthonormality in the same way:

$$\frac{\delta}{\delta \varphi_k^*} \left(-\xi_k \left(\int \sum_{i=1}^N \langle \varphi_i(\vec{r}) | \varphi_i(\vec{r}) \rangle d^3r - N \right) \right) = -\xi_k | \varphi_k(\vec{r}) \rangle \quad (2.62)$$

where $\frac{\delta}{\delta \varphi_k^*} (\xi_k N) = 0$ was used. Also of course,

$$\frac{\delta E_{II}}{\delta \varphi_k^*} = 0. \quad (2.63)$$

Equation (2.60) therefore reduces to

$$0 = -\frac{1}{2} \nabla^2 | \varphi_k(\vec{r}) \rangle + \frac{\delta}{\delta \varphi_k^*} \int V_{KS}(n(\vec{r})) n(\vec{r}) d^3r - \xi_k | \varphi_k(\vec{r}) \rangle \quad (2.64)$$

The chain rule and the explicit expression for $n(\vec{r})$ from (2.45) can now be used to deal with the V_{KS} term:

$$\begin{aligned} & \frac{\delta}{\delta \varphi_k^*} \int V_{KS}(n(\vec{r})) n(\vec{r}) d^3r \\ &= \left(\frac{\delta}{\delta n(\vec{r})} \int V_{KS}(n(\vec{r})) n(\vec{r}) d^3r \right) \frac{\delta n(\vec{r})}{\delta \varphi_k^*} \\ &= V_{KS}(n(\vec{r})) \frac{\delta n(\vec{r})}{\delta \varphi_k^*} \\ &= V_{KS}(n(\vec{r})) | \varphi_k(\vec{r}) \rangle \end{aligned} \quad (2.65)$$

The final step in equation (2.65), $\frac{\delta n(\vec{r})}{\delta \varphi_k^*} = | \varphi_k(\vec{r}) \rangle$, was achieved through orthonormality once again:

$$\frac{\delta n(\vec{r})}{\delta \varphi_k^*} = \frac{\delta}{\delta \varphi_k^*} \sum_{i=1}^N \langle \varphi_i(\vec{r}) | \varphi_i(\vec{r}) \rangle = \frac{\delta}{\delta \varphi_k^*} \langle \varphi_k(\vec{r}) | \varphi_k(\vec{r}) \rangle = |\varphi_k(\vec{r})\rangle. \quad (2.66)$$

Equation (2.64) therefore reads:

$$\left(-\frac{1}{2} \nabla^2 + V_{KS} \right) |\varphi_k(\vec{r})\rangle = \xi_k |\varphi_k(\vec{r})\rangle \quad (2.67)$$

This is a *single particle* Schrödinger equation! The kinetic and potential terms $-\frac{1}{2} \nabla^2$ and V_{KS} add together to give the *Kohn-Sham effective Hamiltonian* H_{KS} , of which $|\varphi_k(\vec{r})\rangle$ is clearly an eigenstate, and where the Lagrange multiplier ξ_k clearly can be interpreted as the *energy eigenvalue* of the state $|\varphi_k(\vec{r})\rangle$. An identical procedure can therefore be carried out for all other k values, to give a system of N *independent equations*:

$$\left(-\frac{1}{2} \nabla^2 + V_{KS} \right) |\varphi_i(\vec{r})\rangle = \xi_i |\varphi_i(\vec{r})\rangle \quad i = 1, \dots, N \quad (2.68)$$

Equations (2.68) are the *Kohn-Sham equations*. They are the foundation of all DFT. Solving them requires only that the total density $n(\vec{r})$ be consistent with the Kohn Sham potential V_{KS} . This is achieved numerically in practice via successive iterations on the form of $n(\vec{r})$ following an initial sensible guess, until the density field $n(\vec{r})$ and V_{KS} are *self-consistent*. For this reason, this manner of solving the KS equations is called the *self-consistent field method* (SCF).

2.5 Periodicity

A very large branch of materials science is concerned with crystalline materials, in which the atomic nuclei are arranged in regular periodic arrays. Periodicity represents a huge simplification, because attention can be confined to one unit cell of such a periodic array without sacrificing much information about the system, allowing calculations which are applied only to a finite volume to shed light on an effectively infinite amount of material. Crystalline materials used in engineering applications are almost always far bigger than the periodic unit cells of which they are made, so exploiting periodicity is a critical ingredient in the extension of computationally-demanding microscopic quantum mechanical calculations to macroscopic

scales. Under periodicity, the surface of the crystal is not defined. However, the Born-von Karman boundary conditions state⁶ that in crystalline materials of all but nanoscale size, the surface effects are negligible and can be justifiably ignored. Furthermore, the periodic framework is such an incredibly effective approach that even *non-periodic* systems, such as low-dimensional materials (chapters 3 and 5) and surfaces (chapter 4) are very often modelled under periodic boundary conditions. See section 2.8 for some examples. Thus despite the title of this thesis, periodicity in 3D is one of its central themes. The goal of this section is therefore to rewrite the KS equations (2.68) under periodic boundary conditions.

Consider an infinite crystal with Bravais lattice vectors \vec{R} given by

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3, \quad (2.69)$$

where \vec{a}_1 , \vec{a}_2 and \vec{a}_3 are the basis vectors of the periodic unit cell and n_1 , n_2 and n_3 are integers. The ground state electron density n_0 and its corresponding, consistent KS potential V_{KS} must be periodic functions which share the periodicity of the crystal, so that we can write

$$n_0(\vec{r}) = n_0(\vec{r} + \vec{R}), \quad (2.70)$$

and

$$V_{KS}(\vec{r}) = V_{KS}(\vec{r} + \vec{R}), \quad (2.71)$$

where the explicit dependence of V_{KS} on $n(\vec{r})$ is dropped in favour of just \vec{r} . It is worth noting that since V_{KS} is periodic, the KS Hamiltonian is also periodic because of the universality of the kinetic term. Since V_{KS} is periodic, it can be expressed as a plane wave series expansion (as can n):

$$V_{KS}(\vec{r}) = \sum_{\vec{G}} V_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \quad (2.72)$$

where $V_{\vec{G}}$ are the coefficients of the Fourier mode $e^{i\vec{G} \cdot \vec{r}}$ corresponding to wavevector \vec{G} . Since $V_{KS}(\vec{r}) = V_{KS}(\vec{r} + \vec{R})$, we can write

$$\begin{aligned} V_{KS}(\vec{r} + \vec{R}) &= \sum_{\vec{G}} V_{\vec{G}} e^{i\vec{G} \cdot (\vec{r} + \vec{R})} = \sum_{\vec{G}} V_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} e^{i\vec{G} \cdot \vec{R}} = V_{KS}(\vec{r}) e^{i\vec{G} \cdot \vec{R}} \\ &\Rightarrow e^{i\vec{G} \cdot \vec{R}} = 1 \end{aligned} \quad (2.73)$$

This implies $\vec{G} \cdot \vec{R} = 2\pi m$ for *all possible Bravais lattice translations* \vec{R} defined in equation (2.69), where m is any integer. The most general way to satisfy this is to write \vec{G} as a linear combination of vectors $\{\vec{b}_i\}$ which satisfy $\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{i,j}$, as such:

$$\vec{G} = m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3, \quad (2.74)$$

where m_1 , m_2 and m_3 are integers. $\vec{G} \cdot \vec{R}$ therefore reads

$$\begin{aligned} \vec{G} \cdot \vec{R} &= (n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3) \cdot (m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3) \\ &= 2\pi(n_1m_1 + n_2m_2 + n_3m_3), \end{aligned} \quad (2.75)$$

where the sum of products of integers in the final set of brackets is clearly the integer m . The set of all possible points \vec{G} is the *reciprocal lattice*.

Equations (2.70) and (2.71) show that the density and KS potential are invariant under Bravais lattice translation operations. The effect of applying this translation operator to any function of space is to translate the argument of the function from \vec{r} to the point $\vec{r} + \vec{R}$. The resultant eigenvalue reveals the amount by which the value of the function at \vec{r} differs from its value at $\vec{r} + \vec{R}$. Let us consider such an operator, call it $\hat{T}(\vec{R})$, and apply it to some KS eigenstate $\varphi(\vec{r})$ with a view to determining the resulting eigenvalue. Let the eigenvalue, which may generally depend on the chosen translation \vec{R} , be denoted $f_{\vec{R}}$, so that

$$\hat{T}(\vec{R})|\varphi(\vec{r})\rangle = |\varphi(\vec{r} + \vec{R})\rangle = f_{\vec{R}}|\varphi(\vec{r})\rangle. \quad (2.76)$$

The magnitude of $f_{\vec{R}}$ must be equal to 1. If not, then $|\varphi(\vec{r} + \vec{R})\rangle$ is unconstrained, and can become arbitrarily large, or zero, at arbitrarily large translations \vec{R} . Expressed another way, $\varphi(\vec{r})$ must be a *bounded* function of \vec{r} . Now consider a second, different translation $\hat{T}(\vec{R}')$, where \vec{R}' is also a Bravais lattice vector. Two translations in succession can be written as a single translation, and they are commutative, so we can write

$$\hat{T}(\vec{R})\hat{T}(\vec{R}') = \hat{T}(\vec{R}')\hat{T}(\vec{R}) = \hat{T}(\vec{R} + \vec{R}') \quad (2.77)$$

which means that

$$\hat{T}(\vec{R})\hat{T}(\vec{R}')|\varphi(\vec{r})\rangle = \hat{T}(\vec{R})(f_{\vec{R}'}|\varphi(\vec{r})\rangle) = f_{\vec{R}'}f_{\vec{R}}|\varphi(\vec{r})\rangle$$

$$= \hat{T}(\vec{R} + \vec{R}')|\varphi(\vec{r})\rangle = f_{\vec{R} + \vec{R}'}|\varphi(\vec{r})\rangle \quad (2.78)$$

which implies that

$$f_{\vec{R}'}f_{\vec{R}} = f_{\vec{R} + \vec{R}}. \quad (2.79)$$

So $f_{\vec{R}}$ must depend on \vec{R} , it must have unit magnitude so that φ is bounded, and it must satisfy equation (2.79). The most general way to write $f_{\vec{R}}$ which satisfies all of these conditions is to force it to lie on the unit circle in the complex plane, so that

$$f_{\vec{R}} = e^{i\vec{q} \cdot \vec{R}} \quad (2.80)$$

where \vec{q} is some vector inserted to make the exponent a scalar whilst retaining the dependence of $f_{\vec{R}}$ on \vec{R} . Equation (2.76) therefore reads:

$$|\varphi(\vec{r} + \vec{R})\rangle = e^{i\vec{q} \cdot \vec{R}}|\varphi(\vec{r})\rangle \quad (2.81)$$

Equation (2.81) shows that the wavefunction in a crystal at point $\vec{r} + \vec{R}$ can be different from the wavefunction at point \vec{r} by a *phase difference* $\vec{q} \cdot \vec{R}$. Thus, generally speaking, $\varphi(\vec{r}) \neq \varphi(\vec{r} + \vec{R})$ and so it doesn't respect the lattice periodicity. Note that, currently at least, there are no restrictions on the form of the vector \vec{q} . The Born-von Karman cyclic boundary conditions,⁶ which are not derived or proved here, can be imposed in order to constrain the possible values of \vec{q} . Consider a *very large but finite* supercell consisting of a periodic $N_1 \times N_2 \times N_3$ array of unit cells within an infinite crystal. The Born-von Karman boundary conditions state that the phase of a wavefunction in this infinite crystal must respect the periodicity of the nominal supercell along each of the three directions parallel to the unit cell lattice vectors, such that

$$\varphi(\vec{r} + N_i \vec{a}_i) = \varphi(\vec{r}), \quad i = 1,2,3. \quad (2.82)$$

So, equation (2.82) adds the constraint that the wavefunction's phase difference $\vec{q} \cdot \vec{R}$ must *also* conform to some periodicity, but this periodicity can be on a length scale which is far greater than the unit cell periodicity of the crystal itself, and that of the cell-periodic functions n_0 and V_{KS} . By substituting the boundary condition (2.82) into equation (2.81), the phase

difference between the wavefunction at \vec{r} and at $\vec{r} + N_i \vec{a}_i$ is forced to be zero, giving

$$\varphi(\vec{r} + N_i \vec{a}_i) = e^{iN_i \vec{q} \cdot \vec{a}_i} \varphi(\vec{r}) = \varphi(\vec{r}), \quad (2.83)$$

implying that $e^{iN_i \vec{q} \cdot \vec{a}_i} = 1$, which results in

$$N_i \vec{q} \cdot \vec{a}_i = 2\pi l_i \quad (2.84)$$

where l_i is an integer, currently unconstrained, and labelled with i in subscript to signify its correspondence to the boundary condition imposed along the direction \vec{a}_i . The vector \vec{q} can always be written as some linear combination of the reciprocal lattice vectors $\{\vec{b}_i\}$:

$$\vec{q} = x_1 \vec{b}_1 + x_2 \vec{b}_2 + x_3 \vec{b}_3 \quad (2.85)$$

where the $\{x_i\}$ are numbers to be determined, not necessarily integers. Choosing to express \vec{q} in terms of the basis vectors $\{\vec{b}_i\}$ is the obvious choice, because the reciprocal lattice relation $\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{i,j}$ can be immediately used to simplify equation (2.84) by filtering out the components of \vec{q} which are perpendicular to \vec{a}_i and therefore do not contribute anything. For the boundary condition along the direction \vec{a}_i , equation (2.84) therefore gives:

$$\begin{aligned} N_i (x_1 \vec{b}_1 + x_2 \vec{b}_2 + x_3 \vec{b}_3) \cdot \vec{a}_i &= N_i (2\pi x_i) = 2\pi l_i \\ \Rightarrow x_i &= \frac{l_i}{N_i}, \quad i = 1, 2, 3. \end{aligned} \quad (2.86)$$

Thus, the range of possible values of the vector \vec{q} under the Born-von Karman boundary conditions is restricted to only those where the $\{x_i\}$ take the values specified in (2.86). Let this particular set of vectors be denoted \vec{k} , so that

$$\vec{k} = \frac{l_1}{N_1} \vec{b}_1 + \frac{l_2}{N_2} \vec{b}_2 + \frac{l_3}{N_3} \vec{b}_3 \quad (2.87)$$

Equation (2.81) can therefore be stated in terms of a wavefunction which respects the boundary conditions by writing

$$\varphi_{\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k} \cdot \vec{R}} \varphi_{\vec{k}}(\vec{r}), \quad (2.88)$$

where the KS eigenstate has a \vec{k} in subscript to denote correspondence to a particular \vec{k} vector. This is *Bloch's theorem*.¹⁶ It states that the amount by which the wavefunction at point \vec{r} can differ from that at $\vec{r} + \vec{R}$ is a multiplicative phase factor $e^{i\vec{k} \cdot \vec{R}}$, where the vectors \vec{k} are restricted to the discrete set of values given by equation (2.87). Even though generally $\varphi_{\vec{k}}(\vec{r}) \neq \varphi_{\vec{k}}(\vec{r} + \vec{R})$, the density $n(\vec{r})$ is still periodic, because the complex phase factor in φ always cancels out upon evaluating quantities of the form $\varphi\varphi^*$ to construct $n(\vec{r})$. Bloch's theorem (2.88) can be written in alternative manner. Consider the function

$$u_{\vec{k}}(\vec{r}) = e^{-i\vec{k} \cdot \vec{r}} \varphi_{\vec{k}}(\vec{r}). \quad (2.89)$$

$u_{\vec{k}}(\vec{r})$ has the periodicity of the lattice:

$$u_{\vec{k}}(\vec{r} + \vec{R}) = e^{-i\vec{k} \cdot (\vec{r} + \vec{R})} \varphi_{\vec{k}}(\vec{r} + \vec{R}) = e^{-i\vec{k} \cdot \vec{r}} \varphi_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r}). \quad (2.90)$$

Because $u_{\vec{k}}(\vec{r})$ is a cell-periodic function, it can be expressed as a plane wave expansion over all possible reciprocal lattice vectors, i.e. \vec{G} vectors, like the KS potential from (2.72):

$$u_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} c_{\vec{G}} e^{i\vec{G} \cdot \vec{r}}, \quad (2.91)$$

The wavefunction $\varphi_{\vec{k}}(\vec{r})$ can therefore be written as

$$\varphi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} c_{\vec{G} + \vec{k}} e^{i(\vec{G} + \vec{k}) \cdot \vec{r}}. \quad (2.92)$$

It can be seen that all information about the allowed \vec{k} vectors is contained within the reciprocal lattice unit cell, or equivalently, the *first Brillouin zone* (BZ). This is because any wavevector \vec{k}' which lays outside the first BZ is equivalent to a vector \vec{k} which does lay within the first BZ, as long as \vec{k} and \vec{k}' are related by any reciprocal lattice vector \vec{G} . Consider two such wavevectors which differ by any reciprocal lattice vector \vec{G} , such that $\vec{G} = \vec{k}' - \vec{k}$. Then, using equation (2.80), and the fact that $e^{i\vec{G} \cdot \vec{R}} = 1$,

$$f_{\vec{R}} = e^{i\vec{k} \cdot \vec{R}} = e^{i\vec{G} \cdot \vec{R}} e^{i\vec{k} \cdot \vec{R}} = e^{i(\vec{G} + \vec{k}) \cdot \vec{R}} = e^{i\vec{k}' \cdot \vec{R}}. \quad (2.93)$$

Thus, by considering translations at wavevector \vec{k}' , no new information is gained from that which is gained by considering translations at wavevector \vec{k} . All wavevectors \vec{k}' which are related to \vec{k} by any combination of reciprocal lattice vectors \vec{G} are therefore redundant, and our attention can be confined to just the \vec{k} vectors which lay within the reciprocal unit cell, or equivalently the first BZ. As a result of this, the values of the integers l_1 , l_2 and l_3 from equations (2.86) and (2.87) can be restricted to the range $0 \leq l_i < N_i$. (Values which exceed N_i give \vec{k} vectors which extend outside the reciprocal unit cell.) Having imposed this restriction, it is therefore easy to see that the total number of allowed l_i values equals N_i , and therefore that the total number of unique \vec{k} vectors within the reciprocal unit cell (first BZ), equals $N_1 N_2 N_3$. Thus, *the number of allowed \vec{k} vectors within the first BZ equals the number of unit cells in the supercell.*

This is an important result, because it means that any intensive property of a crystal, (e.g. the energy per unit cell, E) which is determined by some function such as the electron density $n(\vec{r}) = \sum_{i=1}^N \langle \varphi_{i,\vec{k}}(\vec{r}) | \varphi_{i,\vec{k}}(\vec{r}) \rangle$, can be found by calculating $n(\vec{r})$ for each and every occupied \vec{k} state, and then dividing the sum of these values by $N_1 N_2 N_3$, i.e. an average over the BZ:

$$n(\vec{r}) = \frac{1}{N_1 N_2 N_3} \sum_{\vec{k}(\text{BZ})} \sum_{i=1}^N \langle \varphi_{i,\vec{k}} | \varphi_{i,\vec{k}} \rangle \quad (2.94)$$

where each of the $N_1 N_2 N_3$ \vec{k} vectors within the sum exists within the first BZ, and N is the total number of KS states per unit cell. Now let us consider the limit as $N_1, N_2, N_3 \rightarrow \infty$, i.e. the limit at which the nominal supercell becomes infinitely large. As the limit is approached, the grid of all possible \vec{k} vector endpoints within the BZ becomes an increasingly dense mesh which approaches a continuum. At the limit, the range of possible \vec{k} values becomes the countably infinite set of possible \vec{k} states corresponding to the countably infinite number of possible integer values which l_1 , l_2 and l_3 can take. Equation (2.94) then becomes an integral over the BZ:

$$n(\vec{r}) = \frac{1}{\Omega_{BZ}} \int_{\Omega_{BZ}} \sum_{i=1}^N \langle \varphi_{i,\vec{k}}(\vec{r}) | \varphi_{i,\vec{k}}(\vec{r}) \rangle d^3k \quad (2.95)$$

where Ω_{BZ} is the volume of the first BZ and d^3k are the infinitesimal k -space volume elements integrated over the volume Ω_{BZ} .

Like the modes of any Fourier series, the plane wave basis functions $e^{i\vec{G} \cdot \vec{r}}$ form an orthonormal set, so that integrals of their products are zero for all cases except those in which the \vec{G} vectors are equal:

$$\int_{\Omega_{cell}} e^{-i\vec{G}' \cdot \vec{r}} e^{i\vec{G} \cdot \vec{r}} d^3r = \delta_{\vec{G}', \vec{G}} \quad (2.96)$$

where Ω_{cell} is the Wigner-Seitz cell volume, related to its corresponding reciprocal cell volume, and therefore to Ω_{BZ} , by $\Omega_{cell} = \frac{(2\pi)^3}{\Omega_{BZ}}$. Also, clearly, the orthonormality relation (2.96) is unchanged if identical \vec{k} vectors are added to the pertinent \vec{G} vectors:

$$\int_{\Omega_{cell}} e^{-i(\vec{G}'+\vec{k}) \cdot \vec{r}} e^{i(\vec{G}+\vec{k}) \cdot \vec{r}} d^3r = \int_{\Omega_{cell}} e^{-i\vec{G}' \cdot \vec{r}} e^{i\vec{G} \cdot \vec{r}} d^3r = \delta_{\vec{G}', \vec{G}} \quad (2.97)$$

Having shown this, it is helpful at this point to adopt Dirac notation for the Fourier mode basis functions to reduce the verbosity of the notation of the reasoning that follows, so that

$$|G+k\rangle \stackrel{\text{def}}{=} e^{i(\vec{G}+\vec{k}) \cdot \vec{r}} \quad (2.98)$$

and so that (2.97) can be written as

$$\int_{\Omega_{cell}} \langle G'+k | G+k \rangle d^3r = \int_{\Omega_{cell}} \langle G' | G \rangle d^3r = \delta_{\vec{G}', \vec{G}} \quad (2.99)$$

It is also helpful to adopt a corresponding operator notation for the KS Hamiltonian:

$$\hat{H}_{KS} \stackrel{\text{def}}{=} -\frac{1}{2} \nabla^2 + \hat{V}_{KS}. \quad (2.100)$$

Using this notation, the plane wave expansion over \vec{G} vectors of the i^{th} KS wavefunction from (2.92) can be written as

$$|\varphi_{i,\vec{k}}(\vec{r})\rangle = \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} |G+k\rangle \quad (2.101)$$

and it can be substituted into the KS equations (2.68) to obtain:

$$\hat{H}_{KS} \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} |G+k\rangle = \xi_i \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} |G+k\rangle \quad (2.102)$$

These KS equations (2.102) can be subjected to a Fourier transform now by multiplying from the left by the bra vector $\langle G'+k|$, where \vec{G}' can be any reciprocal lattice vector, and then integrating over the Wigner Seitz cell using the orthonormality property (2.97). It is a good idea to carry this procedure out for both sides of equation (2.102) separately. Firstly, the ξ_i side:

$$\begin{aligned} \int_{\Omega_{cell}} \left(\langle G'+k| \xi_i \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} |G+k\rangle \right) d^3r &= \xi_i \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} \int_{\Omega_{cell}} \langle G'+k|G+k\rangle d^3r \\ &= \xi_i \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} \delta_{\vec{G}',\vec{G}} \\ &= \xi_i c_{i,\vec{G}'+\vec{k}} \end{aligned} \quad (2.103)$$

So, only the $\vec{G}' = \vec{G}$ term in the sum survives due to orthonormality. The \hat{H}_{KS} side of the KS equations (2.102), multiplied by $\langle G'+k|$ and integrated over Ω_{cell} , become

$$\begin{aligned} \int_{\Omega_{cell}} \left(\langle G'+k| \hat{H}_{KS} \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} |G+k\rangle \right) d^3r \\ = \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} \int_{\Omega_{cell}} \left(\langle G'+k| \left(-\frac{1}{2} \nabla^2 + \hat{V}_{KS} \right) |G+k\rangle \right) d^3r. \end{aligned} \quad (2.104)$$

It is best to simplify the $-\frac{1}{2}\nabla^2$ and \hat{V}_{KS} components in equation (2.104) separately. The kinetic term gives

$$\sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} \int_{\Omega_{cell}} \langle G'+k| \left(-\frac{1}{2} \nabla^2 \right) |G+k\rangle d^3r$$

$$\begin{aligned}
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \int_{\Omega_{cell}} \langle G' + k | \frac{1}{2} |\vec{G} + \vec{k}|^2 | G + k \rangle d^3 r \\
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \frac{|\vec{G} + \vec{k}|^2}{2} \int_{\Omega_{cell}} \langle G' + k | G + k \rangle d^3 r \\
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \frac{|\vec{G} + \vec{k}|^2}{2} \delta_{\vec{G}', \vec{G}} \\
 &= \frac{c_{i, \vec{G}' + \vec{k}}}{2} |\vec{G}' + \vec{k}|^2.
 \end{aligned} \tag{2.105}$$

The potential term can be dealt with by using the Fourier expansion (2.72) and placing a double prime on its Fourier modes to distinguish them from \vec{G} and \vec{G}' so that $V_{KS}(\vec{r}) = \sum_{\vec{G}''} V_{\vec{G}''} e^{i\vec{G}'' \cdot \vec{r}}$, and to then use orthonormality as above. It is actually most revealing to revert to the exponential notation as in (2.97) for the Fourier modes in this case. Using this notation from the second line onwards, the potential term simplifies as follows:

$$\begin{aligned}
 &\sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \int_{\Omega_{cell}} \langle G' + k | \hat{V}_{KS} | G + k \rangle d^3 r \\
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \int_{\Omega_{cell}} e^{-i(\vec{G}' + \vec{k}) \cdot \vec{r}} \left(\sum_{\vec{G}''} V_{\vec{G}''} e^{i\vec{G}'' \cdot \vec{r}} \right) e^{i(\vec{G} + \vec{k}) \cdot \vec{r}} d^3 r \\
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \sum_{\vec{G}''} V_{\vec{G}''} \left(\int_{\Omega_{cell}} e^{-i(\vec{G}' - \vec{G}) \cdot \vec{r}} e^{i\vec{G}'' \cdot \vec{r}} d^3 r \right) \\
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} \sum_{\vec{G}''} V_{\vec{G}''} \delta_{(\vec{G}' - \vec{G}), \vec{G}''} \\
 &= \sum_{\vec{G}} c_{i, \vec{G} + \vec{k}} V_{\vec{G}' - \vec{G}}
 \end{aligned} \tag{2.106}$$

Thus, the only non-zero terms of the potential occur when the vectors \vec{G} and \vec{G}' differ by no more than any reciprocal lattice vector \vec{G}'' , so that the coefficients $V_{\vec{G}''}$ can be written as $V_{\vec{G}' - \vec{G}}$.

By substituting the simplified terms from (2.103), (2.105) and (2.106) into the KS equations (2.102), we obtain

$$\frac{c_{i,\vec{G}'+\vec{k}}}{2} |\vec{G}' + \vec{k}|^2 + \sum_{\vec{G}} c_{i,\vec{G}+\vec{k}} V_{\vec{G}'-\vec{G}} = \xi_i c_{i,\vec{G}'+\vec{k}} \quad (2.107)$$

Finally, the notation of (2.107) can be made a bit more compact by reinserting the kinetic term into a sum over \vec{G} vectors with an appropriate Kronecker delta matrix, i.e. by writing it as the penultimate expression in (2.105), so that the entire Hamiltonian in (2.119) can be contained within a summation over \vec{G} , so that (2.119) reads:

$$\sum_{\vec{G}} \left(\frac{|\vec{G} + \vec{k}|^2}{2} \delta_{\vec{G}',\vec{G}} + V_{\vec{G}'-\vec{G}} \right) c_{i,\vec{G}+\vec{k}} = \xi_i c_{i,\vec{G}'+\vec{k}} \quad (2.108)$$

This is the *reciprocal space representation* of the KS equations under periodic boundary conditions. Their form is very similar to their direct space representation from (2.68). The effect of expressing the problem in a periodic potential in reciprocal space has been to replace the problem of evaluating the KS wavefunctions φ_i and KS potential V_{KS} with the problem of evaluating their Fourier coefficients $c_{i,\vec{G}+\vec{k}}$ and $V_{\vec{G}'-\vec{G}}$. It can be seen that the Hamiltonian takes the form of a square matrix where the number of elements is equal to the square of the number of distinct \vec{G} vectors in the sum. Being the experimental observables, the N eigenvalues ξ_i are unaffected by the Fourier transformation. There is an infinite number of possible \vec{k} vectors within the BZ, and an infinite number of possible \vec{G} vectors over the entire crystal in reciprocal space, so solving equations (2.108) requires some sensible constraints to make the problem a finite one. The conventions for this are now explained.

2.6 Practical calculations: finite plane wave basis set

It was shown in equation (2.95) that the electron density $n(\vec{r})$ is found by integrating over all \vec{k} vectors inside the BZ. When using computers to calculate integrals, numerical integration is used. It is a very fortunate fact that the *wavefunctions φ vary slowly with changing \vec{k}* ,⁶ so that the integral in

(2.94) can be accurately calculated numerically with as an average over a discrete set of \vec{k} vectors, instead of the countably infinite number of possible \vec{k} vectors in the BZ, so that we can write

$$\begin{aligned} n(\vec{r}) &= \frac{1}{\Omega_{BZ}} \int_{\Omega_{BZ}} \sum_{i=1}^N \langle \varphi_{i,\vec{k}}(\vec{r}) | \varphi_{i,\vec{k}}(\vec{r}) \rangle d^3k \\ &\approx \frac{1}{\Omega_{BZ}} \sum_{\vec{k}} \sum_{i=1}^N \langle \varphi_{i,\vec{k}}(\vec{r}) | \varphi_{i,\vec{k}}(\vec{r}) \rangle \end{aligned} \quad (2.109)$$

where the sum over \vec{k} only includes a particular set of \vec{k} vectors, or a *k points grid*, chosen to *sample* the BZ. In order to determine whether the k points sample is representative of all possible k points in the BZ, the function of interest, $n(\vec{r})$, and resultant properties such as the total energy per unit cell, must be converged with respect to increasing numbers of k points in the sum. This can be expressed alternatively as converging the energy with respect to decreasing distances between adjacent k points in reciprocal space, the *k point spacings*.

In this thesis, the scheme of Monkhorst and Pack¹⁷ is used, where the k points grid is regular, and defined as the set of points laying at the endpoints of the vectors

$$\vec{k}_{prs} = u_p \vec{b}_1 + u_r \vec{b}_2 + u_s \vec{b}_3 \quad (2.110)$$

with

$$u_p = \frac{2p - q - 1}{2q}, \quad u_r = \frac{2r - q - 1}{2q}, \quad u_s = \frac{2s - q - 1}{2q},$$

$$p, r, s = 1, \dots, q. \quad (2.111)$$

The coefficients u_p , u_r and u_s are just fractional coordinates along the reciprocal lattice vector directions, so the grid is very easy to visualise. The k point spacings along the reciprocal lattice direction \vec{b}_1 is given by

$$(u_{p+1} - u_p) \vec{b}_1 = \left(\frac{2(p+1) - q - 1}{2q} - \frac{2p - q - 1}{2q} \right) \vec{b}_1 = \frac{\vec{b}_1}{q} \stackrel{\text{def}}{=} s^1 \quad (2.112)$$

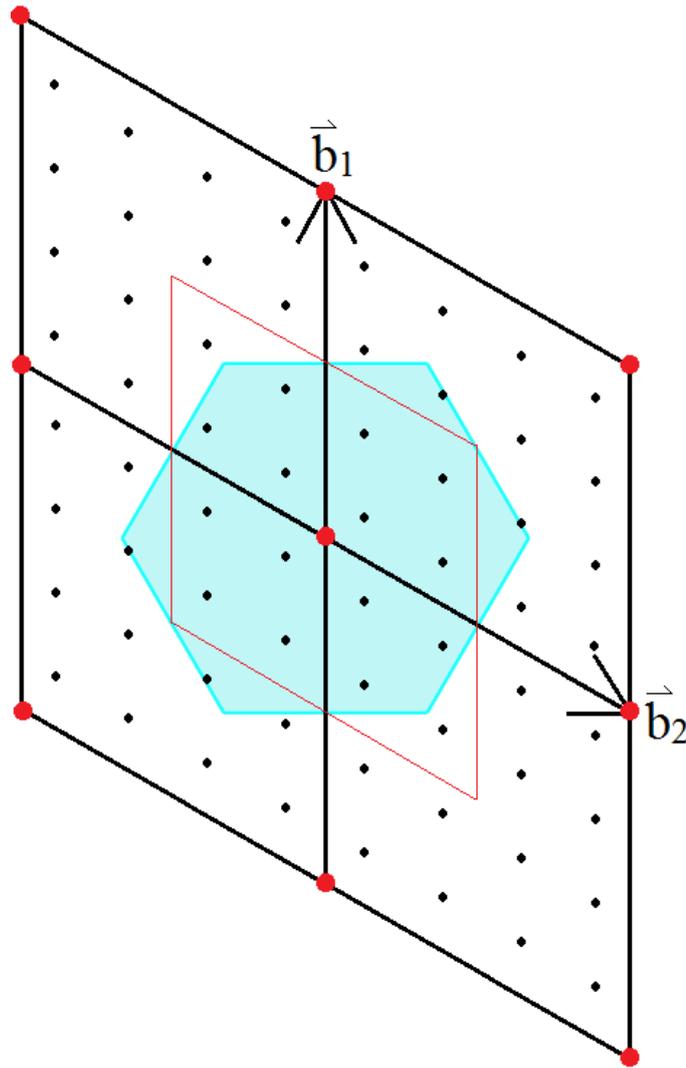


Figure 2.4. A regular 4×4 Monkhorst Pack k points grid defined in a 2 dimensional hexagonal lattice.

The black points are the k points, the red points are the reciprocal lattice points, and the blue region is the first BZ. The red cell is the region within which the fractional coordinates u_p , u_r and u_s lay. It can be seen that the k points which lay within the red cell but outside the 1st BZ have periodic equivalent points which are inside the BZ but outside the red cell. Hence, the number of grid points within the reciprocal unit cell equals the number in the BZ.

The k point spacings s^i are measured in units of (distance)⁻¹. They are the first of the two basis set parameters used throughout this thesis. Figure 2.4 shows a regular Monkhorst Pack grid of k points in 2 dimensions for a hexagonal reciprocal space unit cell (graphene, for example) and corresponding BZ, with $q = 4$.

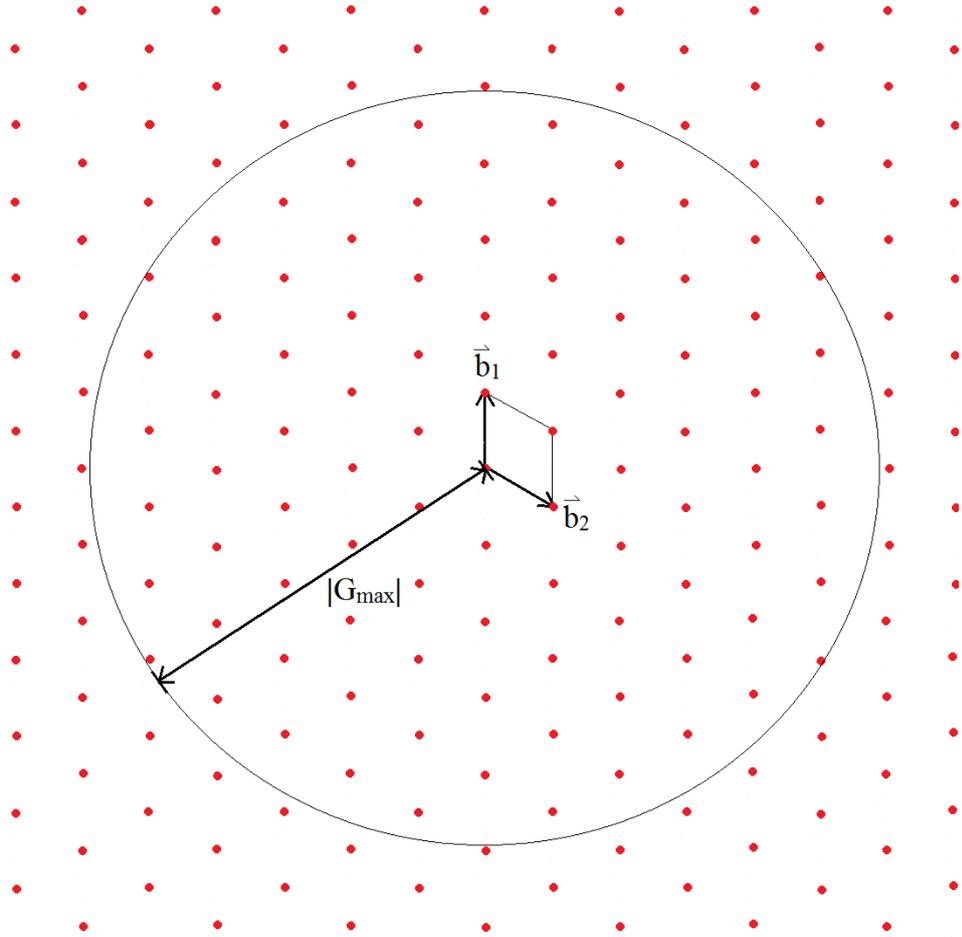


Figure 2.5. *The kinetic energy cutoff circle in 2 dimensions for a reciprocal hexagonal lattice.*

The red points are the set of all possible reciprocal lattice points $\vec{G} = m_1\vec{b}_1 + m_2\vec{b}_2$ as seen in Figure 2.4. The wavefunctions can be represented as a Fourier series whose wavevector modes corresponding to vectors \vec{G} laying within the cutoff circle radius $|\vec{G}_{max}|$, where \vec{G} vectors whose magnitude exceeds $|\vec{G}_{max}|$ are excluded from the series.

The other essential approximation used in numerical evaluation of the density and related quantities is to set a maximum possible magnitude for the \vec{G} vectors in the plane wave expansions of the wavefunction and potential. The Fourier components with the largest \vec{G} vectors represent the highest energy contributions to the wavefunction and potential, where the sharpest features are captured by these highest energy modes. The periodic features of the wavefunction and potential can always be captured to a satisfactory level of accuracy by truncating the plane wave expansion, so as

to exclude Fourier modes whose \vec{G} vectors exceed a given maximum magnitude $|\vec{G}_{max}|$:

$$\varphi_{\vec{k}}(\vec{r}) \approx \sum_{\vec{G}}^{|\vec{G}_{max}|} c_{\vec{G}+\vec{k}} e^{i(\vec{G}+\vec{k}) \cdot \vec{r}} \quad (2.113)$$

The \vec{G} vector cutoff magnitude does not discriminate against any particular reciprocal space direction (unlike the k points sampling grid, whose density can be tailored to individual directions if desired), so the allowed \vec{G} vectors lay within the *cutoff sphere* of radius $|\vec{G}_{max}|$ in reciprocal space. Figure 2.5 shows a schematic example, again for the 2 dimensional hexagonal reciprocal lattice. The cutoff sphere radius, like the k points grid density, is a basis set parameter, and so the system energy needs to be converged with respect to it.

It is clear that the kinetic energy operator applied to the i^{th} KS eigenstate gives

$$\begin{aligned} -\frac{1}{2} \nabla^2 \left(\sum_{|\vec{G}| < |\vec{G}_{max}|} c_{i, \vec{G}+\vec{k}} e^{i(\vec{G}+\vec{k}) \cdot \vec{r}} \right) \\ = \left(\sum_{|\vec{G}| < |\vec{G}_{max}|} \frac{1}{2} |\vec{G} + \vec{k}|^2 c_{i, \vec{G}+\vec{k}} e^{i(\vec{G}+\vec{k}) \cdot \vec{r}} \right) \end{aligned} \quad (2.114)$$

The energies of the largest \vec{G} vector modes in the sum are therefore slightly less than or equal to $\frac{1}{2} |\vec{G}_{max} + \vec{k}|^2$. These modes correspond to the wavevectors whose endpoints lay just within the cutoff circle radius in Figure 2.5. The largest kinetic energy that any mode can possibly have is therefore $\frac{1}{2} |\vec{G}_{max} + \vec{k}|^2$, and is denoted throughout this thesis as E_{cut} :

$$E_{cut} = \frac{1}{2} |\vec{G}_{max} + \vec{k}|^2. \quad (2.115)$$

This is the *kinetic energy cutoff*, measured in units of energy. It is the second basis set parameter against which the total system energy needs to be converged for accurate calculations to be performed. In Figure 2.5, there are 89 reciprocal lattice points which lay within the cutoff circle, so the

corresponding Hamiltonian in (2.108) would become an 89×89 matrix. Truncating the Fourier series expansions using the kinetic energy cutoff (2.115) and taking averages only over selected k points within the BZ with k point spacings (2.112) makes the basis set finite, allowing for calculations to be performed numerically on computers. The total energies of the systems studied throughout chapters 3, 4 and 5 are all converged with respect to these parameters.

2.7 Practical calculations: pseudopotentials

The other widely used simplification in modern electronic structure methods is the *pseudopotential* method. The rigorous theory behind pseudopotentials is not covered in this thesis, but the characteristics of the pseudopotentials used in calculations presented in the later chapters are briefly discussed. The motivation behind the pseudopotential method is reducing the computational cost of solving the KS equations for N Kohn Sham eigenstates by neglecting states which are closely bound to the nuclei, defined as the *core states*, and accounting only for those which are not closely bound, the *valence states*. This is achieved by replacing the external potential of the bare nuclei $V_{ext}(\vec{r})$ with a weaker, *ionic potential* $V_{ps}(\vec{r})$, which treats the nuclei as ions of lesser overall charge, each of which produces a spherically-symmetric potential experienced only by the valence states. Unlike the bare nuclear external potential $V_{ext}(\vec{r})$, the pseudopotential $V_{ps}(\vec{r})$ is not uniquely determined by the electron density, and can take a variety of different forms. In this modified regime, the KS wavefunctions are modified to become pseudowavefunctions, $\varphi_i^{ps}(\vec{r})$. The pseudowavefunctions obey the orthonormality condition, just like their all-electron counterparts from equation (2.46):

$$\langle \varphi_i^{ps}(\vec{r}) | \varphi_j^{ps}(\vec{r}) \rangle = \delta_{i,j} \quad (2.116)$$

The corresponding KS pseudo-Hamiltonian (in direct space) reads:

$$\left(-\frac{1}{2} \nabla^2 + V_{KS}^{ps} \right) | \varphi_i^{ps}(\vec{r}) \rangle = \xi_i^{ps} | \varphi_i^{ps}(\vec{r}) \rangle \quad (2.117)$$

with

$$V_{KS}^{ps} = \int \frac{1}{2} \frac{n_{ps}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' + \varepsilon_{XC}(\vec{r}) + V_{ps}(\vec{r}) \quad (2.118)$$

where the Hartree contribution is integrated over the valence electron density n_{ps} only.

In the context of this discussion, it is useful to remind the reader of the analytic solution to the single-electron Schrödinger equation for the H atom. The solutions for the time-independent single electron states $\psi(\vec{r})$ in a spherically-symmetry external nuclear potential can be straightforwardly separated into radial and angular components, so that they can be written as

$$\psi_{n,l,m_l}(\vec{r}) = \psi_{n,l}(r)Y_{l,m_l}(\theta, \phi) = \frac{\phi_{n,l}(r)}{r}Y_{l,m_l}(\theta, \phi) \quad (2.119)$$

where n , l and m_l are the usual labels denoting the principal quantum number, angular momentum quantum number and magnetic quantum number respectively, $Y_{l,m_l}(\theta, \phi)$ are the purely angular-dependent spherical harmonics, $\psi_{n,l}(r)$ is the purely radial component of $\psi_{n,l,m_l}(\vec{r})$, $\phi_{n,l}$ is given by $\phi_{n,l} = r\psi_{n,l}$, and r , θ and ϕ are spherical coordinates where $r = |\vec{r}|$.

Since each KS equation is a single-electron equation, a similar approach of resolving the wavefunction into angular and radial components can be used for a single, isolated multi-electron atom with a spherical potential. In the case of a fully-occupied outer electron shell, a *closed-shell* atom, the pseudowavefunctions $\phi_{ps}(r)$ and KS potential V_{KS}^{ps} have spherical symmetry and can be resolved into angular and radial components in a manner similar to that shown for the H atom solutions in equation (2.119), with the only conceptual difference being that the external potential needs to be calculated self-consistently along with $n(\vec{r})$. Open-shell atoms are more difficult, and require a careful treatment in terms of multiplets of the total angular momentum because V_{KS}^{ps} does not have spherical symmetry.⁶ Fortunately however, a method due to Slater¹⁸ shows how to express all open-shell calculations in terms of r only, by appealing to symmetries of the angular momentum multiplets. The full details are not given here, but the key point is that all single-atom calculations can be reduced to purely *radial* terms. Let the radial component of a KS pseudoeigenstate be denoted $\phi_{ps}(r)$, and that of the all-electron KS eigenstate be $\phi(r)$.

Norm-conserving pseudopotentials and *ultrasoft pseudopotentials* are used in this thesis. Norm-conserving pseudopotentials must satisfy five criteria, as shown by Hamann, Schluter and Chiang (HSC).¹⁹

1. The all-electron eigenvalues ξ_i and pseudoeigenvalues ξ_i^{ps} must be the same: $\xi_i = \xi_i^{ps}$.

2. The radial component of the all-electron wavefunction $\phi(r)$ and pseudowavefunction $\phi_{ps}(r)$ must agree beyond the *core radius*, r_c , such that $\phi(r)|_{r \geq r_c} = \phi_{ps}(r)|_{r \geq r_c}$.

3. The norm - or integrated electron density - of the radial component of the pseudowavefunction within the core radius must equal that of the all-electron wavefunction, such that $\int_{r < r_c} \phi_{ps}(r) \phi_{ps}^*(r) dr = \int_{r < r_c} \phi(r) \phi^*(r) dr$.

4. The logarithmic derivative of the all-electron wavefunction and that of the pseudowavefunction must be equal at, and beyond, the core radius, such that $\left. \frac{d}{dr} \left(\log \left(\phi_{ps}(r) \right) \right) \right|_{r \geq r_c} = \left. \frac{d}{dr} \left(\log \left(\phi(r) \right) \right) \right|_{r \geq r_c}$.

5. The derivative with respect to variations in the eigenvalue ξ of the logarithmic derivative must be equal beyond the core radius for both wavefunctions, giving $\left. \frac{d}{d\xi} \frac{d}{dr} \left(\log \left(\phi_{ps}(r) \right) \right) \right|_{r \geq r_c} = \left. \frac{d}{d\xi_i} \frac{d}{dr} \left(\log \left(\phi(r) \right) \right) \right|_{r \geq r_c}$.

Norm-conserving pseudopotentials are widely used because of their accuracy and transferability. A sketch illustrating a norm-conserving pseudowavefunction, all-electron wavefunction, pseudopotential and bare nuclear potential is given in Figure 2.6.

Ultrasoft pseudopotentials were devised by Vanderbilt²⁰ and are a variant on norm-conserving pseudopotentials, but with the important difference that the the integrated charge density of the ultrasoft pseudowavefunction within the core radius can differ from that of the all-electron pseudowavefunction, but where they still must agree outside the core radius. The result is that the pseudowavefunctions *violate the norm-conservation condition*, but with the significant practical advantage that they can be constructed so that they have smaller second derivatives with respect to r .

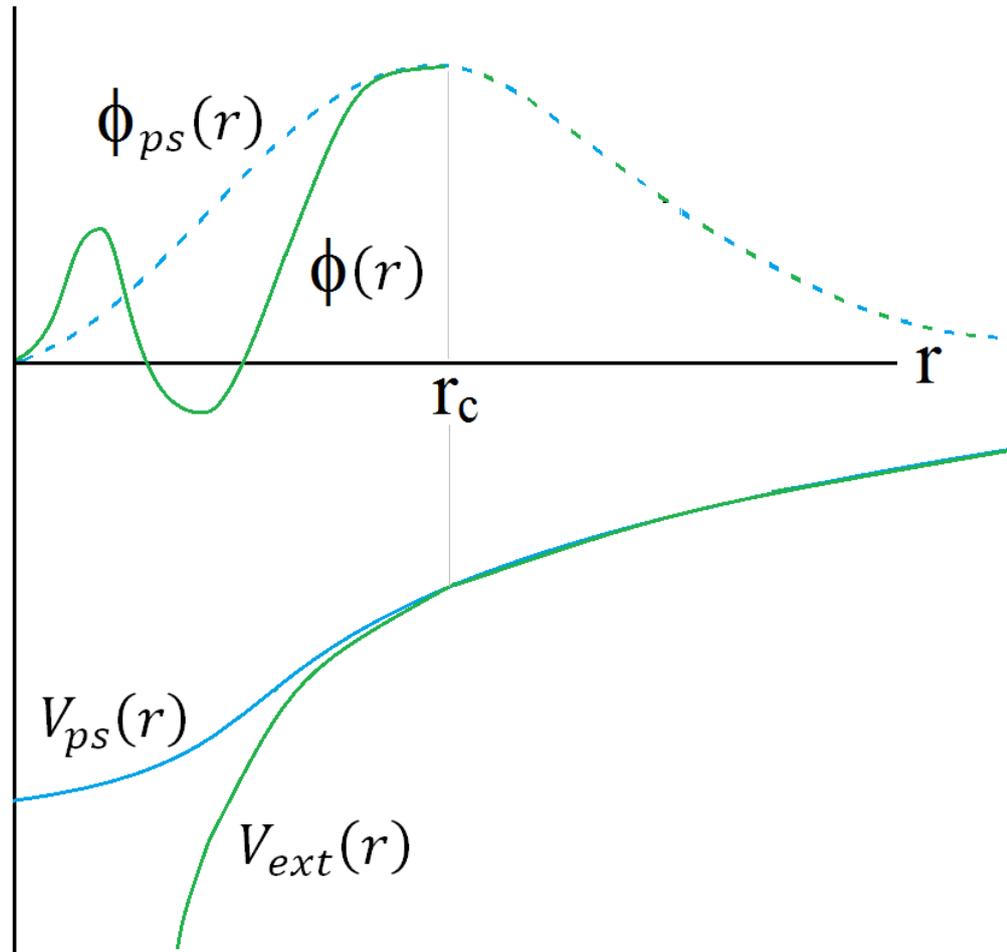


Figure 2.6. All-electron wavefunction and potential compared with pseudo-wavefunction and pseudopotential.

Both potentials and both wavefunctions are the same beyond the core radius r_c , but can differ within the core radius.

Smaller second derivatives means the resulting function has less abrupt peaks and troughs. In this sense, they are "softer" than norm-conserving pseudopotentials, and they are much softer than the very "hard" all-electron wavefunction, which has fast changes in amplitude with changing r within the core radius as illustrated in Figure 2.6. The consequence of this is that a lower cut off energy E_{cut} can be used when representing the pseudopotentials and wavefunctions in a plane-wave basis set because the sharpest features are sufficiently represented with the lower frequency Fourier modes, thus improving the efficiency of the calculation.

2.8 Non-periodic systems under periodic boundary conditions

Even though infinite periodicity is imposed in all 3 dimensions in periodic DFT calculations, it is an extremely effective method of modelling structures of lower dimensionality, such as surfaces and interfaces (2D), line defects such as steps and edge dislocations (1D) and point defects like interstitials, substitutions and vacancies (0D). It is also an effective way to model non-periodic structures such as molecules or just single atoms. The basic requirement in all cases is that appropriately chosen regions of free vacuum space are included in the periodic cells, such that the non-periodic or low-dimensional system of interest is simulated effectively as if the periodic boundary conditions were not imposed at all. Figure 2.7 shows an example for a surface calculation of TiO₂ rutile and a benzene molecule. In the strict technical sense, all such simulations are simulations of exotic bulk materials, and all calculated properties are bulk properties. The "surface energies" calculated in chapter 4, for example, are really just differences between the energies of various exotic bulk unit cells. Similarly, the "binding energies" of metal atoms on graphene calculated in chapter 3 are strictly just the quantities obtained when comparing the energies of three exotic bulk material unit cells. The vast majority of calculations carried out in this thesis are of this type. Approximations of this type, *supercell approximations*, work so well that it is entirely sensible to use terminology which is conventionally used for the non-periodic system being simulated. One very important exception is the class of systems which have a net electrostatic dipole, which can be problematic under PBC because the surface is undefined, and so require correction schemes. An overview of the issue of electric dipoles under PBC is given in section 3.2.4 and encountered directly in calculations described in section 3.3.1. Chapter 3, which follows shortly, contains a detailed description of the procedure for designing the supercell sizes and dimensions required for modelling adatoms on graphene - a 2D material with 0D defects - under 3D periodic boundary conditions. Plane-wave DFT with Vanderbilt ultrasoft pseudopotentials is used throughout chapter 3.

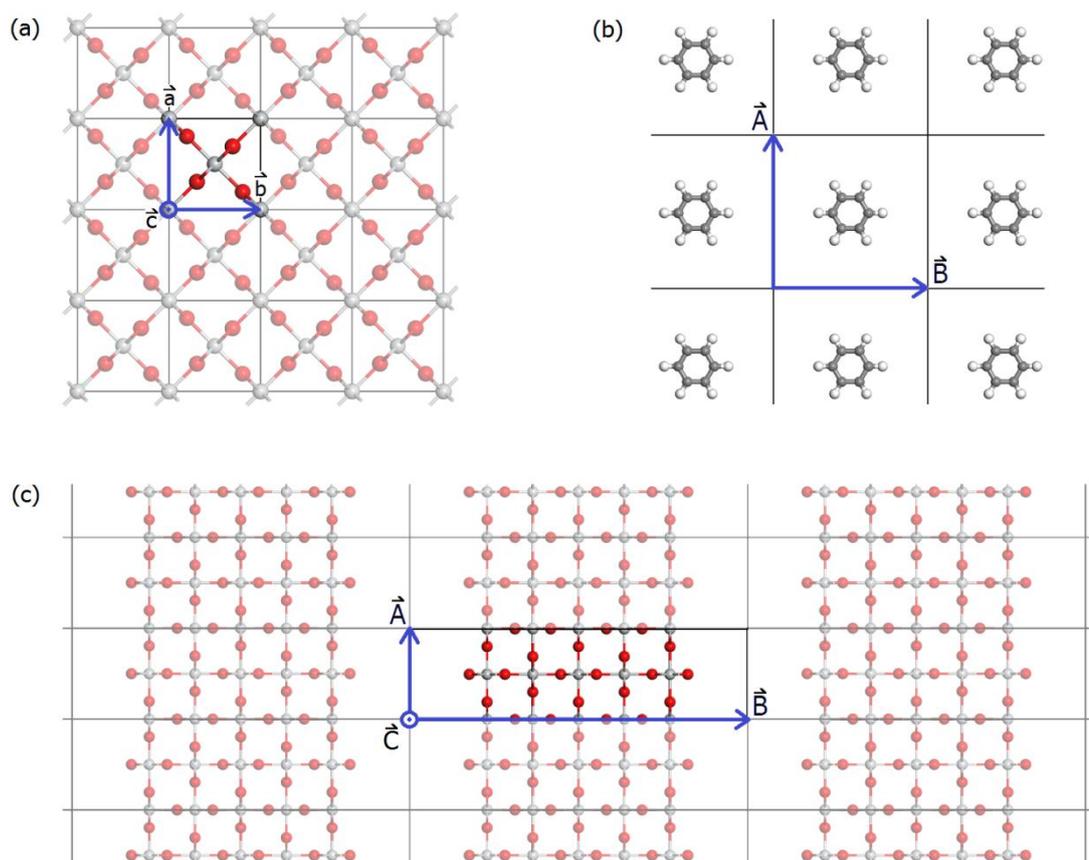


Figure 2.7. *Ball-and-stick representations of periodic and non-periodic systems modelled under periodic boundary conditions.*

(a) Conventional bulk TiO₂ rutile unit cell. (b) Isolated benzene molecule modelled as an exotic bulk material consisting of benzene molecules with vacuum space. (c) A TiO₂ rutile (110) surface modelled as an exotic bulk material. In (a), the lattice vectors \vec{a} , \vec{b} and \vec{c} are written in lower case to signify that they correspond to a normal bulk cell, whereas in (b) and (c), the bulk lattice vectors are written as \vec{A} , \vec{B} and \vec{C} to signify that they are simulating a non-bulk system. There is no technical difference between them.

2.9 Concluding remarks

This chapter has explained the basic theory required for the following chapters, but has barely begun to convey the enormity and diversity of modern DFT research which is a very broad and active research field at the time of writing. The method of actually applying the SCF minimisation approach to solve the KS equations using computer algorithms has not been covered in this chapter, and neither have any of the related tasks which build upon the theoretical foundation of SCF minimisation, such as searching for

energetic saddle points to find chemical reaction transition states or using the Hellmann-Feynman theorem^{21,22} to calculate residual forces on the nuclei following an SCF minimisation. The huge subject of nonlinear optimisation algorithms and intelligent searching schemes used in modern DFT codes has also not been explored.

In the absence of an analytical solution to the many-body problem, DFT provides an extremely practical and remarkably accurate way of studying real materials. It can be very successfully applied to systems of hundreds, and increasingly, thousands of atoms, and is vastly more efficient than the post-Hartree-Fock methods mentioned in section 2.1. At the other extreme, it is significantly less efficient than classical materials simulations based on molecular dynamical force fields, where simulations involving tens or hundreds of thousands of atoms are routine. Besides the practical issues of computational efficiency, DFT can be limited in the more serious and fundamental sense that it can completely fail to simulate certain well-understood and well-defined physical phenomena. One example is van der Waals interactions, which arise from the long-ranged coupling of regions of electron density which have a permanent and/or induced electrostatic dipole, along with higher order multipole moments. Despite their remarkable successes in accurately predicting many material properties, LDA and GGA type functionals both fail to accurately simulate van der Waals forces. Such forces can be the dominant contribution to interactions between molecules and between adsorbates and surfaces. This state of affairs has prompted much of current research to become focused on developing particular forms of the exchange-correlation functional, E_{XC} , designed to suit to particular classes of materials or to simulate particular physical phenomena. The most notable example for the case of van der Waals interactions is the total energy functional of Dion and Rydberg *et al.*²³⁻²⁵ which accounts for non-local van der Waals effects in a seamless fashion, with the total energy functional being the starting point. Alongside this, van der Waals correction schemes have also been designed by Grimme²⁶, Jurečka *et al.*²⁷ and Tkatchenko and Scheffler²⁸ (TS) to be used in conjunction with local and semi-local functionals. Long-ranged van der Waals interaction terms are dominated by $1/|\vec{r}|^6$ terms because the potential of a dipole scales as

$1/|\vec{r}|^3$, so dipole-dipole interactions scale as $1/|\vec{r}|^6$, plus higher order multipole terms. Because of this, the common feature of all these correction schemes is the addition of $C_6/|\vec{r}|^6$ -type terms to the total energy where C_6 are the coefficients whose values determine the strength of the various interatomic pairwise interactions. The Grimme²⁶ and TS²⁸ schemes are both used in this work. They differ slightly in that the coefficients in the Grimme scheme are parameterised according to empirical reference data, whereas in the TS scheme they are determined using a self-consistent field approach. Both schemes are approximations since they neglect all higher order interaction terms beyond pairwise interactions. The key point is that DFT has many great strengths but also some significant limitations, and *choices* must be made when doing calculations. As a consequence of this, it has become the responsibility of theoretical modellers to motivate the choice of functional, along with any correction schemes, when using DFT. In summary, DFT is most effectively deployed for systems for which hundreds of atoms are sufficiently many to build the structures of interest, where the required level of theory extends well into the quantum regime, but where the capabilities of the chosen functional at simulating relevant phenomena is well understood. The remainder of this thesis is devoted to applying modern DFT to some current problems in materials science and nanotechnology.

Chapter 3: Mobility of metal adatoms on graphene

Graphene could form the basis of future nanoelectronic devices. One line of enquiry in this field is the issue of contacting graphene to metals with a view to providing an interface with existing electronics. To begin to investigate whether this may be a real possibility, a theoretical investigation into the basic interaction between graphene and metals is conducted in this chapter.

3.1 Abstract

Recently published aberration-corrected scanning transmission electron microscopy (STEM) images have shown that metal atoms evaporated onto graphene by chemical vapour deposition (CVD) are only ever observed at edge sites and contaminated regions, but not on the pristine regions of graphene. It was hypothesised from these observations that metal adatoms are very mobile on graphene at room temperature and therefore quickly migrate randomly across the lattice until they bind to more energetically-favourable edge sites by the time the samples reach the microscope. To test this hypothesis, plane-wave DFT calculations have been used to optimise the structures of Al, Au and Cr atoms on the adsorption and edge sites of monolayer, bilayer and trilayer graphene, and their energies and bonding characters have been compared. Migration energy barriers between the adsorption sites were then calculated. It was found that Al, Au and Cr atoms form very weak bonds at the adsorption sites but form strong chemical bonds at the edge sites, and the migration activation barriers were all found to be very small: within an order of magnitude of $k_B T$ at $T = 300$ K, where k_B is Boltzmann's constant. It has been concluded from these calculations and the STEM observations that metal adatoms undergo random thermal diffusion on graphene at room temperature until they bind with edge and defect sites, thus verifying the original hypothesis.

3.2 INTRODUCTION

3.2.1 Recent theoretical studies of graphene-metal systems

The original synthesis of graphene²⁹ has subsequently sparked worldwide attention owing to its potential to revolutionise many areas of industry. Nanoelectronics is one such promising area, in which interfacing graphene via metal adatom/cluster contacts is a recurring theme³⁰⁻³⁶. This area of research is still developing and the consequences of particular dopants on the electronic properties of graphene are still being investigated. Widespread implementation of graphene-based electronics will therefore involve developing a more detailed understanding of metal-graphene interactions on a fundamental level. To this end, many theoretical studies using density functional theory (DFT) have already emerged which present predictions of binding energies and relaxed structures of various metal adatoms and clusters on pristine single layer graphene³⁷⁻⁵² and on graphene defect structures⁵³⁻⁵⁹. Potential contacting applications will depend very much on the metal used because vacancy formation energies can be greatly reduced by certain dopants. In 2010, Karuoi *et al.*⁵⁹ predicted that a Ni substrate assists graphene in healing its vacancy defects. In contrast, Boukhvalov and Katsnelson⁵⁴ predicted in 2009 that Fe, Ni and Co adatoms dramatically reduce vacancy formation energies in graphene, destroying it in the process. In this latter study Au atoms were predicted to have almost no effect on graphene vacancy formation energies, thus preserving its strength. Recently, electron microscopy observations were made of nanoscale holes being etched into pristine regions of graphene by various metal adatoms with the exception of Au for which no etching process was seen to occur.⁶⁰

Trends have emerged regarding the preferred binding sites for metal adatoms on graphene at absolute zero. Recent DFT studies^{39,42-44} predict that transition metals generally adsorb at the hollow (H) site (see Figure 3.3). Au atoms have been predicted to adsorb preferentially to the atop (A) site^{39,44,45}. It can be easy to erroneously conclude from these studies that one would expect stable and static configurations for these adatoms to exist on the basis of the local energetic minima predicted by geometry optimisation calculations at absolute zero. However, the calculated absolute

difference in binding energy between adsorption sites is often very small, so it is sensible to suggest from these studies alone that the activation barriers for adatom migration are also small; small enough that the perturbing effects of room temperature, $T \sim 300$ K, cause certain metal adatoms to be highly mobile on graphene at room temperature^{42,44,57}. Indeed, many of these studies state, either directly or indirectly, that metal adatoms migrate on pristine regions of graphene.^{39,42-46,51,55,57} In particular, in the work of Yazyev *et al.*⁵¹, the migration energy barrier for a Co adatom was calculated explicitly using a transition state scheme and found to be 0.4 eV. This is not an insignificant barrier, which suggests that Co migrates slowly, or not at, all on graphene. The underlying concern with results of this type is that the accuracy of such energy values can vary significantly depending on the level of theory used in the calculations. Indeed, it is clear from published results that adatom binding energies are very sensitive to the exchange correlation functional used. To illustrate this, Table 3.1 shows binding energies of a single Au adatom on pristine single layer graphene taken from some recent *ab-initio* DFT studies.^{37,38,40,45,46,53}

Table 3.1. Recently published DFT-calculated Au adatom/graphene binding energies on the 3 high symmetry adsorption sites of single layer graphene.

Negative binding energies signify that the configurations are stable, as per equation (3.1).

Binding energy, E_b , of Au adatom on pristine single layer graphene / eV. All values quoted to 3 d.p. unless otherwise specified									
<i>XC Functional</i>	<i>LDA</i>			<i>GGA PBE</i>			<i>GGA PBE van der Waals-corrected {correction scheme used}</i>		
Author	Atop (A)	Bridge (B)	Hollow (H)	Atop (A)	Bridge (B)	Hollow (H)	Atop (A)	Bridge (B)	Hollow (H)
³⁷ Lima	-	-	-	-0.410	-	-	-	-	-
⁵³ Tang	-	-	-	-0.075	-	-	-	-	-
³⁸ Ding	-0.77 (2d.p.)	-	-0.50 (2d.p.)	-0.16 (2d.p.)	-	-0.16 (2d.p.)	-	-	-
⁴⁰ Varns & Strange	-0.79 (2d.p.)	-0.74 (2d.p.)	-0.52 (2d.p.)	-	-	-	-	-	-
⁴⁵ Chan	-	-	-	-0.096	-0.089	-0.085	-	-	-
⁴⁶ Amft	-0.732	-0.698	-0.451	-0.099	-0.081	no bond	-0.385 {Dion <i>et al.</i> ²³⁻²⁵ }	-0.314 {Dion <i>et al.</i> ²³⁻²⁵ }	-0.322 {Dion <i>et al.</i> ²³⁻²⁵ }
							-0.886 {Grimme ²⁶ }	-0.881 {Grimme ²⁶ }	-0.870 {Grimme ²⁶ }

The local density approximation (LDA) functional is well known to significantly overbind compared to the generalised gradient approximation as parametrised by Perdew, Burke and Ernzerhof⁶¹ (GGA PBE). This is evident from the values shown in Table 3.1. Despite the widespread success of the GGA PBE functional, it fails to accurately simulate non-local correlation effects which dominate in many biological and chemical systems. These systems are characterised by weak long-ranged interactions between instantaneous multipoles occurring in the electron density, collectively and commonly referred to as van der Waals forces. The total energy functional of Dion and Rydberg²³⁻²⁵ accounts for non-local correlations in the form of van der Waals interactions, but this was not used. The GGA PBE functional fails to simulate interlayer interactions in graphite and multilayer graphene, thereby making non-corrected GGA functionals inappropriate for modelling involving structural relaxation of the systems in this chapter. GGA-type van der Waals correction schemes for implementation into DFT codes have been designed by Grimme²⁶, Jurečka *et al.*²⁷ and Tkatchenko and Scheffler²⁸ (TS). These correction schemes allow for new insight to be gained into possible surface physisorption bonding mechanisms between graphene and metal adatoms; an effect which is impossible to probe with the native LDA and GGA PBE functionals employed in virtually all DFT studies so far published. Moreover, van der Waals-corrected DFT sheds light on the graphene-metal interaction, not least because physisorption may be involved, but also because many laboratory synthesis methods produce samples containing regions which are multilayered^{60,62-65,67,68,69} and are therefore graphitic in character. To my knowledge, only two *ab-initio* DFT studies, by Amft *et al.*⁴⁶ and Ming *et al.*⁶⁶, have incorporated non-local correlation effects with graphene/graphite-metal adatom systems. Amft *et al.*⁴⁶ used the GGA-type correction scheme of Grimme²⁶ and the non-local van der Waals functional of Dion and Rydberg²³⁻²⁵ on single layer graphene/metal systems, and their values shown here in Table 3.1 aptly demonstrate the drastic effect of including these interactions.

3.2.2 Electron microscopy studies of graphene-metal systems

Whilst theoretical studies of graphene-metal systems are ubiquitous, the first significant experimental insight of this system that has recently emerged is from a series of images recently published by collaborators^{60,67,68,69} using aberration-corrected scanning transmission electron microscopy (AC STEM) at 60 keV, examples of which are presented in Figure 3.1. In these studies, high angle annular dark field (HAADF) imaging was used to produce images which clearly showed suspended monolayer graphene membranes consisting of pristine regions along with defective and hydrocarbon-contaminated regions, onto which various metal adatoms had been evaporated. As Figure 3.1 shows, Au and Al atoms are observed exclusively at edge sites on the perimeters of etched holes or are clustered at hydrocarbon-contaminated regions.

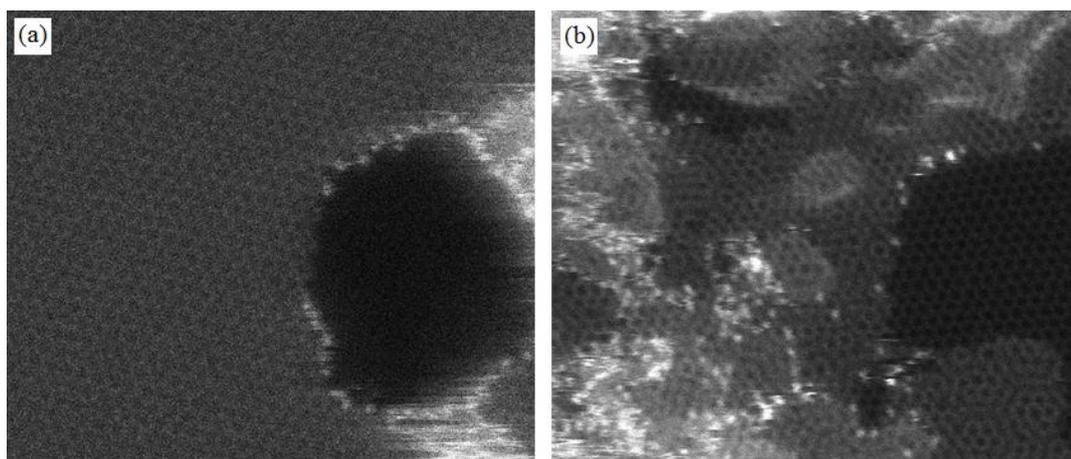


Figure 3.1. STEM HAADF images at 60 keV showing preferential binding of metal atoms to edge defects, hydrocarbon-contaminated regions and metal clusters.

(a) Monolayer graphene sheet with hole, onto which a 2Å layer of Al was evaporated. Al atoms are seen only at edge sites and in clusters near the hole. (b) Monolayer graphene sheet with bilayer and trilayer regions onto which a 5Å layer of Au gold was evaporated. Individual Au atoms and Si contaminants (of less bright contrast) clearly bind preferentially to edge sites. The pristine regions of the lattice are completely devoid of adatoms in both cases.

No metal *adatoms* were ever observed on these samples; only metals bound at edge sites, defect sites and contaminated regions. A sample of variable thickness consisting of monolayer and multilayer regions and evaporated

with Au was also prepared. By following the method of Eberlein *et al.*⁷⁰, collaborators used electron energy loss spectra (EELS) to identify the monolayer, bilayer and trilayer regions unambiguously for this sample. The remaining regions were collectively identified as consisting of 4 or more layers. A very small number of isolated Au adatoms were found momentarily on the pristine regions of this sample, but only on areas whose thickness could be unambiguously identified as 4 or more layers. By using the approximate proportionality of the image intensity to the square of the atomic number, Z , the adsorption positions of these Au adatoms were determined, and found to be consistently at atop sites. In these studies it was speculated that the graphene-metal binding energy may be significantly higher for thicker samples on account of the van der Waals-type contribution from the sublayers. It was also speculated that all metal adatoms were very mobile on all samples and had migrated to defective and contaminated regions, presumed to be more stable, before the samples were characterised in the microscope. The immediate implication from these observations is that fabrication of stable metal-graphene layer-upon-layer type interfaces is unlikely to be possible from this particular synthesis method. However, it cannot be said that *in-plane* metal-graphene interfaces should be ruled out. Indeed, one group has very recently, and very significantly, succeeded in synthesising free-standing atomically-thin Fe films in graphene holes.⁷¹ Thus it appears that atomically-thin, truly 2D, in-plane metal-graphene interfaces are a real possibility. It is important to rationalise all of these observations in theoretical terms in order to identify methods by which metal/graphene systems may be manipulated to fabricate nanostructures for potential deployment in future devices, so the work carried out in this project finds its relevance in providing this foundational theoretical understanding.

The first aim of this chapter is to use van der Waals-corrected DFT to predict the binding energy of selected metal adatoms at the high symmetry sites of pristine regions of graphene, and at the most commonly observed monolayer edge defects, in order to compare the energetic stability of these regions. The second aim is then to investigate adatom mobility on the pristine substrates by directly sampling the energy landscape corresponding to intermediate configurations between high symmetry adsorption sites in

order to locate the transition state saddle points and thus evaluate the migration activation barriers. The van der Waals corrections will produce explicit and original evidence of how adatom binding energy and mobility changes with increasing graphene substrate thickness, if at all. To my knowledge, this is the first DFT study of a multilayer graphene-metal interaction to make a direct comparison with STEM data, the first study to calculate adatom migration barriers on both monolayer and multilayer graphene (real STEM specimens consist of multilayer regions in addition to single layers), and also the first such study to incorporate the van der Waals correction scheme of Tkatchenko and Scheffler²⁸. In addition, the apparent lack of agreement in the fixing of atomic positions during geometry optimisation calculations is addressed, a discussion of which now follows.

3.2.3 Long-ranged lattice perturbations

The essence of the approximation with graphene adsorption studies is attempting to simulate the asymptotic flatness and stiffness of graphene far from the adsorbate, whilst accounting for the fact that adsorbate-induced lattice perturbations can be long-ranged, all under the constraints of finite supercell sizes dictated by the efficient use of shared computing architectures. (Real graphene is known to have ripples under typical laboratory conditions^{72,73}, but these effects are neglected here as the period of these oscillations is relatively large.)

Lambin *et al.*⁷⁴ recently demonstrated that for the case of nitrogen substitutional dopants in graphene with the LDA functional, the calculated local density of states differs significantly for 9×9 and 10×10 supercells. Although adatom-induced lattice perturbations are likely to be smaller than those of substitutional dopants, the convergence of adatom binding energy should ideally be tested with supercell size, or the error due to the use of finite supercell sizes should at least be estimated. Some tests were carried out using the LDA functional with 4×4 supercells with Au and Cr adatoms placed in the centre, in which all carbon atoms were relaxed. It was found that out-of-plane lattice perturbations were significant at the supercell boundaries far from the adatom in response to the localised puckering near

the adatom. This raises the question of whether such undulating structures are a physically meaningful simulation of graphene at all. Also, there is no well-defined way of measuring the distance of the adsorbate above the graphene plane in these systems. Despite it being something of an artifice, it is advocated in this thesis that fixing the positions of selected carbon atoms far from the adatom is a pragmatic way to simulate the stiffness and flatness of pristine graphene far from the adsorbate, but only if the supercells used are large enough to account for lattice perturbations to a justifiable level of energy convergence.

Further on the issue of fixing atomic positions, there appears to be no general consensus on the issue of *which* atomic positions should be fixed. The opportunity is taken now to list the conventions used in recently published studies to illustrate the disparity, and then suggest a simple guiding principle for future studies. In the study by Sargolzaei and Gudarzi⁴¹, the positions of the adatom and the first nearest-neighbour carbon atoms were relaxed, with all other carbon positions fixed. Ding *et al.*³⁸ state that they allowed all atomic positions to relax in the direction normal to the graphene plane, but it is unclear whether they also allowed for in-plane relaxations. Tang *et al.*⁵³ allowed all atoms in the calculation to relax in all directions. Amft *et al.*⁴⁶ appear to have used the still different method of fixing the positions of the adatom and the carbon atoms on the supercell perimeter, whilst all other carbon positions were allowed to relax. Nakada *et al.*⁴² used yet another method and allowed all atoms to relax except for just one carbon atom far from the adatom, with the adatom only allowed to relax in the z (vacuum) direction. Whilst these different choices may or may not result in negligible differences in calculated binding energies for a given supercell size, most of them can introduce the easily avoidable idiosyncrasy of breaking the symmetry of the system. This is illustrated in Figure 3.2 which shows a 32 atom graphene monolayer supercell and the atop adsorption site (A) indicated with a red cross in the centre of the supercell. The C atoms on the supercell perimeter are indicated in blue to signify that their positions are fixed, whilst all remaining C atoms indicated in black are allowed to relax.

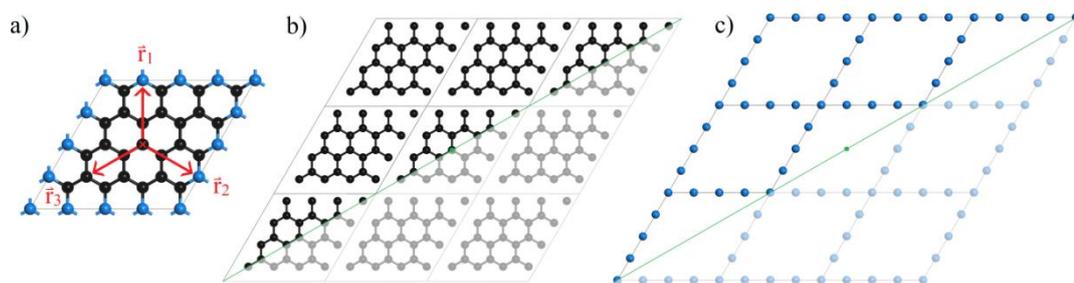


Figure 3.2. *Symmetry-breaking caused by fixing atoms on the supercell perimeter without appealing to lattice symmetries.*

(a) The lattice environment experienced by the adatom along directions \vec{r}_3 is different to that along \vec{r}_1 and \vec{r}_2 , despite these directions being crystallographically equivalent. (b) The resulting 2-fold rotational symmetry of the unfixed carbon sublattice and (c) the 2-fold rotational symmetry of the fixed carbon sublattice about the axis passing through the adsorption site.

By fixing the atoms indicated, the lattice environment encountered along the directions \vec{r}_1 and \vec{r}_2 is not the same as that along the direction \vec{r}_3 , despite the fact that these three directions are all supposed to be crystallographically equivalent. In fact, the resulting sublattices consisting of fixed and unfixed C atoms each have 2-fold rotational symmetry about the adsorption site as shown in Figure 3.2 (b) and (c), in contradiction with the 3-fold rotational symmetry of the complete lattice about the adsorption site. To restore the symmetry and create an environment for the adatom which is unbiased, C atoms are selected to be fixed in the supercells so that i) C atoms which are fixed form a sublattice which shares the rotational symmetry of the complete lattice about the axis passing through the adsorption site of interest and ii) all of the remaining unfixed C atoms form a sublattice which shares the rotational symmetry of the complete lattice about that same axis. This is illustrated in Figures 3.14 and 3.15.

3.2.4 Binding sites, binding energy and dipole corrections

For the pristine regions, attention is confined to the high symmetry points lying at the vertices of the symmetry-reduced Wigner-Seitz cells of the single and multilayer systems as indicated in Figure 3.3 (a) and (b). For lattice edges, the boundaries separating irreducible regions of the “zigzag” and

“armchair” edges indicated in Figure 3.3 (c) and (d) are considered for the monolayer case for each of the 3 metals tested.

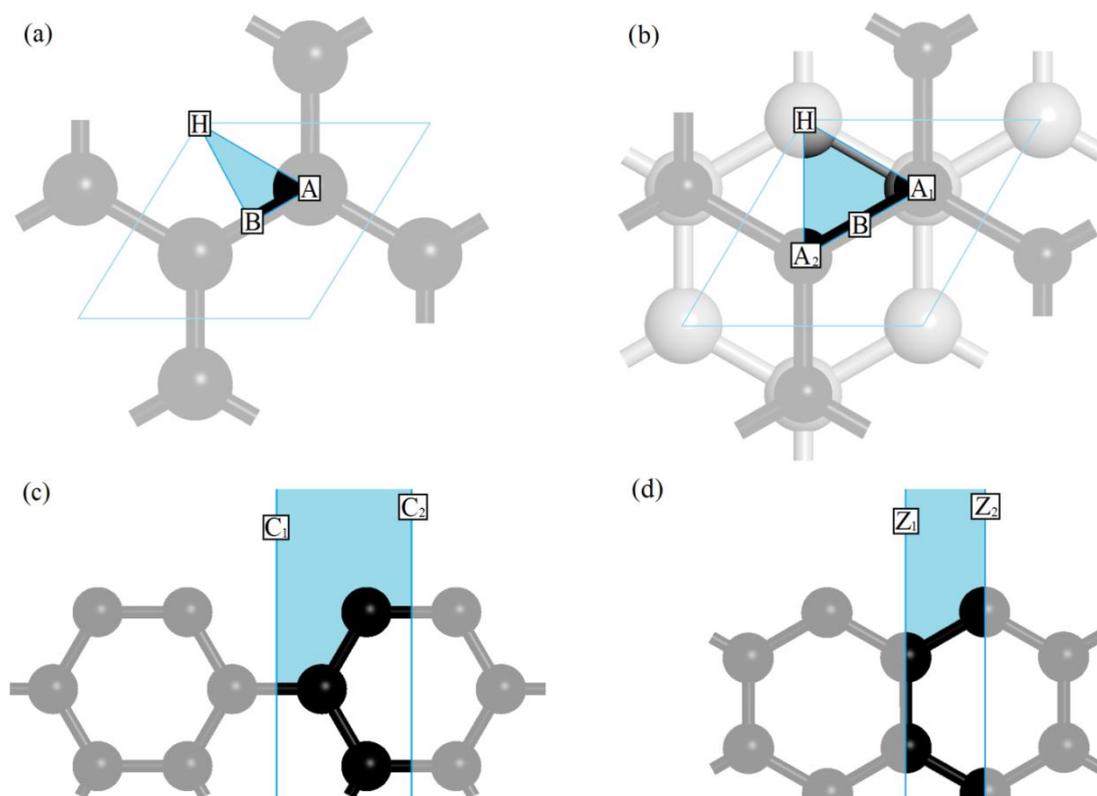


Figure 3.3. The high symmetry adsorption sites located at the vertices of the symmetry-reduced Wigner Seitz cell boundaries.

(a) single layer graphene and (b) 2+ layer graphene, for which AB stacking is assumed. In the multilayer case, the top layer is represented by small black balls and sticks and the sublayer is represented by large grey balls and sticks. (c) & (d) The high symmetry binding sites of the monolayer armchair edge and zigzag edge considered in this chapter.

The binding energy E_b at site X - where X takes the value $X = A, A_1, A_2, B, H$ for adsorption sites, or C_1, C_2, Z_1, Z_2 for edge defect sites as appropriate - is defined in the conventional way as the difference in enthalpy of the composite system supercell and that of the sum of the two isolated system supercells:

$$E_b(X) = E_{C+m}(X) - E_m - E_C \quad (3.1)$$

where E_{C+m} is the TS-corrected enthalpy of the geometry-optimised graphene / metal supercell and E_m and E_C are the TS-corrected enthalpies of the isolated metal and geometry-optimised graphene supercells respectively.

Note that because two metal atoms are present in the edge site supercells, the value of E_b must be divided by two for these cases. (See Figure 3.16.)

One subtle but essential physical ingredient which can interfere with adsorption calculations is that of electrostatic polarity under periodic boundary conditions (PBC). A well-known difficulty which dates back to classical electrostatics is that the polarisation of an ionic crystal can depend on the definition of the (neutral) bulk unit cell if no explicit reference is made to the surface conditions. This has been explained in the context of *ab initio* calculations by Makov and Payne⁷⁵. Under the constraints of PBC, the crystal is infinite so the surface is undefined. Thus, with no surface cell to cancel out the spurious potential produced from unphysical interactions between periodic images of multipole moments in neighbouring supercells, the dipole moment of a neutral polar system can depend on the location of the supercell boundaries, or equivalently, on the placement of the system within the supercell. This positional-dependence of the energy arises because of electron density overlapping with the cell boundary in the direction of the polarity, thus making the total cell dipole sensitive to the placement of the system. Metal adatom-graphene systems, especially adsorption configurations, tend to be polar in the vacuum direction owing to the charge transfer associated with the metal-carbon bond. Hence, it is essential that the systems are placed in the centre of the vacuum slab far from the supercell boundary at each end of the vacuum so as to ensure that the charge density is zero across this boundary. Various dipole correction schemes and studies of the subject have been published⁷⁶⁻⁸². In this chapter the self-consistent electrostatic dipole correction scheme of Neugebauer and Scheffler⁷⁶ as implemented in CASTEP⁸³ is used to ensure that the input files satisfy the condition of zero charge density at the extremities of the vacuum slab.

3.3 METHOD

3.3.1 Basis set parameters

Two van der Waals-corrected cell-optimised geometry optimisation calculations were carried out on the bulk graphite unit cell using the plane

wave density functional theory code CASTEP⁸³ with the GGA PBE functional⁶¹, Vanderbilt ultrasoft pseudopotentials²⁰ and an initially extremely-converged basis set. The TS van der Waals correction scheme²⁸ as implemented in⁸⁴ CASTEP was used for the first calculation and the Grimme scheme²⁶ for the second. The fully-optimised Grimme-corrected final interlayer spacing was found to be 3.27 Å (3 s.f.), whereas the TS-corrected interlayer spacing was found to be 3.32 Å (3 s.f.); considerably closer to the experimentally measured⁸⁵ value of 3.35 Å (3 s.f.). On the basis of this result, the TS correction scheme was selected for all subsequent calculations.

TS-corrected GGA PBE total energy calculations were then carried out to numerically converge the binding energy, E_b , of the two-layer graphene-metal systems shown in Figure 3.4 (d) (with a 24 Å vacuum) with respect to the kinetic energy cutoff E_{cut} and the k-point spacings s^i , where $i = 1,2,3$ denotes correspondence to the reciprocal lattice vector b_i . The binding energies E_b were converged by varying the two basis set parameters E_{cut} and s^i independently. Firstly, in order to converge E_b with respect to the cutoff energy E_{cut} , the energies of supercells identical to the form shown in Figure 3.4 (d) were calculated using fixed and regular $7 \times 7 \times 1$ Monkhorst-Pack¹⁷ k points grids. To ensure that these binding energies were truly independent of the k-points sampling, they were obtained by calculating the energies of the isolated metal atom, isolated graphene structure and metal+graphene structure in the exact locations and in a supercell with exactly the same lattice vectors as that shown in Figure 3.4 (d). This of course means that any artefacts relating to the k-sampling in the choice of supercell cancels out to expose the convergence of all energies with respect to E_{cut} only. The convergence of the total supercell energies is plotted in Figures 3.5 and 3.6, and that of the binding energies in Figure 3.7. On the basis of these calculations, the kinetic energy cutoff E_{cut} was fixed at $E_{cut} = 550$ eV for all subsequent calculations.

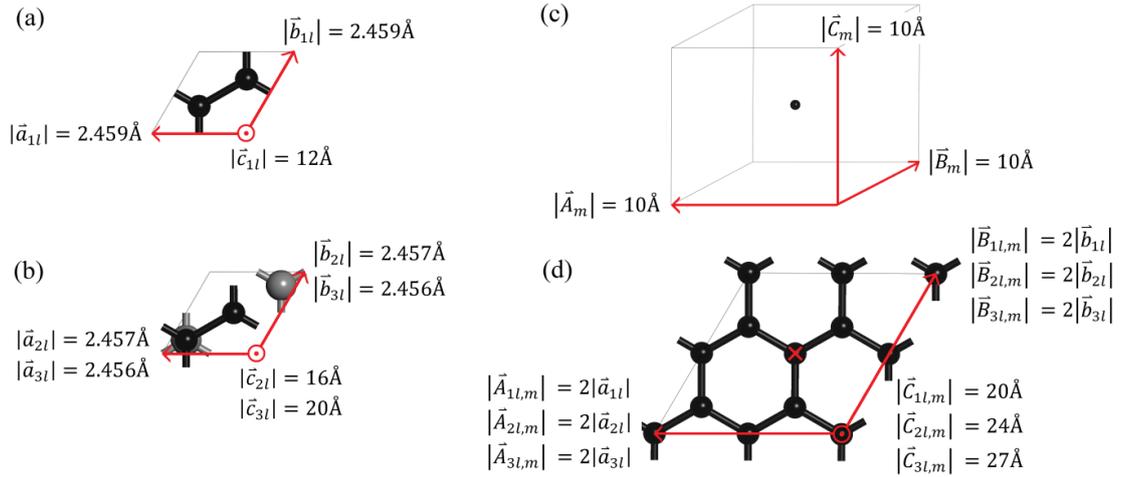


Figure 3.4. Periodic cells used for basis set.

(a) Fully optimised graphene unit cell with relaxed lattice parameters in red. Atoms and bonds are represented by balls and sticks respectively. (b) Fully optimised multilayer graphene unit cell, as in (a). To aid visualisation, the atoms and bonds of the first sublayer are represented with large grey balls and sticks, and those of the top layer with small black balls and sticks. The second sublayer is not indicated owing to the assumed AB stacking structure (c) Isolated metal atom cubic supercell. The lattice parameters shown indicate the smallest supercell size required to decouple all intercellular metal-metal interactions (d) Graphene + metal supercell spanning 2×2 unit cells. The lattice parameters shown indicate the vacuum thicknesses required to decouple intercellular interactions along the vacuum direction only.

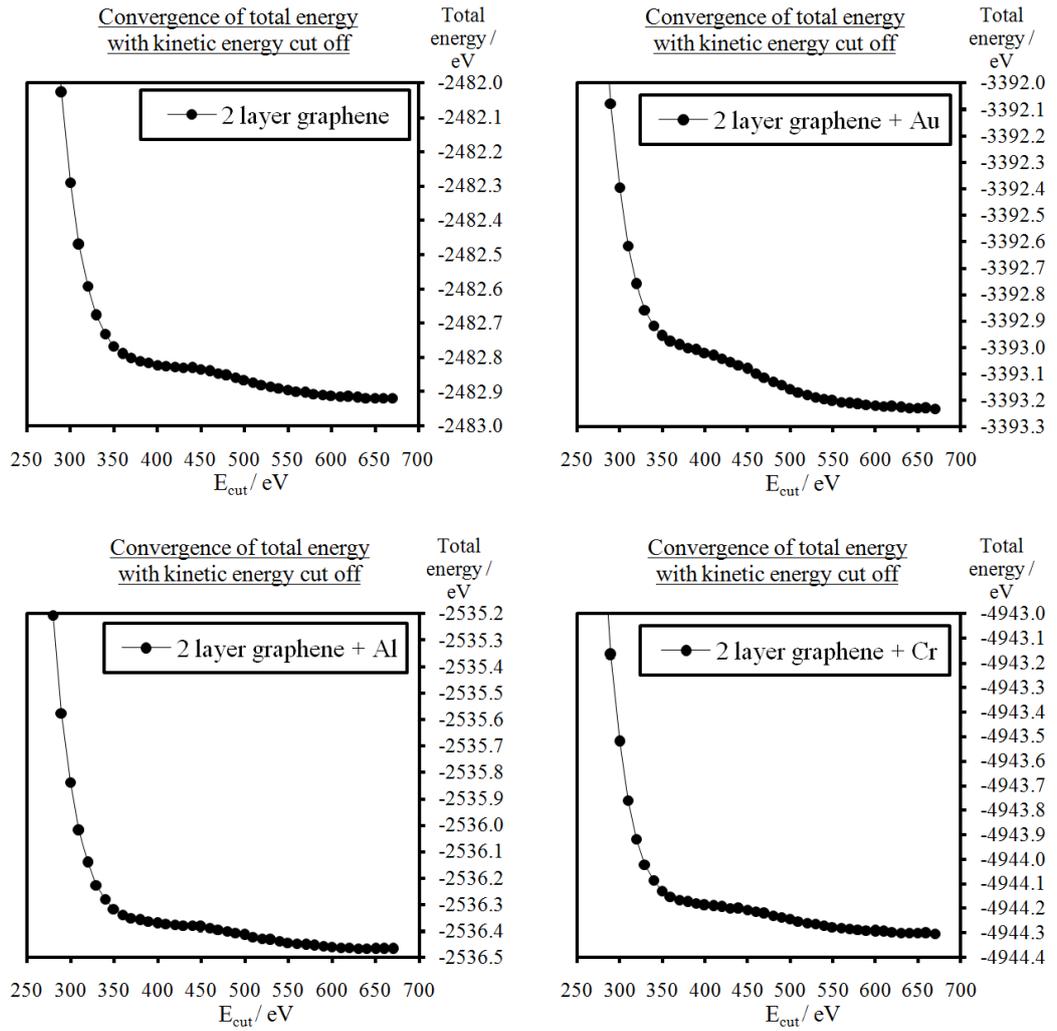


Figure 3.5. Monotonic convergence of total supercell energies with increasing kinetic energy cut off for graphene and metal-graphene systems.

All energies are negative, indicating net stability of the systems in question.

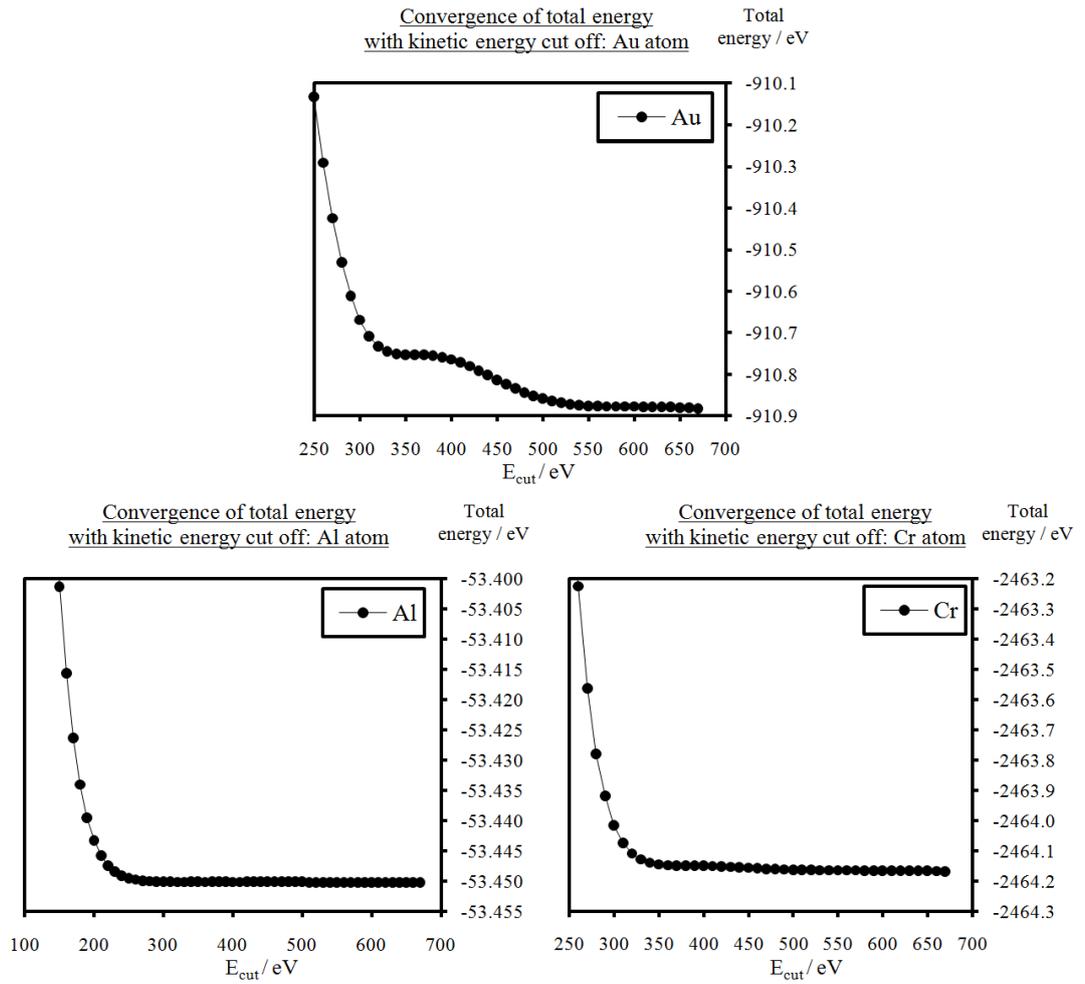


Figure 3.6. Monotonic convergence of total supercell energies with increasing kinetic energy cut off: metal atoms.

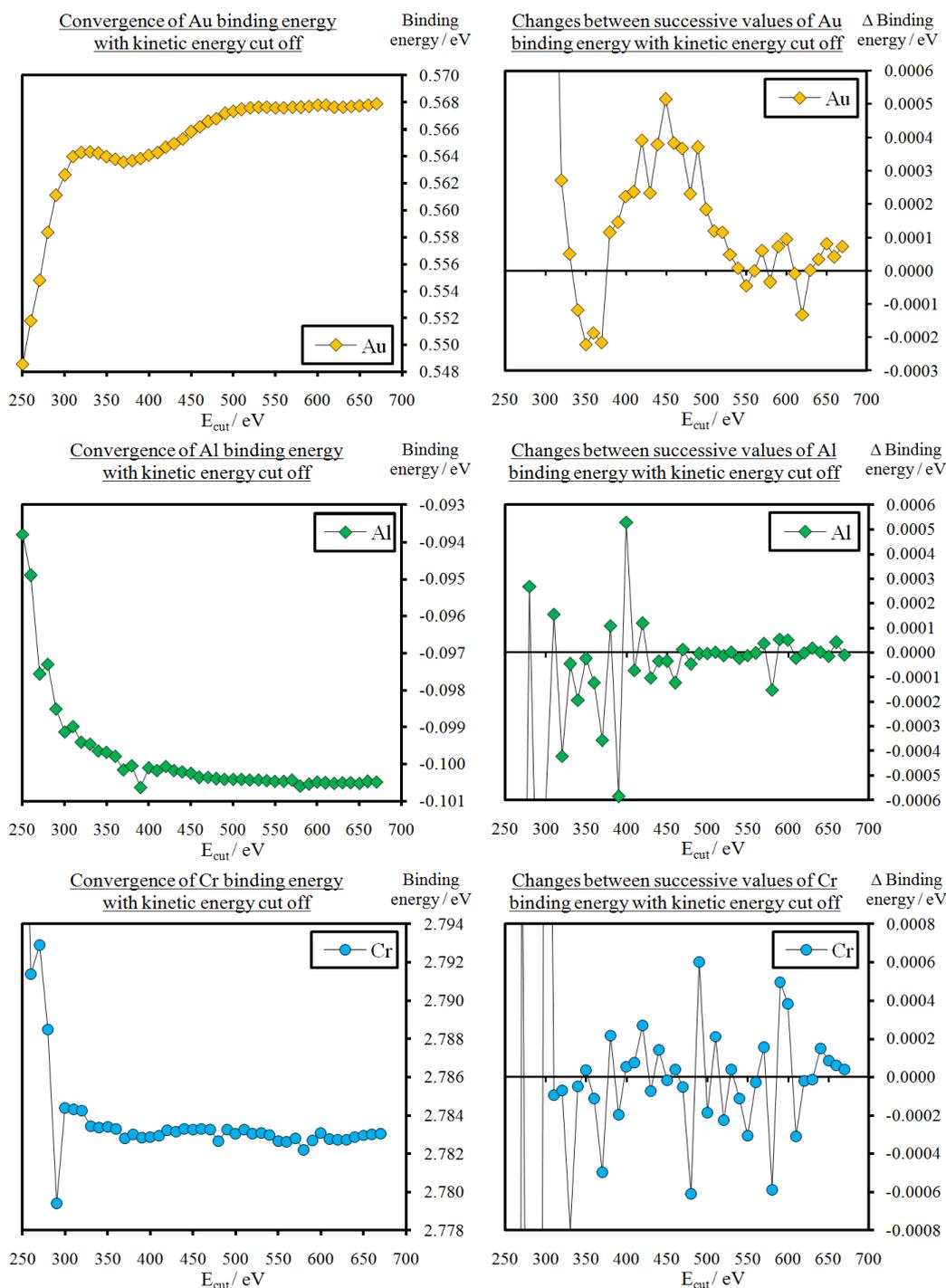


Figure 3.7. Convergence of binding energies with kinetic energy cut off.

Left panels: The binding energies clearly converge to finite values. These binding energies were calculated in initially guessed configurations without optimising the geometry, which is why some (Au and Cr) are positive, indicating energetic unfavourability. Despite the geometry not being optimised, these tests are sufficient to determine the required cut off energy. Right panels: The amounts by which the successive values on the left panels change with increasing cut off energy. Beyond $E_{cut} = 550$ eV, all residual variations are within about ± 0.0001 eV for Au and Al and ± 0.0006 eV for Cr.

Following this, the same supercells were then used to converge the total supercell energies and binding energy E_b with respect to the k-point spacings *in the graphene plane* in a similar manner to the procedure just described for E_{cut} . In-plane k-point spacings are particularly important in graphene: fine convergence is required to capture any subtle behaviour that may be occurring at the Dirac points. The results of these tests are shown in Figures 3.8, 3.9 and 3.10.

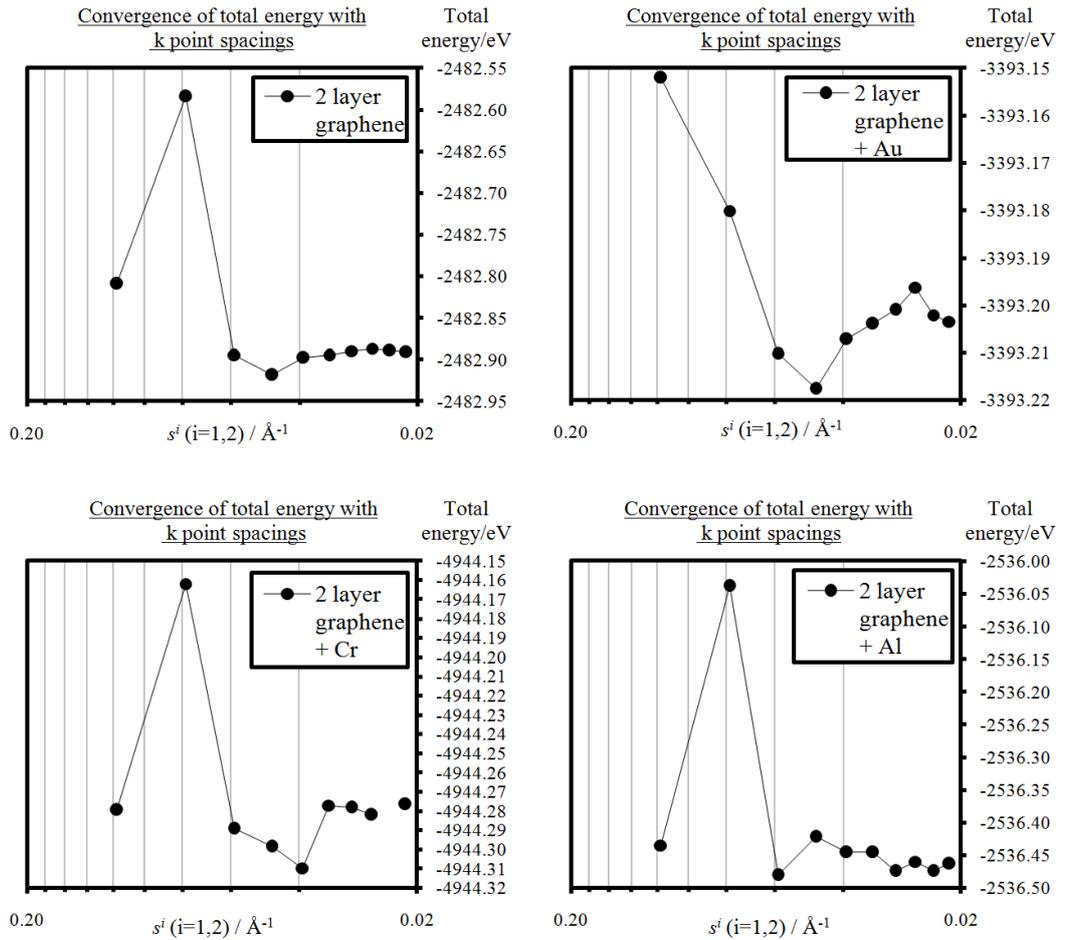


Figure 3.8. Non-monotonic convergence of total supercell energies with decreasing in-plane k point spacings: graphene and metal-graphene systems.

One of the Cr + graphene calculations could not be made to converge which is one data point is missing.

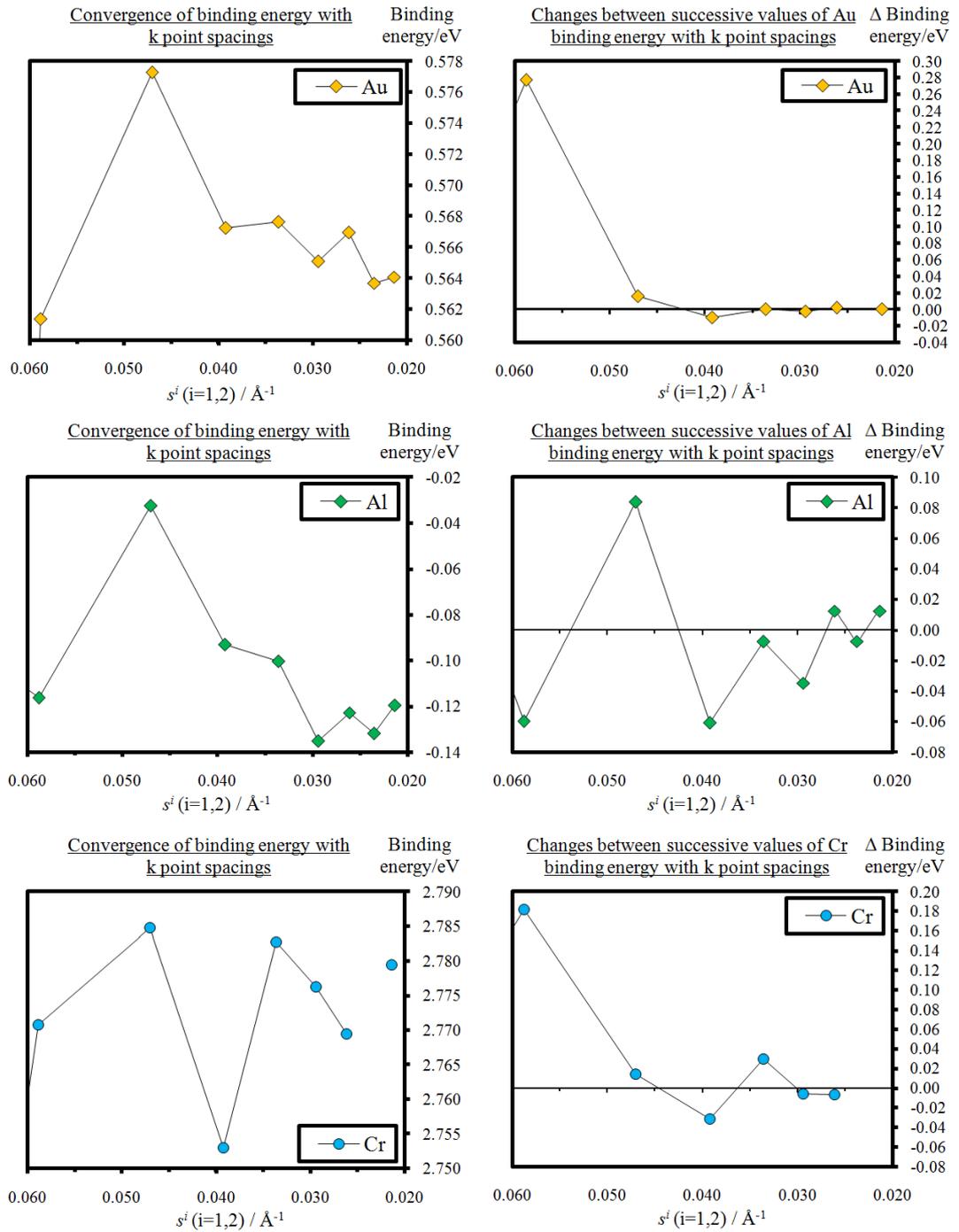


Figure 3.10. Convergence of binding energies with decreasing in-plane k point spacings.

Note that the energies converge towards similar values to those in the left panels of Figure 3.7.

The *out-of-plane* k point spacings along the reciprocal vacuum direction are conceptually slightly different to the in-plane spacings. If the vacuum lattice parameter in direct space is sufficiently large that the intercellular layer-layer interactions are zero, then the variation in the total energy with respect to increasing k points in the reciprocal vacuum direction will also be zero, meaning that a single k point in this direction is sufficient. This is because the bands along the reciprocal vacuum direction are completely flat and discrete from one another, so any extra k points along this direction will simply be sampling the same energy level. It was confirmed with some representative supercells used in this study that this is indeed true; all bands showed no variation in energy with position in reciprocal space, i.e. the bands along the relevant path of the band structure plot were completely flat. It will be stated later in this chapter that 2 k points were used along the vacuum direction for the adsorption supercells, and this is merely because the point just explained was not properly understood at the time these calculations were carried out. Of course, this has no bearing on the validity of the calculation outputs because this lack of insight amounted to nothing more than redundant oversampling. All spacings s^i were subsequently picked such that they always satisfied $s^i < 0.035 \text{ \AA}^{-1}$ for all remaining calculations. Regular Monkhorst Pack grids were used throughout.

The TS / Grimme bulk lattice parameter validation test described earlier was then repeated with this basis set and both correction schemes were verified to produce the same interlayer spacings as before. The TS correction scheme was then chosen along with the established basis set parameters $E_{cut} = 550 \text{ eV}$ and $s^i < 0.035 \text{ \AA}^{-1}$ for all subsequent calculations in the chapter.

3.3.2 Preliminary structure optimisation

Having converged the basis set parameters, the monolayer, bilayer and trilayer graphene unit cells were then fully optimised respect to bond lengths, vacuum thicknesses and layer spacings using geometry optimisation calculations. The energy of isolated metal atom supercells and the binding energy of composite graphene/metal systems were converged with increasing supercell size in order to determine the required supercell

dimensions for each system studied in this chapter. This procedure is now described.

The single layer graphene unit cell shown in Figure 3.4 (a) was constructed, whose initial in-plane lattice vectors $\vec{a}_{1l}^{(i)}$ and $\vec{b}_{1l}^{(i)}$ were left unconstrained and both set initially at the experimentally-measured⁸⁵ bulk graphite value of 2.471 Å. This unit cell was then duplicated, and the vacuum-direction lattice vector $\vec{c}_{1l}^{(i)}$ was fixed at magnitudes increasing in 1 Å increments from $|\vec{c}_{1l}^{(i)}| = 2\text{Å}, \dots, 16\text{Å}$ inclusive, to make a total of 15 unit cells. In all of these, the carbon layer was placed in the centre of the vacuum slab at fractional coordinate $0.5|\vec{c}_{1l}^{(i)}|$. A geometry optimisation calculation was carried out on each of these, in which the atomic positions and lengths $|\vec{a}_{1l}^{(i)}|$ and $|\vec{b}_{1l}^{(i)}|$ were relaxed, all unit cell angles were fixed, and $|\vec{c}_{1l}^{(i)}|$ was fixed at the value appropriate to each case. The fully-optimised TS-corrected enthalpies were plotted against the vacuum thickness $|\vec{c}_{1l}^{(i)}|$ to serve two purposes. Firstly, to identify the smallest value of $|\vec{c}_{1l}^{(i)}|$ for which the undesired inter-cellular interlayer interaction in the vacuum direction had converged to zero. This value of $|\vec{c}_{1l}^{(i)}|$ was named $|\vec{c}_{1l}|$ and identified as $|\vec{c}_{1l}| = 12\text{Å}$. Secondly, it served to calculate the carbon-carbon bond lengths as optimised using the particular choice of functional, along with the corresponding optimised values of $|\vec{a}_{1l}^{(i)}|$ and $|\vec{b}_{1l}^{(i)}|$. These values were named $|\vec{a}_{1l}|$ and $|\vec{b}_{1l}|$ respectively, and recorded at values $|\vec{a}_{1l}| = |\vec{b}_{1l}| = 2.469\text{Å}$ (4 s. f.). A similar procedure was then repeated for the case of 2 layer and 3 layer graphene on the multilayer unit cell shown in Figure 3.4 (b), in which the top carbon layer was placed at the centre of the vacuum similar to above. From similarly designed geometry optimisation calculations, the minimum required vacuum thicknesses for the 2 and 3 layer cases, $|\vec{c}_{2l}|$ and $|\vec{c}_{3l}|$, were identified as $|\vec{c}_{2l}| = 16\text{Å}$ and $|\vec{c}_{3l}| = 20\text{Å}$. The relevant data is plotted in Figure 3.11. The corresponding in-plane lattice parameters were found to be $|\vec{a}_{2l}| = |\vec{b}_{2l}| = 2.467\text{Å}$ (4 s. f.) and $|\vec{a}_{3l}| = |\vec{b}_{3l}| = 2.466\text{Å}$ (4 s. f.),

and the corresponding optimised interlayer spacings were found to be $d_{s(2l)} = 3.360 \text{ \AA}$ (4 s. f.) and $d_{s(3l)} = 3.354 \text{ \AA}$ (4 s. f.).

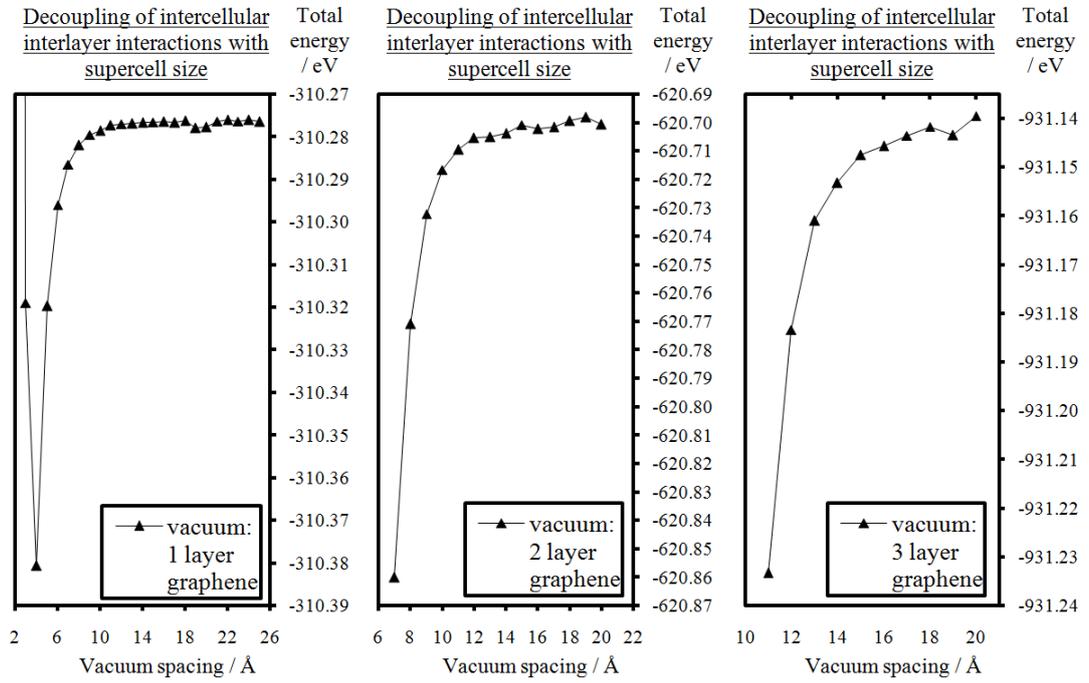


Figure 3.11. Convergence of graphene supercell energies to determine vacuum spacing required to decouple intercellular graphene-graphene interactions along the vacuum direction.

Next, vacuum-filled cubic supercells were constructed containing a metal atom placed directly in the centre as shown in Figure 3.4 (c). Each supercell had lattice parameters fixed at values of $|\vec{A}_m| = |\vec{B}_m| = |\vec{C}_m|$ with $m = Au, Al, Cr$ as appropriate, with $|\vec{A}_m|$ increasing (along with $|\vec{B}_m|$ and $|\vec{C}_m|$) in 1 \AA increments from 2 \AA to 15 \AA inclusive, to make a total of $14 \times 3 = 42$ cubic supercells. TS-corrected total energy calculations were carried out for each, and the supercell energies were converged with respect to the supercell size in order to decouple the intercellular metal-metal interactions. The minimum supercell size required to satisfy the decoupling condition all metals was identified as $|\vec{A}_m| = |\vec{B}_m| = |\vec{C}_m| = 10 \text{ \AA}$. The data from this procedure is plotted in Figure 3.12.

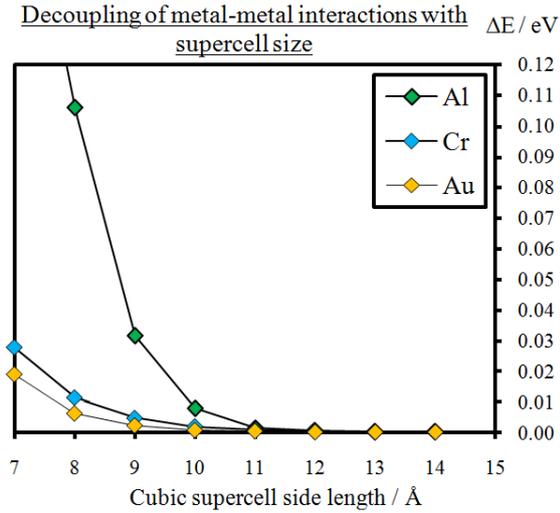


Figure 3.12. The changes in energy between successive total supercell energy calculations of a single metal atom in cubic supercells.

All metal-metal interactions are converged to within 0.01 eV by a side length of 10 Å.

The supercell shown in Figure 3.4 (d) was then constructed by forming a 2×2 array of the fully-optimised single graphene layer unit cells shown in Figure 3.4 (a). The supercell lattice vectors were fixed at values $\vec{A}_{1l,m} = 2\vec{a}_{1l}$ and $\vec{B}_{1l,m} = 2\vec{b}_{1l}$, and the vacuum-direction lattice parameter $\vec{C}_{1l,m}^{(i)}$ was initially fixed at magnitudes increasing in 1Å increments from $|\vec{C}_{1l,m}^{(i)}| = 6 \text{ \AA}, \dots, 27 \text{ \AA}$ inclusive, to make 22 supercells. Into each of these 22 supercells, a metal atom of species $m = Au, Al, Cr$ was placed 2 Å directly above the central carbon atom as indicated in Figure 3.4 (d) by the red cross, thus creating a total of $22 \times 3 = 66$ supercells. In each one of these supercells, the carbon layer was fixed at the centre of the vacuum at fractional coordinate $0.5|\vec{C}_{1l,m}^{(i)}|$, and TS-corrected total energy calculations were carried out to converge intercellular interactions in the direction $\vec{C}_{1l,m}^{(i)}$ to zero. The smallest value of $|\vec{C}_{1l,m}^{(i)}|$ for which the energy plots were deemed to have converged for all metals was named $|\vec{C}_{1l,m}|$ and identified as $|\vec{C}_{1l,m}| = 20 \text{ \AA}$. The convergence of the single layer graphene+metal energies with vacuum spacings is plotted in Figure 3.13.

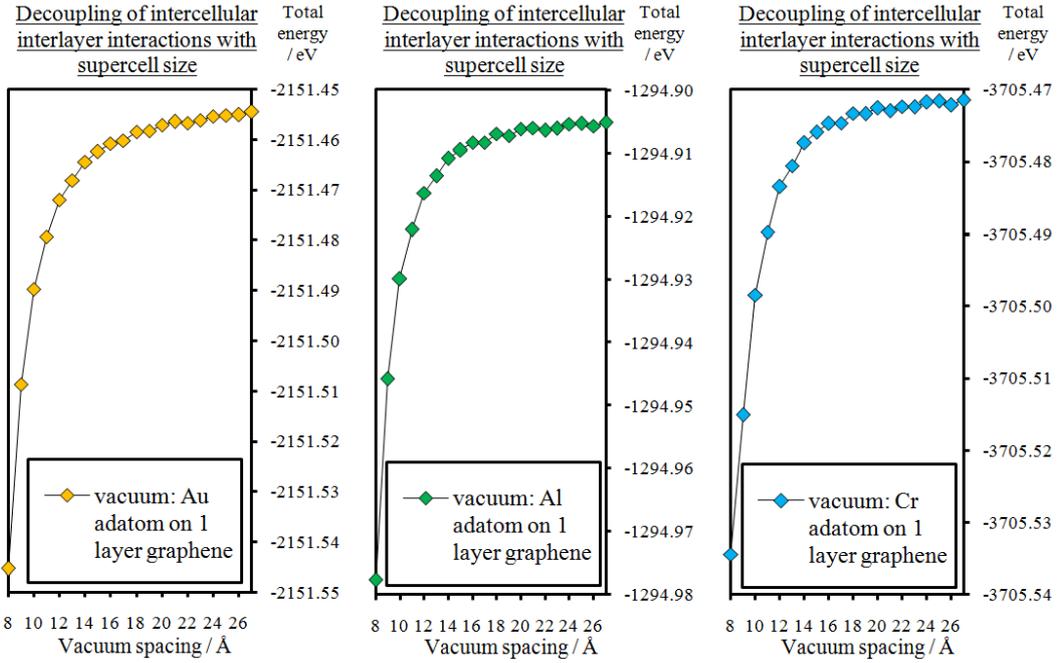


Figure 3.13. The convergence of metal adatom + single layer graphene total supercell energy with vacuum thickness.

For all three metals, the total energies change by no more than 0.005 eV upon extension of the vacuum thickness from 20 Å to 27 Å. The dipole correction scheme was not used for any of these calculations.

A similar procedure (over a shorter range of distances) was repeated for the 2 and 3 layer cases using the optimised interlayer spacings determined above. The top carbon layer was again fixed at the centre of the vacuum and the corresponding required vacuum thicknesses were found to be $|\vec{C}_{2l,m}| = 24 \text{ \AA}$ and $|\vec{C}_{3l,m}| = 27 \text{ \AA}$. The self-consistent electrostatic dipole correction scheme of Neugebauer & Scheffler⁷⁶ was used to verify that any unphysical intercellular dipole-dipole interactions were zero for the vacuum thicknesses chosen.

3 single layer supercells were constructed from 3×3 , 4×4 and 5×5 arrays of the fully relaxed unit cells from Figure 3.4 (a), whose vacuum-direction lattice vector \vec{C}_{1l} fixed at the value $\vec{C}_{1l} = 20 \text{ \AA}$ in all cases. This vacuum thickness was chosen so as to meet the requirements of the individual convergence tests just described; i.e. $|\vec{C}_{1l}| = |\vec{C}_{1l,m}| > |\vec{c}_{1l}|$. A metal atom of species m was placed 2 \AA above the central carbon atom for each of these 3 supercells, with $m = Au, Al$ and Cr as appropriate, thereby

producing a total of $3 \times 3 = 9$ supercells. TS-corrected total energy calculations were carried out for these 9 systems and the total energy was plotted against supercell size for each metal to identify the minimum size required to decouple all intercellular adatom interactions. The 5×5 supercells with lattice parameters $|\vec{A}_{1l}| = |\vec{B}_{1l}| = 5|\vec{a}_{1l}|$ were deemed sufficient for this purpose. It was also verified that this choice satisfied the intercellular metal-metal convergence requirement established above since $|\vec{A}_{1l}| > |\vec{A}_m|$. No similar test of in-plane supercell sizes for multilayer + metal systems was carried out owing to the extensive computational cost involved. Supercell sizes of 5×5 unit cells were used for 2 and 3 layer cases, with $|\vec{A}_{2l}| = |\vec{B}_{2l}| = 5|\vec{a}_{2l}|$ and $|\vec{C}_{2l}| = 24 \text{ \AA}$ for the 2 layer system and $|\vec{A}_{3l}| = |\vec{B}_{3l}| = 5|\vec{a}_{3l}|$ and $|\vec{C}_{3l}| = 27 \text{ \AA}$ for the 3 layer system, with the top carbon layer once again placed in the centre of the vacuum. Measures taken to estimate the error in the calculated binding energies due to lattice perturbations under relaxation and the limited size of the 5×5 supercells are outlined in section 3.3.3.

For the monolayer edge binding supercells, graphene edge slab/vacuum supercells akin to those shown in Figure 3.16 were prepared, in which metal atoms of species $m = Au, Al, Cr$ were placed initially at each end of the slab in the graphene plane as indicated by the red crosses at an initial distance of 2 \AA from the nearest C atom(s). For the zigzag edge slabs, a series of geometry optimisation calculations was carried out in order to converge the total relaxed TS-corrected system enthalpy with respect to the slab width $\vec{A}_{1l,zigzag}$, and the slab thickness + vacuum thickness, whose sum is denoted $\vec{B}_{1l,zigzag}$. The intercellular layer spacings were fixed at the value $|\vec{C}_{1l}| = 20 \text{ \AA}$, thus satisfying the test described earlier. The total binding energy was deemed to have converged for arrays of 5×8 of the appropriate unit cells for sites Z_1 and Z_2 . In an identical manner, an array of 5×7 appropriate unit cells for sites C_1 and C_2 was deemed sufficient. The vacuum thickness in both cases was set at 20 \AA . Recent work has suggested possible reconstruction at graphene edges,⁸⁶ however this extra detail was not deemed necessary for the primary message of this chapter. Also, under typical laboratory conditions, it is not unlikely that H passivation takes place

at the edge sites resulting in another slight discrepancy between the theoretical models and the laboratory samples. Detecting the presence of H at the edge sites using STEM is almost impossible because the scattering is too small. In order to avoid complicating matters by including H passivation into the models in the absence of any actual experimental evidence, this aspect of the system was also neglected in the analysis presented here.

3.3.3 Binding energy calculations

For the adsorption sites, the 3 monolayer graphene supercells shown in Figure 3.14 (a) - (c) were built, and a metal atom of species $m = Au, Al, Cr$ was placed into the centre of each supercell at an initial distance of 2 \AA above the graphene sheet, to make nine supercells. The positions of the carbon atoms indicated in blue were fixed along all directions, and the positions of the adatom and carbon atoms indicated in black were allowed to relax in all directions. 2 and 3 layer versions of the four types of multilayer supercells shown in Figure 3.15 (a) - (d) were then constructed in an identical fashion for each of the three metals $m = Au, Al, Cr$, to make 24 more supercells. The carbon atom positions were fixed in the multilayer cases by simply applying the reasoning used for the monolayer cases independently to each carbon layer. All lattice parameters were fixed at the values indicated in Figures 3.14 and 3.15. For the edge defect sites, 4 supercells like those shown in Figure 3.16 were used, one for each of the 4 edge sites, whose lattice parameters were all fixed, and in which all atomic positions were relaxed with the metal atom placed 2 \AA from the nearest carbon atom(s). These were duplicated into 3 copies, one set for each metal, to make 12 supercells. To curtail the risk of any of these systems failing to relax into an energetic minimum as a consequence of initial high symmetry, all systems were created with $P1$ symmetry and all symmetry finders were disabled. In addition to this, each metal atom was then offset from its initial site by 0.01 \AA in the x direction.

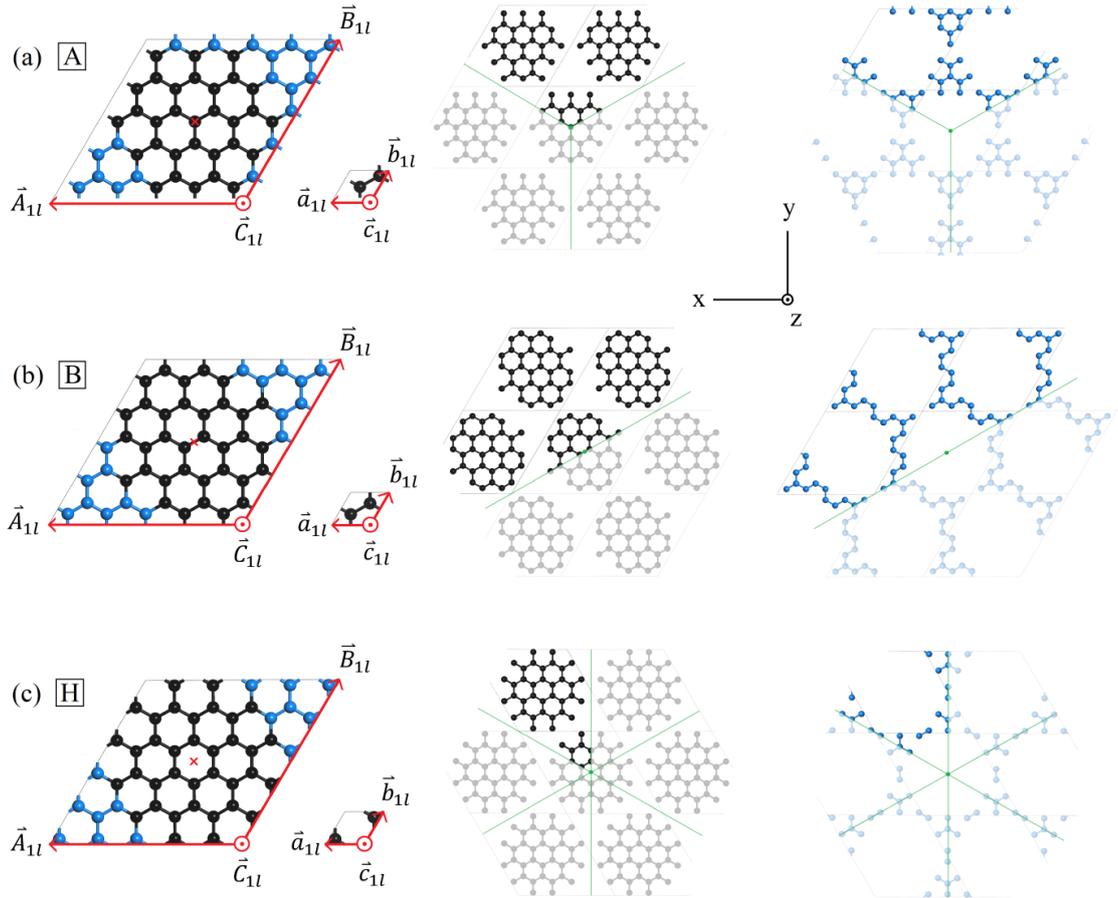


Figure 3.14. The 3 single layer supercells before geometry optimisation used for the adatom + graphene systems for (a) site A, (b) site B and (c) site H.

In all cases, carbon atoms whose positions are fixed are represented in blue and those whose positions are relaxed are represented in black. The corresponding unfixed and fixed sublattices are displayed to the right, in which the green lines show boundaries between segments of which are equivalent by virtue of rotational symmetry about the axis passing through the adsorption site represented by the green dot in the centre. The red cross denotes the initial adatom location. $|\vec{a}_{1l}| = |\vec{b}_{1l}| = 2.469 \text{ \AA}$ (3 d. p.), $|\vec{A}_{1l}| = |\vec{B}_{1l}| = 5|\vec{a}_{1l}| = 12.303 \text{ \AA}$ (3 d. p.) and $|\vec{C}_{1l}| = |\vec{c}_{1l}| = 20 \text{ \AA}$. See Figure 3.3 for nomenclature on adsorption sites.

Without actually doing further calculations, it is not easy to estimate how much the binding energies of these configurations may change if the other constraint schemes^{38,41,42,46,53} were used instead; the essential requirement in all cases is that the supercells are sufficiently large to account for lattice perturbations to a reasonable degree of accuracy. The constraint scheme presented here finds its merit in the sense that it removes the artificial symmetry-breaking influence of periodic boundary conditions, allowing for

the shallow energetic minima associated with weakly-bound adatoms to be reliably identified.

The plane wave DFT code CASTEP⁸³ was then used with the TS van der Waals correction implementation⁸⁴ and Vanderbilt ultrasoft pseudopotentials²⁰ to carry out spin-polarized geometry optimisation calculations for each of these 45 supercells. Valence states incorporated were $2s^2 2p^2$ for C, $5d^{10} 6s^1$ for Au, $3s^2 3p^6 3d^5 4s^1$ for Cr and $3s^2 3p^1$ for Al. To satisfy the k-points spacings convergence criterion, $s^i < 0.035 \text{ \AA}^{-1}$, determined in section 3.3.1., a regular and uniformly-weighted $3 \times 3 \times 2$ Monkhorst-Pack¹⁷ grid of 9 k-points was used to sample the Brillouin zone for the migration supercells illustrated in Figure 3.17 and the adsorption supercells in Figures 3.14 and 3.15. For the edge binding supercells illustrated in Figure 3.16, a $1 \times 2 \times 2$ k points grid was used. For each series of self-consistent field (SCF) cycles used for the electronic minimisation, the exit criterion was imposed that the change in total electron energy between successive SCF cycles be converged to within $5 \times 10^{-7} \text{ eV}$. For the geometry optimisation, the Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimisation algorithm⁸⁷⁻⁹¹ was used with the following three convergence criteria: i) that the maximum force on all atoms be less than 0.01 eV \AA^{-1} , ii) that the maximum change in position for all atoms between successive BFGS steps be less than $5 \times 10^{-4} \text{ \AA}$ and iii) that the maximum change in the total system enthalpy between successive BFGS steps be less than $5 \times 10^{-6} \text{ eV}$ per atom. The final TS-corrected enthalpies of these relaxed structures were then recorded as the values of $E_{C+M}(X)$ for insertion into equation (3.1).

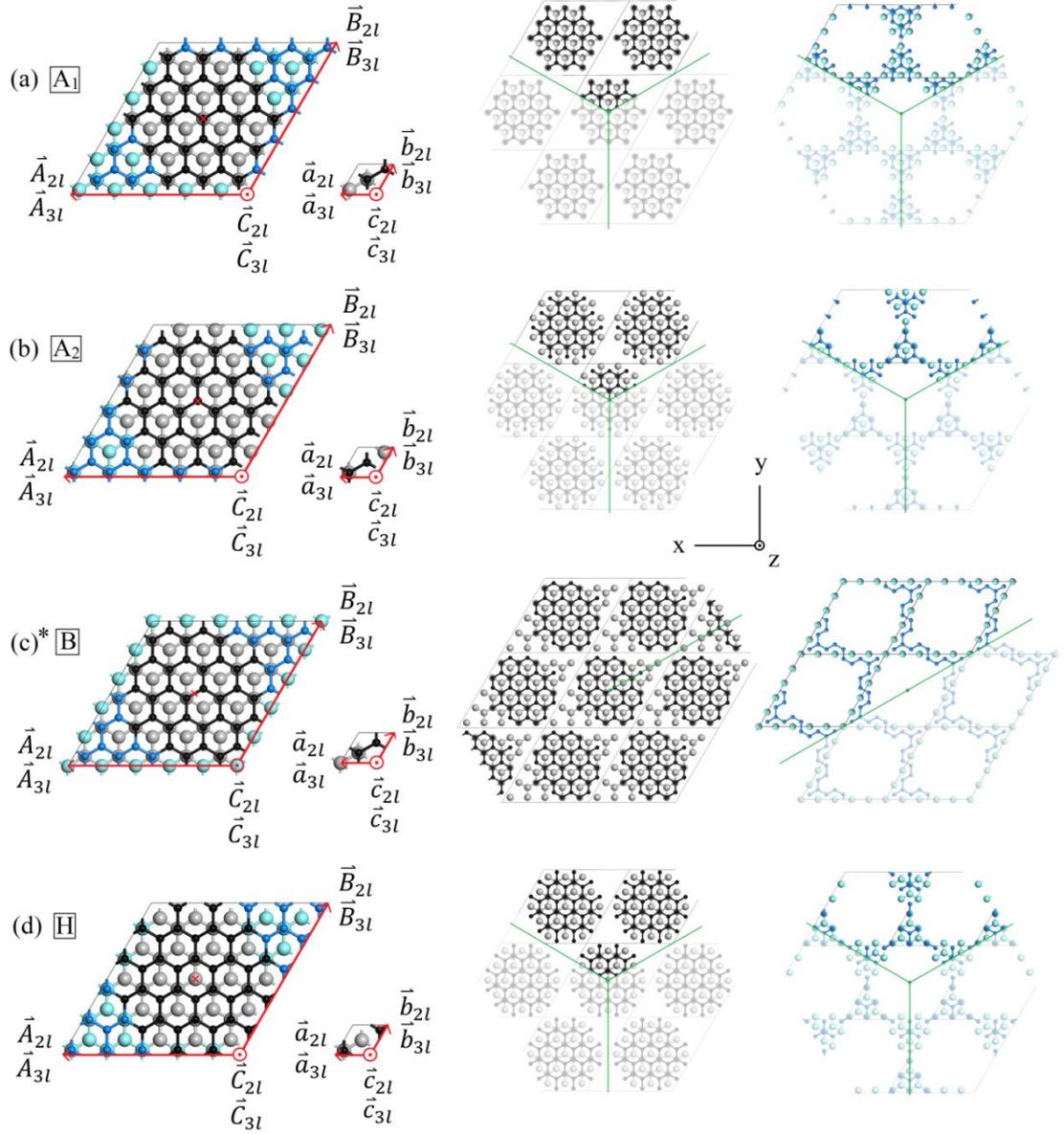


Figure 3.15. The multilayer input supercells for (a) site A_1 , (b) site A_2 (c) site B , and (d) site H .

The top carbon layer is represented by small balls and sticks, and the first sublayer is represented by large balls and sticks. No further sublayers are indicated owing to the assumed AB stacking structure. Fixed top layer and sublayer C atoms are blue and light green respectively. Unfixed top layer and sublayer C atoms are coloured black and grey respectively. As in Figure 3.14, the red cross denotes the initial adatom location. As in Figure 3.14, the unfixed and fixed sublattices are shown below their corresponding supercell, divided into segments which are equivalent by rotational symmetry about the axis passing through the adsorption site. $|\vec{a}_{2l}| = |\vec{b}_{2l}| = 2.467 \text{ \AA}$ (3 d. p.), $|\vec{A}_{2l}| = |\vec{B}_{2l}| = 5|\vec{a}_{2l}| = 12.293 \text{ \AA}$ (3 d. p.), $|\vec{C}_{2l}| = |\vec{c}_{2l}| = 24 \text{ \AA}$, $|\vec{a}_{3l}| = |\vec{b}_{3l}| = 2.466 \text{ \AA}$ (3 d. p.), $|\vec{A}_{3l}| = |\vec{B}_{3l}| = 5|\vec{a}_{3l}| = 12.290 \text{ \AA}$ (3 d. p.) and $|\vec{C}_{3l}| = |\vec{c}_{3l}| = 27 \text{ \AA}$.

*The B site in (c) is the only site for which the rotational symmetry of the first sublayer (and also therefore the complete lattice) is 1-fold. For this case, C positions were fixed on the supercell perimeter on the first sublayer in preference to some other arbitrary selection resulting in 1-fold symmetry, resulting in the 2-fold fixed sublattice rotational symmetry shown. See Figure 3.3 for nomenclature on adsorption sites.

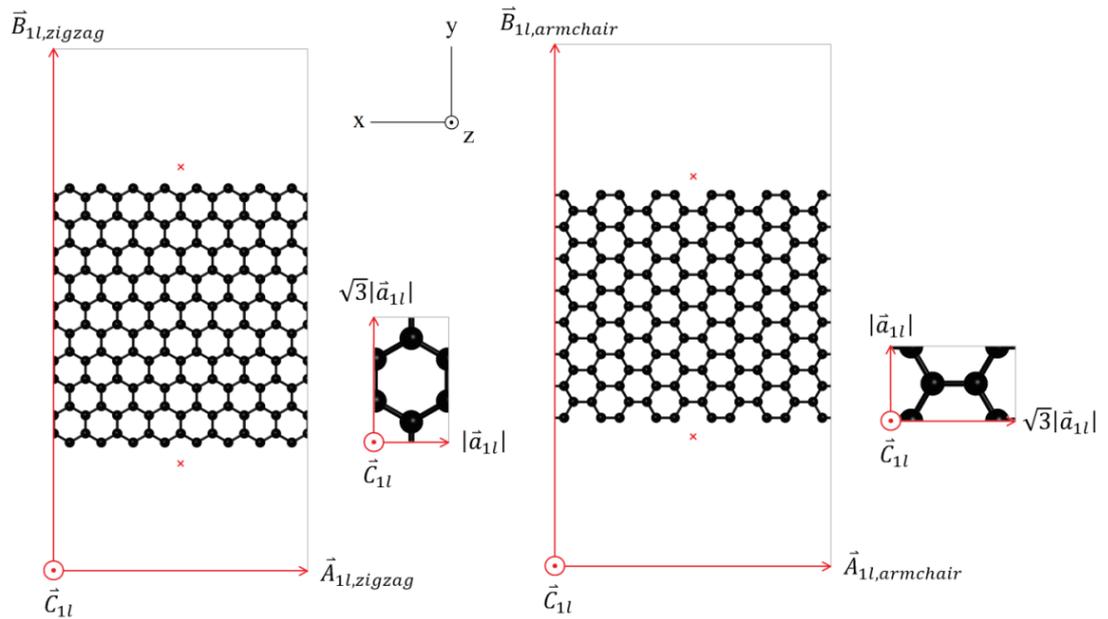


Figure 3.16. Example supercells used for metal binding to monolayer edges with accompanying unit cells.

$$|\vec{A}_{11,zigzag}| = 8|\vec{a}_{11}|, \quad |\vec{B}_{11,zigzag}| = 5\sqrt{3}|\vec{a}_{11}| + 20 \text{ \AA}, \quad |\vec{A}_{11,armchair}| = 5\sqrt{3}|\vec{a}_{11}| \text{ and } |\vec{B}_{11,armchair}| = 7|\vec{a}_{11}| + 20 \text{ \AA}.$$

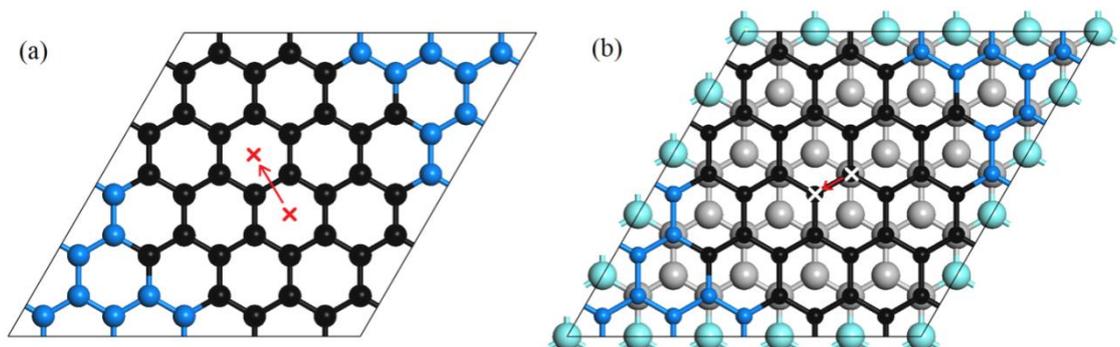


Figure 3.17. Example supercells used for migration activation barrier calculations.

(a) $H \rightarrow B \rightarrow H$ trajectory used for Cr and Al on the monolayer (b) $A_1 \rightarrow B \rightarrow A_2$ trajectory used for Au.

The values E_M and E_C were then calculated for insertion into equation (3.1), and the following measures were taken to exploit k-point error cancellation. Firstly, to evaluate the quantities E_M , the 45 final relaxed structure files were duplicated, and the copies were imported back into the visualisation software. All of the carbon atoms were then deleted, leaving just the metal atom(s) left in its final position in each case, and a spin-polarized TS-corrected total energy calculation was then performed for each of these 45 isolated metal atoms to evaluate the quantity E_M for each supercell separately. To evaluate the quantities E_C in equation (3.1) for the adsorption energies, the spin-polarized TS-corrected isolated graphene energies were calculated using the initial input supercells (i.e. pre-geometry optimisation), from which the metal atom was deleted in each case. For the edge site binding energies, the isolated graphene edge structures were fully relaxed to obtain spin-polarized TS-corrected enthalpy values E_C . No reconstructions of any graphene edges were observed, although this is unsurprising because the energy barrier that needs to be overcome for such a reconstruction to occur is likely to be missed by geometry optimisation algorithms, which, by design, search downhill in the energy landscape. These values were substituted into equation (3.1) to give the relaxed structure binding energies $E_b(X)$, which are plotted in Figure 3.18.

In order to estimate the error in the final values of adsorption binding energies owing to adatom-induced lattice perturbations, the binding energies of three fully relaxed 8×8 supercells were calculated; one for each metal. The binding energies for Al and Cr agreed with those of the 5×5 supercells up to a maximum discrepancy of 0.018 eV and 0.024 eV respectively, with a slightly larger maximum discrepancy of 0.056 eV recorded for the case of Au. In common with the 5×5 supercells, the puckering of the graphene lattice near the adatom was very small in all cases. Besides a weak positive (and probably incidental) correlation of these binding energy discrepancies with the adatoms' atomic masses, there is no obvious way to rationalise these discrepancies without carrying out a more exhaustive range of energy vs. supercell size calculations. Nevertheless, these tests confirmed that whilst the calculated binding energies were likely to be somewhat underestimated owing to the limited supercell sizes employed in this study,

the lattice perturbations for the adatoms studied were not significant enough to have an overriding influence on the main conclusions.

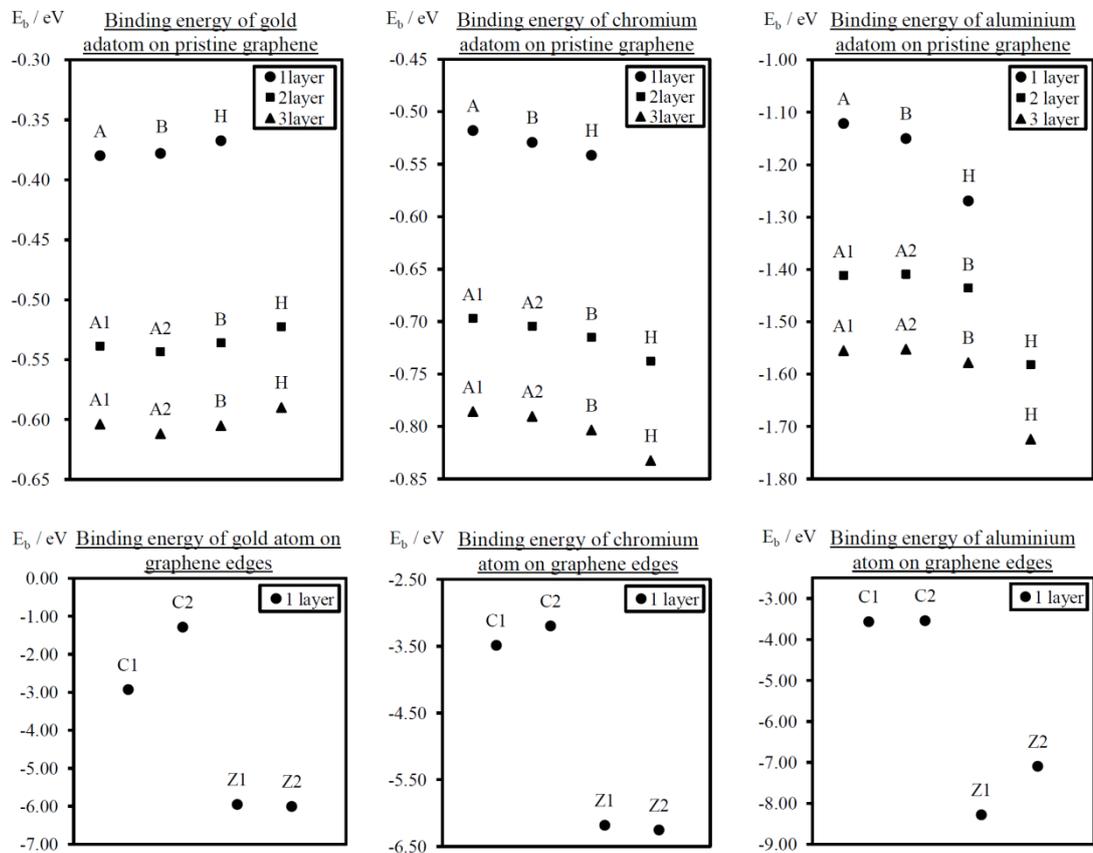


Figure 3.18. The calculated binding energy for metal atoms adsorbed on the pristine substrates and bound at monolayer edge sites.

The energetic ordering of the adsorption sites is seen to remain the same for increasing thicknesses. See Figure 3.3 for nomenclature of binding sites.

3.3.4 Migration energy calculations

Using the adsorption binding energy results of the next section, the migration pathways $H \rightarrow B \rightarrow H$ on all substrates were identified as obvious candidates for initial guesses of the lowest energy adatom diffusion pathways for Cr and Al, along with the paths $A \rightarrow B \rightarrow A$ ($A_1 \rightarrow B \rightarrow A_2$) on the monolayer (multilayer) substrates for Au. These paths were nominated because they comprise sites which give the lowest combination of binding energies which can be joined by a path traversing the entire unit cell. The established linear/quadratic synchronous transit (LST/QST) scheme of

Halgren and Lipscomb⁹² for determining reaction pathways, as modified to include conjugate gradient refinements and generalised to include periodic systems by Govind *et al.*⁹³ and implemented in CASTEP was used to locate the transition state configurations and thus evaluate the adatom migration activation barriers associated with these paths. The reactant and product states were first obtained by carrying out ultrafine geometry optimisation calculations with the pertinent adatoms at the path endpoints as indicated for the supercells illustrated in Figure 3.17. The reaction trajectory joining these reactant and product states was initially guessed by using the LST interatomic distance interpolation scheme,⁹² and the midpoint of this trajectory was used as the intermediate state to define the initial three-point QST pathway. A series of conjugate gradient minimisations and QST cycles were carried out from this point to locate the energy saddle point until the root mean square (RMS) of all atomic forces were converged to within 0.05 eV\AA^{-1} . All other calculation input parameters were the same as in section 3.3.2. The resulting energy barriers are tabulated in Table 3.3.

3.4 RESULTS AND DISCUSSION

3.4.1 Binding energies and bond distances

All metal adatoms and edge atoms settled onto the sites they were initially placed into, confirming that local energy minima exist for all configurations studied. The calculated binding energies E_b corresponding to the fully optimised configurations are tabulated in Table 3.2 and plotted in Figure 3.18. Figure 3.19 shows total electron density slices for Au adsorbed onto monolayer and trilayer graphene, and bound to the monolayer edge sites C_1 , C_2 , Z_1 and Z_2 . For all adsorption states for all metals, structural perturbations to the graphene lattice were small, as Figure 3.19 shows for the case of Au.

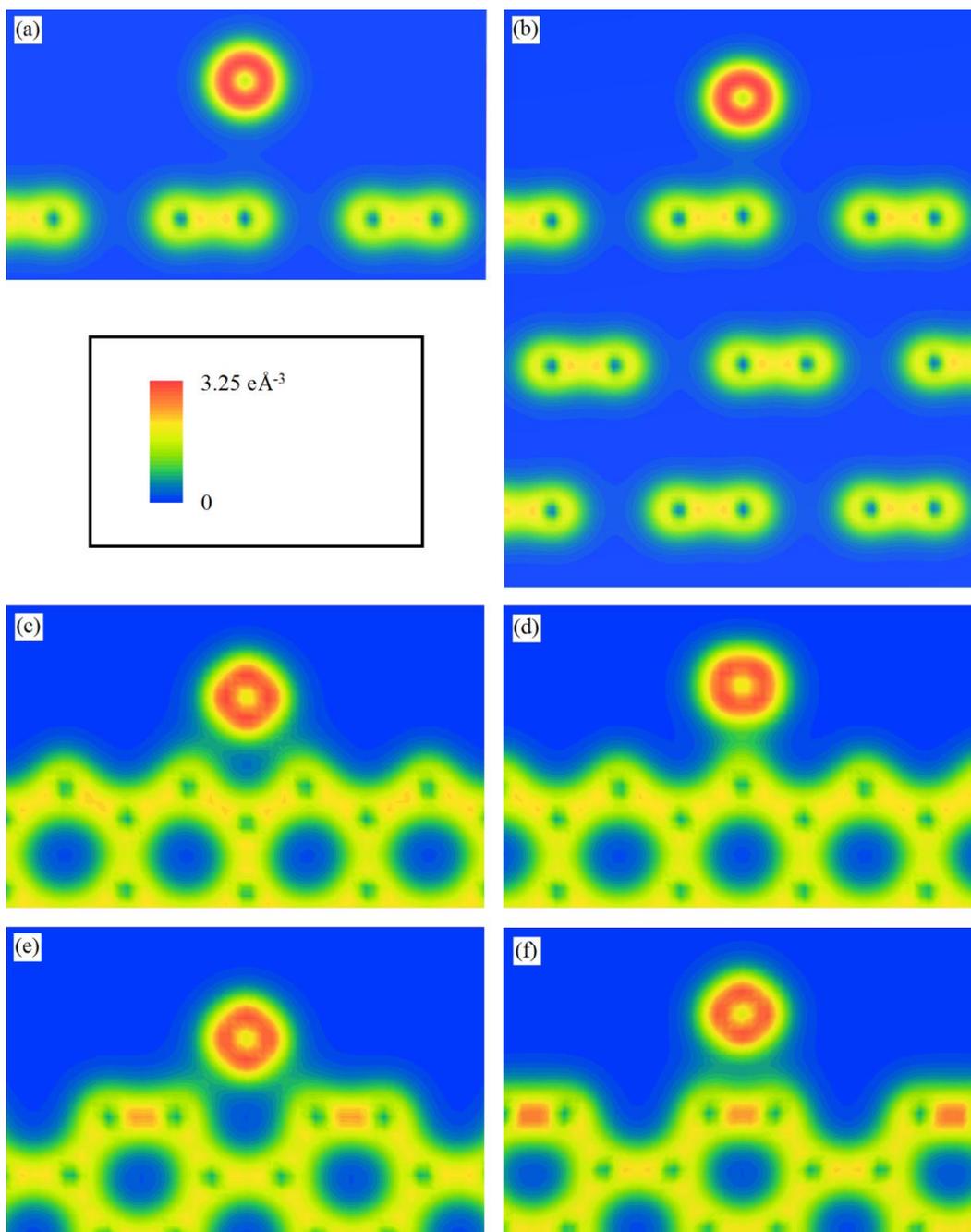


Figure 3.19. *Electron density images showing the difference in bonding character between adsorption and edge sites for Au.*

(a) Cross section of the total electron density field shown in colour units of electrons / \AA^3 for Au at adsorption site A for the fully relaxed monolayer. (b) Corresponding trilayer image, showing Au at site A_1 . The cross sections shown intersect the graphene along the “armchair” direction, thus showing the carbon-carbon bonds for comparison. The bonding character is seen to be consistent with physisorption in both cases, though a slightly more substantial bond is evident for the trilayer case. Au atom binding to the edge sites (c) Z_1 (d) Z_2 (e) C_1 and (f) C_2 . Clear and substantial regions of electron density are observable in all four cases, consistent with a covalent metal carbide bond. See Figure 3.3 for nomenclature of binding sites.

The adsorption bonding character is seen to be consistent with physisorption. The increased adsorption energy for the trilayer case is evident from the smaller Au-graphene surface distance and the increased electron density between the Au and the binding carbon atom. In contrast, the binding energies at monolayer edge sites are much higher than the adsorption energies in all cases, confirming that these defect sites are much more stable, consistent with collaborators' STEM observations. These results are clearly supported by Figure 3.19, which shows a substantial region of electron density in between the C and Au at the edge sites, which warrants interpretation as a substantially stronger covalent chemical bond.

In Table 3.2, three distances are tabulated for the adsorption calculations: i) the distance(s) along the binding direction(s) from the adatom to the nearest carbon atoms (1 for sites A , A_1 and A_2 , 2 for site B and 6 for site H) ii) the distance along the z direction from the adatom to the nearest carbon atom(s) and iii) the distance along the z axis from the adatom to the fixed carbon atoms of the top layer. The difference between the latter two of these three distances is equal to the amount by which the top graphene layer had puckered out-of-plane. These puckering distances are all small, indicating that all metal adatoms do little to interfere with the structural integrity of the lattice. The binding energies are seen to significantly increase for increasing layer numbers for all of the metals tested, adding credibility to the notion that the van der Waals interaction with the sublayers accounts for a significant proportion of the metal-graphene binding energy in real laboratory samples.

For each and every adsorption site and metal studied, the energy difference between the 2 and 3 layer cases is smaller than the difference between the 1 and 2 layer cases. This certainly seems like an intuitive result and it suggests that the binding energy converges towards that of the bulk graphite (0001) surface as the thickness is increased beyond 3 layers. Further calculations for higher numbers of graphene layers could be carried out to predict the thickness required to recover the behaviour of the bulk graphite (0001) surface, although it may be wise to resort to using symmetry finders to make such calculations computationally efficient, depending on the scaling behaviour of the code used.

Table 3.2. (Ad)atom binding energies metal-carbon distances associated with the fully relaxed structures.

The differences between values in the two rightmost columns indicate the graphene lattice puckering distance in each adsorption case. Distances which are identical by definition are highlighted in grey pairs. See Figure 3.3 for nomenclature of binding sites.

Metal / graphene system	Site	Binding energy, E_b / eV (3 d.p.)	Distance from metal (ad)atom to nearest carbon atom(s) along bond direction(s) / Å (3 d. p.)	Distance along z axis from metal adatom to nearest carbon atom(s) / Å (3 d. p.)	Distance along z axis from metal adatom to fixed top layer carbon atoms / Å (3 d. p.)
Au					
1 layer adsorption sites	A	-0.380	3.082	3.082	3.095
	B	-0.378	3.291	3.215	3.217
	H	-0.367	3.700	3.421	3.408
2 layer adsorption sites	A ₁	-0.539	3.008	3.008	3.025
	A ₂	-0.543	2.671	2.671	2.741
	B	-0.536	3.283	3.207	3.209
3 layer adsorption sites	H	-0.522	3.665	3.383	3.379
	A ₁	-0.604	2.739	2.739	2.817
	A ₂	-0.612	2.712	2.712	2.775
1 layer edge sites	B	-0.605	3.233	3.156	3.159
	H	-0.590	3.650	3.366	3.360
	C ₁	-2.937	2.125	-	-
1 layer edge sites	C ₂	-1.284	2.154	-	-
	Z ₁	-5.950	2.171	-	-
	Z ₂	-6.003	2.004	-	-
Cr					
1 layer adsorption sites	A	-0.518	2.309	2.309	2.332
	B	-0.529	2.369	2.260	2.274
	H	-0.542	2.509	2.056	2.063
2 layer adsorption sites	A ₁	-0.697	2.305	2.305	2.318
	A ₂	-0.704	2.298	2.298	2.309
	B	-0.715	2.366	2.257	2.264
3 layer adsorption sites	H	-0.738	2.503	2.049	2.048
	A ₁	-0.786	2.294	2.294	2.310
	A ₂	-0.790	2.289	2.289	2.301
1 layer edge sites	B	-0.803	2.354	2.241	2.253
	H	-0.832	2.483	2.025	2.020
	C ₁	-3.485	2.036	-	-
1 layer edge sites	C ₂	-3.090	1.942	-	-
	Z ₁	-6.181	1.827	-	-
	Z ₂	-6.252	1.893	-	-
Al					
1 layer adsorption sites	A	-1.121	2.287	2.287	2.251
	B	-1.150	2.357	2.246	2.245
	H	-1.269	2.573	2.127	2.135
2 layer adsorption sites	A ₁	-1.411	2.280	2.280	2.230
	A ₂	-1.409	2.282	2.282	2.240
	B	-1.435	2.359	2.249	2.225
3 layer adsorption sites	H	-1.582	2.578	2.135	2.133
	A ₁	-1.555	2.281	2.281	2.232
	A ₂	-1.552	2.282	2.282	2.222
1 layer edge sites	B	-1.578	2.359	2.250	2.226
	H	-1.724	2.580	2.137	2.133
	C ₁	-3.564	2.002	-	-
1 layer edge sites	C ₂	-3.539	1.892	-	-
	Z ₁	-8.280	1.976	-	-
	Z ₂	-7.095	1.935	-	-

A further important conclusion of the results is that the absolute difference in binding energy between the A₁, A₂, B and H adsorption sites remains unchanged for the various studied thicknesses. This is evident from the energy trends in Figure 3.18, in which it can be seen that the 2 layer and 3

layer data points for a given metal are all approximately related by a rigid translation along the energy axis. So despite the proportional contribution of the sublayers to the total binding energy being very significant, the energetic ordering of the adsorption sites is actually predicted to be independent of the thickness. This indicates that at absolute zero, the short-range carbon-metal binding occurring on the top layer takes precedence over the van der Waals contribution from the sublayers, giving rise to static configurations for which the energetically favoured adsorption site is unchanged, irrespective of how thick the sample is.

3.4.2 Mobility of metal adatoms

Despite the prediction that stable configurations exist for all metals and sites at 0 K , this is in contrast to the published STEM observations^{60,67-69} at room temperature, $T \sim 300\text{ K}$. It is now argued that the migration activation barrier calculations summarised in Table 3.3 strongly support the notion that thermal effects cause the adatoms to be mobile along in-plane directions. Room temperature corresponds to a thermal energy of $k_B T = 0.026\text{ eV}$ (3 d. p.), where k_B is the Boltzmann constant. The equipartition theorem of statistical mechanics states that for every spatial degree of freedom possessed by a particle, there exists a contribution of $k_B T/2$ to its average thermal energy. The metal atoms can be considered to be confined to the 2D graphene lattice, so they therefore have two degrees of freedom and a corresponding average thermal energy of $k_B T$. The magnitude of the metal-carbon binding energy in all cases is much larger than $k_B T$ at room temperature, so an argument based on thermal bond breaking cannot be invoked to account for the continual absence of adatoms on clean regions. However, the migration activation barriers presented in Table 3.3 for the case of monolayer substrates at 0 K are well below $k_B T$ for Au and Cr at $T \sim 300\text{ K}$, indicating that these adatoms are likely to be extremely mobile on all the substrates studied at room temperature. For Al, the activation barriers are between 0.166 and 0.197 eV (3 d. p.), within one order of magnitude of $k_B T$. This suggests that Al adatoms are likely to migrate at a slower rate than Au and Cr, although the barrier is nonetheless trivially small. The results

predict that lattice edge sites would result in significantly stronger adatom binding than the pristine, clean regions of the lattice. This is demonstrated very clearly by recent collaborators' STEM observations presented in Figure 3.1. It is noteworthy that one recent study⁹⁴ used DFT to predict binding energies of Au atoms at different types of edge site to be between 3.1 eV and 6.4 eV, in good agreement with the values here.

Table 3.3. *Calculated migration barriers for Au, Cr and Al on the lowest energy migration pathways on pristine monolayer, bilayer and trilayer graphene.*

Adatom	Substrate	Path	Migration barrier ΔE / eV (3.d.p.)
Au	1 layer	A → B → A	0.007
	2 layer	A ₁ → B → A ₂	0.008
		A ₂ → B → A ₁	0.024
	3 layer	A ₁ → B → A ₂	0.019
		A ₂ → B → A ₁	0.025
Cr	1 layer	H → B → H	0.022
	2 layer	H → B → H	0.021
	3 layer	H → B → H	0.022
Al	1 layer	H → B → H	0.166
	2 layer	H → B → H	0.178
	3 layer	H → B → H	0.197

Thus, it is concluded that all adatoms in the samples in Figure 3.1 had migrated across the clean regions of the lattice into more stable defective or contaminated regions within a short timeframe as a result of the statistical inevitability associated with perturbing thermal effects at $T \sim 300$ K. This migration occurred because of the small adatom migration barriers for samples of all thicknesses.

3.4.3 Possible effects of the electron beam

The effect of the STEM electron beam on the experimental specimens also needs to be considered. A very small number of Au atoms were observed by STEM on clean regions of four or more layers thickness some time after deposition of the adatoms. Whilst the published STEM images confirm that the beam does little or nothing to affect the integrity of the graphene itself at 60 keV, it is necessary to rationalise the circumstances whereby single metals were observed on clean regions of graphene. In most instances, regardless of specimen thickness and the metal species, the adatoms exist

as clusters at defected or hydrocarbon contaminated regions of the sample. This can be understood by the high mobility of the adatoms, and the greater energetic stability of adatom binding at these sites. It is conjectured that in cases where Au atoms were observed by STEM on clean thicker regions⁶⁰, the beam may have displaced these Au atoms from clusters in the more stable regions during the scanning process. Considerations based on a recent quantitative study of beam damage in graphene⁹⁵ could be used to test these remarks. This is not to suggest that the possibility of knock-on damage in the microscopy experiments, local heating effects (which are arguably negligible⁹⁶) or the temporary localised accumulation of negative charge around the beam are not also recognised. These effects were not however explicitly considered in this study. In summary, it is considered that migration effects are the pivotal reason why adatoms are not generally observed on clean graphene regions, with electron beam knock-on effects being a secondary consideration. Indeed, the beam itself is attributed as a possible reason for the observation of single Au atoms on clean regions due to displacement effects.

3.5 CONCLUSIONS

DFT calculations have been presented for the binding energy of Au, Al and Cr atoms bound at graphene edge sites and adsorbed on monolayer, bilayer and trilayer graphene using the van der Waals-correction scheme of Tkatchenko and Scheffler²⁸ for the first time. The contribution to the total binding energy from graphene sublayers was predicted to be very significant, although the edge binding energies were found to be substantially higher for all atoms in all cases. Migration activation barriers for these adatoms on monolayer, bilayer and trilayer graphene were then calculated and shown to be smaller than or within one order of magnitude of $k_B T$ at room temperature in all cases, implying that these adatoms are extremely mobile on graphene at room temperature. It was concluded from this that graphene samples doped with Au, Cr and Al should be seen to be completely devoid of these dopants on the pristine regions, with the dopants binding preferentially to the edge defect sites. This was shown to be in striking agreement with the STEM data presented in Figure 3.1 of this study,

along with recently published STEM observations^{60,67-69}. Additionally, a brief review of atomic position fixing conventions adopted in recently published calculations was presented, and a simple guiding principle based on lattice symmetries was suggested for future studies. The case has been made that this constraint scheme is arguably an improvement over published constraint schemes on the grounds that it eliminates the problem of symmetry-breaking, allowing for local energetic minima - often very shallow for adatoms on substrates - to be unambiguously identified.

Chapter 4: Steps on the TiO₂ rutile (110) surface

"God made the bulk, but surfaces were invented by the Devil."

Wolfgang Pauli

Surface science is a persistently important area of materials science, and defects are an inevitable component of any material surface. In particular, step defects comprise a significant proportion of many solid material surfaces. The atomic arrangements and free energies of steps can significantly influence the overall surface chemistry of a material, so gaining a proper understanding of steps is an important part of surface characterization. The energy of a step can be calculated within a density functional theory (DFT) periodic slab/vacuum supercell framework by calculating the total surface energy of a high-index slab which simulates terraces bound by the steps of interest, and then extracting the component of this total high-index surface energy which is solely attributable to the step. However, for certain materials, such as TiO₂ rutile, there is a severe difficulty associated with this approach. For TiO₂ rutile, it is well known that surface energies calculated using DFT slab models converge in a very slow and oscillatory manner with increasing slab thickness. This is really just a consequence of the fact that slabs of computationally practical size are effectively thin films, and are therefore afflicted with slowly decaying surface-surface interactions. This oscillation occurs because the electron orbitals in the ground state hybridize differently depending on whether the slab in question has an even number or an odd number of oxide layers in it. This makes it very difficult to reliably extract the step energy from a high-index slab because the step energy is very small compared to the total surface energy of the slab, and yet this surface energy depends very sensitively on the slab chemical formula. A new and improved method is therefore called for, and this chapter presents the development and application of such a

method to the (110) surface of TiO₂ rutile in detail. In principle, it can be applied to any crystalline surface.

4.1 Abstract

In this chapter, a systematic new approach has been developed to reliably calculate the energies of step defects and step-step interaction energies on general crystalline solid surfaces. It has been argued that existing and widely published statistical approaches to evaluating surface energies based on linear regression can be dangerously misleading for some materials and should generally be avoided in favour of the new method presented here. The method has been applied to the TiO₂ rutile (110) surface and the calculated step energies have been used to predict island shapes which show strong agreement with islands observed in published scanning tunneling microscopy (STM) data.

4.2 INTRODUCTION

4.2.1 The importance of steps on TiO₂

The surface science of TiO₂ is of persistent and widespread interest in materials science because of the many industrial applications which exploit its properties as a photocatalyst⁹⁷, including bactericidal coatings⁹⁸⁻¹⁰⁰, air and water purification^{101,102} and photo-induced superhydrophilic surfaces.¹⁰³⁻¹⁰⁵ Alongside technological motivations, TiO₂ surfaces are of interest in the geosciences because step defects are thought to act as bonding sites for organic molecules in aqueous solutions providing a possible basis for the origins of primordial life.¹⁰⁶⁻¹¹¹ The (110) surface of the rutile phase is particularly important because rutile is the most abundant macroscopic phase of TiO₂ and the (110) surface is the most stable surface of this phase,^{112,113} comprising the majority of the total surface area in laboratory samples. STM studies¹¹⁴⁻¹¹⁸ have emerged since the early 1990's which invariably show that extended step defects comprise a very significant proportion of the overall surface area of rutile (110). Thus, a proper understanding of steps on the atomic level is crucial in moving towards a complete characterisation of this extremely important system, which has in

many ways become the archetypal representative for transition metal oxide surfaces. The surface of metal oxides is where many of the physical and chemical processes of interest to existing and emerging technologies takes place, and the specific atomic configurations of step structures can have a very significant effect on the surface chemistry of metal oxides and their affinity for bonding with adsorbates.¹¹² This is especially true for crystalline nanoparticles on which steps are likely to dominate the surface morphology. Theoretical studies of the energy and structure of steps therefore present opportunities to not only optimise existing metal oxide surface processes but also to predict surface-adsorbate bonding processes to be exploited for new technologies.

4.2.2 Existing studies of steps on TiO₂

In published STM studies¹¹⁴⁻¹¹⁸ of vacuum-annealed rutile (110) surfaces, terraces bound by extended step defects running almost exclusively along the $\langle 1\bar{1}1 \rangle$ and $\langle 001 \rangle$ directions are observable, where the surface roughness is seen to decrease with higher annealing temperatures. Among these studies, Onishi *et al.*¹¹⁶ was the first to discover a new surface phase of highly organised “double-strand rows” after annealing at 1150 K, although that phase is not considered here. These studies have also shown that the vacuum-annealing preparation process for rutile (110) samples creates a small concentration of O vacancies and so these samples are slightly substoichiometric in O. In this chapter, the case of stoichiometric models is focused on in common with conventional surface energy calculations. This is because attempts to calculate surface energies for non-stoichiometric selections of binary compounds immediately result in an *a priori* conflict with the definition of surface energy itself. A detailed discussion of O substoichiometry on rutile (110) samples is given in section 4.4.6. Despite these STM studies showing that steps are clearly extended along well-defined crystallographic directions, the exact atomic arrangement along the steps is difficult to determine theoretically with certainty. The difficulty is that the annealing or cleaving process may provide the energetic perturbations required for surface atoms to overcome local reconstruction activation

barriers and relax into a configuration which cannot be obtained by cleaving the bulk structure. However, Martinez *et al.*¹¹⁹ at Aarhus University recently harnessed the predictive power of a genetic algorithm with density functional theory (DFT) to tackle this difficulty for $\langle 1\bar{1}1 \rangle$ and $\langle 001 \rangle$ type steps on rutile (110).

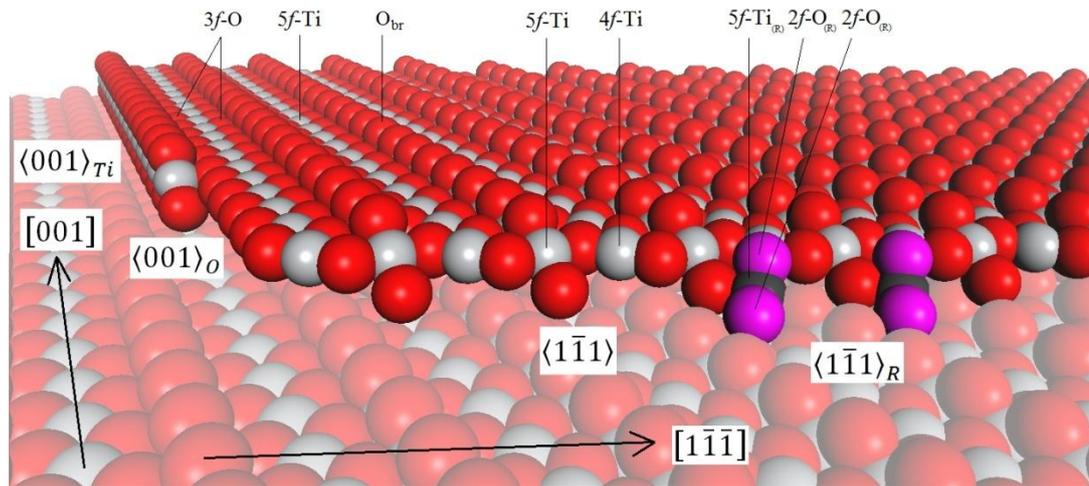


Figure 4.1. The edge of a terrace island bound by $\langle 001 \rangle$ - and $\langle 1\bar{1}1 \rangle$ -direction step structures on the stoichiometric (110) surface of TiO_2 rutile.

The structures $\langle 001 \rangle_{Ti}$, $\langle 001 \rangle_O$, $\langle 1\bar{1}1 \rangle$ and $\langle 1\bar{1}1 \rangle_R$, are the lowest energy stoichiometric structures theoretically predicted by Martinez *et al.*¹¹⁹ The extra TiO_2 units of the $\langle 1\bar{1}1 \rangle_R$ structure are depicted in purple and dark grey to signify that their positions do not correspond to frozen bulk truncations. The surface atoms are labelled according to their coordination numbers. For example, 5f-Ti refers to a five-fold coordinated Ti atom.

By sampling a large number of candidate step structures on a stoichiometric (451) slab, the most stable atomic arrangement which they recovered for $\langle 1\bar{1}1 \rangle$ steps is found to be *not* obtainable from a bulk truncation, but is instead formed by adding an extra TiO_2 unit to the surface cell, a structure which they coined $\langle 1\bar{1}1 \rangle_R$. Their study clearly demonstrates the importance of exhaustive structure searching methods in the characterisation of annealed surfaces. The most energetically stable $\langle 001 \rangle$ step structures that they recovered were two stoichiometric bulk-truncated terminations, both of which are considered in this study. The two most stable $\langle 1\bar{1}1 \rangle$ -type steps and $\langle 001 \rangle$ -type steps are illustrated in Figure 4.1. In their study, they calculated total supercell energies which reveal the overall stability of vicinal surfaces.

However, the individual energetic components of high-index surfaces can not be extracted from such comparisons of total supercell energy.

4.2.3 Step energy calculations are difficult because of quantum size effects

Theoretical calculations of the properties of material surfaces using periodic slab models of computationally practical size are often strongly influenced by quantum size effects (QSE).¹²⁰⁻¹³³ For the case of TiO₂ rutile (110), Bredow *et al.*¹³¹ showed that dominant QSE exist in the form of alternating modes of Ti 3*d* and O 2*p* electronic orbital hybridisation caused by the alternating odd/even number of oxide layers in slabs of increasing thickness.¹³¹ All ground-state material properties are derived from the electron density in DFT calculations, so Bredow's paper is significant because it is indicative of the very general and widespread issue of dealing with QSE when modelling any surface property for materials in which QSE are strong. It is well established from published DFT calculations that TiO₂ rutile surface energies^{120,121-124, 126,132,133}, surface vacancy formation energies^{123,125,126} and other surface energetic properties such as molecular adsorption energies¹²⁷⁻¹³⁰ all converge in a slow and oscillatory way with increasing slab thickness. This general situation means that calculating step energies accurately using high-index surfaces using slab models is difficult. This is because the step energy is tiny compared to the surface energy, and the surface energy (itself tiny compared to total supercell energies) is very sensitive to the choice of slab structure.

4.2.4 Linear fitting methods

There is one commonly used statistical method of estimating surface energies which has been used in the literature^{120,134-137}, which is to plot the total energies of increasingly large slab/vacuum supercells against the number of atoms or formula units in the slab, and to then use linear regression to draw a best fit line and interpret the small positive intercept of this best fit line as the surface energy at infinite slab thickness. It is absolutely critical that this approach is *avoided* for materials in which surface energies converge in a slow and oscillatory manner with increasing slab

size, such as TiO₂ rutile. For such systems, this method can give vastly different intercepts depending on which slab energy data points are chosen to be included in the fit, and therefore vastly different estimates of the surface energy. This is a philosophically unacceptable state of affairs. The "surface energy" cannot be allowed to depend so dramatically on the arbitrary choice of which slab energies are included in the data range. If this fitting approach is used on high-index surface slabs with the intention of extracting the tiny and extremely sensitive energies of steps, it fails spectacularly, and for the same reason. Alongside this obvious practical limitation, it is also argued that using a linear fitting approach in this context has some questionable underlying assumptions because it fails to properly recognize surface energy oscillations in TiO₂ rutile as a non-random, reproduceable and physically significant manifestation of alternating configurations of electron density, as already explained by Bredow *et al.*,¹³¹ and instead wrongly treats them as random deviations from a linear trend characterised by a Gaussian distribution. From a practical point of view, it is worth noting that linear fitting methods work without problems for materials in which surface energies can easily be fully converged with slab thickness (such as metals) as long as the unconverged surface energies from the very thinnest slabs are not included in the fit. There are examples in the literature of robust uses of linear fitting on ledge energies^{138,139} (step + interaction energy), linear fitting on high-index surface energies¹⁴⁰, and direct extraction of step energies from vicinal slabs^{141,142}, where, in all cases, it was ensured that total surface energies were fully converged at the slab thicknesses used. Remarkably, the Aarhus group¹⁴³ independently published their own method for calculating step energies, and *also* applied it to steps on TiO₂ rutile, only 65 days after the method in this chapter was published. They took a practical approach to overcoming the difficulties with linear fitting by simply applying a linear fit to the total energies of *unrelaxed* slabs, for which it is known that the surface energy oscillations decay much more quickly.¹³² This group also noted that linear fitting would only work on fully relaxed slabs of intractable size. The difficulty with systems like TiO₂ rutile is that linear fitting techniques can be misleading, and the "brute force" approach of using arbitrarily large slabs does not work because of the very long-ranged nature

of surface-surface interactions. In Appendix B, it is shown explicitly that the linear fitting approach is flawed using example data from this chapter, and a nonlinear fitting method, along with a discussion of its own limitations, is offered as an alternative.

Despite the apparent hindrance of QSE to slab calculations for systems like TiO_2 , the slow convergence of surface energies with slab sizes actually constitutes a rich source of data which has apparently not yet been exploited in any published articles. In this article, a new systematic approach to extracting the energies of steps is demonstrated which fully uses this data to cross-check and make robust what is otherwise a potentially unreliable method. In order that the suggested methodology be open to scrutiny and improvement, and for the sake of reproducibility, the elementary details are laid out in section 4.2.5. This methodology is then applied to the rutile (110) surface using DFT and the calculated step energies are used to predict island shapes as a function of slab thickness for comparison with published STM images showing terrace islands on this surface. The atomic configuration difficulty described above is tackled by focusing on the lowest energy step structures recently predicted by Martinez *et al.*¹¹⁹; specifically, the two most stable structures along the $\langle 1\bar{1}1 \rangle$ direction and the two most stable structures along the $\langle 001 \rangle$ direction, shown in Figure 4.1. A similar notation is adopted for the step energies as those used in 2006 for the case of steps on TiO_2 anatase (101) by Gong *et al.*,¹³⁶ from which much of the inspiration for this project originated. In their study, the authors calculated the energies of steps theoretically using supercells with DFT. However, the authors of this study used linear fitting to obtain their surface energies, and comprehensive technical details of their methodology and data analysis were not provided, so it is unclear whether they encountered any difficulties relating to QSE, and how they handled them if so.

4.2.5 Methodology

To calculate the energies of steps and their interactions, high-index vicinal surfaces with Miller indices (hkl) which lie at a small angle θ_{hkl} to the (110) plane are modelled. For a vicinal surface, the total surface energy per unit

area (of the plane (hkl)), γ_{hkl} , can be expressed in the form used by Gong *et al.*¹³⁶:

$$\frac{\gamma_{hkl}}{\cos\theta_{hkl}} = \gamma_{110} + \frac{\beta}{d_{hkl}} + \frac{q}{d_{hkl}^3}, \quad (4.1)$$

where γ_{110} is the surface energy per unit area of a pristine (110) terrace, β is the energy per unit length along an isolated step, q is the interaction constant representing the strength of the step-step interaction and d_{hkl} is the step-step separation distance measured in the (110) plane.

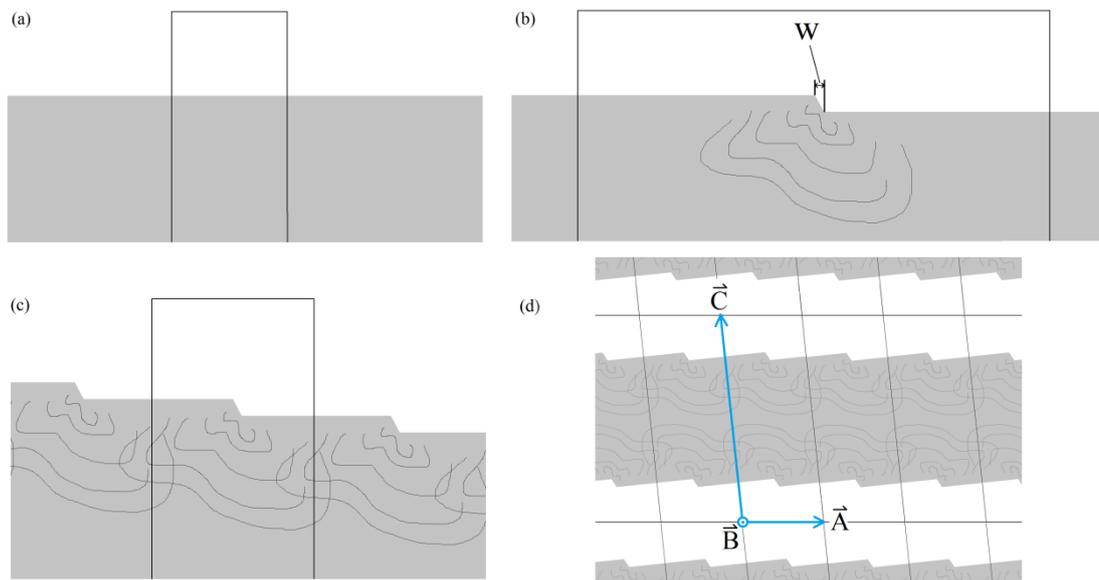


Figure 4.2. *Cartoon illustrations of the physical features corresponding to the terms in equation (4.1).*

(a) A pristine low-index semi-infinite terrace surface. (b) A low-index semi-infinite surface with a single isolated step. The contours illustrate the strain field created by the step. The width of the step is given by w , and is a well-defined quantity in the non-atomistic continuum solid surface model illustrated here. (c) A high-index semi-infinite stepped surface, where the terraces are the same terraces as in (a) and (b). (d) A stepped surface modelled under PBC. The total surface energy is now given by the terrace, step and step-interaction terms plus contributions resulting from the interactions between opposing surfaces within the same supercell. The vacuum gap must be large enough that surfaces cannot interact across the supercell boundaries.

Note the important distinction between these elastic step-step interactions associated with strain, and the entirely different step-step repulsion

associated with the presence of kinks due to configurational entropy at non-zero temperature.¹⁴⁴⁻¹⁴⁶ Note also that this equation can be applied to *any* low-index surface; not just (110). In the general case, θ_{hkl} denotes the angle between the low-index plane of interest and the relevant high-index plane (hkl). The notation used by Gong *et al.*¹³⁶ is adopted for the terrace, step and step-step interaction terms for consistency. By writing the interaction term in the form q/d^3 , the $\sim 1/d^2$ interaction potential of Marchenko and Parshin¹⁴⁷ (MP) derived from the general surface stress tensor for continuum solids has provisionally been adopted.

Each term in equation (4.1) can be understood by referring to Figure 4.2 which gives simplistic illustrations of the surface features from which the terms originate. Let the total surface free energy attributable to the region of material contained within the surface cell drawn in black in Figures 4.2 (a), (b) and (c) be denoted ε (eV). In Figure 4.2 (a), the total surface energy within the box is attributable to the existence of the terrace and is simply given by $\varepsilon = \varepsilon_t$, where ε_t is the terrace energy. In Figure 4.2 (b), a single isolated step of width w is shown on an otherwise pristine terrace as shown in 4.2 (a). The contour lines illustrate the strain field created by the step. The total surface energy is $\varepsilon = \varepsilon_t + \varepsilon_s$, where ε_s is the step free energy contribution. The total surface energy of the box drawn on the high-index semi-infinite stepped surface in Figure 4.2 (c) now includes a step interaction term in the form of overlapping strain fields emanating from adjacent steps, so is given by $\varepsilon = \varepsilon_t + \varepsilon_s + \varepsilon_{ss}$, where ε_{ss} is the step-step interaction energy. There are two important points which must be understood if equation (4.1) is to be validly applied to atomistic stepped surface models. Firstly, equation (4.1) is applicable to *semi-infinite surfaces*, i.e. where only one surface is defined, and makes no account of surface-surface interactions encountered in slab models in which the two opposing surfaces are only separated by a few atomic layers. Figure 4.2 (d) illustrates such a typical stepped slab system under PBC. Therefore, when attempting to apply equation (4.1) to slab models, it is necessary to converge all surface energy quantities with respect to increasing slab thickness to try to eliminate the influence of surface-surface interactions. Secondly, the width of the step, w , must ideally be significantly smaller than the step-step distance d_{hkl} , so that $w \ll d_{hkl}$. If

this condition is not satisfied, then the assignment of the terms from equation (4.1), designed for a simplistic continuum surface, to the more complicated arrangements of atoms on an atomistic surface, becomes invalid. This is one of the difficulties of this method, and the extent to which the small step width approximation is valid for the atomistic models considered in this chapter is assessed in the later sections.

Consider a semi-infinite atomistic vicinal surface like that illustrated in Figure 4.3, in which the supercell lattice vectors \vec{A} and \vec{B} lie in the high-index plane (hkl) with \vec{B} lying parallel with the step, and in which \vec{C} is understood to extend infinitely far into the bulk material. Let the total surface free energy per supercell, ε_{hkl} (eV), be expressed as the sum of three contributing components: ε_t , the free energy of a (110) terrace region, ε_s , the free energy of the step structure and ε_{ss} , the energy associated with the strain field due to step-step interactions. The approximation that the step is of insignificant width is used so that the surface area of the (110) terrace region per supercell, S' , is given by the projection of the area $S = |\vec{A}||\vec{B}|\sin\theta_{AB}$ onto the (110) plane with $S' = S\cos\theta$, where θ is the angle subtended by the high-index plane with the (110) plane, and θ_{AB} is the angle subtended by the lattice vectors \vec{A} and \vec{B} . Since the lattice vector \vec{B} lies parallel with the step, the step-step distance d measured perpendicular to the steps and lying in the (110) plane is found by finding the projection of the vector \vec{A} onto the (110) plane, $\vec{A}_p = \vec{A}\cos\theta$, and then extracting the magnitude of the component of \vec{A}_p which is perpendicular to \vec{B} by multiplying \vec{A}_p by $\sin\theta_{AB}$. The general expression for the step spacing is therefore $d = |\vec{A}|\cos\theta\sin\theta_{AB}$. The inverse square MP¹⁴⁷ step-step interaction potential can then be used to write $\varepsilon_{ss} = \frac{U}{d^2}$, where U is some interaction constant to be found. By combining all of these considerations with some straightforward manipulations, the total surface energy per unit area of the (hkl) plane, γ_{hkl} (eVÅ⁻²), can be written as

$$\gamma_{hkl} = \frac{\varepsilon_{hkl}}{S} = \frac{\varepsilon_t + \varepsilon_s + \varepsilon_{ss}}{S} = \cos\theta \left(\frac{\varepsilon_t}{S'} + \frac{\varepsilon_s}{d|\vec{B}|} + \frac{U}{d^3|\vec{B}|} \right). \quad (4.2)$$

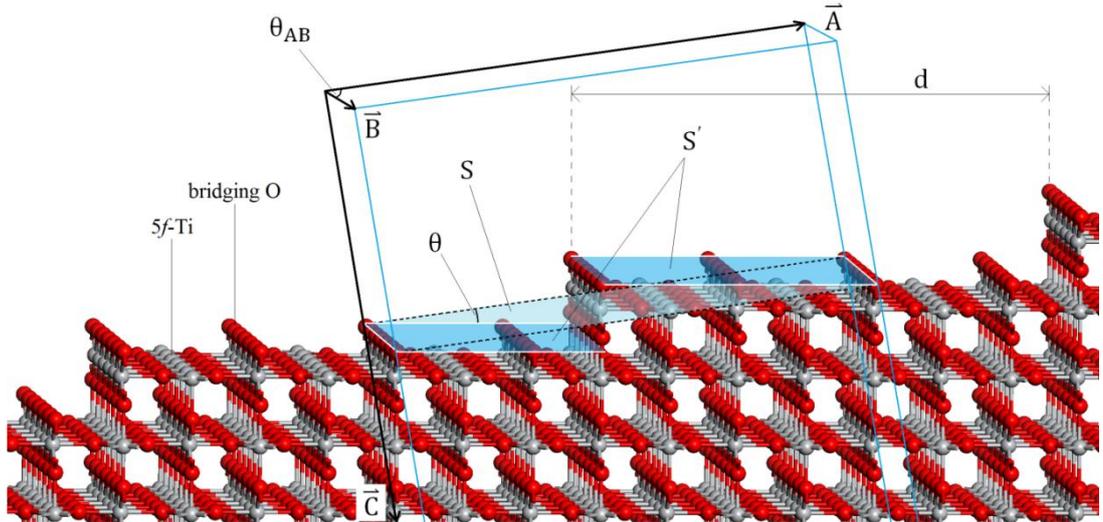


Figure 4.3. Ball-and stick model of a semi-infinite periodic (430) surface of TiO_2 rutile, consisting of (110) terraces and $\langle 001 \rangle_{\text{Ti}}$ steps.

The region shaded in transparent blue with white borders lies in the (110) plane and has area S' , equal to the projection of the area $S = |\vec{A} \times \vec{B}|$ onto the (110) plane. The step-step distance d lies in the (110) plane and perpendicular to \vec{B} . Equation (4.2) can be readily applied to the semi-infinite surface illustrated here, but care must be taken when applying it to slab models because the opposing surfaces of double-ended slabs interact over a long range.

The term $\frac{\varepsilon_t}{S'}$ is identified as the (110) terrace surface energy per unit area, γ_{110} ($\text{eV}\text{\AA}^{-2}$), and $\frac{\varepsilon_s}{|\vec{B}|}$ as the energy per unit length along the step, β ($\text{eV}\text{\AA}^{-1}$).

The term $\frac{U}{|\vec{B}|}$ is also identified as a constant representing the strength of the step-step interaction which is named q ($\text{eV}\text{\AA}$) in order to recover equation (4.1). The form of equation (4.1) knows nothing of atoms, so it is a macroscopic equation and can only be validly applied to atomic step structures as long as the small step width approximation holds. The extent to which this approximation is valid for the atomic step structures studied in this thesis is evaluated throughout sections 4.4.1 to 4.4.5 inclusive.

By using stoichiometric slab/vacuum supercells, numerical values for γ_{hkl} can be calculated using DFT in the conventional way by evaluating the difference between the geometry-optimised supercell energy and that of the corresponding quantity of bulk material:

$$\gamma_{hkl}(m) = \frac{1}{2S} (E_{hkl}(m) - m/2 E_B), \quad (4.3)$$

where E_B is the energy of the Ti_2O_4 bulk rutile unit cell, E_{hkl} is the geometry-optimised total energy of a stoichiometric and symmetric Ti_mO_{2m} slab/vacuum supercell whose lattice vectors \vec{A} and \vec{B} lie in the high-index plane (hkl) and $2S$ is the total surface area per (symmetric and double-ended) supercell, given by $S = |\vec{A} \times \vec{B}|$. The essential requirement for substituting surface energies calculated using equation (4.3) into the semi-infinite surface energy equation (4.1) is that the slab thickness be systematically increased until the resultant energies are converged so that surface-surface interactions are accounted for. The dependence of γ_{hkl} on the slab chemical formula Ti_mO_{2m} is emphasised by writing $\gamma_{hkl} = \gamma_{hkl}(m)$. To create a structurally consistent set of high-index slabs, a convention is chosen whereby the number of layers in a given slab, L , is defined in terms of m and so L is therefore used as the independent variable.

By calculating the surface energy per unit area, γ_{hkl} , of three progressively higher index slab/vacuum supercells of L layers thickness using equation (4.3), and then substituting these values into equation (4.1), a system of three linear simultaneous equations is obtained whose solutions are $\gamma_{110}(L)$, $\beta(L)$ and $q(L)$. For example, (341), (451) and (561) slabs of L layers thickness hosting the $\langle 1\bar{1}1 \rangle$ step can be used to yield the surface energy per unit area values $\gamma_{341}(L)$, $\gamma_{451}(L)$ and $\gamma_{561}(L)$, which can then be substituted into equation (4.1) along with the corresponding step spacings d_{341} , d_{451} , d_{561} and vicinal surface misorientation angles θ_{341} , θ_{451} and θ_{561} taken directly from the supercell dimensions. The resulting system of three linear simultaneous equations written in matrix form is

$$\begin{pmatrix} \gamma_{341}(L)/\cos\theta_{341} \\ \gamma_{451}(L)/\cos\theta_{451} \\ \gamma_{561}(L)/\cos\theta_{561} \end{pmatrix} = \begin{pmatrix} 1 & d_{341}^{-1} & d_{341}^{-3} \\ 1 & d_{451}^{-1} & d_{451}^{-3} \\ 1 & d_{561}^{-1} & d_{561}^{-3} \end{pmatrix} \begin{pmatrix} \gamma_{110}(L) \\ \beta(L) \\ q(L) \end{pmatrix}, \quad (4.4)$$

and the solutions are

$$\gamma_{110}(L) = \frac{\det G(L)}{\det D}, \quad \beta(L) = \frac{\det B(L)}{\det D}, \quad q(L) = \frac{\det Q(L)}{\det D}, \quad (4.5)$$

where D is the matrix from equation (4.4) and $G(L)$, $B(L)$ and $Q(L)$ are matrices identical to D but whose first, second and third columns respectively have been replaced with the column vector of high-index surface energies from the left-hand side of equation (4.4), as per Cramer's determinant rule. This process can then be repeated for increasingly thick slabs of higher L , where the slabs' chemical formulae are increased according to a prescribed arithmetic rule, thereby exposing the behaviour of $\gamma_{110}(L)$, $\beta(L)$ and $q(L)$ with increasing L individually and consistently. The oscillatory behaviour of $\gamma_{hkl}(L)$ with increasing L propagates through to the solutions $\gamma_{110}(L)$, $\beta(L)$ and $q(L)$ so they also oscillate with increasing slab thickness and depend sensitively on L . The significant advantage of this approach is that the solutions $\gamma_{110}(L)$ can be cross-checked, for each and every value of L , with values calculated directly and separately using (110) slabs. These directly calculated values are denoted $\tilde{\gamma}_{110}(L)$. The residual differences $\gamma_{110}(L) - \tilde{\gamma}_{110}(L)$ over the full range of L can then be analysed in order to determine whether the terrace component of the total high-index surface energy has been consistently extracted while naturally accounting for quantum size effects, therefore leaving the remaining step energies and interaction energies open to self-contained analysis. Expressed another way, it allows one to determine whether the macroscopic $\sim 1/d^2$ MP interaction potential accurately captures the true energetics of the atomistic slab models. This is important, because whilst the MP interaction potential appears to have been adopted by Gong *et al.*¹³⁶ and is frequently assumed and observed in the literature¹⁴⁸⁻¹⁵¹, there are also many instances in the literature¹⁵²⁻¹⁵⁷ where atomic step-step interaction energies for various material surfaces have been observed or suggested to deviate from the $\sim 1/d^2$ behaviour due to atomistic effects.

4.3 METHOD

4.3.1 Basis set, exchange correlation functional, bulk reference state and vacuum thickness

The plane wave density functional theory code CASTEP⁸³ was used with the GGA PBE exchange correlation functional⁶¹ and Vanderbilt ultrasoft

pseudopotentials²⁰ to converge the total energy of the bulk TiO₂ rutile unit cell with respect to the k point spacings s^i (where i denotes correspondence to the reciprocal vector \vec{b}_i) and the kinetic energy cutoff E_{cut} , where each was varied independently. This procedure yielded basis set values satisfying $s^i < 0.035 \text{ \AA}^{-1}$ and $E_{cut} = 400 \text{ eV}$, which were then used along with the GGA PBE functional for all subsequent calculations in the chapter. Following this, a cell-optimised geometry optimisation calculation was carried out on the bulk unit cell with P4₂/mnm symmetry imposed, in which the atomic forces were converged to within 0.01 eV \AA^{-1} . This stringent tolerance on the bulk cell was motivated by the high sensitivity of γ_{hkl} to the value of E_B in equation (4.3). The bulk unit cell lattice parameters were recorded as $|\vec{a}| = 4.644 \text{ \AA}$ (3 d. p.) and $|\vec{c}| = 2.976 \text{ \AA}$ (3 d. p.), in good agreement with the experimentally measured¹⁵⁸ quantities $|\vec{a}|_{exp} = 4.594 \text{ \AA}$ (3 d. p.) and $|\vec{c}|_{exp} = 2.969 \text{ \AA}$ (3 d. p.). A three layer stoichiometric and symmetric (110) slab/vacuum Ti₆O₁₂ supercell was then built, whose surface-plane lattice vectors were fixed at values derived directly from the optimised bulk cell, and for which the vacuum thickness was fixed at magnitudes increasing in 1 Å increments from 3 Å to 12 Å inclusive, to make a total of 10 supercells. The geometry-optimised system energy was calculated for each, and a minimum vacuum thickness of 10 Å was identified in order to decouple all intercellular interactions along the vacuum direction. In all cases, the slab was placed in the centre of the vacuum.

4.3.2 Slab calculations

For the $\langle 001 \rangle_{Ti}$ step structure, (210), (320) and (430) slabs of 3,..., 10 layers ($L = 3, \dots, 10$) and $P 1 2/m 1$ symmetry were built to make a total of $8 \times 3 = 24$ slabs on which the terminations were cleaved manually to produce (110) terraces bounded by the $\langle 001 \rangle_{Ti}$ steps. This procedure was repeated for the $\langle 001 \rangle_O$ structure to make 24 more slabs, also of $P 1 2/m 1$ symmetry. For the $\langle 1\bar{1}1 \rangle$ and $\langle 1\bar{1}1 \rangle_R$ steps, (341), (451) and (561) slabs of 3,..., 10 layers ($L = 3, \dots, 10$) and $P\bar{1}$ symmetry were built to make $24 \times 2 = 48$ further slabs. For the $\langle 1\bar{1}1 \rangle_R$ slabs, the extra TiO₂ units were added manually to each end.

Alongside these vicinal surfaces, 8 increasingly-thick symmetric and stoichiometric (110) slabs of 3,..., 10 layers were built for the consistency check. The total set of input files thus consisted of 104 slab/vacuum supercells. The projections onto the (110) plane of the supercells are shown in Figure 4.4. For each and every one of these 104 supercells, the surface-plane lattice vectors, \vec{A}_{hkl} and \vec{B}_{hkl} were fixed at values derived directly from the optimised bulk lattice parameters $|\vec{a}| = 4.644 \text{ \AA}$ (3 d. p.) and $|\vec{c}| = 2.976 \text{ \AA}$ (3 d. p.) as appropriate to each case. The surface-normal lattice vector \vec{C}_{hkl} was fixed at a value such that the vacuum thickness always equalled or exceeded 10 \AA and the crystal lattice was re-oriented with respect to the Cartesian coordinate system so that \vec{C}_{hkl} was always aligned with the z axis.

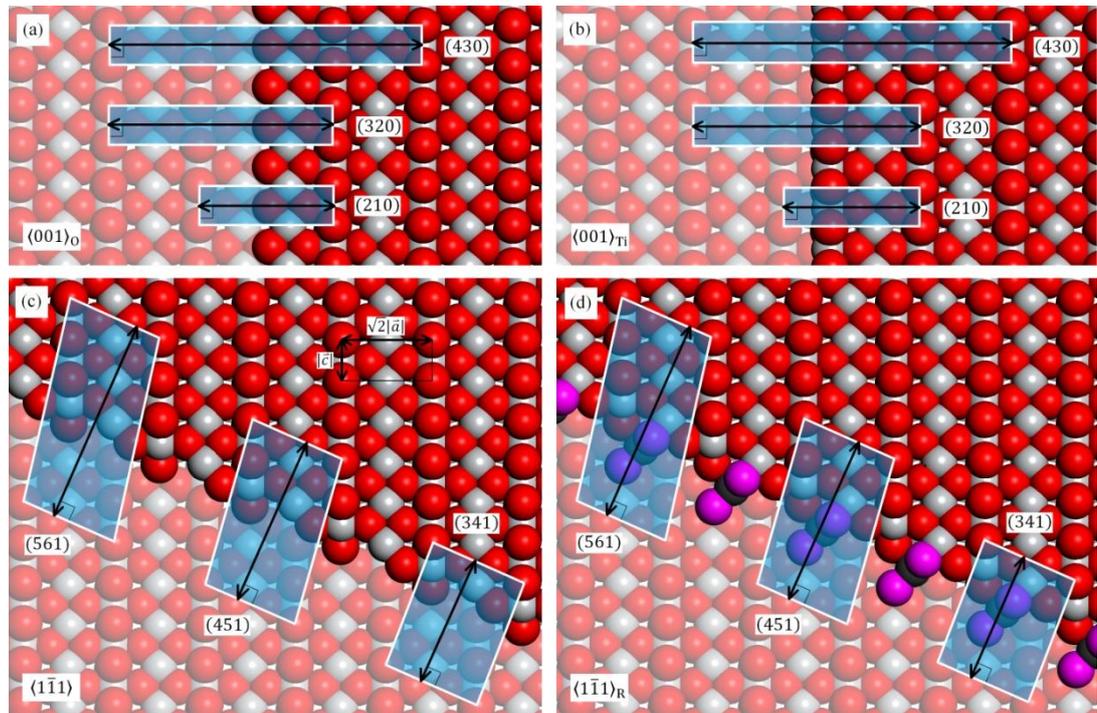


Figure 4.4. Top-down views of the projection of the high-index supercells onto two terraces of a semi-infinite (110) surface bounded by (a) $\langle 001 \rangle_O$ (b) $\langle 001 \rangle_{Ti}$ (c) $\langle 1\bar{1}1 \rangle$ and (d) $\langle 1\bar{1}1 \rangle_R$ steps.

The extra TiO_2 units of the $\langle 1\bar{1}1 \rangle_R$ structure are depicted in purple and dark grey as in Figure 4.1. The distance measurements lying in the (110) plane indicated with black arrows are the step-step spacings d_{hkl} and the step height in all cases is $\sqrt{2}|\vec{a}|$. The supercells used in the corresponding DFT calculations were built double-ended with vacuum slabs in the infinite periodic regime like those illustrated in Figure 4.5.

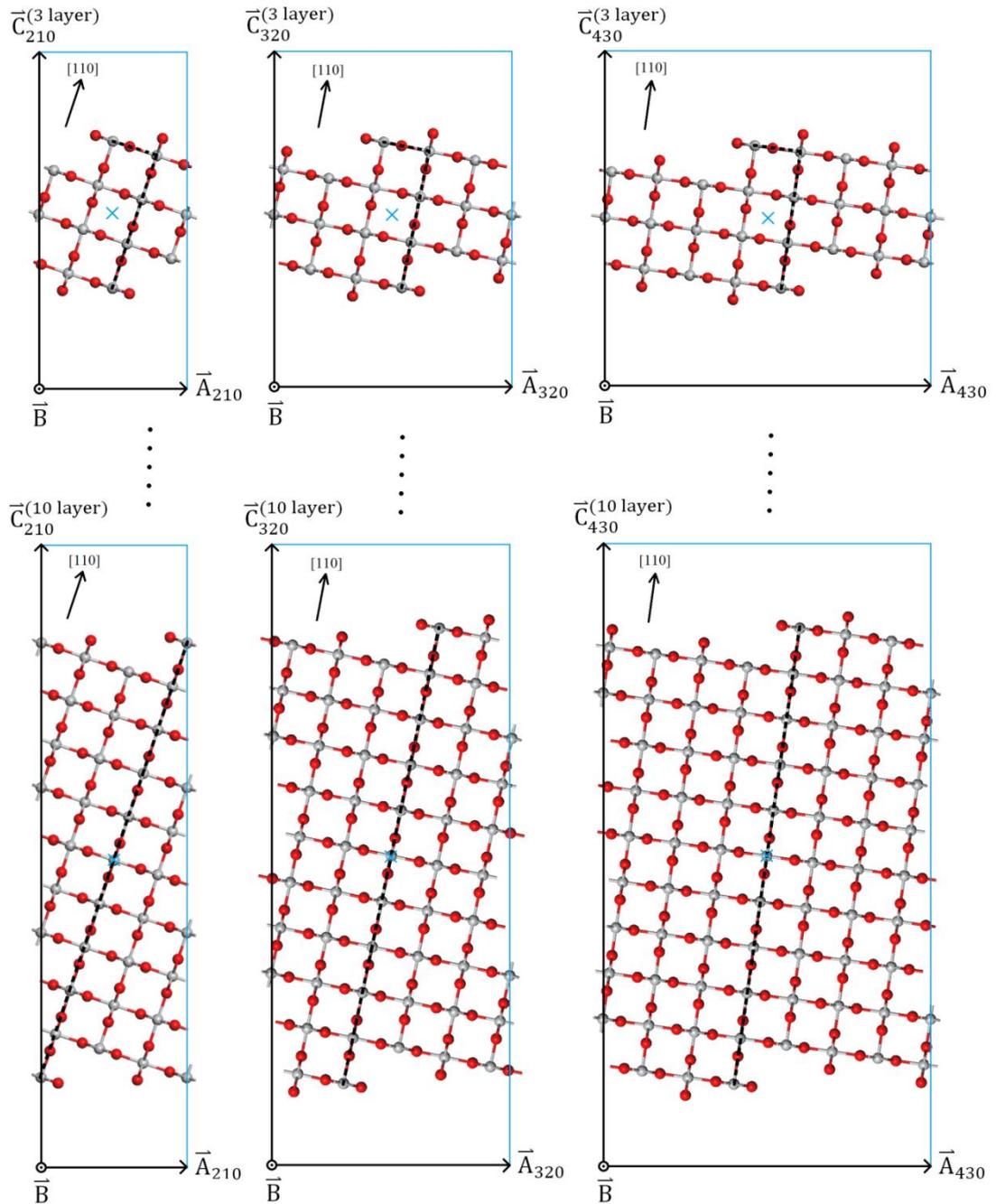


Figure 4.5. Side views of the set of 24 supercells used for the $\langle 001 \rangle_0$ step using ball and stick representation.

All supercells have inversion symmetry about the centroid located at fractional coordinates $(0.5, 0.5, 0.5)$, indicated in each supercell with the blue cross. The two steps on the opposing ends of every slab in the entire set of input files are consistently located either i) directly opposite each other along the direction $[110]$, for $L = \text{even}$ cases, or ii) opposite each other along the direction $[110]$, up to a discrepancy of a single unit cell for $L = \text{odd}$ cases. This is indicated by the black dotted lines joining the opposing steps for the supercells shown.

To maintain structural consistency, all high-index slabs were built such that the step on the top end of each slab was made to be situated directly opposite the step on the bottom end along the direction perpendicular to the (110) plane (up to a discrepancy of a bulk unit cell due to the AB stacking structure) as illustrated for the case of the $\langle 001 \rangle_{Ti}$ step in Figure 4.5. This structural convention was encoded arithmetically in terms of the slabs' chemical formulae and summarised in Table 4.1. CASTEP⁸³ was then used with the GGA PBE functional⁶¹ and Vanderbilt ultrasoft pseudopotentials²⁰ to carry out a geometry optimisation calculation on these 104 supercells in which all atomic positions were allowed to relax in all directions. For each series of self-consistent field (SCF) cycles used for the electronic minimisation, the exit criterion was imposed that the change in total electron energy between successive SCF cycles be less than 5×10^{-7} eV three times in succession.

Table 4.1. *The convention used to define the number of layers, L , in terms of the number of TiO_2 units, m .*

For example, a 6 layer (320) slab hosting a $\langle 001 \rangle_{Ti}$ step on each end has chemical formula $Ti_{31}O_{62}$. By using this convention, all four sets of slabs were structurally consistent with each other as exemplified in Figure 4.5.

Step	Slab Miller index	Number of layers, L , defined by chemical formula $Ti_m O_{2m}$	
		Even L	Odd L
$\langle 001 \rangle_{Ti}$	(210)	$Ti_{3L+1}O_{2(3L+1)}$	$Ti_{3L+2}O_{2(3L+2)}$
	(320)	$Ti_{5L+1}O_{2(5L+1)}$	$Ti_{5L+2}O_{2(5L+2)}$
	(430)	$Ti_{7L+1}O_{2(7L+1)}$	$Ti_{7L+2}O_{2(7L+2)}$
$\langle 001 \rangle_o$	(210)	$Ti_{3L+1}O_{2(3L+1)}$	$Ti_{3L+2}O_{2(3L+2)}$
	(320)	$Ti_{5L+1}O_{2(5L+1)}$	$Ti_{5L+2}O_{2(5L+2)}$
	(430)	$Ti_{7L+1}O_{2(7L+1)}$	$Ti_{7L+2}O_{2(7L+2)}$
$\langle 1\bar{1}1 \rangle$	(341)	$Ti_{7L+3}O_{2(7L+3)}$	$Ti_{7L+2}O_{2(7L+2)}$
	(451)	$Ti_{9L+2}O_{2(9L+2)}$	$Ti_{9L+3}O_{2(9L+3)}$
	(561)	$Ti_{11L+3}O_{2(11L+3)}$	$Ti_{11L+2}O_{2(11L+2)}$
$\langle \bar{1}\bar{1}1 \rangle_R$	(341)	$Ti_{7L+5}O_{2(7L+5)}$	$Ti_{7L+4}O_{2(7L+4)}$
	(451)	$Ti_{9L+4}O_{2(9L+4)}$	$Ti_{9L+5}O_{2(9L+5)}$
	(561)	$Ti_{11L+5}O_{2(11L+5)}$	$Ti_{11L+4}O_{2(11L+4)}$

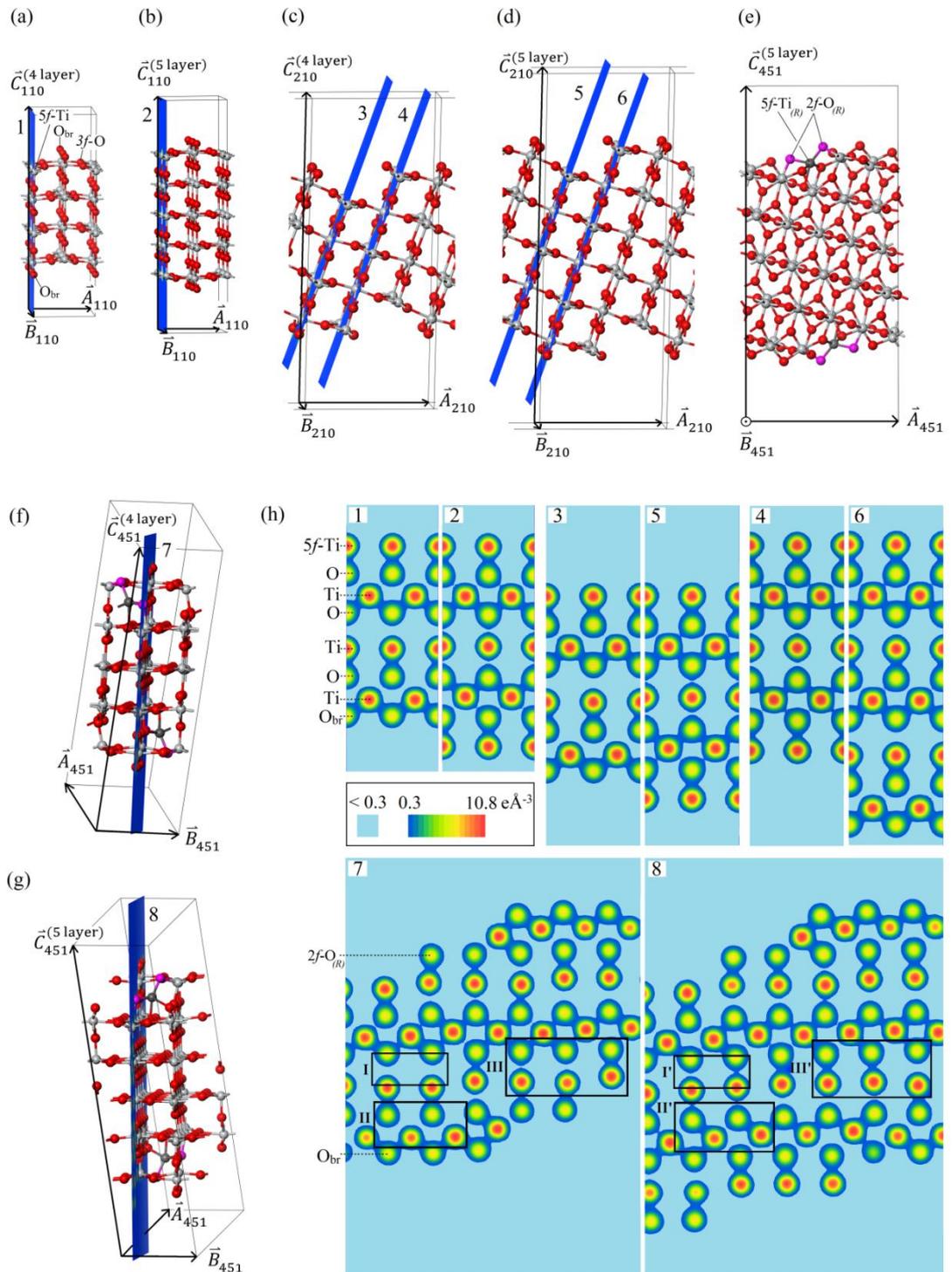


Figure 4.6. The physical reason for the oscillations in the energetic components plotted in Figures 4.8, 4.9 and 4.10: thickness-dependent alternating configurations of electron density.

(a) Ball and stick models of (a) 4 layer and (b) 5 layer (110) slabs with electron density slices passing through 5f-Ti surface atoms. (c) Ball and stick model of a 4 layer (210) / $\langle 001 \rangle_O$ slab, (d) a 5 layer (210) / $\langle 001 \rangle_O$ slab. (e) Side view of a 5 layer (451) / $\langle 1\bar{1}1 \rangle_R$ slab. (f) Ball and stick model of (f) 4 layer (451) / $\langle 1\bar{1}1 \rangle_R$ slab and (g) 5 layer (451) / $\langle 1\bar{1}1 \rangle_R$ slab. In all ball and stick models, Ti atoms are illustrated in grey and O in red. For the $\langle 1\bar{1}1 \rangle_R$ models, the extra Ti atoms are shown in dark grey, and the extra O in purple. (h) Electron density slices from (a) - (f) in units of electrons / \AA^3 . For slices 1 and 2, the interlayer distances between the 4th/5th and 5th/6th atomic layer from the top are quite different. The interlayer distances from slices 3 and 5, and from slices 4 and 6 in the (210) slabs show similar traits. In slices 7 and 8, the regions I/I', II/II' and III/III' are outlined where slab thickness-dependent alternating modes of electron density are most evident. As in Figure 4.1, selected surface atoms are labelled according to their coordination numbers.

For the geometry optimisation, the low memory¹⁵⁹ Broyden–Fletcher–Goldfarb–Shanno (LBFGS) optimisation algorithm⁸⁷⁻⁹¹ was used with the following 3 exit criteria: i) that the maximum force on all atoms be less than $0.03 \text{ eV}\text{\AA}^{-1}$, ii) that the maximum change in position for all atoms between successive LBFGS steps be less than 10^{-3} \AA and iii) that the maximum change in the total system energy between successive LBFGS steps be less than 10^{-5} eV per atom.

The total electron density slices of the fully relaxed slabs were then examined, some of which are presented in Figure 4.6. The 96 fully-optimised supercell energies E_{hkl} were inserted as appropriate with the bulk energy E_B into equation (4.3) to yield the surface energy per unit area of the high index surface for each slab, $\gamma_{hkl}(L)$, and these values were plotted against L in Figure 4.7 and tabulated in Table 4.2. These values were then substituted as appropriate into equation (4.4), and the resulting systems of simultaneous equations were solved to find the solutions $\gamma_{110}(L)$, $\beta(L)$ and $q(L)$ with increasing L for each of the four steps as explained in section 4.2.5. For each step, the solutions $\gamma_{110}(L)$ were plotted against the $\tilde{\gamma}_{110}(L)$ values calculated directly using the (110) slabs in Figure 4.8 for the consistency check, along with the residuals $\Delta\gamma_{110}(L) = \gamma_{110}(L) - \tilde{\gamma}_{110}(L)$. The solutions $\beta(L)$ and $q(L)$ were plotted in Figures 4.8 and 4.9 respectively and tabulated in Table 4.2.

4.4 RESULTS AND ANALYSIS

4.4.1 Initial remarks

All 104 slabs relaxed into a configuration corresponding to an energetic minimum. It can be seen from the total electron density slices in Figure 4.6 that the interlayer interactions (and interlayer distances) depend critically on

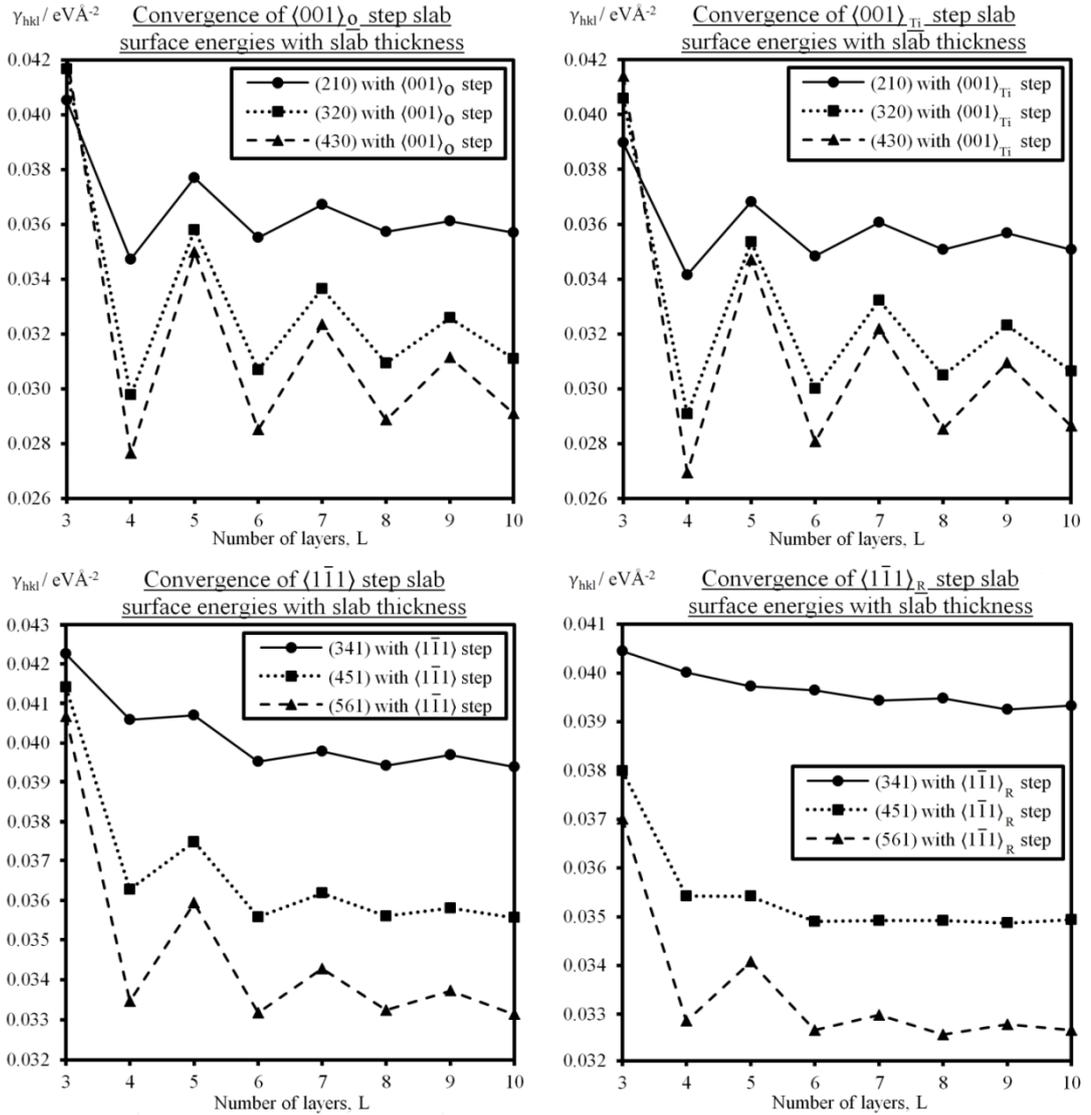


Figure 4.7. The convergence of surface energy per unit area with slab thickness.

Despite the clear oscillations due to alternating modes of electron density, the surface energies clearly converge to finite values. In all but a few $L = 3$ cases, higher index surfaces have lower surface energy per unit area.

the number of layers in the slab. The dominant electronic orbitals which contribute to this effect are Ti 3d and O 2p states, which was confirmed for all of the slabs with angular momentum-resolved density of states plots and orbital density plots. This is exactly what was observed and explained in detail for the case of TiO₂ rutile (110) slabs by Bredow *et al.*¹³¹ Electron density slices 1 and 2 shown in Figure 4.6 (h) for (110) slabs look virtually identical to those presented by Bredow *et al.* The high-index surface energies $\gamma_{hkl}(L)$ plotted in Figure 4.7 display oscillatory convergence with increasing L . This is the direct consequence of the alternating configurations of electron density as seen in Figure 4.6. It is also the physical reason why the method of using linear regression to estimate the surface energy from the intercept of a plot of total supercell energies against m is unsuitable for the case of TiO₂ rutile, or indeed any material in which surface-surface interactions are strong. It is emphasised at this point that the oscillations in Figure 4.7 are not merely a consequence of artificial aspects of the calculation; in particular, the different k points sampling grid used for the calculation of the bulk unit cell energy, E_B , from that of each total supercell energy, E_{hkl} . This was verified directly by calculating values of E_B using new (and substantially larger) bulk unit cells for selected slabs, where the lattice parameters of these bulk cells were systematically chosen so that the k point sampling grids in each case was identical to that of the corresponding slab/vacuum supercell. The resulting surface energies were the same to 7 significant figures and showed identical oscillations. This confirms, for the avoidance of any doubt, that the different sampling grid used for the bulk cell is of no consequence, and moreover that the basis set used throughout all these calculations is robust and fully converged. It can be seen that these oscillations propagate through to the solutions γ_{110} , β and q plotted in Figures 4.8, 4.9 and 4.10 respectively, and the magnitude of the residual oscillations in these solutions is large compared to the values themselves. The γ_{110} , β and q solutions were then substituted appropriately back into the high-index surface energy equation (4.1), and both sides were multiplied by $\cos\theta_{hkl}$ and then multiplied by the total slab surface area $2S = 2|\vec{A} \times \vec{B}|$ in each case to obtain the total surface energy ε_{hkl} of each high-index slab resolved into terrace, step and step-interaction components. These total

surface energy components were plotted in Figures 4.11 and 4.12 against slab thickness. Currently, *ab initio* calculations are generally impractical for supercells exceeding the size of those used in this study, and so to draw comparisons between the solutions γ_{110} , β and q with experimental observations of laboratory samples of finite but macroscopic thickness, one must rely on an extrapolation of the plots along the direction $+L$ towards $L \rightarrow \infty$. This demonstrates the challenge involved with calculating accurate step energies for macroscopically thick crystals. There is also a small statistical uncertainty in the fully optimised supercell energies E_{hkl} which originates from the finite tolerances used in the geometry optimisation. This uncertainty is common to the $\langle 1\bar{1}1 \rangle$, $\langle 1\bar{1}1 \rangle_R$, $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ systems.

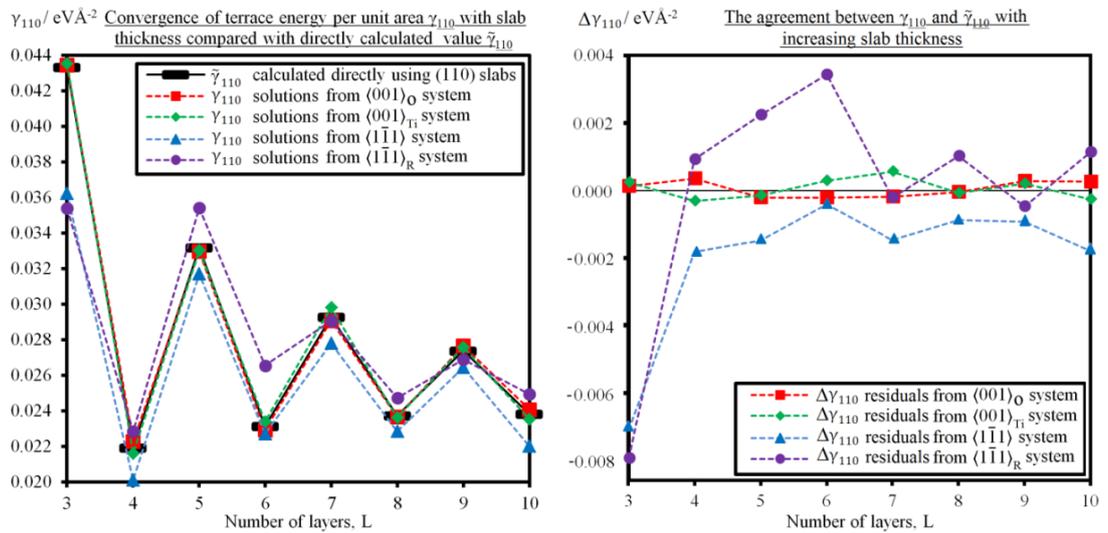


Figure 4.8. Layer-by-layer cross-check.

Left panel: The convergence of the (110) terrace energy per unit area for the four steps, compared with the values calculated directly using (110) slabs. **Right panel:** The residual differences $\Delta\gamma_{110}(L) = \gamma_{110}(L) - \tilde{\gamma}_{110}(L)$ which show the validity of applying equation (1) to the atomistic supercells used.

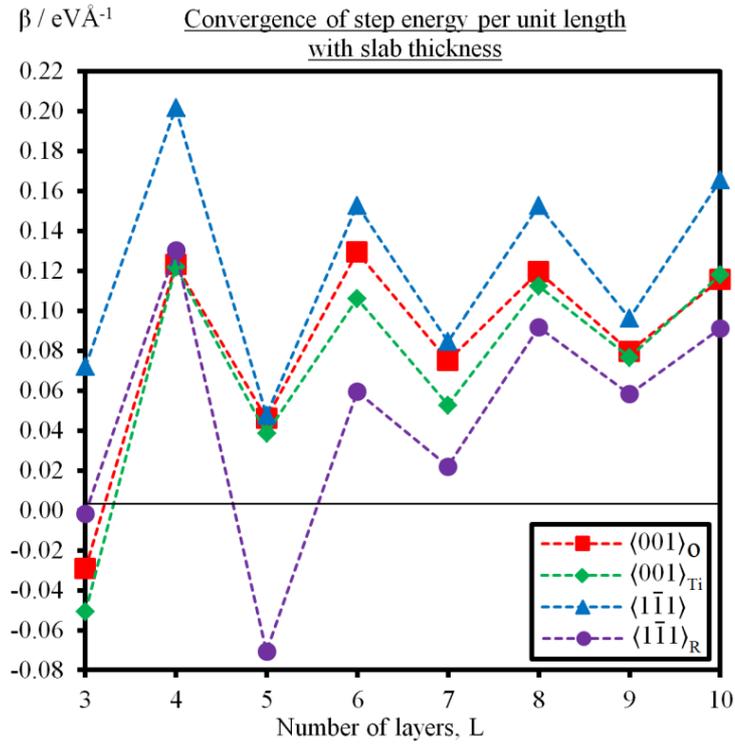


Figure 4.9. *The convergence of step energy per unit length with increasing slab thickness for the four steps studied.*

Despite the large oscillations in the solutions, it is quite clear that the $\langle 1\bar{1}\bar{1} \rangle_R$ step is the most stable.

The $\langle 1\bar{1}\bar{1} \rangle$ and $\langle 1\bar{1}\bar{1} \rangle_R$ slabs are substantially lower in symmetry than those of the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ systems, so this source of uncertainty is likely to be more severe for these systems. The reader is referred to Appendix A for a discussion of these sources of statistical uncertainty. Note that the values $\tilde{\gamma}_{110}(L)$ agree well with published calculations. [120,132,133](#)

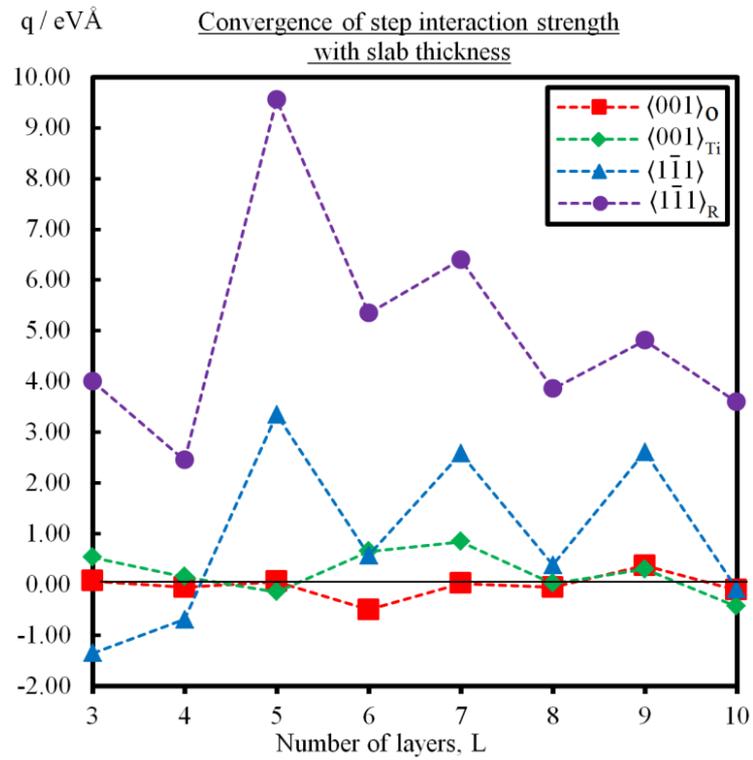


Figure 4.10. *The convergence of the step-step interaction constant with increasing slab thickness for the four steps studied.*

Despite the large oscillations in the plot, the $\langle 001 \rangle_O$ and $\langle 001 \rangle_{Ti}$ steps clearly have the weakest interactions, and the $\langle \bar{1}\bar{1}\bar{1} \rangle_R$ step has the strongest interactions.

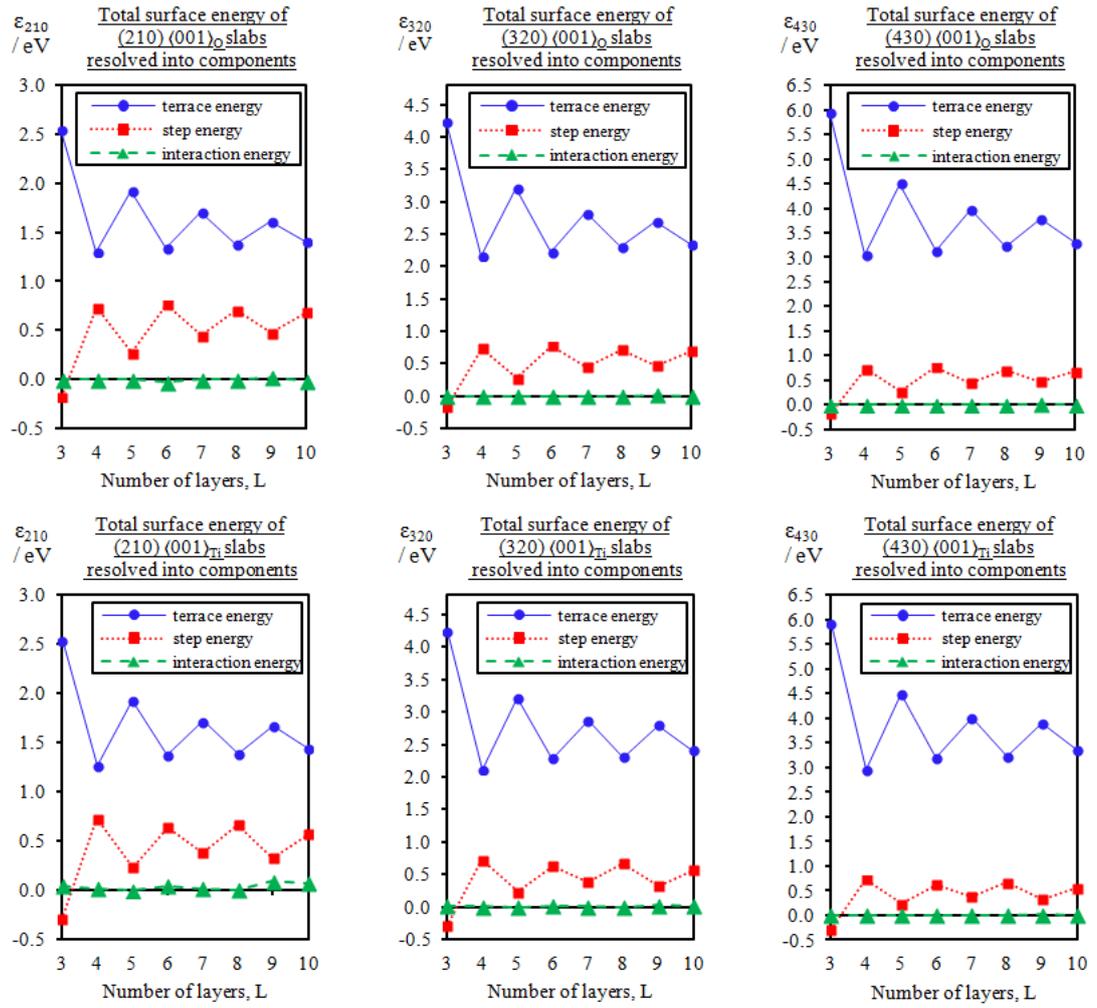


Figure 4.11 The total surface energy of the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ slabs resolved into terrace, step and step-step interaction components.

In all cases, the terrace energy is the dominant component of the total surface energy, with the step-step interaction energy contributing the least.

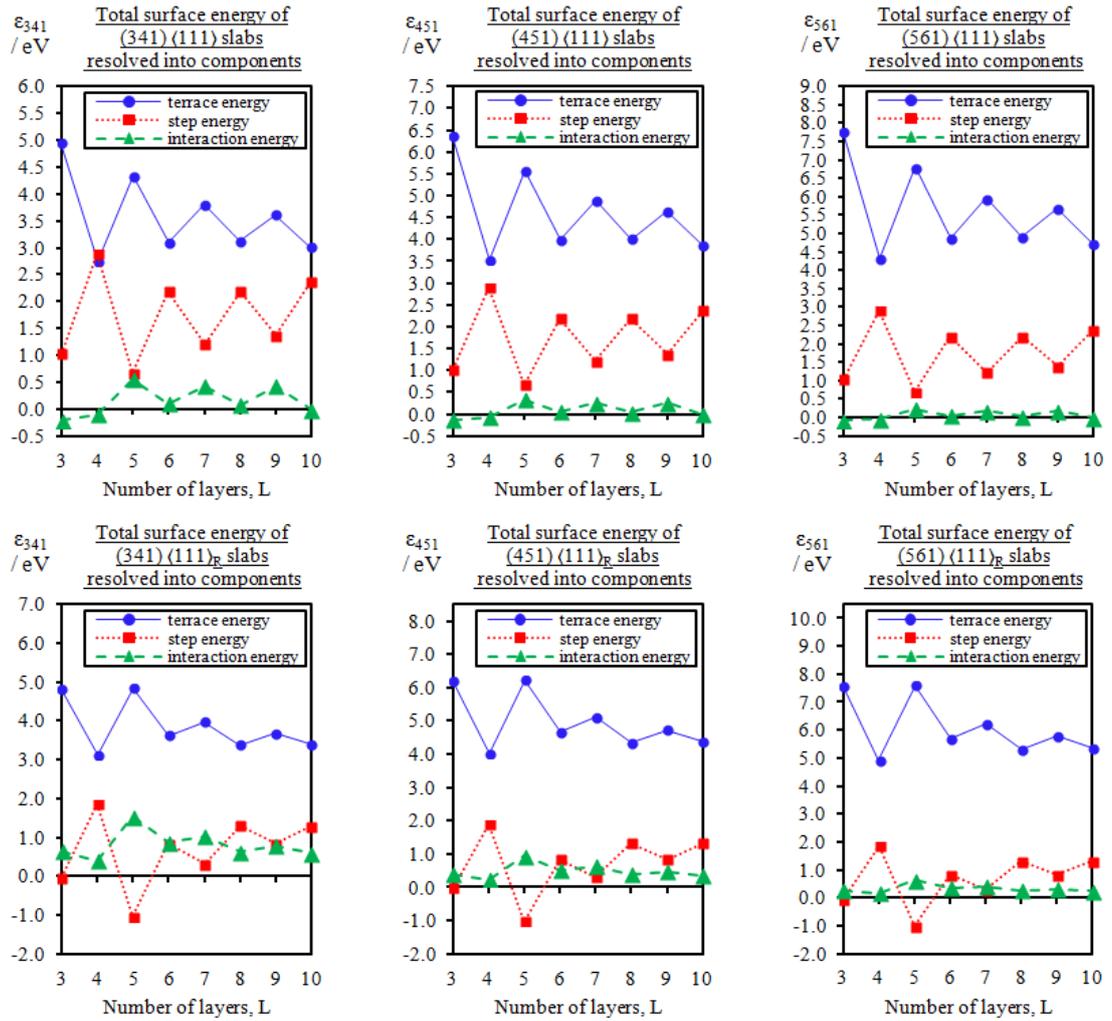


Figure 4.12 The total surface energy of the $\langle 1\bar{1}1 \rangle$ and $\langle 1\bar{1}1 \rangle_R$ slabs resolved into terrace, step and step-step interaction components.

The terrace energy dominates in all systems, with the step-interaction energy generally being the smallest component at largest L . In the $(341)\langle 1\bar{1}1 \rangle_R$ systems, the step interaction energy is only just larger than the step energy at the largest values of L . This shows that the trend of the interaction energy being the smallest component is preserved across the whole data set, even for the $(341)\langle 1\bar{1}1 \rangle_R$ slabs for which the small step width approximation is comparatively poor. Surface-surface interactions clearly have an overriding influence on the data at smaller L .

Table 4.2. The $\langle 110 \rangle$ terrace surface energy per unit area, step energy per unit length and step-step interaction constant on the four vicinal surface types.

Step	Number of layers, L	Slab surface energy per unit area of (hkl) plane, $\gamma_{hkl} / \text{eV}\text{\AA}^{-2} \times 10^{-2}$ (3 d. p.)			$\langle 110 \rangle$ terrace surface energy per unit area, $\gamma_{110} / \text{eV}\text{\AA}^{-2} \times 10^{-2}$ (3 d. p.)	Step energy per unit length, $\beta / \text{eV}\text{\AA}^{-1}$ (3 d. p.)	Step-step interaction constant, $q / \text{eV}\text{\AA}$ (3 d. p.)
		γ_{210}	γ_{320}	γ_{430}			
$\langle 001 \rangle_o$	3 layer	3.845	4.086	4.175	4.344	-0.029	0.066
	4 layer	3.294	2.919	2.736	2.229	0.123	-0.057
	5 layer	3.577	3.510	3.463	3.298	0.046	0.052
	6 layer	3.370	3.009	2.822	2.292	0.129	-0.498
	7 layer	3.483	3.300	3.202	2.908	0.075	0.024
	8 layer	3.390	3.033	2.857	2.367	0.120	-0.057
	9 layer	3.427	3.196	3.084	2.766	0.080	0.370
	10 layer	3.387	3.048	2.881	2.408	0.116	-0.110
$\langle 001 \rangle_{Ti}$	3 layer	3.697	3.980	4.098	4.355	-0.051	0.535
	4 layer	3.240	2.852	2.667	2.162	0.122	0.155
	5 layer	3.492	3.467	3.436	3.304	0.039	-0.149
	6 layer	3.305	2.942	2.779	2.341	0.106	0.649
	7 layer	3.421	3.258	3.187	2.983	0.053	0.843
	8 layer	3.327	2.990	2.825	2.364	0.112	0.021
	9 layer	3.385	3.169	3.063	2.759	0.077	0.299
	10 layer	3.328	3.005	2.836	2.355	0.118	-0.431
		γ_{341}	γ_{451}	γ_{561}			
$\langle 1\bar{1}1 \rangle$	3 layer	3.992	3.998	3.972	3.624	0.072	-1.352
	4 layer	3.834	3.503	3.266	2.011	0.201	-0.690
	5 layer	3.846	3.620	3.509	3.171	0.048	3.351
	6 layer	3.735	3.436	3.238	2.271	0.153	0.574
	7 layer	3.759	3.494	3.347	2.780	0.085	2.580
	8 layer	3.725	3.438	3.245	2.285	0.153	0.383
	9 layer	3.751	3.457	3.292	2.645	0.096	2.612
	10 layer	3.721	3.434	3.234	2.203	0.165	-0.095
$\langle 1\bar{1}1 \rangle_R$	3 layer	3.774	3.619	3.565	3.539	-0.002	4.000
	4 layer	3.733	3.372	3.160	2.286	0.130	2.452
	5 layer	3.706	3.371	3.279	3.542	-0.071	9.553
	6 layer	3.698	3.322	3.142	2.655	0.059	5.350
	7 layer	3.678	3.323	3.172	2.906	0.022	6.400
	8 layer	3.683	3.323	3.133	2.474	0.092	3.853
	9 layer	3.661	3.318	3.153	2.691	0.058	4.808
	10 layer	3.668	3.325	3.142	2.495	0.091	3.599
		$\tilde{\gamma}_{110}$					
pristine $\langle 110 \rangle$ surface	3 layer	4.332	-	-	-	-	-
	4 layer	2.194	-	-	-	-	-
	5 layer	3.319	-	-	-	-	-
	6 layer	2.313	-	-	-	-	-
	7 layer	2.927	-	-	-	-	-
	8 layer	2.373	-	-	-	-	-
	9 layer	2.739	-	-	-	-	-
	10 layer	2.382	-	-	-	-	-

4.4.2 $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_o$ steps

The γ_{110} solutions extracted from the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_o$ systems show excellent agreement with the $\tilde{\gamma}_{110}$ values which suggests that the q/d^3 interaction term in equation (4.1) was a good guess. This result shows that the $\langle 110 \rangle$ terrace component of the total high-index surface energy has been

accurately extracted for the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ steps for all slab thicknesses, leaving the corresponding solutions $\beta(L)$ and $q(L)$ open to physical analysis. Martinez *et al.*¹¹⁹ suggested from their total supercell energy calculations that the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ steps were very similar in energy. Figure 4.9 shows that the step energies per unit length are likely to converge towards about $\beta = 0.10 \pm 0.02 \text{ eV\AA}^{-1}$ for both steps as $L \rightarrow \infty$, where these precision uncertainties have been estimated by merely examining the plots around the region $L = 9,10$. They are indeed therefore very similar, at least up to $L = 10$, and so this suggests that both structures may be likely to occur on laboratory samples in equal proportions. Diebold observed some instances of reconstructions on $\langle 001 \rangle$ -type steps¹¹⁷ on annealed rutile (110) and this observation could perhaps be due to symmetry breaking caused by a stalemate in the formation of the energetically very similar $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ steps. The $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ structures clearly have very weak interactions with themselves, indicating that both structures induce a negligible strain field throughout the surface region of the crystal. The residual variations in the q solutions are of magnitude $\sim 1 \text{ eV\AA}$. It can be seen in Figure 4.11 that these q values correspond to a total interaction energy per supercell which lies well within the range $\pm 0.1 \text{ eV}$, implying a negligibly weak step-step interaction. These interaction energies are substantially smaller than all step and terrace energies, confirming that the series used in equation (4.1) is convergent. With regards these variations, it is not possible to reliably rule out the possible intrinsic sources of (small) uncertainty in the DFT calculations, including the slabs relaxing into metastable minima and the statistical uncertainty of the final energy values owing to the finite energy change tolerance of 10^{-5} eV per atom. It is therefore concluded that the interaction energies of the $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$ steps are zero to the level of precision attainable from these calculations.

4.4.3 $\langle 1\bar{1}1 \rangle$ and $\langle 1\bar{1}1 \rangle_R$ steps: general comments

Steps extended along the directions $\langle 1\bar{1}1 \rangle$ comprise the vast majority of all step defects on rutile (110), so determining which of these two structures is the most energetically stable is of particular interest. The solutions γ_{110}

extracted from the $\langle 1\bar{1}1 \rangle$ and $\langle 1\bar{1}1 \rangle_R$ systems recover the oscillatory manner of convergence but they display some slight deviations from the $\tilde{\gamma}_{110}$ data. The corresponding sets of supercells used for these two systems have identical dimensions, and they differ only in that they contain different arrangements of atoms at the steps. Therefore, besides considerations relating to the small uncertainty resulting from the geometry optimised energy values, the distinct behaviour of these two systems is completely attributable to the two distinct steps. The task at hand therefore is to attempt to itemise and disentangle the factors of physical origin in order to try and explain the discrepancies and gain an indication of the true energies of these steps. In the following two sections, the data for the $\langle 1\bar{1}1 \rangle_R$ step and $\langle 1\bar{1}1 \rangle$ step are discussed separately. Firstly, $\langle 1\bar{1}1 \rangle_R$:

4.4.4 $\langle 1\bar{1}1 \rangle_R$ step: detailed analysis

The trend in the residual differences $\Delta\gamma_{110}$ for the $\langle 1\bar{1}1 \rangle_R$ step show that the γ_{110} values are a very slight overestimation of $\tilde{\gamma}_{110}$ for the thickest slabs and severe underestimation of $\tilde{\gamma}_{110}$ for the $L = 3$ slabs. Two possible contributing factors of physical origin have been identified: *i*) interactions between neighbouring steps which were not accounted for by equation (4.1), and which may themselves depend on L in a non-simple way, and *ii*) surface-surface interactions between steps on opposing surfaces of the slab which also depend non-trivially on L . Both of these factors are likely to be making significant contributions to the true interaction energy and therefore producing a systematic error resulting in overvaluations and undervaluations of the γ_{110} solutions.

The obvious candidate for the origin of atomistic contributions to the interaction between neighbouring steps is the set of (341) slabs with the $\langle 1\bar{1}1 \rangle_R$ step. The small step width approximation is less well founded for these systems, which can be appreciated by referring to Figure 4.4 (d) which shows that the width of the step structure is significant compared to the size of the lower (110) terrace region. This is in contrast to the $\langle 001 \rangle$ systems for which the approximation is clearly a good one. In fact, there is an indication with the data available in Figure 4.8 that interactions between step on

opposing surfaces of the slab are likely to be the dominating factor, especially at small L . Observe the differences between $\gamma_{110}(L)$ and $\tilde{\gamma}_{110}(L)$ at $L = 3$: both the $\langle 1\bar{1}1 \rangle_R$ and the $\langle 1\bar{1}1 \rangle$ systems produce values of $\gamma_{110}(3)$ which are very large underestimates of $\tilde{\gamma}_{110}(3)$. The sets of $\langle 1\bar{1}1 \rangle_R$ and $\langle 1\bar{1}1 \rangle$ slabs at $L = 3$ of course have different steps, but they have small surface-surface distances in common. Furthermore, in the higher L range $L = 7, \dots, 10$, the agreement between γ_{110} and $\tilde{\gamma}_{110}$ for the $\langle 1\bar{1}1 \rangle_R$ step is quite good. The residuals $\Delta\gamma_{110}$ at large L are just about small enough to be plausibly accounted for by the intrinsic uncertainties in the supercell energies. The indication therefore is that choosing the q/d^3 interaction term in equation (4.1) results in an accurate simulation of the interactions between $\langle 1\bar{1}1 \rangle_R$ steps, but only once surface-surface interactions have been decoupled at large slab thicknesses. The instances where β is predicted to be negative ($L = 3, 5$) indicate that equation (4.1) fails for the purposes of extracting the energetics of neighbouring steps because of the significant influence of surface-surface step interactions in the thinner slabs.

In summary, Figure 4.9 shows that β for the $\langle 1\bar{1}1 \rangle_R$ step on macroscopically thick samples is likely to fall approximately within the range $\beta = 0.07 \pm 0.02 \text{ eV\AA}^{-1}$ for $L \rightarrow \infty$, and Figure 4.10 similarly shows that these steps on vicinal surfaces exhibit a repulsive interaction consistent with a positive interaction constant q falling in the range of about $4 \pm 2 \text{ eV\AA}$ as $L \rightarrow \infty$.

4.4.5 $\langle 1\bar{1}1 \rangle$ step: detailed analysis

Now observe the γ_{110} solutions for the $\langle 1\bar{1}1 \rangle$ step: some variation remains, but each and every one of the 8 solutions γ_{110} is undervalued compared to $\tilde{\gamma}_{110}$. Surface-surface interactions are probably making major contributions at small L due to the clear disagreement of $\gamma_{110}(3)$ with $\tilde{\gamma}_{110}(3)$ as explained in the previous section. However the underestimation of the value of γ_{110} at $L = 10$ value is probably too large for intrinsic the uncertainty in the supercell energy to be a plausible cause. Besides the possibility of an unlikely and spurious cancellation of other unconsidered factors, this indicates that atomistic interaction terms may be contributing to the interaction energy for

the $\langle 1\bar{1}1 \rangle$ step, resulting in a systematic error originating from the false assumption that the step interaction is completely described by a term of the form q/d^3 . To provide an approximate indication of what the true overall interaction scaling behaviour is, equation (4.1) was rewritten with an interaction term of the form q/d^p and successive iterations on the value of p were carried out to minimise the sum of the squares of the residual differences between the solutions $\gamma_{110}(L)$ and the values $\tilde{\gamma}_{110}(L)$ over the range $L = 7, \dots, 10$. This range was chosen in a basic attempt to try and filter out surface-surface interactions. The best fit result was $p = 2.14$ (3 s. f.), consistent with an overall step-step interaction energy which scales as $1/d^{1.14}$. (This was then cross-checked by carrying out the same procedure on the $\langle 1\bar{1}1 \rangle_R$ data, and the result was $p = 3.13$, implying an overall interaction energy scaling behaviour of $\sim 1/d^{2.13}$ which is in good agreement with the originally nominated $\sim 1/d^2$ MP interaction potential.) The corresponding values of β for $\langle 1\bar{1}1 \rangle$ with this $q/d^{2.14}$ interaction term remained relatively unchanged and were still substantially larger than those from the $\langle 1\bar{1}1 \rangle_R$ system. Therefore, no pivotal implications for the theoretical probing of STM images emerge from this because the $\langle 1\bar{1}1 \rangle_R$ step is still implied to be vastly energetically favourable.

4.4.6 Comparison with published STM data

In Figure 4.13, STM data from Martinez *et al.*¹¹⁴ is reproduced which shows the annealed rutile (110) surface. Alongside this, a series of 2D Wulff constructions with increasing L is presented in which two island morphologies have been assumed: *i*) a (110) terrace bound by $\langle 1\bar{1}1 \rangle$ steps and $\langle 001 \rangle_{Ti}$ steps and *ii*) a (110) terrace bound by $\langle 1\bar{1}1 \rangle_R$ steps and $\langle 001 \rangle_{Ti}$ steps. By drawing vectors from a common origin whose directions are perpendicular to the step in question and whose lengths are drawn in proportion to the corresponding β values, the island shapes predicted to occur on the surface as a function of increasing slab thickness can be directly compared with the island seen in the STM image. An island of unquestionable similarity on an annealed rutile (110) sample was also observed by Diebold *et al.*¹¹⁷ It is evident that the island shape predicted with

a $\langle 1\bar{1}1 \rangle / \langle 001 \rangle_{Ti}$ morphology is inconsistent with the observed island shape for all values of L .

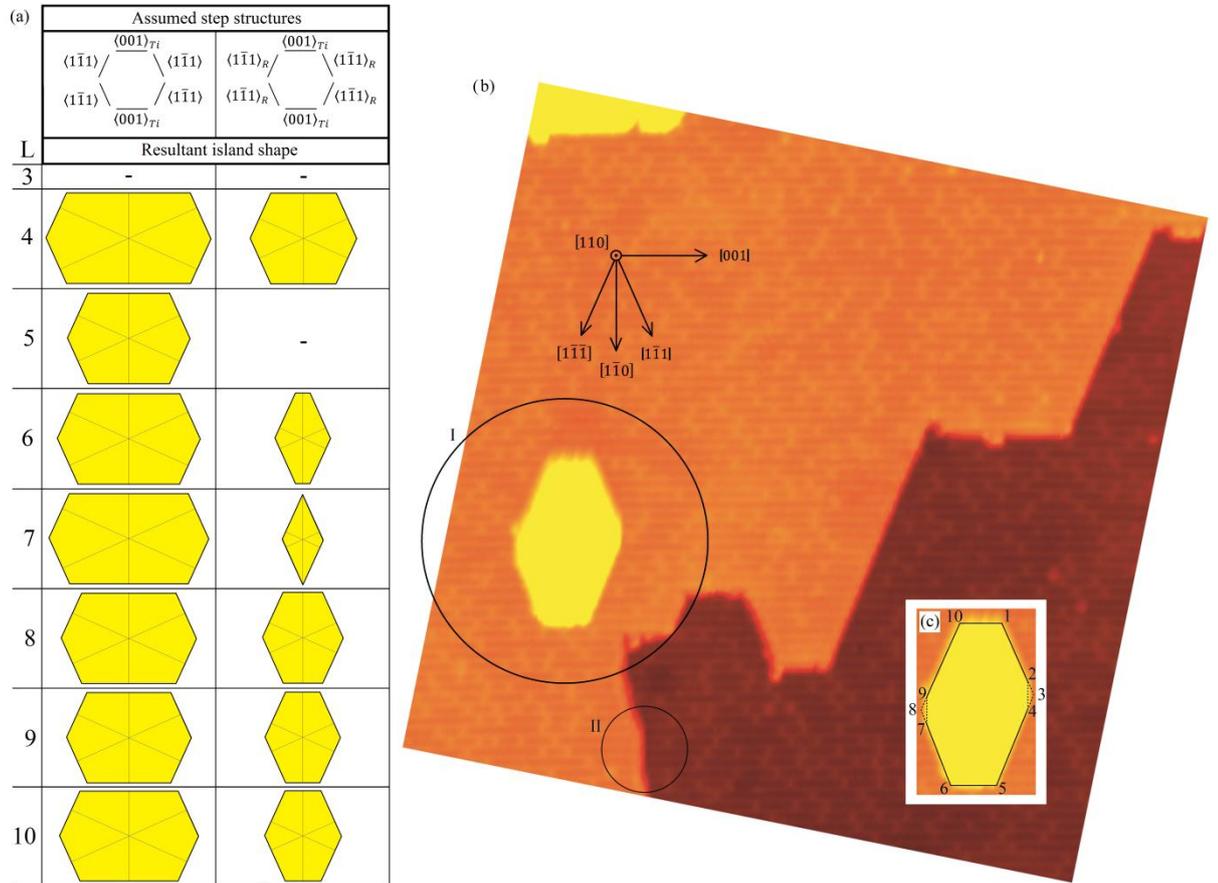


Figure 4.13. Experimental scanning tunnelling electron microscopy data.

(a) Terrace island shapes predicted using Wulff constructions in 2 dimensions using the $\beta(L)$ values ($L = 3, \dots, 10$) plotted in Figure 4.9 assuming a (110) terrace bound by i) $\langle 001 \rangle_{Ti}$ with $\langle 1\bar{1}1 \rangle$ steps and ii) $\langle 001 \rangle_{Ti}$ with $\langle 1\bar{1}1 \rangle_R$ steps. (b) Terrace island on rutile (110) surface (circle I.) observed using STM carried out in a UHV chamber at 120 - 140 K in constant current mode with a tunneling voltage of approximately + 1.25 V. An instance of an apparent $\langle 1\bar{1}0 \rangle$ -type step is indicated in circle II. (c) Further features of the island not accounted for by this work. The kinks labeled 1 - 10 and their locations relative to each other show the apparent chirality of the island, and the possibility of $\langle 1\bar{1}0 \rangle$ -type edges existing on the island. STM image reproduced from Ref. 114 with kind permission of S. Wendt.

In contrast, the $\langle 1\bar{1}1 \rangle_R / \langle 001 \rangle_{Ti}$ morphology (or indeed $\langle 1\bar{1}1 \rangle_R / \langle 001 \rangle_O$, which are not drawn but look very similar) produces an island shape which converges quite convincingly to that seen in experiments^{114,116,117} by

approximately 8 layers thickness and beyond. The predictions are clearly less reliable for slab thicknesses of less than about 7 layers. This demonstrates the predictive power of the methodology and adds credibility to the existence of the $\langle 1\bar{1}1 \rangle_R$ step structure on laboratory samples. Instances which included negative β values for one or both steps (3 and 5 layer cases) were not included in the Wulff constructions.

It is clear that there are features on the surface shown in Figure 4.13 (b) which are not predicted or accounted for by the level of theory presented in this chapter. It is necessary to review these features to validate any claims of correspondence between theory and experiment, and also to try to anticipate how this approach may be refined. Firstly, by basing this approach on surface energies calculated via reference to a stoichiometric bulk state as in equation (4.3), any deviations from stoichiometry have not been accounted for. It is well established¹¹⁴⁻¹¹⁸ that the ion sputtering / high-temperature vacuum-annealing preparation process causes rutile (110) samples to become slightly reduced, with the creation of a small but non-zero concentration point defects in the form of bridging oxygen vacancies on the surface and Ti interstitials in the bulk of the sample. The O vacancies are widely believed to significantly affect the surface chemistry of rutile (110) compared to stoichiometric surfaces due to the resultant under-coordinated surface Ti atoms at these vacancy sites. Such vacuum-annealed samples typically show O vacancy coverages of about 2 - 12% depending on the preparation conditions.¹⁶⁰ The exact physical process which causes the vacancies to appear during annealing is not known, although molecular desorption experiments combined with STM have been used to characterise these defects¹⁶¹ and shed new light on their origin.¹⁶² It is also well known that surface vacancies are healed by exposing the sample to O₂, although, depending on the temperature and partial O₂ pressure, this process is usually accompanied by further adsorption of O to surface 5f-Ti sites¹⁶³ and the appearance of additional O-based surface adstructures.^{117,160,164,165} It is worth mentioning that there is one STM study which apparently shows a perfectly stoichiometric rutile (110) surface.¹⁶⁶ Despite the sensitive dependence of the surface structure on the preparation conditions, the fact that the surface vacancies are healed upon exposure to O₂ confirms that the

stoichiometric surface is energetically favourable over O-reduced surfaces as one might reasonably expect.

Alongside surface O vacancies, the influence of bulk Ti interstitials on the appearance of the (110) surface must also be considered. Investigations into the behaviour of Ti interstitials and their implications for the surface chemistry of rutile are still ongoing,^{159,167} and it is not yet entirely clear whether Ti interstitials located in the bulk of the material have any significant consequences for the energetic stability of step structures occurring on the surface. There is recent evidence which suggests that bulk Ti interstitials do not influence (110) terraces in a significant way: In a recent study by Bechstein *et al.*,¹⁶⁷ the Ti interstitial defect formation energy was calculated at (110) terrace sites and found to be 1.2 eV higher than the formation energy at bulk Ti interstitial sites. This suggests that for mostly flat (110) samples, like that shown in Figure 4.13 (b), Ti interstitials are likely to quickly diffuse away from the surface and into the bulk of the material. O vacancies are therefore likely to be the dominant form of point defects which contribute to the overall surface chemistry on (110) terraces with low step density. However, it is not completely clear as to whether Ti interstitials are repelled by steps in the same way as they are by (110) terraces. In Bechstein's article,¹⁶⁷ it was theoretically predicted that for rutile surfaces with a high step density, Ti interstitials are *more* stable at $\langle 1\bar{1}1 \rangle_R$ step sites than they are at bulk sites by about 0.5 eV. This prediction was then used to argue in favour of a mechanism by which O-deficient strand structures extended along $\langle 001 \rangle$ directions are formed during annealing on high-index surfaces, where Ti interstitials migrate to the surface to form the strands, with the $\langle 1\bar{1}1 \rangle_R$ steps as their exit points. This was corroborated with STM data which showed a rutile (771) surface of high step density after annealing at about 900 K, where strand structures extended along $\langle 001 \rangle$ directions were seen to be protruding from the steps. A mostly flat (110) surface was shown in the same article, in which only a tiny number of such strand structures were seen. The authors argued quite plausibly that more building material is available on surfaces with a high step density due to the larger proportion of undercoordinated surface atoms, which explains the propensity of highly-stepped surfaces to form such strands. In addition to this, it is also

speculated in this thesis that the net diffusion tendencies of Ti interstitials from the surface may also be a question of the competition between the repulsive influence of (110) terraces and the attractive influence of steps, and therefore dependent on step density as well as temperature. With slowly decaying surface-surface interactions, gaining a fuller picture of the energetic landscape and migration pathway barriers for Ti interstitials is not likely to be a trivial task using a DFT approach. Full DFT calculations of Ti interstitial migration paths and preferred binding sites as a function of surface roughness would require supercells which are currently prohibitively large, especially for surfaces of low step density. Nevertheless, the good agreement of the stoichiometric theoretical models used here with experimentally observed slightly nonstoichiometric samples suggests that the O vacancies and Ti interstitials can be regarded as an artefact of the annealing process which otherwise have no effect on overall island shapes.

Secondly, on visual inspection, it appears that the left and right edges of the island are bevelled by $[1\bar{1}0]$ -type steps, where the corresponding kinks are labelled 2, 4, 7 and 9 in a rudimentary sketch in Figure 4.13 (c). Another instance of an apparent $[1\bar{1}0]$ -type structure is highlighted in the smaller of the two circled regions in Figure 4.13 (b). The apparent $[1\bar{1}0]$ -like features seen on the island are tiny; no more than 2 bright Ti rows in extent, or 4 atoms. Because of the periodic boundary conditions used in the calculations in this chapter, the simulated steps are infinitely long. This means that these calculations do not account for subtleties such as the free energy of kinks, their interactions among themselves and with steps, and the atomic structure of kinks. This therefore means that the step energies can only reliably be compared to steps which are of appreciable length in STM images, at least for the purposes of using Wulff constructions. It is therefore doubtful that an infinitely long $\langle 1\bar{1}0 \rangle$ -type step under periodic boundary conditions would be representative of the $[1\bar{1}0]$ -like structure which is only 4 atoms wide, the atomic configuration and energy of which will be largely determined by the fact that at least 2 of these atoms must form bonds with the two adjoining steps. It is also worth pointing out that the island observed by Diebold *et al.*¹¹⁷ did not exhibit these apparent $[1\bar{1}0]$ -type steps. In any case, the $[1\bar{1}0]$ -like features comprise a negligible proportion of the total population of steps

across the whole image and so they were deemed too insignificant to warrant further investigation.

Thirdly, the island habit has an apparent chirality, in that the kinks labelled 3 and 8 (or equivalently, 2, 4, 7 and 9) are not located exactly opposite each other along the direction [001]. This chirality is also quite apparent in Diebold's image.¹¹⁷ The main comment to be made on this is that the Wulff construction is a scheme for predicting the habit of crystals in the macroscopic regime, whereas the island is microscopic in size. It is known that the bright rows observed with STM correspond to the 5-fold coordinated Ti atoms and the dark rows are bridging O atoms¹⁶⁶ (see Figure 4.1). With close inspection of the lower terrace from the STM image of Figure 4.13 (b), it can be verified that the island spans no more than about 15 such rows, and therefore 15 surface cells. Given that the island's entire extent is only one order of magnitude larger than the size of one of its irreducible surface cell component parts, and that no account has been made of the possible atomic arrangements at the kinks, it is entirely plausible that the experimentally observed island habit at such microscopic scales is subject to aliasing of this kind, resulting in kinks 3 and 8 being misaligned by one or two unit cells. It is also unsurprising due to initial small and random configurational perturbations almost certainly being present in the initial growth environment. Fourthly, no accounts of the energetic contributions of the kinks or step-step interactions were accounted for by the Wulff constructions.

4.5 CONCLUSIONS

A DFT/supercell methodology for calculating the terrace free energy per unit area, step free energy per unit length and step-step interaction energy of high-index vicinal surfaces of crystalline solids using slab models has been presented. The slow oscillatory convergence of the individual energetic components (terrace, step and step-step interactions) of vicinal surfaces with increasing slab dimensions due to QSE has been demonstrated for the case of TiO₂ rutile. For such systems in which QSE are strong, it has been argued that using a statistical linear fit of total slab energy against increasing slab

size is an unreliable and misleading approach to calculating surface energies, and therefore an unreliable way of evaluating the energies of steps. For this reason, it is argued instead that step energies and step interaction energies should all be individually plotted against slab size and checked for consistency via a layer-by-layer cross-check of the low-index terrace surface energy, calculated separately using the appropriate low-index slabs. This method has been applied to two $\langle 1\bar{1}1 \rangle$ -direction steps and two $\langle 001 \rangle$ -direction steps on the rutile (110) surface using DFT with the GGA PBE functional, following published STM studies¹¹⁴⁻¹¹⁸ which have repeatedly shown that step defects are extended solely along the directions $\langle 1\bar{1}1 \rangle$ and $\langle 001 \rangle$. The specific atomic arrangements of the steps studied were those recently predicted to be the most energetically stable overall by Martinez *et al.*¹¹⁹ The two $\langle 001 \rangle$ -direction steps, $\langle 001 \rangle_{Ti}$ and $\langle 001 \rangle_O$, were found to be very similar in energy and so it was concluded that these two structures are likely to occur on laboratory samples with approximately equal probability. The $\langle 1\bar{1}1 \rangle_R$ structure was found to be significantly more stable than the $\langle 1\bar{1}1 \rangle$ structure. 2D Wulff constructions were used with the theoretically calculated step energies to show that the shape of islands with $\langle 1\bar{1}1 \rangle_R$ steps compare well with STM data, whereas the shape of an island with $\langle 1\bar{1}1 \rangle$ steps does not stand up to experimental evidence.

Chapter 5: Ordering of substitutional N dopants in graphene

5.1 Abstract

Recent electronic structure studies of substitutional N-doped graphene have shown that the energetic and electronic properties depend sensitively on how the dopants are arranged.^{168,169} This has led to suggestions that the band structure of graphene could be systematically manipulated by tailoring the arrangements of N dopants according to various superlattice and sublattice arrangements. Using the hybrid exchange-correlation functional of Heyd, Scuseria and Ernzerhof (HSE)¹⁷⁰, Park *et al.*¹⁶⁹ predicted that when two nearby N dopants are placed in the same sublattice, a non-zero band gap results at the Dirac point, whereas no such band gap is created for two nearby dopants placed in different sublattices. In addition to these recently published theoretical suggestions, experimental STM^{3,4} and STEM⁵ studies have recently emerged which show an apparent tendency for N dopants to be mainly observed in only one of the two carbon sublattices. A representative illustration is given here in Figure 5.1. This is very significant because successful fabrication of potential graphene-based devices which derive their functionality from sublattice/superlattice dopant configurations, as alluded to by Park *et al.*¹⁶⁹, would rely on the ability to harness energetically favourable processes resulting in the desired dopant arrangements. This is also interesting from a purely scientific point of view because it suggests that currently unidentified physical effects might be influencing their distribution. To try to identify any possible physical causal factors which may be causing N dopants to collect into the same sublattice during fabrication, spin-polarised density functional theory (DFT) calculations have been performed of the total energies of substitutional N-N configurations over an exhaustive range of initial spin states using a range of exchange-correlation functionals.

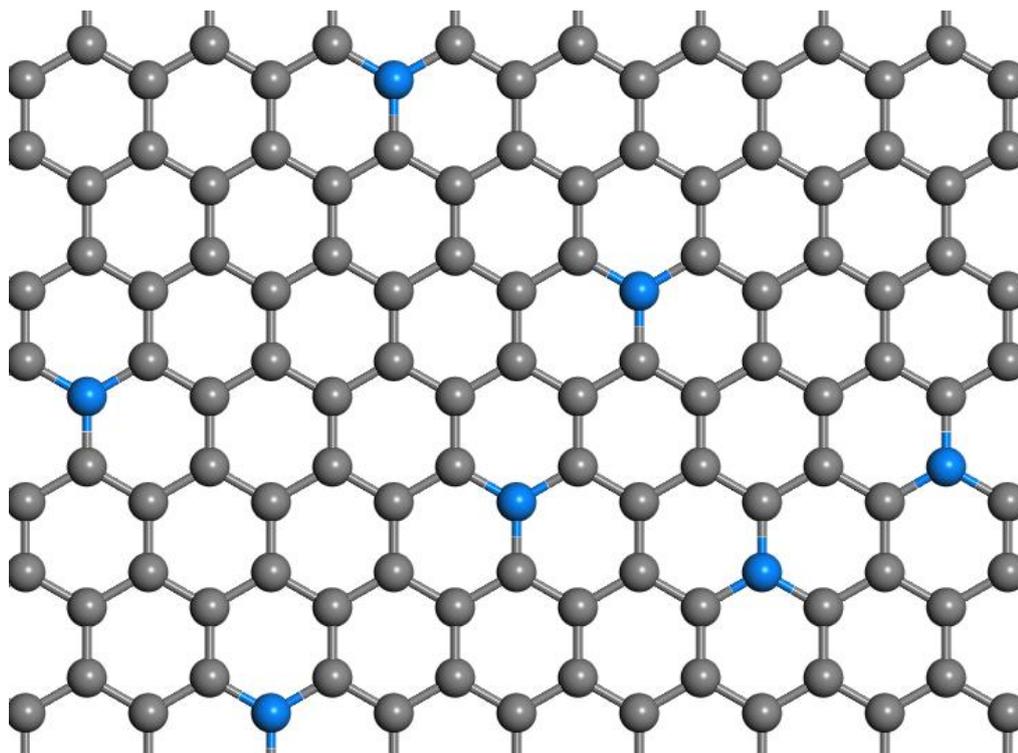


Figure 5.1. A ball-and-stick illustration of substitutional nitrogen dopants in graphene, with nitrogen indicated in blue. In this example, 5 out of 7 of the dopants lay in the same sublattice.

At the rather coarse level of precision achieved in the calculations so far performed, all results for all functionals over all spin states used show that N dopants experience no significant energetic advantage by existing in one particular sublattice. In fact, the configurations in which the nitrogen atoms are in different sublattices are found to be consistently and slightly energetically favourable. This suggests, but by no means proves, that the experimental observations might be attributable to a phenomenon which is unrelated to the intrinsic energetics of the various configurations. At the end of the chapter, current ideas for refining the calculations and advancing the project are outlined.

5.2 INTRODUCTION

5.2.1 Experimental observations and suggested theoretical investigations

Tuning the band gap in graphene using substitutional dopants to create n-type and p-type semiconductors is currently one of the most important objectives in graphene research. n-type modified graphene containing substitutional N dopants has recently been synthesised using various techniques and observed using atomic-scale STM^{3,4} and STEM.⁵ One very interesting peripheral observation from these studies is that when many N substitutional atoms are visible in the same image and on the same grain, the majority of them are often found to be occupying sites in only one of the two carbon sublattices. Specifically, see Figure S6 from Ref. 3, Figures 2c and S3 from Ref. 4, and Figure 4b from Ref. 5.

In Figure S6 from Ref. 3, there are 9 substitutional nitrogen dopants in total in the image, 8 of which lay in the same sublattice. In Figure 2c from Ref. 4, there appear to be 13 identifiable N dopants which *all* lay in the same sublattice (although the authors claim that there are 14 in total). In Figure 4b from Ref. 5, 5 out of the 5 dopants are in the same sublattice. In Figure S3 from Ref. 4 (which has two panels), the ratios are less drastic: 13 / 23 and 27 / 39 N dopants lay in the same sublattice in the two images shown. However, the visual evidence from Figure S3 of Ref. 4 strongly suggests that N dopants do apparently show sublattice ordering at short range (which the authors acknowledge), giving rise to “domains” in which an apparent sublattice preference occurs. At the time of writing, two collaborators (Q. M. Ramasse, D. Kepaptsoglou) are not completely convinced that such domains exist. Given the currently limited amount of experimental data, this is an entirely reasonable point of view. Another important question is whether these ratios are *statistically representative of the population of all dopants over the entire grain of the sample* in each case. This question is very difficult to answer. A self-contained theoretical approach could shed new light on this situation, which is the motivating premise of this chapter.

One way to try to identify whether a specific phenomenon may be responsible is to use DFT. The challenge faced by any practitioner of DFT when approaching a new problem is to motivate the choice of XC functional used in the calculations. For systems in which the phenomenon which contributes to a particular aspect of the observed behaviour is not properly understood or even identified, a sensible approach is to use many functionals at progressively higher levels of theory, and compare the results. Doped semiconductors, by design, are characterised by unpaired electrons or holes. With this conventional wisdom, the probable existence of unpaired spins originating from the N dopants in the ground state is declared as a possible origin of the observed effect. This initially leads to the hypothesis that the tendency of substitutional N dopants to collect into one carbon sublattice is caused by a significant energetic advantage associated with particular lattice configurations caused by long-ranged interactions between the unpaired spins from the N dopants. Interactions in the form of long-ranged ferro- or antiferromagnetic spin orderings are investigated using the LDA, PBE, HSE¹⁷⁰ and sX-LDA¹⁷¹ functionals over an exhaustive range of spin states. However, all of the results of these calculations indicate that the hypothesis is false; no energetic advantage of collective sublattice binding was found to emerge from any of these investigations. This theoretical evidence suggests either that the apparently observed effect might emerge during fabrication, that effects unrelated to magnetism may be significant, or that the published experimental data gives an incomplete picture.

5.3 METHOD

5.3.1. Basis set and computational considerations

The purpose of this chapter is to compare the energetic stability of various configurations of N dopants in graphene using 6×6 supercells. The figure of merit, therefore, is the difference between total energies of supercells containing the relevant configurations. For this purpose, the plane-wave density functional theory code CASTEP⁸³ was used with norm-conserving pseudopotentials¹⁹ to converge the differences between total spin-polarized energies of two C_6N_2 2×2 supercells, designed to resemble the 6×6

supercells used in the full calculations, with respect to the plane-wave kinetic energy cutoff, E_{cut} , and Monkhorst-Pack k-points grid¹⁷ spacings, s^i . Density mixing was used for the electronic minimisation in all cases for these 2×2 supercells, and all systems were treated as metallic by using variable electronic occupancies when finding the ground state. It was found that the relevant energy differences changed by no more than 0.021 eV upon increasing the cutoff energy from 540 eV to 800 eV for all four functionals and two dopants, and E_{cut} was fixed at $E_{cut} = 540$ eV for all subsequent calculations. The other limitation in the precision of the energies in this chapter originates from the k points grid. In a plane wave basis set, the hybrid functional HSE and the non-local functional sX-LDA require an *enormous* amount of computer memory which scales non-linearly with the number of k-points. Upon decreasing the in-plane k-point spacings from 0.078 \AA^{-1} down to 0.039 \AA^{-1} (by increasing the number of k-points with fixed supercell size), supercell energy differences were found to consistently differ by no more than ± 0.1 eV *with identical k-points sampling* on the supercells shown in Figure 5.2 (b) for the six functionals. Figure 5.2 (a) shows this energy convergence. A k points grid of a single gamma point was therefore selected for use in all 6×6 supercell calculations, corresponding to in-plane k-point spacings of $s^i = 0.078 \text{ \AA}^{-1}$ (3 d. p.). This means that all energies calculated using the same functional can be reliably compared up to a precision of ± 0.1 eV, but finer comparisons within this energy range should be treated with caution. To ensure that this level of precision was maintained, all supercells to be used in the full calculations were fixed at $P1$ symmetry. This choice of $P1$ symmetry is a very important step in optimizing the precision, because any given pair of supercells containing two dopant atoms rarely have the same lattice symmetries, and imposing symmetry operators usually changes the k-points sampling. Using a single gamma point with $P1$ symmetry therefore facilitates identical k-points sampling across all supercells, resulting in reliable comparisons of total supercell energies. It was found that the alternative approach of using a finer k-points grid and restoring lattice symmetry operators does not provide a practical advantage; the extra computational cost associated with using more than one k-point is much greater than that incurred by disabling the symmetry

finders and using a single gamma point, and the resultant precision is less consistent and more difficult to determine.

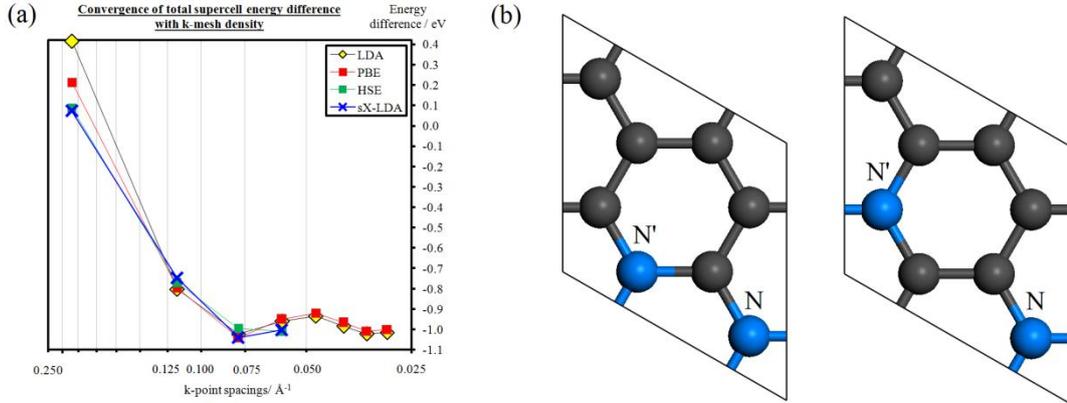


Figure 5.2. Convergence of supercell energy differences with k points sampling.

(a) The convergence of the difference in total energy of two supercells with different configurations of two nitrogen dopants. (b) The two supercells used. In both cases, the initial spin value of N was set at $1\uparrow_B$, and that of N' was $-1\uparrow_B$.

For the pristine graphene unit cells, the required vacuum thickness $|\vec{c}|$ and unit cell lattice parameters $|\vec{a}|$ were then evaluated for the four functionals (in that order, and using finer k-spacings for these calculations only). For all four functionals, the total system energy changed by no more than 0.03 eV upon extension of the vacuum thickness from 8 Å to 13 Å. The vacuum was fixed at $|\vec{c}| = 8$ Å. As expected, the relaxed lattice parameters $|\vec{a}|$ were found to depend on the functional. The results were $|\vec{a}|_{LDA} = 2.457$ Å, $|\vec{a}|_{PBE} = 2.456$ Å, $|\vec{a}|_{HSE} = 2.431$ Å and $|\vec{a}|_{sX-LDA} = 2.442$ Å, all to 3. d. p.

5.3.2. Calculation details: Single N dopant

6×6 graphene supercells were constructed, in which the C atom in the centre was substituted for a N atom. Unlike the calculations from the 2×2 supercells, the electronic minimisations were performed on all 6×6 systems using the ensemble DFT¹⁷² (EDFT) approach rather than density mixing. The EDFT approach, which is much slower, was resorted to because it was

found to be significantly more stable with respect to total energy convergence for these larger supercells. The sX-LDA calculations proved to be particularly difficult to converge using density mixing for the 6×6 systems in some, but not all cases, with no apparent predictability. One of the disadvantages of using EDFT is that magnetic moments cannot be assigned to individual atoms; only to the entire supercell. $P1$ symmetry was retained during all electronic minimisations, in which the initial total magnetic moment of the supercell was set at the 3 initial values, $-1, 0, +1$ (in units of Bohr magnetons, μ_B). With the 4 functionals, this resulted in a total of 12 calculations. This range of initial magnetic moments was used in order to significantly increase the likelihood of reaching the spin-polarized ground state by distinguishing between local and global minima. The supercell lattice parameters $|\vec{B}| = |\vec{A}| = 6|\vec{a}|$ were fixed in each case according to the unit cell lattice parameters derived using the relevant functionals in the previous section. Because of the vast computational cost incurred with the hybrid and non-local functionals, no structural relaxation was allowed to occur; only the electronic minimizations were performed. This approach was taken to allow for valid comparison of the results found using different functionals, despite full geometry optimisation being well within the practical reach of the vastly more efficient LDA and PBE functionals. The relative total supercell energies - calculated relative to the final energy of the initial-zero-spin calculation for each functional - were tabulated in Table 5.1, along with the root mean square (RMS) of the residual inter-ionic forces and final total magnetic moments.

5.3.3. Calculation details: Two N dopants

6×6 supercells of chemical formula $C_{70}N_2$ were prepared, in which the site A0 was always occupied by a substitutional N atom, and in which a C atom laying at one of the 15 sites A1, ..., A7 (sublattice **A**), B1, ..., B9 (sublattice **B**), as shown in Figure 5.3, was substituted for a second N atom, thus resulting in 15 unique $C_{70}N_2$ configurations. In a similar manner to the single dopant calculations described in the previous section, the total magnetic moment of each of these 15 supercells was set at three initial values:

$\mu_B = -2, 0, +2$, to make 45 calculations for each of the 4 functionals. CASTEP was used to carry out EDFT electronic minimizations for all 180 supercells.

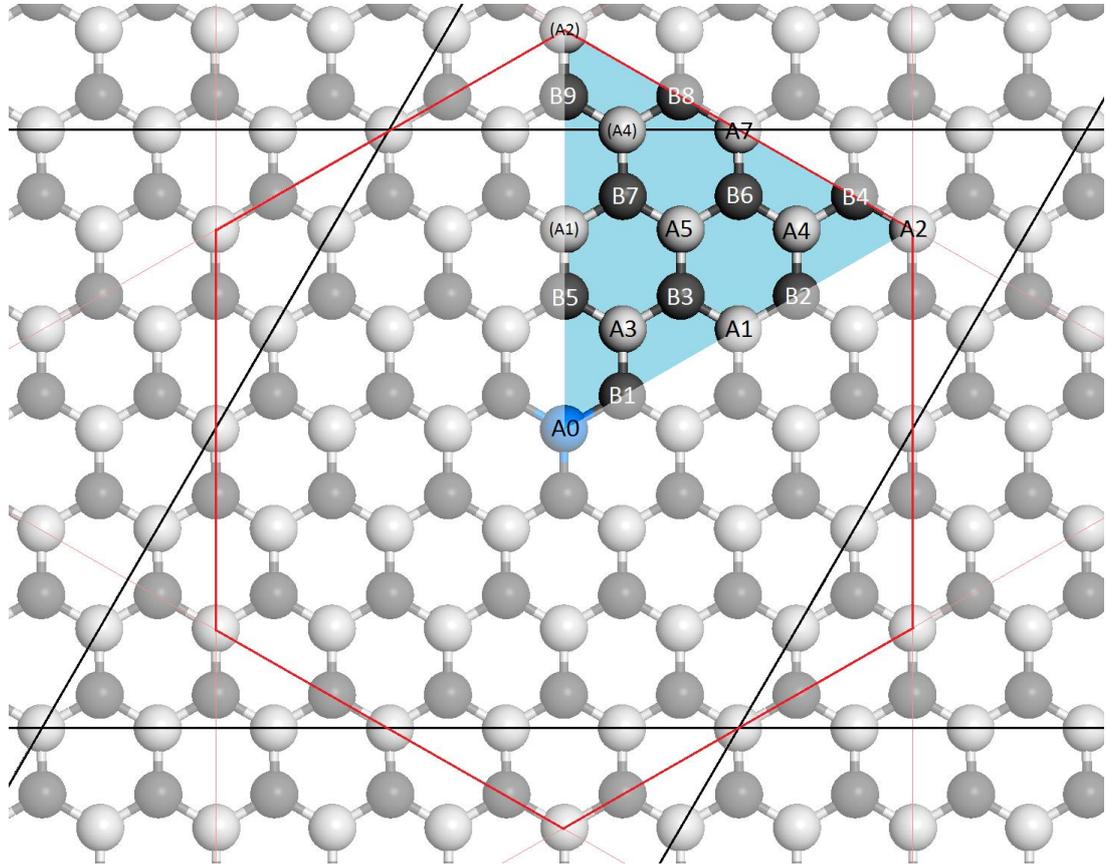


Figure 5.3. *The irreducible Wigner-Seitz cell, whose boundaries are indicated in red, of the 6x6 supercell showing all unique locations for the second substitutional N dopant.*

The periodic supercell boundaries are indicated in black. The first N dopant is located at site A0, and there are 15 unique possible locations for the second N dopant, labelled A1 - A7 and B1 - B9, where A and B distinguish between the two carbon sublattices. The blue shaded region indicates the irreducible Wigner-Seitz cell. Sites denoted (A1), (A4) and (A2) are equivalent to sites A1, A4 and A2 respectively by symmetry.

The notation is used where, for example, A5(+2) refers to the A2 configuration supercell whose total initial magnetic moment is $\mu_B = +2$. The inspiration behind the choices of $-2, 0, +2$ for the initial value of μ_B is that these values would constitute the range of total initial magnetic moments obtained if all possible combinations $\mu_B = \pm 1$ were applied to the two N

dopants. The relative total energies of the final electronically-minimized configurations were resolved into the two sublattices **A** and **B** and initial spin states and plotted against N-N distance in Figure 5.5. The dependence of the final energies on the final total magnetic moments of these configurations were plotted in Figure 5.6.

5.4 RESULTS

5.4.1. Single N dopant: initial remarks

The PBE, HSE and sX-LDA calculations all converged to states with magnetic moments whose absolute magnitudes are very close to $1 \mu_B$, each of which can be reliably interpreted as corresponding to one unpaired electronic spin state. In contrast, the LDA calculations predict no significant magnetic moment at all. In the opinions of two collaborators (K. Refson and B. Montanari) the discrepancy between the final LDA and PBE magnetic states is an unusual and surprising result which should be investigated further. Another pressing remark to be made is that the HSE results are *highly suspect*. The HSE functional is the only functional to predict that states with different final magnetic moments ($+1.000 \mu_B$ (3 d.p.) and $-1.000 \mu_B$ (3 d.p.)) have final energies which differ by over 0.5 eV. This is a rather peculiar result and casts doubt on these particular results. If these HSE calculations are indeed robust, then it follows that they have arrived at two non-equivalent energy minima despite having arrived at states whose magnetic moment is of identical magnitude, which would require an unusual explanation.

Generally, there is no correlation between the initial dopant magnetic moment and magnitude of the final total magnetic moment, or its sign. This immediately exposes the sensitivity of the final ground state spin configuration to the chosen functional and initial spin state, and therefore reiterates the importance of using many functional over a range of spin states and comparing the results.

The residual inter-ionic forces are all small and approximately the same. This confirms that the unit cell lattice parameters derived in section 5.3.1

from the unit cell geometry optimisations were all calculated precisely, and more importantly, that the significant extra computational cost of optimising the structures to minimise the residual forces would probably be unlikely to change the relative total supercell energies significantly, although of course this cannot be guaranteed.

Examples of the spin density fields taken from the four calculations with initial magnetic moment of $+1\mu_B$ are shown in Figure 5.4. The strongest clue which relates to preferential sublattice binding is that the spin density shows a clear and dramatic asymmetry with respect to the two carbon sublattices. Note the different colour scales applicable to each panel; the sublattice asymmetry with respect to spin is predicted to be very subtle by the LDA, moderate by PBE and HSE, and comparatively dramatic by sX-LDA.

Table 5.1. Single N dopant: final magnetic moments, energies and forces.

Substitutional dopant	XC Functional	Initial total magnetic moment / μ_B	Final total magnetic moment / μ_B (3 d. p.)	Relative total energy / eV (3 d. p.)	RMS forces / $eV\text{\AA}^{-1}$ (3 d. p.)
N	LDA	+1	0.037	0.000	0.256
		0	-0.022	$\stackrel{\text{def}}{=} 0$	0.256
		-1	-0.033	0.000	0.256
	PBE	+1	-0.978	0.000	0.260
		0	-0.976	$\stackrel{\text{def}}{=} 0$	0.260
		-1	-0.977	0.000	0.260
	HSE	+1	1.000	+0.005	0.309
		0	0.999	$\stackrel{\text{def}}{=} 0$	0.301
		-1	-1.000	+0.516	0.310
	sX-LDA	+1	1.000	0.000	0.251
		0	-1.000	$\stackrel{\text{def}}{=} 0$	0.251
		-1	1.000	0.000	0.251

The common theme with the spin density distributions shown in Figure 5.4 is the dramatic asymmetry with respect to the two carbon sublattices. The LDA, HSE and sX-LDA plots show that an excess of up spin has delocalised across the sublattice in which the N dopant lies, where an excess of down spin has delocalised across the other sublattice. The PBE plot shows the opposite. Note the very different values used on the four colour scales. If

these colour plots are made absolutely comparable, then the only plot which shows any significant sublattice spin density alternation is the sX-LDA plot, with slight alternation visible for HSE, and nothing at all visible for the LDA and PBE plots. Beyond this, it is clear that there is further detail in the spin fields which requires further investigation. For example, there is one significant qualitative difference which can immediately be identified between the spin field found using the sX-LDA functional and those from the LDA, PBE and HSE functionals. Observe the sX-LDA field from Figure 5.4 (d): all carbon nuclei lying in the N dopant's sublattice have an approximately *equal* accumulation of spin on them. In contrast, in the LDA, PBE and HSE spin fields from Figures 5.4 (a), (b) and (c), *two thirds* of all of the carbon nuclei which lie in the same sublattice as the N dopant have *more spin localised on them than the other one third* of all of the carbon nuclei from that sublattice.

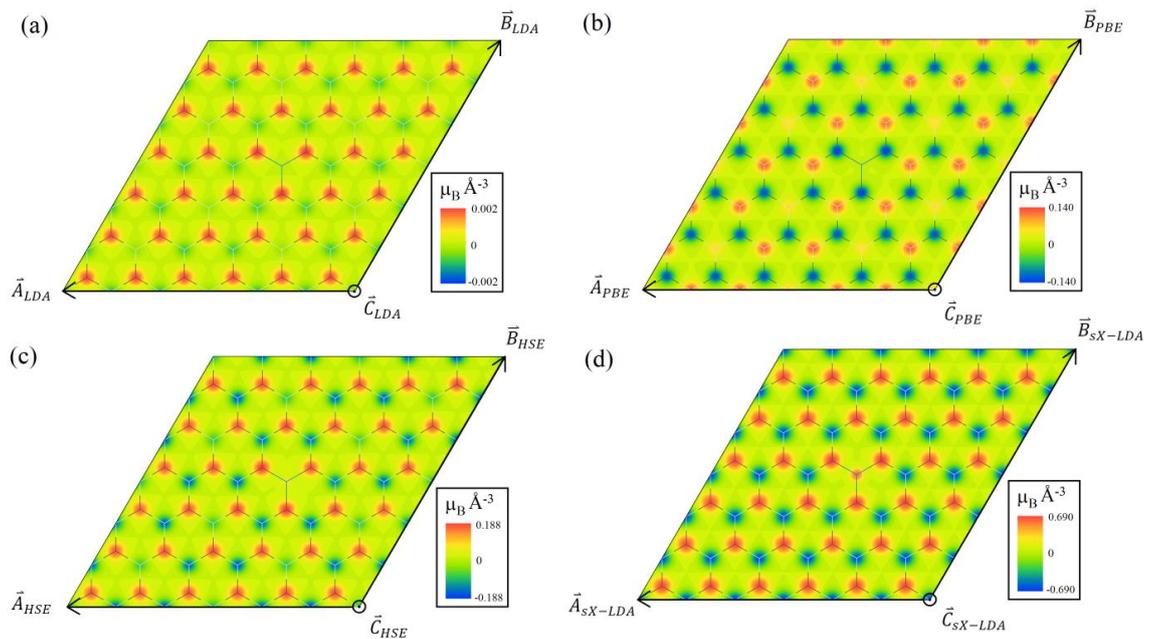


Figure 5.4. Spin density slices showing sublattice asymmetry for (a) LDA, (b) PBE, (c) HSE and (d) sX-LDA.

In all supercells, the blue vertex in the centre denotes the location of N nucleus, and the white and black vertices denote carbon sublattices A and B respectively. In all cases, the majority of the excess spin has delocalised away from the dopant and onto the carbon sublattices. All spin density slices are located 0.376 Å above the graphene plane. The plots are not absolutely comparable because of the different colour scales except at zero density.

There is therefore a clear asymmetry with respect to the spin distribution even *within the one sublattice* occupied by the N dopant for the LDA, PBE and HSE calculations. This distribution forms a pattern of its own, so the dopant sublattice therefore can be further divided into its own sublattices. In all cases, the carbon nuclei lying in the *different* sublattice to that of the N dopant have an equal accumulation of spin on them, so no such division into further sublattices appears to be occurring for this sublattice. In summary, this evidence suggests that finite size effects due to the limited supercell size may be affecting the spin distribution patterns; it would perhaps be unsurprising if the patterns form a trend with supercell size.

5.4.2. Two N dopants: there is no apparent intrinsic energetic advantage to N dopants collecting in one sublattice

Figure 5.5 shows the relative final energies and final spins of all of the two-dopant configurations and functionals studied, resolved into the two sublattices and initial total spin states. In all cases, the highest energy state was found to be a B1 configuration, so the energies of these configurations, as found using the appropriate functional, were selected as the reference energy to which all other supercell energies found using that same functional are compared.

The HSE results, like those from the single dopant case, require careful further investigation, because they appear to consistently suggest that the systems whose initial spin configurations are equivalent by spin-exchange symmetry have arrived at significantly different locations in the energy landscape. In particular, note the bottom left panel of Figure 5.6 which shows that the states with a final positive magnetic moment are significantly more stable than those with zero and negative moments. The corresponding relative HSE energies in Figure 5.5 show this dramatic energy asymmetry with respect to the initial magnetic moments.

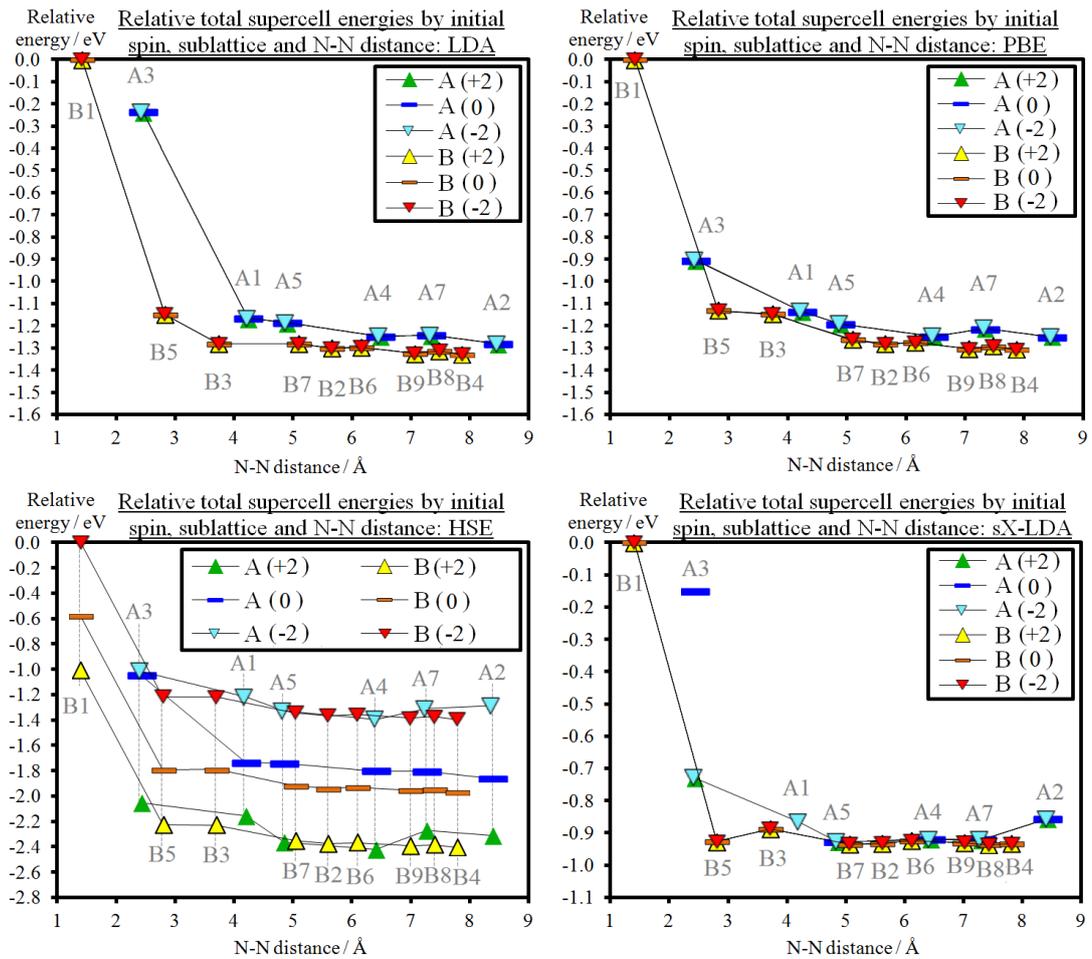


Figure 5.5. Relative total energies of the various configurations resolved into sublattices **A** and **B**.

The HSE results with zero and positive initial spins are highly suspect and likely to be affected by a bug.

At the time of writing this thesis, it is not entirely clear exactly what kind of investigation should be carried out to try to understand this. It should be noted that a bug in spin-polarised HSE was discovered in versions of CASTEP prior to version 7.0.1, which was announced on 3rd December 2013. This bug was found to affect the ability of the electronic minimiser to locate the ground state. However, these calculations were carried out using CASTEP 7.0.1, and so are *not* affected by this particular bug. As a simple diagnostic, the spin-polarised ground state of a single H atom was calculated using the HSE functional using initial magnetic moments of $\mu_B = \pm 1$. This is a sensible test to carry out because, as long as the final spin states are equal and opposite, the corresponding total energies must of course be

identical by spin-exchange symmetry. It was found that the calculation with initial moment $\mu_B = +1$ converged to a total energy of -13.6 eV (3 s. f.) and a final spin of $+1\mu_B$ as expected. However, the $\mu_B = -1$ calculation was found to consistently fail to converge with repeated attempts using density mixing and EDFT, which, given the simplicity of the H atom, is very strong evidence that a bug still exists in the code for the case of the spin-polarised HSE functional. It must be concluded that the HSE calculations of all but positive initial magnetic moment are very likely to be invalid and should be discarded, which is a very plausible explanation for the unusual HSE results.

The second key observation from these plots is that there is no energetic favourability to the configurations in which the two N dopants occupy the same sublattice. In fact, the most stable configurations tend to be those in which the second N lay in sublattice **B**. Thus, despite the strong and long-ranged sublattice spin-asymmetry which characterises the electronic environment around a single N dopant as shown in Figure 5.4, the apparently-observed tendency of N dopants to collect into one sublattice *cannot* be attributed to the intrinsic energies of the various substitutional N configurations, at least to the level of theory presented here. This might, or might not, be different if 5x5 or 7x7 supercells were used.

There is another interesting point: the relative energies of the A and B sublattice configurations predicted by the LDA and PBE functionals are extremely similar (i.e. the top two panels of Figure 5.5 have very similar trends), but the final magnetic moments predicted by these two functionals are very different, as seen in Figure 5.6. Therefore, the widespread energetic preference for the B configurations, as seen in all panels in Figure 5.5, is *not* due to magnetism. This suggests other possibilities like charge effects. A systematic investigation of the dependence of the Mulliken and Hirshfeld charge populations on supercell size would perhaps shed more light on this.

In addition, it is clear that the plots show a repulsive N-N interaction which decays quickly beyond 4 Å of separation. All LDA and PBE energies are seen to be independent of the initial spins, showing that no metastable spin-polarized electronic states exist within or near the range of the initial spin

values studied. The same can be said for the sX-LDA data with the exception of configuration A3 - which has the smallest N-N distance of all the A sublattice configurations - for which one metastable state appears to have been found about 0.6 eV above that arrived at in the other two A3 calculations.

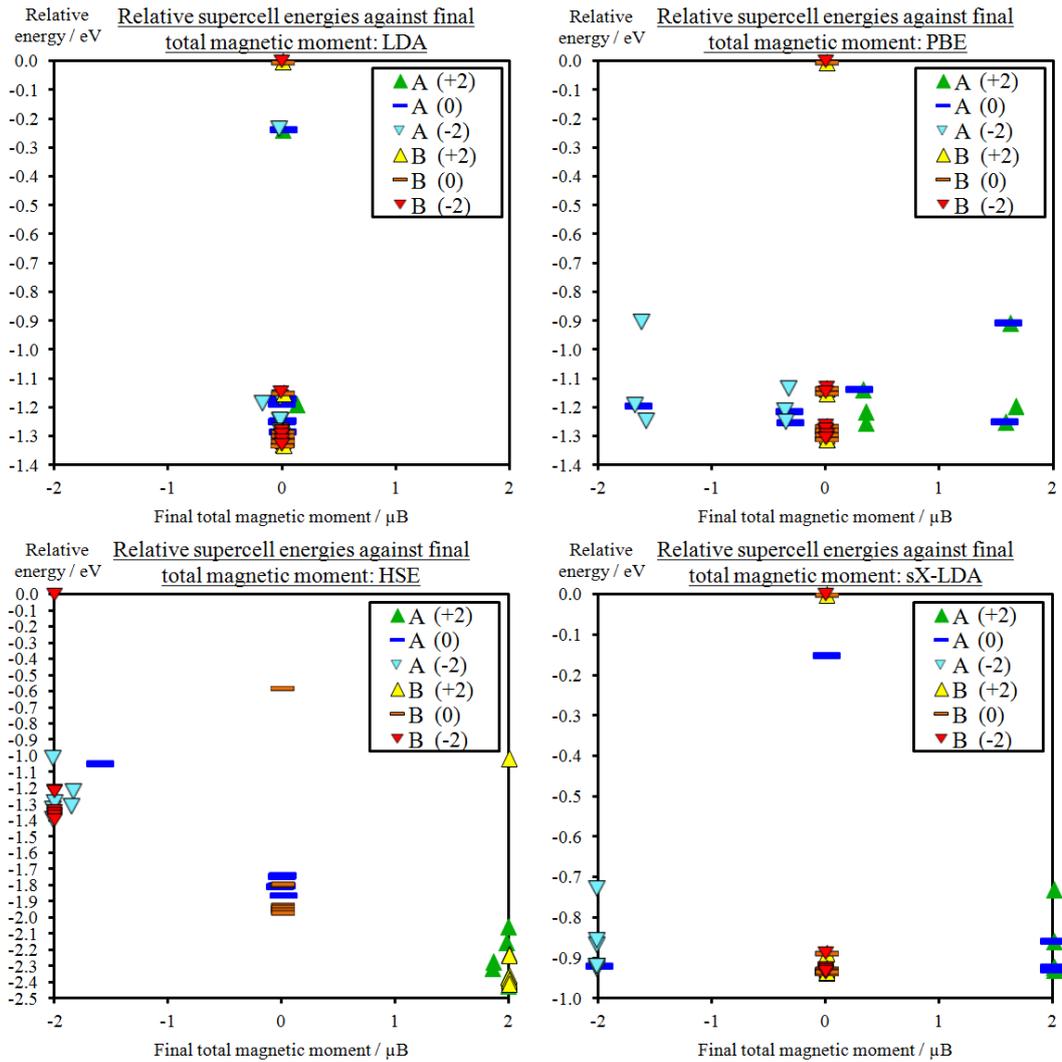


Figure 5.6. Final magnetic moments of the various configurations resolved into sublattices A and B.

All LDA results show negligible final spin, whereas PBE, HSE and sX-LDA show significant final magnetic moments. The HSE results require further investigation because they apparently suggest that the states with an excess of up spin are energetically favourable over those with an excess of down spin, which suggests either a very unusual energy landscape, or, more likely, a bug in the code.

To summarise, the results of the calculations presented in this chapter, and the implications that they have for the current pool of experimental observations, are not yet properly understood. Many possible factors still need to be ruled out before more robust conclusions can be reached. More precisely converged calculations are required, and the intended next steps are briefly outlined in section 5.5.

5.5. Proposed improvements to this project

The theoretical evidence so far gathered throughout this chapter can be used to inform the ways in which this project can be advanced. The first notable point is that the magnetic states predicted by the LDA and PBE functionals are markedly different, so a fuller survey of the results predicted by the other semi-local functionals over a larger range of initial spin states may shed more light on the system. Alongside the PBE parameterisation of the GGA, the revised PBE (rPBE) parameterisation of Hammer, Hansen and Norskov,¹⁷³ the parameterisation of Perdew *et al.* for atoms, molecules, solids, and surfaces¹⁷⁴ (PW91), the GGA for solids of Wu and Cohen¹⁷⁵ (WC) and the “PBE for solids” (PBEsol) of Perdew *et al.*¹⁷⁶ will be used to carry out single dopant geometry optimisation calculations at increasingly large supercell sizes such as 5x5, 6x6,..., 12x12. These are all semi-local functionals and so are quite computationally practical. Including geometry optimisation allows for long-ranged lattice perturbations to be assessed. In addition, the consequences of using density mixing and assigning the initial spins to the N dopant, and of using EDFT and assigning the initial spin to the entire supercell, will be assessed. The final Mulliken spin and charge populations of the N dopant and the C atoms in the fully optimised systems, and the total final magnetic moments of the supercells, will offer more information than that which is currently available. If the scenario should occur that trends are obtained for some functionals but not others, then this would pave the way for a systematic cross-examination of the results whereby a unique feature of one functional may be found to be responsible for the trend, or lack of, which may shed light on the basic underlying physics. Based on some initial calculations, there is already strong evidence

that finite size effects due to the supercell size are playing a role. The data from these preliminary investigations is tabulated in Table 5.2.

Table 5.2 Preliminary single N dopant calculations carried out to try to identify themes which could be influencing the results.

Finite size effects due to supercell size appear to be playing a role. For example, the N dopant Mulliken populations for the PBE calculations seem to depend on whether 5x5 or 6x6 supercells are used.

Supercell size	electronic minimisation method	XC	initial spin	Final energy/eV (not geometry-optimised) (9 d. p.)	N atom final Mulliken population		final magnetic moment / μ_B (5 d. p.)	modulus of final magnetic moment (5 d. p.)	k points grid	
					charge	spin				
5x5	density mixing	LDA	+1	-7874.622334522	-0.29	0.03	0.38837	0.47203	Γ	
			0	-7874.614929909	-0.29	0.00	-0.00491	0.00601	Γ	
			-1	-7874.622341246	-0.29	-0.03	-0.38872	0.47248	Γ	
		PBE	+1	-7899.046000085	-0.28	0.08	0.97322	1.46456	Γ	
			0	-7899.045988264	-0.28	0.08	0.97309	1.46200	Γ	
			-1	-7899.046000182	-0.28	-0.08	-0.97322	1.46446	Γ	
		HSE	+1	-7890.411692603	-0.34	0.09	1.00011	2.48187	Γ	
			0	-7890.411849805	-0.34	0.09	1.00006	2.48989	Γ	
			-1	-7889.889020245	-0.34	-0.09	-0.99987	2.44644	Γ	
		sX-LDA	+1	-7884.162906756	-0.44	0.02	1.00097	27.2731	Γ	
			0	-7884.160428829	-0.44	-0.02	-0.99958	27.2638	Γ	
			-1	-7884.150730743	-0.44	-0.02	-0.95568	27.1249	Γ	
6x6	density mixing	LDA	+1	-11289.58506555	-0.27	0.00	0.02433	0.05058	Γ	
			0	-11289.58505357	-0.27	0.00	-0.00747	0.01569	Γ	
			-1	-11289.58507815	-0.27	0.00	-0.03224	0.06709	Γ	
		PBE	+1	-11324.68852917	-0.27	-0.01	0.97191	5.45911	Γ	
			0	-11324.68863378	-0.27	0.01	-0.97409	5.46136	Γ	
			-1	-11324.68861832	-0.27	0.01	-0.97390	5.46152	Γ	
		HSE	+1	<i>not performed</i>	-	-	-	-	-	
			0	-11313.42118013	-0.32	-0.01	2.92276	4.07697	Γ	
			-1	-11312.53089238	-0.32	0.00	-0.99977	8.42361	Γ	
		sX-LDA	+1	<i>not performed</i>	-	-	-	-	-	
			0	<i>not performed</i>	-	-	-	-	-	
			-1	<i>not performed</i>	-	-	-	-	-	
5x5	EDFT	LDA	+1	<i>not performed</i>	-	-	-	-		
			0	<i>not performed</i>	-	-	-	-		
			-1	<i>not performed</i>	-	-	-	-		
		PBE	+1	<i>not performed</i>	-	-	-	-	-	
			0	<i>not performed</i>	-	-	-	-	-	
			-1	<i>not performed</i>	-	-	-	-	-	
		HSE	+1	<i>not performed</i>	-	-	-	-	-	
			0	<i>not performed</i>	-	-	-	-	-	
			-1	<i>not performed</i>	-	-	-	-	-	
		sX-LDA	+1	-7884.177778918	-0.44	0.02	1.00000	27.3360	Γ	
			0	-7884.177778914	-0.44	0.02	1.00000	27.3360	Γ	
			-1	-7884.177778827	-0.44	-0.02	-1.00000	27.3372	Γ	
6x6	EDFT	LDA	+1	-11289.58509690	-0.27	0.00	0.03657	0.07532	Γ	
			0	-11289.58506750	-0.27	0.00	-0.02157	0.04469	Γ	
			-1	-11289.58508921	-0.27	0.00	-0.03335	0.06806	Γ	
		PBE	+1	-11324.68669671	-0.27	0.01	-0.97846	5.30088	Γ	
			0	-11324.68686510	-0.27	0.01	-0.97579	5.33712	Γ	
			-1	-11324.68721144	-0.27	0.01	-0.97689	5.36019	Γ	
		HSE	+1	-11313.09967766	-0.32	0.00	0.99999	7.38326	Γ	
			0	-11313.10438967	-0.33	-0.01	0.99909	7.23237	Γ	
			-1	-11312.59801970	-0.32	0.00	-0.99999	6.98535	Γ	
		sX-LDA	+1	-11305.41781677	-0.43	0.02	1.00000	38.6458	Γ	
			0	-11305.41781677	-0.43	-0.02	-1.00000	38.6458	Γ	
			-1	-11305.41781672	-0.43	0.02	1.00000	38.6445	Γ	

Similarly precise calculations could be carried out using a range of hybrid functionals using Gaussian-like basis functions, which is a more efficient approach than using a plane-wave basis set, especially considering that the majority of the supercells used consist of vacuum space. The energies obtained so far have all been obtained at a relatively coarse level of precision, so a renewed emphasis on stringent convergence will be used for the new calculations. Alongside single dopant calculations, configurations involving 2 or 3 N dopants could be systematically investigated. Using 3 dopants is likely to give rise to a very large number of unique configurations, however, especially for the larger supercells. A more systematic approach based on dopant concentration or mean N-N distance, or even a more rigorous group-theory approach could be employed. No multiple-dopant calculations will be carried out until the single dopant calculations described above have been completed and fully analysed.

In summary, the objective at this stage is to examine the consequences of using a particular functional more closely by using a wide range of functionals with a much higher level of precision. The information gathered will be used to inform the subsequent investigations of this project.

Chapter 6: Concluding remarks

In this chapter, the key achievements of the projects presented in this thesis are summarised.

In chapter 3, the behaviour of metal adatoms on graphene substrates was explained theoretically using DFT. When the STEM images shown in Figure 3.1 were originally obtained, it was not fully understood exactly why the metals appeared at defects, contaminated regions and edge sites, but not on the pristine lattice regions. One of the biggest barriers to properly understanding what was being observed in the images was not knowing the energy landscape experienced by an adatom at the various bonding sites on graphene, or indeed whether energy barriers associated with adatoms' motion across the lattice were playing a role. Before this investigation was carried out, speculations were centred mainly on the thickness of the underlying substrate, and the possibility that metal adatoms might form stable bonds on thicker substrates owing to extra contributions to the binding energy from van der Waals interactions with the deeper sublayers. What this investigation has demonstrated is that van der Waals interactions from sublayers do indeed have a major influence on the total binding energy, but that this is not responsible for the phenomena seen in the microscope. To the best of my knowledge, this information was completely new at the time of publication. It has been shown that Cr, Al and Au adatoms (and probably other chemically similar metals) evaporated onto graphene by CVD are likely to diffuse on multilayered graphene substrates at room temperature, even on those of different thicknesses, due to the migration energy barriers being both *i)* small, and *ii)* relatively unaffected by substrate thickness. It is certainly fair to say that the calculations carried out in this project were inspired by previously published calculations,³⁷⁻⁵² but they also significantly build on these published works. One specific advancement of this project over these previously published works was the explicit calculation of migration barriers of the adatoms considered. Transition state calculations of adatom migration barriers on graphene are apparently rare in the literature; only the work of Yazyev *et al.*⁵¹ was found during the literature review, in

which the migration barrier for Co was calculated. It is certainly quite reasonable to anticipate from the binding energies in Figure 3.18 alone that the migration barriers between these sites are extremely small, simply because the energy differences are small. However, by going one step further in this thesis and calculating these barriers directly, the qualitative argument for adatom migration based on energy barrier estimates - a suggestion made in previous studies^{39,42-46,55,57} - has been advanced to a *quantitative* argument based on the precisely and explicitly calculated barriers tabulated in Table 3.3. This is new information regarding the energetic environment experienced by metal adatoms on graphene which did not exist before this project was published. Nevertheless, credit is very much due to the authors of these studies^{39,42-46,55,57} despite the absence of actual energy barrier calculations, because they all clearly identified the phenomenon of adatom migration. Another important remark is that many of these published studies repeatedly used phrases like "atom X prefers to adsorb at site Y" in their abstracts. This project has placed fresh emphasis on the fact that phrases of this sort have the potential to be quite misleading because they can easily lead to incorrect interpretations of metal adatoms on graphene being static and stably bonded at room temperature. The continued emphasis on random *thermal migration* at room temperature inferred from theory and seen directly in collaborators' experiments helps to clarify current understanding of the behaviour of metals on graphene. Alongside this, to the very best of my knowledge, the idea of fixing substrate carbon atom positions according to the lattice symmetries (see Figures 3.14 and 3.15) and investigating the effects of varying substrate thickness with van der Waals corrections had not been published elsewhere at the time this project was published in April 2013. As for the original aim of investigating possibilities of interfacing graphene to metal contacts, the primary conclusion of this project is that CVD of single metal atoms is unlikely to be the synthesis technique of choice for building metal-graphene interfaces. Other synthesis methods may well result in different behaviour. Indeed, the study by Zhao⁷¹ mentioned in chapter 3 which shows the first images of a freestanding 2-D Fe film in a graphene hole is a potentially very significant demonstration that building 2-D interfaces may be the most feasible

approach. Looking forward in a general sense, much of nanotechnology relies on the passive manipulation of physical matter on the microscopic level by means of harnessing energetically favourable processes under controlled conditions. Preferential edge-site binding of adatoms on graphene substrates, or 2D materials in general, is one such process which could perhaps be exploited in contexts such as patterned nanoscale devices, systematic edge-decoration of 2D nanoribbons and other nanoscale constructions where site-dependent bonding tendencies are an important ingredient in the fabrication process.

In chapter 4, a new method was developed to reliably calculate the energies of step defects and step-step interaction energies on crystalline solid surfaces. The work presented in this chapter is useful from the perspective of theoretical methods development because it highlights the potentially misleading nature of existing published methods and provides a basis for more robust investigations into the energies of steps in nanoscale systems in which the interactions between opposing surfaces are strong. Having conducted a very thorough literature search, it can be said with a high degree of confidence that, for systems in which surface energies converge slowly with slab thickness, this method should be used in favour of all previously published methods because it is significantly more robust. Before this project was published, the only serious attempt at *ab initio* calculation of step and step interaction energies was the *Nature Materials* paper by Gong *et al.*¹³⁶, which makes use of linear regression to evaluate high-index surface energies; a method which has been shown in this thesis to have the *potential* to be seriously misleading. (Once again, it should be emphasised that this does not necessarily invalidate Gong's results, because explicit details of their methodology was not given.) In section 4.2.4 and appendix B.1 of this thesis, it has been argued directly that the linear fitting method is flawed for TiO₂ rutile surface energies, so the fact that the Aarhus group¹⁴³ has also independently identified this same difficulty adds a great deal of credibility to the arguments given in this thesis. They did not account for step-step interaction energies in their method, but their step energies and consequent island shape agree well (see Figure 8 from their paper) with those predicted in this project, which is good evidence that both

methods are robust. The long-term impact of the method developed in chapter 4 is likely to depend on the importance of the systems for which it may find application. As it stands, this method can currently be reliably applied to high-index *macroscopic* surfaces, i.e. high-index flat crystal facets with steps, or truly high-index vicinal facets. Having applied this method to TiO₂ rutile (110), the agreement of the island shape with STM data has strongly suggested that the method works, and it has added credibility to the existence of the atomic arrangements used in the model on the real laboratory sample. Further lines of enquiry could attempt to start focusing on how to extend this method to nanoparticle surfaces, taking into account the surface curvature on the nanoscopic/mesoscopic scale. Nanoparticle-based technologies are becoming increasingly important in emerging fields of engineering, so it is easy to see where the motivation lies for such a method. It must be conceded that accounting for nanoparticle surface curvature would be quite a serious challenge under periodic boundary conditions, but the method developed in chapter 4 is certainly an improvement over published methods of step energy calculation. In contrast to nanoparticles, this method could of course be immediately applied to the case of high-index vicinal films, although these are rather more exotic entities which have not found such widespread application as nanoparticles. In summary, the method developed throughout chapter 4 makes a significant new contribution to theoretical methods in surface science, and also contributes to the overall characterisation of the extremely important system that is the (110) surface of TiO₂ rutile.

In chapter 5, current progress on an investigation into the viability of systematically doping graphene with substitutional nitrogen atoms in only one carbon sublattice was presented. Despite being incomplete, it is apparent from the investigations conducted so far that there is unlikely to be a significant energetic advantage to nitrogen dopants collecting into one particular sublattice. Specifically, there is apparently very little dependence of the energy trends on the exchange-correlation functional used. Moreover these trends predicted by all functionals appear to unanimously imply that the pairs of substitutional dopants studied are more stable when laying in *different* sublattices. As it stands, the greatest source of uncertainty in this

data is the possibility that finite size effects may be significantly influencing the trends. The main robust and significant conclusion drawn from the investigations so far is that magnetic effects appear to not be playing any role at all, because while final magnetic moments depend sensitively on the functional, the energy trends do not. Clear trends do appear to be present in the relative total energies plotted Figure 5.5, which is indicative of some underlying effect, although it cannot be ruled out as a calculation artifact and so cannot currently be identified, even if it is physically meaningful. In summary, the progress in chapter 5 sheds significant light on how to proceed with the project, but from the information so far obtained in this thesis, it must be concluded that no clear implications for graphene-based technologies exploiting patterned-doping of nitrogen are clearly identifiable at the time of writing.

List of References

- ¹P. Hohenberg and W. Kohn *Physical Review* **136**, B864 (1964)
- ²W. Kohn and L. J. Sham *Physical Review* **140**, A14433 (1965)
- ³R. Lv, Q. Li, A. R. Botello-Méndez, T. Hayashi, B. Wang, A. Berkdemir, Q. Hao, A. L. Elías, R. Cruz-Silva, H. R. Gutiérrez, Y. A. Kim, H. Muramatsu, J. Zhu, M. Endo, H. Terrones, J.-C. Charlier, M. Pan and M. Terrones *Scientific Reports* **2**, 586 (2012)
- ⁴L. Zhao, R. He, K. T. Rim, T. Schiros, K. S. Kim, H. Zhou, C. Gutiérrez, S. P. Chockalingam, C. J. Arguello, L. Pálová, D. Nordlund, M. S. Hybertsen, D. R. Reichman, T. F. Heinz, P. Kim, A. Pinczuk, G. W. Flynn, A. N. Pasupathy *Science* **333**, 999 (2011)
- ⁵U. Bangert, W. Pierce, D. M. Kepaptsoglou, Q. M. Ramasse, R. Zan, M. H. Gass, J. A. Van den Berg, C. B. Boothroyd, J. Amani and H. Hofsäss *Nano Letters* **13**, 4902 (2013)
- ⁶R. M. Martin *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press (2004)
- ⁷M. Born and J. R. Oppenheimer *Annalen der Physik* **84**, 457 (1927)
- ⁸D. R. Hartree *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 89 (1928)
- ⁹J. C. Slater *Physical Review* **35**, 210 (1930)
- ¹⁰J. C. Slater *Physical Review* **34**, 1293 (1929)
- ¹¹V. Fock *Zeitschrift für Physik* **61**, 126 (1930)
- ¹²C. Møller and M. S. Plesset *Physical Review* **46**, 618 (1934)
- ¹³J. Čížek *The Journal of Chemical Physics* **45**, 4256 (1966)
- ¹⁴S. R. Langhoff and E. R. Davidson *International Journal of Quantum Chemistry* **8**, 61 (1974)
- ¹⁵J. A. Pople, M. Head-Gordon and K. Raghavachari *The Journal of Chemical Physics* **87**, 5968 (1987)
- ¹⁶F. Bloch *Zeitschrift für Physik* **52**, 555 (1928)
- ¹⁷H. J. Monkhorst and J. D. Pack *Physical Review B* **13**, 5188 (1976)
- ¹⁸J. C. Slater *Physical Review* **81**, 385 (1951)
- ¹⁹D. H. Hamalm, M. Schluter, and C. Chiang *Physical Review Letters* **43**, 1494 (1979)
- ²⁰D. Vanderbilt *Physical Review B* **41**, 7892 (1990)
- ²¹H. Hellmann *Einführung in die Quantumchemie*, Franz Deutsche, Leipzig (1937)

- ²²R. P. Feynman *Physical Review* **56**, 340 (1939)
- ²³H. Rydberg, M. Dion, N. Jacobson, E. Schröder, P. Hyldgaard, S. Simak, D. Langreth and B. Lundqvist *Physical Review Letters* **91**, 126402 (2003)
- ²⁴M. Dion, H. Rydberg, E. Schröder, D. Langreth and B. Lundqvist *Physical Review Letters* **92**, 246401 (2004)
- ²⁵M. Dion, H. Rydberg, E. Schröder, D. Langreth and B. Lundqvist *Physical Review Letters* **95**, 109902 (2005)
- ²⁶S. Grimme *Journal of Computational Chemistry* **27**, 1787 (2006)
- ²⁷P. Jurečka, J. Černý, P. Hobza and D. R. Salahub *Journal of Computational Chemistry* **28**, 555 (2007)
- ²⁸A. Tkatchenko and M. Scheffler *Physical Review Letters* **102**, 073005 (2009)
- ²⁹K. S. Novoselov, A. K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I. V. Grigorieva and A.A. Firsov *Science* **306**, 666 (2004)
- ³⁰F. Schwierz *Nature Nanotechnology* **5**, 487 (2010)
- ³¹Y. Wu, V. Perebeinos, Y-M. Lin, T. Low, F. Xia and P. Avouris *Nano Letters* **12**, 1417 (2012)
- ³²K. Nagashio, T. Nishimura, K. Kita and A. Toriumi *Japanese Journal of Applied Physics* **49**, 051304 (2010)
- ³³K. Nagashio, T. Nishimura, K. Kita and A. Toriumi *Applied Physics Letters* **97**, 143514 (2010)
- ³⁴S. Russo, M. F. Craciun, M. Yamamoto, A. F. Morpurgo and S. Tarucha *Physica E: Low-dimensional Systems and Nanostructures* **42**, 677 (2010)
- ³⁵P. A. Khomyakov, A. A. Starikov, G. Brocks and P. J. Kelly *Physical Review B* **82**, 115437 (2010)
- ³⁶R. S. Sundaram, M. Steiner, H-Y. Chiu, M. Engel, A. A. Bol, R. Krupke, M. Burghard, K. Kern and P. Avouris *Nano Letters* **11**, 3833 (2011)
- ³⁷M. P. Lima, Antônio J. R. da Silva and A. Fazzio *Physical Review B* **84**, 245411 (2011)
- ³⁸J. Ding, Z Qiao, W. Feng, Y. Yao and Q. Niu *Physical Review B* **84**, 195444 (2011)
- ³⁹H. Valencia, A. Gil and G. Frapper *The Journal of Physical Chemistry C* **114**, 14141 (2010)
- ⁴⁰R. Varns and P. Strange *Journal of Physics: Condensed Matter* **20**, 225005 (2008)
- ⁴¹M. Sargolzaei and F. Gudarzi *Journal of Applied Physics* **110**, 064303 (2011)
- ⁴²K. Nakada and A. Ishii *Solid State Communications* **151**, 13 (2011)
- ⁴³C. Cao, M. Wu, J. Jiang and H-P. Cheng *Physical Review B* **81**, 205424 (2010)
- ⁴⁴A. Ishii, M. Yamamoto, H. Asano and K. Fujiwara *Journal of Physics: Conference Series* **100**, 052087 (2008)

- ⁴⁵K.T. Chan, J. B. Neaton and M. L. Cohen *Physical Review B* **77**, 235430 (2008)
- ⁴⁶M. Amft, S. Lebègue, O. Eriksson and N. V. Skorodumova *Journal of Physics: Condensed Matter* **23**, 395001 (2011)
- ⁴⁷Z. M. Ao, Q. Jiang, R. Q. Zhang, T. T. Tan and S. Li *Journal of Applied Physics* **105**, 074307 (2009)
- ⁴⁸L. Hu, X. Hu, X. Wu, C. Du, Y. Dai and J. Deng *Physica B: Physics of Condensed Matter* **405**, 3337 (2010)
- ⁴⁹H. Johll and H.C. Kang *Physical Review B* **79**, 245416 (2009)
- ⁵⁰Y. Mao, J. Yuan and J. Zhong *Journal of Physics: Condensed Matter* **20**, 115209 (2008)
- ⁵¹O. V. Yazyev and A. Pasquarello *Physical Review B* **82**, 045407 (2010)
- ⁵²J. K-Hwan, C. S-Myeong and J. S-Hoon *Physical Review B* **82**, 033414 (2010)
- ⁵³Y. Tang, Z. Yang and Xianqi Dai *The Journal of Chemical Physics* **135**, 224704 (2011)
- ⁵⁴D. W. Boukhvalov and M. I. Katsnelson *Applied Physics Letters* **95**, 023109 (2009)
- ⁵⁵S. Malola, H. Hakkinen and P. Koskinen *Applied Physics Letters* **94**, 043106 (2009)
- ⁵⁶E. J. G. Santos, A. Ayuela and D. Sánchez-Portal *New Journal of Physics* **12**, 053012 (2010)
- ⁵⁷O. Cretu, A. V. Krasheninnikov, J. A. Rodríguez-Manzo, L. Sun, R. M. Nieminen and F. Banhart *Physical Review Letters* **105**, 196102 (2010)
- ⁵⁸E. J. G. Santos, D. Sánchez-Portal and A. Ayuela *Physical Review B* **81**, 125433 (2010)
- ⁵⁹S. Karoui, H. Amara, C. Bichara and F. Ducastelle *ACS Nano* **4**, 6114 (2010)
- ⁶⁰R. Zan, Q. M. Ramasse, U. Bangert, K. S. Novoselov *Nano Letters* **12**, 3936 (2012)
- ⁶¹J. P. Perdew, K. Burke and M. Ernzerhof *Physical Review Letters* **77**, 3865 (1996)
- ⁶²L. G. De Arco, Y. Zhang, A. Kumar, and C. Zhou *IEEE Transactions on Nanotechnology* **8**, 135 (2009)
- ⁶³D. Wei, Y. Liu, H. Zhang, L. Huang, B. Wu, J. Chen and G. Yu *Journal of the American Chemical Society* **131**, 11147 (2009)
- ⁶⁴A. Malesevic, R. Vitchev, K. Schouteden, A. Volodin, L. Zhang, G. Van Tendeloo, A. Vanhulsel and C. Van Haesendonck *Nanotechnology* **19**, 305604 (2008)
- ⁶⁵A. Reina, X. Jia, J. Ho, D. Nezich, H. Son, V. Bulovic, M. S. Dresselhaus and J. Kong *Nano Letters* **9**, 30 (2009)

- ⁶⁶G. Ming, J. J. Belbruno, S. D. Kenny and R. Smith *Surface Science* **541**, 91 (2003)
- ⁶⁷R. Zan, U. Bangert, Q. M. Ramasse and K. S. Novoselov *Nano Letters* **11** 1087 (2011)
- ⁶⁸R. Zan, U. Bangert, Q. M. Ramasse and K. S. Novoselov *Small* **7** 2868 (2011)
- ⁶⁹Q. M. Ramasse, R. Zan, U. Bangert, D.W. Boukhvalov, Y-W. Son and K. S. Novoselov *ACS Nano* **6**, 4063 (2012)
- ⁷⁰T. Eberlein, U. Bangert, R. R. Nair, R. Jones, M. Gass, A. L. Bleloch, K. S. Novoselov, A. Geim and P. R. Briddon *Physical Review B* **77**, 233406 (2008)
- ⁷¹J. Zhao, Q. Deng, A. Bachmatiuk, G. Sandeep, A. Popov, J. Eckert and M. H. Rümmeli *Science* **343**, 1228 (2014)
- ⁷²J. C. Meyer, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. J. Booth and S. Roth *Nature* **446**, 60 (2007)
- ⁷³A. Fasolino, J. H. Los and M. I. Katsnelson *Nature Materials* **6**, 858 (2007)
- ⁷⁴P. Lambin, H. Amara, F. Ducastelle and L. Henrard *Physical Review B* **86**, 045448 (2012)
- ⁷⁵G. Makov and M. C. Payne *Physical Review B* **51**, 4014 (1995)
- ⁷⁶J. Neugebauer and M. Scheffler *Physical Review B* **46**, 16067 (1992)
- ⁷⁷L. Bengtsson *Physical Review B* **59**, 12301 (1999)
- ⁷⁸L. N. Kantorovich *Physical Review B* **60**, 15476 (1999)
- ⁷⁹P. Schultz *Physical Review B* **60**, 1551 (1999)
- ⁸⁰I.-C. Yeh and M. L. Berkowitz *Journal of Chemical Physics* **111**, 3155 (1999)
- ⁸¹T. M. Nymand and P. Linse *Journal of Chemical Physics* **112**, 6152 (2000)
- ⁸²I. Hamada, M. Otani, O. Sugino and Y. Morikawa *Physical Review B* **80**, 165411 (2009)
- ⁸³S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson and M. C. Payne *Zeitschrift fuer Kristallographie* **220**, 567 (2005)
- ⁸⁴E. R. McNellis, J. Meyer, and K. Reuter *Physical Review B* **80**, 205414 (2009)
- ⁸⁵P. Trucano and R. Chen *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics* **72**, 446 (1976)
- ⁸⁶P. Koskinen, S. Malola and H. Häkkinen *Physical Review Letters* **101**, 115502 (2008)
- ⁸⁷C. G. Broyden *IMA Journal of Applied Mathematics* **6**, 76 (1970)
- ⁸⁸C. G. Broyden *IMA Journal of Applied Mathematics* **6**, 222 (1970)
- ⁸⁹R. Fletcher *The Computer Journal* **13**, 317 (1970)
- ⁹⁰D. Goldfarb *Mathematics of Computation* **24**, 23 (1970)
- ⁹¹D. F. Shanno *Mathematics of Computation* **24**, 647 (1970)

- ⁹²T. A. Halgren and W. N. Lipscomb *Chemical Physics Letters*, **49** 225 (1977)
- ⁹³N. Govind, M. Peterson, G. Fitzgerald, D. K-Smith and J. Andzelm *Computational Materials Science* **28**, 250 (2003)
- ⁹⁴H. Wang, K. Li, Y. Cheng, Q. Wang, Y. Yao, U. Schwingenschlögl, X. Zhang and W. Yang *Nanoscale* **4**, 2920 (2012)
- ⁹⁵J. C. Meyer, F. Eder, S. Kurasch, V. Skakalova, J. Kotakoski, H. J. Park, S. Roth, A. Chuvilin, S. Eychus, G. Benner, A. V. Krasheninnikov and U. Kaiser *Physical Review Letters* **108**, 196102 (2012)
- ⁹⁶R. F. Egerton, P. Li and M. Malac *Micron* **35**, 399 (2004)
- ⁹⁷A. Fujishima, T. N. Rao and D. A. Tryk *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **1**, 1 (2000)
- ⁹⁸Z. Huang, P.-C. Maness, D. M. Blake, E. J. Wolfrum, S. L. Smolinski and W. A. Jacoby *Journal of Photochemistry and Photobiology A: Chemistry* **130**, 163 (2000)
- ⁹⁹P. Maness, S. Smolinski, D. M. Blake, Z. Huang, E. J. Wolfrum and W. A. Jacoby *Applied Environmental Microbiology* **65**, 4094 (1999)
- ¹⁰⁰K. Sunada, Y. Kikuchi, K. Hashimoto and A. Fujishima *Environmental Science & Technology* **32**, 726 (1998)
- ¹⁰¹D. F. Ollis, E. Pelizzetti and N. Serpone *Environmental Science & Technology* **25**, 1522 (1991)
- ¹⁰²U. I. Gaya and A. H. Abdullah *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **9**, 1 (2008)
- ¹⁰³R. Wang, K. Hashimoto, A. Fujishima, M. Chikuni, E. Kojima, A. Kitamura, M. Shimohigoshi and T. Watanabe *Nature* **388**, 431 (1997)
- ¹⁰⁴R. Wang, K. Hashimoto, A. Fujishima, M. Chikuni, E. Kojima, A. Kitamura, M. Shimohigoshi and T. Watanabe *Advanced Materials* **10**, 135 (1998)
- ¹⁰⁵Y. Paz, Z. Luo, L. Rabenberg and A. Heller *Journal of Materials Research* **10**, 2842 (1995)
- ¹⁰⁶K. J. T. Livi, B. Schaffer, D. Azzolini, C. R. Seabourne, T. P. Hardcastle, A. J. Scott, D. Sverjensky, R. M. Hazen, J. D. Erlbacher and R. M. D. Brydson *Langmuir* **29**, 6876 (2013)
- ¹⁰⁷M. L. Machesky, D. J. Wesolowski, D. A. Palmer and M. K. Ridley *Journal of Colloid and Interface Science* **239**, 314 (2001)
- ¹⁰⁸M. K. Ridley, M. L. Machesky, D. J. Wesolowski and D. A. Palmer *Geochimica et Cosmochimica Acta* **68**, 239 (2004)
- ¹⁰⁹H. J. Cleaves II, A. M. Scott, F. C. Hill, J. Leszczynski, Nita Sahai and R. Hazen *Chemical Society Reviews* **41**, 5502 (2012)
- ¹¹⁰S. Bahri, C. M. Jonsson, C. L. Jonsson, D. Azzolini, D. Sverjensky and R. M. Hazen *Environmental Science & Technology* **45**, 3959 (2011)
- ¹¹¹M. Machesky, D. Wesolowski, J. Rosenqvist, M. Předota, L. Vlček, M. Ridley, V. Kohli, Z. Zhang, P. Fenter, P. Cummings, S. Lvov, M. Fedkin, V. R.-Santiago, J. Kubicki and A. Bandura *Langmuir* **27**, 4585 (2011)

- ¹¹²U. Diebold *Surface Science Reports* **48**, 53 (2003)
- ¹¹³M. Ramamoorthy, D. Vanderbilt and R. D. King-Smith *Physical Review B* **49**, 16721 (1994)
- ¹¹⁴U. Martinez, J. Ø. Hansen, E. Lira, H. H. Kristoffersen, P. Huo, R. Bechstein, E. Lægsgaard, F. Besenbacher, B. Hammer and S. Wendt *Physical Review Letters* **109**, 155501 (2012)
- ¹¹⁵H. Onishi and Y. Iwasawa *Chemical Physics Letters* **226**, 111 (1994)
- ¹¹⁶H. Onishi, K. Fukui and Y. Iwasawa *Bulletin of the Chemical Society of Japan* **68**, 2447 (1995)
- ¹¹⁷U. Diebold, J. Lehman, T. Mahmoud, M. Kuhn, G. Leonardelli, W. Hebenstreit, M. Schmid and P. Varga *Surface Science* **411**, 137 (1998)
- ¹¹⁸M. Li, W. Hebenstreit, L. Gross, U. Diebold, M. A. Henderson, D. R. Jennison, P. A. Schultz and M. P. Sears *Surface Science* **437**, 173 (1999)
- ¹¹⁹U. Martinez, L. B. Vilhelmsen, H. H. Kristoffersen, J. Stausholm-Møller and B. Hammer *Physical Review B* **84**, 205434 (2011)
- ¹²⁰A. Kiejna, T. Pabisiak, and S.W. Gao *Journal of Physics: Condensed Matter* **18**, 4207 (2006)
- ¹²¹X. Wu, A. Selloni and S. K. Nayak *The Journal of Chemical Physics* **120**, 4512 (2004)
- ¹²²S. J. Thompson and S. P. Lewis *Physical Review B* **73**, 073403 (2006)
- ¹²³K. J. Hameeuw, G. Cantele, D. Ninno, F. Trani and G. Iadonisi *The Journal of Chemical Physics* **124**, 024708 (2006)
- ¹²⁴S. P. Bates, G. Kresse and M. J. Gillan *Surface Science* **385**, 386 (1997)
- ¹²⁵M. D. Rasmussen, L. M. Molina and B. Hammer *The Journal of Chemical Physics* **120**, 988 (2004)
- ¹²⁶J. Oviedo, M. A. San Miguel and J. F. Sanz, *The Journal of Chemical Physics* **121**, 7427 (2004)
- ¹²⁷C. Zhang and P. J. D. Lindan *The Journal of Chemical Physics* **118**, 4620 (2003)
- ¹²⁸P. J. D. Lindan and C. Zhang *Physical Review B* **72**, 075439 (2005)
- ¹²⁹L. A. Harris and A. A. Quong *Physical Review Letters* **93**, 086105 (2004)
- ¹³⁰H. Perron, J. Vandenborre, C. Domain, R. Drot, J. Roques, E. Simoni, J.-J. Ehrhardt and H. Catalette *Surface Science* **601**, 518 (2007)
- ¹³¹T. Bredow, L. Giordano, F. Cinquini, and G. Pacchioni *Physical Review B* **70**, 035419(2004)
- ¹³²H. Perron, C. Domain, J. Roques, R. Drot, E. Simoni and H. Catalette *Theoretical Chemistry Accounts* **117**, 565 (2007)
- ¹³³T. P. Hardcastle, R.M.D. Brydson, K.J.T. Livi, C.R. Seabourne and A.J. Scott *Journal of Physics: Conference Series* **371**, 012059 (2012)
- ¹³⁴J. C. Boettger *Physical Review B* **53**, 13133 (1996)

- ¹³⁵J. G. Gay, J. R. Smith, R. Richter, F. J. Arlinghaus and R. H. Wagoner *Journal of Vacuum Science and Technology A* **2**, 93 (1984)
- ¹³⁶X.-Q. Gong, A. Selloni, M. Batzill and U. Diebold *Nature Materials* **5**, 665 (2006)
- ¹³⁷S. M. Koslov, F. Vin, N. Nilius, S. Shaikhutdinov, K. M. Neyman *The Journal of Physical Chemistry Letters* **3**, 1956 (2012)
- ¹³⁸B. Li, A. Michaelides and M. Scheffler *Physical Review B* **76**, 075401 (2007)
- ¹³⁹B. D. Yu and M. Scheffler *Physical Review B* **55**, 13916 (1997)
- ¹⁴⁰J. W. M. Frenken and P. Stoltze *Physical Review Letters* **82**, 3500 (1999)
- ¹⁴¹C. Barreateau, F. Raouafi, M.C. Desjonquères and D. Spanjaard *Journal of Physics: Condensed Matter* **15**, S3171–S3196 (2003)
- ¹⁴²F. Raouaf, C. Barreateau, F. Raouafi, M.C. Desjonquères and D. Spanjaard *Surface Science* **505**, 183 (2002)
- ¹⁴³J. Stausholm-Møller, H. H. Kristoffersen, U. Martinez and B. Hammer *The Journal of Chemical Physics* **139**, 234704 (2013)
- ¹⁴⁴E. D. Williams *Surface Science* **299**, 502 (1994)
- ¹⁴⁵E. E. Gruber and W. W. Mullins *Journal of Physics and Chemistry of Solids* **28**, 875 (1967)
- ¹⁴⁶C. Jayaprakash, C. Rottman and W. F. Saam *Physical Review B* **39**, 6549 (1984)
- ¹⁴⁷V. I. Marchenko and A. Y. Parshin *Soviet Journal of Experimental and Theoretical Physics* **79**, 257 (1981)
- ¹⁴⁸L. Barbier, L. Masson, J. Cousty and B. Salanon *Surface Science* **345**, 197 (1996)
- ¹⁴⁹L. Masson, L. Barbier, J. Cousty and B. Salanon *Surface Science* **317**, L1115 (1994)
- ¹⁵⁰N. García and P. A. Serena *Surface Science* **330**, L665 (1995)
- ¹⁵¹M. Nowicki, C. Bombis, A. Emundts, H. P. Bonzel and P. Wynblatt *New Journal of Physics* **4**, 60 (2002)
- ¹⁵²R. Najafabadi and D. J. Srolovitz *Surface Science* **317**, 221 (1994)
- ¹⁵³A. C. Redfield and A. Zangwill *Physical Review B* **46**, 4289 (1992)
- ¹⁵⁴R. V. Kukta, A. Peralta and D. Kouris *Physical Review Letters* **88**, 186102 (2002)
- ¹⁵⁵K. Swamy, E. Bertel and I. Vilfan *Surface Science* **425**, L369 (1999)
- ¹⁵⁶S. van Dijken, H. J. W. Zandvliet and B. Poelsema *Physical Review B* **55**, 7864 (1997)
- ¹⁵⁷J. Frohn, M. Giesen, M. Poensgen, J. F. Wolf and H. Ibach *Physical Review Letters* **67**, 3543 (1991)
- ¹⁵⁸W. H. Baur *Acta Crystallographica* **9**, 515 (1959)

- ¹⁵⁹D. C. Liu and J. Nocedal *Mathematical Programming* **45**, 503 (1989)
- ¹⁶⁰E. Lira, P. Huo, J.Ø. Hansen, F. Rieboldt, R. Bechstein, Y. Wei, R. Streber, S. Porsgaard, Z. Li, E. Lægsgaard, S. Wendt and F. Besenbacher *Catalysis Today* **182**, 25 (2012)
- ¹⁶¹E. Lira, S. Wendt, P. Huo, J. Ø. Hansen, R. Streber, S. Porsgaard, Y. Wei, R. Bechstein, E. Lægsgaard, and F. Besenbacher *Journal of the American Chemical Society* **133**, 6529 (2011)
- ¹⁶²S. Wendt, R. Schaub, J. Matthiesen, E. K. Vestergaard, E. Wahlström, M. D. Rasmussen, P. Thostrup, L. M. Molina, E. Lægsgaard, I. Stensgaard, B. Hammer and F. Besenbacher *Surface Science* **598**, 226 (2005)
- ¹⁶³Y. Du, Z. Dohnálek and I. Lyubinetsky *The Journal of Physical Chemistry C* **112**, 2649 (2008)
- ¹⁶⁴L.-Q. Wang, D.R. Baer and M.H. Engelhard *Surface Science* **320**, 295 (1994)
- ¹⁶⁵J.-M. Pan, B. L. Maschhoff, U. Diebold and T. E. Madey *Journal of Vacuum Science and Technology A* **10**, 2470 (1992)
- ¹⁶⁶U. Diebold, J. F. Anderson, K.-O. Ng and D. Vanderbilt *Physical Review Letters* **77**, 1322 (1996)
- ¹⁶⁷R. Bechstein, H. H. Kristoffersen, L. B. Vilhelmsen, F. Rieboldt, J. Stausholm-Møller, S. Wendt, B. Hammer, and F. Besenbacher *Physical Review Letters* **108**, 236103 (2012)
- ¹⁶⁸S. Casolo, R. Martinazzo and G. F. Tantardini *The Journal of Physical Chemistry C* **115**, 3250 (2011)
- ¹⁶⁹H. Park, A. Wadehra, J. W. Wilkins and A. H. Castro Neto *Physical Review B* **87**, 085441 (2013)
- ¹⁷⁰J. Heyd, G. E. Scuseria and M. Ernzerhof *Journal of Chemical Physics* **118**, 8207 (2003)
- ¹⁷¹D. M. Bylander and L. Kleinman *Physical Review B*, **41**, 7868 (1990)
- ¹⁷²N. Marzari, D. Vanderbilt and M. C. Payne *Physical Review Letters* **79**, 1337 (1997)
- ¹⁷³B. Hammer, L. B. Hansen and J. K. Nørskov *Physical Review B*, **59**, 7413 (1999)
- ¹⁷⁴J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais *Physical Review B* **46**, 6671 (1992)
- ¹⁷⁵Z. Wu and R. E. Cohen *Physical Review B* **73**, 235116 (2006)
- ¹⁷⁶J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke *Physical Review Letters* **100**, 136406 (2008)

Bibliography

- R. M. Martin *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press (2004)
- P. D. Haynes, *PhD thesis "Linear-scaling methods in ab initio quantum-mechanical calculations"* (1998)
- S. J. Clark, *PhD thesis "Complex Structure in Tetrahedral Semiconductors"* (1994)

Appendix A : Uncertainty analysis for data in chapter 4

A.1 Statistical uncertainty in step energies

For a large geometry optimisation problem, there is a degree of uncertainty in the value of the final energy. For most purposes, this uncertainty is ignored because it is small. However, the γ_{110} , β and q values are quite sensitive to the slab/vacuum supercell energies E_{hkl} and so it is wise to carry out an uncertainty analysis. This is difficult to do analytically because it is highly unlikely that a repeated optimisation of a given supercell would yield a distribution of final energies characterised by a Gaussian curve. The optimisation algorithm searches for the variationally lowest energy state so one might expect an asymmetric distribution with an abrupt lower bound on the energies. One major source of uncertainty is that of systems becoming trapped in configurations corresponding to local metastable energetic minima which lie close to or even far from the global minimum. Also, the specification of finite tolerances introduces systematic errors into the distribution. Nevertheless, it is easy to calculate how sensitive γ_{110} , β and q are to changes in the supercell energies E_{hkl} , which is now shown.

Consider three high-index supercells of L layers thickness with total surface areas $2S_1$, $2S_2$ and $2S_3$ and step-step distances d_1 , d_2 and d_3 . Suppose that these supercells and a bulk unit cell are optimised to give energies E_1 , E_2 , E_3 and E_B . Suppose then that these optimisations are repeated identically to obtain energies E'_1 , E'_2 and E'_3 and E'_B , where any differences between the two sets of results are denoted as $\delta E_1 = E'_1 - E_1$, $\delta E_2 = E'_2 - E_2$, $\delta E_3 = E'_3 - E_3$ and $\delta E_B = E'_B - E_B$. The corresponding γ_{110} and γ'_{110} solutions are

$$\gamma_{110} = \frac{1}{\det D} \begin{vmatrix} \frac{1}{2S_1} \left(E_1 - \frac{m_1}{2} E_B \right) & d_1^{-1} & d_1^{-3} \\ \frac{1}{2S_2} \left(E_2 - \frac{m_2}{2} E_B \right) & d_2^{-1} & d_2^{-3} \\ \frac{1}{2S_3} \left(E_3 - \frac{m_3}{2} E_B \right) & d_3^{-1} & d_3^{-3} \end{vmatrix}, \quad (\text{A.1})$$

$$\gamma'_{110} = \frac{1}{\det D} \begin{vmatrix} \frac{1}{2S_1} \left(E'_1 - \frac{m_1}{2} E'_B \right) & d_1^{-1} & d_1^{-3} \\ \frac{1}{2S_2} \left(E'_2 - \frac{m_2}{2} E'_B \right) & d_2^{-1} & d_2^{-3} \\ \frac{1}{2S_3} \left(E'_3 - \frac{m_3}{2} E'_B \right) & d_3^{-1} & d_3^{-3} \end{vmatrix}, \quad (\text{A. 2})$$

and the absolute difference between the two solutions is $\delta\gamma_{110} = \gamma'_{110} - \gamma_{110}$, given by

$$\delta\gamma_{110} = \frac{1}{\det D} \begin{vmatrix} \frac{1}{2S_1} \left(\delta E_1 - \frac{m_1}{2} \delta E_B \right) & d_1^{-1} & d_1^{-3} \\ \frac{1}{2S_2} \left(\delta E_2 - \frac{m_2}{2} \delta E_B \right) & d_2^{-1} & d_2^{-3} \\ \frac{1}{2S_3} \left(\delta E_3 - \frac{m_3}{2} \delta E_B \right) & d_3^{-1} & d_3^{-3} \end{vmatrix}. \quad (\text{A. 3})$$

By similar reasoning,

$$\delta\beta = \frac{1}{\det D} \begin{vmatrix} 1 & \frac{1}{2S_1} \left(\delta E_1 - \frac{m_1}{2} \delta E_B \right) & d_1^{-3} \\ 1 & \frac{1}{2S_2} \left(\delta E_2 - \frac{m_2}{2} \delta E_B \right) & d_2^{-3} \\ 1 & \frac{1}{2S_3} \left(\delta E_3 - \frac{m_3}{2} \delta E_B \right) & d_3^{-3} \end{vmatrix}, \quad (\text{A. 4})$$

$$\delta q = \frac{1}{\det D} \begin{vmatrix} 1 & d_1^{-1} & \frac{1}{2S_1} \left(\delta E_1 - \frac{m_1}{2} \delta E_B \right) \\ 1 & d_2^{-1} & \frac{1}{2S_2} \left(\delta E_2 - \frac{m_2}{2} \delta E_B \right) \\ 1 & d_3^{-1} & \frac{1}{2S_3} \left(\delta E_3 - \frac{m_3}{2} \delta E_B \right) \end{vmatrix}. \quad (\text{A. 5})$$

To gain a sense of the order of magnitude of the quantities $\delta\gamma_{110}$, $\delta\beta$, and δq , a three layer $\text{Ti}_{26}\text{O}_{52}$ (341) slab/vacuum supercell with the $\langle 1\bar{1}1 \rangle_R$ step was optimised to an energetic minimum four times with identical input parameters to those listed in the method. The four final energies were -64513.033697 eV, -64513.034118 eV, -64513.031639 eV, and -64513.033121 eV, all to 6 decimal places. The range of this sample is 2.489×10^{-3} eV (3. d. p.), or about 3×10^{-5} eV per atom. For the sake of example, let us prescribe an uncertainty of $\pm 1.5 \times 10^{-5}$ eV per atom for all supercell energies. The bulk unit cell was also optimised repeatedly, but the range was found to be $< 10^{-6}$ eV due to the comparatively stringent

tolerances used and so $\delta E_B = 0$ is used. Let us now consider the worst case scenario and borrow the dimensions and total atom numbers from the biggest slabs of lowest symmetry: the 10 layer $\langle 1\bar{1}1 \rangle_R$ $\text{Ti}_{74}\text{O}_{148}$ (341), $\text{Ti}_{94}\text{O}_{188}$ (451) and $\text{Ti}_{114}\text{O}_{228}$ (561) slabs. Let us denote their energies E_{341} , E_{451} and E_{561} respectively, to produce numerical values for δE_{341} , δE_{451} and δE_{561} . Using the nominated uncertainty range of $\pm 1.5 \times 10^{-5}$ eV per atom, this gives $\delta E_{341} = \pm 3.33 \times 10^{-3}$ eV, $\delta E_{451} = \pm 4.23 \times 10^{-3}$ eV and $\delta E_{561} = \pm 5.13 \times 10^{-3}$ eV. The functions $\delta\gamma_{110}$, $\delta\beta$ and δq each depend on the three variables δE_{341} , δE_{451} and δE_{561} and so can't be straightforwardly plotted, so instead they are tabulated in Table A.1 with all possible combinations of the upper and lower limits of δE_{341} , δE_{451} and δE_{561} to show the upper and lower limits of the intervals within which $\delta\gamma_{110}$, $\delta\beta$ and δq lie. For this particular case, the maximum deviations are $\delta\gamma_{110}/\gamma_{110} = \pm 2.7\%$, $\delta\beta/\beta = \pm 12.3\%$ and $\delta q/q = \pm 14.5\%$, which are significant, but not pivotal to the conclusions.

Table A.1. The sensitivity of γ_{110} , β and q to changes in supercell energy.

The instances where γ_{110} , β and q change the most are indicated in bold type.

Solutions taken from 10 layer $\langle 1\bar{1}1 \rangle_R$ slabs			Hypothetical supercell energy changes			Resultant changes to γ_{110} , β and q		
γ_{110} / eVÅ ⁻² (3 d. p.)	β / eVÅ ⁻¹ (3 d. p.)	q / eVÅ (3 d. p.)	δE_{341} / eV (3 d. p.)	δE_{451} / eV (3 d. p.)	δE_{561} / eV (3 d. p.)	$\delta\gamma_{110}$ / eVÅ ⁻² (3 d. p.)	$\delta\beta$ / eVÅ ⁻¹ (3 d. p.)	δq / eVÅ (3 d. p.)
2.505 $\times 10^{-2}$	9.111 $\times 10^{-2}$	3.599	+ 3.330 $\times 10^{-3}$	+ 4.230 $\times 10^{-3}$	+ 5.130 $\times 10^{-3}$	0.234 $\times 10^{-4}$	0.000 $\times 10^{-2}$	0.000
			- 3.330 $\times 10^{-3}$	+ 4.230 $\times 10^{-3}$	+ 5.130 $\times 10^{-3}$	- 4.973 $\times 10^{-4}$	0.849 $\times 10^{-2}$	- 0.319
			+ 3.330 $\times 10^{-3}$	- 4.230 $\times 10^{-3}$	+ 5.130 $\times 10^{-3}$	6.628 $\times 10^{-4}$	- 1.186 $\times 10^{-2}$	0.520
			+ 3.330 $\times 10^{-3}$	+ 4.230 $\times 10^{-3}$	- 5.130 $\times 10^{-3}$	- 1.421 $\times 10^{-4}$	0.337 $\times 10^{-2}$	- 0.201
			+ 3.330 $\times 10^{-3}$	- 4.230 $\times 10^{-3}$	- 5.130 $\times 10^{-3}$	4.973 $\times 10^{-4}$	- 0.849 $\times 10^{-2}$	0.319
			- 3.330 $\times 10^{-3}$	+ 4.230 $\times 10^{-3}$	- 5.130 $\times 10^{-3}$	- 6.628 $\times 10^{-4}$	1.186 $\times 10^{-2}$	- 0.520
			- 3.330 $\times 10^{-3}$	+ 4.230 $\times 10^{-3}$	+ 5.130 $\times 10^{-3}$	1.421 $\times 10^{-4}$	- 0.337 $\times 10^{-2}$	0.201
			- 3.330 $\times 10^{-3}$	- 4.230 $\times 10^{-3}$	- 5.130 $\times 10^{-3}$	- 0.234 $\times 10^{-4}$	0.000 $\times 10^{-2}$	0.000

Appendix B : Statistical fitting methods for surface energies

B.1 Using linear regression to evaluate surface energies is flawed

One commonly employed method of calculating surface energies is to plot the total energies of slabs of increasing size against m , and then to use a linear regression fit to identify the intercept of the best fit line as the total surface energy $2S\gamma$. This approach is potentially very misleading if used in situations where the surface energy (as calculated for individual slabs using equation (4.3) - see main text) converges in a slow oscillatory way with increasing slab thickness. To illustrate the point, the total supercell energies E_{hkl} from the $(430) / \langle 001 \rangle_{Ti}$ slabs in Figure B.1 are now plotted along with their corresponding surface energies $2S\gamma_{hkl}$ as found using equation (4.3). Consider a set of \tilde{M} slab/vacuum supercells of increasing size, labelled $1, \dots, i, \dots, \tilde{M}$. By calculating the total energy of the i^{th} slab, E_i , one then subtracts the (negative) energy of the corresponding quantity of bulk material, $(m_i/2)E_B$, to obtain the total (positive) surface energy of that slab, $2S\gamma_i$. Expressed another way, the total slab energies E_i are related to their corresponding surface energies, $2S\gamma_i$, by a linear shear transformation in the positive direction along the energy axis:

$$\begin{pmatrix} m_i/2 \\ E_i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ E_B & 1 \end{pmatrix} \begin{pmatrix} m_i/2 \\ 2S\gamma_i \end{pmatrix}. \quad (\text{B.1})$$

If this shear transformation is applied to the best fit line, then its gradient is changed, but not its intercept. This of course means that fitting a straight line to E_i plotted against n_i gives exactly the same intercept as fitting a straight line to $2S\gamma_i$ plotted against m_i . Therefore, these two methods of estimating the surface energy are entirely equivalent. The intercept from either of the two best fit lines in Figure B.1 give the same estimate of the surface energy, since the two best fit lines are related by the shear matrix and therefore have the same intercept. There is clearly nothing linear about the trend of $2S\gamma_i$ with increasing m_i in Figure B.1, and one could barely justify using the "best

fit" line shown. The best fit line for the E_i data clearly gives a very good estimate of the *gradient* (i.e. an estimate of the bulk energy, E_B), however the point is that even an excellent estimate of the bulk energy can still produce an intercept which is a poor estimate of the surface energy. This is because the intercept is tiny compared to the values E_i .

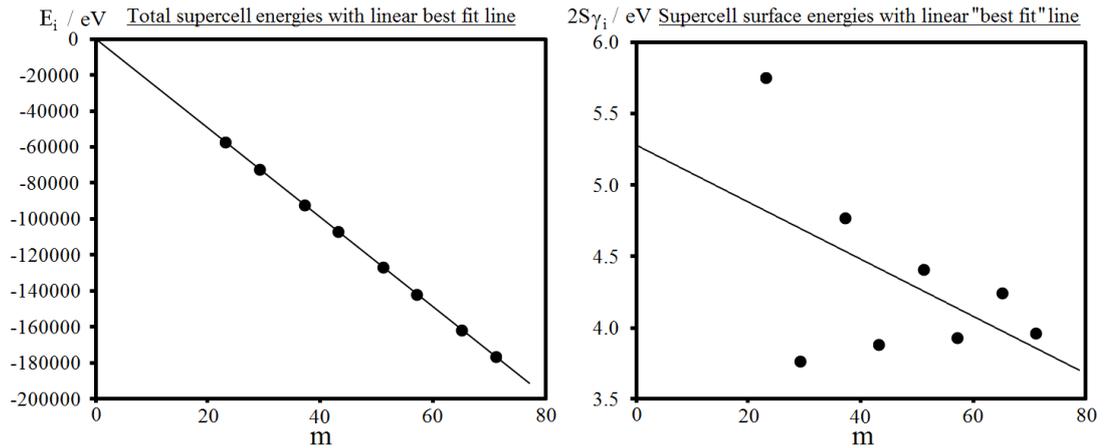


Figure B.1. *The flawed nature of evaluating surface energy using a linear fit of supercell energies against slab size.*

Left panel: Total supercell energies plotted against supercell size with best fit line. **Right panel:** Supercell surface energies plotted against slab size with corresponding best fit line. The two "best fit" lines give the same estimate of the surface energy because they share the same intercept.

The fact that the estimate of the surface energy is poor is made particularly obvious by observing that for the data shown, the intercept is about 5.3 eV, whereas with the surface energy data points $2S\gamma_i$ are quite clearly converging to the region of around 4.1 eV. There are three main sources of motivation behind this extensive discussion. Firstly, the linear fitting approach is based on the false premise that the data points $2S\gamma_i$ are merely a series of measurements of some linear trend subject to random errors characterised by a Gaussian distribution, which is a required assumption in regression models. In fact, the oscillatory behaviour of the surface energies is a physically meaningful, reproduceable, and predictable (in principle) manifestation of alternating modes of electron density configurations. The

linear fitting approach therefore effectively discards \tilde{M} precise and physically insightful surface energy values, $2S\gamma_i$, in favour of one poorly estimated surface energy value obtained via an intercept. Secondly, there is generally no guarantee that selectively omitting data points at small m does in fact completely remove the bias in the calculated intercept. Thirdly, if the linear fitting method is used to calculate the surface energies of high-index slabs, it can potentially give very inaccurate predictions of the energies of steps. If the data points $2S\gamma_i$ oscillate significantly, then the intercept of the "best fit" line also oscillates greatly depending on how many data points and indeed *which* data points are included in the fit. Some examples of this are shown in Table B.1. The instability shown in Table B.1. then propagates through to the solutions γ_{110} , β and q if they are evaluated by using intercept-derived values γ_{hkl} , and these solutions are very sensitive to the intercepts used. In particular, β and q are very small compared to the total surface energy, and they can change by an order of magnitude upon making the most minor of alterations to the data range across which the regression is applied.

Table B.1. *The dependence on the chosen data range of the intercept of a linear fit of total slab energies against the number of atoms.*

The resultant intercept varies dramatically depending on which data points are included in the fit, so the intercept cannot reliably be interpreted as the surface energy.

Slab size (Ti _m O _{2m})	Total supercell energy E_{hkl} / eV	Data range used for linear regression					
		included in fit			not included in fit		
3 layer (Ti ₂₃ O ₄₆)	-57068.4036						
4 layer (Ti ₂₉ O ₅₈)	-71959.2572						
5 layer (Ti ₃₇ O ₇₄)	-91810.0404						
6 layer (Ti ₄₃ O ₈₆)	-106699.8280						
7 layer (Ti ₅₁ O ₁₀₂)	-126551.1090						
8 layer (Ti ₅₇ O ₁₁₄)	-141440.4900						
9 layer (Ti ₆₅ O ₁₃₀)	-161291.9740						
10 layer (Ti ₇₁ O ₁₄₂)	-176181.1710						
Resultant intercept (3 d. p.)		6.299 eV	6.452 eV	3.439 eV	3.937 eV	5.461 eV	3.677 eV

B.2 A nonlinear alternative

Let us consider odd- and even-layered slabs as physically distinct systems, where the total calculated surface energy of a slab consists of a component due to the surface alone, γ , and a component due to surface-

surface interactions which decays exponentially with increasing slab thickness:

$$\gamma^\downarrow(x) = \gamma + ae^{-Bm}, \quad (B.2)$$

$$\gamma^\uparrow(x) = \gamma - ce^{-Dm}, \quad (B.3)$$

where a, B, c, D and γ are all free parameters to be evaluated and m is the number of formula units (which is proportional to slab thickness). These two functions are sketched in Figure B.2.

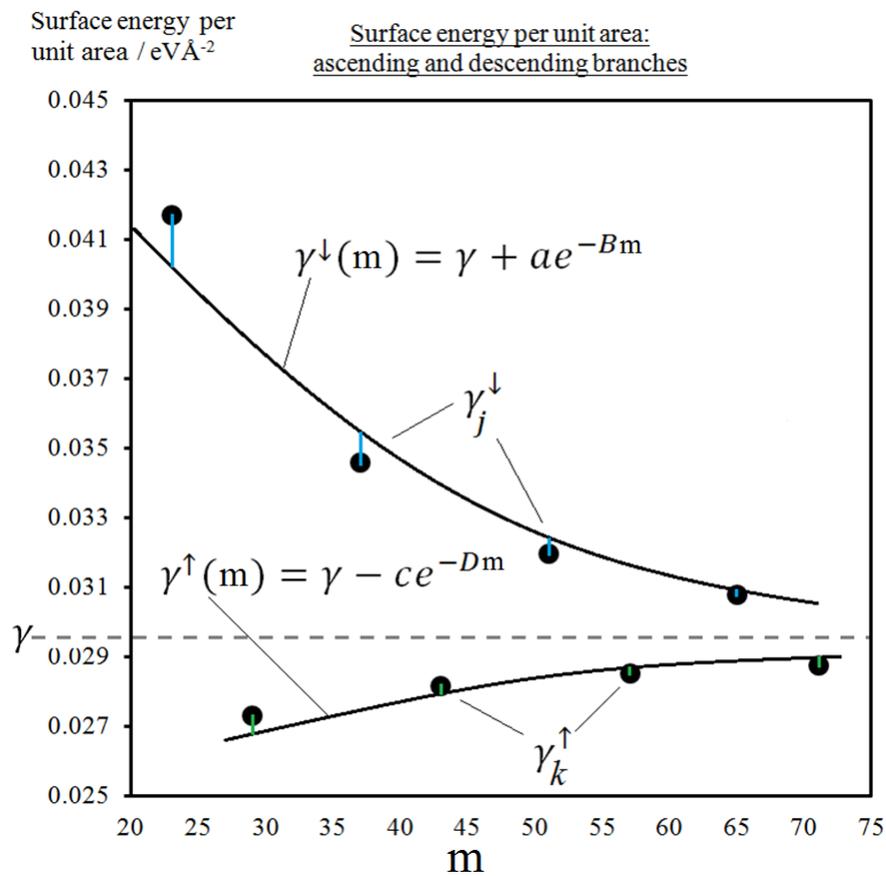


Figure B.2. Supercell slab energies plotted against increasing supercell size with a sketch of the best fit functions $\gamma^\uparrow(m)$ and $\gamma^\downarrow(m)$.

Correspondingly, the \tilde{M} data points γ_i ($i = 1, \dots, \tilde{M}$) can be split into two groups: those on the descending branch, γ_j^\downarrow ($j = 1, \dots, P$), and those on the ascending branch, γ_k^\uparrow ($k = 1, \dots, Q$). Two nonlinear regression procedures can be carried out, whereby the free parameters a, B, c, D and γ are optimised so that the function $\gamma^\downarrow(m)$ is the best fit to the γ_j^\downarrow points and $\gamma^\uparrow(m)$ is the best fit

to the γ_k^\uparrow points. To simplify this process, equations (B.2) and (B.3) can be rearranged and then linearised by taking natural logarithms of both sides:

$$\ln(\gamma^\downarrow - \gamma) = \ln a - Bx, \quad (B.4)$$

$$\ln(\gamma - \gamma^\uparrow) = \ln c - Dx, \quad (B.5)$$

By adopting the notation $\Gamma^\downarrow(m) = \ln(\gamma - \gamma^\downarrow(m))$ and $\Gamma^\uparrow(m) = \ln(\gamma^\uparrow(m) - \gamma)$ and writing $\ln a = A$ and $\ln c = C$, the simple linear form is obtained:

$$\Gamma^\downarrow(m) = A - Bm, \quad (B.6)$$

$$\Gamma^\uparrow(m) = C - Dm. \quad (B.7)$$

The corresponding data points γ_j^\downarrow and γ_k^\uparrow can also be subjected to this linearisation, giving rise to the transformed data points Γ_j^\downarrow and Γ_k^\uparrow :

$$\Gamma_j^\downarrow = \ln(\gamma_j^\downarrow - \gamma) \quad (B.8)$$

$$\Gamma_k^\uparrow = \ln(\gamma - \gamma_k^\uparrow) \quad (B.9)$$

Note that at this point, γ is still a free parameter which has not yet been evaluated, so the numerical values of the data points Γ_j^\downarrow and Γ_k^\uparrow are undetermined despite the values γ_j^\downarrow and γ_k^\uparrow being known. The data points Γ_j^\downarrow and Γ_k^\uparrow can nevertheless be treated as measured values of the quantities modelled by the functions $\Gamma^\downarrow(m)$ and $\Gamma^\uparrow(m)$. The task at hand therefore is to minimise the sum of the squared residuals between the "measured" values Γ_j^\downarrow and the values predicted by the model at the corresponding value of m , $\Gamma^\downarrow(m_j) = A - Bm_j$, and similarly for Γ_k^\uparrow . The total sum of squared residuals over both branches of the plot is therefore:

$$\Sigma = \sum_{j=1}^P (\Gamma_j^\downarrow - A + Bm_j)^2 + \sum_{k=1}^Q (\Gamma_k^\uparrow - C + Dm_k)^2. \quad (B.10)$$

There are five free parameters and five corresponding conditions required to minimise Σ :

$$\frac{\partial \Sigma}{\partial A} = 0, \quad \frac{\partial \Sigma}{\partial B} = 0, \quad \frac{\partial \Sigma}{\partial C} = 0, \quad \frac{\partial \Sigma}{\partial D} = 0, \quad \frac{\partial \Sigma}{\partial \gamma} = 0. \quad (B.11)$$

It is easiest to deal with the A, B, C and D conditions first and then enforce the γ condition afterwards. By evaluating these first four derivatives and setting their values to zero, a system of four linear simultaneous equations is obtained:

$$\frac{\partial \Sigma}{\partial A} = 2 \sum_{j=1}^P (\Gamma_j^\downarrow - A + Bm_j) (-1) = 0 \quad (B.12)$$

$$\frac{\partial \Sigma}{\partial B} = 2 \sum_{j=1}^P (\Gamma_j^\downarrow - A + Bm_j) (m_j) = 0 \quad (B.13)$$

$$\frac{\partial \Sigma}{\partial C} = 2 \sum_{k=1}^Q (\Gamma_k^\uparrow - C + Dm_k) (-1) = 0 \quad (B.14)$$

$$\frac{\partial \Sigma}{\partial D} = 2 \sum_{k=1}^Q (\Gamma_k^\uparrow - C + Dm_k) (m_k) = 0 \quad (B.15)$$

With some minor simplifications, these can be written in matrix form:

$$\begin{pmatrix} \sum_{j=1}^P \Gamma_j^\downarrow \\ \sum_{j=1}^P \Gamma_j^\downarrow m_j \\ \sum_{k=1}^Q \Gamma_k^\uparrow \\ \sum_{k=1}^Q \Gamma_k^\uparrow m_k \end{pmatrix} = \begin{pmatrix} P & -\sum_{j=1}^P m_j & 0 & 0 \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & 0 & 0 \\ 0 & 0 & Q & -\sum_{k=1}^Q m_k \\ 0 & 0 & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} \quad (B.16)$$

The expressions for A, B, C and D can then be found using Cramer's determinant rule:

$$A = \frac{\begin{vmatrix} \sum_{j=1}^P \Gamma_j^\downarrow & -\sum_{j=1}^P m_j & 0 & 0 \\ \sum_{j=1}^P \Gamma_j^\downarrow m_j & -\sum_{j=1}^P m_j^2 & 0 & 0 \\ \sum_{k=1}^Q \Gamma_k^\uparrow & 0 & Q & -\sum_{k=1}^Q m_k \\ \sum_{k=1}^Q \Gamma_k^\uparrow m_k & 0 & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}}{\begin{vmatrix} P & -\sum_{j=1}^P m_j & 0 & 0 \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & 0 & 0 \\ 0 & 0 & Q & -\sum_{k=1}^Q m_k \\ 0 & 0 & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}},$$

$$B = \frac{\begin{vmatrix} P & \sum_{j=1}^P \Gamma_j^\downarrow & 0 & 0 \\ \sum_{j=1}^P m_j & \sum_{j=1}^P \Gamma_j^\downarrow m_j & 0 & 0 \\ 0 & \sum_{k=1}^Q \Gamma_k^\uparrow & Q & -\sum_{k=1}^Q m_k \\ 0 & \sum_{k=1}^Q \Gamma_k^\uparrow m_k & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}}{\begin{vmatrix} P & -\sum_{j=1}^P m_j & 0 & 0 \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & 0 & 0 \\ 0 & 0 & Q & -\sum_{k=1}^Q m_k \\ 0 & 0 & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}},$$

$$C = \frac{\begin{vmatrix} P & -\sum_{j=1}^P m_j & \sum_{j=1}^P \Gamma_j^\downarrow & 0 \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & \sum_{j=1}^P \Gamma_j^\downarrow m_j & 0 \\ 0 & 0 & \sum_{k=1}^Q \Gamma_k^\uparrow & -\sum_{k=1}^Q m_k \\ 0 & 0 & \sum_{k=1}^Q \Gamma_k^\uparrow m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}}{\begin{vmatrix} P & -\sum_{j=1}^P m_j & 0 & 0 \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & 0 & 0 \\ 0 & 0 & Q & -\sum_{k=1}^Q m_k \\ 0 & 0 & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}},$$

$$D = \frac{\begin{vmatrix} P & -\sum_{j=1}^P m_j & 0 & \sum_{j=1}^P \Gamma_j^\downarrow \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & 0 & \sum_{j=1}^P \Gamma_j^\downarrow m_j \\ 0 & 0 & Q & \sum_{k=1}^Q \Gamma_k^\uparrow \\ 0 & 0 & \sum_{k=1}^Q m_k & \sum_{k=1}^Q \Gamma_k^\uparrow m_k \end{vmatrix}}{\begin{vmatrix} P & -\sum_{j=1}^P m_j & 0 & 0 \\ \sum_{j=1}^P m_j & -\sum_{j=1}^P m_j^2 & 0 & 0 \\ 0 & 0 & Q & -\sum_{k=1}^Q m_k \\ 0 & 0 & \sum_{k=1}^Q m_k & -\sum_{k=1}^Q m_k^2 \end{vmatrix}}.$$

(B.17)

Let us rewrite these in a more compact notation:

$$\begin{aligned}
 A &= \frac{\det M_A(\gamma)}{\det M}, & B &= \frac{\det M_B(\gamma)}{\det M}, \\
 C &= \frac{\det M_C(\gamma)}{\det M}, & D &= \frac{\det M_D(\gamma)}{\det M}
 \end{aligned}
 \tag{B.18}$$

where M is the matrix from (B.16) and the γ in brackets indicates dependence on γ . These can be substituted back into the sum of squared residuals expression, which now looks like

$$\begin{aligned}
 \Sigma &= \sum_{j=1}^P \left(\Gamma_j^\downarrow - \frac{\det M_A(\gamma)}{\det M} + \frac{\det M_B(\gamma)}{\det M} m_j \right)^2 \\
 &\quad + \sum_{k=1}^Q \left(\Gamma_k^\uparrow - \frac{\det M_C(\gamma)}{\det M} + \frac{\det M_D(\gamma)}{\det M} m_k \right)^2.
 \end{aligned}
 \tag{B.19}$$

The $\frac{\partial \Sigma}{\partial \gamma} = 0$ condition then reads:

$$\begin{aligned}
 \sum_{j=1}^P \left(\left(\frac{-1}{\gamma_j^\downarrow - \gamma} \right) - \frac{\partial}{\partial \gamma} \left(\frac{\det M_A}{\det M} \right) + \frac{\partial}{\partial \gamma} \left(\frac{\det M_B}{\det M} \right) m_j \right) \\
 + \sum_{k=1}^Q \left(\left(\frac{-1}{\gamma_j^\downarrow - \gamma} \right) - \frac{\partial}{\partial \gamma} \left(\frac{\det M_C}{\det M} \right) + \frac{\partial}{\partial \gamma} \left(\frac{\det M_D}{\det M} \right) m_k \right) = 0.
 \end{aligned}
 \tag{B.20}$$

It is difficult to solve for γ analytically, but easy to solve numerically by carrying out successive iterations on the value of γ . The initial guess of γ must be larger than all values of γ_k^\uparrow and smaller than all values of γ_j^\downarrow otherwise the quantities $\ln(\gamma_j^\downarrow - \gamma)$ or $\ln(\gamma - \gamma_k^\uparrow)$ are undefined. Within this range, there is always a solution for γ . i.e. there is always an asymptote which lies in between the two branches but does not intersect them.

B.3 Application: merits and limitations

This fitting procedure is now carried out on the data presented in Table B.2 and the best fit functions $\gamma^\downarrow(m)$ and $\gamma^\uparrow(m)$ are plotted in Figure B.3 along with the data points γ_j^\downarrow and γ_k^\uparrow . The two functions converge towards the asymptote $\gamma = 0.030076 \text{ eV\AA}^{-2}$ (6 d. p.), which is the desired number.

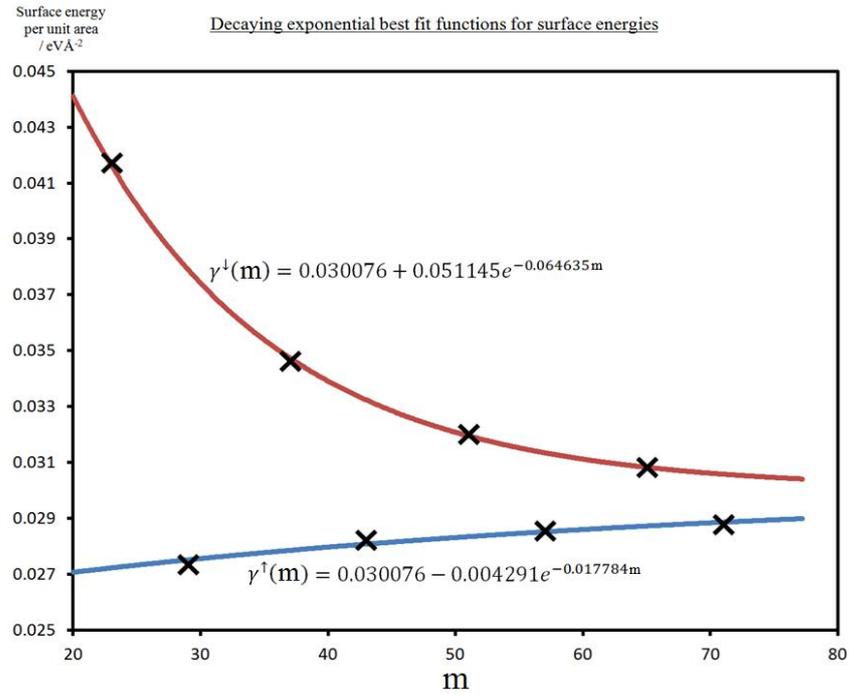


Figure B.3. Supercell slab energies plotted against increasing supercell size with optimised best fit functions $\gamma^{\uparrow}(m)$ and $\gamma^{\downarrow}(m)$.

This method could potentially be useful because it provides a formal method for deciding on the value of γ from a given plot of slowly-converging surface energy data points. This is especially important if the convergence is slow for the largest computationally-practical slabs. The most severe limitation of this method is that there is no clear physical justification for writing $\gamma^{\uparrow}(m)$ and $\gamma^{\downarrow}(m)$ as decaying exponential functions. The method can be easily repeated using hyperbolae, and probably many other arbitrary functions which can always be made to fit. Also, it is not clear how to decide on confidence regions for the optimised parameters, because the deviations of the data points from the fitted functions are unlikely to be characterised by a Gaussian distribution.

Table B.2. Supercell energies calculated using CASTEP divided into two branches.

Surface energy data										
System	m_j	m_k	Supercell energy, E_j / eV (4 d. p.)	Supercell energy, E_k / eV (4 d. p.)	Bulk energy per formula unit, E_B / eV (4 d. p.)	Surface area, $2S$ / \AA^2 (4 d. p.)	Surface energy, $2S\gamma_j^\dagger$ / eV (4 d. p.)	Surface energy, $2S\gamma_k^\dagger$ / eV (4 d. p.)	Surface energy per unit area, γ_j^\dagger / $\text{eV}\text{\AA}^{-2}$ (4 d. p.)	Surface energy per unit area, γ_k^\dagger / $\text{eV}\text{\AA}^{-2}$ (4 d. p.)
TiO₂ rutile (430) slabs	23		-57068.2976		-2481.4804	137.7513	5.7508		0.0417	
		29		-71959.1617	-2481.4804	137.7513		3.7689		0.0274
		37	-91810.0026		-2481.4804	137.7513	4.7709		0.0346	
		43		-106699.7690	-2481.4804	137.7513		3.8867		0.0282
		51	-126551.0880		-2481.4804	137.7513	4.4106		0.0320	
		57		-141440.4450	-2481.4804	137.7513		3.9358		0.0286
		65	-161291.9760		-2481.4804	137.7513	4.2477		0.0308	
	71		-176181.1380	-2481.4804	137.7513		3.9679		0.0288	
Results of best fit analysis										
Parameter					Optimised value (6 d. p.)					
a / $\text{eV}\text{\AA}^{-2}$					0.051145					
B					0.064635					
c / $\text{eV}\text{\AA}^{-2}$					0.004291					
D					0.017784					
γ / $\text{eV}\text{\AA}^{-2}$					0.030076					