Density functional theory (DFT) investigations of defects in three materials systems have enabled an enhanced understanding of how the defects influence the final properties of materials. Point defects, surfaces, and the heterogeneous interface between dissimilar materials are all discussed. The formation energies of point defects in bulk AlN are used to show that the source of the unwanted deep ultraviolet absorption is carbon incorporation during growth. The carbon in these samples is also shown to form a donor-acceptor pair with nitrogen vacancies, leading to a 2.8 eV optical emission. The co-doping of these high-carbon AlN samples with Si and O donors is shown to render the samples optically transparent because the donor level acts as a trapping center for electrons and leads to a long-lived state. Finally, the challenges of n-type doping in AlN are discussed based on the activation energies of both the Si and O donors are discussed in terms of the DX-type transition between a shallow donor state and a deep acceptor state. To study the stable surfaces of polar rocksalt oxides, DFT calculations are extended to experimentally relevant temperatures and pressures using ab initio thermodynamics. This technique allowed the prediction of an experimental growth window where polar surfaces could be stabilized using a hydrogen surfactant which acted as a polarity compensator. The heterogeneous interface between wurtzite GaN and rocksalt MgCaO was examined based on the site preference during the initial stages of epitaxial oxide growth in order to explain the strong observed preference for MCO to have only one rotational configuration on GaN.
Predictive Calculations of Defects in Thin-Film Heterostructures

by
Benjamin Erik Gaddy

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Materials Science and Engineering

Raleigh, North Carolina

2013

APPROVED BY:

__________________________  _________________________
Donald W. Brenner        Jon-Paul Maria

__________________________  _________________________
Ramón Collazo             Phillip Westmoreland

__________________________
Douglas L. Irving
Chair of Advisory Committee
DEDICATION

To my parents,
who encouraged and fostered my interests in science and computers ever since I was a child.

To my siblings,
Kristina
who always inspires and uplifts us,
and Thomas
whose work ethic I aspire to.

To Matney and Miller,
whose boundless love and support make each day a treasure.
BIOGRAPHY

The author was born in Stockholm, Sweden and was raised in Kensington, Maryland.

He attended North Carolina State as a Park Scholar, where, on one cold December saturday in 2004 he ran from the campus bell tower two miles to the Krispy Kreme store to eat a dozen doughnuts – then ran back, thus becoming the first person to ever complete the Krispy Kreme Challenge. He became one of the founders of the event, which has now been run nine times and has raised $347,000 for the North Carolina Children’s Hospital. During his undergraduate career he was also the captain of the NC State Rowing Team.

After graduating in 2007 with a B.S. in Electrical Engineering, he worked as a patent examiner at the United States Patent and Trademark Office. He returned to Raleigh and began his graduate studies in MSE with professor Doug Irving in August of 2009.
ACKNOWLEDGEMENTS

The work presented in this dissertation could not have been done without the help of my advisor, group members, collaborators, mentors, friends, and family, and could not have been completed without generous financial support from a number of sources.

I would like to thank Professor Doug Irving for his advice, support, and leadership during my graduate career, and the members of the Irving group: Carissa Lada, Chris Freeze, Josh Harris, Changning Niu, Xiaoyin Ji, Patrick Chiu, and Nick Restaino for their guidance and helpful discussions. The work in this dissertation has benefitted immensely from close collaborations and discussions with Professors Jon-Paul Maria, Ramón Collazo, Zlatko Sitar, Jim LeBeau, and Mark Losego at NC State, and their students Beth Paisley, Chris Shelton, Zachary Bryan, Isaac Bryan, and James Tweedie. I would also like to thank Ronny Kirste, Jinqiao (Andy) Xie, Seiji Mita, Marc Hoffman, Rafael Dalmau, Baxter Moody, Yoshinao Kumagai, Toru Nagashima, Yuki Kubota, Toru Kinoshita, Akinori Koukitu, Michael Biegalski, and Hans Christen for their contributions to the projects presented here.

My research has been funded thanks to the generous support of the National Defense Science and Engineering Graduate Fellowship, the Department of Education Graduate Assistantship in areas of National Need in Computational Science, and the NCSU startup funds of Professor Irving. Additional project and conference travel support has been provided by the U.S. Navy Office of Naval Research, the NSF CAREER Award of Professor Irving, and NCSU MSE startup funds of Professors Doug Irving and Ramón Collazo, and generous support from Professor Zlatko Sitar.

I would like to thank all of those who reviewed versions of this dissertation: Doug Irving, Carissa Lada, Josh Harris, Changning Niu, and my committee members. Their edits and comments have been invaluable, and any errors that remain are solely my own.

I would like to thank my committee members for generously giving their time and insight into my projects and into this dissertation: Donald Brenner, Jon-Paul Maria, Phillip Westmoreland, Ramón Collazo, and Doug Irving.

I would like to thank my friends and family for their unwavering encouragement and support during my undergraduate and graduate career.
# TABLE OF CONTENTS

LIST OF TABLES ......................................................... ix

LIST OF FIGURES ......................................................... xi

Chapter 1 Introduction ..................................................... 1
  1.1 Defects in Materials ................................................. 1
  1.2 Wide-Bandgap Semiconductors and Their Defects .................. 2
    1.2.1 Summary of Findings ....................................... 3
  1.3 Thin Films and the Challenges of Polarity ....................... 4
    1.3.1 Summary of Findings ....................................... 4
  1.4 Interfaces Between Dissimilar Materials ......................... 5
    1.4.1 Summary of Findings ....................................... 5
  1.5 Overall Summary of Results ...................................... 5

Chapter 2 Density Functional Theory ................................... 7
  2.1 The Basics of DFT ................................................ 7
  2.2 The Schrödinger Equation ....................................... 8
  2.3 The Variational Principle ...................................... 9
  2.4 Many Electrons .................................................. 10
  2.5 The Born-Oppenheimer Approximation ........................... 10
  2.6 Slater Determinants ............................................ 11
  2.7 Thomas-Fermi and the electron density ......................... 12
  2.8 Functionals ..................................................... 13
  2.9 Hohenberg, Kohn, and Sham .................................... 13
  2.10 Hohenberg & Kohn ............................................... 14
  2.11 Kohn & Sham .................................................... 15
  2.12 Exchange-Correlation Functionals .............................. 16
    2.12.1 The Local Density Approximation ....................... 16
    2.12.2 Generalized Gradient Approximation .................... 17
    2.12.3 Post-GGA functionals ................................... 17
    2.12.4 Meta-GGA functionals ................................... 18
    2.12.5 Hyper-GGA functionals .................................. 18
    2.12.6 Semi-local functionals: HSE ............................ 19
  2.13 Pseudopotentials .............................................. 20
  2.14 Bloch’s Theorem ................................................. 20

Chapter 3 DFT in Practice ............................................... 22
  3.1 The anatomy of a DFT calculation .............................. 22
  3.2 Atomic Relaxation ............................................... 23
  3.3 Electronic Structure ............................................. 23
    3.3.1 Density of Electronic States ......................... 24
    3.3.2 Band Structure ........................................... 24
    3.3.3 Charge Density ........................................... 25
3.3.4 Electrostatic potential ........................................... 25

Chapter 4 Thermodynamics ........................................... 26
4.1 Reaction Energetics ............................................. 26
4.2 Enthalpy of Formation ......................................... 27
4.3 Formation Energies of Surfaces .............................. 29
4.4 Temperature and Pressure Dependence ..................... 31
4.5 Formation Energies of Point Defects ....................... 31
4.6 Chemical Boundary Conditions .............................. 33
4.7 Chemical potentials of impurities ......................... 35
4.8 Defect Concentration .......................................... 35
4.8.1 More Complicated Defects and Their Entropies ......... 37

Chapter 5 Calculation Details - Setup ............................. 39
5.1 Equilibrium Geometries ........................................ 40
5.2 Equilibrium Geometry of Tetragonal and Hexagonal Unit Cells 40
5.3 HSE Band Gap Calculations .................................. 43
5.4 Determining High-Symmetry Paths ......................... 44
5.5 Effects of HSE Mixing ......................................... 45
5.6 Relaxation Shell ................................................ 46
5.7 Defects of Interest ............................................ 46
5.8 Surface Geometries ............................................ 47

Chapter 6 Post Processing .............................................. 48
6.1 Point Defects .................................................... 48
6.1.1 Point Defect Formation Energies ....................... 48
6.1.2 Formation Energy Diagrams .............................. 50
6.1.3 Thermodynamic Transitions ............................ 55
6.1.4 Optical Transitions ....................................... 57
6.1.5 Law of Mass Action ....................................... 59
6.1.6 Equilibrium Concentrations of Defects ................ 60
6.2 Surfaces ......................................................... 62
6.2.1 Surface Free Energy Calculations ....................... 62
6.2.2 Extending DFT results with thermochemical tables ... 63
6.2.3 Bader Charge Analysis .................................. 65

Chapter 7 The source of the 4.7 eV absorption band in AlN ......... 67
7.1 Background ..................................................... 68
7.2 Computational Methodology .................................. 68
7.2.1 The “Band Gap Problem” in AlN ....................... 68
7.2.2 Hybrid DFT ................................................ 68
7.3 Defects Studied ............................................... 69
7.4 Results ........................................................ 69
7.5 Summary of Experimental Confirmation ................... 72
7.6 Conclusions .................................................... 73
### Chapter 8  Thermal activation energy of Si and O DX donors in Al\(_x\)Ga\(_{1-x}\)N

- **8.1 Introduction** ................................................................. 77
- **8.2 Results & Discussion** .................................................... 78
  - 8.2.1 Conductivity and activation energy change as a function of Si content 78
  - 8.2.2 Structure of substitutional and DX centers .......................... 78
  - 8.2.3 Self-consistent mass balance approach ............................. 84
  - 8.2.4 Configuration coordinate to determine barriers .................... 85
- **8.3 Conclusions** .............................................................. 86
- **8.4 Computational Setup** .................................................. 86

### Chapter 9  The Donor-Acceptor Pair of Carbon and Nitrogen Vacancies in Bulk AlN

- **9.1 Authors** ................................................................. 88
- **9.2 Introduction** ............................................................. 89
- **9.3 Results and Discussion** .............................................. 90
  - 9.3.1 Photoluminescence Measurements .................................. 90
  - 9.3.2 Equilibrium Defect Concentrations ............................... 92
  - 9.3.3 DAP Model .................................................................. 92
- **9.4 Conclusions** .............................................................. 93
- **9.5 Methods** ................................................................. 94
  - 9.5.1 Formation Energy of Defects ....................................... 94
  - 9.5.2 Mass Balance Equations ............................................ 95

### Chapter 10  Enabling the smooth growth of rocksalt oxides by changing the chemical environment.

- **10.1 Introduction** ............................................................. 98
- **10.2 Computational Methodology** ....................................... 98
- **10.3 Results** ................................................................. 105
- **10.4 Summary of Experimental Confirmation** ......................... 106
- **10.5 Discussion** .............................................................. 106

### Chapter 11  Overcoming the polarization catastrophe in epitaxial thin-film rocksalt oxides (MgO and CaO)

- **11.1 Introduction** ............................................................. 108
- **11.2 The Polarization Catastrophe** ....................................... 109
- **11.3 Calculated Surface Free Energies** ................................. 112
- **11.4 Competition between Compensation Mechanisms** ................ 117
- **11.5 Energy Cost of Electronic Compensation** ....................... 118
- **11.6 Conclusions** ............................................................. 118

### Chapter 12  Understanding the heterogeneous interface between two polar materials.

- **12.1 Introduction** ............................................................. 121
- **12.2 Computational Methodology** ....................................... 122
- **12.3 Rotational Variants** .................................................. 122
12.4 Rotational Preference ...................................................... 123
12.5 Early Growth Preference ................................................. 125
12.6 Consequences of a Step-Edge ............................................ 125
12.7 Conclusions ................................................................. 126

Chapter 13 Conclusions and Future Work ................................. 128
13.1 Conclusions ................................................................. 128
13.2 Future Work ................................................................. 129
  13.2.1 Sodium Defects in Ammonothermal Growth ......................... 129
  13.2.2 Strain cracking in GaN .................................................. 129
  13.2.3 Void formation in AlN .................................................... 129
  13.2.4 The MCO-GaN Interface under varying growth conditions ........ 130

References .............................................................................. 131


### LIST OF TABLES

| Table 5.1 | The first step of the energy minimization procedure involves a coarse sweep based on an initial trial volume. | 42 |
| Table 5.2 | The second step of the energy minimization procedure involves choosing values of \( c/a \) around a central value, which is then used to determine \( a \). | 42 |
| Table 6.1 | The standard thermochemical table values obtained from a reference such as the JANAF tables, shown here for gas-phase molecular \( O_2 \). [Stull and Prophet, 1971] | 64 |
| Table 6.2 | The thermochemical data in the JANAF table of \( O_2 \) has been converted to a form that allows simple calculation of a line of best fit for the temperature dependence. [Stull and Prophet, 1971] | 64 |
| Table 6.3 | Calculated bond-valence parameters for MgO and CaO bonds. The calculated DFT parameters are compared to those given in Ref. [Brese and O’Keeffe, 1991]. | 66 |
| Table 7.1 | Lattice parameters, cohesive energies, and band gaps as calculated using five standard density functional approximations (DFAs). Experimental lattice parameters are taken from [Iwama et al., 1982] and experimental band gap is taken from [Dalmau et al., 2010]. | 69 |
| Table 7.2 | The calculated lattice parameters, formation enthalpy, and band gaps according to a standard GGA density functional (PBE) and the HSE functional with 32% exact exchange, compared with experiment. Experimental lattice parameters are taken from [Iwama et al., 1982] and experimental band gap is taken from [Dalmau et al., 2010]. | 70 |
| Table 7.3 | Summary of the isolated native and impurity defects and defect complexes calculated to determine the origin of the 4.7 eV absorption. 180 calculations were performed for the defects listed in this table. | 71 |
| Table 8.1 | Comparison of energies and nearest-neighbor bond lengths in on-site SiAl (d) and the two DX variants in AlN. | 81 |
| Table 8.2 | Comparison of energies and nearest-neighbor bond lengths for on-site \( O_N \) (d\(^9\)) and the two DX variants in AlN. | 82 |
| Table 8.3 | Equilibrium defect concentrations as determined by the mass-balance procedure described in the text for samples with varying Si concentration and a fixed oxygen background of \( 1 \times 10^{18} \text{cm}^{-3} \). | 85 |
| Table 10.1 | Surface configurations evaluated with ab initio thermodynamics calculations: The 16 surface configurations considered for surface energy calculations and the label by which they are referred in Figure 10.15 | 99 |
| Table 11.1 | Descriptions of the surface configurations explored in this study. The plot labels are used to refer to the surfaces in subsequent discussions. | 113 |
Table 11.2 Total 0K energies (including ground state vibration) for the surfaces considered. All energies are shown at a growth condition midway between Oxygen-rich and Oxygen-poor limits.

Table 11.3 The charge of each layer in the surface caps of MgO with different polarity compensation mechanisms. The regions have been assigned according to Goniakowski’s method. In each case, the sum of the surface caps is zero, indicating that polarity has been compensated. White cells indicate hydrogen surfactant layers, red cells indicate oxygen layers, blue cells indicate magnesium layers, and grey cells indicate a cell that is shared with the center region, according to Goniakowski’s method.

Table 11.4 The charge of each layer in the surface caps of CaO with different polarity compensation mechanisms. The regions have been assigned according to Goniakowski’s method. In each case, the sum of the surface caps is zero, indicating that polarity has been compensated. White cells indicate hydrogen surfactant layers, red cells indicate oxygen layers, blue cells indicate calcium layers, and grey cells indicate a cell that is shared with the center region, according to Goniakowski’s method.
LIST OF FIGURES

Figure 6.1 The formation energy $E_f$ of $V_{\text{Al}}$ defects are shown as a function of the growth conditions $\Delta \mu$ ranging from N-rich to Al-rich, at three different fixed values of the Fermi level: (a) $E_F = 0$, the valence band maximum, (b) $E_F = E_g/2$, at the mid-gap, and (c) $E_F = E_g$, at the conduction band minimum. ................................................. 51

Figure 6.2 The formation energy $E_f$ of $V_{\text{Al}}$ defects are shown as a function of the Fermi level $E_F$ ranging from the valence band maximum ($E_F = 0$) to the conduction band minimum ($E_F = E_g$), at three different fixed growth conditions: (a) $\Delta \mu = +\Delta H$, N-rich conditions, (b) $\Delta \mu = 0$, halfway between the extremes, and (c) $\Delta \mu = -\Delta H$, Al-rich conditions. .......... 52

Figure 6.3 The formation energies $E_f$ of $V_{\text{Al}}$ and $V_{\text{N}}$ defects are shown as a function of the growth conditions $\Delta \mu$ ranging from N-rich to Al-rich. In this figure, only the lowest energy charge state of each defect is shown. Each defect is shown at three different fixed values of the Fermi level: (a) $E_F = 0$, the valence band maximum, (b) $E_F = E_g/2$, at the mid-gap, and (c) $E_F = E_g$, at the conduction band minimum. ................................. 53

Figure 6.4 The formation energies $E_f$ of $V_{\text{Al}}$ and $V_{\text{N}}$ defects are shown as a function of the Fermi level $E_F$ ranging from the valence band maximum ($E_F = 0$) to the conduction band minimum ($E_F = E_g$). In this figure, only the lowest energy charge state of each defect is shown. Each defect is shown at three different fixed growth conditions: (a) $\Delta \mu = +\Delta H$, N-rich conditions, (b) $\Delta \mu = 0$, halfway between the extremes, and (c) $\Delta \mu = -\Delta H$, Al-rich conditions. ................................................. 54

Figure 6.5 Band diagram showing the transitions for the aluminum vacancy in AlN. The transitions between the +1, 0, and -1 charge states are shown. The arrows indicate the acceptor and donor activation energies for the $q = 0$ defect. ................................................................. 56

Figure 6.6 Schematic drawing of band transitions related to optical absorption and emission for the $V_{\text{Al}}$ defect in AlN. ................................................................. 57

Figure 7.1 Change in the conduction band, valence band, and band-gap for GaN as a function of the amount of exact exchange. ......................................................... 70

Figure 7.2 Formation energy for substitutional carbon defects for (a) nitrogen-rich and (b) aluminum-rich growth conditions. ......................................................... 72

Figure 7.3 Comparison of the bonds in bulk AlN (left) to the substitutional carbon defect in the negative (center) and neutral (right) charge states. The significant breathing mode relaxation between the negative and neutral defects corresponds with a large relaxation energy shown in the configuration-coordinate diagram. ......................................................... 73
Figure 7.4 Experimental absorption measurements of samples with varying carbon concentration. a) Absorption at 4.7 eV decreases with decreasing carbon concentration. b) A linear relationship exists between the magnitude of the absorption peak at 4.7 eV and carbon concentration. Figures are from the original manuscript. [Collazo et al., 2012] 74

Figure 7.5 Experimental photoluminescence measurements of samples with varying carbon concentration. High carbon samples show a prominent emission peak at 3.9 eV which corresponds to the carbon acceptor transition. This figure appears in the original publication. [Collazo et al., 2012] 75

Figure 7.6 Configuration coordinate diagram for the optical transition of carbon substitution on the nitrogen site. The energies indicated are in eV. 75

Figure 8.1 Measured compensation ratio (red) and activation energy of donors (blue) for three AlN samples with varying Si concentration. For low and high Si concentration (low and high silane flow) the compensation ratio is high, while it falls to a minimum when Si concentration is on the order of the background oxygen. Above a Si threshold, the activation energy is lowered to about 250 meV. Figure courtesy Collazo and Sitar groups. 79

Figure 8.2 Configurations of Si donor substituting for the III-atom (Al or Ga) in wurtzite AlGaN. Small blue spheres represent N, large green spheres represent the III atom, and the large yellow sphere represents Si. a) shows the standard on-site substitution, b) shows the $\gamma$-BB where the broken bond is in the c-direction, and c) shows the $\alpha$-BB, where the broken bond is in one of the three non-c directions. 80

Figure 8.3 Defect formation energies of dopants in AlN. (a) Substitutional Si on the Al site; (b) substitutional oxygen on the nitrogen site. 83

Figure 8.4 (a) The simulated compensation ratio as a function of Si content is shown for AlN samples with several different concentrations of background oxygen. Compare to experimental results shown in Figure 8.1. In both cases, low and high Si content corresponds to a greater compensation ratio, while there is a minimum for intermediate Si content. (b) The transition between regions where $[\text{Si}] < [\text{O}]$ to where $[\text{Si}] > [\text{O}]$ corresponds to a shift in the equilibrium Fermi level from around 5.55 (near the thermodynamic transition level for O) to 6.07 (above the transition for Si). For high [Si] content, the negatively charged Si DX2- is the energetically preferred charge state. 84

Figure 8.5 Configuration coordinate diagram showing the transitions between the shallow donor state and the deep acceptor state of the Si defect. 86

Figure 9.1 Photoluminescence measurements of AlN samples with varying carbon concentrations shows that the 2.8 eV emission peak is correlated with higher carbon concentrations. Courtesy of Collazo and Sitar groups. 90
Figure 9.2 Formation energy of vacancies and substitutional carbon defects in AlN. Formation energy as a function of the Fermi level is shown for the extreme growth conditions: Nitrogen rich (a) and aluminum rich (b). Experimental conditions used to grow the samples mentioned in this report are closer to the Al rich extreme. 91

Figure 9.3 Equilibrium concentration of defects in AlN as a function of growth conditions. Concentrations are shown as quenched at 300K. 92

Figure 9.4 The vacancy concentration follows the carbon concentration, indicating that the vacancy is the primary compensating defect. 93

Figure 9.5 The configuration-coordinate diagram describes the transitions that would take place in the material when carbon and the nitrogen vacancy are interacting as a DAP. 94

Figure 10.1 Schematic illustration of the cleavage planes of bulk CaO to generate the surface orientations studied. The cleavage planes illustrated are (a) {001}, (b) {011}, and c) {111}. 100

Figure 10.2 Side view of the {100} CaO surface configuration with a bare, unreconstructed surface. 100

Figure 10.3 Side view of the {100} CaO surface configuration with a hydroxyl layer situated atop the surface atoms. 100

Figure 10.4 Side view of the {100} CaO surface configuration with a hydroxyl layer situated above the surface, in a bridge site along the {110} direction. 101

Figure 10.5 Side view of the {100} CaO surface configuration with a half monolayer of associatively adsorbed water molecules. 101

Figure 10.6 Side view of the {100} CaO surface configuration with a full monolayer of associatively adsorbed water molecules. 101

Figure 10.7 Side view of the {100} CaO surface configuration with a p(3x2) adsorbate symmetry, featuring 4 associatively adsorbed water molecules, two hydroxyl groups, and two hydrogens in each repeat unit. 102

Figure 10.8 Side view of the {110} CaO surface configuration with a bare, unreconstructed surface. 102

Figure 10.9 Side view of the {110} CaO surface configuration with a hydroxyl termination. 102

Figure 10.10 Side view of the {111} CaO surface configuration with a bare, unreconstructed surface. 102

Figure 10.11 Side view of the {111} CaO surface configuration with the cation-terminated octopolar reconstruction. 103

Figure 10.12 Side view of the {111} CaO surface configuration with hydroxyl termination. 103

Figure 10.13 Side view of the {111} CaO surface configuration with a hydroxyl termination and a half monolayer of associatively adsorbed water molecules. 104

Figure 10.14 Side view of the {111} CaO surface configuration with a hydroxyl termination and a full monolayer of associatively adsorbed water molecules. 104
Figure 10.15 Temperature dependent comparison of relevant CaO surfaces: Surface free energy values as a function of temperature for a collection of relevant CaO surface orientations and chemical terminations involving bulk, bulk reconstructed, hydroxylated and hydrated. The grey region marks the upper limit of the chemical potential.

Figure 10.16 Leakage current comparison for CaO films prepare in oxygen and water vapor: Leakage current analysis of 4.5 nm CaO capacitors. (a) Surfactant-assisted CaO plus capping layer, and (b) conventional CaO plus capping layer. Each trace is an average of data captured for 20 capacitors. The error bars correspond to the standard deviation values at each voltage. The offset in field corresponds to 0.1 V and is consistent with the Schottky barrier known to exist at the GaN|CaO interface. [Craft et al., 2008] The figure appears in the original paper. [Paisley et al., 2011]

Figure 11.1 The surface energy of various surface configurations of MgO (black squares) and CaO (grey circles) as a function of the number of broken bonds per unit area. The surface labels correspond to Table 2. The trend lines show a fit through the bare, unreconstructed surfaces for both materials.

Figure 11.2 Schematic illustration of the polarization catastrophe. (a) shows a representative polar orientation consisting of 6 planes of opposing charge, where arrows indicate a dipole, dark arrows are uncompensated dipoles, and the arrow length defines the magnitude of the dipole. (b) shows a plot of the electric field (E) and electrostatic potential (V) for the system depicted in (a). (c) shows an example of a compensated polar orientation with half charges on the surfaces, and (d) shows the field and potential for this case.

Figure 11.3 Comparison of the spatial decomposition according to Tasker (left) and Goniakowski (right) of the divergent (top) and convergent (bottom) systems of Figure 11.2.

Figure 11.4 Surface phase diagrams for MgO (left) and CaO (right). The lowest energy surface is shown for each material as a function of the chemical potential of oxygen and water. The chemical potentials are referenced to 0K enthalpies, as described in the text.

Figure 11.5 The density of electronic states is shown for surface terminations of CaO, and compared to the Bulk density of states. Higher energy surfaces, such as the bare unreconstructed \{111\} surface show states pushed into the band gap.

Figure 12.1 STEM micrograph of the GaN—MCO interface. The left and right sections show layer-by-layer epitaxial growth, while the region above the step-edge shows disorder.

Figure 12.2 Side view of the interface between MCO and GaN. The R1 (left) and R2 (right) configurations are compared.
Figure 12.3  The relaxed atomic coordinates of the MCO-GaN interface in the $R_1$ and $R_2$ orientations, superimposed upon a STEM micrograph of the interface. STEM image courtesy J.M. LeBeau. 

Figure 12.4  Potential energy surface of the interface between GaN and MCO during the first layer of growth. 

Figure 12.5  The $R_1$ configuration shown above (left) and below a step edge. 

Figure 12.6  Comparison of the atomic positions of the cations above and below a step edge, as predicted by DFT (left) to the STEM image of the as-grown sample.
Chapter 1

Introduction

1.1 Defects in Materials

The structure and properties of real materials are defined by their defects. Unintentional defects may be detrimental to the performance of the material. The incorporation of point defects in semiconductors can cause a substrate to crack during growth, and self-compensating defect centers can render the material opaque when applications require transparency. Or, the rough surface and accompanying domain boundaries in an insulating oxide buffer layer can provide pathways for detrimental leakage current. These defects often form as a byproduct of the processing conditions under which the material is prepared, handled, or used. In other cases, defects can be purposefully incorporated into a material to imbue a set of desired properties. Dopants in semiconductors, dislocations in ductile metals, catalytic surfaces, and p-n junctions are just a few examples of defects which are critical to the performance of materials. A full understanding of defects – how they form, how they interact, and how they affect the material properties – is critical to the creation of the next generation of advanced materials. Probing these defects using only one method of investigation may not give the full picture of how to enhance the formation of desired defects or to suppress the formation or adverse behavior of undesirable defects. This dissertation presents a framework for integrating computational studies of the formation and structure of defects with experiment to gain a better understanding of how the structure and properties of defects define the properties and functionality of their host material.

Density Functional Theory (DFT) is used as the basis for determining the atomic and electronic structure of defects with tight coupling to experimental investigations. In order to demonstrate the power of this integrated approach for defects of varying dimensionality, this dissertation will discuss defects of three types: point defects, surfaces, and interfaces.

Three specific defect systems that are coupled to experimental investigations are explored:

1. Impurity and native point defects in Bulk AlN. The formation energetics of these defects
are studied to describe and explain the experimentally observed optical and electrical properties, namely: the origin of an unwanted deep-UV absorption, the activation energy of donors in n-type material, and the vacancy-related compensation due to Donor-Acceptor Pair (DAP) formation. The interaction of defects via the Fermi level is studied to predict doping strategies that will overcome the undesirable effects of unintentional dopants.

2. Surfaces of rocksalt oxides MgO and CaO. These surfaces are explored to understand the role of polarity in determining the preference for certain surface terminations. DFT has been coupled to experimental growth conditions via the use of ab-initio thermodynamics (AIT) to predict a stable window where gas-phase surfactants can be used to stabilize otherwise unfavorable habits.

3. Interfaces between a cubic rocksalt oxide and a hexagonal wurtzite nitride. The energetic preferences of atomic configuration at these interfaces are studied using DFT in order to explain the presence of a disordered region near each substrate step-edge.

1.2 Wide-Bandgap Semiconductors and Their Defects

Gallium nitride and aluminum nitride have been proposed as candidate materials for high-power electronics and high-energy optics applications. Alloys between these two nitrides are desirable because of their wide, tunable band gap from 3.5 eV to 6.1 eV. Alloys across the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ composition range also feature low dielectric loss and high carrier mobility. The inability to control the defects precisely in these wide-band-gap semiconductors, however, has proven to be a barrier to their widespread use. The wide electronic band gap makes additional ionization states available, which in turn leads to electronic compensation, unwanted optical absorption, reduced carrier concentration, and reduced carrier mobility. [Collazo et al., 2011] Understanding the behavior of these defects is the first step to enabling an engineering solution that will allow their integration into commercial devices.

Nitride semiconductors must be able to support n- and p-type doping, high carrier mobilities and concentrations, and low carrier activation energies, and they must present desired optical properties such as optical transparency. These properties are governed in large part by the interaction of impurity and native defects via changes in the Fermi level. Point defects, however, are challenging to control because of the difficulty in identifying and measuring them accurately without a well-known reference. There are many experimental techniques which can give a picture of the defect profile in a material, but each has its own set of limitations. Secondary Ion Mass Spectroscopy (SIMS) can identify the impurities in a material, assuming they are above some threshold level, but cannot identify native defects or identify the structural, electronic, or
optical properties of the impurities. Positron Annihilation Spectroscopy can identify negatively charged vacancies and vacancy-impurity complexes but requires theoretical input for reference. Photoluminescence (PL) and cathodoluminescence (CL) provide a detailed signature of the transitions between energy levels in the material but also require a reference input. Electron paramagnetic resonance (EPR) and nuclear magnetic resonance (NMR) can be used to identify some defects if they meet certain criteria such as having an unpaired electron spin. Electronic structure calculations using DFT can provide relative energies of formation of defects, but they are limited by the short length scales that can easily be treated. Each of these techniques can provide invaluable information about the structure and properties of defects, but none alone is sufficient to form a complete understanding of the nature of defects in new materials.

The final challenge to understanding point defects is that the defects are influenced by their environment, meaning that slight differences between the growth techniques, conditions, temperature, and even growth orientation, can lead to significant differences in defect incorporation and behavior. Furthermore, defects interact with each other structurally, through strain, and electronically, through both Coulombic interaction and the law of mass action (for example compensation, which changes the equilibrium Fermi level).

In this dissertation, DFT results using hybrid functionals are coupled with solutions to the mass balance equations to provide a picture of the defects in GaN and AlN. Hybrid functionals, which mix a standard exchange-correlation functional with a fraction of exact Hartree-Fock exchange are used in the calculation of point defects in order to better capture the experimental band gap and enthalpies of formation. These calculation results are closely tied to experimental PL and SIMS measurements, and through this collaboration we have been able to answer several long-standing questions about the influence of defects in III-V Nitrides.

1.2.1 Summary of Findings

Bulk AlN has been known to suffer from a characteristic optical absorption at 4.7 eV. This absorption has limited the use of AlN as a native substrate for AlGaN deep-UV optical devices because the absorption occurs at the edge of the deep-UV spectrum needed for UV sterilization or spectroscopy applications. Presented here are results that demonstrate that carbon is responsible for this absorption and that reducing the amount of carbon in the samples yields a corresponding reduction in the magnitude of the absorption coefficient. AlN that displays the 4.7 eV absorption also displays a 2.8 eV emission whose origin has previously been unknown. Calculations of the optical transitions between a carbon-vacancy DAP, coupled with photoluminescence measurements, indicate that this is the origin of the 2.8 eV emission. Finally, the influence of point defects on the electronic properties of n-type Al$_x$Ga$_{1-x}$N are presented. Experimental observations shows a change in the donor activation energy as a function of Si content.
DFT calculations coupled with solutions to the equilibrium defect concentrations indicate that this change is due to an interaction between shallow donor states and deep acceptor levels of O and Si.

1.3 Thin Films and the Challenges of Polarity

Thin film processing techniques were the enabling technology for silicon-based transistors and the first information revolution, and the field has been evolving ever since to tackle new and challenging materials problems. One active area of exploration is the epitaxial integration of dissimilar materials, which is central to technology such as high-mobility transistors, solid-state lasers and LEDs, semiconductor quantum wells, and superconducting tunnel junctions. [Paisley et al., 2011] When applied to materials with different properties, heteroepitaxy can allow property coupling at the interface which would not be possible with bulk materials or a single thin-film material. [Malin et al., 2007, Murray and Vanderbilt, 2009, Schmehl et al., 2007, Thomas et al., 2010, Zavaliche et al., 2005] The epitaxial integration of functional materials, including oxide ferroelectrics and piezoelectrics as just two examples, depends upon the ability to create smooth interfaces between the active component and the substrate, and the ability to stabilize smooth surfaces.

The creation of smooth, consistent interfaces for novel functional heterostructures introduces an additional set of challenges. The most exciting new properties, such as high-mobility carrier gases, may arise at interfaces between polar materials. The same polar growth orientation that makes these devices interesting also leads to polar surfaces, which are known to be unstable due to the uncompensated electrical dipoles and the buildup of an electrostatic potential as the sample grows thicker. This polarization catastrophe has been shown to determine the structure of (001) LaAlO$_3$/SrTiO$_3$ interfaces, among other materials. [Tasker, 1979, Nakagawa et al., 2006] Recent developments in the growth of polar \{111\} rocksalt oxides MgO, CaO, and Mg$_x$Ca$_{1-x}$O via the use of a hydrogen surfactant have enabled smooth, epitaxial growth of a polar oxide on a polar nitride substrate. This dissertation explains the mechanisms of stabilization via DFT calculations coupled to \textit{ab initio} thermodynamics (AIT)$^1$, which enables the description of surfaces in contact with a gas-phase reservoir to extend the ground-state DFT calculations to experimentally relevant temperatures and pressures.

1.3.1 Summary of Findings

The competition between different polarity compensation mechanisms on CaO and MgO surfaces is explored by extending DFT calculations to experimental growth conditions. The calcu- 

$^1$Also referred to as “Atomistic Thermodynamics”
lations predict the existence of a processing window in which the polar \(\{111\}\) CaO surface could be stabilized by the use of a hydrogen surfactant. This prediction is in good agreement with experimental measurements. The polarity compensation mechanisms of polar \(\{111\}\) surfaces MgO and CaO are compared in terms of their energetics, charge redistribution, and changes to the electronic density of states.

1.4 Interfaces Between Dissimilar Materials

Property coupling between two materials of different structure depends on the ability to reliably create a smooth, atomically flat interface. The performance of this interface depends on the differences in structure, bonding character, and symmetry between the substrate and epitaxial layer. In the case of lattice-matched Mg\(_{x}\)Ca\(_{1-x}\)O (MCO) grown on a GaN substrate, this interfacial structure has been shown to depend on the stepped morphology of the substrate layer. Perfect interfacial registry has been observed by high angle annular dark field scanning transmission electron microscopy (HAADF-STEM) on the substrate terraces but with significant disorder at a substrate step-edge. This observation implies that the oxide layer adopts only one rotational variant with respect to the substrate and that the rotational difference between the substrate above and below a step-edge is responsible for the disordered regions.

1.4.1 Summary of Findings

During the growth of the first oxide monolayer, the presence of hydrogen sets a strong epitaxial preference for a rotational variant that preserves the stacking sequence of the GaN substrate. This preference leads to uniform growth on terraces of the GaN substrate. Nevertheless, due to the presence of growth spirals with \(c/2\) step heights on the substrate, the MCO above a step-edge is rotated \(60^\circ\) with respect to the material below the edge. This difference in rotation leads to a disordered region above the step-edge, which can be seen clearly in HAADF-STEM imaging.

1.5 Overall Summary of Results

This dissertation presents results which show that the framework of DFT calculations closely coupled to experiment can provide valuable insight into the structure and properties of defects in nitride substrates, polar oxide surfaces, and nitride-oxide interfaces.

1. Carbon is identified as the source of the 4.7 eV deep-UV absorption in bulk AlN
2. The 2.8 eV emission that is seen in high-C samples with the 4.7 eV absorption is shown to be due to a donor-acceptor pair formed between the carbon and the nitrogen vacancy.
3. The activation energies of O and Si donors in AlN are determined, the DX transition of these donors from a shallow to a deep level are demonstrated, and the experimentally observed transition between the O donor and Si donor is explained.

4. The experimental growth window that can stabilize a polar \{111\} surface of CaO by the presence of a hydrogen surfactant is predicted.

5. The polarity compensation mechanisms (electronic redistribution, atomic rearrangement, and the adsorption of foreign species) are compared for surfaces of MgO and CaO.

6. The strong preference for a single rotational variant of cubic rocksalt MCO on hexagonal wurtzite GaN is demonstrated as a result of the preference for the first layers of the oxide to follow the bonding template set by the substrate.

The understanding of defects in materials is of vital importance to the ability to reliably create new devices, because while the defects can add functionality that would be otherwise impossible to achieve, they can also render a material useless. Understanding how these defects act to change the structure, electronic and optical properties, and energetics of a material is critical to being able to design better materials. This dissertation aims to provide several specific examples of how iterative feedback loops between experimental and theoretical researchers have enabled quantitative descriptions and predictions of the behavior of defects and the solutions to using those defects to achieve new functionality.
Chapter 2

Density Functional Theory

In the body of this dissertation, the structure and properties of defects are examined from a theoretical perspective and compared to experimental measurements. DFT is an approach to solving the many-electron problem without empirical parameterization. DFT has been used to study metals, insulators, semiconductors, and organic materials. It has been used to describe material properties such as equilibrium structure, elastic properties, and phonons. [Jones and Gunnarsson, 1989, Payne et al., 1992] This technique is also suitable for determining surface reconstructions, surface energies, and configurations at grain boundaries and interfaces.

2.1 The Basics of DFT

This chapter is intended to provide the reader with a brief introduction to DFT and how it can be used to solve problems in Materials Science and Engineering. The structure of the chapter borrows largely from the graduate course developed by Professor Irving with my assistance: “Density Functional Theory in Materials Science and Engineering” Spring 2013. The chapter begins with a discussion of elementary quantum mechanics and the variational principle. For a more thorough background on those topics, Griffiths’ textbook provides an excellent starting point. [Griffiths, 2005] The remainder of the chapter discusses the fundamentals of DFT. For a more detailed look, thorough revies can be found in Density Functional Theory of Atoms and Molecules, The Chemist’s Guide to DFT, and The ABC of DFT. [Parr and Yang, 1989, Koch and Holthausen, 2001, Burke, 2007] Finally, for an overview of the applications and use of DFT, Density Functional Theory: A Practical Introduction is recommended. [Sholl and Steckel, 2009]
2.2 The Schrödinger Equation

The “density” in density functional theory refers to the real space electron density. The properties of atoms, molecules, and materials depend on how electrons interact with the positively charged nuclei in the system, as well as with each other. Understanding the interactions at the scale of electrons and nuclei is the key to understanding the behavior of the system as a whole. It is complicated by the fact that electrons do not behave like the classical particles but also have a wave-like nature. This wave-particle duality means that the position of the electron at a point in time is not as simple as it is for, say, a projectile fired from a cannon. Classical kinematics equations specify the position $x$ of an object at any time $t$ as long as the initial conditions (position, mass, and velocity) and forces $F$ acting on the object are known:

$$x(t) = x_0 + tv_0 + \frac{1}{2} \frac{F}{m} t^2. \quad (2.1)$$

An electron, on the other hand, is described in position and time by a wavefunction $\Psi(x,t)$, which is determined according to Schrödinger’s Equation (in one dimension):

$$i\hbar \frac{\partial \Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x,t)}{\partial x^2} + V\Psi(x,t). \quad (2.2)$$

Where $i = \sqrt{-1}$, $\hbar$ is the reduced Planck constant, $m$ is the electron mass, and $V$ is the potential energy. It is somewhat tricky to develop an intuitive understanding of wavefunctions, but the wavefunction can be related to something a little more common in daily life: a density. The squared amplitude of the wavefunction $|\Psi(x,t)|^2$, which is simply the wave function multiplied by its complex conjugate\(^1\) times a volume $dx$ (in 1-D), also written $|\Psi^*\Psi|dx$, gives us the probability that at time $t$, the electron will be found in the volume $dx$. If we map this probability across real space, we are left with a probability density, or an indication of where we are more or less likely to find the electron at time $t$. It can be conceptually useful to think of this probability density as the density of electrons, or the density of charge, in space.

Although we have a definition of how the wavefunction behaves in Eq. 2.2, we are no closer to understanding how to determine $\Psi$. We can start by making a simplifying assumption by a separation of variables: defining the wavefunction as a product of two functions, one that depends on $t$ and one that depends on $x$:

$$\Psi(x,t) = f(t)\psi(x). \quad (2.3)$$

\(^1\)Let complex number $k = a + bi$, its complex conjugate, $k^* = a - bi$. 
Substituting Eq. 2.3 into Eq. 2.2 and multiplying by \( \frac{1}{f} \) yields

\[
\frac{i\hbar}{f} \frac{df}{dt} = -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V. \tag{2.4}
\]

Notice that the left hand side of the equation depends only upon \( t \) and the right hand side depends only on \( x \). The implication here is that a change to \( t \) with \( x \) held constant cannot change the value of the left hand side of the equation, and *vice versa*. This is a powerful result because it can only be true if both sides are equal to the same constant. Call this constant \( E \). The left side of the equation becomes a differential equation of the form \( i\hbar - Ef = 0 \) which has solutions of the form \( f(t) = \exp \frac{-iEt}{\hbar} \). The right hand side of this equation, the time-independent part, will take more effort to solve. This time-independent Schrödinger equation can be written

\[
\left( -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V \right) \psi = E\psi. \tag{2.5}
\]

Recall that \( V \) is the potential energy, and it turns out that \( -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \) is the kinetic energy. Thus, the term in parentheses represents a Hamiltonian \( H \) of the wavefunction, and implies that \( E \) is the total system energy for the Hamiltonian \( H \) and is an eigenvalue of the Hamiltonian operator. This provides one route to determining the wavefunction. We can define our Hamiltonian by defining the potential energy, \( V \), that corresponds to our system. This gives a differential equation which we can solve, evaluate at our boundary conditions, and finally normalize. In applying this procedure, it quickly becomes apparent that solutions for simple systems such as the Hydrogen atom, with just one electron and a relatively simple potential energy, yield a non-trivial solution. More importantly, there is a fixed set of potentials that can be solved exactly in this way. This brute-force approach cannot be applied to real materials systems.

### 2.3 The Variational Principle

The alternative to brute-force solutions to the time-independent wave equation relies on the variational principle. This principle proves that once we have defined our Hamiltonian \( H \), we can pick any normalized trial wave function \( \Psi_T \) and the resulting energy will be greater than or equal to the ground-state energy \( E_0 \).

\[
\int_{-\infty}^{\infty} \Psi_T^* H \Psi_T \, dx \geq E_0 \tag{2.6}
\]

Furthermore, the only time when our calculated energy \( E \) is exactly equal to the ground state energy \( E_0 \) is when the trial wave function equals the true ground-state wave function, \( \Psi_T = \Psi_0 \).
This is another powerful result because it enables us to make successive guesses at the true wave function, and if our energies are lower, we more closely approximate the true ground-state wave function.

2.4 Many Electrons

As stated previously, it is the interaction of electrons which gives rise to many of the interesting properties of materials. These interactions add a level of complexity to these problems. In the previous section, it was casually mentioned that the solution to the Hydrogen atom with one electron was a solvable problem with an exact solution, though its derivation is not trivial. The helium atom seems not to be much more complicated: it has a nucleus with an extra proton, and it now has two electrons. It turns out, however, that an exact analytical solution to this system remains unknown because of the interactions between the two electrons.

2.5 The Born-Oppenheimer Approximation

Consider the Hamiltonian of a system with \( N_e \) electrons and \( N_\mu \) nuclei:

\[
\mathcal{H}(r_1, r_2, r_3, \ldots, r_{N_e}) = \sum_{i=1}^{N_e} -\nabla_i^2 + \frac{1}{2} \sum_{i,j} 2 \frac{|r_i - r_j|}{|r_i - r_j|} - 2 \sum_{\mu,j} \frac{Z_\mu}{|R_\mu - r_j|}. \tag{2.7}
\]

The solution to the stationary state represented by this Hamiltonian is found by solving the eigenvalue problem

\[
\mathcal{H}(r_1, r_2, r_3, \ldots, r_{N_e}) \psi_j(r_1, r_2, r_3, \ldots, r_{N_e}) = E_j \psi_j(r_1, r_2, r_3, \ldots, r_{N_e}) \tag{2.8}
\]

In this solution, \( E \) depends on the nuclear positions \( E_j = E_j(R_1, R_2, R_3, \ldots, R_{N_\mu}) \). Electrons are approximately \( 1/1800 \) as massive as protons, and thus they can adjust their positions instantaneously compared to the nuclei. Therefore, the positions of the nuclei can be fixed and the state of the electrons can be determined. In a subsequent step, the positions of the nuclei can be updated based on the solutions to the electronic structure, and the procedure can be repeated. This assumption is known as the Born-Oppenheimer approximation and represents a significant step in simplifying the calculation of the many-body wavefunction.

While the system can be solved for the Hydrogen atom, larger values of \( N_e \) (even for \( N_e = 2 \), as in the case of Helium) make an analytical solution intractable. Instead, the variational principle of Eq. 2.6 can be used, written here for our \( N_e \) electron Hamiltonian in \( \mathbb{R}^3 \) and
ensuring normalization:

\[
\int_{-\infty}^{\infty} dr_1 dr_2 dr_3 \cdots dr_{N_e} \Psi_T^*(r_1, r_2, r_3, \ldots, r_{N_e}) \mathcal{H}(r_1, r_2, r_3, \ldots, r_{N_e}) \Psi_T(r_1, r_2, r_3, \ldots, r_{N_e}) \geq E_0
\]

(2.9)

The task at hand now is to use the variational principle to find the ground-state wave function. Notice that our expression for the trial wavefunction \( \Psi_T \) in Eq. 2.9 depends on 3\( N_e \) spatial coordinates.

### 2.6 Slater Determinants

The Pauli exclusion principle requires that we guarantee that this trial wavefunction is antisymmetric under the exchange of any two electrons. To satisfy this requirement, the wavefunction defined by Hartree and Fock was modified by Slater into the single determinental form, called the Slater determinant:

\[
\chi(r_1, r_2, r_3, \ldots, r_{N_e}) = \frac{1}{\sqrt{N_e!}} \begin{vmatrix}
\phi_1(r_1) & \phi_1(r_2) & \phi_1(r_3) & \cdots & \phi_1(r_{N_e}) \\
\phi_2(r_1) & \phi_2(r_2) & \phi_3(r_3) & \cdots & \phi_2(r_{N_e}) \\
\phi_3(r_1) & \phi_3(r_2) & \phi_3(r_3) & \cdots & \phi_3(r_{N_e}) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\phi_{N_e}(r_1) & \phi_{N_e}(r_2) & \phi_{N_e}(r_3) & \cdots & \phi_{N_e}(r_{N_e})
\end{vmatrix}
\]

(2.10)

The Slater determinant naturally gives a wavefunction that is anti-symmetric under exchange of two electrons, which is done by exchanging two columns of the determinant. Furthermore, each element of the determinant is a single-particle eigenfunction, or orbital, which means that we can minimize the energy as a function of the wavefunction \( E(\chi) \) for each single-particle orbital by finding solutions to

\[
\frac{\partial E(\chi)}{\partial \phi_i} = 0 \quad 1 \leq i \leq N_e
\]

(2.11)

This simplifies the set of one-electron equations for the orbital to the form

\[
-\frac{1}{2} \nabla^2 \phi_i(r) + V_{\text{ion}}(r)\phi_i(r) + U(r)\phi_i(r) = \epsilon_i \phi_i(r)
\]

(2.12)
The full equations are given by:

\[
\epsilon_i \phi_i(r) = \left( -\frac{1}{2} \nabla^2 + V_{\text{ion}}(r) + \sum_{j=1}^{N_e} \int dr' \phi_j^*(r') \frac{1}{|r - r'|} \phi_j(r') \right) \phi_i(r) \\
- \sum_{j=1}^{N_e} \delta_{m_i m_j} \left[ \int dr' \phi_j^*(r') \frac{1}{|r - r'|} \phi_i(r') \right] \phi_j(r_i) \quad (2.13)
\]

Note that in the first summation over all electrons \(j\), the Coulombic repulsion between electron \(i\) and \(j\) is computed, but in this case it is computed even when \(i = j\). Thus, this Coulombic term includes self-interaction. The second summation term over \(j\) is a quantum-mechanical expression for the exchange energy, which has a finite number for like spins and is zero for unlike spins. This exact exchange completely eliminates self-interaction in the final solution.

The important results of Hartree-Fock theory are that the solutions \(\epsilon_i\) represent the ionization energies of the electrons (electron removal energies), according to the frozen orbital approximation and Koopmans’ theorem. Second, the solutions to the many-electron problem using Hartree-Fock methods include the exact exchange of electrons, but do not include any electron correlation. Some attempts have been made to add correlation to the Hartree-Fock approach, such as Configuration Interaction, but this approach is prohibitively expensive. Solutions to the Hartree-Fock approach alone are computationally expensive because the solutions require treating \(3N_e\) spatial coordinates and therefore Hartree-Fock calculations are computationally expensive for even some moderate-sized systems.

### 2.7 Thomas-Fermi and the electron density

Recall that the wavefunction, a complex function that is defined by the Schrödinger equation, can, somewhat easily, give us information about the probability of finding the electron at a given place and time. This probability density can be mapped over \(\mathbb{R}^3\) to give a real-space probability density, which describes the electrons in the system using only 3 dimensions, independent of the number of particles. The Thomas-Fermi approach to solving the many-body problem was to treat the electron-electron interactions and electron-nucleus interactions classically, and to treat the kinetic energy piece using quantum statistics. Critically for the development of DFT, their classical treatment of the Coulombic interactions used the real-space electron density \(n(r)\) instead of the electronic wavefunctions. Their kinetic energy term \(T_{TF}\) is given by:

\[
T_{TF}[n(r)] = \frac{3}{10} \left(3\pi^2\right)^{2/3} \int n(r)^{5/3} \, dr \quad (2.14)
\]
Combined with the classical electrodynamics terms for the Coulombic interaction, their expression for the total energy was:

$$E[n(r)] = T_{TF} - Z \int \frac{n(r)}{r} \, dr + \frac{1}{2} \int \int \frac{n(r_1)n(r_2)}{|r_1 - r_2|} \, dr_1 \, dr_2$$  \hfill (2.15)$$

The Thomas-Fermi model is a coarse approximation to the true solution because it neglects exchange and correlation completely. This limitation leads to a catastrophic failure of the Thomas-Fermi approach in that it predicts that no materials or molecules are stable compared to their isolated atomic states. Despite this failure, the approach represented the first attempt at expressing the energy as a functional of the electron density.

### 2.8 Functionals

We mentioned in section Section 2.7 that the Thomas-Fermi model was the first attempt to describe the energy of a system using the electron density. We also used the word “functional” to describe this relationship. The notation used indicates that the energy $E$ is a functional of the density $n$,

$$E[n(r)]$$  \hfill (2.16)$$

A functional is often simply described as a “function of a function”. More specifically, if a function maps a number to another number (for example the function $f(x) = x^2 + 1$ maps a value of $x$ to another number, one we are used to calling $y$) then a functional maps a function to a number. [Burke, 2007] Take the definition of the definite integral:

$$F[f(x)] = \int_{-1}^{1} f(x) \, dx$$  \hfill (2.17)$$

This functional $F$ maps the function $f(x)$ to a single number. In the example given for $f(x) = x^2 + 1$ above, the single value of the functional $F$ is $8/3$. Thus, the functional expression of Thomas and Fermi, despite its several limitations, mapped the real-space electron density to a single value: the total energy of the system $E$.

### 2.9 Hohenberg, Kohn, and Sham

At this point we have defined the “density” and have an understanding of what the “functional” is. This section will discuss in more detail the Theory behind DFT, which was initially developed in two theorems by Hohenberg & Kohn [Hohenberg and Kohn, 1964] and then by Kohn & Sham [Kohn and Sham, 1965]. These major contributions are discussed in detail in the references cited
at the beginning of this chapter, and an effort has been made to summarize those discussions over the next three sections.

2.10 Hohenberg & Kohn

There are two central theorems laid out by Hohenberg and Kohn. The first theorem is a proof of existence. They stated that the external potential energy $V_{\text{ext}}(r)$ uniquely defines an electron density $n(r)$ because $V_{\text{ext}}(r)$ defines the Hamiltonian. They then went further to state that the many-particle ground state is a unique functional of $n(r)$. They proved this by *reductio ad absurdum* by first assuming that two different wavefunctions could yield the same electron density, but by using the two test wave functions as trial wave functions for the other’s Hamiltonian, showed that these wavefunctions can only both give the ground state electron density if they are the same. In other words, the ground-state electron density uniquely defines the external potential $V_{\text{ext}}$. Hohenberg & Kohn also showed that the energy of the system with ground state electron density $n_0$ can be written as

$$ E[n_0] = T[n_0] + E_{\text{e-e}}[n_0] + E_{\text{N-e}}[n_0] \tag{2.18} $$

In equation Eq. 2.18, $T$ is the ground-state kinetic energy, $E_{\text{e-e}}$ is the full (classical and quantum-mechanical) electron-electron interaction, and $E_{\text{N-e}}$ is the electron-nucleus interaction. The expression can be separated into terms that are system dependent, that is they depend on the positions and type of nuclei, and a universal component that is always the same for a given density.

$$ E[n_0] = \int n_0(r)V_{\text{N-e}} \, dr + T[n_0] + E_{\text{e-e}}[n_0] \tag{2.19} $$

This universal piece, which is the same given a density $n_0$ for any system, is the so-called Hohenberg & Kohn functional $F_{\text{HK}}[n_0]$.

$$ F_{\text{HK}}[n_0] = T[n_0] + E_{\text{e-e}}[n_0] \tag{2.20} $$

If the electron density is known, then this HK functional gives the exact energy of the system. Hohenberg and Kohn proved that it existed, but its exact form is still unknown. The functional includes only the kinetic energy of the electrons $T$ and the electron-electron interactions, but unfortunately neither is known exactly. The electron-electron interaction can be broken apart into classical and non-classical pieces:

$$ E_{\text{e-e}}[n_0] = \frac{1}{2} \int \int \frac{n(r_1)n(r_2)}{|r_1 - r_2|} \, dr_1 \, dr_2 + E_{\text{ncl}} \tag{2.21} $$
The integral term on the left is the classical interaction between each pair of particles but includes a particle’s interaction with itself. The term on the right, \( E_{\text{nel}} \) is the non-classical contribution, and includes all of the electron interactions that are not captured in the Coulombic piece (such as exchange and correlation) as well as the self-interaction correction.

The second theorem of Hohenberg and Kohn showed that the application of the variational principle to their energy expression could be used to solve the ground state of the system. They attempted to show that any trial density \( \tilde{n}(\mathbf{r}) \geq 0 \) such that \( \int \tilde{n}(\mathbf{r}) \, d\mathbf{r} = N_e \) satisfied the variational principle:

\[
E_0 \leq E[\tilde{n}(\mathbf{r})] = T[\tilde{n}(\mathbf{r})] + E_{\text{e-e}}[\tilde{n}(\mathbf{r})] + E_{\text{N-e}}[\tilde{n}(\mathbf{r})] \tag{2.22}
\]

### 2.11 Kohn & Sham

Kohn and Sham sought a simpler approach to finding \( F_{\text{HK}} \) and approached the problem by attempting to define a system of interacting electrons in terms of a system of non-interacting electrons that would yield the correct ground-state density. The approach begins by using the Slater determinant expression for the kinetic energy \( T_s \), which is the exact kinetic energy for a system of non-interacting electrons, each having a single-electron wavefunction \( \phi_i \).

\[
T_s [n(\mathbf{r})] = -\frac{1}{2} \sum_{i=1}^{N_e} \langle \phi_i | \nabla^2 | \phi_i \rangle \tag{2.23}
\]

Recall the original version of the Hohenberg and Kohn functional \( F_{\text{HK}} \) has two pieces, the kinetic energy and the electron-electron interactions.

\[
F_{\text{HK}} [n] = T [n] + E_{\text{e-e}} [n] \tag{2.24}
\]

If we define a new quantity that represents the difference between the full, exact kinetic energy \( T \) and the non-interacting kinetic energy \( T_s \), this new quantity \( T_c \) encapsulates the correlated motion of electrons.

\[
T_c [n] = T [n] - T_s [n] \tag{2.25}
\]

If we define another quantity which is the difference between the full electron-electron interaction \( E_{\text{e-e}} \) and the classical part of the interaction including self-interaction \( J \), we obtain the non-classical electron-electron interactions \( E_{\text{nel}} \) which includes the self-interaction correction, exchange, and correlation.

\[
E_{\text{nel}} [n] = E_{\text{e-e}} [n] - J [n] \tag{2.26}
\]

15
If we substitute these terms into equation Eq. 2.24 and separate the classical and quantum mechanical pieces, we arrive at the expression:

\[
F_{HK}[n] = T_C[n] + T_S[n] + E_{\text{incl}}[n] + J[n]
\]

(2.27)
\[
= T_S[n] + J[n] + T_C[n] + E_{\text{incl}}[n] + E_{XC}[n]
\]

(2.28)

Here, the non-classical terms are grouped together in to \( E_{XC} \), which collects all of the terms that are not known. It can be shown that the single-particle effective potentials \( V_S \) for the Kohn-Sham orbitals \( \phi_i \) can be found by

\[
V_S(r) = \int \frac{n(r_j)}{|r_i - r_j|} \, dr_j + V_{XC}(r) - \sum_A^M Z_A r_i A
\]

(2.29)

The potential \( V_{XC} \) is the potential associated with the exchange and correlation functional \( E_{XC} \), and is its functional derivative.

\[
V_{XC} = \frac{\delta E_{XC}}{\delta n}
\]

(2.30)

2.12 Exchange-Correlation Functionals

The result of the Kohn-Sham equations is that if the expression for the exchange and correlation potential \( V_{XC} \) were known, the system could be solved exactly.

2.12.1 The Local Density Approximation

In practice, there is only one system for which the true exchange and correlation potential can be determined, namely the uniform electron gas. One simple approach to defining the unknown \( V_{XC}(r) \) for our real system is to assume that the value of the potential at \( r \) is equal to what the potential would be in a uniform electron gas that has the same density as the density at \( r \).

\[
V_{XC}^{LDA} = V_{XC}^{\text{electron gas}}[n(r)]
\]

(2.31)

Because we are using the local electron density at \( r \) to obtain the exchange-correlation potential, this is known as the Local Density Approximation (LDA). In many cases this approximation works quite well, in part due to a fortuitous cancellation of errors and in part because the valence electron density in most materials varies slowly. It should be noted that LDA calculations tend to be over-binding, that is, they overestimate the cohesive energy. This is, in part, because the LDA is poor at estimating the energy of the unbound state – where the density tends to vary...
more quickly.

2.12.2 Generalized Gradient Approximation

If there are some cases when only looking at the local value of the electron density is insufficient because that density varies quickly, such as the isolated atom, then including the rate of change of the density in the approximation to the exchange and correlation functional may be a good starting point. The Generalized Gradient Approximation (GGA) uses the value of the local density, as well as how the density varies in space (the gradient) in order to determine the exchange-correlation potential. The GGA is an example of a semi-local functional, compared to the fully local LDA.

\[
V_{\text{XC}}^{\text{GGA}} = V_{\text{XC}}[n(r), \nabla n(r)]
\]  

There are two main approaches to how this gradient is included. The first was developed by Perdew and Wang, [Perdew and Wang, 1992] and the second by Perdew, Burke, and Ernzerhof. [Perdew et al., 1996a, Perdew et al., 1997] These are commonly referred to in the literature as the PW91 and PBE functionals. By including the gradient term, PBE is able to better describe lattice parameters of materials compared to LDA but, at the same time, is poorer at describing bulk moduli and the surface energies of transition metals and oxides. Additionally, while LDA is overbinding, GGA can be either under- or over-binding. The decision of whether to use LDA or GGA depends on the specific materials problem at hand and should be made on a case-by-case basis.

2.12.3 Post-GGA functionals

Several attempts to improve GGA have been made, and each is suited to a specific set of problem. The PBE functional does not describe the surface energies of transition metals and oxides as well as LDA, and it also is poorer at describing the jellium surface. Attempts to improve the performance of GGA for these real materials systems have focused on either gradient expansion of the exchange energy and fitting to the exchange-correlation of the jellium surface[Perdew et al., 2008, Perdew et al., 2009], or fitting the jellium surface alone[Armiento and Mattsson, 2005]. Despite the differences in their approaches, both functionals (PBEsol and AM05) show strong agreement in describing lattice parameters and surface energies for a number of solid state systems. Because of the good agreement with experiment of surface energies and lattice parameters, PBEsol is chosen in this dissertation to describe the surfaces of MgO, CaO, and MgCaO.

\footnote{Jellium is simply another name for the uniform electron gas.}
2.12.4 Meta-GGA functionals

Meta-GGA functionals represent another type of non-empirical approximation to the true exchange-correlation functional. In these functionals, the idea of using a gradient of the local density is extended by also considering the laplacian $\nabla^2 n(r)$ which contains the same information as the kinetic energy $\tau(r)$ of the Kohn-Sham orbitals. The functional of Tao, Perdew, Staroverov, and Scusiera (TPSS) uses this formalism, but it should be noted that like LDA and GGA, meta-GGA does not use any empirical fitting or parameterization.

2.12.5 Hyper-GGA functionals

Efforts in creating and modifying exchange-correlation functionals are motivated by the desire to better capture the true exchange-correlation potential of a material which is unknown. Recall, however, from Section 2.6 that the expression for the exchange given in Hartee-Fock theory is the exact exchange of the system. This exchange can be written

$$E_{\text{exchange}}(r) = -\frac{1}{2n(r)} \int d^3r' \left| \sum_{\text{occ. states}} \phi_i^*(r')\phi_i(r) \right|$$

Notice, however, that in order to find this exact exchange at a point $r$, the wavefunctions everywhere else in the system (for all $r'$) must be known. Compare this to the case of LDA or GGA, where only the density at point $r$ must be known. Functionals like these are often called “non-local” for this reason. This non-local nature increases the computational expense of calculations that use the exact exchange term.

Typically the exact portion of the exchange is used in combination with standard local exchange and correlation. For example, the B3LYP functional combines LDA exchange and correlation, exact exchange, and GGA exchange and correlation by using empirical parameters $\alpha_1...3$.

$$V_{\text{XC}}^{\text{B3LYP}} = V_{\text{XC}}^{\text{LDA}} + \alpha_1 (E_{\text{exch.}} - V_X^{\text{LDA}}) + \alpha_2 (V_X^{\text{GGA}} - V_X^{\text{LDA}}) + \alpha_3 (V_C^{\text{GGA}} - V_C^{\text{LDA}})$$

The inclusion of empirical parameters, as in B3LYP [Kim and Jordan, 1994, Stephens et al., 1994], makes a functional empirical, while the inclusion of exact exchange makes the equation non-local. A functional that includes exact exchange is often referred to as “Hyper-GGA” or, because of the use of local “DFT” exchange and correlation combined with non-local “Hartree-Fock” exchange, a “Hybrid” functional. Another empirical, non-local, hybrid functional is PBE0 [Perdew et al., 1996b], whose exchange and correlation potential is given by the full correlation of PBE $V_C^{\text{PBE}}$, and a blend of the PBE exchange $V_X^{\text{PBE}}$ with the exact exchange according to a
parameter $\alpha$:
\[
V_{XC}^{\text{PBE0}} = \alpha E^{\text{exch.}} + (1 - \alpha) V_X^{\text{PBE}} + V_C^{\text{PBE}} \tag{2.35}
\]

2.12.6 Semi-local functionals: HSE

One of several reasons for choosing to use one of the more computational expensive hybrid functionals discussed in section Section 2.12.5 is to overcome one of the principal limitations of DFT: an underestimation of the electronic band gap. DFT in its usual form does not accurately reproduce the position of unoccupied electronic states, placing them too close to the occupied states. In most metals, this is not a severe limitation because the unoccupied states are already near the occupied states. [Pickett, 1989] In narrow-gap semiconductors, however, these unoccupied states may be calculated by DFT to be very close to the occupied states, leading to artificial metallic behavior. In semiconductors and insulators with wider energy gaps, this limitation may manifest as a narrower band gap than experimental measurements indicate. Often, this is not problematic, and DFT is able to reasonably reproduce experimental structural properties and surface energies. This limitation becomes more restrictive when attempting to calculate the transitions between different charge states of defects in materials, because an inaccurate representation of the band gap makes it ambiguous whether a given transition corresponds to a deep or shallow state. Furthermore, defect levels which are close to the band edge may artificially interact with the conduction band states.

Hybrid functionals have been shown to reproduce the band gaps of solids much more accurately than standard DFT functionals. [Muscat et al., 2001] Still, this leaves the problem of extensive computational cost. Recently, Heyd, Scuseria, and Ernzerhof proposed a scheme by which the exact exchange is incorporated only in the short range. [Heyd et al., 2003, Heyd et al., 2006] The expression for the exchange and correlation potential has clear similarities to the expression for PBE0 (Eq. 2.35), only differing in that the entire long-range exchange is taken from PBE, while only the short-range exchange is mixed between PBE and the exact Hartree-Fock exchange:

\[
V_{XC}^{\text{HSE}} = \alpha E^{\text{exch.,SR}} + (1 - \alpha) V_X^{\text{PBE,SR}} + V_X^{\text{PBE,LR}} + V_C^{\text{PBE}} \tag{2.36}
\]

This semi-local hybrid technique has the advantage that the calculations are significantly less expensive than their non-local equivalents. The HSE functional has been tested for a number of semiconductors and insulators and shows considerably better agreement with experimental lattice parameters and band gaps compared to traditional DFT. [Heyd et al., 2005, Paier et al., 2006] The standard implementation of HSE uses $\alpha = 0.25$, based on the $\alpha$ of PBE0 which was chosen based on matching perturbation theory calculations. [Perdew et al., 1996b] Although the semi-local HSE is less computationally demanding than PBE0, it nevertheless is approximately
ten times more expensive than a pure GGA calculation. Because of the ability of HSE to accurately represent the electronic band gap and ionization energies of defects in semiconductors, it is used in this dissertation for calculations on point defects in GaN and AlN. [Lyons et al., 2010, Van de Walle et al., 2010, Janotti et al., 2010]

2.13 Pseudopotentials

Once the desired approximation to the exchange and correlation functional has been chosen, solutions to the ground state of the system requires only solving a Kohn-Sham equation for each electron (orbital) in the system. This number can quickly increase, for example in the 96-atom GaN systems discussed in this dissertation, there are a total of 1824 electrons. In most materials, only the valence electrons participate in bonding, while the core electrons are essentially fixed. Using this “frozen core” approximation, the number of valence electrons in the 96-atom cell can be reduced to 384 electrons. DFT codes often make use of this frozen core approximation through the use of “pseudopotentials” which describe the potential of the core of the ion, including the non-bonding core electrons. There are multiple ways of generating this pseudopotential, two common variants are the ultra-soft pseudopotentials (USPP) and the projector-augmented wave (PAW) pseudopotentials. Note that these pseudopotentials are fitted only to match the behavior of the atom with all of the electron (the all-electron potential), and not to any empirical values. In many cases the all-electron, USPP, and PAW potentials all give similar results, though PAW tends to outperform USPP in materials with strong magnetic moments or when there are large differences in the electronegativity of atoms in the material. In this dissertation, all calculations have been performed with PAW pseudopotentials. Details on which electrons are included in the core and which are included in valence is described in the appropriate sections.

2.14 Bloch’s Theorem

Bloch’s theorem states that the solution to any periodic problem can be expressed as a sum of functions of the type

\[ \phi_k(r) = \exp \left[ i (k \cdot r) \right] u_k(r) \]  

where \( u_k \) is a function that is periodic over the unit cell, and \( \exp \left[ i (k \cdot r) \right] \) describes a plane wave, which is defined in terms of its position in real space \( r \) and its position in reciprocal space \( k \). This theorem enables us to describe the solutions to our Kohn-Sham equations in terms of this set of plane waves. A plane-wave expansion has several advantages for solving problems in solid state materials. First, the plane waves, like other trigonometric functions,
can be used in Fourier expansion to approximate any other function. This means that a given set of planewaves can be used to approximate our orbital wavefunctions. Second, crystalline solids of the sort that this dissertation concerns are periodic, which lets us use symmetry to simplify the calculations. Plane waves form an orthogonal basis set, which makes the calculation of integrals simpler. Plane waves do have a few limitations, namely that everything must be periodic and that extracting localized information can be challenging. In later sections methods to overcome these limitations will be discussed. In principle an infinite sum of Bloch functions would be desired, but for practical reasons this sum is truncated. How the sum is truncated varies depending on the specifics of the calculation, but it is generally chosen such that some numerical convergence is reached.
Chapter 3

DFT in Practice

The basics behind Density Functional Theory (DFT) were introduced in Chapter 2. This chapter will focus on the types of information that can be obtained from a DFT calculation. Specific details on how the calculations discussed in this dissertation were performed will be presented in Chapter 5 and Chapter 6. The first practical step in performing a DFT calculation is selecting the code to be used. All of the DFT calculations presented in this dissertation were performed using the Vienna *ab initio* Simulation Program (VASP) [Kresse and Hafner, 1993, Kresse and Hafner, 1994, Kresse and Furthmuller, 1996b, Kresse and Furthmuller, 1996a], and therefore the VASP input and output standards are used when specifics are required.

3.1 The anatomy of a DFT calculation

Once the desired exchange-correlation functional (see Section 2.12) and pseudopotentials (see Section 2.13) to be used have been decided upon, performing a DFT calculation begins with defining an initial configuration of atoms. These initial positions are often based on experimental or previous theoretical evidence for the structure of a solid. The solution to the ground-state energy of this configuration of atoms is then iteratively solved using the variational principle. The electron density $n(r)$ is found by first solving the Kohn-Sham equations to obtain the single-particle wavefunctions $\psi_j$:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi_j(r) + V_{\text{eff}}(r) \psi_j(r) = \epsilon_j \psi_j(r)$$

(3.1)

These solutions $\psi_j(r)$ are related to the density by:

$$n(r) = \sum_j \psi_j(r) \psi_j^*(r).$$

(3.2)
The system is non-trivial to solve because the effective potential $V_{\text{eff}}$ depends upon the electron density. The approach is to solve the system self-consistently, by making a first guess for the electron density, then using this trial density to construct $V_{\text{eff}}$, then to solve for the wavefunctions $\psi_j$ and finally to determine the new density. If this density is not equal to the trial density, we have not yet solved the problem. In that case, the updated density is used to generate a new effective potential $V_{\text{eff}}$ and the procedure is updated until the density obtained from the solution to the Kohn-Sham equations is equal to the trial density at that step, within some practical numerical cutoff, or equivalently, when the difference in energy between steps approaches zero. At this point, the ground-state electron density and the ground-state Kohn-Sham orbitals for the configuration of atoms we defined has been found. This solution also gives us the ground state energy of the system.

### 3.2 Atomic Relaxation

In the previous section, the ground state energy for a defined atom configuration was solved iteratively using DFT. In many cases, such as surface relaxations and the relaxation of point defects, we are interested in finding a ground state configuration of those atoms based on their electronic and atomic structure. DFT can solve for this equilibrium configuration by minimizing the total system energy (or total forces) as a function of the atomic positions. In this procedure, an initial atomic configuration is chosen and the electronic ground state is solved as in Section 3.1. This ground state can then be used to determine the forces Hellman-Feynman forces acting on the ions due to the other ions and the electrons in the system. These forces are used to update the positions of the atoms based on a chosen minimization scheme. The ground state of the system is now calculated based on these new positions. This procedure is repeated until the system reaches specified convergence criterie, as measured by either convergence of the energy differences between steps or convergence of the forces on atoms.

### 3.3 Electronic Structure

Once the ground state of the system has been determined using DFT, the charge density $n(\mathbf{r})$ and the wavefunctions $\psi_j$ are known. These solutions give information about the electronic ground state, such as the density of electronic states and the electronic band structure, which can be used to give physical insight about the properties of the material. The capabilities and limitations of DFT to provide that information is presented here, while further details about how this information has been used for the specific materials problems in this dissertation are given in later sections.
3.3.1 Density of Electronic States

The density of available electronic states provides information about how the electrons in a system are distributed as a function of the energy. In addition to giving information about whether a material is an insulator or metal and the magnetic character of the material, the density of states can be a powerful tool to evaluate the changes in the electronic structure caused by defects. The density of states \( n(E) \) is essentially the number of electronic states in the energy interval between \( E \) and \( E + dE \), at each energy \( E \). Determining this quantity then depends on determining the position of the electronic states at each energy, which is equivalent to determining the energies of each of the orbital wavefunctions. In section Section 2.14, Bloch’s theorem was used to describe the wavefunctions as a sum of Bloch functions of the form

\[
\phi_k(r) = \exp \left[ i \left( k \cdot r \right) \right] u_k(r).
\]

(3.3)

The energy of electrons described by wavefunctions of this form is

\[
E = \frac{(\hbar k)^2}{2m}
\]

(3.4)

These energies \( E \) are determined based on integrations in reciprocal (\( k \)) space, and so it should be noted that in practice, accurate density of states calculations require a larger basis set than would be required to converge structural parameters or total energies.

3.3.2 Band Structure

The density of states (DOS) is essentially calculated “for free” in a DFT calculation, with the critical caveat that a more complete basis set is needed to converge this density of states. A different picture of the electron distribution in energy space can be obtained by the electronic band structure. The band structure describes the position of the energy levels in Eq. 3.4 at specific \( k \)-points. These special high-symmetry \( k \)-points and the paths between them vary depending on the structure of a material, but provide a picture of the electron energies in \( k \)-space, conveniently projected onto a one-dimensional path. These band diagrams are useful for determining the direct and indirect electronic band gaps, and can be used to determine electron and hole mobilities. Details of how to set up and perform these band structure calculations are presented in Chapter 5, and an excellent reference to paths of interest can be found in the work of Setyawan and Curtarolo. [Setyawan and Curtarolo, 2010]
3.3.3 Charge Density

The iterative solution to the Kohn-Sham equations gives the ground-state orbital wavefunctions and the ground-state electron density. The charge density represents the distribution of electrons in the system and can be used to evaluate the type of bonding in the system. The charge density can also be used to estimate the boundaries of atoms in a solid through the analysis proposed by Bader, which defines the atomic boundaries as the zero flux surfaces of the charge density. [Bader, 1985] This method is useful because it provides the possibility of defining the boundaries of an atom even when covalent bonding is dominant. These atomic volumes can be used to evaluate the density of electronic states associated with specific atoms (such as a defect) in the material. The Bader approach has been implemented for use in plane-wave codes by Henkelman [Henkelman et al., 2006, Sanville et al., 2007, Tang et al., 2009] and Trinkle [Yu and Trinkle, 2011]. Specific uses of this approach in this dissertation are given in Chapter 6.

3.3.4 Electrostatic potential

In the course of solving the Kohn-Sham equations, the density is used to determine the effective potential \( V_{eff} \), which is a sum of the contributions to the potential from the external potential \( v_{ext} \) which incorporates electron-nuclear interactions, the Hartree potential \( v_H \) which includes electron-electron interactions with self-interaction, and the exchange-correlation potential \( V_{XC} \), which is the functional derivative of the exchange-correlation functional.

\[
V_{eff} = v_{ext} + v_H + \frac{\delta E_{XC} [n(r)]}{\delta n(r)}
\]  

(3.5)

In general, DFT reports the energies of the Kohn-Sham orbitals without a common reference. The lack of a common reference makes direct comparison between the energy levels reported in two calculations problematic, but a comparison of the local potential allows these energies to be referenced to a common state. The local potential is also useful in defining the work function of the material; in this case it is preferable only to use the electrostatic potential, that is, only the external potential and the hartree potential, because the exchange-correlation potential converges more slowly in the vacuum. VASP gives the option of reporting the local potential as the electrostatic potential (the sum of the ionic and Hartree potential), or the total potential (the sum of the ionic, Hartree, and exchange-correlation potential). The electrostatic potential converges more rapidly to the vacuum level compared to the total potential. Details on how the local potential is used in defect calculations appear in Chapter 6.
Chapter 4

Thermodynamics

The structure of defects, and the influences of this structure on the properties of the material can often be described by the energetic changes between a perfect system and a defected system. These energy changes from processes such as the incorporation of a substitutional impurity defect, creation of a free surface from a bulk material, or chemical adsorption of a foreign species on the surface can readily be obtained by running DFT calculations of the initial and final systems. This chapter will cover some of the basics of thermodynamics as they apply to evaluating energy differences in DFT.

4.1 Reaction Energetics

The difference in total internal energy $U$ between two states of a closed system is given by

$$\Delta U = U_{\text{final}} - U_{\text{initial}}.$$  \hspace{1cm} (4.1)

Combining the first and second laws of thermodynamics enables the equation to be rewritten

$$\Delta U = T\Delta S - P\Delta V,$$  \hspace{1cm} (4.2)

where $T$ is temperature, $\Delta S$ is the change in entropy, $P$ is pressure, and $\Delta V$ is the change in volume. It is useful to define three additional terms, the Helmholtz free energy\footnote{The Helmholtz free energy is often referred to by the variable $A$, but $F$ is used here to avoid confusion between the free energy and the surface area, which will be denoted $A$ in this dissertation.} $F$, the Gibbs
free energy $G$, and the enthalpy $H$:

$$F = U - TS$$  \hspace{1cm} (4.3)  
$$G = U + PV - TS$$  \hspace{1cm} (4.4)  
$$H = U + PV$$  \hspace{1cm} (4.5)  

The Gibbs free energy $G$ of Eq. 4.4 can be combined with the expression for enthalpy $H$ in Eq. 4.5 to give a useful expression of the system energy

$$G = H - TS,$$  \hspace{1cm} (4.6)  

and thus the difference in the Gibbs free energy between two states of a closed system under constant temperature is given by

$$\Delta G = \Delta H - T \Delta S.$$  \hspace{1cm} (4.7)  

If we consider an open system under constant temperature, one which can exchange species with a reservoir, the difference in the Gibbs free energy is

$$\Delta G = \Delta H - T \Delta S + \sum_{i} \mu_i \Delta n_i,$$  \hspace{1cm} (4.8)  

Where $\mu_i$, the chemical potential of species $i$, is the partial molar Gibbs free energy, or the partial derivative of the Gibbs free energy with respect to the change in the number of particles (moles) of the species $n_i$:

$$\mu_i = \left( \frac{\partial G}{\partial n_i} \right)_{T,p,n_j,...}.$$  \hspace{1cm} (4.9)  

The chemical potential can be thought of as the change in (Gibbs) energy of the system with the addition of another particle (mole) of the species.

### 4.2 Enthalpy of Formation

In DFT calculations it is often useful to compare the calculated enthalpy of formation $\Delta H^f$ to some known experimental quantity. The enthalpy of formation describes the change in energy associated with the reaction of species in their chemical reservoirs (some chosen reference state) to being associated in a compound or bulk material. As an example we can consider the formation of magnesium oxide (MgO) from bulk magnesium ($\text{Mg}_{\text{bulk}}$) and oxygen gas ($\text{O}_2$). The energy difference $\Delta H$ is given by subtracting the energies of the initial configuration from
the final configuration:

\[
\Delta H_{\text{total}} = H_{\text{final}}(T, \{p_i\}, n_i) - H_{\text{initial}}(T, \{p_i\}, n_i),
\]

(4.10)

where \(p_i\) is the partial pressure of species \(i\). By stating that the initial state of our system consists of the elements in their reference states, which are not interacting, the equation can be written in terms of the separated enthalpies:

\[
\Delta H_{\text{total}} = H_{\text{MgO}}^{\text{final}}(T, \{p_i\}, n_i) - n_{\text{Mg}} H_{\text{Mg}}(T, \{p_i\}, n_i) - n_{\text{O}} H_{\text{O}}(T, \{p_i\}, n_i),
\]

(4.11)

where \(n_{\text{O}}\) and \(n_{\text{Mg}}\) are the number of each species in the final calculated compound. This equation describes the change in energy of bringing \(n_i\) of each species \(i\) from its reservoir (where it has initial reference energy \(H_i\)) to the final compound. We are interested in determining this enthalpy of formation from a DFT calculation, which gives energies at 0 K. The enthalpy was given as:

\[
H = U + PV,
\]

and so the change in enthalpy under constant pressure can be written

\[
\Delta H = \Delta U + P\Delta V,
\]

(4.12)

The Helmholtz energy \(F\), can be written in terms of \(U\) as

\[
F = U - TS,
\]

(4.13)

and at 0K, the internal energy \(U\) is simply equivalent to the Helmholtz energy \(F\). The \(P\Delta V\) term is small compared to the Helmholtz energy and so can be ignored. The Helmholtz energy \(F\) can be written as the sum of the energy at 0 K, plus the vibrational contributions:

\[
F = E_{\text{total}} + F_{\text{vib}},
\]

(4.14)

where the vibrational contributions have a 0 K part \(E_{\text{vib}}\) and a temperature-dependent entropic part \(-TS_{\text{vib}}\). In most cases, the zero-point energy can be ignored for differences between bulk systems. [Reuter and Scheffler, 2001] At this point, the enthalpies of our initial and final systems have been reduced to the Helmholtz energy at 0 K, which is directly obtained from the DFT calculations. Our enthalpy of formation can be written:

\[
\Delta H_{\text{total}} = E_{\text{MgO}} - n_{\text{Mg}} E_{\text{Mg}} - n_{\text{O}} E_{\text{O}},
\]

(4.15)
where $E$ will denote an energy obtained from DFT. These per-atom reference energies $E_{\text{Mg}}$ and $E_{\text{O}}$ are obtained from DFT calculations of bulk Mg and molecular $\text{O}_2$ and it is worth noting that they are equivalent to the chemical potential $\mu_i$ of the species in the reference state, which will be denoted $\mu_i^\circ$. Thus, the enthalpy of formation can finally be written:

$$\Delta H^{\text{total}} = E_{\text{MgO}} - n_{\text{Mg}}\mu_{\text{Mg}}^\circ - n_{\text{O}}\mu_{\text{O}}^\circ. \quad (4.16)$$

A result that we will make further use of later is that when the species are in equilibrium with the bulk, their chemical potentials (different from their reference chemical potentials) are limited by the energy of the bulk, by:

$$g_{\text{MgO}} = n_{\text{Mg}}\mu_{\text{Mg}} + n_{\text{O}}\mu_{\text{O}}. \quad (4.17)$$

### 4.3 Formation Energies of Surfaces

When using DFT to understand the formation of surfaces of a material, the energies reported for the surface configuration are compared to a reference state. It is convenient that this reference state be the bulk material, without defects. The total energy of this reference state is easy to compute in DFT, since it is periodic. Computing the total energy of the surface slabs has certain inherent complexity that is discussed further in Chapter 5. Once the total energies of the surface system and reference system have been computed, however, the task remains to determine the surface energy. As in the case of the formation enthalpy of a solid, the formation energy of the surface will take the form of an energy difference between the final state (surface) and the initial state (bulk compound). In the case of surfaces, there is interest in determining how the energy of surfaces changes as a function of the growth conditions, namely temperature and pressure. Because the effects of temperature and pressure are important, the Gibbs free energy is the natural auxiliary function of choice.

$$\Delta G = -S\Delta T + V\Delta P + \sum_i \mu_i \Delta n_i \quad (4.18)$$

In the case of the formation energy of a surface, the difference in energy can be expressed by:

$$\Delta G^{\text{total}} = G_{\text{final}}(T, \{p_i\}, n_i) - G_{\text{initial}}(T, \{p_i\}, n_i), \quad (4.19)$$

where $p_i$ is the partial pressure of species $i$. By stating that the initial state, the reference state for this comparison, is the energy of the bulk system, we can use the relation defined in Eq. 4.17 that the total energy of our system in the bulk is equal to the sum of the chemical potentials of the constituent species in equilibrium. The energy change can be written in terms of the
chemical potentials of the species, $\mu_i$:

$$\Delta G_{\text{total}} = G_{\text{final}}(T, \{p_i\}, n_i) - \sum_i n_i \mu_i(T, p_i), \quad (4.20)$$

This equation describes the change in energy of bringing $n_i$ of each species $i$ from its reservoir (where it has reference energy $\mu_i$) to the final compound. In the specific case of MgO, this equation takes the form:

$$\Delta G_{\text{total}} = G_{\text{MgO}}(T, \{p_i\}, n_i) - n_{\text{Mg}} \mu_{\text{Mg}} - n_{\text{O}} \mu_{\text{O}}. \quad (4.21)$$

Using Eq. 4.17 again, and substituting for $\mu_{\text{Mg}}$, the equation can be rewritten:

$$\Delta G_{\text{total}} = G_{\text{MgO}}(T, \{p_i\}, n_i) - n_{\text{Mg}} \left( g_{\text{MgO}}^{\text{bulk}}(T, p) - \mu_{\text{O}}(T, p) \right) \mu_{\text{Mg}} - n_{\text{O}} \mu_{\text{O}}(T, p) \quad (4.22)$$

This is useful because it means that the surface energy can be written in terms of the bulk energy and the chemical potential of only one of the species. As in the case of determining the enthalpy of formation, ground-state DFT energies are to be used to determine changes in energy. The Gibbs free energy $G$ is equivalent to the Helmholtz energy $F$ under the assumption that the $PV$ contribution is small, especially when normalized by the surface area. The Gibbs free energy is then

$$G = F \quad (4.23)$$

$$= E_{\text{total}} + F_{\text{vib}} \quad (4.24)$$

$$= E_{\text{total}} + E_{\text{vib}} - TS_{\text{vib}} \quad (4.25)$$

In this case, however, the vibrational contributions to the Helmholtz energy cannot be so easily neglected. Reuter and Scheffler showed that in the case of a surface energy, the difference between the vibrational contributions of surface atoms and bulk atoms in RuO$_2$ was negligible. [Reuter and Scheffler, 2001] However, in later work they found that the vibrational contributions of adsorbed foreign species on the surface, especially when the species contain hydrogen, are significant and cannot be ignored. [Sun et al., 2003] In the results presented in this dissertation, the vibrational contributions to the surface energy for adsorbed species are included, and are summed according to

$$F_{\text{vib}} = \sum \frac{1}{2} \hbar \omega \quad (4.26)$$

Where $\omega$ is the energy of the vibrational contribution as calculated using DFT. Details on how these calculations are performed is presented in Chapter 6. The final Gibbs free energy change,
normalized by the surface area $A$ is represented by:

$$
\Delta G^{\text{total}} = \frac{1}{A} \left( E^{\text{total}} - n_{\text{Mg}} g_{\text{MgO}}^{\text{bulk}} + (n_{\text{Mg}} - n_{\text{O}}) \mu_{\text{O}}(T,p) + \sum \frac{1}{2} \hbar \omega \right) \quad (4.27)
$$

At this point, each term in the equation is obtainable by DFT calculations except for the chemical potential of oxygen $\mu_{\text{O}}$.

### 4.4 Temperature and Pressure Dependence

If ideal-gas behavior for oxygen is assumed, we can define the chemical potential at arbitrary temperature and pressure by

$$
\mu_{\text{O}}(T,p) = \frac{1}{2} \mu_{\text{O}_2}(T,p) = \mu_{\text{O}}(T,p^\circ) + \frac{1}{2} kT \ln \left( \frac{p}{p^\circ} \right), \quad (4.28)
$$

where $p^\circ$ is the standard pressure and $k$ is the Boltzmann constant. If a reference state where $\mu_{\text{O}}(0K,p) = \frac{1}{2} E^{\text{DFT}}_{\text{O}_2}$ is chosen, then the temperature and pressure dependence with respect to this zero can be described by

$$
\mu'_{\text{O}}(T,p^\circ) = \mu_{\text{O}}^{\text{O-rich}}(0K,p) + \frac{1}{2} \Delta G (\Delta T, p^\circ, \text{O}_2) \\
= 0 + \frac{1}{2} \left[ \Delta H - T \Delta S \right] \\
= \frac{1}{2} \left[ H(T,p^\circ, \text{O}_2) - H(0K,p^\circ, \text{O}_2) \right] - \frac{1}{2} T \left[ S(T,p^\circ, \text{O}_2) - S(0K,p^\circ, \text{O}_2) \right], \quad (4.29)
$$

and the full value of the chemical potential is found by

$$
\mu_{\text{O}}(T,p) = \frac{1}{2} \left( \mu_{\text{O}_2}(T,p) + E^{\text{DFT}}_{\text{O}_2} \right) \quad (4.30)
$$

The quantities needed in this expression can be obtained from thermochemical tables. Details on how these tables are used can be found in Chapter 6.

### 4.5 Formation Energies of Point Defects

DFT calculations of point defects aim to calculate the defect formation energy. This formation energy is useful because it allows comparison between different defects to determine which is more likely to form, and it allows the calculation of the concentration of defects in the sample. The general formation energy of the defect is given by:

$$
E_f = E_{\text{final}} - E_{\text{initial}} \quad (4.31)
$$
The equation is an expression of the cost of going from an initial system, which will be defined as an unperturbed, charge-neutral bulk system to a final defected system. Compared to the reference system, this defected system may be missing host atoms (as in the case of vacancies), may have additional atoms (as in the case of interstitial defects), and may be in a different charge state. A system with a substitutional defect will be missing host atoms and will have additional atoms. When considering the “initial” and “final” states in Eq. 4.31, it is useful to think of Hess’ law as it is usually presented in introductory chemistry courses: “products minus reactants”. [Hess, 1840] If we gained a species in our defected configuration, then our initial state should include a species of that type in a reservoir. If a species was lost from the initial host, then in the final state, that species should be in a reservoir. If any electrons were gained or lost, they must come from and return to their reservoir. The energy of a species in its reservoir is described by its chemical potential \( \mu_i \). The chemical potential of the electron is discussed later in the text.

An important part of determining the energy difference is determining what the limits of the chemical potentials of the atomic species are. In general, the defect formation energy for the change between an initial system that loses \( n_p \) particles of type \( p \) and gains \( n_r \) particles of type \( r \) to form the defected system is given by:

\[
E_f = \left( E_{\text{defected}}^{\text{DFT}} + \sum_p n_p \mu_p \right)_{\text{products}} - \left( E_{\text{bulk}}^{\text{DFT}} + \sum_r n_r \mu_r \right)_{\text{reactants}} \tag{4.32}
\]

Recall that the products of the reaction are the final defected system, plus the species \( p \) that end up in a reservoir, and that the reactants are the initial bulk system, plus the species \( r \) that begin in a reservoir and are then incorporated into the defect system. Because the initial state is a bulk system, \( p \) can only represent a chemical species of the host material, that is either one of the chemical elements that comprise it, or an electron. Assume that the change in number of electrons is captured in \( p \) (i.e. when \( p \) represents electrons, \( n_p \), the number of electrons lost from the system, can be negative if the product gains electrons). In this case, \( r \) contains all of the chemical elements that join the defected cell, and nothing else. Note that in the case of an interstitial, \( r \) can still describe atoms of the type in the host lattice, but they have nevertheless entered the system from a reservoir. Let \( n_e \) represent the number of electrons lost from the initial (neutral) state to the final state (i.e. \( p = e \)). The charge (\( q \)) of the final system is then \( n_e \). The defect formation energy equation can be rewritten by separating the products \( p \) into host species \( s \) and electrons \( e \), and replacing \( r \) with \( d \) to indicate that these are all defects that
enter the system:

\[ E_f = \left( E_{\text{DFT defected}} + \sum_s n_s \mu_s \right) + q\mu_e - \left( E_{\text{DFT bulk}} + \sum_d n_d \mu_d \right) \]  \hspace{1cm} (4.33)

The total energies from DFT calculations that enter into this equation are obtained with relative ease, though some minor caveats are addressed in Chapter 5. The chemical potentials of the \( s \) and \( d \) species are addressed in Section 4.6 and the chemical potential of the electron \( \mu_e \) is addressed in Chapter 6.

\section*{4.6 Chemical Boundary Conditions}

While the chemical potentials \( \mu \) in point defect and surface formation energies (e.g. Eq. 4.33 or Eq. 4.27) can in principle be varied from \(-\infty \) to \(+\infty \), there exist practical boundaries on the values that these chemical potentials can take. Recall that the chemical potentials of species in equilibrium with a bulk compound must obey the relationship:

\[ g_{\text{bulk}} = \sum_i n_i \mu_i \]  \hspace{1cm} (4.34)

Since the energy of the compound \( g \) is invariant, the chemical potentials of the components must be linked. Consider the defect formation energy of an aluminum vacancy in AlN, where the initial system has lost one Al atom to the reservoir, and assume that there is no change in charge:

\[ E_{\text{vac}}^f = E_{\text{DFT defected}} + \mu_{\text{Al}} - E_{\text{DFT bulk}} \]  \hspace{1cm} (4.35)

The chemical potential \( \mu_{\text{Al}} \) is the energy of the Al atom in the reservoir, and we see from Eq. 4.34 that this chemical potential is related to the energy of bulk AlN and the chemical potential of N:

\[ g_{\text{AlN bulk}}^{\text{AlN}} = \mu_{\text{Al}} + \mu_N \]  \hspace{1cm} (4.36)

Examination of Eq. 4.35 shows that higher values of \( \mu_{\text{Al}} \) lead to a higher formation energy, which means the defect is less likely to form. Higher values of the chemical potential indicate that the species is more prevalent or more reactive. This implies that under Al-rich conditions, when \( \mu_{\text{Al}} \) is high, the aluminum vacancy is less likely to form. The practical maximum to this chemical potential in equilibrium is where Al metal precipitates on the surface. At this point, the chemical potential is equal to the chemical potential of bulk Al

\[ \mu_{\text{Al max}} = \mu_{\text{Al}} = \mu_{\text{Al bulk}} \]  \hspace{1cm} (4.37)
This upper bound on $\mu_{Al}$ implies a lower bound on $\mu_N$ according to Eq. 4.36

$$\mu_{N_{\text{min}}} = g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^{\text{max}}$$

$$= g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^o$$

There is a similar upper bound on the chemical potential of N, which occurs when the $\mu_N$ is equal to its value in molecular N$_2$

$$\mu_{N_{\text{max}}} = \mu_{N}^o = \frac{1}{2}\mu_{N_2}$$

This represents the Nitrogen-rich growth extreme.

In order to quantify the growth conditions easily it is useful to define a quantity $\Delta \mu$ which represents how far away the chemical potentials of the species are compared to the extreme conditions.

$$\Delta \mu = (\mu_{Al} - \mu_N) - (\mu_{Al}^o - \mu_N^o)$$

The boundary conditions on $\Delta \mu$ are defined by the Al-rich and N-rich extremes.

Under Al-rich conditions, where $\mu_{Al} = \mu_{Al}^o$ and $\mu_N = g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^o$

$$\Delta \mu = (\mu_{Al} - \mu_N) - (\mu_{Al}^o - \mu_N^o)$$

$$= (\mu_{Al}^o - (g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^o)) - (\mu_{Al}^o - \mu_N^o)$$

$$= - (g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^o) + \mu_N^o$$

The enthalpy of formation of Eq. 4.16 can be written for AlN as

$$\Delta H_{\text{AlN}}^{\text{AlN}} = g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^o - \mu_N^o$$

$$\Delta H_{\text{AlN}}^{\text{N}} + \mu_N^o = g_{\text{bulk}}^{\text{AlN}} - \mu_{Al}^o$$

Substituting the left hand side of this equation into Eq. 4.44 allows the equation for $\Delta \mu$ in the Al-rich extreme to be simplified further:

$$\Delta \mu = - \left( \Delta H_{\text{AlN}}^{\text{AlN}} + \mu_N^o \right) + \mu_N^o$$

$$= - \Delta H_{\text{AlN}}$$

The same procedure can be repeated for N-rich conditions and the appropriate substitutions, which yields $\Delta \mu = +\Delta H_{\text{AlN}}$, and it can be shown that for $\Delta \mu \in (-\Delta H, +\Delta H)$. Thus, $\Delta \mu$
provides a parameter that can be used to uniquely define $\mu_{Al}$ and $\mu_{N}$:

$$\Delta \mu = (\mu_{Al} - \mu_{N}) - (\mu_{Al}^{0} - \mu_{N}^{0})$$  \hspace{1cm} (4.49)

$$\mu_{Al} = \Delta \mu + \mu_{N} + (\mu_{Al}^{0} - \mu_{N}^{0})$$  \hspace{1cm} (4.50)

$$= \Delta \mu + \left( g_{\text{AlN}}^{\text{bulk}} - \mu_{Al} \right) + (\mu_{Al}^{0} - \mu_{N}^{0})$$  \hspace{1cm} (4.51)

$$= \frac{1}{2} \left( \Delta \mu + g_{\text{AlN}}^{\text{bulk}} + (\mu_{Al}^{0} - \mu_{N}^{0}) \right)$$  \hspace{1cm} (4.52)

Thus, $\mu_{Al}$ can be determined from a given $\Delta \mu$ and other known quantities, and can be used to determine the chemical potential of $N$, if required: $\mu_{N} = g_{\text{AlN}}^\text{bulk} - \mu_{Al}$.

### 4.7 Chemical potentials of impurities

For defects that contain foreign species, the chemical potential of that species enters into the formation energy equation. The value of this chemical potential can also be defined as a range of accessible chemical potentials. Determining an upper limit to the chemical potential requires considering the possible solubility limiting phases. In the case of Be incorporation in GaN, for instance, the chemical potential of Be in the Be-rich case $\mu_{Be}$ is limited by the formation of Be$_3$N$_2$. The upper limit of the chemical potential of Be is then connected to the chemical potential of N, which varies from N-rich to Ga-rich depending on the growth conditions. The lower limit of Be is unbounded, so it is often useful to examine the system at a specified value of $\mu_{Be}$ for comparison to other types of defects. The values of the specific chemical potentials must be evaluated to determine whether a compound or the bulk elemental phase is limiting under each set of growth conditions.

### 4.8 Defect Concentration

In Section 4.5, the formation energy of point defects given in Eq. 4.33 was only the 0K contribution to the total energy, that is, the enthalpy change to form one defect $\Delta H$. The change in Gibbs free energy from a perfect crystal to one with $n$ defects $\Delta G_f(n)$ is given by:

$$\Delta G_f(n) = n\Delta H - T\Delta S.$$  \hspace{1cm} (4.53)

Where the entropy $S$ includes multiple separate contributions to the total entropy. If the entropic effects other than configurational entropy are treated separately, then $\Delta S = \Delta S_{\text{conf}} + \Delta S_{\text{other}}$. Configurational entropy is a measure of the number of ways a system can be configured, given by Boltzmann's formula:

$$S = k_B \ln W,$$  \hspace{1cm} (4.54)
where $W$ is the number of ways in which the system can be configured. In the case of the final system, a host crystal with simple substitutional defects, the $W_f$ of interest is the number of ways $n$ defects can be arranged on $N_s$ possible sites

$$W_f = \binom{N_s}{n} \quad (4.55)$$

The initial system, a perfect crystal with no defects has only one possible configuration, while a system with one defect $n = 1$ has as many configurations as there are available sites $N_s$ for the defect

$$W_i = \binom{N_s}{1} \quad (4.56)$$

The number of these configurations increases with a second and third defect, but will reach a maximum when the number of defects equals the number of available sites. Conceptually, a material consisting solely of one type of defect has only one way in which it can be configured. There exists therefore some optimal number of defects which maximizes the number of configurations of the system. Assuming that defects are dilute enough that $\Delta H$ is constant as a function of $n$, there is an enthalpic cost to adding each additional defect. The entropic contribution has some maximum, and since this positive entropic term is subtracted from the enthalpic term, there is some number of defects that minimizes the total energy. This minimum can be found by taking the derivative with respect to $n$ (assuming that the additional contributions to the entropy do not vary significantly with respect to $n$) and finding where it equals zero:

$$\frac{\partial \Delta G_f(n)}{\partial n} = 0 = \Delta H + \frac{n \partial \Delta H}{\partial n} - T \frac{\partial \Delta S_{\text{conf}}}{\partial n} \quad (4.57)$$

If we have defined $S_{\text{conf}} = k_B \ln W$, then the derivative of the difference between the two states can be written:

$$\frac{\partial \Delta S_{\text{conf}}}{\partial n} = \frac{\partial}{\partial n} \left( k_B \ln \binom{N_s}{n} - k_B \ln \binom{N_s}{1} \right) \quad (4.58)$$

$$= \frac{\partial}{\partial n} \left( k_B \ln \frac{N_s!}{n!(N_s-n)!} - k_B \ln \frac{N_s!}{1!(N_s-1)!} \right) \quad (4.59)$$

$$= \frac{\partial}{\partial n} k_B \left( \ln N_s! - \ln n! - \ln(N_s-n)! \right) \quad (4.60)$$
This can be simplified by Stirling’s approximation \( \ln x! = x \ln x - x \) under the conditions that \( x \gg 1 \)

\[
\frac{\partial \Delta S_{\text{conf}}}{\partial n} = \frac{\partial}{\partial n} k_b (\ln N_s - \ln n - \ln (N_s - n)!) \\
= \frac{\partial}{\partial n} k_b ((N_s \ln N_s - N_s) - (n \ln n - n!) - ((N_s - n) \ln (N_s - n) - (N_s - n)]) \\
= k_b \left( \frac{\partial}{\partial n} (N_s \ln N_s) - \frac{\partial}{\partial n} (n \ln n) - \frac{\partial}{\partial n} ((N_s - n) \ln (N_s - n)) \right) \\
= k_b (- (\ln n + 1) - (- \ln (N_s - n) - 1)) \\
= k_b (\ln (N_s - n) - \ln(n)) \\
= k_b \ln \frac{N_s - n}{n} 
\]

Inserting this into our minimization equation, simplification and the assumption that \( N_s/n \gg 1 \):

\[
0 = \Delta H - Tk_b \ln \frac{N_s - n}{n} \\
\frac{\Delta H}{Tk_b} = \ln \frac{N_s - n}{n} \\
\frac{N_s - n}{n} = \exp \frac{\Delta H}{k_b T} \\
n = N_s \exp \frac{-\Delta H}{k_b T} 
\]

Typically the number of sites \( N_s \) is normalized as sites per unit volume, usually \( \text{cm}^{-3} \), which gives the concentrations of defects \( n \) in the same units.

### 4.8.1 More Complicated Defects and Their Entropies

In the discussion above, it was noted that the defects must be of the simple substitutional. If further degrees of freedom are introduced, these must be taken into consideration in counting the number of ways the \( n \) defects can be distributed. Two specific examples of relevance in this dissertation are DX centers, which are off-site substitutional defects, and ionized donor defects are used to illustrate this point.

A DX center is a substitutional defect that has a broken bond with one of its nearest neighbors. In wurtzite structures, this means that four possible broken bond configurations exist. If the formation energy for all four directions is equal, then the defect is equally likely to form in each. This means that for each site that the defect can incorporate on, there are four possible configurations. This leads to the total number of ways the system can be configured to
be

\[ W_f = \binom{N_s}{n} 4^n \]  

(4.63)

Using this expression for \( W_f \) and repeating the above procedure to find the value \( n \) which minimizes the total free energy leads to the solution:

\[ n = 4N_s \exp \frac{-\Delta H}{k_b T} \]

(4.64)

Note that the factor of four that appears is the same as the number of structural configurations for each defect. Some defects may have a multiplicity other than four, so in general, this structural multiplicity will be called \( N_{\text{struct}} \).

Consider the case of a neutral donor defect. When this defect has only one electron, the electron may be spin-up or spin-down. In this case, the number of ways the spins can be configured \( N_{\text{elec}} \) for \( n \) sites is \( 2^n \), making the total number of configurations

\[ W_f = \binom{N_s}{n} 2^n \]  

(4.65)

The general expression for the equilibrium number of defects is then given by

\[ n = N_{\text{elec}} N_{\text{struct}} N_{\text{sites}} \exp \frac{-\Delta H}{k_b T} \]

(4.66)

In most cases \( N_{\text{struct}} = 1 \), while for neutral donors and acceptors \( N_{\text{elec}} = 2 \) and for ionized donors and acceptors \( N_{\text{elec}} = 1 \).
Chapter 5

Calculation Details - Setup

The main results of investigations of surfaces, interfaces, and point defects are presented in paper chapters later in this dissertation. Each paper chapter includes some details on the calculations that were performed to obtain those results, but this chapter and the chapter that follows will include a more thorough level of detail on the methods used.

An investigation of any materials problem using DFT will require three main steps: pre-processing, “production” runs, and post-processing.

Pre-processing covers the steps necessary to set up a successful DFT study of the problem at hand, including choosing an exchange-correlation functional and pseudopotential, determining the initial structure of the system, and deciding which variants of the relevant defect (which surfaces? which adsorbates? which impurities on which sites?) to investigate.

The “production” runs are the main DFT calculations that follow the prescription defined in Section 3.1. These calculations determine a final equilibrium atomic configuration and minimum energy of the system. Often, these calculations will produce other outputs\(^1\) that are used in the final analysis. In other cases, further DFT calculations based on the final configurations of these “production” runs are needed.

These secondary DFT calculations that may be needed to converge, e.g., charge densities, density of states, projected density of states, or constrained-geometry calculations in different charge states, are included in the post-processing steps. Post-processing also includes the extraction of information reported by DFT in order to populate equations, such as formation energy equations. In a more complex example of post-processing, this dissertation discusses the solutions to a set of self-consistent mass balance equations to determine the equilibrium defect concentrations based on the calculated DFT outputs.

Portions of these steps are generic to DFT, while others are specific to the implementation

---

\(^1\)These outputs include charge density, local potential, and electronic density of states. For further discussion of these outputs, see Chapter 3
used; in this dissertation VASP is used to describe the calculations. Some steps are applicable to any problem in materials science to which DFT is being applied, and others are specific to the materials system or type of defect studied.

5.1 Equilibrium Geometries

Once the desired exchange-correlation functional and pseudopotentials have been chosen, the equilibrium geometry of the system must be found. The equilibrium geometry is the one that minimizes the energy as a function of atomic positions. It has already been shown that the main iterative loops of atomic relaxation in DFT find a minimum energy with respect to atomic coordinates in a unit cell, but these relaxations typically assume a fixed unit cell size. The true minimum energy geometry must also find the minimum with respect to the volume of the system. There are sophisticated built-in procedures for an automated volume relaxation in VASP, but these are not discussed here. The simpler approach is to vary the cell size and determine, according to a preferred equation of states (EOS) (such as the Birch-Murnaghan EOS), the cell volume which minimizes the energy of the system. This approach is straightforward for cubic cells where the lattice parameter in all three dimensions is the same \((a = b = c)\), but is not as simple for systems where even one of the three basis vectors is a different length such as a hexagonal \((a \neq c)\) or tetragonal cell \((a = b \neq c)\).

5.2 Equilibrium Geometry of Tetragonal and Hexagonal Unit Cells

As discussed in Section 5.1, the equilibrium geometries of the bulk materials of interest must be determined before calculations on the defects can begin. In calculations of point defects that depend on accurately capturing the ionization energies of defects, hybrid functionals have been shown to reproduce experimental observations, and are used in the results presented in this dissertation. These hybrid functionals, however, are approximately ten to forty times more computationally expensive than traditional DFT calculations. Several DFT calculations are typically required in order to find the equilibrium geometry, and the complexity of this problem is increased when hybrid functionals that vary \(\alpha\), the amount of exact exchange, are used, because each \(\alpha\) has a different minimum geometry. Thus, while adjusting the amount of mixing in order to find the \(\alpha\) that optimizes the electronic band gap, many rounds of expensive calculations must be performed. It is therefore sensible to find an efficient procedure for determining the equilibrium geometry so that the “production” calculations can be performed. To find the equilibrium geometry, that is: the set of atomic coordinates that minimize the total
energy of the system, there are two first-order degrees of freedom: the size and shape of the cell (determined by the three basis vectors) and the atomic positions within the cell. For a cubic cell, the cell volume has only one degree of freedom, defined by the lattice parameter. Thus, varying the lattice parameter defines unique unit cells of different volumes whose energy can be determined, and the energy versus volume relationship can be fitted using an appropriate equation of states, such as the Birch-Murnaghan EOS. For a tetragonal or hexagonal unit cell, where there exist two degrees of freedom (tetragonal: \(a = b \neq c\), hexagonal \(a \neq c\)), this procedure cannot be applied immediately. Mehl, et al., have shown that for a tetragonal unit cell, an equation of states that relates energy as a function of volume can be used, provided that the energy is the minimum energy for the given volume. [Mehl et al., 1993] This provides an additional complication because the volume is not uniquely defined by one value as in the case of a cubic material. The prescription defined by Mehl, et al., is to select a set of volumes, and for each volume vary one lattice parameter (say, \(a\)) and then to use \(V = a \cdot a \cdot c\) to define the second parameter \(c\), and determine the final energy of each of these configurations. Then, for each volume \(V\), determine the combination of \(a\) and \(c\) which minimizes \(E(a, c)|_{V = V'}\), and use this minimum \(E(V')\) in the equation of state.

Presented here is a procedure for determining the combination of \(a\) and \(c\) which minimize the energy of the cell using a set of 26 DFT calculations where traditional brute-force sweeps would require \(\approx 150\) calculations. First, trial values of the lattice parameters \(a_T\) and \((c/a)_T\) are chosen, based on theoretical or experimental input. Based on these initial parameters, five DFT calculations are run using the lattice constants shown in Table 5.1. As seen in the table, these five runs have five different volumes. A first guess for the minimum volume \(V_{BM}\) of the system is obtained by fitting the energies of these five systems using a Birch-Murnaghan fit. Because of the nature of this approximation, the midpoint energy \(E(a' = a_T, (c/a)') = (c/a)_T\) must be the lowest of the five energy points. If not, the initial search grid should be redefined with the combination of \(a\) and \(c/a\) which gave the lowest energy as the center of expansion. At this point, three volumes are chosen \((0.975V_{BM}, V_{BM}, \text{ and } 1.025V_{BM})\), and for each volume a set of five isochoric (volume conserving) calculations are prepared by choosing five values of \((c/a)_{j} =\), which are shown in Table 5.2, and using these values to determine \(a_{ij}\) for each volume \(V_i\). This gives three sets of volumes, each with five combinations of \(a\) and \(c/a\) whose energy can be calculated. For each volume, the minimum energy is determined by a fit to a quadratic function. These three minimum energies for the three volumes are used to fit a Birch-Murnaghan equation of states to provide a refined estimate of the final volume \(V_{BM2}\). This volume is used for a new isochoric calculation, using five \((c/a)\) values as described previously. These five energies are fit using a quadratic function to determine the \((c/a)\) (and therefore \(a\)) that minimizes the energy at this volume. A final calculation is performed at this minimum \((c/a)\) and \(a\) combination to give the final minimum energy of the system.
Table 5.1: The first step of the energy minimization procedure involves a coarse sweep based on an initial trial volume.

<table>
<thead>
<tr>
<th>Run No.</th>
<th>( \frac{a'}{a_T} )</th>
<th>( \frac{(c/a)'}{(c/a)_T} )</th>
<th>( V' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.95</td>
<td>0.95</td>
<td>0.8145</td>
</tr>
<tr>
<td>2</td>
<td>0.95</td>
<td>1.05</td>
<td>0.9002</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.00</td>
<td>1.0000</td>
</tr>
<tr>
<td>4</td>
<td>1.05</td>
<td>0.95</td>
<td>1.0997</td>
</tr>
<tr>
<td>5</td>
<td>1.05</td>
<td>1.05</td>
<td>1.2155</td>
</tr>
</tbody>
</table>

Table 5.2: The second step of the energy minimization procedure involves choosing values of \( c/a \) around a central value, which is then used to determine \( a \).

<table>
<thead>
<tr>
<th>Run No.</th>
<th>( (c/a)_j - (c/a)_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.04</td>
</tr>
<tr>
<td>2</td>
<td>-0.02</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>+0.02</td>
</tr>
<tr>
<td>5</td>
<td>+0.04</td>
</tr>
</tbody>
</table>
5.3 HSE Band Gap Calculations

The standard VASP procedure for calculating the band gap involves running a self-consistent calculation to converge the wavefunctions and charge density. The charge density of a VASP calculation is stored in the CHGCAR file, and the next geometry is stored in the CONTCAR file. This charge density and final equilibrium geometry is then used to start a new non-self-consistent calculation where the wavefunctions are evaluated at a series of special \( k \)-points along the paths of interest. The energies of the plane waves at these \( k \)-points are evaluated and used to generate the band diagram.

It was discussed in Section 2.12.6 that the HSE hybrid functional better represented the electronic band gaps of many materials, and that it has been used with success to describe defect ionization energies in semiconductors. It stands to reason then, that it would be useful to calculate the electronic band diagram when using the HSE functional in order to verify that the functional is capturing important aspects of the band structure. The implementation of band structure calculations using HSE does not proceed in the same manner as standard DFT calculations. This is because the standard VASP procedure for calculating band gaps or band structure using local or semi-local functionals relies on restarting a calculation with the converged local charge density, but the contributions of the exact exchange are not stored in the CHGCAR file. If the standard DFT procedure is used, the initial charge density will not be based on the final charge density, and will not be updated. Instead, the correct procedure using HSE is to run the self-consistent HSE calculation, and then use the converged positions as the starting point of a second self-consistent calculation. The original KPOINTS file identifies the size of the \( k \)-point mesh to be used in the calculation, as well as the algorithm by which the \( k \)-points in the irreducible Brillouin zone are chosen. The \( k \)-points that are actually used in the self-consistent calculation are recorded by VASP in the IBZKPT file. These \( k \)-points are represented by four values: their three reciprocal-space coordinates, and their weight in the calculation. For the band-structure calculation, the IBZKPT file is copied to the KPOINTS file (that is, the \( k \)-points for the calculation are manually specified and not generated automatically. Then, the special \( k \)-points from the paths of interest are appended to the file. Importantly, the weight for these additional \( k \)-points should be set to zero, so that they are not included in the criteria for convergence. The second line of the explicitly defined KPOINTS file contains the number of \( k \)-points, and this number must be updated for the total number of \( k \)-points (the original number of points in the IBZ and the number of special points). Then, a new self-consistent calculation must be run. Be advised that the calculation will stop when the wavefunctions have converged at the original (weighted) \( k \)-points, and no guarantee is made about what the occupancies at the special zero-weight \( k \)-points are. Thus, it is necessary to require several self-consistent steps, regardless of the convergence at the weighted \( k \)-points by specifying NELMIN=5 in the INCAR file.
Even after five self-consistent steps, it may be necessary to check that the non-self-consistent k-points are converged, by checking that they have integral occupancies.

5.4 Determining High-Symmetry Paths

Determining which paths in reciprocal space are interesting to report in a band-structure calculation is not always straightforward. The first priority is the ability to locate the direct and indirect band gaps. A second priority is to present the same information that has been found by previous experimental and theoretical investigations. Typically the high-symmetry k-points of a given crystal structure are well known, and the paths reported in the band diagram should connect these points. The work of Setyawan and Curtarolo [Setyawan and Curtarolo, 2010] provides an excellent resource for a standardized set of k-paths. For instance, in the case of the hexagonal lattice (the lattice of AlN AlN and GaN GaN), the recommended path is \( \Gamma - M - K - \Gamma - A - L - H - A | L - M | K - H \). The paper also defines the reciprocal space coordinates of these special points, and is therefore immensely useful for these sorts of calculations.

As an example of how to perform the HSE-style band structure calculation for the first part of the hexagonal path, from \( \Gamma - M \), using four interpolation points, the first step is to determine the reciprocal space coordinates of the special points: \( \Gamma = (0, 0, 0) \) and \( M = (\frac{1}{2}, 0, 0) \). This short path is then represented by the set of six k-points:

\[
\begin{align*}
0.0 & 0.0 0.0 0.0 \ #\Gamma \\
0.1 & 0.0 0.0 0 \\
0.2 & 0.0 0.0 0 \\
0.3 & 0.0 0.0 0 \\
0.4 & 0.0 0.0 0 \\
0.5 & 0.0 0.0 0 \ #M
\end{align*}
\]

If the final IBZKPT file of the first self-consistent calculation was:

<table>
<thead>
<tr>
<th>Automatically generated mesh</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
</tr>
<tr>
<td>Reciprocal lattice</td>
</tr>
<tr>
<td>0.2500000000000000</td>
</tr>
<tr>
<td>0.2500000000000000</td>
</tr>
<tr>
<td>0.2500000000000000</td>
</tr>
<tr>
<td>0.2500000000000000</td>
</tr>
<tr>
<td>0.2500000000000000</td>
</tr>
<tr>
<td>0.2500000000000000</td>
</tr>
</tbody>
</table>

Then this file would be copied to KPOINTS and the six k-points would be appended to the file, and the total number of k-points in the second line of the file would be updated:

<table>
<thead>
<tr>
<th>Automatically generated mesh</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
</tr>
</tbody>
</table>
At this point, the second self-consistent run is performed. It is recommended to use the Davidson diagonalization scheme (IALGO=N) to converge the wavefunctions, since it will converge the zero-weight \( k \)-points faster.

### 5.5 Effects of HSE Mixing

As stated in Section 2.12.6, including a fraction of exact exchange has been shown to improve the calculated electronic band gap of many materials. In the original HSE formulation [Heyd et al., 2003, Heyd et al., 2006] the mixing parameter \( \alpha \), the amount of exact exchange, was defined to be 0.25 based on the definition in the PBE0 functional [Perdew et al., 1996b], where it was determined by perturbation theory. Studies of a set of molecular species showed that the \( \alpha = 0.25 \) selection yielded low mean absolute error, and a study of the enthalpy of formation, lattice parameter and band gap of 40 semiconductors showed similar low average error when using this mixing parameter. [Heyd et al., 2003, Heyd et al., 2005] It can be shown that the effect of the mixing parameter on the band gap is linear. The parameter has been used to empirically fit the experimental band gap in GaN [Lyons et al., 2010] and in so doing has been shown to accurately reproduce experimental ionization energies of defects. In the calculations presented in this dissertation, the amount of mixing has been adjusted to match the experimental band gap.

The equilibrium geometry, which is defined by a geometric configuration that minimizes the system energy as described in Section 5.1 is a function of the mixing parameter. Thus, the geometry minimization must be made for each mixing parameter \( \alpha \) and then the band gap must be calculated for this \( \alpha \) according to the procedure in Section 5.3. This procedure is somewhat simplified by the fact that the energy gaps changes approximately linearly with the mixing parameter \( \alpha \).
5.6 Relaxation Shell

When running a DFT calculation on a system with a point defect, some care must be taken to account for the fact that the defect concentration in the $\sim 100$ atom supercell\(^2\) is two to six orders of magnitude higher than in real materials. The defect formation energy that should be calculated is the formation energy in the dilute limit. This is accomplished within the constraints of a periodically repeated supercell by first using relatively large supercells (about 100 atoms for wurtzite GaN and AlN). Second, atoms near the defect are allowed to relax, while atoms further away are constrained in their bulk-like positions. The appropriate number of neighbor shells to relax will change with different materials and types of defects, as will the size of the supercell needed to converge the defect formation energy. For a new materials system, careful convergence tests should be performed to ensure the spurious interactions between defects are minimized. In the case of GaN and AlN, relaxing the four nearest-neighbor shells around the defect has been shown to yield convergence. [Lyons et al., 2010, Collazo et al., 2012] Note that this relaxation shell will differ for defects that sit at different sites, and so a set of calculations of, e.g., nitrogen vacancies will not have the same atoms fixed as a calculation of an aluminum vacancy (though each will have the same number of atoms fixed).

5.7 Defects of Interest

All that remains is to determine the defects of interest. Native defects, such as vacancies, interstitial, and anti-sites, will often be of interest. Many materials will have multiple interstitial configurations available. Impurity defects can be considered based on experimental input such as SIMS profiles of the impurities present. Once the defects have been decided, it is also necessary to select the likely defect charge states. This can be accomplished by comparing the bonding character of the defect to the original site. In the case of a group IV element in AlN, for instance, it may substitute on either the III site (bringing one extra electron compared to the number needed to satisfy the local bonds) or the V site (bringing one less electron). When the IV element substitutes on the III site, the sensible guess is that the dominant charge state will be the +1 state. In wide band gap materials, however, there exists ample opportunity for other charge states to become favorable depending on the value of the Fermi level, and so charge states above and below the state of interest should be examined (0 and +2). As a general guideline, if it is later found that either of these additional states show a transition within the band gap, the next charge state in either direction should also be examined to ensure that no favorable states are missed. Examples of these transitions and how they are determined after the DFT calculations have been run are described in Chapter 6.

\(^2\)One vacancy in a 96 atom AlN supercell corresponds to a defect concentration of $1 \times 10^{21}$ cm$^{-3}$
5.8 Surface Geometries

Selecting surfaces of interest is based on input from previous experimental and theoretical studies. Interesting surfaces will typically be those that are often grown or observed, or those for which attempts at growth have been made but are not yet successful. Once the set of surfaces has been decided, the supercells needed for a DFT calculation can be constructed. The goal of the surface calculations in this work are to represent the surface energy of a thin-film material. The thin-films to which these surfaces are being compared are hundreds of nanometers thick, whereas a large supercell in DFT is less than 10nm in one dimension. These calculations must be set up in a way that uses an ultra-thin film to reproduce the behavior of the surfaces of a much thicker film or bulk material. A slab morphology, where the atoms of interest are continuously periodically repeated in two dimensions, but has free surfaces in the other dimension is often used in plane-wave implementations. Due to the periodic nature of plane-wave implementations, the slab supercell is also repeated in the direction above and below the surface. Care must be taken to ensure that the slabs to not interact with their periodic images across the vacuum. Typically this can be accomplished by ensuring that the periodic images are at least 10Å away from each other. Polar surfaces, however, provide an additional challenge because depending on how the crystal is terminated, the two surfaces may have different terminations, leading to an electric field in the vacuum. This directional electric field exists across the periodic boundary and leads to calculated energies that change with vacuum distance. This electric field can be avoided if the slab is symmetric along the surface axis and terminated by the same species on each side. This is feasible for \{111\} surfaces of rocksalt oxides, and indeed this is advantageous because the calculated surface energy will correspond to the creation of two identical surfaces. Note that surface configurations for materials that do not exhibit inversion symmetry (such as the \{0001\} surfaces of wurtzite GaN or ZnO) require additional considerations that are not discussed here.

The expression for calculating the surface free energy of Eq. 4.27 requires the DFT energy of the slab system, as well as the energy of a formula unit of the bulk compound. It is good practice to calculate this bulk energy in a supercell of the same size and orientation, and with the same basis set, as is used for the slab. This is made easier by first constructing the bulk supercell which has been repeated in the desired surface direction to be used in the calculation, and then removing approximately half of the atomic layers. The number of layers used in the calculation will determine how closely the configuration reproduces bulk-like surface behavior. Common practice is to test convergence of the surface energy as a function of slab thickness, typically with several middle layers of the slab frozen in their bulk-like positions. Calculation results on the surfaces presented in this dissertation are typically performed with 16 layers in a slab, and middle five layers frozen.
Chapter 6

Post Processing

Once the DFT calculations have finished and the final system energies are reported, the post-processing phase includes the steps necessary to take the DFT outputs discussed in Chapter 3 and draw conclusions about the defect properties.

6.1 Point Defects

Calculations of point defects seek to find the equilibrium concentration of defects, thermodynamic transitions, and optical transitions.

6.1.1 Point Defect Formation Energies

All of these depend on the formation energy of defects, as derived in Eq. 4.5. Eq. 4.33 is reproduced here:

$$E_f = \left( E_{\text{DFT defected}} + \sum_s n_s \mu_s \right) + q \mu_e - \left( E_{\text{DFT bulk}} + \sum_d n_d \mu_d \right).$$

If our DFT calculations were set up properly, the energy of the bulk system and the energy of the defected cell can be used directly for $E_{\text{DFT defected}}$ and $E_{\text{DFT bulk}}$. Based on the defect cell of interest, $n_s$, $n_d$, and $q$ are known, and the chemical potentials of the species $\mu_i$ can be tailored based on the growth conditions $\Delta \mu$ as described in Section 4.6 and Section 4.7. The remaining unknown in this equation is $\mu_e$, the chemical potential of the electron in its reservoir. The standard definition of the chemical potential of the electron applies here: $\mu_e$ is the Fermi level of the material. For the purposes of comparison, it is useful to reference the chemical potential to the top of the valence band, which is obtained from the DFT calculated value of the highest occupied state in the defect-free bulk calculation, and denoted VBM. However, because of the lack of a common
reference between two DFT calculations, this chemical potential must be adjusted so that the
two states can be compared. The misalignment between two states can be calculated from the
differences between the electrostatic potentials of the two systems. The electrostatic potential
(which excludes the exchange-correlation potential) is used because it converges more rapidly
compared to the total potential. This potential was discussed in Section 3.3.4, and its value at
discrete points in the supercell is reported in the VASP output file LOCPOT. The value of the local
potential that is desired is potential at the vacuum level, but since these cells do not contain
a free surface, the vacuum level cannot be computed directly. Comparing this local potential
between the bulk cell and the defect cell shows that it varies most significantly near the defect,
and that the values will differ at positions near atoms which have relaxed to different positions
between the two cells. The difference between local potential value closest to the differences
in the vacuum level is therefore found in an interstitial site (further from atoms) far away
from the defect, where the surrounding atoms were not allowed to relax during the calculation,
as discussed in Section 5.6. The value of the local potential as reported in the LOCPOT file is
reported at discrete points, as defined by the real-space discretization mesh used by the charge
density, defined by the VASP flags NGX, NGY, NGZ. Typically the distance between these discrete
points is 0.18 Å, while the distance between the interstitial site and the surrounding atoms is
2.03 Å. To reduce noise in the comparison, the value of the local potential is averaged at 27
points that form a cube at the center of the interstitial site. The difference between these two
averages is denoted ∆V. The value for the chemical potential of the electron reservoir is then
given by
\[ \mu_e = \text{VBM} + E_F + \Delta V, \]  
so that when \( E_F \) is zero, the electron is aligned with the valence band maximum, and when
\( E_F \) is equal to the band gap \( E_g \), the electron is aligned with the conduction band minimum.
In this manner \( E_F \) can be varied from \((0, E_g)\) to calculated the defect formation energy as a
function of the position of the Fermi level in the band gap. It was also shown in Section 4.6 that
\( \Delta \mu \) could be varied from \((-\Delta H, +\Delta H)\) to represent the growth conditions. The total defect
formation energy can be calculated as a function of any combination of these values:
\[ E_f (\Delta \mu, E_F) = \left( E_{\text{DFT, defected}} + \sum_s n_s \mu_s \right) + q (\text{VBM} + E_F + \Delta V) - \left( E_{\text{DFT, bulk}} + \sum_d n_d \mu_d \right). \]  
In this sense, the growth conditions (represented by \( \Delta \mu \)) and the Fermi level (represented by
\( E_F \)) can be adjusted at will to determine the defect formation energy under a desired condition,
even though only one DFT calculation was run for the defect in charge state \( q \).\(^1\)

\(^1\)Note that a separate calculation must be run for every desired charge state \( q \).
formation energy for the vacancy in charge state $q$ takes the form

$$E_f[V_{Al}^q](\Delta\mu, E_F) = \left( E_{DFT}[V_{Al}^q] + \mu_{Al} \right) + q(VBM + E_F + \Delta V) - E_{bulk}^{DFT} \quad (6.3)$$

because the host has lost one Al to the reservoir. Recall that the chemical potential of Al is connected to the growth conditions via $\Delta\mu$ according to Eq. 4.52, and the formation energy can be written as a function of $\Delta\mu$ and $E_F$ with all other quantities known:

$$E_f[V_{Al}^q](\Delta\mu, E_F) = \left( E_{DFT}[V_{Al}^q] + \frac{1}{2} \left( \Delta\mu + g_{AlN}^{bulk} + (\mu_{Al}^{\circ} - \mu_{N}^{\circ}) \right) \right) + q(VBM + E_F + \Delta V) - E_{bulk}^{DFT} \quad (6.4)$$

6.1.2 Formation Energy Diagrams

In order to explore the energies of the defects in each charge state, either the Fermi level can be fixed and the growth conditions varied, or vice versa. Figures Figure 6.1 and Figure 6.2 show the formation energy of the aluminum vacancy for $q \in \{1, 0, -1, -2, -3\}$. For all three plots in Figure 6.1, the growth condition is varied from N-rich ($\Delta\mu = +\Delta H_{AlN}$) to Al-rich ($\Delta\mu = -\Delta H_{AlN}$), while the Fermi level is fixed at the valence band maximum ($E_F = 0$, Figure 6.1a), the mid-gap ($E_F = E_g$, Figure 6.1b), and at the conduction band minimum ($E_F = E_g$, Figure 6.1c). When examining only one defect, the plots of this type are of only limited use. In each plot, the slope of all the $V_{Al}$ defects are the same (the slope is proportional to $\partial\mu_{Al}/\partial\Delta\mu$), and so across $\Delta\mu$ the favorability of one charge state over another is constant (the lowest energy charge state is the most likely to form). It is evident, however, that depending on the value of the Fermi level (plot a vs. b vs. c) that this most favorable defect changes. This indicates that as the Fermi level moves from the VBM to the CBM (from p-type material to n-type material), there is a change in favorability between the charge states. This can be more clearly seen when the formation energies are shown with continuously varying Fermi level, as is shown in Figure 6.2.

In contrast, Figure Figure 6.2 shows the formation energy of the same charge states of $V_{Al}$ for a continuously varying Fermi level at three fixed growth conditions: N-rich (Figure 6.2a), midway between the extremes $\Delta\mu = 0$ (Figure 6.2b), and Al-rich (Figure 6.2c). In these figures, the changes in which charge state is lowest in energy (and therefore more favorable) are more clearly seen. Furthermore, this figure allows for the identification of the Fermi level at which the most favorable defect changes charge state. Note that these transitions, known as the thermodynamic transitions, do not change based on the growth conditions of the material.

It is also useful to compare the formation energies of different defects using the two styles of formation energy diagram mentioned above. Figures Figure 6.3 and Figure 6.4 show the
Figure 6.1: The formation energy $E_f$ of $V_{Al}$ defects are shown as a function of the growth conditions $\Delta \mu$ ranging from N-rich to Al-rich, at three different fixed values of the Fermi level: (a) $E_F = 0$, the valence band maximum, (b) $E_F = E_g/2$, at the mid-gap, and (c) $E_F = E_g$, at the conduction band minimum.
Figure 6.2: The formation energy $E_f$ of $V_{Al}$ defects are shown as a function of the Fermi level $E_F$ ranging from the valence band maximum ($E_F = 0$) to the conduction band minimum ($E_F = E_g$), at three different fixed growth conditions: (a) $\Delta \mu = +\Delta H$, N-rich conditions, (b) $\Delta \mu = 0$, halfway between the extremes, and (c) $\Delta \mu = -\Delta H$, Al-rich conditions.
formation energies of $V_{Al}$ and $V_N$. Figure Figure 6.3 is constructed similar to figure Figure 6.1 which only showed the formation energy of charge states of $V_{Al}$. In Figure Figure 6.3, only the lowest energy charge state is shown, as labeled. The figure shows that the formation energy of the nitrogen vacancies and aluminum vacancies have opposite slopes as a function of the growth conditions, as expected based on thermodynamics arguments, and that the slope of each should vary as the partial derivative of its chemical potential with respect to the growth conditions.

Figure 6.3: The formation energies $E_f$ of $V_{Al}$ and $V_N$ defects are shown as a function of the growth conditions $\Delta \mu$ ranging from N-rich to Al-rich. In this figure, only the lowest energy charge state of each defect is shown. Each defect is shown at three different fixed values of the Fermi level: (a) $E_F = 0$, the valence band maximum, (b) $E_F = E_g/2$, at the mid-gap, and (c) $E_F = E_g$, at the conduction band minimum.

The second type of formation energy diagram can also be used to compare defects of different types. Figure Figure 6.4 shows the formation energy of the two vacancies as a function of Fermi level.

This figure is similar to Figure 6.2, where the formation energies of several $V_{Al}$ charge states were shown, except that in this case only the most favorable (lowest energy) charge state of
Figure 6.4: The formation energies $E_f$ of $V_{\text{Al}}$ and $V_{\text{N}}$ defects are shown as a function of the Fermi level $E_F$ ranging from the valence band maximum ($E_F = 0$) to the conduction band minimum ($E_F = E_g$). In this figure, only the lowest energy charge state of each defect is shown. Each defect is shown at three different fixed growth conditions: (a) $\Delta \mu = +\Delta H$, N-rich conditions, (b) $\Delta \mu = 0$, halfway between the extremes, and (c) $\Delta \mu = -\Delta H$, Al-rich conditions.
each defect is shown. These figures can be used to show how the favorability of each charge state changes as a function of the Fermi level, and that these favorabilities are different for different growth conditions. In general, defects that require a species to leave the host are more favorable when the prevalence of that species in the environment is poor. Or, in other words, when the environment is enriched with one species, that species is less likely to leave the lattice.

The formation energies of defects and the Fermi level are clearly linked: a change in the Fermi level will change the formation energy of any charged defect, and therefore will change its concentration in the lattice, as shown in Section 4.8. What is not yet apparent is that the addition of species (a change in concentration of defects) will have an influence on the Fermi level at a finite temperature. This concept is central to much of the analysis reported in this dissertation, and is discussed more fully in Section 6.1.6.

6.1.3 Thermodynamic Transitions

The exact value of the transitions between favorable charge states can be found by finding the value $E_I$ of the Fermi level $E_F$ required to set the formation energy $E_f$ of two charge states $q_1$ and $q_2$ equal. For a generic defect $X$, the transition between $q_1$ and $q_2$, denoted $\epsilon(q_1|q_2)$ is found via:

$$E_f[X^{q_1}](\Delta \mu, E_I) = E_f[X^{q_2}](\Delta \mu, E_I)$$

$$0 = E_f[X^{q_1}](\Delta \mu, E_I) - E_f[X^{q_2}](\Delta \mu, E_I)$$

$$0 = \left( E^{DFT}[X^{q_1}] + \sum_s n_s \mu_s \right) + q_1 (VBM + E_I + \Delta V_1) - \left( E^{DFT}_{bulk} + \sum_d n_d \mu_d \right)$$

$$- \left( E^{DFT}[X^{q_2}] + \sum_s n_s \mu_s \right) + q_2 (VBM + E_I + \Delta V_2) - \left( E^{DFT}_{bulk} + \sum_d n_d \mu_d \right)$$

$$0 = \left( E^{DFT}[X^{q_1}] + \sum_s n_s \mu_s \right) + q_1 (VBM + \Delta V_1) - \left( E^{DFT}_{bulk} + \sum_d n_d \mu_d \right)$$

$$- \left( E^{DFT}[X^{q_2}] + \sum_s n_s \mu_s \right) + q_2 (VBM + \Delta V_2) - \left( E^{DFT}_{bulk} + \sum_d n_d \mu_d \right)$$

$$E_f[X^{q_1}]_{E_F=0} - E_f[X^{q_2}]_{E_F=0} + (q_1 - q_2) E_I$$

$$0 = E_f[X^{q_1}]_{E_F=0} - E_f[X^{q_2}]_{E_F=0} + (q_1 - q_2) E_I$$

$$E_I = \frac{E_f[X^{q_1}]_{E_F=0} - E_f[X^{q_2}]_{E_F=0}}{q_2 - q_1}$$

(6.5)
In the specific case where \( q_1 > q_2 \), \( E_I \) gives us the acceptor ionization energy of the defect \( q_1 \). For instance, if \( q_1 = 0 \), and \( q_2 = -1 \), then \( E_I = \epsilon(0|1) = E_{A}^{q=0} \), the acceptor level of the neutral defect. If, instead, \( q_1 < q_2 \), as in the case where \( q_1 = 0 \) and \( q_2 = 1 \), then \( E_I = \epsilon(0|1) = E_{D}^{q=0} \) is the donor level for the neutral defect. It should be apparent that \( \epsilon(0|1) = \epsilon(1|0) \) is also the acceptor level of the positive defect \( E_{A}^{q=1} \). The acceptor level \( E_A \), referenced to the valence band maximum, is the acceptor ionization energy \( E_a \). The donor ionization energy, that is, the energy difference between the donor level \( E_D \) and the conduction band minimum, can be expressed as \( E_d = E_q - E_D \), when referenced to the valence band maximum. Therefore, in general for a defect in charge state \( q \):

\[
E_A = \epsilon(q|q - 1) \quad \text{(6.6)}
\]
\[
E_D = \epsilon(q|q + 1) \quad \text{(6.7)}
\]
\[
E_a = E_A \quad \text{(6.8)}
\]
\[
E_d = E_q - E_D \quad \text{(6.9)}
\]

These defect levels are shown in Figure 6.5 for the case of the aluminum vacancy \( V_{Al} \), based on the formation energy diagram shown in Figure 6.2. To illustrate the principle, only the \( q = 1, 0, -1 \) charge states are shown, but it should be noted that other charge states are critical to describing the complete behavior of the defect.

---

**Figure 6.5:** Band diagram showing the transitions for the aluminum vacancy in AlN. The transitions between the +1,0, and -1 charge states are shown. The arrows indicate the acceptor and donor activation energies for the \( q = 0 \) defect.
6.1.4 Optical Transitions

In addition to determining the thermodynamic transitions between charge states, the formation energies of a defect can be used to predict optical transitions, such as photon absorption and photon emission. These can be directly compared to experimental measurements of transmission and luminescence. The starting point for these optical transitions are the thermodynamic transitions in Section 6.1.3. In general, optical transitions are possible when two charge states have a thermodynamic transition. In the simplest case of optical transitions, it is useful to think of a specific example, such as the Aluminum vacancy $V_{Al}$. Inspection of Figure 6.2 shows that for values of the Fermi level close to the conduction band (n-type material), the defect is favorable in the $q = -3$ charge state.

Suppose that this defect undergoes an optical transition. The charged defect then may absorb a photon to promote an electron from its defect level in the gap to the conduction band. It is often helpful to visualize these transitions with the help of a band-gap diagram, as shown in Figure 6.6a. By inspecting the crossings of the lines representing defect formation energies in Figure 6.2, it is clear that there is a thermodynamic transition between the $q = -3$
charge state and the \( q = -2 \) charge state where the two lines cross, near the middle of the band gap. Calculation of this transition proceeds by applying Eq. 6.5, and the transition level \( E_I(-3| -2) = 3.08 \) eV. Conceptually, this means that for values of the Fermi level below this transition level, the \( q = -2 \) charge state is favorable compared to the \( q = -3 \) state, and for values of the Fermi level higher than this \( E_I \), the opposite is true. This level is typically thought of as the ionization energy of the defect, that is, when the Fermi level is high enough to support the presence of the \( q = -3 \) state, the highest occupied energy state is situated \( E_I \) above the valence band, and a distance of \( E_g - E_I \) below the conduction band.

The process of a simple photo-absorption/photo-emission process would proceed as follows. The charged defect is in charge state \( q = -3 \) with its highest electron sitting at the defect level corresponding to \( E_I \) in the gap, as shown by column A of Figure 6.6a, at the far left. The absorption of a photon promotes an electron to the conduction band, as shown part B of Figure 6.6a. Note that according to the Franck-Condon principle [Stoneham, 1975], an optical transition will occur much more quickly than a lattice relaxation. The defect is now in the \( q = -2 \) charge state, with an electron in the conduction band, but remains in its equilibrium geometry determined when \( q = -3 \). For most defects, the geometries of different charge states will be slightly different. A lattice relaxation process occurs, while the charge state remains in the \( q = -2 \) condition, the lattice relaxes to the equilibrium geometry of \( q = -2 \), part C of Figure 6.6a. This is an excited state – the system’s ground state configuration is such that \( V_{Al} \) prefers the \( q = -3 \) state. At some later time, the electron falls from the conduction band to the defect level. This process emits a photon. Again, this process is faster than the corresponding lattice relaxation, so the defect has returned to its \( q = -3 \) charge state, but remains in the \( q = -2 \) geometric configuration, shown in part D of Figure 6.6a. A lattice relaxation will occur, bringing the geometry to the \( q = -3 \) geometry, and the system has now returned to its original state. The photon that is absorbed, then, must have the requisite energy to promote an electron from the defect band to the conduction band, without a lattice relaxation. Similarly, the photon that is emitted must have the energy corresponding from an electron falling from the conduction band to the defect level, with no lattice relaxation.

The magnitudes of these energies are presented in a configuration-coordinate diagram, as shown in Figure 6.6b. The configuration coordinate diagram shows two curves, one representing the \( q = -3 \) ground state and one representing the excited state where \( q = -2 \) with an additional electron in the conduction band. The alphabetic labels in Figure 6.6b correspond to the same labels in Figure 6.6a. The photon absorption, then, is represented by the distance between A and B, while the photon emission is represented by the distance between C and D. The lattice relaxations in the excited state occur between A and B, and in the ground state between C and A. Each of these relaxation energies is a Franck-Condon shift, and summed together they are the Stokes shift for the absorption and emission process.
In order to construct the configuration coordinate diagram, point A provides a logical starting place. It has been stated that point A in Figure 6.6b is the minimum energy configuration of the $V_{Al}^{-3}$ defect, which was calculated using the formation energy equation (Eq. 6.4) based on the energy of a DFT relaxation calculation. Notice that point A is at the minimum of the curve corresponding to the $q = -3$ charge state. The upper curve in the diagram corresponds to the $q = -2$ charge state with an electron in the conduction band, and the minimum of the curve will represent the relaxed configuration of the $V_{Al}^{-2}$ defect with an electron in the conduction band. In order to represent this using ground-state energies obtained with DFT, we define this minimum with respect to the minimum of $V_{Al}^{-3}$ (point A) according to the definition of the thermodynamic transition level $E_I$ between the two states. The energy of the equilibrium $V_{Al}^{-2}$ lies at an energy $E_I$ below $V_{Al}^{-3}$. Adding an electron to the conduction band of $V_{Al}^{-2}$ corresponds to adding the energy of the band gap. Therefore, the energy difference between the $V_{Al}^{-3}$ and the $V_{Al}^{-2} + e^-$ is equal to $E_g - E_I$, which is the vertical distance (on the energy scale) between point C and point A. Point B corresponds to the energy of the aluminum vacancy in the -3 geometry, but with a charge state of -2, and an electron in the conduction band. Calculating the energy of this point requires an additional DFT calculation. Using the final relaxed geometry from the $V_{Al}^{-3}$, the number of electrons (NELECT) in the INCAR file is reduced by one, and a single-point energy (no relaxation) calculation is performed. The defect formation energy of this configuration is calculated as normal according to Eq. 6.4. This calculated defect formation energy $E'_{f}$ [$V_{Al}^{-2}$] in the non-equilibrium geometry does not yet include the electron in the conduction band. The contribution to the total formation energy of the addition of an electron in the conduction band is equivalent to the energy to move an electron from the valence band to the conduction band, which is defined as the band gap $E_g$. The final energy of point B, then is the energy of the $q = -2$ defect in the $V_{Al}^{-3}$ geometry, plus the value of $E_g$. Similarly, the energy of point C is the energy of the equilibrium $V_{Al}^{-2}$ (which has been calculated for), plus an additional $E_g$ to account for the electron in the conduction band. Point D represents the geometric configuration of the $V_{Al}^{-2}$ defect in the $q = -3$ charge state. Similar to the manner in which the energy for point B was calculated, the final relaxed geometry of the $V_{Al}^{-2}$ calculation is used for a single-point energy DFT calculation where the number of electrons has been increased by 1. These four DFT calculations – the original two relaxations plus two additional single-point energies – are the only DFT inputs needed to construct this simple configuration coordinate diagram and determine the optical transitions and lattice relaxation energies.

6.1.5 Law of Mass Action

Some of the plots presented display a negative formation energy for a defect. See, for instance, Figure 6.2. If the system were truly at a value of the Fermi level and chemical environment where
the energy of the Al vacancy is negative, this would imply that the evaluated concentration of vacancies based on (4.62) would be greater than the number of aluminum sites present. Thermodynamically, this means that the enthalpy of formation is negative and therefore the formation would be favorable even at 0K. At finite temperature, the configuration entropy term would drive the reaction forward further, and the material would be expected to decompose. This is not seen in real materials, because in practice, this combination of growth conditions and Fermi level are not accessible. The reason for this non-accessibility arises because of the laws of mass action, which state that in the closed system charge and matter must be conserved. This concept, and its implications on the system of defects is explored in more detail in Section 6.1.6. Several of the results presented in this dissertation rely on using the mass balance equations to determine the equilibrium Fermi level in the system at finite temperature.

6.1.6 Equilibrium Concentrations of Defects

In Section 4.8, Eq. 4.62 gave the equilibrium concentration \( C_i \) of defect \( i \) as a function of the enthalpy of formation of the defect. It was also shown in Eq. 6.4 that the enthalpy of formation of the defect depends on the chemical environment (growth conditions) and the Fermi level. It was mentioned briefly in Chapter 6 that the incorporation of a defect would affect the prevalence of other defects by virtue of changes in the Fermi level. This concept is explored further in this chapter, and a system by which the equilibrium Fermi level, and therefore the concentration of defects in the system can be determined.

The law of mass action states that in order to preserve the charge neutrality of the system, the creation of electrons must be balanced by the creation of holes. The electrons and holes can either come from thermally ionized carriers in the valence band and conduction band, or can come from thermally ionized donors and acceptors. The mass balance equation for a semiconductor can be written

\[
n + N_{a}^- = p + N_{d}^+,
\]

where \( n \) and \( p \) are the number of electrons and holes thermally ionized from the bands, and \( N_{a}^- \) and \( N_{d}^+ \) are the number of thermally ionized acceptors and donors. Zhang and Northrup assumed a fixed value for the number of ionized donors in an n-type material and used this value to determine the equilibrium Fermi level and defect concentrations. [Zhang and Northrup, 1991]

The solution presented here calculates the number of ionized defects explicitly. These terms of the mass balance equation depend on the Fermi level and the temperature, as well as on the total concentrations of defects \( N_{a,i} \) and \( N_{d,j} \), and the ionization energies of the acceptors \( E_{a,i} \).
and donors $E_{d,j}$.

$$n = N_C \cdot \exp \left( \frac{E_F - E_g}{k_B T} \right)$$

$$p = N_V \cdot \exp \left( \frac{-E_F}{k_B T} \right)$$

$$N_a^- = \sum_i \frac{N_{a,i}}{1 + 4 \exp(\frac{E_{a,i} - E_F}{k_B T})}$$

$$N_d^+ = \sum_j \frac{N_{d,j}}{1 + \frac{1}{2} \exp(\frac{-(E_{d,j} - E_F)}{k_B T})}$$

In the expressions above, $N_C$ and $N_V$ are the density of states in the conduction and valence bands. Both are material-specific properties with a temperature dependence of $\propto T^{3/2}$. For a given temperature $T$ and for each value of the growth conditions $\Delta \mu$, the task is to find the value of the Fermi level $E_F$ which satisfies the mass balance equation. The Fermi level enters the equation both in the exponential terms above, as well as in the defect formation energies that determine $N_{a,i}$ and $N_{d,j}$. The solution to the complete mass balance equation relies on finding the value of the Fermi level that satisfies

$$0 = N_C \cdot \exp \left( \frac{E_F - E_g}{k_B T} \right) + \sum_i \frac{N_{s,i} \cdot \exp \left( \frac{-(E_{DFT,i} - E_{bulk} + \sum_k n_k \mu_k + q(E_F + \text{VBM} + \Delta V_i))}{k_B T} \right)}{1 + 4 \exp(\frac{E_{a,i} - E_F}{k_B T})} - N_V \cdot \exp \left( \frac{-E_F}{k_B T} \right) - \sum_j \frac{N_{s,j} \cdot \exp \left( \frac{-(E_{DFT,j} - E_{bulk} + \sum_k n_k \mu_k + q(E_F + \text{VBM} + \Delta V_j))}{k_B T} \right)}{1 + \frac{1}{2} \exp(\frac{-(E_{d,j} - E_F)}{k_B T})}$$

(6.15)

Solving for the Fermi level proceeds iteratively, by solving the equation for each value of the growth conditions $\Delta \mu$ separately. For each value of $\Delta \mu$, the Fermi level is varied in discrete steps from $E_F = 0$ to maximum $E_F = E_g$ until the value of $E_F$ which satisfies the mass balance equation is found. This $E_F$ is recorded, the value of $\Delta \mu$ is iterated, and the solution progresses. The solution proceeds in two parts, first at high temperature $T_{hi}$ which represents the growth temperature, when the defects are first incorporated. A rapid quench is assumed, such that defects stay on the sites at which they incorporated, but electronic configurations are allowed to relax. Based on this low-temperature condition, and with fixed defect concentrations based on the high temperature solution, the mass balance equations are solved once again, and the low temperature equilibrium Fermi level is determined.
6.2 Surfaces

The goal of DFT calculations of surfaces is typically to determine which surfaces will be most favorable under a given set of circumstances. This favorability is determined by comparing their surface energies. The surface energetics of a given material can depend on the surface orientation, termination, surface relaxations, reconstructions, and the chemical environment. This section will discuss some of the details of analyzing the output of DFT calculations on surfaces to determine surface energies as a function of temperature and pressure, and to investigate the competition between compensation mechanisms that arise at the surfaces of polar materials.

6.2.1 Surface Free Energy Calculations

The expression for the surface free energy was presented earlier in Eq. 4.27, reproduced here for convenience:

\[
\Delta G^{\text{total}} = \frac{1}{A} \left( E^{\text{total}} - n_{\text{Mg}} g_{\text{MgO}}^{\text{bulk}} + (n_{\text{Mg}} - n_{\text{O}}) \mu_{\text{O}}(T,p) + \sum \frac{1}{2} \hbar \omega \right)
\]

The total energy \( E \) is obtained directly from the DFT calculation, just as \( g_{\text{MgO}}^{\text{bulk}} \) is obtained from the bulk calculation, normalized to be the energy per formula unit. Remaining to be calculated are the vibrational modes \( \omega \) and the chemical potential as a function of \( T \) and \( p \). The temperature and pressure dependent chemical potentials are discussed in Section 6.2.2.

The vibrational modes \( \omega \) are calculated by a secondary DFT calculation. For the surfaces discussed in this dissertation, it was determined that the atoms of the terminated bulk crystal had vibrational modes similar to modes in the bulk, so the contributions of these atoms to the vibrational free energy cancel by subtraction. The modes of surfactant species such as H\(_2\)O or OH, however, are significant, in agreement with the findings of Sun, Reuter, and Scheffler for surface of RuO\(_2\). [Sun et al., 2003] As an example consider the OH-capped MgO \{001\} surface. The surface of the unit cell consists of a plane containing an oxygen and a magnesium, with one hydrogen situated atop each surface oxygen, and a hydroxyl group sitting above the surface between Mg neighbors in the [110] direction. The bonds that are likely to contribute to the vibrational modes are the O-H bonds, and the bonds between the H-capped oxygens and their neighboring Mg atoms. Once the surface has relaxed to its low energy configuration, a second DFT calculation is run using the relaxed coordinates as a starting point for a set of finite difference calculations. In this calculation, the atoms that are not of interest are fixed, while the atoms whose vibrational modes are to be calculated are allowed to relax using the Selective Dynamics tag in the VASP POSCAR file. The finite difference calculation is specified in the INCAR file, using IBRION=5. During this calculation, each atomic step consists of displacing each ion, one at a time, by a finite distance and calculating the total energy self consistently. The Hessian
matrix, which is the matrix of second derivatives of energy with respect to the atomic positions, is determined based on the change in forces between atoms for each of the perturbations. The final vibrational modes $\omega$ are reported in the OUTCAR file and can be summed according to Eq. 4.27.

### 6.2.2 Extending DFT results with thermochemical tables

The final component to calculating the temperature and pressure dependent surface energies is to extend the chemical potential by use of the thermochemical tables. Combining Eq. 4.28 and Eq. 4.29 yields the expression of the chemical potential of O$_2$ at any temperature and pressure, based on the ideal gas approximation:

$$
\mu_O(T, p) = \frac{1}{2} [H(T, p^R, O_2) - H(0K, p^R, O_2)] - \frac{1}{2} T [S(T, p^R, O_2) - S(0K, p^R, O_2)] + \frac{1}{2} kT \ln \left( \frac{p}{p^R} \right) \quad (6.16)
$$

A standard thermochemical table is shown for molecular O$_2$ in Table 6.1. The table gives a number of thermodynamic values, but the columns of interest at present are the entropy ($S^R$) and the enthalpy $H - H^R (T_r)$. Obtaining the necessary entropy term ($(S(T, p^R, O_2) - S(0K, p^R, O_2))$) is straightforward because $S(0K) = 0$. The equation requires $H(T) - H(0K)$ whereas the table gives values for $H(T) - H(298.15K)$. The value of $H(0) - H(T_r)$ is subtracted from each of the values in the column $H(T) - H(T_r)$ to satisfy the equation. This is shown in the left columns of Table 6.2. These values are then converted into $T\Delta S$ and $H$ in units of eV, as shown in the right columns of Table 6.1. It is useful in computing the temperature dependence to find a line of best fit for the two terms as a function of $T$, though care must be taken to ensure that the data support a linear fit for the species of choice. These parameters allow the temperature and pressure dependence to be defined as:

$$
\Delta H(T) = m_H T + b_H \quad (6.17)
$$

$$
T\Delta S(T) = m_S T + b_H \quad (6.18)
$$

and since our fitted values for $m_H$, $b_H$, $m_S$, and $b_S$ were based on thermochemical data which had reference pressure $p^R = 0.1MPa$, Eq. 6.16 can be rewritten

$$
\mu_O(T, p) = \frac{1}{2} (m_H T + b_H) - \frac{1}{2} (m_S T + b_H) + \frac{1}{2} kT \ln \left( \frac{p}{0.1MPa} \right) . \quad (6.19)
$$
Table 6.1: The standard thermochemical table values obtained from a reference such as the JANAF tables, shown here for gas-phase molecular O\textsubscript{2}. [Stull and Prophet, 1971]

<table>
<thead>
<tr>
<th>T/K</th>
<th>Cp\textsuperscript{o}</th>
<th>S\textsuperscript{o}</th>
<th>-[G\textsuperscript{o} - H\textsuperscript{o} (T\textsubscript{r})] /T</th>
<th>H - H\textsuperscript{o} (T\textsubscript{r})</th>
<th>fH</th>
<th>fG</th>
<th>log Kf</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>INFINITE</td>
<td>-8.683</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>29.106</td>
<td>173.307</td>
<td>231.094</td>
<td>-5.779</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200</td>
<td>29.126</td>
<td>193.485</td>
<td>207.823</td>
<td>-2.868</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>250</td>
<td>29.201</td>
<td>199.990</td>
<td>205.630</td>
<td>-1.410</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>298.15</td>
<td>29.376</td>
<td>205.147</td>
<td>205.147</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>300</td>
<td>29.385</td>
<td>205.329</td>
<td>205.148</td>
<td>0.054</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>350</td>
<td>29.694</td>
<td>209.880</td>
<td>205.506</td>
<td>1.531</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>400</td>
<td>30.106</td>
<td>213.871</td>
<td>206.308</td>
<td>3.025</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>450</td>
<td>30.584</td>
<td>217.445</td>
<td>207.350</td>
<td>4.543</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>500</td>
<td>31.091</td>
<td>220.693</td>
<td>208.524</td>
<td>6.084</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.2: The thermochemical data in the JANAF table of O\textsubscript{2} has been converted to a form that allows simple calculation of a line of best fit for the temperature dependence. [Stull and Prophet, 1971]

<table>
<thead>
<tr>
<th>T (K)</th>
<th>S(T) - S(0K)</th>
<th>H(T) - H(0K)</th>
<th>T \cdot (S(T) - S(0K))</th>
<th>H(T) - H(0K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>100</td>
<td>173.307</td>
<td>2.904</td>
<td>0.180</td>
<td>0.030</td>
</tr>
<tr>
<td>200</td>
<td>193.485</td>
<td>5.815</td>
<td>0.401</td>
<td>0.060</td>
</tr>
<tr>
<td>250</td>
<td>199.990</td>
<td>7.273</td>
<td>0.518</td>
<td>0.075</td>
</tr>
<tr>
<td>298.15</td>
<td>205.147</td>
<td>8.683</td>
<td>0.634</td>
<td>0.090</td>
</tr>
<tr>
<td>300</td>
<td>205.329</td>
<td>8.737</td>
<td>0.638</td>
<td>0.091</td>
</tr>
<tr>
<td>350</td>
<td>209.880</td>
<td>10.214</td>
<td>0.761</td>
<td>0.106</td>
</tr>
<tr>
<td>400</td>
<td>213.871</td>
<td>11.708</td>
<td>0.886</td>
<td>0.121</td>
</tr>
<tr>
<td>450</td>
<td>217.445</td>
<td>13.226</td>
<td>1.014</td>
<td>0.137</td>
</tr>
<tr>
<td>500</td>
<td>220.693</td>
<td>14.767</td>
<td>1.143</td>
<td>0.153</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m</th>
<th>m\textsubscript{S}</th>
<th>m\textsubscript{H}</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b\textsubscript{S}</td>
<td>b\textsubscript{H}</td>
</tr>
</tbody>
</table>
6.2.3 Bader Charge Analysis

The real-space charge density is solved during the course of a DFT calculation. This density can be used to provide information about the bonding character in the material. Bader proposed a method of specifying the charge associated with each atom in the system by delineating the boundaries between atoms based on zero flux surfaces of the charge density.[Bader, 1985] After an ionic relaxation is completed with VASP, a second calculation is performed which doubles the density of real-space sampling points in each Cartesian coordinate direction and the resulting calculated charge density is post-processed to determine the Bader charge of each atom.[Tang et al., 2009, Henkelman et al., 2006, Sanville et al., 2007] This post-processing step utilizes the tools developed by Henkelman. First, the charge density of the cores can be included by specifying LAECHG=.TRUE. in the INCAR file, which creates two new files after the run AECCAR0 and AECCAR2. These files can be summed using the chgsum.pl routine which takes the two filenames as input. The chgsum.pl routine produces a summed file CHGCAR_sum which is used in the calculation of the Bader charges: bader CHGCAR -ref CHGCAR_sum -vac auto, where -vac auto is appropriate for calculations with surfaces. This Bader analysis produces a file called ACF.dat which contains the charge of each atom, listed as the total number of electrons associated with the atom. This Bader-counted number of electrons can be compared to the number of electrons expected to determine the effective Bader charge of the atom.

**Bond Valence**

In addition to Bader analysis, the changes to local charges on atoms in the system can also be measured by using bond-valence parameters.[Brese and O’Keeffe, 1991] This method does not use the charge density, relying only on the final positions of the atoms in the defect structure compared to their positions in the bulk material. This method defines the valence $V_i$ of an atom as the sum of the valences of its bonds with other atoms, according to the equation

$$V_i = \sum_j v_{ij}.$$  \hspace{1cm} (6.20)

The equation sums the bond valence $v_{ij}$ of the bonds between atoms $i$ and each of its neighbors $j$. The bond valence, in turn, depends on the bond length $d_{ij}$ according to

$$v_{ij} = \exp((R_{ij} - d_{ij})/b),$$  \hspace{1cm} (6.21)

where $b$ is a universal constant and $R_{ij}$ is the bond-order parameter for the $i$-$j$ bond. The bond-order parameter is, in turn, calculated based on the DFT results from the bulk material. Calculating the bond order-parameter from a DFT calculation involves first defining the for-
Table 6.3: Calculated bond-valence parameters for MgO and CaO bonds. The calculated DFT parameters are compared to those given in Ref. [Brese and O’Keeffe, 1991].

<table>
<thead>
<tr>
<th>Bond</th>
<th>$R_{ij}$ (DFT)</th>
<th>$R_{ij}$ ([Brese and O’Keeffe, 1991])</th>
<th>$d_{ij}$ (DFT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mg-O</td>
<td>1.70</td>
<td>1.69</td>
<td>2.11</td>
</tr>
<tr>
<td>Mg-O</td>
<td>1.98</td>
<td>1.96</td>
<td>2.39</td>
</tr>
<tr>
<td>O-H</td>
<td>1.98</td>
<td>1.96</td>
<td>2.39</td>
</tr>
</tbody>
</table>

The total valence $V_i$ of each atom, then determining the bond valences $v_{ij}$. If the bulk is a binary compound, the neighboring species will all be of the same type, so the bond valence is simply the total valence divided by the number of bonds. Then, the distances between atoms in the bulk as calculated by DFT are measured, and the bond-order parameter is solved. Shown in Table 6.3 are the calculated bulk bond-order parameters as computed from DFT calculations, compared to previous results. The bond-order parameter calculated for the bulk bond is then used to calculate the valence of surface bonds, and in turn the valence of the surface atoms, which differ from bulk atoms based on both under-coordination (cleaved bonds) and structural bond-length changes.
Chapter 7

The source of the 4.7 eV absorption band in AlN

This chapter presents an extended discussion of my DFT calculations on the electronic, thermodynamic, and optical properties of point defects in AlN. My calculations were undertaken to determine the origin of the 4.7 eV optical absorption band that has plagued bulk AlN grown by physical vapor transport (PVT) for the past 20 years.


I would like to acknowledge the contributions of my experimental collaborators who provided experimental data and helpful discussions in guiding my calculations: Ramón Collazo1, Jinqiao Xie2, Zachary Bryan1, Ronny Kirste1, Marc Hoffmann1, Rafael Dalmau2, Baxter Moody2, Yoshinao Kumagai3, Toru Nagashima4, Yuki Kubota4, Toru Kinoshita4, Akinori Koukitu3, and Zlatko Sitar1.

1Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, USA
2HexaTech, Inc., 991 Aviation Pkwy., Suite 800, Morrisville, North Carolina 27560, USA
3Department of Applied Chemistry, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan
4Tsukuba Research Laboratories, Tokuyama Corporation, Tsukuba, Ibaraki 300-4247, Japan
7.1 Background

Pseudomorphic AlN substrates allow for the epitaxial growth of low strain AlGaN overlayers, which can be used for high-energy optics and power applications. For deep UV optics applications, this substrate must be optically transparent. It has been well-known for the past 20 years that bulk AlN has been plagued by a prominent (intensity greater than 1000 cm\(^{-1}\)) deep-UV absorption (4.7 eV). [Strassburg et al., 2004, Bickermann et al., 2006] In order to improve the extraction efficiency of optical devices grown on these substrates, this absorption must be suppressed.

The source of the absorption remained unknown, but experimental SIMS data of a non-transparent PVT-grown sample measured by our collaborators at NCSU showed elevated concentrations of carbon ([C] > 2 \times 10^{19}) and moderately high concentrations of silicon and oxygen ([Si] \approx [O] \approx 5 \times 10^{18}).

7.2 Computational Methodology

DFT calculations using hybrid functionals as described in earlier chapters were used to determine the formation energies, thermodynamic transitions, and optical signatures of a host of native and impurity point defects in bulk AlN, in order to determine the cause of the 4.7 eV absorption band.

7.2.1 The “Band Gap Problem” in AlN

As mentioned in previous sections, DFT suffers from the so-called “band-gap problem”, in that it often underpredicts the electronic band gap of a material. This underprediction in AlN is due to an incorrect placement of the conduction band levels. This limitation is particularly problematic in defect calculations, because the relative position of the defect level within the band gap may also be incorrectly predicted. Table 7.1 shows the calculated lattice parameters, cohesive energies, and band-gaps as calculated with five different approximations to the density functional. The structural parameters as calculated with DFT agree with experiment to within 1%, but the electronic band gap is underestimated by approximately 30%.

7.2.2 Hybrid DFT

Hybrid DFT, where a fraction of short-range exact exchange is included has been shown to better reproduce the experimental band gap. In its initial formulation, HSE06 [Heyd et al., 2003, Heyd et al., 2006] chose 25% exact exchange based on the prior PBE0 functional. Adjusting the fraction of exact exchange changes both the relative position of the conduction band and the
Table 7.1: Lattice parameters, cohesive energies, and band gaps as calculated using five standard density functional approximations (DFAs). Experimental lattice parameters are taken from [Iwama et al., 1982] and experimental band gap is taken from [Dalmau et al., 2010].

<table>
<thead>
<tr>
<th>Functional</th>
<th>a (Å)</th>
<th>c/a</th>
<th>u</th>
<th>$E_{coh}$ (eV/pair)</th>
<th>$E_g$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE</td>
<td>3.13</td>
<td>1.60</td>
<td>0.38</td>
<td>-14.89</td>
<td>4.05</td>
</tr>
<tr>
<td>PBEsol</td>
<td>3.11</td>
<td>1.60</td>
<td>0.38</td>
<td>-14.72</td>
<td>4.11</td>
</tr>
<tr>
<td>PW91</td>
<td>3.13</td>
<td>1.60</td>
<td>0.38</td>
<td>-15.00</td>
<td>4.09</td>
</tr>
<tr>
<td>AM05</td>
<td>3.11</td>
<td>1.60</td>
<td>0.38</td>
<td>-14.55</td>
<td>4.17</td>
</tr>
<tr>
<td>LDA</td>
<td>3.09</td>
<td>1.60</td>
<td>0.38</td>
<td>-16.43</td>
<td>4.39</td>
</tr>
<tr>
<td>Exper.</td>
<td>3.11</td>
<td>1.60</td>
<td></td>
<td></td>
<td>6.09</td>
</tr>
</tbody>
</table>

valence band, and yields a linear increase in the band gap. This is shown for AlN in Figure 7.1. Based on these tests, it as determined that the fraction of exact exchange which best predicted the band gap was 32%. The predicted band gap of 6.09 eV and lattice constants $a=3.11$ Å, $c=4.96$ Å, and $u=0.382$, are in good agreement with experimental measurements. [Iwama et al., 1982, Dalmau et al., 2010] A comparison between the parameters of standard DFT, HSE-.32, and experiment are listed in Table 7.2.

7.3 Defects Studied

The experimental SIMS data showed a significant concentration of C, Si, and O. These impurities, as well as vacancies, were studied on both Al and N sites, in a number of different charge states. A table showing the full set of 180 defects studied is shown in Table 7.3.

7.4 Results

Over 180 defect calculations were analyzed to determine whether any defects may be responsible for the 4.7 eV absorption. The substitutional carbon acceptor (C$_N$) was determined to be a likely candidate based on its favorable formation, deep acceptor level, and calculated optical properties. The calculated formation energy diagrams are presented in Figure 7.2, where the formation energy for the substitutional carbon defects and the vacancies are shown in the N-rich and Al-rich limits. The other defects investigated (Si and O) did not have thermodynamic transitions that would lead to a 4.7 eV absorption. The results indicate that in the Al-rich growth environment typical of PVT, carbon readily incorporates on the nitrogen site. Furthermore, for
Figure 7.1: Change in the conduction band, valence band, and band-gap for GaN as a function of the amount of exact exchange.

Table 7.2: The calculated lattice parameters, formation enthalpy, and band gaps according to a standard GGA density functional (PBE) and the HSE functional with 32% exact exchange, compared with experiment. Experimental lattice parameters are taken from [Iwama et al., 1982] and experimental band gap is taken from [Dalmau et al., 2010].

<table>
<thead>
<tr>
<th></th>
<th>PBE</th>
<th>HSE 32</th>
<th>Exper.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (Å)</td>
<td>3.13</td>
<td>3.10</td>
<td>3.11</td>
</tr>
<tr>
<td>c/a</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
</tr>
<tr>
<td>c (Å)</td>
<td>5.02</td>
<td>4.96</td>
<td>4.98</td>
</tr>
<tr>
<td>u</td>
<td>0.38</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Eg (eV)</td>
<td>4.05</td>
<td>6.09</td>
<td>6.09</td>
</tr>
<tr>
<td>Hf (eV/pair)</td>
<td>-2.90</td>
<td>-3.21</td>
<td>-3.30</td>
</tr>
</tbody>
</table>
Table 7.3: Summary of the isolated native and impurity defects and defect complexes calculated to determine the origin of the 4.7 eV absorption. 180 calculations were performed for the defects listed in this table.

<table>
<thead>
<tr>
<th>Defect</th>
<th>Charge States</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_N</td>
<td>-4.4</td>
</tr>
<tr>
<td>C_Al</td>
<td>-4.4</td>
</tr>
<tr>
<td>Si_N</td>
<td>-4.4</td>
</tr>
<tr>
<td>Si_Al</td>
<td>-4.4</td>
</tr>
<tr>
<td>V_N</td>
<td>-4.4</td>
</tr>
<tr>
<td>V_Al</td>
<td>-4.4</td>
</tr>
<tr>
<td>O_N</td>
<td>-4.4</td>
</tr>
<tr>
<td>O_Al</td>
<td>-4.4</td>
</tr>
<tr>
<td>CD_X_N</td>
<td>-1.1</td>
</tr>
<tr>
<td>CD_X_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>OD_X_N</td>
<td>-1.1</td>
</tr>
<tr>
<td>OD_X_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>Si_DX_N</td>
<td>-1.1</td>
</tr>
<tr>
<td>Si_DX_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>C_N−C_Al</td>
<td>-4.1</td>
</tr>
<tr>
<td>C_N−Si_Al</td>
<td>-3.4</td>
</tr>
<tr>
<td>Si_N−C_Al</td>
<td>-1.4</td>
</tr>
<tr>
<td>C_N−1_V_Al</td>
<td>-5.1</td>
</tr>
<tr>
<td>C_N−2_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>C_N−3_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>Si_N−1_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>Si_N−2_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>Si_N−3_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>O_N−1_V_Al</td>
<td>-4.1</td>
</tr>
<tr>
<td>O_N−2_V_Al</td>
<td>-6.1</td>
</tr>
<tr>
<td>O_N−3_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>Si_Al−V_N</td>
<td>-3.1</td>
</tr>
<tr>
<td>Si_Al−O_N</td>
<td>1.3</td>
</tr>
<tr>
<td>C_Al−O_N</td>
<td>1.2</td>
</tr>
<tr>
<td>1_O_N−1_V_Al</td>
<td>-2.2</td>
</tr>
<tr>
<td>2_O_N−1_V_Al</td>
<td>-2.1</td>
</tr>
<tr>
<td>3_O_N−1_V_Al</td>
<td>-1.1</td>
</tr>
<tr>
<td>C_Al−C_N_vert</td>
<td>-2.2</td>
</tr>
<tr>
<td>Si_Al−Si_N</td>
<td>-2.2</td>
</tr>
<tr>
<td>Si_Al−Si_N_vert</td>
<td>-2.2</td>
</tr>
</tbody>
</table>
Figure 7.2: Formation energy for substitutional carbon defects for (a) nitrogen-rich and (b) aluminum-rich growth conditions.

A wide range of the Fermi level (from $2 \leq E_F \leq E_g$) the $C_N$ defect acts as an ionized acceptor. The optical transition of this $C_N^-$ defect was determined according to the procedure in Chapter 6. The corresponding configuration coordinate diagram is shown in Figure 7.6. As indicated by the diagram, the ionized carbon acceptor will absorb a 4.7 eV photon. This absorption is consistent with the experimental observation of the deep UV 265 nm (4.7 eV) absorption. The photon absorption is followed by a lattice relaxation which has a maximum energy reduction of 0.5 eV. This lattice relaxation is followed by a minimum optical emission of 3.5 eV, and a final lattice relaxation of 0.7 eV. The total Stokes’ shift of this optical process has a maximum value of 1.2 eV. These lattice relaxation energies are large because of the significant bond length changes between these defect charge states, as shown in Figure 7.3. In the negative charge state, the carbon atom is only slightly perturbed off of the nitrogen lattice site. The c-axis carbon-aluminum bond is 0.5% larger than the bulk bond distance. After the photoexcitation process, the neutral defect prefers a substantial breathing-mode relaxation where the a-axis bonds lengthen by 2.6% and the c-axis bond lengthens by an additional 8.5%. These large relaxations indicate that a de-excitation is likely to occur before full relaxation can take place. Thus, the observed emission is expected to be higher in energy than the predicted minimum 3.5 eV emission.

### 7.5 Summary of Experimental Confirmation

The theoretical evidence that carbon is the source of the 4.7 eV absorption was experimentally investigated by comparing the absorption spectrum of a sample grown by PVT (containing high
carbon concentration) to three carefully controlled carbon-doped AlN samples grown by hydride vapor phase epitaxy (HVPE). The PVT sample was grown by our collaborators at NCSU and HexaTech, Inc., and the HVPE samples were grown by our collaborators at Tokuyama Corp. The SIMS and optical measurements were taken by our collaborators at NCSU and Tokyo University of Agriculture and Technology. SIMS analysis of the HVPE samples showed carbon concentrations in the samples of $1 \times 10^{19} \text{cm}^{-3}$, $7 \times 10^{17} \text{cm}^{-3}$, and $< 2 \times 10^{17} \text{cm}^{-3}$.

The optical absorption of these four samples demonstrated a linear relationship between the 4.7 eV absorption intensity and the carbon concentration, as seen in Figure 7.4. The sample with the lowest carbon is effectively transparent in the deep UV range. The measured photoluminescence (PL) of these four samples is shown in Figure 7.5. The samples with high carbon concentrations ([C]$ \sim 10^{19} \text{cm}^{-3}$) have a prominent emission peak at 3.9 eV. This emission corresponds to the predicted minimum 3.5 eV emission, and is not present for samples with lower carbon concentrations.

### 7.6 Conclusions

A theoretical model was constructed based on DFT calculations of the formation energy and transitions of point defects in AlN. In the proposed model, the ionized carbon acceptor $\text{C}_N^-$ absorbs a photon at 4.7 eV and emits a photon with an energy greater than 3.5 eV. The good agreement between the theoretical predictions and the experimental measurements of both absorption and emission support the idea that the ionized carbon acceptor is responsible for the deep UV emission. The combined theoretical and experimental approach reveals for the first time that removal of carbon from the growth environment would yield optically transparent bulk AlN substrates.
Figure 7.4: Experimental absorption measurements of samples with varying carbon concentration. a) Absorption at 4.7 eV decreases with decreasing carbon concentration. b) A linear relationship exists between the magnitude of the absorption peak at 4.7 eV and carbon concentration. Figures are from the original manuscript. [Collazo et al., 2012]
Figure 7.5: Experimental photoluminescence measurements of samples with varying carbon concentration. High carbon samples show a prominent emission peak at 3.9 eV which corresponds to the carbon acceptor transition. This figure appears in the original publication. [Collazo et al., 2012]

Figure 7.6: Configuration coordinate diagram for the optical transition of carbon substitution on the nitrogen site. The energies indicated are in eV.
Chapter 8

Thermal activation energy of Si and O DX donors in Al$_x$Ga$_{1-x}$N

Benjamin E. Gaddy, Douglas L. Irving, Zachary A. Bryan, Isaac Bryan, Ronny Kirste, Ramón Collazo, Zlatko Sitar
Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, USA
Jinqiao Xie, Seiji Mita, Baxter Moody
HexaTech, Inc., 991 Aviation Pkwy., Suite 800, Morrisville, North Carolina 27560, USA

This article presents results on the computational and experimental investigation of the donor activation energies of Si and O in AlN. Experimental conductivity measurements were performed in the Collazo and Sitar groups. These results show a change in the donor activation energy and the defect compensation ratio as a function of Si content. My DFT calculations of the defect formation energies and electron capture barriers demonstrate that this behavior is due to the formation of deep DX centers, and the interaction between the O and Si impurities.

This article is being prepared for submission to *Applied Physics Letters*. 
8.1 Introduction

Alloys between GaN and AlN have become appealing candidates for high power electronic devices and deep-UV light emitting diodes and lasers because of their wide, tunable bandgap ranging from 3.5 eV in GaN to 6.09 in AlN. [Dalmau et al., 2010] Widespread use of $\text{Al}_x\text{Ga}_{1-x}\text{N}$-based devices requires reliable n-type and p-type doping. Controlling this doping has been a challenge because the wide bandgap which makes these alloys attractive also provides additional opportunities for electronic compensation by native and unintentional impurity defects. To achieve n-type $\text{Al}_x\text{Ga}_{1-x}\text{N}$, Si and O have been frequently tested as intentional donor species.

Wetzel, et al. found that oxygen was an effective donor in GaN, but that under pressure a deep level appeared. [Wetzel et al., 1997] The authors predicted that oxygen would undergo a similar transformation in AlGaN with increasing Al content above about $x=0.40$. They did not observe the same behavior for Si donors in GaN under pressure and thus predicted that Si could be used as a donor in higher Al content $\text{Al}_x\text{Ga}_{1-x}\text{N}$. McCluskey, et al. showed that for $x=.44$, the carrier concentration was approximately equal to the Si doping level. [McCluskey et al., 1998] However, in samples with higher Al content, the activation energy of the Si donor has been shown to increase with increasing Al content. Activation energies as high as 320 meV have been reported for pure AlN. [Stutzmann et al., 1997, Polyakov et al., 1998, Skierbiszewski et al., 1999b, Collazo et al., 2011]

This increase in activation energy has been linked to the appearance of the deep localized state of O and Si, due to the formation of a DX center [Wetzel et al., 1997, McCluskey et al., 1998, Skierbiszewski et al., 1999b, Skierbiszewski et al., 1999a, Goennenwein et al., 2001, Zeisel et al., 2000] similar to the one first proposed in $\text{Al}_x\text{Ga}_{1-x}\text{As}$. [Lang et al., 1979, Chadi and Chang, 1988] The DX center is a substitutional point defect which undergoes a lattice relaxation. [Lang, 1992] This relaxation may lead to metastable and stable atomic configurations that can capture electrons. First-principles theoretical investigations have proposed that oxygen substituting for nitrogen in the lattice forms a stable DX center in zinc-blende AlN, but not in zinc-blende GaN, [Mattila and Nieminen, 1996, Park and Chadi, 1997] while other investigations have suggested that no DX center is formed in zinc-blende AlGaN. [Van de Walle, 1998] There is consensus, however, that an oxygen DX forms in wurtzite AlN. [Mattila and Nieminen, 1996, Park and Chadi, 1997, Van de Walle, 1998, Van de Walle et al., 1998, Stampfl and Van de Walle, 1998, Stampfl et al., 1999]

There is still considerable debate about the structure of the Si DX center in the wurtzite nitrides. Reports using quantum molecular dynamics calculations [Bogusawski and Bernholc, 1997] and density functional theory [Van de Walle, 1998] predict that Si substituting for Al (or Ga) will form a stable DX center in AlN but not in GaN, while other density functional theory (DFT) results indicate that Si remains a shallow donor across the composition range.
In this report we present experimental and theoretical evidence for the DX formation of Si in AlGaN grown on single crystal AlN substrates. Our results show that Si and O both have DX character in AlN, while they act as shallow donors in GaN. Resistivity measurements of samples with varying Al concentration show that there is a significant increase in the activation energy of Si donors when the Al content is above about x=0.8. Experimental measurements of the conductivity of AlN show that the carrier concentration and defect activation energy depend on the amount of Si and O in the sample because of the interaction of these donor defects. State of the art hybrid DFT methods, which we previously used to explain the deep UV absorption in AlN in close agreement with experiment and which others have used to describe optical properties of defects in III-V nitrides[Collazo et al., 2012, Lyons et al., 2010], confirm that when the Fermi level is close to the conduction band, the shallow donor character of the Si and O defects is overcome by the energetic favorability of the DX centers. Our results show that the oxygen DX has a thermal activation energy of 550 meV and the silicon DX has an activation energy of 200 meV. Based on these theoretical results, we confirm that the low conductivity observed in high Al-content alloys is due to the interaction between these two defects.

8.2 Results & Discussion

8.2.1 Conductivity and activation energy change as a function of Si content

The compensation ratio and activation energy of these samples is shown in Figure 8.1. Low silane flow corresponds to lower Silicon incorporation in the lattice. The low-silane sample showed a high compensation ratio (13) and an activation energy of 267 meV. The middle sample showed a low compensation ratio (4) and an activation energy of 247 meV. The high-silane sample showed a high compensation ratio (10.6) and an activation energy of 251 meV. These results suggest that additional Si is compensated by oxygen transitioning between a shallow donor state and a deeper acceptor, consistent with DX behavior. As silicon is further increased, the self-compensation of Si via the Si DX center becomes evident.

8.2.2 Structure of substitutional and DX centers

Silicon readily substitutes for the group III element (Al or Ga) in III-V nitrides while oxygen favors nitrogen substitution. The structures of three common configurations of substutional Si\textsubscript{Al} defects are shown in Figure 8.2. Substitutional defects may either incorporate on the original lattice site with little lattice relaxation (Figure 8.2a) or, as is characteristic of DX centers, may prefer a lattice relaxation. In DX centers, this relaxation is characterized by breaking the bond...
Figure 8.1: Measured compensation ratio (red) and activation energy of donors (blue) for three AlN samples with varying Si concentration. For low and high Si concentration (low and high silane flow) the compensation ratio is high, while it falls to a minimum when Si concentration is on the order of the background oxygen. Above a Si threshold, the activation energy is lowered to about 250 meV. Figure courtesy Collazo and Sitar groups.
between the defect and a nearest neighbor atom. In the wurtzite structure, this broken bond can either be oriented in the [0001] direction (Figure 8.2b) or one of the three other equivalent bond directions (Figure 8.2c). By convention these are referred to as $\gamma$-BB (c-axis) and $\alpha$-BB (three-fold axes) [Park and Chadi, 1997], but for simplicity we will refer to them as $DX1$ and $DX2$, while we refer to the onsite substitutional defect as $d$. The oxygen donor ($O_N$) has three analogous configurations, with the O sitting on an N site, and the broken bond between the O and a neighboring Ga/Al atom.

![Figure 8.2: Configurations of Si donor substituting for the III-atom (Al or Ga) in wurtzite AlGaN. Small blue spheres represent N, large green spheres represent the III atom, and the large yellow sphere represents Si. a) shows the standard on-site substitution, b) shows the $\gamma$-BB where the broken bond is in the c-direction, and c) shows the $\alpha$-BB, where the broken bond is in one of the three non-c directions.](image)

The configurations shown in Figure 8.2 served as the initial geometries for the DFT relaxation. Similar initial geometries were created for the oxygen defects. Some configurations were unstable, that is, they relaxed to one of the other configurations, typically the on-site substitutional. For those defects which retained DX character, the relative stability of the defect is
Table 8.1: Comparison of energies and nearest-neighbor bond lengths in on-site SiAl (d) and the two DX variants in AlN.

<table>
<thead>
<tr>
<th>Neighbor Distances (Å)</th>
<th>stability</th>
<th>∆ E (eV)</th>
<th>d1</th>
<th>d2</th>
<th>d3</th>
<th>d4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al Site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d⁺ ground</td>
<td></td>
<td>0</td>
<td>1.77</td>
<td>1.77</td>
<td>1.77</td>
<td>1.77</td>
</tr>
<tr>
<td>DX1⁺ meta</td>
<td></td>
<td>1.23</td>
<td>2.00</td>
<td>1.82</td>
<td>1.81</td>
<td>1.81</td>
</tr>
<tr>
<td>DX2⁺ unstable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d⁰ meta</td>
<td></td>
<td>0.44</td>
<td>1.78</td>
<td>1.77</td>
<td>1.77</td>
<td>1.77</td>
</tr>
<tr>
<td>DX1⁰ meta</td>
<td></td>
<td>0.30</td>
<td>1.77</td>
<td>1.78</td>
<td>1.76</td>
<td>1.76</td>
</tr>
<tr>
<td>DX2⁰ ground</td>
<td></td>
<td>0</td>
<td>1.80</td>
<td>1.90</td>
<td>1.79</td>
<td>1.79</td>
</tr>
<tr>
<td>d⁻ meta</td>
<td></td>
<td>0.97</td>
<td>1.88</td>
<td>1.85</td>
<td>1.95</td>
<td>1.95</td>
</tr>
<tr>
<td>DX1⁻ meta</td>
<td></td>
<td>1.52</td>
<td>1.86</td>
<td>1.79</td>
<td>1.82</td>
<td>1.82</td>
</tr>
<tr>
<td>DX2⁻ ground</td>
<td></td>
<td>0</td>
<td>1.81</td>
<td>2.52</td>
<td>1.80</td>
<td>1.80</td>
</tr>
</tbody>
</table>

determined according to the formation energy of the defect, according to the equation

\[
E^f_1 (\Delta \mu, \mu_e) = E_{def} - E_{bulk} + \sum \mu_S - \sum \mu_D + q (E_F + VBM + \Delta V_q) \tag{8.1}
\]

Where \(E_{def}\) and \(E_{bulk}\) are the DFT-calculated energy of the defected and perfect supercell. The stoichiometry of the equation is balanced by the chemical potentials of the defect atom and replaced atom, given by \(\mu_D\) and \(\mu_S\). The charge of the equation is balanced by the final term, which represents the energy of the electron reservoir multiplied by the charge \(q\). This term depends on the Fermi level \(E_F\), referenced to the valence band maximum (VBM) and a final term \(\Delta V_q\) which accounts for finite size effects in the supercell approximation of DFT.[Persson et al., 2005]

In GaN, none of the DX1 or DX2 centers tested for Si or O were stable ground state configurations. Furthermore, only the Si\(_{Ga}^+\) and O\(_{N}^+\) defects are favorable, and no transitions between charge states are seen.

In AlN, several metastable and stable off-site configurations were found. These are summarized in Table 8.1 for Si and Table 8.1 for O.

In the case of Si, the onsite configuration was the ground state only for the positively charged donor (d⁺) while the c-axis DX center (DX1⁺) provided a metastable configuration with an elongation of the c-axis bond by 6% compared to an Al-N bond, and a slight breathing mode contraction in the other bonds by about 4%. In the neutral charge state, the off-axis DX2 center lengthens one of the three-fold bonds by 1%, while the other three bonds contract by 4%. The
Table 8.2: Comparison of energies and nearest-neighbor bond lengths for on-site O\textsubscript{N} (d\textsuperscript{q}) and the two DX variants in AlN.

<table>
<thead>
<tr>
<th>N Site</th>
<th>Stability</th>
<th>Δ E (eV)</th>
<th>d\textsubscript{1}</th>
<th>d\textsubscript{2}</th>
<th>d\textsubscript{3}</th>
<th>d\textsubscript{4}</th>
</tr>
</thead>
<tbody>
<tr>
<td>d\textsuperscript{+}</td>
<td>ground</td>
<td>0.00</td>
<td>1.99</td>
<td>1.95</td>
<td>1.95</td>
<td>1.95</td>
</tr>
<tr>
<td>DX1\textsuperscript{+}</td>
<td>unstable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DX2\textsuperscript{+}</td>
<td>unstable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d\textsuperscript{0}</td>
<td>meta</td>
<td>0.20</td>
<td>1.86</td>
<td>1.95</td>
<td>1.97</td>
<td>1.97</td>
</tr>
<tr>
<td>DX1\textsuperscript{0}</td>
<td>meta</td>
<td>0.29</td>
<td>2.67</td>
<td>1.82</td>
<td>1.83</td>
<td>1.83</td>
</tr>
<tr>
<td>DX2\textsuperscript{0}</td>
<td>ground</td>
<td>0.00</td>
<td>1.87</td>
<td>1.85</td>
<td>2.02</td>
<td>2.01</td>
</tr>
<tr>
<td>d\textsuperscript{−}</td>
<td>meta</td>
<td>0.32</td>
<td>1.75</td>
<td>2.00</td>
<td>2.01</td>
<td>2.01</td>
</tr>
<tr>
<td>DX1\textsuperscript{−}</td>
<td>meta</td>
<td>0.19</td>
<td>2.91</td>
<td>1.82</td>
<td>1.82</td>
<td>1.82</td>
</tr>
<tr>
<td>DX2\textsuperscript{−}</td>
<td>ground</td>
<td>0.00</td>
<td>1.75</td>
<td>2.81</td>
<td>1.83</td>
<td>1.83</td>
</tr>
</tbody>
</table>

DX1 configuration relaxes to a state only slightly perturbed from the on-site position, but is 300 meV higher in energy than the DX2 center. The on site position is a further 140 meV higher in energy. Finally, for the negative charge state the DX2 center is considerably more stable than the other metastable configurations. In this case, one of the three-fold bonds has lengthened by 34% while the others contract by about 4%. This significant relaxation is accompanied by an energy reduction of 0.97 eV from the d\textsuperscript{−}, which is itself slightly perturbed off of the host lattice site, because two of the four-fold bonds are lengthened by 4% and the other is shortened by 2%.

For the substitutional oxygen defect in AlN, there is no stable DX formation in the positive charge state. In the neutral charge state, however, the DX2\textsuperscript{0} configuration is stable compared to the on-site defect by 0.2 eV. The DX2\textsuperscript{0} configuration features the lengthening of two of the three-fold bonds and a shortening of the third. The metastable DX1\textsuperscript{0} configuration features a significant lengthening of the c-axis bond by 42%, pushing the oxygen into the plane of the three other nearest neighbor aluminum atoms. In the negative charge state, the DX2\textsuperscript{−} configuration is the ground state. This DX2\textsuperscript{−} is more similar to the conventional α-BB, where one three-fold bond has been broken. The DX1\textsuperscript{−} shows a broken c-axis bond like the prototypical γ-BB defect. The on-site configuration is slightly distorted off the host lattice site, the c-axis bond is shortened by 7%. The DX2\textsuperscript{−} configuration is 0.32 eV more favorable than the on-site substitution.

Whether these stable DX configurations will be present in the material depends on the total formation energy of the defect, which depends on the growth conditions and Fermi level. The
formation energies of the Si\textsubscript{Al} and O\textsubscript{N} defects are shown in Figure 8.3. In each plot, the formation energies of the stable and metastable defects are shown as a function of the Fermi level. For the purposes of plotting in Figure 8.3, the chemical potentials of the host species are referenced to growth conditions between the Al-rich and N-rich extremes, and the chemical potentials of Si and O are with respect to experimental enthalpies of formation of their solubility limiting phases Si\textsubscript{3}N\textsubscript{4} and Al\textsubscript{2}O\textsubscript{3}.[Lide, 2009] The choice of these chemical potentials for Figure 8.3 does not affect the transitions between charge states. At a given value of the Fermi energy, the most favorable charge state is the one with the lowest formation energy. The plots show that both silicon and oxygen undergo a transition from the positively charged state (shallow donor) to the negatively charged states (deep acceptor) when the Fermi level is near the conduction band minimum. This transition area can be seen in greater detail in the insets of the figures. The thermodynamic transition level between Si\textsubscript{Al}\textsuperscript{+}(d\textsuperscript{+}) and Si\textsubscript{Al}\textsuperscript{−}(DX\textsubscript{2}\textsuperscript{−}) occurs when the Fermi level is 5.89 eV, 200 meV below the conduction band. The transition between O\textsubscript{N}\textsuperscript{+}(d\textsuperscript{+}) and O\textsubscript{N}\textsuperscript{−}(DX\textsubscript{2}\textsuperscript{−}) occurs at 5.54 eV, 550 meV below the conduction band.

![Si\textsubscript{Al} Defects](image-a)

![O\textsubscript{N} Defects](image-b)

**Figure 8.3:** Defect formation energies of dopants in AlN. (a) Substitutional Si on the Al site; (b) substitutional oxygen on the nitrogen site.
8.2.3 Self-consistent mass balance approach

In order to explain the experimental changes in compensation and activation energy, we implement a self-consistent approach to determining the concentration of each defect in each charge state. The details of our implementation of this approach are published elsewhere, and so are only briefly summarized here. The relative concentrations of the defects in the material will depend on growth conditions and the Fermi level. The equilibrium Fermi level, in turn, can be determined based on a solution to the mass balance equation

\[ n + N_{a^-} = p + N_{d^+}. \]

In this expression, \( n \) and \( p \), the number of electrons and holes that are activated, are determined at the temperature of interest. Similarly, the number of ionized acceptors and donors, \( N_{a^-} \) and \( N_{d^+} \), is determined by the defect concentrations and the thermodynamic conditions and value of the Fermi level. The concentrations of defects are determined by the DFT-calculated formation energies of the defects.

Figure 8.4: (a) The simulated compensation ratio as a function of Si content is shown for AlN samples with several different concentrations of background oxygen. Compare to experimental results shown in Figure 8.1. In both cases, low and high Si content corresponds to a greater compensation ratio, while there is a minimum for intermediate Si content. (b) The transition between regions where [Si]<[O] to where [Si]>[O] corresponds to a shift in the equilibrium Fermi level from around 5.55 (near the thermodynamic transition level for O) to 6.07 (above the transition for Si). For high [Si] content, the negatively charged Si DX2- is the energetically preferred charge state.
Table 8.3: Equilibrium defect concentrations as determined by the mass-balance procedure described in the text for samples with varying Si concentration and a fixed oxygen background of $1 \times 10^{18} \text{cm}^{-3}$

<table>
<thead>
<tr>
<th>Sample</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>[O]</td>
<td>$1 \times 10^{18}$</td>
<td>$5 \times 10^{17}$</td>
<td>$1 \times 10^{18}$</td>
<td>$5 \times 10^{17}$</td>
<td>$1 \times 10^{18}$</td>
</tr>
<tr>
<td>[Si]</td>
<td>$2 \times 10^{17}$</td>
<td>$5 \times 10^{17}$</td>
<td>$1 \times 10^{18}$</td>
<td>$5 \times 10^{17}$</td>
<td>$1 \times 10^{19}$</td>
</tr>
<tr>
<td>$E_F$ (eV)</td>
<td>5.53</td>
<td>5.54</td>
<td>5.6</td>
<td>6.04</td>
<td>6.06</td>
</tr>
<tr>
<td>SiDX2^0_{Al}</td>
<td>$3 \times 10^{17}$</td>
<td>$10^{17}$</td>
<td>$4 \times 10^{10}$</td>
<td>$2 \times 10^{18}$</td>
<td>$7 \times 10^{18}$</td>
</tr>
<tr>
<td>Si^{+}_{Al}</td>
<td>$1 \times 10^{17}$</td>
<td>$5 \times 10^{17}$</td>
<td>$1 \times 10^{18}$</td>
<td>$3 \times 10^{18}$</td>
<td>$3 \times 10^{18}$</td>
</tr>
<tr>
<td>ODX2^-_{N}</td>
<td>$5 \times 10^{17}$</td>
<td>$7 \times 10^{17}$</td>
<td>$1 \times 10^{18}$</td>
<td>$1 \times 10^{18}$</td>
<td>$1 \times 10^{18}$</td>
</tr>
<tr>
<td>O^+_N</td>
<td>$5 \times 10^{17}$</td>
<td>$3 \times 10^{17}$</td>
<td>$5 \times 10^{15}$</td>
<td>$7 \times 10^{00}$</td>
<td>$2 \times 10^{00}$</td>
</tr>
</tbody>
</table>

In order to understand the changes in compensation, the concentration of Si and O defects in their many charge states are shown in Table 3 for five impurity profiles ranging from $[\text{Si}] \ll [\text{O}]$ to $[\text{Si}] \gg [\text{O}]$. In the lowest Si case (A), the oxygen donor is self-compensated by the DX center, according to the reaction: $2d^0 \rightarrow d^+ + DX2^-$. [Chadi and Chang, 1988] The Si donors are being compensated by the cumulative effect of several low-concentration compensating centers such as vacancies which are not shown in the table. In example (B), $6 \times 10^{17} \text{cm}^{-3}$ oxygen donors participate in the DX reaction, creating $3 \times 10^{17} \text{cm}^{-3}$ each of $d^+$ and $DX2^-$. The $5 \times 10^{17} \text{cm}^{-3}$ Si donors are all compensated, $4 \times 10^{17} \text{cm}^{-3}$ of them contribute an electron to ODX2^- states and $1 \times 10^{17} \text{cm}^{-3}$ are compensated by the same centers as in example (A). In example (C), the Fermi level is above the transition level for the oxygen donor, and all of the oxygen donors are compensated by Si. In examples (D) and (E) the oxygen is fully compensated, but Si is only partially compensated.

### 8.2.4 Configuration coordinate to determine barriers

In the doped AlN samples, as illustrated by the examples in Table 1, each occupied SiDX2^0_{Al}^- state has two electrons that can be ionized. These electrons can be promoted into the conduction band either via a photon absorption process, or by thermal excitations. The configuration coordinate diagrams for thermal excitations from these defect states is shown in Figure 8.5. The optical ionization energy is 1.25 eV and the zero-phonon line of the thermal ionization is 200 meV. As indicated in the figure, there is a barrier to the thermodynamic transition between the DX2- acceptor (at r$_1$) and the d+ donor (at r$_2$), in good agreement with the experimentally measured 250 meV activation energy.
Figure 8.5: Configuration coordinate diagram showing the transitions between the shallow donor state and the deep acceptor state of the Si defect.

8.3 Conclusions

Silicon-doped AlN samples doped exhibit a 250 meV activation energy for n-type conductivity. This activation energy is a result of the stability of a deep DX acceptor when the Fermi level is close to the conduction band. The DX reaction prevents Si from acting as a shallow donor as it does in GaN. The zero-phonon thermodynamic transition level between the DX acceptor and the on-site donor is calculated to be 200 meV. A thermally activated transition between the two charge states would require additional energy to overcome the thermodynamic barrier. This barrier is estimated to be on the order of 50 meV, accounting for the higher measured activation energy.

8.4 Computational Setup

Atomic structures were relaxed and total energies of pure and defected GaN and AlN supercells were calculated by solving the many-electron problem according density functional theory [Payne et al., 1992, Jones and Gunnarsson, 1989] using pseudopotentials and a plane wave basis set as implemented in VASP [Kresse and Hafner, 1993, Kresse and Hafner, 1994, Kresse and Furthmuller, 1996b, Kresse and Furthmuller, 1996a], and using the HSE[Heyd et al., 2003, Heyd et al., 2006] refinement to the PBE [Perdew et al., 1996a, Perdew et al., 1997] generalized gradient approximation to the exchange and correlation functional, by mixing in a fraction of local exact Hartree-Fock exchange. The fraction of exact exchange for AlN and GaN were chosen to best match the experimental atomic and electronic structure, yielding a fraction of 0.32 for
AlN and 0.31 for GaN. Structures were simulated using a 96-atom wurtzite supercell, where bulk configurations were approximated by freezing atoms beyond the fourth shell of nearest neighbors.[Collazo et al., 2012, Lyons et al., 2010] Relaxations and total energy calculations were performed with a 2 x 2 x 2 Monkhorst-Pack k-points grid and an energy cutoff of 500 eV. The equilibrium lattice parameters for the wurtzite structure were found to be $a=3.11$, $c=4.96$, and $u=0.382$ for AlN and $a=3.21$, $c=5.20$, and $u=0.378$. The theoretical band gaps were 6.09 eV for AlN and 3.5 eV for GaN.
Chapter 9

The Donor-Acceptor Pair of Carbon and Nitrogen Vacancies in Bulk AlN

9.1 Authors

Benjamin E. Gaddy, Zachary A. Bryan, Isaac Bryan, Ronny Kirste, Zlatko Sitar, Ramón Collazo, Douglas L. Irving
Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, USA
Seiji Mita, Baxter Moody
HexaTech, Inc., 991 Aviation Pkwy., Suite 800, Morrisville, North Carolina 27560, USA
Yoshinao Kumagai, Akinori Koukitu
Department of Applied Chemistry, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan
Toru Nagashima, Yuki Kubota, Toru Kinoshita
Tsukuba Research Laboratories, Tokuyama Corporation, Tsukuba, Ibaraki 300-4247, Japan
This article presents DFT calculations which explain the source of a previously unidentified 2.8 eV emission peak detected by PL spectroscopy. Experimental luminescence and absorption measurements were performed in the Collazo and Sitar groups on substrates grown at HexaTech, NCSU, and Tokuyama Corporation. My DFT calculations on the formation energies, optical transitions, and equilibrium concentrations of several defects were used to construct a model of the DAP recombination that would lead to a 2.8 eV emission. PLE measurements taken by the Collazo and Sitar groups confirmed that the 2.8 eV emission was the results of DAP recombination.

This article is being prepared for submission to Applied Physics Letters.
9.2 Introduction

The wide, tunable bandgap of Al\textsubscript{x}Ga\textsubscript{1-x}N will enable the widespread and inexpensive creation of deep-UV LEDs and lasers for applications such as disinfection and data storage. The efficiency of these devices is greatly enhanced by using a pseudomorphic substrate, such as AlN. Bulk AlN has been shown to suffer from an optical absorption at 4.7 eV, the edge of deep UV, thus reducing the overall extraction efficiency of the device. Our recent work has shown that this absorption is due to the incorporation of substitutional carbon defects on the nitrogen site in the bulk AlN substrate. [Collazo et al., 2012] The same work showed that the ionized carbon acceptor was responsible for the absorption, and was coupled to an optical emission at 3.9 eV. AlN samples with varying concentrations of carbon confirmed the theoretical model. Photoluminescence results of these carbon-containing samples show the appearance of a lower emission peak at 2.8 eV. In this paper, we present results based on photoluminescence (PL) and photoluminescence excitation (PLE) spectroscopy, coupled with state of the art hybrid density functional theory (DFT) formation energies, and solutions to a set of self-consistently solved mass balance equations that show that this 2.8 eV emission peak is caused by a donor-acceptor pair (DAP) between substitutional carbon (C\textsubscript{N}) and the nitrogen vacancy (V\textsubscript{N}).

PL spectroscopy results presented here show that the 2.8 eV emission peak is associated with the carbon concentration in the samples. A 2.78 eV emission peak has previously been identified as related to the aluminum vacancy (V\textsubscript{Al}) transition between the -3/-2 states. [Sedhain et al., 2009] Previous theoretical investigations of vacancies in AlN by Laaksonen et al. show qualitative similarities to the results reported here, [Laaksonen et al., 2009] but the quantitative thermodynamic and optical transition levels differ because the hybrid functional used in this work (HSE) has been shown to better capture transitions for wide-bandgap materials, and better predict experimental transitions. [Lyons et al., 2010, Collazo et al., 2012] Previous theoretical results on native defects in AlN have similarly suggested that nitrogen vacancies are unlikely to form in strongly n-type conditions. The samples used in that study featured high oxygen concentrations, while the present results are for samples with [O] < 5×10\textsuperscript{17} cm\textsuperscript{-3}. Furthermore, our results show that the environmental conditions under which the present samples were grown do not support the formation of appreciable aluminum vacancies.

Our self-consistent mass balance equations are informed by DFT calculations of defect formation energies using state-of-the-art hybrid functionals. By solving the mass balance equations, the concentration of defects in the samples can be predicted. These predictions show that the concentration of positively charged nitrogen vacancies increases with the concentration of C defects. We present a model of the optical transitions between these defects that show that the C\textsubscript{N}\textsuperscript{-} and V\textsubscript{N}\textsuperscript{+} act as a DAP that corresponds to the experimental PLE spectroscopy results.
9.3 Results and Discussion

9.3.1 Photoluminescence Measurements

The HVPE-grown samples with controlled carbon concentrations showed that the presence of carbon was related to a 3.9 eV emission peak that was not present in low-carbon samples. [Kumagai et al., 2010, Collazo et al., 2012] Further PL measurements, shown in Figure 9.1 demonstrate that high carbon samples also have a prominent emission peak at 2.8 eV. The

![Figure 9.1: Photoluminescence measurements of AlN samples with varying carbon concentrations shows that the 2.8 eV emission peak is correlated with higher carbon concentrations. Courtesy of Collazo and Sitar groups.](image)

origin of this emission is not immediately clear: experimental SIMS measurements show that the concentration of other defects is low, and theoretical calculations of the optical transitions of native any common impurity defects show no relevant transition at 2.8 eV. Previous studies had suggested a 2.78 eV peak related to the aluminum vacancy. DFT calculations of the formation energies of vacancies ($V_N, V_{Al}$) and substitutional carbon defects ($C_N, C_{Al}$) are shown in Figure 9.2 as a function of the Fermi level, at the limits of the accessible growth conditions. The samples measured in this case are formed in aluminum-rich growth conditions (corresponding to Figure 9.2b), under which the aluminum vacancy is unlikely to form, as seen by the lower formation energy of the nitrogen vacancy across the range of the Fermi level.
Figure 9.2: Formation energy of vacancies and substitutional carbon defects in AlN. Formation energy as a function of the Fermi level is shown for the extreme growth conditions: Nitrogen rich (a) and aluminum rich (b). Experimental conditions used to grow the samples mentioned in this report are closer to the Al rich extreme.
9.3.2 Equilibrium Defect Concentrations

The formation energy of defects as calculated using DFT can be used to determine the equilibrium concentration of each defect, if the equilibrium Fermi level is known. The equilibrium Fermi level, in turn, depends on the concentration of defects and their ionization energies. Our approach to determining this equilibrium level involves extending the methods of Jansen and Sankey and Zhang and Northrup by determining the equilibrium Fermi level that satisfies the law of mass action, for each value of the growth conditions. [Jansen and Sankey, 1989, Zhang and Northrup, 1991] The self-consistently calculated equilibrium defect concentrations for a sample with high carbon concentration is shown in Figure 9.3. In Al-rich growth conditions, the $V_N$ concentration is approximately as high as the carbon concentration. Figure 9.4 shows the relationship between $V_N$ concentration and carbon concentration at extreme values of the chemical potential.

![Figure 9.3: Equilibrium concentration of defects in AlN as a function of growth conditions. Concentrations are shown as quenched at 300K.]

9.3.3 DAP Model

The defect formation energies and solutions to the equilibrium Fermi level establish a clear link between the presence of carbon and the presence of the nitrogen vacancy. The optical transitions caused by the simple $V_N^{(+1,+2)}$ transition would emit at 4.12 eV, not the 2.8 eV seen experimentally. Based on the opposite charge states of the $C_N^-$ and $V_N^+$, there exists the possibility of a compensating interaction between the two defects. In order to test this possibility, we have developed a model for the relevant DAP interaction, which is shown in
Figure 9.4: The vacancy concentration follows the carbon concentration, indicating that the vacancy is the primary compensating defect.

Figure 9.5. At equilibrium, the substitutional carbon and the nitrogen vacancy are ionized as $C_N^- \text{ and } V_N^+$ due to the compensation between them. Under illumination, the $C_N^-$ center (point A in Figure 9.5) absorbs a photon with energy 4.7 eV and excites an electron to the conduction band (B). This photo-absorption process can be seen in transmission measurements of bulk AlN crystals. [Collazo et al., 2012] The excited electron can be coulombically attracted to the positively charged vacancy $V_N^+$ (C). At this point, the system can reduce its energy via a photon emission of 0.81 eV, returning the nitrogen vacancy to its neutral state $V_N^0$ (D). According to the Franck-Condon principle, this emission is followed by a lattice relaxation which reduces the system energy by 0.7 eV (E). The vacancy and carbon defects are now both neutrally charged. At this point, the carbon defect compensates the nitrogen vacancy, resulting in the emission of a photon with a minimum emission of 2.55 eV (F). This 2.55 eV minimum calculated emission corresponds to the measured 2.8 eV emission – due to the significant lattice relaxation in the $V_N^0$ state, the emission occurs before full relaxation, as seen in GaN, ZnO, and AlN. [Alkauskas et al., 2012, Collazo et al., 2012]. Following this emission, a final lattice relaxation occurs to bring the system back to its initial starting configuration at (A).

9.4 Conclusions

The 2.8 eV emission peak that has been associated with high carbon concentrations in bulk AlN crystals is caused by a transition between a Donor-Acceptor pair formed between the substitutional carbon and the nitrogen vacancy. Density Functional Theory shows that for values of the Fermi level near the middle of the band, these two defects have opposite charge states:
Figure 9.5: The configuration-coordinate diagram describes the transitions that would take place in the material when carbon and the nitrogen vacancy are interacting as a DAP.

carbon is an ionized acceptor and the nitrogen vacancy is an ionized donor. A self-consistent solution to the equilibrium Fermi level demonstrates that increasing carbon concentrations lead to a change in the Fermi level which supports the formation of additional vacancies. A DFT-based optical transition model shows that the donor-acceptor pair transition has an absorption at 4.7 eV and a minimum predicted emission at 2.55 eV, corresponding to the experimentally observed 2.8 eV emission. Temperature-dependent PL and PLE measurements confirm that the observed emission is a DAP transition.

9.5 Methods

9.5.1 Formation Energy of Defects

The equilibrium concentration \( C_i \) of a defect \( i \) in a material can be determined based on the formation energy according to the expression

\[
C_i = N_{\text{conf}} \exp \left( -\frac{E_{f,i}}{k_B T} \right),
\]

where \( E_{f,i} \) is the formation energy of the defect, \( N_{\text{conf}} \) is the number of possible configurations of the defect, \( k_B \) is Boltzman’s constant, and \( T \) is the temperature. The formation energy of a defect (D) on a given atomic site (S) in charge state q is calculated via the equation:

\[
E_f^q (\Delta \mu, \mu_e) = E_{\text{def}} - E_{\text{bulk}} + \sum \mu_S - \sum \mu_D + q (\mu_e + VBM + \Delta V_q)
\]
In this expression, $E_{\text{def}}$ and $E_{\text{bulk}}$ are the energy of the defected and perfect supercell, respectively, as calculated by DFT. The terms $\mu_D$ and $\mu_S$ are the chemical potentials of the defect atom and the atom replaced at site S, respectively, and are required to balance the stoichiometry of the equation. The final term balances the charge of the equation. If the charge ($q$) of the defect is non-zero, the term in parentheses represents the energy of the electron reservoir, which depends on the Fermi level referenced to the valence band maximum (VBM) and a final term $\Delta V_q$ which accounts for finite size effects in the supercell approximation of DFT.[Persson et al., 2005]

In AlN, $\mu_S$ will always be $\mu_{Al}$ or $\mu_{N}$, thus the formation energy of a given defect depends on the reference state chosen for the components. Experimental growth often occurs between Al-rich conditions (where $\mu_{Al} = \mu_{Al\text{bulk}}$) and N-rich conditions (where $\mu_{N} = \frac{1}{2}\mu_{N_2}$). Quantitative measurement of these chemical conditions is difficult, but the experimental growth conditions for the samples here fall nearer the Al-rich extreme. We can, nevertheless, derive a variable which represents how far we are from the limits of the accessible chemical potentials as:

$$\Delta \mu = (\mu_{Al} - \mu_{N}) - \left(\mu_{Al(\text{bulk})} - \frac{1}{2}\mu_{N(N_2)}\right)$$

The value of $\Delta \mu$ is bounded by the enthalpy of formation of AlN ($\pm \Delta H_{\text{AlN}}$), which we calculate to be -3.281 eV.

### 9.5.2 Mass Balance Equations

Jansen and Sankey presented a method for calculating the concentration of native defects as a function of stoichiometry in several compound semiconductors.[Jansen and Sankey, 1989] Their results fixed the Fermi level at n-type and p-type extremes. Later, Zhang and Northrup extended this concept by defining an effective doping level and solving for the Fermi level that satisfied the doping level, then using the equilibrium Fermi level to determine the defect formation energies and concentrations.[Zhang and Northrup, 1991] We extend these methods by iteratively solving for the Fermi level which satisfies the mass balance equations for each value of the chemical potential of the growth environment, making no a priori assumptions about which defects are ionized. We solve the mass balance equation $n + N_a^- = p + N_d^+$, where $n$ and $p$ are the number of thermally activated electrons and holes, respectively, and $N_a^-$ and $N_d^+$ are the number of ionized acceptors and donors in the system. The number of ionized acceptors, for example, is determined by the equation $N_a^- = \frac{N_o}{1 + 4e^{\frac{-E_a - \mu_e}{kT}}}$, where $N_o$ is the number of acceptors, $E_a$ is the activation energy of the acceptor, and $\mu_e$ is the Fermi level. The number of acceptors and the activation energy are determined based on the DFT results. The Fermi level is determined according to our iterative procedure.
For each value of the chemical potential, the Fermi level is swept across the range of its possible values, from the valence band to the conduction band. For each pair of chemical potential and Fermi level, the formation energy of each defect is updated, as is the concentration of the defect. The number of ionized defects, as well as the number of electrons and holes in the conduction and valence bands are also a function of the Fermi level, and these values are also updated. The Fermi level is adjusted until the mass balance equation is satisfied. The next value of the chemical potential is chosen, and the equations are updated and resolved. This leads to a description of the equilibrium Fermi level as a function of growth conditions, and thus enables a prediction of the equilibrium concentration of defects as a function of the growth conditions. In our procedure, we solve this set of equilibrium concentrations at an elevated temperature representative of the high temperatures during experimental growth. We then solve the equations again at low temperature, using a simulated quenching where electronic rearrangement is allowed but atomic diffusion is constrained. Thus a non-equilibrium concentration of defects is present at low temperatures.
Chapter 10

Enabling the smooth growth of rocksalt oxides by changing the chemical environment.

This chapter presents a summary of my \textit{ab initio} thermodynamics results which predict an experimental growth window where polar \{111\} rocksalt CaO can be grown epitaxially with a smooth surface.

My calculations were part of a collaborative effort between the J.P. Maria thin-film oxides group and the Collazo and Sitar wide-bandgaps laboratory.

The collaborative research article was published in \textit{Nature Communications}, vol. 2, article 461 on September 6, 2011.

I would like to thank my experimental collaborators who grew and characterized the GaN substrates and the epitaxial oxide thin films, and with whom the results and research directions were discussed: Elizabeth A. Paisley, Mark. D. Losego, James S. Tweedie, Ramón Collazo, Zlatko Sitar, Jon-Paul Maria.
10.1 Introduction

The heteroepitaxial integration of dissimilar materials may provide a pathway to new functionality at the interface that is not accessible in the bulk or single-material thin film. This new functionality may arise from the property coupling between the advanced functional materials on either side of the interface. Previously this property coupling had typically only been possible for materials with the same structure (wurtzite grown on wurtzite, perovskite on perovskite). The set of possible interfaces between dissimilar materials can be expected to provide an enhanced set of properties, if the challenges to this integration can be overcome. The two main challenges are epitaxial integration – finding common symmetry between the two different materials – and a smooth growth mode. The experiment discussed in this chapter uses cubic, rocksalt CaO and hexagonal, wurtzite GaN as an example of this integration. The first challenge is overcome by exploiting the hexagonal symmetry of the \{111\} plane of the CaO. The hexagonal surface of the GaN \{0001\} substrate provides an epitaxial seed for the continued low-strain \{111\} growth of the CaO. This \{111\} orientation of the rocksalt is an unfavorable polar orientation. The implications of this polarity, and how it can be mitigated, are primary concern of the results in this section.

Attempts to grow smooth \{111\} CaO surfaces by molecular beam epitaxy (MBE) were undertaken in the JP Maria group. Reflection high-energy electron diffraction (RHEED) measurements provided quantitative measurement of the roughness of the surfaces. When the CaO was grown using a traditional molecular oxygen flux (O\textsubscript{2}) the RHEED patterns indicated a rough morphology after only a few layers of growth. The significant experimental achievement of our collaborators was the observation that growth using water vapor flux (H\textsubscript{2}O) led to smooth growth for several layers, as confirmed by RHEED and atomic force microscopy (AFM). The high formation energy of the \{111\} surface, compared to the lower energy \{001\} or \{011\} surfaces was well known, but it was unclear why the use of water vapor to the growth environment would change this behavior. To determine the role of hydrogen and/or water in the growth environment, I coupled DFT calculations of the ground state surface formation energies of 19 different low-index surface terminations with DFT calculations of the zero-point energy vibrations and extended the 0 K calculations to experimentally relevant temperatures and pressures via \textit{ab initio} thermodynamics (AIT) using the JANAF thermochemical tables.

10.2 Computational Methodology

The surface energies of low-index CaO surfaces (\{001\}, \{011\}, and \{111\}) were compared in order to determine the energetic penalties associated with the polar \{111\} surface. The labels and descriptions of the 15 different surface terminations I explored are listed in Table 10.1.
Table 10.1: Surface configurations evaluated with ab initio thermodynamics calculations: The 16 surface configurations considered for surface energy calculations and the label by which they are referred in Figure 10.15

<table>
<thead>
<tr>
<th>Surface/Plot Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(001) clean CaO</td>
<td>CaO (001) surface</td>
</tr>
<tr>
<td>(001) 1ML H2O CaO</td>
<td>CaO (001) hydrated with one monolayer H2O</td>
</tr>
<tr>
<td>(001) 1/2 ML H2O CaO</td>
<td>CaO (001) hydrated with a 1/2 monolayer H2O</td>
</tr>
<tr>
<td>(001) OH a CaO</td>
<td>CaO (001) hydroxylated with OH atop each Ca, H atop each O</td>
</tr>
<tr>
<td>(001) OH b CaO</td>
<td>CaO (001) hydroxylated with OH between Ca neighbors along [110]</td>
</tr>
<tr>
<td>(001) p(3x2) CaO</td>
<td>CaO (001) with 4 of 6 Ca bonded to H2O, 2 of 6 Ca bonded to OH, 2 of 6 surface O bonded to H from hydroxylation</td>
</tr>
<tr>
<td>(111) clean CaO</td>
<td>CaO (111) with Ca termination</td>
</tr>
<tr>
<td>(111) clean O</td>
<td>CaO (111) with O termination</td>
</tr>
<tr>
<td>(111) Ca Oct CaO</td>
<td>CaO (111) with octopolar reconstruction Ca termination</td>
</tr>
<tr>
<td>(111) O Oct CaO</td>
<td>CaO (111) with octopolar reconstruction O termination</td>
</tr>
<tr>
<td>(111) H CaO</td>
<td>CaO (111) with full hydroxylation</td>
</tr>
<tr>
<td>(111) H 1ML H2O CaO</td>
<td>CaO (111) with full hydroxylation and hydrated with one monolayer</td>
</tr>
<tr>
<td>(111) H 1/2 ML H2O CaO</td>
<td>CaO (111) with full hydroxylation and hydrated with a 1/2 monolayer</td>
</tr>
<tr>
<td>(011) clean CaO</td>
<td>CaO (011) surface</td>
</tr>
<tr>
<td>(011) OH CaO</td>
<td>CaO (011) with full hydroxylation</td>
</tr>
</tbody>
</table>

These surfaces were chosen based on evaluation of experimental and theoretical results for CaO and MgO, another rocksalt oxide. In order to compare surface energies of these configurations, the bare surfaces of interest were constructed by terminating a bulk supercell which had the desired orientation in the Z-direction. This is illustrated in Figure 10.1, which shows the three cleavage planes for the initial bare, unreconstructed surfaces.

The creation of these surface geometries for DFT calculations was described in detail in Chapter 6. For each orientation of interest (001), (b) (011) (c) (111)) bare surface configurations were created. For those orientations where surfactants and reconstructions were studied, those configurations were also constructed and the energies were calculated. Side views of the surfaces of interest are shown in Figure 10.2 to Figure 10.14.

Each of the surfaces of interest were relaxed and their final surface energy determined VASP [Kresse and Furthmuller, 1996b, Kresse and Furthmuller, 1996a, Kresse and Hafner, 1994, Kresse and Hafner, 1993] and the PBEsol generalized gradient approximation to the exchange and correlation functional. [Perdew et al., 2008, Perdew et al., 1996a, Perdew et al.,
Figure 10.1: Schematic illustration of the cleavage planes of bulk CaO to generate the surface orientations studied. The cleavage planes illustrated are (a) \{001\}, (b) \{011\}, and c) \{111\}.

Figure 10.2: Side view of the \{100\} CaO surface configuration with a bare, unreconstructed surface.

Figure 10.3: Side view of the \{100\} CaO surface configuration with a hydroxyl layer situated atop the surface atoms.
Figure 10.4: Side view of the \{100\} CaO surface configuration with a hydroxyl layer situated above the surface, in a bridge site along the \{110\} direction.

Figure 10.5: Side view of the \{100\} CaO surface configuration with a half monolayer of associatively adsorbed water molecules.

Figure 10.6: Side view of the \{100\} CaO surface configuration with a full monolayer of associatively adsorbed water molecules.
Figure 10.7: Side view of the \{100\} CaO surface configuration with a p(3x2) adsorbate symmetry, featuring 4 associatively adsorbed water molecules, two hydroxyl groups, and two hydrogens in each repeat unit.

Figure 10.8: Side view of the \{110\} CaO surface configuration with a bare, unreconstructed surface.

Figure 10.9: Side view of the \{110\} CaO surface configuration with a hydroxyl termination.
Figure 10.10: Side view of the \{111\} CaO surface configuration with a bare, unreconstructed surface.

Figure 10.11: Side view of the \{111\} CaO surface configuration with the cation-terminated octopolar reconstruction.

Figure 10.12: Side view of the \{111\} CaO surface configuration with hydroxyl termination.
Figure 10.13: Side view of the \{111\} CaO surface configuration with a hydroxyl termination and a half monolayer of associatively adsorbed water molecules.

Figure 10.14: Side view of the \{111\} CaO surface configuration with a hydroxyl termination and a full monolayer of associatively adsorbed water molecules.
1997] PBEsol was chosen to best reproduce surface energies and lattice constants compared to experiment. All of the terminations and coverages for a given surface orientation were computed in identical supercells. The thickness of the slabs, the kinetic energy cutoff (1000 eV), and the k-point mesh (which was supercell dependent) were chosen to converge the bulk cohesive energy to within 6.5 meV/atom and to converge the surface energies to within 4.5 mJ/m^2. The frozen phonon vibrational modes were calculated by using a finite displacement method, using four displacements in each direction and the same basis set as the relaxation and total energy calculations.

Figure 10.15: Temperature dependent comparison of relevant CaO surfaces: Surface free energy values as a function of temperature for a collection of relevant CaO surface orientations and chemical terminations involving bulk, bulk reconstructed, hydroxylated and hydrated. The grey region marks the upper limit of the chemical potential.

10.3 Results

The calculated surface free energies for several of the surface terminations of interest are shown in Figure 10.15. The energies are a function of both temperature and pressure. In the figure, the pressure is fixed based on the experimental growth conditions \( p_{H_2O} = 5 \times 10^{-4} \) Torr, and the surface energies are presented over a temperature range of 0 to 800 K. The grey region of the figure represents a chemical potential of water above its upper limit, where the ideal gas approximation is not expected to hold. Examination of these results reveal three distinct growth regimes. In the first regime, below about 350 K, the \( \{111\}\)H surface is thermodynamically
stable, as indicated by its low formation energy. In the second regime, from 350 K to about 725 K, the \{111\}H surface is still the lowest energy \{111\} surface, but is higher in energy than the \{001\} surface. In the third growth regime, above 725 K, the octopolar reconstruction of the \{111\} surface is more favorable than the hydrogen terminated surface, and the \{001\} surface is lower still. These results indicate the thermodynamic competition between the surface energies of different terminations, and are consistent with experimental observations that in dry environments the \{001\} surface is preferred while in humid environments the hydroxylated \{111\} surface is favored. [Hacquart and Jupille, 2007]

10.4 Summary of Experimental Confirmation

The DFT results presented in the previous section suggest three distinct growth modes. The first mode (low temperature) is not experimentally relevant for MBE growth of CaO. Experimental measurements of growth at several different temperatures were studied by RHEED. These results indicate that when the CaO is grown with water vapor flux, smooth, layer-by-layer growth occurs for temperatures below about 725 K, and then at higher temperatures the surface roughens after only a few layers.

10.5 Discussion

The experimental and theoretical results taken together suggest a significant difference between the intermediate temperature regime and the high temperature regime. The \textit{in situ} RHEED analysis of the rough surfaces (both those grown in oxygen and those grown in water vapor at high temperatures) indicates that the initial growth is a few layers of \{111\} oriented CaO. This orientation is due to the matched symmetry of the \{0001\} GaN substrate. Rough growth then begins to occur after a few layers. Thus, at intermediate temperatures, the \{111\} growth is seeded epitaxially and even though the \{001\} surface is lower in energy, there is a significant barrier to reconstruction. On the other hand, at higher temperatures, the \{111\} surface would be expected to reconstruct into the lower energy octopolar reconstruction. The octopolar reconstruction exposes \{001\} nanofacets, which then grow favorably to create a rough surface with exposed \{001\} planes. The combined experimental and theoretical approach confirms that the smooth, layer-by-layer growth of CaO films on GaN is possible by controlling the experimental growth conditions. Figure 10.16 shows the leakage current of capacitors fabricated using the two CaO films, with and without the surfactant. The smooth \{111\} surface provides a 1000x reduction in the leakage current of a planar capacitor compared to a rough \{111\} surface, illustrating the need for a better understanding of how the growth conditions can be tuned to stabilize otherwise unfavorable surfaces. In Chapter 11 the mechanisms behind the stabilization of polar
surfaces are explored further, and in Chapter 12 the epitaxial integration of the lattice-matched MgCaO alloy system on GaN is discussed.

Figure 10.16: Leakage current comparison for CaO films prepare in oxygen and water vapor: Leakage current analysis of 4.5 nm CaO capacitors. (a) Surfactant-assisted CaO plus capping layer, and (b) conventional CaO plus capping layer. Each trace is an average of data captured for 20 capacitors. The error bars correspond to the standard deviation values at each voltage. The offset in field corresponds to 0.1 V and is consistent with the Schottky barrier known to exist at the GaN|CaO interface. [Craft et al., 2008] The figure appears in the original paper. [Paisley et al., 2011]
Chapter 11

Overcoming the polarization catastrophe in epitaxial thin-film rocksalt oxides (MgO and CaO)

Benjamin E. Gaddy, Elizabeth A. Paisley, Jon-Paul Maria, Douglas L. Irving

This chapter presents results from DFT calculations and related analysis of polar rocksalt oxide surfaces (MgO and CaO). The work arises from a fruitful collaboration with Elizabeth Paisley and Jon-Paul Maria, who we greatly acknowledge for helpful discussions.

This chapter is being prepared for submission to Physical Review B.

11.1 Introduction

Epitaxial integration of heterogeneous materials provides a rich opportunity for property coupling between dissimilar materials. Several new challenges to this integration must be overcome before the property coupling can be explored. New functionality at heterogeneous interfaces can be explored if the two main challenges to integrating dissimilar materials can be overcome. As an example of heterogeneous integration of dissimilar materials, the cubic rocksalt oxides MgO and CaO, have been epitaxially grown on GaN, a hexagonal wurtzite crystal. [Paisley et al., 2011, Paisley et al., 2012] The large bandgaps of these oxides allow them to be used as a buffer layer in electronic devices, and their good band offsets compared with GaN allow them to be used as tunneling barriers. [Craft, 2006, Craft et al., 2008] The interface between these oxides and GaN may also give rise to a two-dimensional electron gas (2DEG). These materials also provide a prototype for the heteroepitaxial integration of different advanced functional materials such as ferroelectrics and piezoelectrics.

The first challenge to integration is the ability to create a smooth, abrupt interface between
materials that do not share a common crystal structure. One successful approach to overcoming this challenge relies on exploitation of a common planar symmetry. In the case of rocksalt oxides, the underlying cubic structure provides hexagonal symmetry on the \{111\} planes. These hexagonal planes can be favorably aligned with the \{0001\} hexagonal wurtzite planes. Thus, growth of MgO and CaO in the \{111\} direction is seeded by the hexagonal symmetry of the GaN substrate, leading to a smooth, atomically abrupt interface. The second challenge to integration is the ability to reliably create a smooth surface. To date, this has been difficult because the same \{111\} orientation chosen to enable epitaxy exposes a high-energy surface. The unfavorable surface lowers its energy by faceting to expose lower energy \{001\} surfaces. The surface roughness that results from this faceting has been shown to increase leakage currents and worsen device performance. [Paisley et al., 2011, Paisley et al., 2012]

### 11.2 The Polarization Catastrophe

A portion of surface energy – the energetic penalty for creating a free surface – can be attributed to broken interatomic bonds at the surface that leave some unsatisfied charge. This relationship is illustrated by Figure 11.1, which shows the DFT calculated surface energy of low-index MgO and CaO surfaces as a function of the number of broken bonds. The trend lines in the figure demonstrate that for non-polar surfaces the primary contribution to surface energy is the unsatisfied bonding configuration at the surface. In contrast, the \{111\} surfaces have a significant additional energy penalty. This additional energy penalty arises as a consequence of polarity.

A slab of \{111\} rocksalt MgO or CaO can be visualized as a stack repeat units, where each unit consists of a layer of cations (Ca\(^{2+}\), Mg\(^{2+}\)) above a layer of anions (O\(^{2-}\)). The opposite charges of each adjacent plane induce an electric dipole between them. The repeating structure leads to a finite electric field in each repeat unit. As the slab thickness increases, the effect of this electric field accumulates and induces an electrostatic potential at the surface which diverges with sample thickness. This is the well-known polarization catastrophe. A schematic of a 6-layer polar surface is shown in Figure 11.2a, where the dark arrows represent uncancelled dipoles and light arrows are dipoles that cancel. Figure 11.2b shows the electric field (E, dotted line) at each point in the slab and the electrostatic potential (V, dash-dotted line), as a function of distance in the z-direction of the crystal.

In an early analysis of polar surfaces, Tasker proposed that if the charges on the outer layers could be reduced to 1/2 the charge of interior layers, the net dipoles would cancel, resulting in a non-diverging surface potential. [Tasker, 1979] This polarity compensation is illustrated schematically in Figure 11.2c. The charges on the outer planes are half of the charge of interior planes. In Figure 11.2d, the electric field is shown (dotted line) and the electrostatic potential
Figure 11.1: The surface energy of various surface configurations of MgO (black squares) and CaO (grey circles) as a function of the number of broken bonds per unit area. The surface labels correspond to Table 2. The trend lines show a fit through the bare, unreconstructed surfaces for both materials.
Figure 11.2: Schematic illustration of the polarization catastrophe. (a) shows a representative polar orientation consisting of 6 planes of opposing charge, where arrows indicate a dipole, dark arrows are uncompensated dipoles, and the arrow length defines the magnitude of the dipole. (b) shows a plot of the electric field (E) and electrostatic potential (V) for the system depicted in (a). (c) shows an example of a compensated polar orientation with half charges on the surfaces, and (d) shows the field and potential for this case.

is shown, which now oscillates between zero and a finite value depending on the slab thickness. Goniakowski, et al. [Goniakowski et al., 2008] described the difference between compensated and uncompensated polarity more generally by dividing the slab into two surface parts (left and right) and a center part. The differences between the two constructions are illustrated in Figure 11.3, which shows the example system of Figure 11.2 with each plane shown as a layer of atoms. Red (light) atoms represent anions (O$^{2-}$) and blue (dark) atoms represent cations (Mg$^{+2}$ or Ca$^{+2}$). The left side of the figure shows Tasker’s partitioning and the right side shown Goniakowski’s partitioning.

In Tasker’s model, the inner central region is a charge-neutral ($\sigma = 0$) repeat unit which contains a dipole ($\mu \neq 0$) perpendicular to the surface. In the divergent Tasker system (Figure 11.3a), the left and right regions are identical to the bulk regions. The compensated Tasker system (Figure 11.3b) has a surface “cap” at each surface whose charge sums to half the charge of a bulk plane.

In Goniakowski’s model, the center region is constructed according to two requirements: it must be charge neutral ($\sigma_C = 0$) and have no net dipole $\mu_C = 0$. The center region of the example system shown in Figure 11.3 (right) can be defined to start in the middle of atomic plane 2 and end in the middle of atomic plane 4. The divergent system (Figure 11.3c) has surface caps which have a net charge, while the convergent system (Figure 11.3d) has a surface “cap”
whose charge sums to zero. It can be shown that in the Goniakowski partition, any system is fully compensated if both of the left and right parts are also charge neutral. It is worth noting that both constructions lead to equivalent descriptions of the systems, and the choice of one over the other is a matter of convenience.

There are three mechanisms by which a slab can satisfy the condition of charge neutrality in the surface regions: electronic redistribution (where the electron density moves away from surface atoms towards the center of the slab), atomic redistribution (where atoms in outer planes are missing), and adsorption of foreign species (such as the chemical or physical adsorption of atoms or molecules). [Noguera, 2000] The system shown in Figure 11.3 is an example of atomic redistribution. In real systems, the energetic competition between these compensation mechanisms ultimately determines the nature of the polar surface.

### 11.3 Calculated Surface Free Energies

In order to understand how the compensation mechanisms discussed in the previous section compete to stabilize polar surfaces, *ab initio* thermodynamics (AIT) calculations of the temperature- and pressure-dependent surface energies of several configurations of the low index surfaces of MgO and CaO are compared. Configurations of polar \{111\} were chosen to compare the three polarity compensation mechanisms. A bare, undecorated surface is most often characterized by an electronic redistribution which leads to changes in the surface charge density. The octopolar reconstruction is an example of the atomic reconstruction mechanism, and was originally
Table 11.1: Descriptions of the surface configurations explored in this study. The plot labels are used to refer to the surfaces in subsequent discussions.

<table>
<thead>
<tr>
<th>Surface</th>
<th>Surface Configuration</th>
<th>Plot Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>{100}</td>
<td>Bare Unreconstructed</td>
<td>(001) clean</td>
</tr>
<tr>
<td></td>
<td>H2O Monolayer</td>
<td>(001) 1ML H2O</td>
</tr>
<tr>
<td></td>
<td>H2O Half Monolayer</td>
<td>(001) 1/2ML H2O</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation – atop</td>
<td>(001) OH a</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation – bridge</td>
<td>(001) OH b</td>
</tr>
<tr>
<td></td>
<td>Mixed Hydrox. &amp; H2O</td>
<td>(001) p(3x2)</td>
</tr>
<tr>
<td>{110}</td>
<td>Bare Unreconstructed</td>
<td>(011) clean</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation</td>
<td>(011) OH</td>
</tr>
<tr>
<td></td>
<td>Bare – Cation Terminated</td>
<td>(111) clean Mg/Ca</td>
</tr>
<tr>
<td></td>
<td>Bare – Anion Terminated</td>
<td>(111) clean O</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation</td>
<td>(111) H</td>
</tr>
<tr>
<td>{111}</td>
<td>H2O Monolayer</td>
<td>(111) 1ML H2O</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation + H2O Monolayer</td>
<td>(111) H 1ML H2O</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation + H2O Half Monolayer</td>
<td>(111) H 1/2ML H2O</td>
</tr>
<tr>
<td></td>
<td>Cation Terminated Octopolar Reconstruction</td>
<td>(111) Mg/Ca Oct</td>
</tr>
<tr>
<td></td>
<td>Anion Terminated Octopolar Reconstruction</td>
<td>(111) O Oct</td>
</tr>
<tr>
<td>{211}</td>
<td>Bare Unreconstructed</td>
<td>(112) clean</td>
</tr>
<tr>
<td></td>
<td>Hydroxylation</td>
<td>(112) H</td>
</tr>
</tbody>
</table>
proposed by Wolf, et al. for NaCl. [Wolf, 1992] The reconstruction is defined by having 3/4 of the surface atoms and 1/4 of the second-layer atoms missing, leaving a rough, pyramidal morphology. The final mechanism, foreign species adsorption, was examined by comparing surfaces decorated with hydroxyl groups or associated water molecules. The polar configurations were compared to decorated and undecorated non-polar \{001\}, \{011\}, and \{112\} surfaces. Descriptions of all the surfaces studied, along with the labels used to identify them in the text, are listed in Table 11.1. DFT-calculated ground-state energies and vibrational contributions are tabulated in Table 11.2. The ground state energies are calculated with a plane-wave basis set and the PBEsol refinement [Perdew et al., 2008, Perdew et al., 2009] to the generalized gradient approximation functional, [Perdew et al., 1996a] as implemented in VASP. [Kresse and Hafner, 1993, Kresse and Hafner, 1994, Kresse and Furthmuller, 1996b] The k-point mesh for each simulation cell was chosen such that the total cohesive energy was converged to within 0.0065 eV/atom and the surface and hydration energies were converged to within 0.0045 J/m². [Monkhorst and Pack, 1976] Vibrational energies were determined using a finite-difference approach.

The AIT method extends these DFT calculations of the ground state energy and the 0K vibrational energy contributions to experimentally relevant temperatures and pressures via thermochemical data. [Reuter and Scheffler, 2001, Reuter and Scheffler, 2003] AIT has been used to demonstrate that the favorability of surface terminations changes as a function of growth conditions for surfaces of a number of different materials. [Sun et al., 2003, Reuter and Scheffler, 2001, Reuter and Scheffler, 2003, Rohrer and Hyldgaard, 2010, Paisley et al., 2011] The surface free energy $\Gamma$ is given by the equation

$$\Delta \Gamma^{total} = \frac{1}{A} \left( E^{total} - n_X g_{XO}^{bulk} (T, p) + (n_X - n_O) \mu_O (T, p) + n_{H_2O} \mu_{H_2O} (T, p) + F_{vib} \right),$$

(11.1)

where $E^{total}$ is the Helmholtz free energy of the surface system without vibrational contributions, and is obtained directly from DFT calculations, $g_{XO}$ is the energy per formula unit of a bulk crystal, $n_X$ and $n_O$ are the number of cation and oxygen atoms, respectively, present in the surface slab calculation, $\mu_O$ is the chemical potential of oxygen, $n_{H_2O}$ is the number of water molecules in the system, $\mu_{H_2O}$ is the chemical potential of water, $F_{vib}$ are the vibrational contributions to the free energy and $A$ is the surface area.

These surface energies agree well with previously calculated ground-state energies for low-index MgO and CaO surfaces. Among the undecorated surfaces, the \{100\} has the lowest energy for both oxides. [de Leeuw and Parker, 1999, de Leeuw and Purton, 2001] Full adsorption of water on the \{100\} surface is unfavorable, but a partially dissociated arrangement with
Table 11.2: Total 0K energies (including ground state vibration) for the surfaces considered. All energies are shown at a growth condition midway between Oxygen-rich and Oxygen-poor limits.

<table>
<thead>
<tr>
<th>Surface</th>
<th>Plot Label</th>
<th>MgO $\Gamma_{ijk}(0K,p^\circ)$</th>
<th>CaO $\Gamma_{ijk}(0K,p^\circ)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>{100}</td>
<td>(001) clean</td>
<td>1.02</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>(001) 1ML H2O</td>
<td>2.15</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>(001) 1/2ML H2O</td>
<td>1.58</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>(001) OH a</td>
<td>2.9</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td>(001) OH b</td>
<td>2.22</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>(001) p(3x2)</td>
<td>0.62</td>
<td>0.2</td>
</tr>
<tr>
<td>{110}</td>
<td>(011) clean</td>
<td>2.33</td>
<td>1.59</td>
</tr>
<tr>
<td></td>
<td>(011) OH</td>
<td>1.96</td>
<td>1.23</td>
</tr>
<tr>
<td>{111}</td>
<td>(111) clean Mg/Ca</td>
<td>5.065</td>
<td>3.97</td>
</tr>
<tr>
<td></td>
<td>(111) clean O</td>
<td>6.925</td>
<td>5.44</td>
</tr>
<tr>
<td></td>
<td>(111) Mg/Ca+1ML H2O</td>
<td>5.465</td>
<td>3.61</td>
</tr>
<tr>
<td></td>
<td>(111) Mg/Ca+1/2ML H2O</td>
<td>2.37</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>(111) O 1ML H2O</td>
<td>6.2</td>
<td>6.01</td>
</tr>
<tr>
<td></td>
<td>(111) O 1/2ML H2O</td>
<td>4.35</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>(111) H</td>
<td>-0.43</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td>(111) H 1ML H2O</td>
<td>0.25</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(111) H 1/2ML H2O</td>
<td>0.44</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(111) Mg/Ca Oct</td>
<td>2.64</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>(111) O Oct</td>
<td>2.51</td>
<td>1.78</td>
</tr>
<tr>
<td>{112}</td>
<td>(112) clean</td>
<td>3.28</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>(112) H</td>
<td>0.66</td>
<td>0.01</td>
</tr>
</tbody>
</table>
p(3x2) symmetry is seen, in agreement with previous reports for MgO. [Scamehorn et al., 1993, Giordano et al., 1998, Odelius, 1999, Delle Site et al., 2000, Giordano et al., 2000, Lynden-Bell et al., 2002, Ferry et al., 1998] The octopolar reconstruction of the \{111\} surface reduces the surface energy compared to the bulk-terminated surface. In a humid environment, the \{111\} surface prefers full hydroxylation over the bare surface, reconstruction, or partial water coverage. [Refson et al., 1995, de Leeuw and Parker, 1999, Hacquart and Jupille, 2007] The energetic competition between the hydroxylated \{111\} surface and the clean \{001\} surface is illustrated in Figure 3 where the lowest-energy surface is shown as a function of the chemical potential of oxygen and water. For both MgO (Figure 3a) and CaO (Figure 3b) the \{111\} H surface is favorable when the chemical potential of water is near its maximum, that is, in humid conditions. As seen in the plot, the \{111\} H surface is favored when the differential chemical potential of water $\Delta \mu_{\text{H}_2\text{O}}$ is greater than -0.9 eV, in all but the most extreme oxygen-poor conditions. Using the framework of AIT, the chemical potentials can be expressed in terms of their dependence on temperature ($T$) and pressure $p$ by an expression of the form $\Delta \mu_{\text{H}_2\text{O}}(T, p) = \mu_{\text{H}_2\text{O}}(T, p^\circ) + \frac{1}{2} kT \ln \left( \frac{p}{p^\circ} \right)$, where $\mu_{\text{H}_2\text{O}}(T, p^\circ)$ is obtained from thermochemical tables. [Wang et al., 2000, Stull and Prophet, 1971]

![Surface Phase Diagram for MgO](image1.png)

![Surface Phase Diagram for CaO](image2.png)

Figure 11.4: Surface phase diagrams for MgO (left) and CaO (right). The lowest energy surface is shown for each material as a function of the chemical potential of oxygen and water. The chemical potentials are referenced to 0K enthalpies, as described in the text.
Table 11.3: The charge of each layer in the surface caps of MgO with different polarity compensation mechanisms. The regions have been assigned according to Goniakowski’s method. In each case, the sum of the surface caps is zero, indicating that polarity has been compensated. White cells indicate hydrogen surfactant layers, red cells indicate oxygen layers, blue cells indicate magnesium layers, and grey cells indicate a cell that is shared with the center region, according to Goniakowski’s method.

<table>
<thead>
<tr>
<th>Surface:</th>
<th>{111}O</th>
<th>{111}Mg</th>
<th>{111}OH</th>
<th>{111}MgH</th>
<th>{111}OctO</th>
<th>{111}OctMg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charge sum:</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Layer</td>
<td>Adsorbate</td>
<td>0.32</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(-0.58)</td>
<td>(0.63)</td>
<td>(-0.82)</td>
<td>(0.87)</td>
<td>(-0.24)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>2</td>
<td>(1.00)</td>
<td>(-1.12)</td>
<td>(1.00)</td>
<td>(-1.04)</td>
<td>(0.74)</td>
<td>(-0.73)</td>
</tr>
<tr>
<td>3</td>
<td>(-0.93)</td>
<td>(1.00)</td>
<td>(-1.00)</td>
<td>(1.00)</td>
<td>(-1.00)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>4</td>
<td>(1.00)</td>
<td>(-1.00)</td>
<td>(1.00)</td>
<td>(-1.00)</td>
<td>(1.00)</td>
<td>(-1.00)</td>
</tr>
<tr>
<td>5</td>
<td>(-0.99)</td>
<td>(1.00)</td>
<td>(-1.00)</td>
<td>(1.00)</td>
<td>(-1.00)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>6</td>
<td>(0.50)</td>
<td>(-0.50)</td>
<td>(0.50)</td>
<td>(-0.50)</td>
<td>(0.50)</td>
<td>(-0.50)</td>
</tr>
</tbody>
</table>

11.4 Competition between Compensation Mechanisms

Each of the polar configurations studied undergoes at least one of the three compensation mechanisms. The planar-averaged atomic charge in the atomic layers near the \{111\} surfaces of interest are shown in Table 11.3 and Table 11.4. The atomic charges are defined by Bader’s method, which delineates the boundaries between atoms based on zero flux surfaces of the charge density. [Bader, 1985] The charge of each atom is calculated by summing the total charge density in each Bader volume, and comparing that to the total charge density of an atom in the bulk. [Tang et al., 2009, Henkelman et al., 2006, Sanville et al., 2007] In each case, the charge-neutrality requirement is satisfied, and the polarity of all three surfaces is compensated. This leads us to conclude that the difference in surface energy between the polar surfaces is not due to the polarity itself (since the polarity is fully compensated), but to the compensation mechanism itself. It is interesting to note that the oxygen layers change their charge state more freely than the cation layers. This phenomenon is also seen in the local density of electronic states (discussed in the next section) where the primary changes to the DOS occurs at the top of the valence band, which in these materials is composed of oxygen 2p states.
Table 11.4: The charge of each layer in the surface caps of CaO with different polarity compensation mechanisms. The regions have been assigned according to Goniakowski’s method. In each case, the sum of the surface caps is zero, indicating that polarity has been compensated. White cells indicate hydrogen surfactant layers, red cells indicate oxygen layers, blue cells indicate calcium layers, and grey cells indicate a cell that is shared with the center region, according to Goniakowski’s method.

<table>
<thead>
<tr>
<th>Layer Adsorbate</th>
<th>Layer Charge sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>{111}O</td>
<td>-0.62</td>
</tr>
<tr>
<td>{111}Ca</td>
<td>1.00</td>
</tr>
<tr>
<td>{111}OH</td>
<td>1.00</td>
</tr>
<tr>
<td>{111}OctO</td>
<td>-0.98</td>
</tr>
<tr>
<td>{111}OctCa</td>
<td>0.50</td>
</tr>
</tbody>
</table>

11.5 Energy Cost of Electronic Compensation

The compensation mechanism that is observed in a real system is the one with the lowest energy penalty. Undecorated, unreconstructed {111} surface compensates its polarity by a redistribution of atomic charge, as was shown in Table 11.3 and Table 11.4. The redistribution comes at a high price: the surface energy of the bare {111} surface is significantly higher (2-3 eV) than the octopolar reconstruction or any of the non-polar surfaces. Figure 11.5 shows the calculated local density of electronic states (LDOS) of the atoms near the surface of the {111} surface, and is compared to the bulk DOS. Electronic states are pushed into the bulk band gap. Both the octopolar reconstruction and the hydroxylated surface of the {111} direction have states much closer to the bulk DOS, especially with respect to preserving the gap of unavailable states.

11.6 Conclusions

Our results show that although the polar catastrophe predicts infinite surface energies for unreconstructed polar surfaces, real systems undergo one of several compensation mechanisms to avoid the catastrophe. The compensation mechanism exhibited will depend on the environmental conditions. Adsorption of a foreign species, such as the adsorption of hydrogen for the rock-salt oxides studied here, is a favorable compensation mechanism which leaves the {111} surface...
Figure 11.5: The density of electronic states is shown for surface terminations of CaO, and compared to the Bulk density of states. Higher energy surfaces, such as the bare unreconstructed \{111\} surface show states pushed into the band gap.

smooth, provided that there is ample hydrogen in the growth environment. Even for unreconstructed, undecorated surfaces, the electronic charge will redistribute in order to overcome the polarity. This electronic reconstruction is costly in MgO and CaO, because the reconstruction pushes electronic states into the band-gap.
Chapter 12

Understanding the heterogeneous interface between two polar materials.

This chapter presents my DFT calculations of the atomic and electronic structure of the heterogeneous interface of an MgCaO alloy grown on polar \{0001\} GaN. This work is the result of a collaboration between our group and the thin-film oxides group of J.P. Maria, the Collazo and Sitar wide-bandgaps laboratory, and the J.M. LeBeau STEM group at NCSU, and the Center for Nanophase Materials Science at Oak Ridge National Laboratory.

The full collaborative article has been submitted to *Advanced Functional Materials*.

I would like to thank my experimental collaborators: Elizabeth A. Paisley, Christopher T. Shelton, Michael D. Biegalski, Hans M. Christen, Mark D. Losego, and Jon-Paul Maria grew and characterized the oxide thin films; Seiji Mita, Ramón Collazo, and Zlatko Sitar grew and characterized the GaN substrate; and James M. LeBeau prepared the samples for STEM microscopy and captured the micrographs.
12.1 Introduction

The previous two chapters (Chapter 10 and Chapter 11) have discussed the nature of polar rock-salt surfaces MgO and CaO in order to explain the competition between polarity compensation mechanisms. The results presented previously indicate that the polar \{111\} surface can be stabilized by the use of a hydrogen surfactant with careful control of the environmental growth conditions. In order to maximally exploit novel property coupling at the interface, the nature of the interface between the two materials must be understood. In the present experiment, an alloy of Mg_{0.5}Ca_{0.5}O is chosen to minimize the lattice strain between the GaN substrate and the oxide epitaxial layer. Atomic force microscopy and RHEED demonstrated thick, layer-by-layer growth of the MCO using a hydrogen surfactant, similar to the results shown for MgO and CaO. Furthermore, STEM microscopy, shown in Figure 12.1, of the as-grown samples revealed a smooth, epitaxial interface on terraces of the GaN substrate, while a disordered region was present in the vicinity of a substrate step-edge. In order to explain the occurrence of this disorder,

![Figure 12.1: STEM micrograph of the GaN—MCO interface. The left and right sections show layer-by-layer epitaxial growth, while the region above the step-edge shows disorder.](image)

DFT calculations compared the stability of a number of interfaces both for thick MCO layers as well as layer-by-layer in the early stages of growth. The disordered region occurs because 1) the GaN substrate consists of c/2 step edges, where the terraces above and below the step are rotated 180° from each other and 2) calculations indicated that the \{111\} MCO layer takes only one orientation with respect to the underlying substrate symmetry. These two factors taken together indicate that the MCO grown above and below a step-edge will be rotated 180° from each other. The system responds to this rotational symmetry by the formation of a disordered
region above the step edge.

12.2 Computational Methodology

Two types of calculations were performed: “thick” films, which approximate an epitaxial growth of several layers, and “thin” films, which simulate the initial layer-by-layer growth of the oxide. The thick film interfaces were calculated with a slab morphology and five bilayers of GaN. The bottom surface of the GaN was passivated by pseudohydrogen to satiate dangling bonds. The active surface (top surface) of the GaN was capped with five bilayers of \{111\} MCO. Each plane in the supercell approximation contained two atoms, and the 50/50 MCO alloy ratio was satisfied using one Mg and one Ca atom in each cation plane. The outer plane of the MCO layer was terminated with hydrogen to stabilize the polar surface. A minimum of 20 Å of vacuum spacing is included between periodic representations of the cell, which is shown sufficient for convergence. For the thick films, four starting configurations of MCO on the Ga-polar, Ga-terminated GaN substrate were tested. The two configurations which had an oxygen-gallium interface (the two differed by their rotational orientation with respect to the GaN) were energetically favored compared to the cation-gallium interface. The differences between the rotational variants will be discussed further below. The second type of interface, the thin film, was generated by using the same GaN substrate and calculating the relaxed ionic positions of successive layers of MCO. For example, the first phase consisted of a monolayer of oxygen or a monolayer of hydroxyl groups, followed by a cation layer, etc. For each successive layer, several initial positions were tested to determine the preferred thermodynamic pathways. Atomic relaxations and total energies were calculated using VASP [Kresse and Hafner, 1993, Kresse and Hafner, 1994, Kresse and Furthmüller, 1996b] and a plane wave basis set using the generalized gradient approximation [Perdew et al., 1996a] in the embodiment of the PBEsol functional [Perdew et al., 2008, Perdew et al., 2009], using a kinetic energy cutoff of 1000 eV and a 3x3x1 Gamma-centered Monkhorst-Pack k-point mesh. [Monkhorst and Pack, 1976]

12.3 Rotational Variants

Given that the interface is a gallium-oxygen interface, there exist two possible rotational variants to the MCO orientation on GaN. These two possibilities are shown in Figure 12.2. The rotations can be best described by considering the underlying symmetry of the cation planes of the (0001) GaN substrate. These cation planes have an ABAB... stacking sequence. If the top gallium atom sits at a B site, our DFT calculations reveal that the first oxygen layer sits directly atop it. At this point, the first oxide cation plane can either sit in an A site or a C site. This orientation is called the R_1 orientation. Based on the stacking sequence of the \{111\} rocksalt system, the
next cation planes will sit at B, then C. The stacking sequence across the interface is then ABAB|ABC. Thus, in the R$_1$ configuration, the stacking sequence of the substrate is preserved across the interface and the new plane, sitting in the C site, does not appear until the third layer away from the interface. The R$_1$ configuration can also be thought of as preserving the dihedral angle of the underlying GaN into the first layer of the MCO, as indicated by the black line in the figure. The second configuration, called R$_2$, features the first cation layer sitting at the C site instead of the A site. The rocksalt stacking sequence then requires the next layers to sit the the B site, followed by the A site. The stacking sequence across an R$_2$ interface is then ABAB—CBA, providing an immediate break in the GaN symmetry at the interface. In addition, the GaN dihedral angle is broken across the interface.

### 12.4 Rotational Preference

Comparison of the relaxed atomic positions as determined by DFT to the STEM image of the as-grown sample, shown in Figure 12.3 reveals that the R$_1$ orientation is favored. The energetic preference for of a single rotational variant was evaluated by calculating the interface energy for both thick-film configurations. These calculations revealed that the R$_1$ interface is preferred by 60 mJ/m$^2$. This relatively small preference is expected, since the differences between the two configurations is only present in the third-nearest-neighbor interaction. The weak energetic preference does not, however, satisfactorily explain the experimental observation that only the R$_1$ configuration is seen.
Figure 12.3: The relaxed atomic coordinates of the MCO-GaN interface in the $R_1$ and $R_2$ orientations, superimposed upon a STEM micrograph of the interface. STEM image courtesy J.M. LeBeau.
12.5 Early Growth Preference

In order to determine why $R_1$ is preferred, the interface was studied in its thin-film configuration. The GaN substrate was capped with oxygen atop the surface Ga atoms, and the Mg/Ca atoms were placed at several locations, including above the A site and the C site. The energies of these configurations were calculated and used to construct the potential energy surface seen in Figure 12.4. These calculations reveal that there is a 500 mJ/m$^2$ energy difference between the favorable $R_1$ and unfavorable $R_2$ configurations, and, furthermore, that once the system finds the stable $R_1$ configuration, there is a 300 mJ/m$^2$ barrier to escaping this configuration. As the system grows thicker, the energetic benefit of the $R_1$ configuration is reduced, but every subsequent layer of growth increases the energetic barrier to rearrangement. Thus, our DFT calculations reveal the preference for the $R_1$ configuration is set during the first layer of cation growth, and is locked in by the continued growth.

![Figure 12.4: Potential energy surface of the interface between GaN and MCO during the first layer of growth.](image)

12.6 Consequences of a Step-Edge

The implications of a preference for a single rotational variant in the presence of a step-edge are illustrated in Figure 12.5. The step-edge on the GaN substrate is a half-unit cell step, or a so-called c/2 step. The stacking sequence of GaN in the (0001) direction leads to a 180° rotation between the GaN above and below the step. If the region above the step terminates in a B plane, the region below the step terminates in an A plane. The $R_1$ configuration is still preferred, but
Figure 12.5: The $R_1$ configuration shown above (left) and below a step edge.

The MCO epitaxial regions are also rotated 180°. Thus, the first cation plane above the step edge will flow into the cation plane below the edge because both are A planes. However, the next two planes differ. In the vicinity of the step edge, the system can overcome this difference either by forming a stacking fault in one of the regions, or by forming a disordered region where the dissimilar planes meet above the step edge. The experimental evidence indicates that the system prefers to create a disordered region, as shown in Figure 12.6.

Figure 12.6: Comparison of the atomic positions of the cations above and below a step edge, as predicted by DFT (left) to the STEM image of the as-grown sample.

12.7 Conclusions

The experimental observation of a region of disorder above a substrate step-edge is explained by the preference of the epitaxial \{111\} rocksalt oxide to adopt only one orientation with respect to the underlying GaN substrate. This rotational preference, along with the $c/2$ step
edge leads to a 180° rotation between the epitaxial regions on either side of the step edge. The rotational preference arises during the first layer of growth, where the system prefers to preserve the underlying stacking sequence of the substrate across the interface. This preference is then locked in as subsequent layers of the oxide grow. In the vicinity of a step edge, the system must choose between a disordered region and the formation of a stacking fault. Experimental evidence indicates that the disordered region is preferred, since no stacking faults are observed. Smooth, epitaxial growth across a wide area is then limited by the smoothness of the substrate. Widening the terraces in the substrate would reduce the number of step-edges and the frequency of disordered regions in the oxide layer.
Chapter 13

Conclusions and Future Work

13.1 Conclusions

Computation integrated with experiment has enabled an improved understanding of defects in nitride semiconductors, at oxide surfaces, and at the heterogeneous interface between nitrides and oxides.

The properties of point defects in nitride semiconductors was explored by extending standard DFT calculations using state-of-the-art hybrid functionals to more accurately capture the formation enthalpy and ionization energies of native and impurity defects. These methods were used to explain the long-standing question of the 4.7 eV deep UV absorption in bulk AlN substrates as arising from the unintentional incorporation of carbon during growth. The formation energies obtained via DFT were extended by solving the mass-balance equations to determine the equilibrium Fermi level and the ionization of defects, so that the interaction between point defects in the material could be captured. This method enabled the determination of a compensating donor-acceptor pair between carbon and the nitrogen vacancy. This method also was able to identify the activation energy of the oxygen and silicon donors and identify the presence of a previously unknown Si DX center in AlN.

Ab-initio thermodynamics were used to couple DFT results to experimental temperatures and pressures to examine the stability of oxide surfaces in contact with a gas phase. These results enabled the prediction of an experimental growth window where the high-energy \{111\} polar rocksalt oxide surface could be stabilized with the use of a hydrogen surfactant. This stabilization was understood within the framework of polarity compensation mechanisms. Several of these compensation mechanisms were compared in terms of their energetics and electronic structure.

The interface between a lattice-matched rocksalt oxide and a wurtzite nitride substrate was examined in terms of its preference to choose one rotational configuration. This rotational
invariance was explored using DFT to determine the energetics of the first stages of epitaxial growth and to show that based on the growth conditions, there is a strong preference for the rotational variant that preserves the substrate stacking sequence across the interface. The rotational invariance corresponds to a region of disorder that would occur in the vicinity of a c/2 step edge on the substrate. This disordered region was experimentally confirmed using HAADF-STEM imaging.

13.2 Future Work

The integrated computational and experimental framework that has been used in this dissertation to explore defects can be applied to other defects in similar materials.

13.2.1 Sodium Defects in Ammonothermal Growth

Ammonothermal growth uses a liquid sodium flux to support the growth of GaN, but defect incorporation prevents this type of growth from achieving optics-grade crystals. The presence of Na defects in the crystals may cause the detrimental properties directly, or the Na defects might cause the enhanced incorporation of other native and impurity defects. Understanding the role of Na defects in GaN may allow the tailoring of growth conditions such that the ammonothermal process can be used for large-scale nitrides growth.

13.2.2 Strain cracking in GaN

The presence of Si and Ge dopants in GaN has been shown to cause cracking in certain cases. Some groups have reported that only Ge causes this cracking, while others have shown that Si has the same effect. Whether the presence of these defects impedes dislocation motion, or whether the presence of these defects encourages the formation of vacancies is still unknown. DFT calculations formation energies and solutions to the mass balance equations can determine how the observation of strain-related cracking and the defect incorporation are related.

13.2.3 Void formation in AlN

The creation of UV-transparent AlN bulk crystals by the addition of Si and O has led to crystals with high void concentrations. The presence of these voids may be due to the increased likelihood of vacancy-complex formation. The formation energies of these complexes can be explored and compared for samples with low Si and O concentrations.
13.2.4 The MCO-GaN Interface under varying growth conditions

The nature of the dissimilar interface between MCO and GaN is complex because of the lack of symmetry between the crystals, the polarity of the two materials in the direction normal to the interface, and because of the different bonding character across the interface. The properties of this interface are expected to change as a function of the growth conditions. These changes can be explored by examining the layer by layer growth of the oxide on the nitride substrate using \textit{ab initio} thermodynamics.
REFERENCES


Geysermans et al., 2009 Geysermans, P., Finocchi, F., Goniakowski, J., Hacquart, R., and Jupille, J. (2009). Combination of (100), (110) and (111) facets in MgO crystals shapes from dry to wet environment. Physical chemistry chemical physics : PCCP, 11(13):2228–33.


136


138