Investigation of defect properties of Li in ZnO

A first-principles study of vibrational frequencies, transition rates and diffusion profiles of Li and Li-complexes in ZnO

Roger Wold

Master Thesis at the Department of Physics under the Program: “Materials, Energy and Nanotechnology”

UNIVERSITY OF OSLO

December 2013

2013

Investigation of defect properties of Li in ZnO - *A first-principles study of vibrational frequencies, transition rates and diffusion profiles of Li and Li-complexes in ZnO*

Roger Wold

http://www.duo.uio.no/

Print: Reprosentralen, University of Oslo

Abstract

ZnO is a promising material for modern technological applications. Due to its large direct bandgap (3.4 eV) and large exciton binding energy (60 meV), the compound is suggested as a material for optoelectronic devices, such as light emitting diodes, lasers etc. Moreover, the piezoelectric and pyroelectric properties make ZnO a promising candidate for sensors, actuators and energy generators. ZnO is also suggested for spintronics, and in other application areas in science like biotechnology and for medical applications. ZnO is a transparent wide bandgap semiconductor, and the electronic conductivity can be controlled by doping. External defects, however, can have both positive and negative effects on the material properties. Thus, understanding the defect and dopant properties in ZnO is vitally important for developing ZnO-based devices.

In this thesis, bulk ZnO has been analysed and effects due to defects/dopants in ZnO are considered. The calculations have been performed by first-principles atomistic modeling within the density functional theory (DFT), employing the local density approximation (LDA) as well as the generalized gradient approximation (GGA) to describe the effective potential in the Kohn-Sham equation.

The presence of three elements, hydrogen, lithium and nickel in ZnO is investigated individually, with a primarily focus on Li in ZnO. The vibrational frequencies have been calculated for H and Li, and the results are consistent with earlier reports. However, in this work it is found that the choice of the effective potential affects the vibrational frequencies. Calculations of the frequencies within the GGA show better accuracy relative to the experimental values.

Interstitial Li (Li_i) is most stable in octahedral position in ZnO. In this work the diffusion barriers of axial diffusion and basal diffusion have been calculated to $\Delta E_{a,\parallel}$ = 0.67 eV and $\Delta E_{a,\perp}$ = 0.72 eV, respectively. As a result, Li_i shows slightly anisotropic transition rates as the diffusion barriers are directional dependent. The corresponding diffusion coefficients are also carried out. Li_i is found to be mobile at room temperature and can easily diffuse in the material at normal conditions. Substitutional Li (Li_{Zn}) is suggested as an acceptor in ZnO to establish stable *p*-type ZnO. However, the presence of donors, like interstitial H (H_i) and interstitial Li, are compensating dopants making the *p*-type behaviour difficult. We find that Li_{Zn} can act as a trapping center for Li_i diffusion. In this work, the kick-out barrier for a Li_{Zn} – Li_i pair has been calculated to $\Delta E_{a,Li_{Zn}-Li_i} \leq 0.37$ eV, a significant decrease from the direct diffusion of Li_i. This indicates that if a Li_{Zn} is present, the Li_i will more likely diffuse through a kick-out mechanism than through a direct interstitial diffusion. Dependent on the Li concentration, the kick-out mechanism is expected to be present to large extent at room temperature. The kick-out mechanism where a Li_i atom kick out a host Zn atom has been investigated. However, the configuration of an interstitial Zn is unstable and the kick-out mechanism exhibits a large barrier. The kick-out mechanism is found to be reversible at normal conditions. The present study indicates that transport of Li_{Zn} needs to be assisted by the presence of interstitial Li atoms or by intrinsic Zn vacancies in the structure. The relaxed structure of the Li_{Zn} – Li_i pair and the Li_{Zn} – Zn_i pair has been investigated. In both cases, the formation energy of the two-defect pair were lower than the separated single-defects formation energy, implying formation of pairs, with the interstitial atoms located closer to the substitutional Li. The formation energy is calculated to be $\Delta H_f^{rel} = -1.62$ eV for the Li_{Zn} – Li_i pair relative to formation of the two defects separately, which strongly indicates the formation of a Li-pair in the structure.

Ni is a magnetic element and hence it provides magnetic properties in ZnO. Vibrational frequencies of substitutional Ni (Ni_{Zn}) have been studied as well as investigation of Ni-doped ZnO co-doped with Li_{Zn}. Li_{Zn} is most stable close to the substitutional Ni for the formation of a $Ni_{Zn} - Li_{Zn} - O$ cluster. The vibrational mode of Ni_{Zn} is found to be modified to higher frequencies by the presence of Li_{Zn}.

This thesis can lead to a better understanding of the defect properties of Li in ZnO. The slightly anisotropic transition rates indicate that interstitial diffusion is preferred through the axial direction for Li_i, and Li_i is found to be mobile at room temperature. However, it is suggested that Li_{Zn} acts as a trapping center for the Li_i and prevent further diffusion as a result of the low kick-out barrier for the $Li_{Zn} – Li_i$ pair.

Preface

This thesis has been carried out at the Structure Physics group and the LENS (Light and Electricity from Novel Semiconductors) group connected to the Department of Physics at UiO under the supervision of Clas Persson. Most of all, I want to thank my supervisor Clas Persson for comprehensive and useful discussions the last two years. Your course in FYS-MENA4111 woke my interest for theoretical modeling of nanomaterials which turned into this thesis. Your expertise and guidance through my master degree have been crucial for this thesis. I also give thanks to my co-supervisor Bengt Svensson. Although we did not have much to do with each other in daily life it was a pleasure being seated at the LENS' group with nice surroundings and a great office my MSc colleagues only could imagine. I want to give big thanks to Girma Alemgasha Gardew for sharing office with me the last two years. Your knowledge about both the theoretical and practical background of theoretical modeling has been valuable, and our discussions have been truly fruitful to me. In addition, our weekly table tennis games have been really fun and a nice way to change the focus of a stressful day. The rest of Clas' group who have been sharing office with me during different periods in the last two years also deserves thanks for a nice collaborative environment; Maofeng Dou, Dan Huang, Gustavo Baldissera and Rongzhen Chen. Also thanks to Mukesh Kumar for taking care of me during my short trip to Stockholm, and to our newest member at UiO Alexander Hupfer for some interesting dialogues about NEB calculations. Thanks to the LENS group and Klaus Magnus Johansen for discussing my results compared with experimental results and the willingness to collaborate. Through my bachelor and master studies I have taken many courses and worked together with many students. I want to thank all my colleagues in the MENA program I've worked with throughout the years. Thanks for our numerous valuable workshops and for maintaining my competitive spirit. Last, but not least, I want to thank my family and especially my wife, Christina Elise Wold, and our beloved son, Wilhelm Elis. Thanks for always believing in me and for your smiles and hugs that always meet me at the door frame.

Abbreviations

ALD – Atomic Layer Deposition

APW – Augmented Plane Wave

BOA – Born-Oppenheimer Approximation

BZ – Brillouin Zone

ccp – cubic close packed

cNEB – Climbing Nudged Elastic Band method

CVD – Chemical Vapour Deposition

DFT – Density Functional Theory

EB – Elastic band method

e.g. – exempli gratia (“for example”, common Latin abbreviation)

et al. – et alii (“and others”, common Latin abbreviation)

etc. – et cetera (”and so on”, common Latin abbreviation)

fcc – face centred cubic

FTIR – Fourier Transform InfraRed

GGA – General Gradient Approximation

hcp – hexagonal close packed

HDF – Hybrid Density Functional

HF – Hartree-Fock

i.e. – id est (“that is”, common Latin abbreviation)

IBZ – Irreducible Brillouin Zone

LAPW – Linearized Augmented Place Wave

LCAO – Linear Combination of Atomic Orbitals

LDA – Local Density Approximation

LENS - Light and Electricity from Novel Semiconductors

LSDA – Local Spin Density Approximation

LVM – Local Vibrational Modes

MENA – Materials, energy and nanotechnology

MEP – Minimum Energy Path

NEB – Nudged Elastic Band method

PBE – Perdew Burke Ernzerhof

PW – Plane Waves

RT – Room Temperature

SIMS – Secondary Ion Mass Spectrometry

TCO – Transparent Conducting Oxide

TF – Thomas Fermi

TST – Transition state theory

UiO – University of Oslo

VASP – Vienna Ab initio Simulation Package

VLS – Vapor-Liquid-Solid method

vs. – versus (“against”, common Latin abbreviation

Table of contents

1. Introduction

The importance of studying materials at atomic level has strongly increased the past decades through the introduction of nanotechnology, and the possibility to achieve new properties by manipulating materials at the atomic level. The main problem related to treatment of materials at the atomic level is the complexity and the large number of atoms even in nanomaterials, which leads to huge numerical problems. This is in quantum mechanics referred to as the many-body problem. One of the most common methods for solving the many-body problem is the density functional theory (DFT). DFT employs density functionals to reduce the many-body problem into a three-dimensional problem, and the theory has proven to be an impressive theory for calculating different properties of materials. It has even received international attention through the Nobel Prize in Chemistry awarded Walter Kohn in 1998 for "*his development of the density-functional theory*" (Nobelprize.org, 1998). DFT is the theory behind most atomistic models today, and the practical application of DFT is through numerous simulation packages and codes.

This thesis has used DFT to calculate mainly the vibrational frequencies and diffusion properties of impurities in ZnO. In this thesis the *Vienna Ab initio Simulation Package* (VASP) has been used for the DFT calculations (Hafner, 2008), but several other simulation packages are also available and applicable. This first chapter will contain a short introduction about the motivation and the prior art behind this study. Chapter 2 contains a detailed background study of semiconductors with a special focus on the present material, ZnO. At first, an introduction of electronic band structures and the basics in solid state physics is presented, before ZnO is presented both at the macroscopic level in a historical perspective, and at the atomic level along with its technological applications. In chapter 3 the theory behind DFT is presented, from the many-body problem to the Kohn-Sham equation. Unfortunately, the DFT is not accurate (to this date) and the available approximations will be presented and discussed. Following that, chapter 4 focuses on the details behind the practical approach of DFT. DFT is a general theory and many approximations and simplifications must be considered in order to perform calculations on realistic materials, and to make the computations efficient. It also includes the theory behind vibrational calculations and diffusion properties, which are the main topics in the present study. Chapter 5 is the results and discussion section and will include the results obtained from the present work. Three different impurities, H, Li and Ni in ZnO have been studied, with a primarily focus on Li. First, the host structure is calculated to obtain the most stable configuration, before impurities are introduced and the vibrational frequencies and the diffusion mechanisms are presented and discussed. In chapter 6 there will be a discussion and comparison of the major results from the result section, and how the present study can lead to a better understanding of primarily Li in ZnO. Chapter 6 will be concluded by a short summary and a brief discussion of further work from the present study. The two last chapters include the references and an appendix, respectively. In the appendix the proof of the two Hohenberg-Kohn theorems is presented, as the theorems lays the foundations of the DFT.

1.1 Motivation

During my bachelor studies I was more attracted to theoretical subjects like quantum physics rather than experimental subjects including laboratory work. Through my academic career I have always

enjoyed working with tough algebraic equations more than wearing a lab coat. Therefore, I searched for a theoretical study for my master studies. Being inspired by an introduction course in computational modeling of nanomaterials I realized that studying materials in a theoretical way would be something for me.

Zinc oxide has been a subject for investigation throughout the years. As a material with numerous applications, it has been relevant for many different fields in science. As depicted from figure 1, the number of publications considering ZnO is still increasing. It has been an interesting material since the beginning of the 20th century. By the introduction of nanoscience and nanotechnology, understanding the material at an atomic level can predict many of the material´s properties and give important information related to how the material reacts in a given environment. Knowledge about the atomic structure of a material has also made it possible to manipulate materials at atomic levels and thereby give various, desired properties.

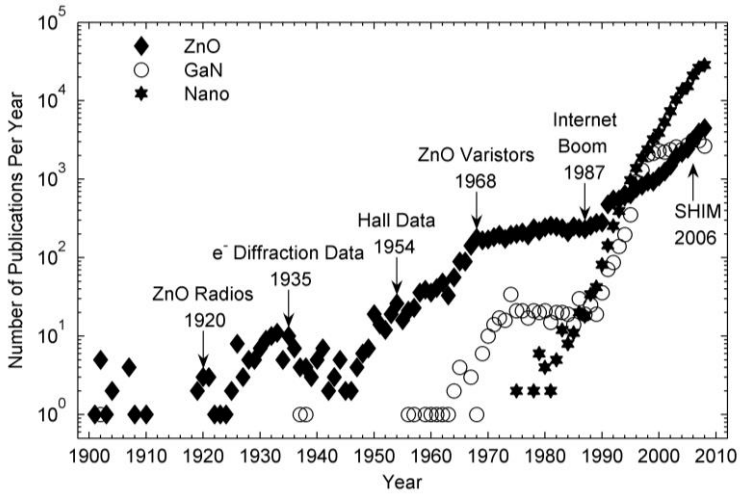

Figure 1: An overview of the number of publications for ZnO, GaN and "Nano" per year. ZnO has been an interesting material for decades, and the yearly publication number is still increasing. After the introduction of nanoscience, the number of publications has increased with a large amount the last 20 years. Figure adapted from ref. (sites.google.com).

Hydrogen and lithium, and to some extent nickel, are all common impurities in ZnO. Due to its small size and appearance in all kind of environments, H can easily be introduced in interstitial sites in ZnO. Li can be introduced during hydrothermal growth of ZnO or introduced intentionally as a dopant to achieve new properties. As Ni is a magnetic element, Ni-doped ZnO films have shown ferromagnetic behavior, although at very low temperatures, and are promising for magnetic transport properties for use in modern devices (Wakano et al., 2001). A comprehensive study of how H and Li behave in the structure is important to understand the native properties and the properties obtained through doping of ZnO. Vibrational frequencies are often of great interest, as they can be measured experimentally by spectroscopy, and thereby identify impurities and defects in a structure (Sholl and Steckel, 2009) p.114. Diffusion profiles can be an important tool to understand which configurations are the most stable for the impurities of interest. A diffusion profile will give information about the energy cost to move an impurity through energy barriers in the material and identify the most energetically favorable path between different positions. From this, one can also consider the possibility of the impurity to be trapped in the structure.

1.2 Prior art

From figure 1 one can see that ZnO has been an interesting compound for decades. Numerous articles and books have been published about this material, and the physical and chemical properties of bulk crystalline ZnO are well understood. After the introduction of nanotechnology, the possible use of ZnO in technological applications like optoelectronic devices has increased the focus on this material even further. To use ZnO in modern technology, knowledge about the atomic structure and the atomic properties are required. Although the perfect structure of ZnO is studied to a large extent, there are possible impurities and defects that might change the atomic properties significantly. The impurities might occur unintentionally, or one can dope the material in order to obtain new and improved properties. As different elements can occur in ZnO and affect the host structure, the focus on ZnO have been put in a new light. To understand how all dopants affect ZnO give rise to a new, advanced and huge research field which might result in exciting technological developments.

In this thesis hydrogen, lithium and nickel in ZnO have been investigated, with a primarily focus on Li. As the most common impurity in ZnO, H in ZnO has been studied precise and comprehensive in a number of papers. From reports, the different positions of H in ZnO with their formation energies are discussed and the vibrational frequencies are calculated both theoretically and measured experimentally (Limpijumnong and Zhang, 2005, Van de Walle, 2000, Lavrov et al., 2002). In the present study, the theoretical paper (Limpijumnong and Zhang, 2005) is used as a reference to obtain stable and precise values for the relaxed positions and the vibrational frequencies of H in ZnO. The paper uses the LDA, which is an old, but useful approximation, and the present study improved the calculated values by a more recent approximation, the GGA.

How Li behaves in ZnO is somewhat less described than hydrogen, but incorporating Li has still received much attention due to its possible technological applications and the possibility to make stable *p*-type ZnO. This thesis gives a precise description of the vibrational frequencies of Li in ZnO performed by theoretical calculations. The frequencies are calculated and discussed by Wardle (Wardle et al., 2005b), but the calculations are performed by another computational code, AIMPRO. Another recent paper, ref. (Vidya et al., 2012), contains a precise study of the vibrational frequencies obtained from the VASP code. This study therefore relates the present results with the two mentioned references. Another interesting topic that is treated in this paper is the diffusion mechanisms of Li and Li-complexes in ZnO. The diffusion mechanisms are briefly presented by ref. (Wardle et al., 2005b), but are treated more detailed in other papers, like ref. (Huang et al., 2009, Carvalho et al., 2009). The diffusion profiles and the activation barriers of interstitial Li are also investigated experimentally (Lander, 1960, Knutsen et al., 2013).

Ni in ZnO has been studied in the literature primarily because of its interesting magnetic properties and the possibility to yield ferromagnetism in ZnO. Previous studies indicate that Ni substitutes for Zn cations in ZnO (Yin et al., 2005). The relaxed structure of Ni_{Zn}-doped ZnO, as well as co-doped with Li_{Zn} have been investigated and compared with previous reports. Jayakumar et. al (Jayakumar et al., 2010) states that the co-doped Li are most stable in the presence of the substitutional Ni, forming small clusters of $Ni_{Zn} - Li_{Zn}$. The presence of Li_{Zn} in Ni_{Zn}-doped ZnO is shown to affect the ferromagnetic properties of ZnO. The magnetic coupling of Ni_{Zn} in ZnO co-doped with Li has been studied by Persson et. al, indicating a deviation from free-hole doping at higher Li concentrations (Persson et al., 2011). Experimental studies have reported a vibrational shift in the frequencies

related to the introduction of Ni in ZnO (Malaeru et al., 2012). In the present study, the vibrational modes are calculated as well as how Li_{Zn} affects the local vibrational modes of Ni in ZnO.

As many of the listed references to this study are recent papers, this indicates that the present study is highly relevant, at the edge of the ZnO research.

2. General theory

In this chapter some of the general theory of solid state physics will be presented, as a background for this thesis. A short preview of the concept of the unit cell and the reciprocal space followed by the theory behind electronic band structures and Bloch waves will begin this chapter. This is followed by a general introduction about wide bandgap semiconductors, their properties and applications, before a detailed description of the main material in this thesis, ZnO, is presented. The history of ZnO is presented, together with the basic material properties, its applications, and synthesis methods. After that, ZnO is described at atomic scale. This thesis will focus on atomic modeling of ZnO and impurities, so a description of the atomic structure of ZnO is important. Finally, a brief preview of the promising and interesting applications of ZnO in modern technological devices is presented.

2.1 Unit cell and Brillouin Zone

The first sections will give an introduction in solid-state physics, and is based primarily on C. Kittel: *Introduction to Solid State Physics* (Kittel, 2005)

All crystals are made of atoms in periodic order. The atoms may be arranged in a short-arranged or a long-arranged order with different symmetric properties. Many of the crystal's properties can be predicted by knowing the periodic order, and hereby, the atomic structure of the crystal. Whether the crystal is short- or long-arranged, it can always be divided into a repetitive structure which is repeated throughout the crystal. The repetitive arrangement in the crystal can be described by the unit cell, which is the building block of the crystal. The crystal is put together by numerous unit cells in all directions.

A Bravais lattice, as an infinite array of discrete lattice points, gives the periodicity of a crystal. In order to be periodic, the Bravais lattice is equivalent under a set of linear translation operations given by the lattice vector (Thijssen, 2007) p. 123

$$\boldsymbol{R} = n_1\boldsymbol{a_1} + n_2\boldsymbol{a_2} + n_3\boldsymbol{a_3} \ ,$$

where n_i are any integers, and $\boldsymbol{a_i}$ are the primitive vectors that span the lattice. Considering symmetry, it can be shown that there are 14 possible Bravais lattice in the three-dimensional space. A crystal is created by a periodic arrangement of one or more atoms repeated at every lattice point in the Bravais lattice creating a three-dimensional crystal structure. The repeating unit is called the basis, and a lattice plus the basis at each lattice site yields the crystal structure. Therefore, the crystal structure is equivalent within any choice of the lattice vector $\boldsymbol{R}$.

The unit cell can be described and defined by several means, but a common technique for making a primitive unit cell is by a technique from Wigner and Seitz. The primitive cell is the smallest possible unit cell and contains only one lattice point. The Wigner-Seitz cell is created by first choosing a lattice point and then draw a vector from the chosen lattice point to all its nearest neighbors, as can be seen in figure 2. Furthermore, one constructs a plane that is perpendicular to and passes through the midpoint of the lines just drawn. From this, the generated enclosed area is defined as the Wigner-Seitz cell.

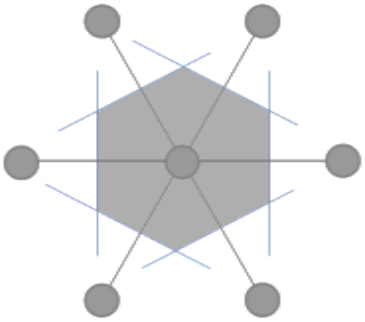

Figure 2: A schematic presentation of how to create a two-dimensional Wigner-Seitz cell from a hexagonal close-packed (hcp) structure. The grey enclosed area indicates the Wigner Seitz cell.

For several physical problems, e.g. in diffraction, where the reflecting wave is the Fourier transform of the crystal lattice, there might be difficult to treat the problem in the real space. Therefore, there exists a reciprocal space which is defined as the reciprocal real space; namely the Fourier transform of the crystal lattice. The reciprocal space is also called the k-space or the momentum space. In periodic crystals, the potentials and the wavefunctions are described as periodic functions. As a periodic function can be written as a Fourier series of the form $f(\boldsymbol{r}) = \sum_{\boldsymbol{G}} f_{\boldsymbol{G}}\, e^{i\boldsymbol{G}\cdot\boldsymbol{r}}$ (Thijssen, 2007) p. 125, and when the Fourier series are considered in the reciprocal space, this may simplify the calculations. This is described more detailed in section 2.3.

A compact mathematical formalism of the reciprocal space is given as

$$[\boldsymbol{b_1}\ \boldsymbol{b_2}\ \boldsymbol{b_3}]^T = 2\pi[\boldsymbol{a_1}\ \boldsymbol{a_2}\ \boldsymbol{a_3}]^{-1}\,.$$

The vectors $\boldsymbol{b_i}$ are the primitive vectors of the reciprocal space, while $\boldsymbol{a_i}$ are the primitive vectors of the three-dimensional real space. The term 2π is disregarded in crystallography, but is included in structure physics because of the periodic manner of the crystal structures.

The three reciprocal primitive vectors are defined as follow (Thijssen, 2007) p. 124

$$\boldsymbol{b_1} = 2\pi \frac{\boldsymbol{a_2} \times \boldsymbol{a_3}}{\boldsymbol{a_1} \cdot (\boldsymbol{a_2} \times \boldsymbol{a_3})}$$

$$\boldsymbol{b_2} = 2\pi \frac{\boldsymbol{a_3} \times \boldsymbol{a_1}}{\boldsymbol{a_1} \cdot (\boldsymbol{a_2} \times \boldsymbol{a_3})}$$

$$\boldsymbol{b_3} = 2\pi \frac{\boldsymbol{a_1} \times \boldsymbol{a_2}}{\boldsymbol{a_1} \cdot (\boldsymbol{a_2} \times \boldsymbol{a_3})}$$

Similar to the lattice vector in the real space, the reciprocal lattice vector in terms of the primitive vectors in the reciprocal space can be defined by

$$\boldsymbol{G} = h\boldsymbol{b_1} + k\boldsymbol{b_2} + l\boldsymbol{b_3}$$

The integers *h*, *k*, *l* refers to the vector coordinates (*h*, *k*, *l*) of the reciprocal space described by Miller indices (Miller, 1839). The vector coordinate (*h*, *k*, *l*) in the reciprocal space refers to a plane in the real space which is the normal vector to the [*h*, *k*, *l*] direction in the real space.

As in the real space, one can define a primitive unit cell also in the reciprocal space. Although not so important in the real space, the Wigner-Seitz cell in the reciprocal lattice in k-space is called the Brillouin zone, which plays an important role in solid-state physics. The modulus of the reciprocal lattice vectors is 2π/a, and is related to the wave vector $\boldsymbol{k}$. The first Brillouin zone therefore has k-values defined by $-\frac{\pi}{a} < k < \frac{\pi}{a}$. All important information about the crystal, i.e. crystal momentum and energies, can be described within the first Brillouin zone, as the calculated values for higher zones can be mapped into the first Brillouin zone.

2.2 Electronic band structure

The atom can be divided in two parts, the nucleus and the surrounding electrons. Electrons are divided into core electrons which are tightly bounded to the nucleus and valence electron which are the outermost electrons. In this thesis, the ion core refers to the inner part of the atom, thus the nucleus and the core electrons. The valence electrons in the material are responsible for most of the material's properties like bonding structure and electronic properties, so a description of how the valence electrons behave are of high importance. The simplest way to describe the valence electrons is by an assumption that the valence electrons are completely decoupled from their ions and behaving like a free electron gas. In addition, one assumes that the valence electrons are non-interacting. The wavefunction and the energy values for a free electron can be solved by the time-independent Schrödinger equation for a single particle with $V(\boldsymbol{r}) = 0$. The wavefunction in the free electron model is given as $\psi_k(\boldsymbol{r}) = Ae^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ where A is a normalization factor and $\boldsymbol{k}$ is the wave vector related to the wave number k. The wave vector is coupled to the momentum for a free electron by the relation $\boldsymbol{p} = \hbar\boldsymbol{k}$, where $\boldsymbol{p}$ is the momentum of the free particle. The energy of a free particle is described by

$$E = \frac{\boldsymbol{p}^2}{2m_e} = \frac{\hbar^2\boldsymbol{k}^2}{2m_e} \ .$$

A wave described in this manner is referred to as a plane wave. As the shape of the wavefunction implies, free valence electrons propagate through the crystal as a periodic wave. Despite its simplicity, the free electron model has proven to be a good model for metals to describe i.e. the electrical and thermal conductivity and the electronic density of states. However, electrons are not completely free from their nucleus and one should include an interaction with the ion cores. The nearly free electron model, or more general, the empty lattice approximation includes a weak perturbation by introducing a small periodic potential to model the interaction between valence electrons and the inert ion cores. The time-independent Schrödinger equation for a free electron in a weak periodic potential described by $V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R})$ is

$$H\psi(\boldsymbol{r}) = \left[-\frac{\hbar^2\nabla^2}{2m_e} + V(\boldsymbol{r})\right]\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) \ ,$$

where the solutions can be written as

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}u_{n\boldsymbol{k}}(\boldsymbol{r}) \ .$$

Wavefunctions written in this way are called Bloch waves, derived by F. Bloch (Bloch, 1929). Electrons in a weak periodic potential can therefore be written as a product between a travelling

plane wave and a periodic function $u_{n\boldsymbol{k}}(\boldsymbol{r})$, where the latter part has the same periodicity as the crystal. Bloch functions will be discussed somewhat more detailed in the next section. Considering the energy eigenvalues, these are found as a function of the wave vector $\boldsymbol{k}$, $E(\boldsymbol{k})$. The energy states are occupied in continuously energy bands. However, energies between these energy bands are found to be forbidden, resulting in bandgaps. The reason why some energies are not allowed can be explained by Bragg reflections from interaction between valence electrons and ion cores in the crystal, or by further treatment of the empty lattice approximation by solving the central equation. The wavefunctions at $k = \pm\frac{\pi}{a}$ cannot be described by travelling waves, but by standing waves which yields two different values for the wavefunctions and a bandgap occur (Kittel, 2005) ch. 7. Another approach to explain the origin of bandgaps is to use a linear combination of the atomic orbitals (LCAO). Electrons in atoms are allowed to obtain certain discrete energy states in the orbit around the nucleus. When adding multiple numbers of atoms together the electrons occupy energy levels close to each other and in accordance to Pauli Exclusion Principle. In real materials, the size of a crystal is in terms of Avogadro's number, $\sim 10^{23}$ which results in a closely spaced region of energies that can be treated as a continuous band. However, there are energy levels between these bands that are not available to be occupied by the electrons which are called bandgaps, illustrated in figure 3.

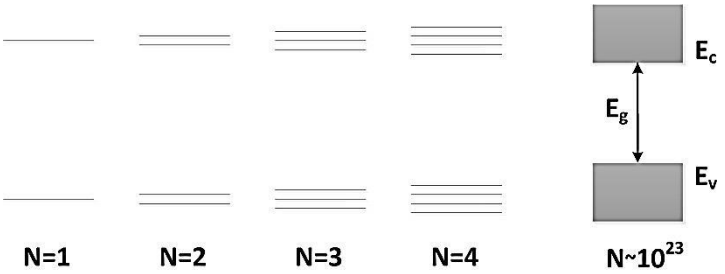

Figure 3: A schematic description of a crystal built atom by atom in the Linear Combination of Atomic Orbitals (LCAO) approach. Two states cannot obtain the same quantum state, and the energy levels are separated by a finite value for the increasing number of atoms. For large number of atoms, the energy levels are separated by infinite small steps and can be seen as a continuous band. However, certain regions are not allowed and are termed energy bandgaps, E_g. E_v is the valence band and E_c is the conduction band.

For energy bands near the ion cores (the lowest energy bands) the energy levels are narrow as they are more localized to the cores. The higher energy bands are delocalized and more wide-spread through the crystal leading to a broad energy range of allowed energy levels. As a result the bandgaps between the lower energy bands (i.e. core bands) are larger than for the higher energy bands (i.e. valence bands) where the bandgaps eventually disappears (i.e. for conduction bands). As the electrons search for the lowest energy, they will occupy energy states in bands from the lower according to the Pauli Exclusion Principle. In the ground state, at 0 K, the electrons will be filled in the lowest possible energy configuration. The highest occupied band in terms of energy is called the valence band and the lowest unoccupied band is called the conduction band. The bandgap therefore refers to the energy difference between the lowest conduction band and the topmost valence band. The bandgap is the energy required to jump from the occupied valence band to the free conduction band.

A material's band structure can give information whether it is a metal, insulator or semiconductor. Metals are electronic conductors and electrons must therefore be free to move in the band. If the valence band is partly filled, only a small thermal contribution (at finite temperatures) will excite the

electron to a new higher energy state and hence the material is a conducting metal. The excited electrons are termed charge carriers as they are responsible for the conductivity in the material. If the valence band is completely filled, the electrons must be excited through the bandgap in order to conduct current. However, the size of the bandgap determines whether a material easily can excite electrons and be an electronic conductor. Insulators are non-conducting materials and semiconductors can conduct current in certain environments. Therefore, materials that have a large bandgap are termed insulators, and materials with smaller bandgaps are defined as semiconductors. The distinction between insulators and semiconductors is somewhat undefined and blurry. The distinction is given in separate ways by different authors. In general, a semiconductor conducts current as a result of somewhat low energetic excitations i.e. thermal excitations at room temperatures while insulators cannot thermally excite electrons under normal conditions. The distinction then corresponds to bandgaps ~4 eV (Strehlow and Cook, 1973).

In semiconductors, one separates between direct and indirect semiconductors. The energy $E(\boldsymbol{k})$ is a function of the wave vector $\boldsymbol{k}$ called the dispersion relation, and the lowest conduction band might be at a different wave vector than the highest valence band. In order to excite these electrons a change of wave vector, hence a change in the electron's momentum, is required. A semiconductor which has the same wave vector in the lowest unoccupied state as the highest occupied state is termed a direct semiconductor, and if a change in the wave vector is needed it is called an indirect semiconductor, see figure 4. Thermal excitations occur through the presence of an incoming photon. In an indirect semiconductor, the process needs to be assisted through a phonon with the exact wave vector, a phonon-assisted process. Excitations in an indirect semiconductor are somewhat more difficult than in direct semiconductors as an incoming photon of a specific energy needs to be assisted by a phonon of a specific energy to transfer the electron to the right wave vector. An indirect semiconductor includes thus more particles in the process and is therefore less probable to excite.

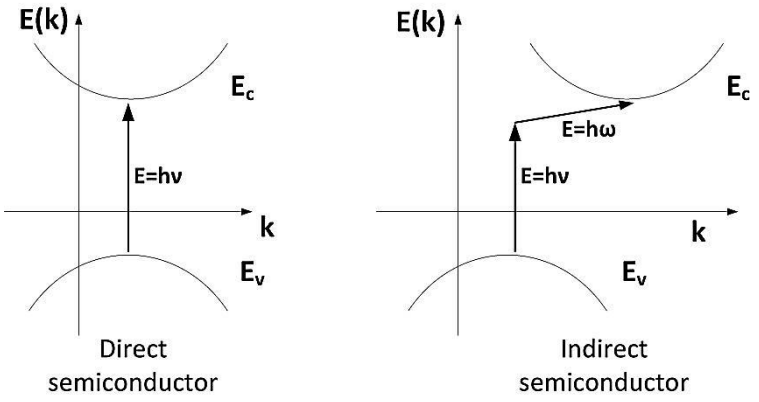

Figure 4: Left: A schematic presentation of a direct semiconductor. The highest occupied state and the lowest unoccupied state has the same momentum, hence electrons can be excited through thermal excitations of an incoming photon with energy $E = h\nu$ that equals the bandgap. Right: A schematic presentation of an indirect semiconductor. A change in momentum is required for the excitation across the bandgap. The incoming photon is assisted by a phonon with frequency ω, in order to change the momentum.

Excitations occur naturally in semiconductors when the thermal energy is higher than the bandgap. Semiconductors which perform electrical conductivity without any external dopants and only through excitations in the pure material are called intrinsic semiconductors. There is a probability that above absolute zero some of the electrons are excited to the conduction band, and the probability increases with temperature. When an electron excites across the bandgap it leaves an

electron vacancy behind in the valence band. An electron vacancy in the electronic band structure is called a hole, and for intrinsic semiconductors the number of electrons in the conduction band is equal to the number of holes in the valence band, $n = p$. Under appliance of voltage, intrinsic semiconductors can provide a small current caused by both electrons and holes as charge carriers. As mentioned above, the distribution of electrons in the band structure searches to achieve the lowest possible energy distribution. Electrons are distributed by the Fermi-Dirac distribution for identical fermions given by

$$f(\varepsilon_i) = \frac{1}{e^{(\varepsilon_i - \mu)/k_B T} + 1} \quad ,$$

Here, $f(\varepsilon_i)$ is the probability to find the state i with the given energy ε_i occupied at the absolute temperature T. k_B is the Boltzmann's constant and μ is the chemical potential, often called the Fermi level, E_F, defined as the energy level where the probability to find the state occupied equals ½. At $T = 0$ K, the distribution function is a step function where all the states below the Fermi level are occupied and none of the electrons are excited, in accordance to the requirement of minimization of energy. At finite and increasing temperatures, more and more electrons are excited through the bandgap by thermal excitations and motions in the crystal and the probability function is smeared out. The probability function is symmetric at all temperatures as the excited electrons leave holes behind, and the probability of finding a hole in the energy state ε_i at the temperature T is given as $1 - f(\varepsilon_i)$. The charge carriers are the electrons in the conduction band and the holes in the valence band. The number of charge carriers can be found as integration over the density of electronic states multiplied with the distribution functions;

$$n = \int_{E_c}^{\infty} f(\varepsilon_i) N(\varepsilon_i) d\varepsilon_i \quad ,$$

$$p = \int_{-\infty}^{E_v} \big(1 - f(\varepsilon_i)\big) N(\varepsilon_i) d\varepsilon_i \, ,$$

where $N(\varepsilon_i)$ is the density of states at the given energy, ε_i. The temperatures of interest are regions where $\varepsilon_i - \mu \gg k_B T$ as only few electrons are able to pass the energy gap. By that, one assumes that the Fermi level is well below the conduction band edge. For large temperatures a large amount of electrons can excite through the bandgap and give large charge carrier concentration and metal-like behavior. By the assumption that $\varepsilon_i - \mu \gg k_B T$, the distribution function reduces to $f(\varepsilon_i) \cong e^{-(\varepsilon_i - \mu)/k_B T}$. The integrations given above can therefore be calculated in the low-temperature range as

$$n = \int_{E_c}^{\infty} f(\varepsilon_i) N(\varepsilon_i) d\varepsilon_i \cong N_c e^{-\frac{(E_c - \mu)}{k_B T}} \quad ,$$

$$p = \int_{-\infty}^{E_v} \big(1 - f(\varepsilon_i)\big) N(\varepsilon_i) d\varepsilon_i \cong N_v e^{\frac{(E_v - \mu)}{k_B T}} \, .$$

Here, N_c is the effective density of states, which is the number of states per unit volume at the conduction band edge, and similarly for N_v as the the number of states per unit volume at the valence band edge. As the number of holes and electrons are equal in an intrinsic semiconductor $n = p = n_i$, one obtains the relation called the law of mass action (Dilli, 2008)

$$np = n_i^2 \ ,$$

where n_i is called the intrinsic carrier concentration. By combining the previous equations one can obtain a relation for the intrinsic carrier concentration in the low-temperature region

$$n_i = \sqrt{np} \cong \sqrt{N_c e^{-\frac{(E_c-\mu)}{k_B T}} N_v e^{\frac{(E_v-\mu)}{k_B T}}} = \sqrt{N_c N_v e^{\frac{-(E_c-E_v)}{k_B T}}} = \sqrt{N_c N_v} e^{-\frac{E_g}{2k_B T}} \ ,$$

where E_g is the energy of the bandgap. Therefore, the intrinsic carrier concentration is independent of the Fermi level and only dependent on the size of the bandgap and the temperature. The bandgap of Si is 1.1 eV (Chiang and Himpsel) and the intrinsic carrier concentration in crystalline Si at room temperature is $n_i \sim 10^{10}$ cm^{-3} (Sproul and Green, 1991) which implies that a fraction of only one electron per 10^{10} silicon atoms will excite through the bandgap at room temperature.

However, the carrier concentration can be enhanced by defects related to external dopants. An extrinsic semiconductor is a doped semiconductor. There are two types of extrinsic semiconductors; *n*-type and *p*-type semiconductors. If the dopant element has more valence electrons than the host structure the dopant introduce excessive electrons which are loosely bound to the ion core. The dopant can then easily be ionized releasing the extra electrons to delocalize in the crystal. For large number of dopants, many electrons are delocalized and act as charge carriers in the structure. A dopant that has more electrons than the host element is referred to as a donor, as it donates electrons to the crystal. In terms of the electronic band structure, a donor introduces electrons in a new occupied energy level in the crystal´s bandgap close to the conduction band. For a semiconductor to conduct, the electrons in the donor state must be excited to the conduction band. However, this can easily happen through a small energy contribution since the ionization energy are much lower than the intrinsic bandgap energy, as indicated in figure 5. If the semiconductor is dominated by electrons as the major charge carriers it is assigned an *n*-type semiconductor. In an *n*-type semiconductor, the Fermi energy is increased due to the occupied donor state compared to an intrinsic semiconductor. A common example of an *n*-type semiconductor is phosphorus-doped silicon. Phosphorus contains five valence electrons compared to four of silicon, and therefore phosphorus introduces one extra delocalized electron that can act as a charge carrier.

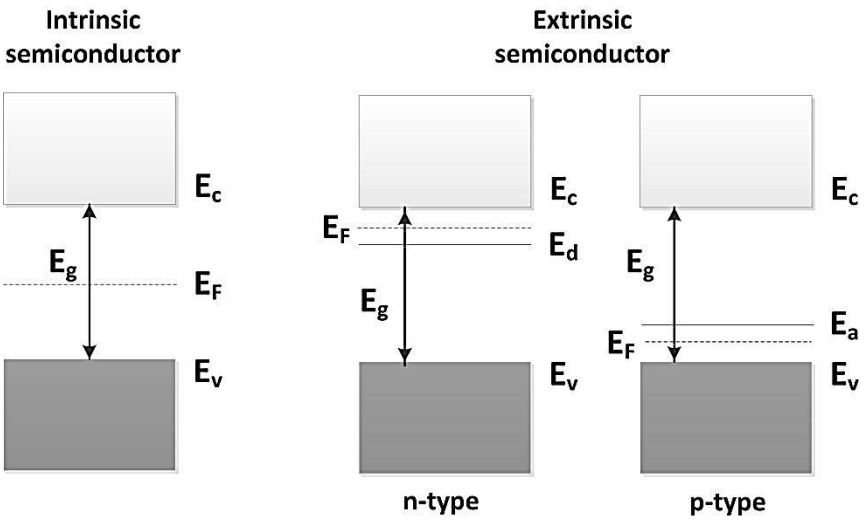

Figure 5: Band structure of an intrinsic (left) and an *n*-type and a *p*-type extrinsic semiconductor (right). E_c and E_v is the conduction band and the valence band, respectively. E_g is the bandgap, and the dashed line assigns the Fermi level, E_F. An *n*-type semiconductor introduces an occupied donor state, E_d, close to the conduction band. This results in an increase of the Fermi level. A *p*-type semiconductor introduces an unoccupied acceptor state, E_a, close to the valence band. This results in a decrease of the Fermi level.

On the other hand, if the dopant atom has fewer electrons than the host structure, holes are introduced in the structure free to move through the crystal, similarly to the delocalized electrons in *n*-type semiconductors. Dopants that have fewer electrons than the host element are referred to as acceptors. Acceptors introduce a new unoccupied energy level in the bandgap close to the occupied valence band. Only small excitations are therefore required to make it a conductor. The acceptors then accept electrons to occupy the introduced acceptor state. As a result, the Fermi level is lowered than in the case of an intrinsic semiconductor. When the semiconductors are dominated by holes as the major charge carriers it is assigned a *p*-type semiconductor. Boron-doped silicon is an example of a *p*-type semiconductor as boron is a trivalent element and introduces holes to the silicon crystal. One divides between shallow and deep doping levels. A deep acceptor indicates that the acceptor level is introduced far from the valence band edge, where a shallow acceptor level is placed closer to the band edge. A shallow acceptor level is therefore desirable from doping, as it requires smaller excitation energy. The similar terms are used for donors, where a shallow donor introduces a donor level close to the conduction band edge.

The carrier concentration of electrons in doped semiconductors can be found in terms of the relation

$$np = n_i^2 \Rightarrow n = \frac{n_i^2}{p}, p = \frac{n_i^2}{n}.$$

In an *n*-type semiconductor with only donor atoms assuming that all dopant are ionized, electron and hole concentration in the low-temperature region are given as (Dilli, 2008)

$$n = N_d \, , p = \frac{n_i^2}{N_d},$$

and for a *p*-type semiconductor

$$p = N_a \, , n = \frac{n_i^2}{N_a}.$$

Here, N_d is the donor concentration and N_a is the acceptor concentration. Normal doping concentrations in silicon and other semiconductors are in the range $10^{15} - 10^{18}$ cm^{-3}, so in the case of silicon, the intrinsic charge carrier concentration for a normal doping concentration will be $\sim 10^5$ cm^{-3} or lower.

2.3 Bloch waves

As been discussed, crystals are organized by atoms in periodic order. With atoms placed in this manner, the Schrödinger equation needs to be solved with a lattice-periodic potential, i.e. $V(\boldsymbol{r} + \boldsymbol{R}) = V(\boldsymbol{r})$. The wavefunction of an electron placed in a periodic potential can be described by a Bloch wave, after F. Bloch (Bloch, 1929). This is described by Bloch's theorem, restated by Kittel in "Introduction to solid state physics"(Kittel, 2005): "The eigenfunctions of the wave equation for a periodic potential are the product of a plane wave $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ times a function $u_{n\boldsymbol{k}}(\boldsymbol{r})$ with the periodicity of the crystal lattice." Therefore, the wavefunction for a particle in a periodic potential is given as

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_{n\boldsymbol{k}}(\boldsymbol{r}) \ ,$$

where $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ describes the behavior of a free electron in terms of a plane wavefunction. $u_{nk}(\boldsymbol{r})$ is a periodic function which to some extent describes the behavior of the atomic orbitals in the crystal lattice and has the same periodicity as the crystal lattice;

$$u_{n\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = u_{n\boldsymbol{k}}(\boldsymbol{r}) \ .$$

By that, it follows that the periodic function will be invariant under the translation of a lattice vector, $\boldsymbol{R}$. The subscript $n\boldsymbol{k}$ refers to two different quantum numbers. A $\boldsymbol{k}$-state is a state in the reciprocal space. $\boldsymbol{k}$ is often referred to as the wavenumber from the expression $k = 2\pi/\lambda$ and multiplied by Planck's constant, $\hbar\boldsymbol{k}$ is known as the momentum for a free particle. The dimension of k is reciprocal length, so it is coupled with the frequency. n refers to the index number of the electronic energy bands (Persson, 2012). When combining the two equations above, one gets the following property for the wavefunction

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = e^{i\boldsymbol{k}\cdot(\boldsymbol{r}+\boldsymbol{R})} u_{n\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}} u_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \psi_{n\boldsymbol{k}}(\boldsymbol{r}) \ .$$

The wavefunction are similar within a phase shift under the translation of a lattice vector $\boldsymbol{R}$. However, the probability density is unchanged under translation of $\boldsymbol{R}$ as the plane waves cancel out.

$$|\psi_{n\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R})|^2 = |\psi_{n\boldsymbol{k}}(\boldsymbol{r})|^2 \ .$$

In the reciprocal space, the periodic function can be described as a Fourier sum $u_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} u_{n\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}$, and the wavefunction can be then be expressed by (Thijssen, 2007) p. 125

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} u_{n,\boldsymbol{k}+\boldsymbol{G}} e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} \ ,$$

where $\boldsymbol{G}$ runs over all lattice vectors. The advantage of calculations in the reciprocal space is to reduce the problem from calculating an infinite number of wavefunctions, to a finite number of

wavefunctions for a finite number of reciprocal lattice vectors in the first Brillouin zone as the highest terms in the summation are often neglected. The last $\boldsymbol{G}$ included in the summation is referred to as $\boldsymbol{G}_{\text{cut}}$ and corresponds to a given energy cutoff. This is described in detail in section 4.2.2. By some algebra one can obtain an expression for $\psi_{n,\boldsymbol{k}+\boldsymbol{G}}(\boldsymbol{r})$ (Kittel, 2005) p. 173, and for simplification, $\boldsymbol{k}' \equiv \boldsymbol{k} + \boldsymbol{G}$.

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \psi_{n,\boldsymbol{k}'-\boldsymbol{G}}(\boldsymbol{r}) = e^{i(\boldsymbol{k}'-\boldsymbol{G})\cdot\boldsymbol{r}} u_{n,\boldsymbol{k}'-\boldsymbol{G}}(\boldsymbol{r}) = e^{i\boldsymbol{k}'\cdot\boldsymbol{r}} e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} u_{n,\boldsymbol{k}'-\boldsymbol{G}}(\boldsymbol{r}) \quad .$$

The phase factor shifts the Bloch periodic part and one can define $u_{n\boldsymbol{k}'}(\boldsymbol{r}) = e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} u_{n,\boldsymbol{k}'-\boldsymbol{G}}(\boldsymbol{r})$ (Kittel, 2005) p. 173. Inserted in the equation above one obtains an expression for $\psi_{n,\boldsymbol{k}+\boldsymbol{G}}(\boldsymbol{r})$.

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}'\cdot\boldsymbol{r}} u_{n\boldsymbol{k}'}(\boldsymbol{r}) = \psi_{n\boldsymbol{k}'}(\boldsymbol{r}) = \psi_{n,\boldsymbol{k}+\boldsymbol{G}}(\boldsymbol{r})$$

The wavefunction is therefore invariant within a reciprocal lattice vector $\boldsymbol{G}$. The crystal momentum related to the crystal outside the first Brillouin zone can therefore be mapped into the first Brillouin zone without loss of information. A similar expression can be obtained for the energy, as

$$E(\boldsymbol{k}) = E(\boldsymbol{k} + \boldsymbol{G}) \quad .$$

These results underpin the fact that the first Brillouin zone yields information about the whole structure, as results from higher Brillouin zones can be mapped into the first Brillouin zone by the relations given above.

2.4 Wide bandgap semiconductors

Wide bandgap semiconductors have paid great attention the last years due to the possible use in electronic, optoelectronic and power devices. Wide bandgap semiconductors are useful for applications involving stability at high temperatures and high voltages (Tolbert et al., 2003). Larger bandgaps imply higher electric breakdown fields, and the good electrical and thermal properties of wide bandgap materials overcome silicon by large amount, which is the leading material in technological devices today. Wide bandgap semiconductors show promising improvement in cost, size and lifetime compared to the existing technology. The threshold where bandgap is considered as large is somewhat unclear, and depends on the application. Wide bandgap semiconductors are defined to be substantially larger than commonly used semiconductors, and some define bandgap energies larger than 2 eV as wide (Takahashi et al., 2007) p. 2. Absorption and emission spectra are directly related to the size of the bandgap (Takahashi et al., 2007) p. 1. Therefore, semiconductors of this type can emit wavelengths in the visible spectrum from green/blue to the UV spectral range and hence the special interest for use in optoelectronic devices as light-emitting diodes (LEDs), lasers and optical sources for electronic devices. The bandgap energies in wide bandgap semiconductors correspond to wavelengths in the range 250 nm to 400 nm. As the human eye can observe wavelengths between 380 nm and 750 nm (Starr et al., 2010) p. 94, photons with wavelengths in the visible spectrum are not absorbed by these semiconductors and wide bandgap semiconductors are often transparent. Wide bandgap semiconductors consist mainly of three groups; group III-nitrides like GaN, group II-oxides like ZnO and group II-chalcogenides like ZnSe (Takahashi et al., 2007) p. 2. The bandgap is dependent of atomic sizes and electronegativity, and light and electronegative elements found in the upper range of the periodic table are more likely to be wide bandgap semiconductors. As group II-oxides have large electronegativity difference they show strong ionic

bonding leading to large bandgap energies (Takahashi et al., 2007) p. 3. However, the strong bonding in II-VI semiconductors also has some disadvantages in terms of crystal growth. For use in optoelectronic devices there is a need for high-purity single crystals, as the properties are drastically affected by defects and impurities, which is difficult to obtain through the strong forces between the atoms. Another issue related to the use in modern devices is the lack of conductivity control related to doping. The wide bandgap semiconductors often exhibit one kind of conductivity, either *n*-type or *p*-type depending on the material, due to native defects of donor or acceptor character that naturally occurs in these materials. Wide bandgap semiconductors single crystals for technological devices can be grown epitaxial, which means that the crystal are grown in the same direction as the underlying layers, to create thin films, or by bulk crystal growth through chemical or physical vapor transport (Kasap and Capper, 2006) p. 339. As wide bandgap semiconductors have high melting points, which are positive for the applications, the crystal growth of these materials from melt are somewhat difficult due to the high vapor pressure at the melting points. However, growth is more easily performed in the vapor phase (Kasap and Capper, 2006) p. 326.

2.5 Material properties of ZnO

ZnO is a II-VI semiconductor with a wide (~3.4 eV) direct bandgap (Pearton et al., 2005). The compound is transparent, has good electron mobility, and has thus several interesting application areas, ranging from windows to LEDs and nanotransistors. However, ZnO is also an interesting compound for several other non-technological areas, from hygiene products to rubber.

ZnO has been a popular compound throughout the years due to its large width in applications areas, however mostly used in medicinal ointments and in paintings. One of the earliest known applications of ZnO is “pushpanjan” which is mentioned in an Indian medical text dated around 500 BC (Craddock, 1998) p. 27. “Pushpanjan” was a salve for eyes and open wounds. ZnO has also played an important role in the brass production during the first centuries. It is assumed that the Romans reduced ZnO to zinc gas that dissolved into copper to create the alloys. (Oleson, 2008) p. 110-112. At the year of 1025 AD, Avicenna mentioned, in his famous “The Canon of Medicine”, ZnO as a good ointment for skin treatments, and also as a treatment for skin cancer. ZnO is no longer used to cure skin cancer, but are still widely used in skin treatments, such as baby powders, shampoos, sunscreen etc. During the 14th century, zinc became recognized as a metal from zinc mines in India. ZnO were also synthesized as a byproduct of the smelting process. The Europeans imported zinc from China in the 17th century and zinc was introduced in the periodic table of elements by Antoine Lavoisier in 1789. (Partington, 1989). Before the technological applications arrived in the 1900’s, ZnO was used in the 1800’s as a pigment for painting due to its clear, white color. ZnO is today also a big part of the rubber industry, i.e. as an additive in the fabrication of car tires, as it together with stearic acid take part in the vulcanization of rubber. (Klingshirn, 2007) Figure 7 gives an overview of today’s worldwide applications of ZnO where one sees that the rubber industry is the most important application. In addition, ZnO is used in numerous fields today; as a source to make ceramics, as a compound in animal feed to control the zinc intake and in pharmaceutical applications, to name a few (Zinc.org, 2011).

Figure 6: ZnO as a white powder. Figure adapted from ref. (CCE-Group).

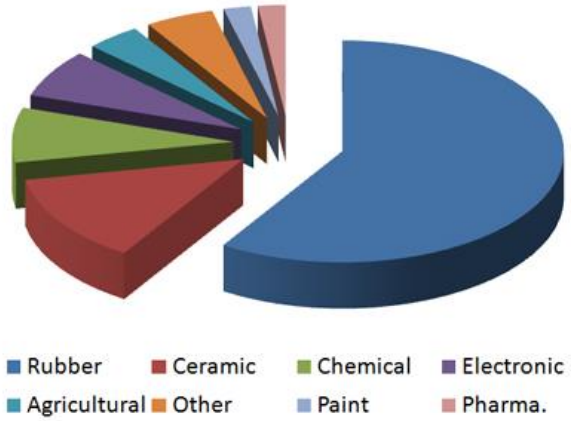

Figure 7: An overview of the worldwide ZnO applications today. As one can see, ZnO is primarily used in the rubber industry. Figure adapted from ref. (Zinc.org, 2011).

Zinc occurs natural as zinc ores available from zinc mines, and zinc ores are often used as a source in the production of ZnO. Bulk ZnO is mainly manufactured from three different processes. The most popular production process is the indirect process, or French process. In the indirect process, metallic zinc are melted and vaporized before it reacts with oxygen in air to make ZnO by the reaction

$$2\,\mathrm{Zn} + \mathrm{O_2} \rightarrow 2\,\mathrm{ZnO} \quad .$$

ZnO particles made from the indirect process are in the micrometer range. Another common process is the direct process, or American process. In the direct process, ZnO is processed from oxidic zinc ores. The oxidic zinc ores are reduced by CO to zinc vapor, and then reoxidized with oxygen in air as in the indirect process. Due to the low purity of the raw material in this process, the direct process produces lower quality ZnO compared to the indirect process. The last of the common ZnO processes is the chemical process. In this process, zinc hydrosulfite is synthesized in an aqueous solution. After that, the salt is neutralized by sodium hydroxide leaving ZnO as a co-product. ZnO can also be manufactured by many other techniques for small-scale productions, like hydrothermal growth of single-crystals, chemical vapor deposition (CVD) and atomic layer deposition (ALD) for thin films of ZnO, and nanostructured from the vapor-liquid-solid (VLS) method, which is a common technique to produce one-dimensional nanostructures (Klingshirn, 2007, Zinc.org, 2011).

Table 1: Basic material constants of ZnO.

Material constant	Value
Molar mass	81.4 g/mol
Density	5.6 g/cm^3
Melting point	1975°C (decomposes)
Dielectric constant	8.8 (at RT)
Standard enthalpy of formation	-321 kJ/mol (at RT)
Heat capacity	41.1 J/(Kmol) (at RT)
Refractive index	2.0
Bulk modulus	142 GPa
Bandgap	3.4 eV (direct bandgap)
Crystal structure	Wurtzite (at RT)

Refs. (Kobiakov, 1980, Morkoç and Özgür, 2009, Pearton et al., 2005).

2.6 ZnO at atomic scale

The unit cell of ZnO can be described in several ways and also exhibit different crystal structures. ZnO crystallizes into two main forms, face-centered cubic (fcc), so-called zincblende, and hexagonal structure dependent of temperature, pressure etc. The hexagonal crystal structure (hcp), so-called wurtzite, is the most common at room temperature, and was the one considered in this thesis. Zincblende consists of three different planes with an ABCABC arrangement which means that every third layer is directly above each other. Wurtzite consists of an alternating ABAB arrangement which means that all the C-sites are empty and available for interstitial impurities.

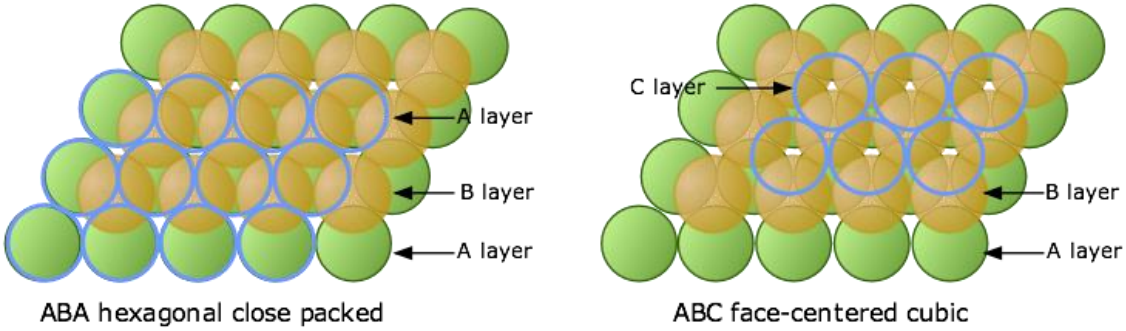

Figure 8: Schematic presentation of the two different stacking possibilities related to close-packed structures. Left: The hexagonal close-packed (hcp) structure consists of a third layer directly above the first A layer leading to an ABAB arrangement. In the hcp structure all the C positions (octahedral positions) are empty. Right: In the cubic close-packed (ccp/fcc) structure the third layer is placed in the C-positions leading to an ABCABC arrangement. The figure is adapted from ref. (Davis).

As said, there are different ways of describing the hexagonal structure, but in this thesis the unit cell was defined with the lattice vectors *a* and *b* standing 120° apart. Vector *c* was perpendicular to the plane spanned by the vectors *a* and *b*. The unit vectors were defined as $a = (a, 0,0)$, $b = (-\frac{a}{2}, \frac{\sqrt{3}}{2}a, 0)$ and $c = (0,0,c)$. *c* is therefore parallel to the z-axis. The lattice constants of ZnO have been studied comprehensively. As initial values for this thesis the following experimental lattice constants were used; $a = 3.25$ Å and $c = 5.21$ Å (Morkoç and Özgür, 2009) p. 12. This gave a unit cell axial ratio of $\frac{c}{a} = 1.60$, somewhat lower than the ideal hcp axial ratio of $\frac{c}{a} = 1.63$. The Zn-atoms were placed in the positions $\left(\frac{1}{3}, \frac{2}{3}, 0\right)$ and $\left(\frac{2}{3}, \frac{1}{3}, \frac{1}{2}\right)$, while the O-atoms were positioned in $\left(\frac{1}{3}, \frac{2}{3}, u\right)$ and $\left(\frac{2}{3}, \frac{1}{3}, \frac{1}{2} + u\right)$ relative to the lattice vectors, where $u \approx \frac{3}{8}$ (Kisi and Elcombe, 1989). The initial value of u was based on the ideal würtzite structure and geometrical considerations, so that the oxygen atoms had approximately the same distance to the surrounding zinc atoms. The exact values of the lattice constants and atomic coordinates were found during ionic relaxation in section 5.1. The unit cell used in this thesis is shown in figure 9, seen along *a*, and in 3D, respectively. The point group for ZnO is 6 mm and the space group is $P6_3mc$ (Klingshirn, 2007).

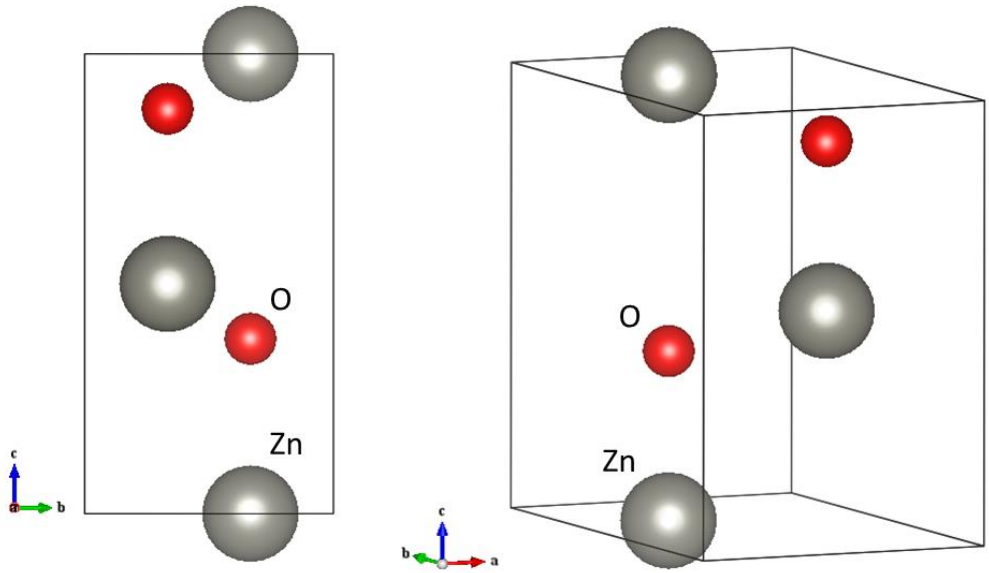

Figure 9: Left: Unit cell for ZnO seen along *a*-direction. The large (grey) balls are the Zn atoms, and the smaller (red) are the O atoms. Right: Unit cell for ZnO seen in 3D.

2.7 Technological applications

ZnO has been used widely in many areas throughout the years. Its applications fields range from electronic devices to cosmetic products. Some of the applications have been presented in an earlier sections, but for technological applications, ZnO has received much attention for its promising role in modern technology as a cheap and conducting oxide.

The introduction of nanotechnology and development of the growth mechanisms of ZnO nanostructures have introduced new and interesting applications by the use of single crystals, epitaxial layers and nanoparticles of ZnO. One of its most promising applications is in the light industry as a material for laser diodes and light emitting diodes (LEDs). ZnO has a large exciton (electron-hole pair) binding energy, above 60 meV (Klingshirn, 2010) p. 1, and emits blue and ultraviolet radiation at room temperature. Therefore, it is therefore an attractive material for blue and UV operating optical devices. Together with GaN, which holds similar properties, ZnO is suggested as the next main light source (Jagadish and Pearton, 2006) ch. 1. One-dimensional nanostructures of ZnO have proven interesting properties for use in field emitters and thin-film transistors (Li et al., 2004, Heo et al., 2004). ZnO is non-centrosymmetric, thus lack of inversion symmetry, and shows a spontaneous polarization. Hereby, ZnO possess piezoelectric and pyroelectric properties (Klingshirn, 2007). It has one of the highest piezoelectric tensors among the tetrahedrally bonded semiconductors, providing ZnO as an preferable mechanical material for use in sensors, actuators and power sources where mechanical stress generates an electrical charge (Dalcorso et al., 1994). Aluminum-doped ZnO is used in the fabrication of transparent conducting oxides (TCO), as a cheaper and less poisonous alternative to indium tin oxide (ITO). Conduction oxides show a wide range of commercially applications like displays, sensors and photovoltaics, i.e. solar cells (Klingshirn, 2007). A suggested technological area for ZnO is in spintronics applications, like magnetic data storage, as ZnO doped with magnetic transition metals have shown ferromagnetism at room

temperature (Dietl et al., 2000). Also, as ZnO is highly resistant against high-energy radiation it is a candidate for space applications (Jagadish and Pearton, 2006) ch. 1.

Although, the development of nanotechnology has shown new possibilities for ZnO at the edge of today's technology, it has also improved the existing applications of ZnO through greater efficiency and smarter designs, like sunscreens and coatings. Currently, the main problem in ZnO is to achieve a reliable and reproducible *p*-type doping, and this problem limits the extensive use of ZnO for technological devices. (Klingshirn, 2007) Nevertheless, not all technological applications are in need for a stable *p*-type semiconductor, so one expects an increased use of ZnO independent of the outcome of the *p*-type problem.

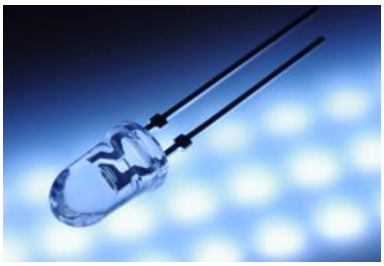

Figure 10: ZnO is a promising material for use in light emitting diodes (LEDs) Figure adapted from ref. (Zinc.org, 2011).

3. First-principles methods

In this chapter, the theory behind the computational software of VASP which is used for the calculations in this thesis will be presented. VASP is based upon calculations from "First-principles methods", namely the density functional theory (DFT). First-principles methods, or ab-initio methods, are terms to describe methods that do not involve any experimental or empirical parameters. In our approach it means that these methods are based upon the quantum theory from the Schrödinger equation with some approximations, but without any fitting parameters to the experimental data.

This chapter will present the steps from the many-body problem to the density functional theory and the Kohn-Sham equation. The many-body problem is an extremely complex problem, but through the Born-Oppenheimer approximation and the Hartree-Fock approximations, Hohenberg and Kohn managed to present two famous theorems (Hohenberg and Kohn, 1964) which reduced the great complexity of the many-body problem into a three-dimensional problem. DFT is an exact theory if the exchange-correlation energy is known. However, the exact description of this energy is, to this date, not known and the last section will therefore present some of the approximations of the exchange-correlation energy for practical approaches.

3.1 The many-body problem

To describe the universe, and the behavior of all the atoms within it, is a fascinating thought. Unfortunately there is a huge numerical problem to be solved before that is a reality. From Avogadro's number, one can approximate the number of atoms in one cubic centimeter to be in order of ~10^{23}. As all these atoms interact with each other, there is a huge numerical problem to overcome, and this is still without any external potentials.

To describe a material one needs to solve the many-body Schrödinger equation. The many-body Schrödinger equation can be seen as an extension of the two-particle Schrödinger equation, but now summing up all the particles. The eigenvalue equation of the many-body time-independent Schrödinger equation can be written as

$$H^{en}\,\Psi_{\kappa}^{en}(\boldsymbol{r},\boldsymbol{R}) = E_{\kappa}\,\Psi_{\kappa}^{en}(\boldsymbol{r},\boldsymbol{R}) \quad .$$

Here, $\Psi_{\kappa}^{en}(\boldsymbol{r},\boldsymbol{R}) = \Psi_{\kappa}^{en}(\boldsymbol{r}_1,\boldsymbol{r}_2,\dots,\boldsymbol{r}_{N_e},\boldsymbol{R}_1,\boldsymbol{R}_2,\dots,\boldsymbol{R}_{N_n})$ is the many-body wavefunction assuming that there are N_e electrons, and N_n nuclei. E_{κ} is the total energy of the system in the given κ-th eigenstate (Persson, 2012). κ is only used a reference index. The superscript *en* is to denote that it considers both electronic and nuclei part, respectively.

If relativistic effects and magnetic interactions are not taken into account, the full Hamiltonian is given as (Thijssen, 2007) p. 44, (Persson, 2012), (Zherebetskyy, 2010) p. 18

$$H^{en} = -\sum_{j=1}^{N_e}\frac{\hbar^2\nabla_j^2}{2m_e} - \sum_{\alpha=1}^{N_n}\frac{\hbar^2\nabla_\alpha^2}{2m_\alpha} + \sum_{j<j'}^{N_e}\frac{e^2}{|\boldsymbol{r}_j-\boldsymbol{r}_{j'}|} + \sum_{\alpha<\alpha'}^{N_n}\frac{e^2Z_\alpha Z_{\alpha'}}{|\boldsymbol{R}_\alpha-\boldsymbol{R}_{\alpha'}|} - \sum_{j=1}^{N_e}\sum_{\alpha=1}^{N_n}\frac{e^2Z_\alpha}{|\boldsymbol{r}_j-\boldsymbol{R}_\alpha|} \quad ,$$

where j runs over all the electrons, and α runs over all the nuclei. Z_α and m_α are the atomic number and the atomic mass of the α-th nuclei, respectively. In SI-units, the above equation should include

$\frac{1}{4\pi\varepsilon_0}$, but this is set to unity for simplicity. As m_j is the mass of electron j, one set $m_j = m_e = 9.11 \times 10^{-31}$ kg, as one assumes that all the electrons have the same mass.

The five terms in the full Hamiltonian are (Persson, 2012),

$T_j = -\sum_{j=1}^{N_e} \frac{\hbar^2 \nabla_j^2}{2m_e}$: Kinetic operator for all the electrons

$T_\alpha = -\sum_{\alpha=1}^{N_n} \frac{\hbar^2 \nabla_\alpha^2}{2m_\alpha}$: Kinetic operator for all the nuclei

$U_{jj} = +\sum_{j<j'}^{N_e} \frac{e^2}{|\boldsymbol{r}_j - \boldsymbol{r}_{j'}|}$: Repulsive electron-electron interactions (because of equal charges)

$U_{\alpha\alpha} = +\sum_{\alpha<\alpha'}^{N_n} \frac{e^2 Z_\alpha Z_{\alpha'}}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_{\alpha'}|}$: Repulsive nuclei-nuclei interactions (because of equal charges)

$U_{j\alpha} = -\sum_{j=1}^{N_e}\sum_{\alpha=1}^{N_n} \frac{e^2 Z_\alpha}{|\boldsymbol{r}_j - \boldsymbol{R}_\alpha|}$: Attractive electron-nuclei interactions (beacuse of opposite charges)

3.2 Born-Oppenheimer approximation

In order to approach the full Schrödinger equation, one searches for approximations that solve the huge numerical problem from the electron-electron interaction in the full Hamiltonian. The form of $\Psi_\kappa^{en}(\boldsymbol{r}, \boldsymbol{R})$ is highly complex, since it includes all the positions of all the electrons and the nuclei. In 1927 (Born and Oppenheimer, 1927), Born and Oppenheimer published a paper where they described the separation of electronic motion, nuclear vibrations and molecular rotation. However, the separation was not done clearly, and it was not until eight years later that the separation was done properly and clear. Eckart (Eckart, 1934) separated internal coordinates (vibration) from external coordinates (rotation and translation), at least minimized the coupling between the two types of coordinates. By this separation of coordinates, the famous Born-Oppenheimer approximation (BOA) occurred as a way to separate the full wavefunction into two parts, one electronic wavefunction, describing all the electrons, and one nuclear wavefunction, describing all the nuclei. The total wavefunction in the simplest BOA can therefore be written as (Zherebetskyy, 2010) p. 18

$$\Psi_\kappa^{en}(\boldsymbol{r}, \boldsymbol{R}) = \psi_\kappa^e(\boldsymbol{r}, \boldsymbol{R})\, \phi_\kappa^n(\boldsymbol{R}) \quad ,$$

with $\psi_\kappa^e(\boldsymbol{r}, \boldsymbol{R}) = \psi_\kappa^e(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{N_e}, \boldsymbol{R}_1, \boldsymbol{R}_2, \dots, \boldsymbol{R}_{N_n})$, as the electronic wavefunction, and $\phi_\kappa^n(\boldsymbol{R}) = \phi_\kappa^n(\boldsymbol{R}_1, \boldsymbol{R}_2, \dots, \boldsymbol{R}_{N_n})$, as the nuclear wavefunction. The success of the BOA is mainly because the nuclei are treated as point charges. The nuclei mass is at least three orders of magnitude higher than the electron mass and it is reasonable to predict that $m_\alpha \gg m_e$. Thereby one can treat the nuclei as point charges with fixed positions since the electrons will be affected instantaneously when the nuclei coordinates are changed. Because of that, the nuclear wavefunction only depends on the coordinates of the nuclei. The electrons feel the potential from the nuclei and are therefore dependent of where the nuclei are situated. However, the nuclei are fixed and therefore the electrons are only indirectly independent of nuclei coordinates (Thijssen, 2007) p. 45. A little note about the fixed nuclear position is that it leads to a contradiction with Heisenberg's uncertainty principle. Therefore, to assume the nuclei to be fixed in the Born-Oppenheimer approximation is not

a pure quantum mechanical approximation, but rather a classical approximation (Zherebetskyy, 2010) p. 19.

Inserted in the eigenvalue many-body equation one obtains

$$H^{en}\psi_\kappa^e(\boldsymbol{r},\boldsymbol{R})\phi_\kappa^n(\boldsymbol{R}) = E_\kappa\psi_\kappa^e(\boldsymbol{r},\boldsymbol{R})\phi_\kappa^n(\boldsymbol{R}) \quad .$$

The full Schrödinger equation can then be separated into an electronic and a nuclear Schrödinger equation, with the total electronic energy; ε_κ^e as the coupling term between the equations (Persson, 2012), (Zherebetskyy, 2010) p. 18

Electronic Schrödinger equation:

$$\left(T_j + U_{jj} + U_{j\alpha}\right)\psi_\kappa^e(\boldsymbol{r},\boldsymbol{R}) = \varepsilon_\kappa^e\,\psi_\kappa^e(\boldsymbol{r},\boldsymbol{R}) \quad .$$

Nuclear Schrödinger equation:

$$(T_\alpha + U_{\alpha\alpha} + \varepsilon_\kappa^e)\phi_k^n(\boldsymbol{R}) = E_\kappa\,\phi_\kappa^n(\boldsymbol{R}) \quad ,$$

where E_κ is the total energy of the system.

However, the nuclear-nuclear interaction is often put into the electronic Schrödinger equation. This is because the nuclei are fixed, and therefore the kinetic term T_α is of small interest. By including the nuclear-nuclear interaction term into the electronic Schrödinger equation, only the electronic equation is to be solved in order to get the total energy of the system, ε_κ^e (Persson, 2012). The reason for neglecting the nuclear equation is based upon two factors. First, the thermal wavelength for a nucleus of mass m_α is $\lambda_T = \frac{\hbar}{\sqrt{m_\alpha k_B T}}$, so if the distance between the nuclei is exceeding the thermal wavelength with a significant value there will be no quantum phase coherence. For the lightest atom, hydrogen (which will have the highest thermal wavelength), at room temperature, the thermal wavelength is $\lambda_T \approx 0.4$ Å. The normal spacing between atoms are ≈ 1 Å, so for most practical purposes, the quantum phase coherence can be neglected for the nuclei. In addition, the potential energy for the nuclei, including the ε_κ^e, are often rigid and the nuclei can be treated as localized particles as described above. At the same time, one cannot neglect the quantum nuclear effects in all systems, and there are several examples where one needs to take the nuclear Schrödinger equation into account (Wilson, 2003) Vol. 2. p. 532-568.

As seen from the separated Schrödinger equation, the two equations are coupled together by the ε_κ^e term. This term describes the total energy of the electronic system; however, the nuclei will depend on the electronic configuration and therefore ε_κ^e is a potential term in the nuclear Schrödinger equation. The two terms $U_{\alpha\alpha}$ and ε_κ^e in the nuclear equation are often merged into one term, called the potential energy surface, $\varepsilon_p = U_{\alpha\alpha} + \varepsilon_\kappa^e$ (Persson, 2012), (Zherebetskyy, 2010) p. 18. This potential energy surface is of interest when analyzing crystal structures and atomic configuration of materials. Usually one solves the electronic Schrödinger equation for a fixed nuclear configuration. Then by solving for another configuration one can see which has the lowest energy, and hereby identify the most stable configuration. By doing this for a large number of atomic configurations, one can get the energy as a function of the nuclear positions, which is called the potential energy surface, often denoted as $\varepsilon_p(\boldsymbol{R})$.

From the potential energy surface, $\varepsilon_p(\boldsymbol{R})$, one can directly by Newton's classical models, derive the forces acting on each atom in the system. Let us consider a nucleus in the system with mass, m_α. By Newton's second law the force acting on that nucleus is given by $\boldsymbol{F}_\alpha = m_\alpha \boldsymbol{a}_\alpha = m_\alpha \frac{\partial^2 \boldsymbol{R}_\alpha}{\partial t^2}$. The force on the nucleus can also be derived from the classical formula for work, $W = -\Delta V = \int \boldsymbol{F}_\alpha \, d\boldsymbol{R}_\alpha$. In the nuclear Schrödinger equation, the change in the potential energy is given by the change in the potential energy surface, $\varepsilon_p(\boldsymbol{R})$, as defined above. Therefore one can determine the force acting on the nucleus by (Persson, 2012)

$$\boldsymbol{F}_\alpha = -\nabla_\alpha \varepsilon_p(\boldsymbol{R}_\alpha) \quad .$$

An analogous way of achieving this formula for the force from a quantum mechanical approach is by Ehrenfest's theorem (Griffiths, 2005) p. 18, which couples the quantum mechanics with the formulas from the classical mechanics. From Ehrenfest's theorem, the quantum mechanical version of Newton's second law is expressed as (Wilson, 2003) Vol 2. p. 534

$$m_\alpha \frac{\partial^2 \langle \boldsymbol{R}_\alpha \rangle}{\partial t^2} = -\langle \nabla_{\boldsymbol{R}_\alpha} \varepsilon_p(\boldsymbol{R}) \rangle \quad .$$

If one assumes that the nuclear wave packet is sufficiently localized as discussed above, one can represent the nuclear wavefunctions as a product of Dirac's delta functions. Furthermore, one obtains the expectation value of the nucleus to be the classical position; $\langle \boldsymbol{R}_\alpha \rangle = \boldsymbol{R}_\alpha^{classic}$. This assumption is called the classical nuclei approximation. The expectation value of the potential energy surface can then be calculated as $\langle \nabla_{\boldsymbol{R}_\alpha} \varepsilon_p(\boldsymbol{R}) \rangle = \nabla \varepsilon_p(\boldsymbol{R}^{classic})$. The expression is therefore only valid when the nuclear positions are represented by delta functions.

3.3 Hartree and Hartree-Fock approximation

So far, one has only talked about the complexity of $\Psi_\kappa^{en}(\boldsymbol{r}, \boldsymbol{R})$, but nothing about its actual form. The main problem related to the many-body equation is not necessarily the complexity of the full wavefunction, but to understand how the full wavefunction are described in terms of the single-electron wavefunctions, and how the single-electron Hamiltonians work on the full wavefunction. However, this has been a subject for discussion throughout the years, since the development of the Schrödinger equation in 1926. One of the first physicists to dwell with this problem was Douglas Hartree. Only the year after Schrödinger's equation, Hartree formed the self-consistent field (SCF) theory based upon how the electrons are moving in a central field, also called the Hartree method. (Froese Fischer, 2003) p. 41.

Consider the electronic Schrödinger equation, and assume a system of N_e electrons. Hartree assumed, as an initial guess, that the electrons had no effect on each other. This is of course a crude model, since it would only be true if the electrons lived in separate universes. If that is true one can write the full Hamiltonian as a sum of N_e single-electron Hamiltonians, $H = \sum_{i=0}^{N_e} H_i$ (Sholl and Steckel, 2009) p. 19. Thereby one achieves a simple eigenvalue-equation for the one-electrons, $H_i \psi_i = E_i \psi_i$ with the eigenfunctions ψ_i called spin orbitals. When the total Hamiltonian is summed over all single-electron operators H_i one can write the full wavefunction as a product of all one-electron states:

$$\psi_\kappa^e(\boldsymbol{r}) = \psi_1^e(\boldsymbol{r}_1)\psi_2^e(\boldsymbol{r}_2) \dots \dots \dots \psi_{N_e}^e(\boldsymbol{r}_N)$$

The energy of this wavefunction can be generated from the single-electron states through the summation of the single-electron energies. (Sholl and Steckel, 2009) p.20. To achieve the exact total Hartree energy, an electronic interaction term also needs to be taken into account. One major disadvantage with the simple Hartree method is that it does not obey the Pauli Exclusion Principle. Since electrons are fermions the total wavefunction must change sign when two electrons exchange, otherwise the wavefunction will cancel. This is due to the fact that two fermions may not occupy the same quantum state simultaneously, and thereby has to have an anti-symmetric wavefunction.

From the Hartree product; substituting two electrons will not change any signs in the full wavefunction. In 1930 (Fock, 1930), Slater and Fock independently pointed out that the Hartree method did not satisfy the anti-symmetry principle. However, by making a linear combination of the one-electron orbitals in a two-particle Hartree product with a minus sign between the factors when exchanging the two electrons, one sees that the anti-symmetry principle is satisfied. It was then suggested that a Slater determinant could be used as a trial function for the wavefunction. A Slater determinant still obtains the single-electron wavefunctions used in the Hartree method, but express the wavefunction as the determinant of a matrix built by these one-electron wavefunctions. For a two-electron system, the full wavefunction is expressed by (Thijssen, 2007) p. 53

$$\psi_\kappa^e(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1^e(\boldsymbol{r}_1) & \psi_2^e(\boldsymbol{r}_1) \\ \psi_1^e(\boldsymbol{r}_2) & \psi_2^e(\boldsymbol{r}_2) \end{vmatrix} = \frac{1}{\sqrt{2}} \left[\psi_1^e(\boldsymbol{r}_1)\psi_2^e(\boldsymbol{r}_2) - \psi_1^e(\boldsymbol{r}_2)\psi_2^e(\boldsymbol{r}_1)\right] \ ,$$

and for a N-electron system as

$$\psi_\kappa^e(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_N) = \frac{1}{\sqrt{N}} \, det \begin{bmatrix} \psi_1^e(\boldsymbol{r}_1) & \cdots & \psi_N^e(\boldsymbol{r}_1) \\ \vdots & \ddots & \vdots \\ \psi_1^e(\boldsymbol{r}_N) & \cdots & \psi_N^e(\boldsymbol{r}_N) \end{bmatrix} \ ,$$

with $\psi_\kappa^e(\boldsymbol{r}_n)$ as the one-electron wavefunction with the electron positioned in position n. The factor in front is a normalization factor to make the total wavefunction normalized. The introduction of a Slater determinant improved the accuracy, but largely increased the computational time. As seen from the two-particle system, when the two electrons are interchanged, the sign of the wavefunction will change. In addition, the Slater determinant has other advantages; when two electrons are in the same positions or if two electrons obtain the same quantum state, the total wavefunction will be zero, as required from Pauli Exclusion Principle (Sholl and Steckel, 2009) p. 20. Hartree-Fock is a more accurate and hence a more computational demanding method than the earlier Hartree method. The electron-electron interaction is taken into account in the Hartree-Fock approach through the Hartree potential defined as

$$V_H(\boldsymbol{r}) = e^2 \int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' - e^2 \int \frac{n_j(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' \ ,$$

where the last term is the self-interacting term which need to be subtracted from the potential. The self-interaction term is often neglected in calculations as the Hartree potential will then be equal for all electrons leading to more efficient calculations (Persson, 2012). In addition, the factor e^2 will be set to unity throughout this thesis for simplification.

Within the Hartree-Fock method, the Schrödinger equation for a single electron is written as (Sholl and Steckel, 2009) p. 21

$$\left[-\frac{\hbar^2}{2m}\nabla_i^2 + V_{en}(\boldsymbol{r}) + V_H(\boldsymbol{r})\right]\psi_i^e(\boldsymbol{r}) = \varepsilon_i\psi_i^e(\boldsymbol{r}) \quad .$$

Here, $V_H(\boldsymbol{r})$ is the Hartree potential defined above and $V_{en}(\boldsymbol{r})$ is the external potential, often referred to the nuclear-electron interaction as defined in the full Schrödinger equation.

The Hartree potential describe the repulsive force between electrons as an average potential integrated over all the electrons. The description of the wavefunction as a single Slater determinant is only an approximation and will not generate the exact energy as it for instance does not include Coulomb correlation between the electrons (Chakravorty and Clementi, 1989). A difference between the exact energy and the Hartree-Fock energy will therefore occur, referred to as the *correlation energy* (Handy and Cohen, 2001). The correlation energy is one of the main reasons why DFT is not an accurate method, to this date, and the description of the correlation energy is thoroughly studied.

3.4 Density functional theory

Thomas and Fermi worked out a theory for calculating the electronic structure of many-body systems, often known as the Thomas-Fermi (TF) model. In 1927 (Thomas, 1927, Fermi, 1928), they independently presented many-body models based upon the electronic density, apart from earlier wavefunction theories. In this respect, the TF model introduced the concepts of the density functional theory.

DFT is based upon two theorems, called the Hohenberg-Kohn theorems. Kohn restates the first theorem in his Nobel lecture from 1999;

Theorem: "The ground state density $n(\boldsymbol{r})$ of a bound system of interacting electrons in some external potential $v(\boldsymbol{r})$ determines this potential uniquely" (Kohn, 1999)

The external potential, $v(\boldsymbol{r})$, defined by Hohenberg-Kohn is the same as the external potential $V_{en}(\boldsymbol{r})$ introduced in the previous section. This means that if one knows the electronic ground state density $n(\boldsymbol{r})$, one can determine the external potential uniquely (except an additive constant), and from that one can determine all the ground state properties of the system. From the fact that if the external potential $V_{en}(\boldsymbol{r})$ is known, the full Hamiltonian is known, and by that one can solve the Schrödinger equation to get the eigenfunctions $\Psi_\kappa^e(\boldsymbol{r}, \boldsymbol{R})$. This is an extremely powerful theorem; Instead of working with 3N spatial coordinates, one has reduced the problem into 3 spatial coordinates with the fact that all ground state properties can be found from the electronic ground state density $n(\boldsymbol{r})$. The proof is rather simple, and can be seen in the appendix. The proof is based upon a non-degenerate ground state, i.e. a unique ground state density. For the degenerate case, one can lift the degeneracy by perturbation theory, shown by M. Levy, which proves that the theorem also holds for the degenerate ground state (Levy, 1979). The ground state density $n(\boldsymbol{r})$, then refers to *any* ground state density $n(\boldsymbol{r})$ (Kohn, 1999). The first theorem states that there exists a functional of the electron density that can be used to solve the Schrödinger equation. However, the first theorem does not say anything about the form of this functional. A mathematical description of a functional is that a functional takes a function as the argument and returns a scalar. As a general

function inserts a variable as the argument, a functional inserts a function $f(x)$ into the functional and returns a scalar.

The second theorem gives an important property of the functional. Hohenberg and Kohn (Hohenberg and Kohn, 1964) states that since every ground state property is a functional of the ground state density, there must exist a universal functional, $F[n(\boldsymbol{r})] = \langle\Psi|T + U_{ee}|\Psi\rangle$, where T and U_{ee} is the system's kinetic and potential energy, respectively, for all number of particles and any external potential. By knowing the universal functional, one can obtain the ground state energy, by applying the universal functional, $F[n(\boldsymbol{r})]$, to the density $n(\boldsymbol{r})$. As Hohenberg and Kohn writes in their original paper "Inhomogeneous electron gas" from 1964: "If $F[n(\boldsymbol{r})]$ were a known and sufficiently simple functional of $n(\boldsymbol{r})$, the problem of determining the ground-state energy and density in a given external potential would be rather easy since it requires merely the minimization of a functional of the three-dimensional density function." (Hohenberg and Kohn, 1964). Again; this theorem is also extremely powerful, since for a given external potential, $V_{en}(\boldsymbol{r})$, the exact ground state energy and the ground state density, can in principle be found easily by minimizing this functional within the variational principle.

The Hohenberg-Kohn functional, $F[n(\boldsymbol{r})]$, is independent of external potentials and should therefore, according to the theorem, have the same form in all electronic systems. The major problem of these theorems is that, up to this date, the exact form of the universal functional $F[n(\boldsymbol{r})]$ is not known. This is due to the huge complexity of this functional. Since $F[n(\boldsymbol{r})]$ must include all the interactions and complexities from the many-electron problem one understands that to design this universal functional is a difficult task. The proof of the second theorem is also rather simple, and can be seen in the appendix.

Density functional theory, in its original form from Hohenberg and Kohn, is an exact theory for describing the ground state. From theory, the ground state is obtained only at $T = 0\ K$, when all the electrons are in their lowest possible energy state. Therefore, the original DFT cannot describe electronic excitations and systems where $T > 0\ K$. However, in the years after 1964, DFT was extended to include more than only the ground state. The year after Hohenberg and Kohn's famous article, Mermin (Mermin, 1965) developed the DFT for higher temperatures. In the years following, DFT has been extended into relativistic DFT (Rajagopa.Ak and Callaway, 1973), and to describe superconductivity (Oliveira et al., 1988). DFT has also been extended to involve time, the so-called time-dependent density functional theory (TDDFT) based upon the Runge-Gross theorem (Runge and Gross, 1984). TDDFT is analogous to the concepts of original DFT by using the same coupling between the time-dependent wavefunction and the time-dependent electronic density, and from there derive the effective potential. This is done by the use of molecular dynamics and taking the vibrations of the atoms into account. TDDFT calculations are much more complex than for DFT since the time-dependent effective potential is in principle dependent on the form of the electronic density from all previous times (Marques et al., 2011a). Lately, there has been some research of constructing the DFT without using the Hohenberg-Kohn theorems, but instead using Legendre transforms. This is called the "effective action approach", and was introduced first by Fukuda in 1994 (Fukuda et al., 1994).

3.5 Kohn-Sham equation

The idea of the Kohn-Sham equations is to achieve the exact ground state density $n_0(\boldsymbol{r})$ by replacing the real system with another auxiliary system that can be solved more easily. The Kohn-Sham equation is the Schrödinger equation of a many-electron interacting system mapped into a fictious non-interacting system. Instead, all the many-body terms are put into a local effective potential, called the "Kohn-Sham potential", $v_{eff}(\boldsymbol{r})$. The ansatz of the Kohn-Sham method is that the ground state density in a fictious non-interacting system is the same ground state density as would be in the real many-electron interacting system. The Kohn-Sham formulations are therefore an exact method (based on the fact that $E_{xc}[n(\boldsymbol{r})]$ has to be exact), and the method is the most common implementation of the DFT today, (Martin, 2004) p. 135, (Lalena and Cleary, 2010) p. 198.

As mentioned above, the Kohn-Sham idea is based on electrons in a non-interacting system and therefore it is convenient to start with a Hartree-like many-electron wavefunction, i.e.

$$\psi_\kappa^e(\boldsymbol{r}) = \psi_1^e(\boldsymbol{r_1})\psi_2^e(\boldsymbol{r_2}) \dots\dots\dots \psi_N^e(\boldsymbol{r_N})$$

From this one can easily calculate the density as

$$n(\boldsymbol{r}) = \sum_{k=0}^{N} |\psi_\kappa^e(\boldsymbol{r})|^2$$

The Kohn-Sham equations are often represented by the following single-electron eigenvalue equation within the Hartree approximation (Sholl and Steckel, 2009) p. 12

$$\left[-\frac{\hbar^2}{2m}{\nabla_i}^2 + v_{eff}(\boldsymbol{r})\right]\psi_i^e(\boldsymbol{r}) = \varepsilon_i\psi_i^e(\boldsymbol{r}) \quad ,$$

where the effective potential $v_{eff}(\boldsymbol{r})$ consists of three different terms

$$v_{eff}(\boldsymbol{r}) = V_H(\boldsymbol{r}) + V_{xc}(\boldsymbol{r}) + V_{en}(\boldsymbol{r}) \quad .$$

Here, $V_H(\boldsymbol{r})$ is the Hartree-potential defined in section 3.3, $V_{xc}(\boldsymbol{r})$ is the exchange-correlation potential which will be described later and $V_{en}(\boldsymbol{r})$ is a known factor derived from the electron-nucleus interaction, $U_{en}[n(\boldsymbol{r})] = \int V_{en}(\boldsymbol{r})\, n(\boldsymbol{r})d\boldsymbol{r}$, in the many-body Schrödinger equation.

As seen above, the Kohn-Sham equations look similar to the earlier Hartree-Fock equations, but they differ in the underlying theories of how they are achieved. In the HF equations, the variational principle is used to minimize the energy with respect to the single particle wavefunctions to achieve the Slater determinant which in turn describes the full many-body wavefunction. The KS equations have the same idea, but instead of minimizing with respect to single particle wavefunctions, the minimization is done with respect to the density. The KS equations are also known as a "Beyond Hartree-Fock" method since they include correlation effects.

Hohenberg and Kohn (Hohenberg and Kohn, 1964) stated that the ground-state energy of an interacting inhomogeneous electron gas in an external static potential $V_{en}(\boldsymbol{r})$ is

$$E[n(\boldsymbol{r})] = T[n(\boldsymbol{r})] + U_{ee}[n(\boldsymbol{r})] + U_{en}[n(\boldsymbol{r})] = F[n(\boldsymbol{r})] + \int V_{en}(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r}$$

where $n(\boldsymbol{r})$ is the electronic density of the system, and $F[n(\boldsymbol{r})]$ is the universal functional introduced by Hohenberg and Kohn. The universal functional can be divided into two terms;

$$F[n(\boldsymbol{r})] = T[n(\boldsymbol{r})] + U_{ee}[n(\boldsymbol{r})]$$

where $T[n(\boldsymbol{r})]$ is the kinetic energy of a system with non-interacting electrons, $U_{ee}[n(\boldsymbol{r})]$ is the electron-electron interaction term. All terms are functionals of the electronic density to the system, $n(\boldsymbol{r})$.

The Kohn-Sham ansatz is to create another Kohn-Sham system of non-interacting particles to make the calculation easier. Hereby, the expression for the total energy can be rewritten as

$$E[n(\boldsymbol{r})] = T_s[n(\boldsymbol{r})] + U_s[n(\boldsymbol{r})] + U_{en}[n(\boldsymbol{r})] + (T[n(\boldsymbol{r})] - T_s[n(\boldsymbol{r})]) + (U_{ee}[n(\boldsymbol{r})] - U_s[n(\boldsymbol{r})]) \quad ,$$

where $T_s[n(\boldsymbol{r})]$ and $U_s[n(\boldsymbol{r})]$ are the kinetic and the static potential energy of the Kohn-Sham system, respectively. These can easily be calculated within the Hartree approximation, and are defined by

$$T_s[n(\boldsymbol{r})] = \sum_{j=1}^{N_e} \int \psi_i^*(\boldsymbol{r}) \frac{-\hbar^2 \nabla^2}{2m_e} \psi_i(\boldsymbol{r}) d\boldsymbol{r} \quad ,$$

$$U_s[n(\boldsymbol{r})] = \frac{1}{2} \iint \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}' \quad .$$

The external potential, denoted as $V_{en}(\boldsymbol{r})$, which leads to the $U_{en}[n(\boldsymbol{r})]$ term is unaffected by the change into the Kohn-Sham system. The two parentheses in the expression for the total energy, which are defined as the difference in the kinetic and the potential energy between the real interacting system and the Hartree-like system, are collected into one term defined as the exchange-correlation energy,

$$E_{xc}[n(\boldsymbol{r})] = (T[n(\boldsymbol{r})] - T_s[n(\boldsymbol{r})]) + (U_{ee}[n(\boldsymbol{r})] - U_s[n(\boldsymbol{r})]) \quad .$$

After rewriting, the total energy functional can be expressed as

$$E[n(\boldsymbol{r})] = T_s[n(\boldsymbol{r})] + U_s[n(\boldsymbol{r})] + U_{en}[n(\boldsymbol{r})] + E_{xc}[n(\boldsymbol{r})]$$

$$E[n(\boldsymbol{r})] = \int \sum_i \psi_i^*(\mathbf{r}) \left\{ -\frac{\hbar^2}{m_e} \nabla^2 \right\} \psi_i(\mathbf{r}) d\boldsymbol{r} + \frac{1}{2} \iint \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}' + \int V_{en}(\boldsymbol{r}) n(\boldsymbol{r})\, d\boldsymbol{r} + E_{xc}[n(\boldsymbol{r})]$$

As the first three terms can be calculated easily within the Hartree approximation, one obtains an exact solution of the total energy if the exchange-correlation energy is known. If the system is non-interacting, $E_{xc}[n(\boldsymbol{r})] = 0$, and the system will have the Hartree total energy, which is the total energy based on non-interacting particles.

In order to find the ground state one obtains the variational method to search for solutions that fulfill the following relation (Persson, 2012) and

$$\frac{\partial E[n(\boldsymbol{r})]}{\partial \psi_i^*(\boldsymbol{r})} = 0 \quad .$$

The reason to take the functional derivative with respect to the complex conjugate wavefunction is to avoid an equation with the complex conjugate in the end. In addition, the solutions need to fulfill an orthonormality requirement defined by

$$g = \sum_i \lambda_i \left\{ \int \psi_i^*(\boldsymbol{r}) \psi_i(\boldsymbol{r}) d\boldsymbol{r} - 1 \right\} = 0 \ .$$

The two requirements can be summarized in one equation

$$\frac{\partial (E[n(\boldsymbol{r})] - g)}{\partial \psi_i^*(\boldsymbol{r})} = 0 \quad ,$$

$$\frac{\partial E[n(\boldsymbol{r})]}{\partial \psi_i^*(\boldsymbol{r})} = \frac{\partial g}{\partial \psi_i^*(\boldsymbol{r})} \quad .$$

From the equation of the orthonormality requirement one can easily take the functional derivative and obtain an equation for the energy functional.

$$\frac{\partial E[n(\boldsymbol{r})]}{\partial \psi_i^*(\boldsymbol{r})} = \frac{\partial g}{\partial \psi_i^*(\boldsymbol{r})} = \lambda_i \psi_i(\boldsymbol{r})$$

The energy functional can be expressed in terms of a kinetic, an electron-electron interaction, an electron-nuclei, and an exchange-correlation part,

$$\frac{\partial}{\partial \psi_i^*(\boldsymbol{r})} \{T_s[n(\boldsymbol{r})] + U_s[n(\boldsymbol{r})] + U_{en}[n(\boldsymbol{r})] + E_{xc}[n(\boldsymbol{r})]\} = \lambda_i \psi_i(\boldsymbol{r}) \quad .$$

As the terms are functionals of the electronic density one applies the chain rule to take the functional derivative with respect to density for the last three terms.

$$\frac{\partial}{\partial \psi_i^*(\boldsymbol{r})} T_s[n(\boldsymbol{r})] + \frac{\partial}{\partial n(\boldsymbol{r})} \{U_s[n(\boldsymbol{r})] + U_{en}[n(\boldsymbol{r})] + E_{xc}[n(\boldsymbol{r})]\} \psi_i(\boldsymbol{r}) = \lambda_i \psi_i(\boldsymbol{r})$$

The last factor on the left-hand side is from the chain rule and it is now clear why one should take the functional derivative with respect to the complex conjugate wavefunction. Within the Hartree approximation, the kinetic part and the electron-electron interaction can be calculated and give

$$\left\{ \frac{-\hbar^2 \nabla_i^2}{2m_e} + \int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' + V_{en}[n(\boldsymbol{r})] + V_{xc}[n(\boldsymbol{r})] \right\} \psi_i(\boldsymbol{r}) = \varepsilon_i \psi_i(\boldsymbol{r}) \quad ,$$

which can be recognized as the single particle KS equation presented earlier in this section (Persson, 2012). λ_i is set equal to ε_i to clarify it as an energy term. From classical physics, the potential energy is given as the derivative of the total energy. Therefore the exchange-correlation potential energy is found to be the functional derivative of the exchange-correlation energy (Thijssen, 2007) p. 119.

$$V_{xc}[n(\boldsymbol{r})] = \frac{\partial E_{xc}[n(\boldsymbol{r})]}{\partial n(\boldsymbol{r})}$$

From the single electron KS equation one multiply by $\int \sum_i \psi_i^*(\mathbf{r})$ from left to obtain an expression for the single-electron energies.

$$\int \sum_i \psi_i^*(\mathbf{r}) \left\{ -\frac{\hbar^2 \nabla_i^2}{m_e} + V_H(\boldsymbol{r}) + V_{en}(\boldsymbol{r}) + V_{xc}(\boldsymbol{r}) \right\} \psi_i(\mathbf{r}) d\boldsymbol{r} = \int \sum_i \psi_i^*(\mathbf{r})\, \varepsilon_i \psi_i(\mathbf{r}) d\boldsymbol{r}$$

$$\sum_i \varepsilon_i = \int \sum_i \psi_i^*(\mathbf{r}) \left\{ -\frac{\hbar^2 \nabla_i^2}{m_e} \right\} \psi_i(\mathbf{r}) d\boldsymbol{r} + \iint \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}' + \int V_{en}(\boldsymbol{r}) n(\boldsymbol{r})\, d\boldsymbol{r} + \int V_{xc}(\boldsymbol{r}) n(\boldsymbol{r})\, d\boldsymbol{r}$$

If one rewrites the latest equation as an expression for the kinetic part and insert it in the expression for the total energy functional one gets a relation between the energy functional and the single-electron energies from the KS equation (Thijssen, 2007) p. 119.

$$E[n(\boldsymbol{r})] = \sum_i^N \varepsilon_i - \frac{1}{2} \iint \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}' + E_{xc}[n(\boldsymbol{r})] - \int V_{xc}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r}$$

The total energy can be generated from the Hartree-like single-electron energies, a self-interacting term and the exchange-correlation functional. The exact energy is therefore generated if the exact form of the exchange-correlation functional is found.

3.6 Exchange-correlation energy

Exchange interaction is related to the quantum mechanical effect between identical particles in a system in accordance to Pauli Exclusion Principle. Correlation energy is related to the interactions between electrons in a system, hence the influence a given electron exerts on all the other electrons in the system. The exchange interaction are taken into account in the Slater-determinant of the HF method, however the exact correlation energy is not treated exactly and the correlation energy is often defined as the energy difference between the HF energy and the exact non-relativistic energy of a system. The exact form of the exchange correlation energy has been a subject in theoretical material science for decades and, to this date, there exist no exact solution to the $E_{xc}[n(\boldsymbol{r})]$ term. For numerical calculations it has been important to make good approximations for this unknown energy term in order to approach the exact solutions. In the previous section, the ground state energy of an electron gas in a static potential was written as

$$E[n(\boldsymbol{r})] = F[n(\boldsymbol{r})] + \int V_{en}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} \quad .$$

From the previous section, the universal function $F[n(\boldsymbol{r})]$ can be expressed as

$$F[n(\boldsymbol{r})] = T_s[n(\boldsymbol{r})] + U_s[n(\boldsymbol{r})] + E_{xc}[n(\boldsymbol{r})] \quad .$$

One of the great advantages with the Kohn-Sham equation is that an approximation for the exchange-correlation energy can be included directly in the equation without changing any of the theories and assumptions behind the KS equation. However, this has led to a huge amount of different approximation methods with different grades of success (Persson, 2012).

3.6.1 Local density approximation

The first famous approximation, which is still in use today because of its simplicity and success of describing various material properties, is the local density approximation (LDA). The LDA was presented by Kohn and Sham in 1965, in the paper "Self-Consistent equations including exchange and correlation effects" (Kohn and Sham, 1965).

The LDA is based upon an assumption that the exchange-correlation energy of one electron in a many-body system can locally be approximated by the exchange-correlation energy for a free electron gas. Therefore, the exchange-correlation energy can be calculated by summing up the contributions from all the electrons in the system;

$$E_{xc}^{LDA}[n(\boldsymbol{r})] = \int \varepsilon_{xc}^{gas}(\boldsymbol{r})\, n(\boldsymbol{r}) d\boldsymbol{r} \quad .$$

$\varepsilon_{xc}^{gas}(\boldsymbol{r})$ is the exchange-correlation energy of an electron in the free electron gas, which can be computed exactly from Quantum Monte Carlo methods (Becke, 1993b). It is common to separate the exchange energy and the correlation energy;

$$E_{xc}[n(\boldsymbol{r})] = E_x[n(\boldsymbol{r})] + E_c[n(\boldsymbol{r})] \ .$$

The electronic density in a homogenous electron gas is constant, $n(\boldsymbol{r}) = n = \frac{N}{V}$, and the term $\varepsilon_{xc}^{gas}(\boldsymbol{r})$ can be found analytically as (Persson, 2012)

$$\varepsilon_{xc}^{gas}(\boldsymbol{r}) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n^{\frac{1}{3}} + \begin{cases} Alnr_s + B + Cr_s lnr_s + Dr_s & r_s \leq 1 \\ \dfrac{\gamma}{1 + \beta_1\sqrt{r_s} + \beta_2 r_s} & r_s > 1 \end{cases}; \quad \text{where} \quad r_s = \left(\frac{3}{4\pi n}\right)^{\frac{1}{3}} \ .$$

By comparing the above equations, one gets the exchange energy term in the LDA as (Becke, 1988a)

$$E_x^{LDA}[n(\boldsymbol{r})] = \int \frac{-3e^2}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\boldsymbol{r})^{\frac{1}{3}} n(\boldsymbol{r})\, d\boldsymbol{r} = -\int \frac{3e^2}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\boldsymbol{r})^{\frac{4}{3}} d\boldsymbol{r} \quad .$$

Furthermore, the corresponding exchange potential is found by the functional derivative of the exchange energy with respect to the electronic density.

$$V_x^{LDA}[n(\boldsymbol{r})] = \frac{\partial E_x^{LDA}[n(\boldsymbol{r})]}{\partial n(\boldsymbol{r})} = -e^2\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\boldsymbol{r})^{1/3}$$

Likewise, the analytic term $E_c^{LDA}[n(\boldsymbol{r})]$ in the high- and low density limit from the above equation is expressed as

$$E_c^{LDA}[n(\boldsymbol{r})] = \int \begin{cases} (Alnr_s + B + Cr_s lnr_s + Dr_s)\, n(\boldsymbol{r}) d\boldsymbol{r} & r_s \leq 1 \\ \left(\dfrac{\gamma}{1 + \beta_1\sqrt{r_s} + \beta_2 r_s}\right) n(\boldsymbol{r}) d\boldsymbol{r} & r_s > 1 \end{cases} \ .$$

The LDA is only exact if the electronic system is uniform, which for practical means never happen. However, it is found that LDA works appropriate for slowly varying densities where the uniform density approach is more convenient. As mentioned above, the LDA has turned out to work

surprisingly well for many different systems, and to describe many material properties by an impressive accuracy, despite its great simplicity compared to other complex approximations.

However, there are also serious drawbacks with the local density approximation. One of the main problems is that it highly underestimates the bandgap energies in semiconductors. As a rule of thumb, the bandgap is normally underestimated by about 40 % within the LDA (Perdew and Levy, 1983). Considering the exchange-correlation energy, the LDA underestimates the exchange energy with by roughly 10 % (Becke, 1988b), and overestimates the correlation energy (Becke, 1997) leading to a distorted balance between the exchange-correlation effects. Another problem associated with LDA is that lattice constants of solids are underestimated and are worse for group IIA and 3d elements. However, the underestimation is in most cases a few percent. (Haas et al., 2009)

The LDA is based on spin-unpolarized systems, but there is also some extension of the LDA to consider spin-polarized systems. This is called the local-spin-density approximation (LSDA), and was derived by Oliver and Perdew in 1979 (Oliver and Perdew, 1979). They show that the exchange-correlation energy is extended to spin-polarized systems by a similar equation (Becke, 1993b)

$$E_{xc}^{LSDA}[n_{\uparrow}(\boldsymbol{r}), n_{\downarrow}(\boldsymbol{r})] = \int \varepsilon_{xc}^{gas}(n_{\uparrow}(\boldsymbol{r}), n_{\downarrow}(\boldsymbol{r}))n(\boldsymbol{r})d\boldsymbol{r} \quad ,$$

where $\varepsilon_{xc}^{gas}(n_{\uparrow}(\boldsymbol{r}), n_{\downarrow}(\boldsymbol{r}))$ is the exchange-correlation energy for the system with the two spin-densities, up and down. The extension of the exchange energy is a simple task since the spin-scaling is known, however, for the correlation energy the extension is somewhat more difficult.

3.6.2 Generalized gradient approximation

Another approximation to the exchange-correlation energy for solving the Kohn-Sham equation is the generalized gradient approximation (GGA). The GGA is a continuation and improvement from the LDA; as LDA includes the electronic density in its exchange-correlation energy functional, the GGA exchange correlation functional is dependent of both the electronic density and the gradient of the electronic density (Perdew et al., 1996).

$$E_{xc}^{GGA}[n_{\uparrow}(\boldsymbol{r}), n_{\downarrow}(\boldsymbol{r})] = \int f(n_{\uparrow}(\boldsymbol{r}), n_{\downarrow}(\boldsymbol{r}), \boldsymbol{\nabla} n_{\uparrow}(\boldsymbol{r}), \boldsymbol{\nabla} n_{\downarrow}(\boldsymbol{r}))n(\boldsymbol{r})d\boldsymbol{r}$$

The GGA has shown great improvements compared to the LSDA considering the accuracy of total ground-state energies, atomization energies, binding energies, energy barriers, band structures and molecular geometries. Some known GGA functionals are the PW91 (Perdew et al., 1992), the PBE (Perdew et al., 1996), the RPBE (Hammer et al., 1999) and the WC (Wu and Cohen, 2006).

There exist also meta-GGA functionals, which is an extension to the GGA functionals. This is an even more accurate method as it also includes second order derivatives of the electronic density in its functionals through the use of Laplacian operators (Tao et al., 2003). However, the meta-GGA functionals have significant higher computational requirements.

Later there have been introduced semi-empirical exchange-correlation parameters to achieve quantities and properties for the exchange-correlation functionals, such as LDA+U and GGA+U. The effective parameter, U, introduces a corrective functional to calibrate the error in the approximation. From this, the highly underestimated bandgap from LDA can be corrected to its experimental value

within the LDA+U. The introduction of the effective parameter, U, has shown good improvement for electronic and magnetic ground state properties within both the LDA and the GGA (Himmetoglu et al., 2013).

3.6.3 Hybrid density functionals

While the earlier mentioned functionals are based purely on the DFT, the hybrid density functionals (HDF) are based on a mixture between the DFT exchange-correlation functionals and the Hartree-Fock exchange functionals. The idea comes from the fact that for several material properties and quantities the experimental values lie between the calculated DFT value and the calculated HF value. In example, the DFT functionals from the LDA often strongly underestimates the bandgap, while the HF functionals overestimates the bandgap by a factor 2. (Marques et al., 2011b). Therefore, by taking the HF functionals into account and use the right mixing between the two types of functionals one can obtain the experimental bandgap.

The energy functional in the hybrid functionals has the following form (Becke, 1993b)

$$E_{xc}^{hybrid} = aE_{xc}^{HF}[\psi_i] + (1-a)E_{xc}^{DFT}[n(\boldsymbol{r})] \quad ,$$

with a as the *mixing parameter* which describes to what extent the different functionals are weighted. The HF energy functionals are functionals of the wavefunction, while the DFT functionals are functionals of the electronic density. Examples of some hybrid functionals are the B3LYP (Becke, 1993a, Stephens et al., 1994), PBE0 (Adamo and Barone, 1999) and TPSSh (Staroverov et al., 2003).

HDF have become a great success the recent years by an impressive accuracy in a wide range of molecular properties compared to experimental studies. One of the major drawbacks of the hybrid functionals is the demand for computational time. This is due to the long-range order of the Hartree-Fock exchange which causes high computational effort for periodic systems, and especially in periodic systems which use dense Brillouin zone measurements, such as metallic systems. However, some of the new hybrid functionals, such as HSE03 (Heyd et al., 2003) evaluate the Hartree-Fock exchanges in the real space rather than in the reciprocal space to overcome the problem related to the dense Brillouin zone measurements. (Paier et al., 2006).

A notification about the HDF is that they use experimental data to tune the mixing parameter in order to obtain the experimental value. In this respect, HDF is not an *ab-initio* method, but rather a semi-empirical approach. A problem related to the HDF is how the mixing parameter is to be determined. The mixing parameter is in principle only set for a unique system and the mixing parameter is different whether it is tuned against the bandgap or another desired quantity. Hence, the mixing parameter needs to be determined exact for each system and for a specific quantity in order to be reproducible and consistent, in analogue to the U parameter for the previous approximations.

4. Atomistic modeling of solids

In this chapter some of the theory and methods behind atomistic modeling of solids will be presented. There is a long way from the physics of the density functional theory to the practical interpretation in VASP. Even though DFT simplified the many-body problem, the calculations are time-consuming and several approximations must be considered in order to be able to do calculations within reasonable computational time. This chapter presents some of the parameters and simplifications used in numerical calculations for the practical approach of DFT. At first, the Schrödinger equation in reciprocal space is presented. In most DFT calculations the reciprocal space is the reference space for calculations as it simplifies the integrals in the calculations. Next, the accuracy of numerical calculations can to a large extent be decided by convergence parameters and a compromise must be taken between accuracy and computational demand. In addition, the concepts of pseudopotentials and basis sets are presented. By the fact that ion cores often behave as an inert term and to a large extent do not take part in chemical bonding, the core can be approximated by several means in order to reduce computational time. Next, the theory behind vibrational models and transition state theory by the elastic band methods are presented. A large part of this thesis involves calculations of vibrational frequencies and diffusion profiles, and these sections describe the theory and the practical interpretation of the theory. Last, some general theory of defects and impurities are presented, as well as the formation energy and ionization energy related to the introduction of defects.

4.1 Schrödinger equation in reciprocal space

Due to the periodicity related to crystalline materials, numeric calculations are often performed in the reciprocal space. In a DFT calculation, where the Kohn-Sham equation needs to be solved, this is done in the reciprocal space. The reciprocal space is the Fourier transform of the real space, and one can write the Fourier transform of a periodic potential as $V(\boldsymbol{r}) = \sum_{\boldsymbol{G}} V_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}$. This relation has the same periodicity as the crystal $V(\boldsymbol{r}+\boldsymbol{R}) = V(\boldsymbol{r})$ from the relation $e^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1$, because of the orthogonal relationship between the real and the reciprocal lattice unit vectors. The wavefunction can be written in the reciprocal space as a Bloch wave with the Bloch periodic part as a summation over all reciprocal lattice vectors $\psi_{nk}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{\boldsymbol{G}} u_{n\boldsymbol{k}+\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}$. Starting from the Schrödinger equation in the real space, one can easily derive the Schrödinger equation in the reciprocal space. The derivation follows the steps given by ref. (Persson, 2012).

$$\left(-\frac{\hbar^2\nabla^2}{2m_e} + V(\boldsymbol{r})\right)\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = E_{n\boldsymbol{k}}\,\psi_{n\boldsymbol{k}}(\boldsymbol{r})$$

$$\left(-\frac{\hbar^2\nabla^2}{2m_e} + \sum_{\boldsymbol{G}'} V_{\boldsymbol{G}'} e^{i\boldsymbol{G}'\cdot\boldsymbol{r}}\right)\sum_{\boldsymbol{G}} u_{n\boldsymbol{k}+\boldsymbol{G}} e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} = E_{n\boldsymbol{k}} \sum_{\boldsymbol{G}} u_{n\boldsymbol{k}+\boldsymbol{G}} e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}$$

The kinetic energy is calculated by a double derivation, and the potential energy is calculated by a double summation.

$$\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2\sum_{\boldsymbol{G}}u_{n\boldsymbol{k}+\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}+\sum_{\boldsymbol{G}\boldsymbol{G}'}V_{\boldsymbol{G}'}u_{n\boldsymbol{k}+\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G}+\boldsymbol{G}')\cdot\boldsymbol{r}}=E_{n\boldsymbol{k}}\sum_{\boldsymbol{G}}u_{n\boldsymbol{k}+\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}$$

The vector sum $\boldsymbol{G}+\boldsymbol{G}'$ is just another reciprocal lattice vector. Therefore one can rewrite the summation in the potential term as $\boldsymbol{G}+\boldsymbol{G}'=\boldsymbol{G}''\rightarrow\boldsymbol{G}$ to obtain

$$\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2\sum_{\boldsymbol{G}}u_{n\boldsymbol{k}+\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}+\sum_{\boldsymbol{G}\boldsymbol{G}'}V_{\boldsymbol{G}'}u_{n\boldsymbol{k}+\boldsymbol{G}-\boldsymbol{G}'}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}=E_{n\boldsymbol{k}}\sum_{\boldsymbol{G}}u_{n\boldsymbol{k}+\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}\quad.$$

As the summation over all $\boldsymbol{G}$ is a common factor in all terms it is placed in front;

$$\sum_{\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}\left[\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2u_{n\boldsymbol{k}+\boldsymbol{G}}+\sum_{\boldsymbol{G}'}V_{\boldsymbol{G}'}u_{n\boldsymbol{k}+\boldsymbol{G}-\boldsymbol{G}'}\right]=\sum_{\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}\left[E_{n\boldsymbol{k}}u_{n\boldsymbol{k}+\boldsymbol{G}}\right]\quad.$$

When the summation is taken over all the reciprocal lattice vectors, one can look at the terms in the brackets as the coefficients for a given plane wave $e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}$. In order to fulfill the equation the coefficients need to be equal for all $\boldsymbol{G}$. From that requirement, the summation over plane waves can be eliminated from both sides, as the brackets must be equal for every $\boldsymbol{G}$.

$$\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2u_{n\boldsymbol{k}+\boldsymbol{G}}+\sum_{\boldsymbol{G}'}V_{\boldsymbol{G}'}u_{n\boldsymbol{k}+\boldsymbol{G}-\boldsymbol{G}'}=E_{n\boldsymbol{k}}u_{n\boldsymbol{k}+\boldsymbol{G}}$$

Again, the summation $\boldsymbol{G}-\boldsymbol{G}'$ is just another reciprocal lattice vector and can be rewritten as $\boldsymbol{G}-\boldsymbol{G}'=\boldsymbol{G}''\rightarrow\boldsymbol{G}'$ to compress the equation. The index of summation is not needed to change as the summation is taken over all reciprocal vectors and it would only change the direction of summation (Persson, 2012).

$$\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2u_{n\boldsymbol{k}+\boldsymbol{G}}+\sum_{\boldsymbol{G}'}V_{\boldsymbol{G}'}u_{n\boldsymbol{k}+\boldsymbol{G}-\boldsymbol{G}'}=E_{n\boldsymbol{k}}u_{n\boldsymbol{k}+\boldsymbol{G}}$$

$$\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2u_{n\boldsymbol{k}+\boldsymbol{G}}+\sum_{\boldsymbol{G}'}V_{\boldsymbol{G}-\boldsymbol{G}'}u_{n\boldsymbol{k}+\boldsymbol{G}'}=E_{n\boldsymbol{k}}u_{n\boldsymbol{k}+\boldsymbol{G}}$$

$$\sum_{\boldsymbol{G}'}\left(\frac{\hbar^2}{2m_e}(\boldsymbol{k}+\boldsymbol{G})^2\delta_{\boldsymbol{G}\boldsymbol{G}'}+V_{\boldsymbol{G}-\boldsymbol{G}'}\right)u_{n\boldsymbol{k}+\boldsymbol{G}'}=E_{n\boldsymbol{k}}u_{n\boldsymbol{k}+\boldsymbol{G}}$$

The last equation is a way to describe the Schrödinger equation in periodic coordinates, thus the reciprocal space. From this equation one achieves the eigenvalues $E_{n\boldsymbol{k}}$ as the total energy of the system, and the coefficients $u_{n\boldsymbol{k}+\boldsymbol{G}}$ to generate the corresponding eigenfunctions.

4.2 Convergence parameters

When performing numerical calculations on materials, convergence is a keyword. The basic idea in numerical calculations is to make qualified guesses, and then improve and optimize the initial guesses. A calculation is converged when the next optimization step does not change the previous calculation larger than an error bar specified by the convergence parameters. There are several

parameters that are of special importance in numerical calculations and two of them are presented in the following sections, the number of k-points and the energy cutoff.

4.2.1 k-point grid

In a DFT calculation one needs to handle integrals that run over all k-points in the Brillouin zone. To be correct, as the number of k-points is discrete, the calculation is summations over all the discrete k-points, but from interpolation and smearing one expresses the summations as integrals. As the integrals run over all k-points it is reasonable to calculate the integrals in k-space. A k-point grid defines a three-dimensional grid in the reciprocal space in which the integrals will be evaluated. The most common grid in numerical calculations today is the Monkhorst-Pack k-point grid (Monkhorst and Pack, 1976) which is uniformly spaced in the Brillouin zone. The grids are given as $M_1 \times M_2 \times M_3$ to specify the number of k-points in each reciprocal direction, and a larger number of k-points gives more accurate solutions to the integrals. For systems where the real lattice vectors have the same length in all directions, i.e. cubic systems, the length of the reciprocal lattice vectors will also be equal. In this case it is common to use the same number of k-points in each direction (Sholl and Steckel, 2009) p. 55. For non-cubic structures, individual assumptions are made for M_1, M_2, and M_3. The computational cost is in principle linear with respect to the total number of k-points. As a result, doubling the number of k-points in each direction, $2M_1 \times 2M_2 \times 2M_3$, implies an 8 times more expensive calculation than a $M_1 \times M_2 \times M_3$ grid (Kitchin, 2012) p. 92. As the relation between the real space and k-space is reciprocal, a larger supercell indicates that fewer k-points are needed in order to obtain convergence since the volume of the reciprocal supercell will then decrease.

The numerical calculations involve integrals in the reciprocal space. However, for supercells that have a high degree of symmetry, simplification is possible in order to increase the computational efficiency. In such supercells, the integrals can be evaluated in a reduced area of the Brillouin zone and then extended to the entire Brillouin zone by exploiting the symmetry of the cell. The reduced region that holds the least number of k-points needed to evaluate the whole Brillouin zone is called the *irreducible Brillouin zone* (IBZ). Higher degree of symmetry implies a smaller IBZ which shortens the computational time. For a non-symmetric supercell, the number of k-points in IBZ approaches half the number of k-points in the full BZ. The full BZ can always be divided by a factor 2 by the relation between the states at **k** and **–k** (Martin, 2004) p. 92. Therefore, non-symmetric supercells are more difficult to converge and require a larger k-point grid than supercells that show some degree of symmetry. Depending on the symmetry of the cell, for even or odd numbers of k-points in each direction some k-points will lie on the boundary of IBZ. The number of k-points in each direction can then be increased by one without any increased computational time, but with slightly improved convergence as all the k-points will now be inside IBZ. This effect is smaller for larger k-point grids and hence well-converged values (Sholl and Steckel, 2009) p. 57.

A property, e.g. total energy, is said to be converged when an increase in the number of k-points does not change the total energy within a specified value. Therefore, one must compromise between accuracy and computational demand. The same errors and inaccuracies in numerical calculations often occur in similar calculations. Therefore it is important to use the same parameters, e.g. k-point grid, when comparing data. In material science, comparing calculated properties between different structures i.e. total energy, density of states, bandgap etc. are of major interest. One can thereby divide between relative and absolute convergence. Since the same errors are repeated in the same type of calculations, the relative differences in the calculated properties tends to converge faster

than the absolute values. In example; changing between different potentials provides a substantial discrepancy in the free energy. The absolute value of the total free energy calculated by VASP is therefore difficult to predict. However, the relative difference in total energy between structures can predict the most stable configuration, and give important information of the energy differences between related structures. In terms of the discussion above, the relative difference in energy between two structures will therefore converge with a smaller k-point grid than the absolute value. Therefore, one can obtain relevant information about non-symmetric supercells at smaller k-point grids than what discussed above.

In VASP one defines the number of k-points in each direction and whether an original Monkhorst-Pack grid or a Γ–centered Monkhorst Pack grid is used. There is also possible to make a shift in the automatic generated k-mesh and to create the k-point grid manually in terms of Cartesian or reciprocal coordinates. For all calculations in this thesis an automatic k-mesh is generated with a Γ-centered Monkhorst-Pack grid unless otherwise stated. Due to the large supercell used in this thesis most of the calculations are performed with a $2 \times 2 \times 2$ k-point grid as an appropriate grid.

4.2.2 Energy cutoff

As the Bloch periodic part $u_{nk}(\boldsymbol{r})$ from Bloch's theorem possesses the same periodicity as the crystal, the Bloch periodic part can be described as a Fourier series where a discrete set of plane waves are summed over all the reciprocal lattice vectors $u_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}$. From Bloch's theorem one can write

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{\boldsymbol{k}+\boldsymbol{G}} e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} \quad .$$

The kinetic energy of these solutions are taken as the double derivative and are given as

$$E_{kin} = \frac{\hbar^2}{2m} |\boldsymbol{k} + \boldsymbol{G}|^2 \quad .$$

To avoid the summation over all the reciprocal lattice vectors one introduces an energy cutoff corresponding to a reciprocal cutoff vector. The lowest values of $\boldsymbol{G}$ correspond to the first terms in the Fourier series. Furthermore, they correspond to the lowest energies which will be more physically important than the higher terms. Therefore, it is reasonable to stop the calculation after a certain value in order to save computational time without losing any physical properties (Sholl and Steckel, 2009) p. 62. The energy cutoff corresponds to a cutoff reciprocal lattice vector by the expression

$$E_{cut} = \frac{\hbar^2}{2m} G_{cut}^2 \quad .$$

The summation above will only be evaluated up to the given reciprocal vector G_{cut} corresponding to E_{cut}. The energy cutoff defines the size of the basis set, and the total energy is a function of the plane wave energy cutoff. The plane waves in the basis set meet the requirement

$$E_{cut} \geq \frac{\hbar^2}{2m} |\boldsymbol{k} + \boldsymbol{G}|^2 \quad .$$

Pseudopotentials calculated in the literature often introduce adequate values for the energy cutoff given in eV. It is recommended to keep the value fixed for calculations performed on similar systems. As for k-points, when different systems are compared it is reasonable to use the same energy cutoff in order to keep the values as comparable as possible.

The choice of energy cutoff is important in terms of the computational time as high energy cutoff is expensive. A doubling of the energy cutoff may increase the computer time three or four times. For practical issues, only a certain amount of computer time is available for each computational science group and one needs to choose cutoff carefully to achieve applicable accuracy.

4.3 The effective potential

In order to treat the interactions between the nuclei and the electrons in VASP, several methods are possible. Two famous approaches are the use of pseudopotentials and the projector augmented wave (PAW) method. The PAW method is discussed in section 4.4.3.

4.3.1 All-electron potential vs. pseudopotentials

Without further treatment, one can imagine the computational effort of calculating the nuclei-electrons interactions by taking care of the motions of the nuclei and the core electrons in a large supercell. The *all-electron* potential, represented by a strong Coulomb potential, describes the nucleus and the highly localized core-electrons by individual wavefunctions. However, the all-electron potential is extremely time-consuming to treat computationally with a discrete set of plane waves as described by Bloch's theorem. This is due to the rapid oscillations of wavefunctions close to the ion core region. Because of the reciprocal relation between the real space and the k-space, large cutoff energies are needed in order to include the short wave lengths oscillations associated with the localized core electrons (Sholl and Steckel, 2009) p. 63. The rapid oscillations in the core regions appear in order to maintain the orthogonality of the wavefunctions between the localized ion cores and the valence states.

A pseudopotential is an approximation to simplify this problem. The idea to use pseudopotentials as a replacement for the core-electron potentials was first introduced by Hans Hellmann (Hellmann, 1935). In order to maintain the orthogonality between the core electrons and the valence electrons, Hellmann created a pseudopotential which was strongly repulsive to the inner electrons to keep the valence electrons out of the core (Schwerdtfeger, 2011). A pseudopotential is an effective potential which describes the behavior of the nucleus and the core electrons, so that the calculations can be focused on the valence electrons. The ion cores are treated as inert terms, so the method is also called the frozen core approximation. The valence electrons are dependent on the ion cores, however, the core electrons screen the nucleus, so the valence electrons feel an effective potential from the ion core which is smoother than the potential described by the nucleus individually. Pseudopotentials can therefore replace the narrow oscillating all-electron potential by a smooth potential which approximates the strong Coulomb potentials inside the core regions. Therefore, the pseudopotentials are said to be *softer* than the full potential. The pseudopotential is constructed to yield the correct shape of the wavefunction beyond a certain radii, r_c, (Hamann et al., 1979)

$$\psi_i^{PP}(\boldsymbol{r}) = \psi_i^{AE}(\boldsymbol{r}) \quad , \quad r > r_c \quad ,$$

where $\psi_i^{PP}(\boldsymbol{r})$ is the pseudo wavefunction, and $\psi_i^{AE}(\boldsymbol{r})$ describes the all-electron wavefunction. Figure 11 illustrates the concept behind the pseudopotentials.

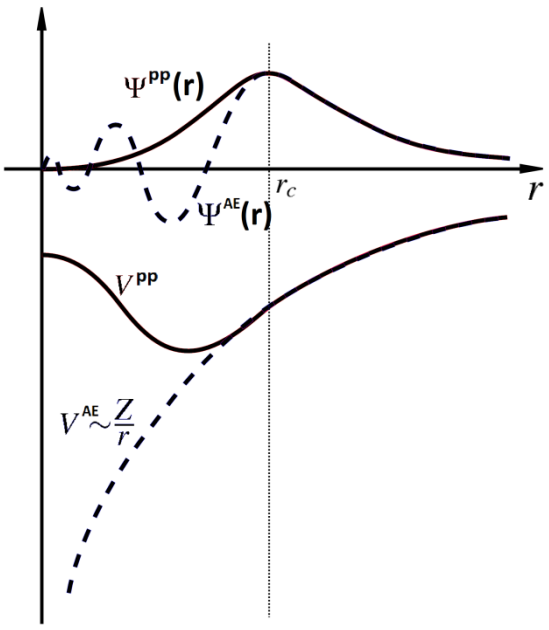

Figure 11: A schematic illustration of the concept of pseudopotentials. The dashed lines represent the full potential and the full wavefunction whereas the solid lines represent the pseudopotential and the corresponding pseudo wavefunction. In the core region the full Coulomb potential is replaced by a smooth pseudopotential which approximates the behavior of the ion cores, and above a certain radius, r_c, the two potentials coincides, resulting in the same wavefunctions. Figure adapted from ref. (Wikipedia.org), but modified.

The success of the frozen core approximation is due to the fact that the valence electrons represent most of the physical and chemical properties of the material, i.e. chemical bonding. Moreover, the tightly bound core electrons are nearly chemical inactive in the reactions. For example, the approximation is shown to give an error less than 5% in calculations of transformation energies (von Barth and Gelatt, 1980). A pseudopotential replaces the full Hamiltonian with a simpler Hamiltonian which reproduces (or at least as close as possible) the same eigenvalues and eigenvectors of the real system. (Schwerdtfeger, 2011) This will to a large extent decrease the computational effort due to reduction of plane waves in the basis set and the reduction of calculating heavy integrals associated with the ion cores. Another criterion for the pseudopotentials is that the scattering properties of the pseudopotential must be conserved for $r > r_c$. The scattering properties are given as the logarithmic derivative of the pseudo wavefunction, and shall be the same as for the all-electron case (Hamann et al., 1979).

$$\frac{d}{dr}\ln\left(\psi_i^{PP}(r)\right) = \frac{d}{dr}\ln\left(\psi_i^{AE}(r)\right) \quad , \quad r > r_c \quad .$$

The pseudopotentials make it easier to include relativistic effects directly into the potentials. However, some properties are associated with the all-electron wavefunction inside the core regions and are to some extent lost, e.g. hyperfine parameters (Bachelet et al., 1982).

Each element has its own pseudopotential, and some sources have given the full periodic table with the pseudopotentials developed for each atom, for use in ab-initio modeling (i.e. the Vanderbilt group: http://www.physics.rutgers.edu/~dhv/uspp/). Figure 12 shows a flow chart related to how

pseudopotentials are generated. The pseudopotentials act as an external background potential in the calculations. However, a pseudopotential is not unique, and numerous pseudopotentials are available for each element developed from different methods and designed for different simulation packages. The ease of creating new pseudopotentials is of course a big advantage, but the fact that different authors perform calculations with numerous pseudopotentials makes it difficult to reproduce and compare results.

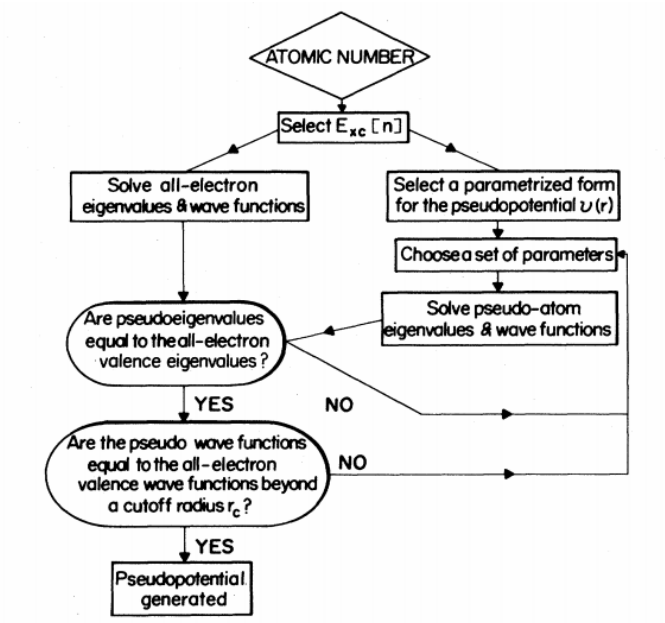

Figure 12: A schematic description of how pseudopotentials are generated. Figure adapted from ref. (Payne et al., 1992).

4.3.2 Norm-conserving and ultrasoft pseudopotentials

The most important criteria of a pseudopotential like conservation of eigenvalues and identical wavefunctions for radius longer than the cutoff radius are presented above. Another conservation criterion which is discussed in the theory is the norm-conserving condition. The early generated pseudopotentials with a hard and repulsive core failed to meet this condition. In this criterion, the total charge density of the pseudopotential needs to agree outside the core region with the all-electron potential (Hamann et al., 1979). The norm-conservation criterion is fulfilled by taking the norm of the wavefunctions,

$$\int_0^{r_c} \left|\psi_i^{AE}(\boldsymbol{r})\right|^2 d\boldsymbol{r} = \int_0^{r_c} \left|\psi_i^{PP}(\boldsymbol{r})\right|^2 d\boldsymbol{r} \quad .$$

The electrons will feel the total charge from the core, but in accordance to Gauss law, the electrons are not dependent of the distribution of charges inside the core as long as the total charge is conserved. The norm-conserving pseudopotentials (NCPP) are divided into two parts, a local and a non-local term. The non-local term can be written as a short-ranged term which gives different potential for every angular momentum.

$$V = V_{LOC} + V_{NL} = V_{LOC} + \sum_{lm} |Y_{lm}\rangle\, V_l \langle Y_{lm}| \quad ,$$

with $|Y_{lm}\rangle$ as the spherical harmonics (Kleinman and Bylander, 1982, Hamann et al., 1979). The main advantage with the norm-conserving pseudopotentials is the high degree of transferability. Transferability is a parameter to define to what extent the scattering properties are conserved, not only in the presence configuration, but in all other chemical environments (Goedecker and Maschke, 1992).

In 1990, Vanderbilt presented another approach to construct the pseudopotentials that employ a lower energy cutoff (Vanderbilt, 1990). The norm-conserving pseudopotentials fail to construct softer potentials for some of the important orbitals (O 2p and Ni 3d) under the condition that the total pseudocharge inside the core region need to match the all-electron core charge. Vanderbilt showed that the norm-conserving criterion is unnecessary by introducing a generalized eigenvalue formalism to define the charge density. Hereby the pseudo wavefunctions are split into two parts; One wavefunction that do not meet the norm-conserving condition and one wavefunction that includes a core augmentation charge to overcome the charge deficiency in the core region. The augmentation charge is defined as the charge difference between the all-electron wavefunctions and the pseudo wavefunctions (Kresse and Joubert, 1999). Pseudopotentials based on this approach can choose cutoff radius beyond the radial wavefunction maximum, as shown in figure 13. The pseudopotentials created by this method are called ultrasoft potentials (USPP). In comparison with the norm-conserving pseudopotential, USPP increases the iterations needed per computational cycle, but the basis set is reduced so the overall computational cost decreases, especially for the real-space projection techniques (Payne et al., 1992).

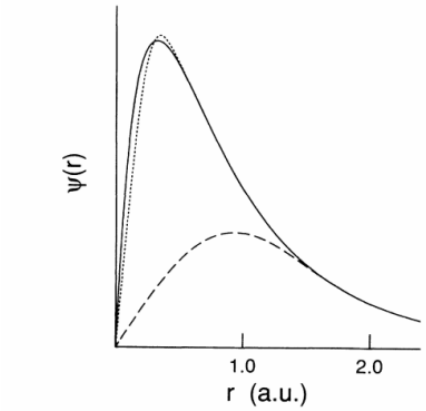

Figure 13: Schematic illustration of a USPP wavefunction (dashed line) compared with the all-electron wavefunction (solid line) and a norm-conserving potential (dotted line) for the O 2p-orbital. One of the advantages of USPP is the ability to use a longer cutoff radius. Figure adapted from ref. (Vanderbilt, 1990).

4.4 Basis set

In computational material science the molecular orbitals and the atomic potentials are described by linear combinations of mathematical functions, called *basis set*.

4.4.1 Plane waves

When the Schrödinger equation is solved for a zero or constant potential since the reference point can be chosen arbitrary, the solutions can be written as a complex plane wave, $\psi(\boldsymbol{r}, t) = Ae^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$. Plane waves are therefore suitable for calculations that involve periodic boundary conditions. In regions far from the nuclei, referred to as the interstitial region, electrons can be treated as free

particles, where plane waves are an appropriate description. However, in regions closer to the nuclei the electrons cannot be described by linear combinations of simple plane waves as the strong Coulomb potential make the norm-conservation and continuity requirement difficult in the core region and the overlapping region. To treat this problem one might introduce the idea of pseudopotentials as previously described, or or one can divide the basis set into two different regions.

4.4.2 APW, LAPW, local orbitals

As pseudopotentials separate between the core and valence regions, this separation can be done within the basis sets itself without manipulating the potential. The augmented plane wave method (APW) divides space into two regions; one region describes the sphere, S, around each atom, called the muffin tin region, where the simple plane wave description is not adequate, and the remaining region is called the interstitial region, I, where a plane wave description is applicable. In the original APW, the potential is therefore described as constant in the interstitial region, and as atomic-like, nearly symmetric, within the muffin tin region (Slater, 1953). The idea is then to describe the orbitals as a simple oscillating function in the interstitial region and a more atomic-like form within the muffin tin regions. One can obtain the following expression for the single-particle orbitals (Cottenier, 2002) p. 15;

$$\phi_{\boldsymbol{G}}^{\boldsymbol{k}}(\boldsymbol{r},E)=\begin{cases}\frac{1}{\sqrt{\Omega}}\sum_{\boldsymbol{G}}c_{\boldsymbol{G}}e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} & \boldsymbol{r}\in I\\ \sum_{l,m}A_{lm}^{\alpha,\boldsymbol{k}+\boldsymbol{G}}u_l^{\alpha}(r',E)Y_l^m(\theta',\varphi') & \boldsymbol{r}\in S_{\alpha}\end{cases}.$$

In the interstitial region, the orbitals are described by linear combinations of plane waves where Ω is the volume of the unit cell. Inside the sphere region the orbitals are described in terms of the spherical harmonics, $Y_l^m(\theta',\varphi')$ and $u_l^{\alpha}(r',E)$ which are the radial part solutions to the Schrödinger equation for a free atom α with energy E. $\boldsymbol{r}'$ is the position within the sphere, and can be written in terms of the center of each sphere $\boldsymbol{r}'=\boldsymbol{r}-\boldsymbol{r}_{\alpha}$ (see figure 14). r' is the length of $\boldsymbol{r}'$. The angles θ' and φ' specify the direction of $\boldsymbol{r}'$. $A_{lm}^{\alpha,\boldsymbol{k}+\boldsymbol{G}}$ is a complex coefficient which assures norm-conservation and meet boundary conditions related to continuous wavefunctions throughout the whole r-space.

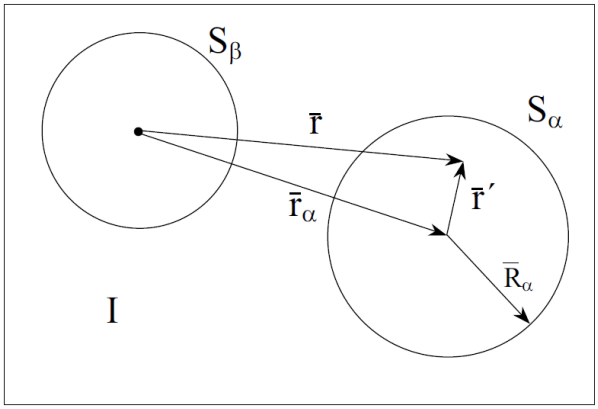

Figure 14: Schematic illustration of the two different regions in the APW approach. The parameters used in the equations are defined in the figure. S_{α} and S_{β} defines the muffin-tin region related to the core region of atom α and β, respectively. I defines the interstitial region. The black dot defines the origin of the axis system. Figure adapted from ref. (Cottenier, 2002) p. 16.

To describe the orbitals accurately within APW, the parameter E shall be equal to the band energy, $E_{n\boldsymbol{k}}$. However, $E_{n\boldsymbol{k}}$ is not known as it is the value to be calculated. Therefore, E needs to be guessed and optimized through iterations. APW is not widely used today because of slow computational time. Although it has a decreased basis set, the eigenvalues are found by matrix diagonalization for each eigenvalue and not a single diagonalization as for the plane waves.

An improved approach to consider the problems related to the unknown parameter E in the radial part is the linearized augmented plane wave method (LAPW). The method introduced by Andersen (Andersen, 1975) calculates the radial part for a given energy E_l, called the linearized parameter. As E_l is slightly different from the real band energy $E_{n\boldsymbol{k}}$, the radial wavefunction can be Taylor expanded around E_l (Cottenier, 2002) p. 21,

$$u_l^\alpha(r', E_{n\boldsymbol{k}}) = u_l^\alpha(r', E_l) + (E_l - E_{n\boldsymbol{k}})\dot{u}_l^\alpha(r', E_l) + O(E_l - E_{n\boldsymbol{k}})^2 \quad .$$

where $\dot{u}_l^\alpha(r', E_l)$ is the energy derivative of the radial part solution. LAPW therefore includes both the radial function and the derivative with respect to the energy parameter E_l in the basis functions. The basis set in the LAPW method can then be described in the following way (Cottenier, 2002) p. 22

$$\phi_{\boldsymbol{G}}^{\boldsymbol{k}}(\boldsymbol{r}, E) = \begin{cases} \frac{1}{\sqrt{\Omega}} \sum_{\boldsymbol{G}} c_{\boldsymbol{G}} e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} & \boldsymbol{r} \in I \\ \sum_{l,m} \left(A_{lm}^{\alpha,\boldsymbol{k}+\boldsymbol{G}} u_l^\alpha(r', E_l) + B_{lm}^{\alpha,\boldsymbol{k}+\boldsymbol{G}} \dot{u}_l^\alpha(r', E_l) \right) Y_l^m(\theta', \varphi') & \boldsymbol{r} \in S_\alpha \end{cases} ,$$

with $B_{lm}^{\alpha,\boldsymbol{k}+\boldsymbol{G}}$ as the complex coefficients related to the energy derivative, analogous to $A_{lm}^{\alpha,\boldsymbol{k}+\boldsymbol{G}}$. An advantage of the LAPW method is that both the function and the derivative are continuous at the muffin tin surface. This involves more plane waves, and a larger energy cutoff must be used. In the interstitial region, APW and LAPW are similar. In the LAPW method the energy bands can be calculated by a single diagonalization which makes it much faster than the APW (Singh and Nordström, 2006) p. 47-48.

A way to improve the LAPW is by introducing local orbitals within the muffin tin region, a method called LAPW+LO. The local orbital is zero outside the spheres, but includes another basis function inside the spheres (Cottenier, 2002) p. 23.

$$\phi_{\alpha,lm}^{LO}(\boldsymbol{r}) = \begin{cases} 0 & \boldsymbol{r} \in I \\ \left(A_{lm}^{\alpha,LO} u_l^\alpha(r', E_l) + B_{lm}^{\alpha,LO} \dot{u}_l^\alpha(r', E_l) + C_{lm}^{\alpha,LO} u_l^\alpha(r', E_l') \right) Y_l^m(\theta', \varphi') & \boldsymbol{r} \in S_\alpha \end{cases}$$

The local orbital is also zero within the muffin tin sphere of other atoms beside the one of interest. The local orbital is specified for a given l and m, for a given atom α. E_l' is an energy term to describe the more localized states, and $C_{lm}^{\alpha,LO}$ are the coefficients, analogue to the earlier defined coefficients, which are all zero at the muffin tip surface to meet the local orbital boundary conditions.

In LAPW, the valence states are considered, while the core states are assumed to behave as in a free atom due to the strong bonding to the nucleus. When introducing local orbitals in the basis set, one can consider semi-core states more accurate. The LAPW+LO is effective when different atoms require different energy cutoffs. When two different elements are represented in a unit cell, the highest energy cutoff of the two elements will be used in the calculation. Therefore, an impurity requiring high energy cutoff in a low energy cutoff host structure is expected to make the calculation slower.

To introduce local orbitals augmentation for the impurity can therefore reduce the basis set size and decrease the computational cost significantly. Due to the increased accuracy, the local orbitals are always included in calculations today (Singh and Nordström, 2006) p. 49.

4.4.3 PAW method

The projector augmented wave (PAW) method, introduced by Blöchl, is another approach to improve the computational efficiency in numerical calculations (Blöchl, 1994). The method connects and generalizes the use of pseudopotentials and the linearized augmented plane wave (LAPW) method and achieves the exact all-electron formalism (Kresse and Joubert, 1999). The PAW method is often used in conjunction with the frozen core approximation which treats the ion cores as rigid spheres. The PAW method applies a linear transformation operator to map the pseudo wavefunctions onto the all-electron wavefunctions. Spoken in terms of the overlying quantum mechanics, the transformation is similar to "*the change from a Schrödinger to a Heisenberg picture*" (Blöchl, 1994) where the Schrödinger picture holds the observables constant and let the quantum states evolve in time, while the Heisenberg picture keeps the quantum states constant and let the observables evolve in time (Sudarshan et al., 1961).

$$|\Psi\rangle = \mathcal{T}\left|\widetilde{\Psi}\right\rangle \quad ,$$

where $|\Psi\rangle$ and $\left|\widetilde{\Psi}\right\rangle$ are the all-electron and the pseudo wavefunctions, respectively. The linear transformation operator is given as (Blöchl, 1994)

$$\mathcal{T} = 1 + \sum_i \left(|\phi_i\rangle - \left|\widetilde{\phi}_i\right\rangle\right)\langle\tilde{p}_i| \quad ,$$

where $|\phi_i\rangle$ and $\left|\widetilde{\phi}_i\right\rangle$ are the all-electron partial waves and the pseudo partial waves, respectively. The partial waves are typically described by polynomials or Bessel functions. The transformation operator is 1 outside the augmented sphere region that encloses each atom so that the pseudo wavefunction equals the AE wavefunction outside the sphere region as described in the pseudopotential approach. $\langle\tilde{p}_i|$ is a projector function which is defined for each pseudo partial wave. The projector functions are localized in the augmented region and stand perpendicular to the pseudo partial waves. One of the challenges associated with the PAW method is the continuation of partial waves at the edge of the spheres, especially when the augmentation spheres are overlapping. The PAW method has become popular due to its great accuracy, and the method is widely used in numerical calculations today.

4.5 Dirac equation and scalar-relativistic approximation

As electrons are fermions with spin-1/2, the Schrödinger equation presented earlier is actually not valid for electrons as it is an equation for spin-less particles. In addition, the Schrödinger equation is a non-relativistic equation which means that relativistic effects are not taken into account. Dirac (Dirac, 1928) developed, a few years after the introduction of the Schrödinger equation, a similar equation that included the relativistic effects for spin-1/2 particles. The stationary Dirac equation obtains the following form

$$\left(c\boldsymbol{\alpha}\boldsymbol{p} + \boldsymbol{\beta}m_e c^2 + V(\boldsymbol{r})\right)\tilde{\psi}_k(\boldsymbol{r}) = E_k\tilde{\psi}_k(\boldsymbol{r}) \quad ,$$

where $\boldsymbol{p}$ is the momentum operator similar to the Schrödinger equation. The operators $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are 4x4 matrices related to the Pauli matrices $\boldsymbol{\sigma}$, which complicates the equation computationally. The wavefunction can be written as $\tilde{\psi}_k(\boldsymbol{r}) = \{\psi_{L,\uparrow}(\boldsymbol{r}), \psi_{L,\downarrow}(\boldsymbol{r}), \psi_{S,\uparrow}(\boldsymbol{r}), \psi_{L,\downarrow}(\boldsymbol{r})\}$ where the arrows refer to the spin-up and the spin-down components, and the letters refer to "Large" and "Small" components of the wavefunction.

As the Dirac equation is somewhat tough to use in atomic modeling due to its complexity and the complex form of the wavefunction, another approximation was developed to include relativistic effects without changing the computational demand significantly from the non-relativistic equations. This method is called the scalar-relativistic approximation and is a simplification of the relativistic Schrödinger equation (Koelling and Harmon, 1977). The scalar-relativistic approximation reduces the wavefunction from the Dirac equation into a one-component wavefunction by neglecting the small components of the wavefunction as they are only important at velocities approaching the speed of light. In addition, it assumes the spin-up and the spin-down component to be equal. In order to do so, the approximation omits the spin-orbit interaction, but keeps the relativistic effects. The spin-orbit coupling is a pure quantum mechanic effect as it is the interaction between a particle's spin and its motion which might cause shifts in its internal energy bands. From the electron's point of view, the positive nucleus is moving in orbit around the electron. From this, a magnetic field is created which affects the spinning electron. The spin-orbit coupling therefore makes everything somewhat more difficult as it introduces a coupling between the angular momentum and the spin (Griffiths, 2005) p. 271-273. However, the spin-orbit interaction can be added to the approximation as a perturbation after the bands and wavefunctions are calculated. A mathematical formula of the stationary scalar-relativistic approximation can hereby be written as

$$\left(-\frac{\hbar^2\nabla^2}{2m_e} + V(\boldsymbol{r}) - \frac{\hbar^4\nabla^4}{8m_e^3c^2} - \frac{\hbar^2\big(\nabla V(\boldsymbol{r})\big)\cdot\nabla}{4m_e^2c^2}\right)\psi_k(\boldsymbol{r}) = E_k\psi_k(\boldsymbol{r})$$

with the two last terms as the relativistic effects where the third term is the mass-velocity correction and the last term is called the Darwin term (Persson, 2012). Most codes that utilize atomistic modeling include relativistic effects i.e. from the scalar-relativistic approximation.

4.6 Vibration models

Traditional DFT calculations are based on a zero-temperature approach. However, at finite temperatures, atoms vibrate spontaneously as a result of thermal energy. Vibrations in materials are crucial for properties like thermal conductivity and electrical conductivity. Atomic vibrations are also important to identify elements in a structure. Every atom has a characteristic frequency related to the element in the given environment. Knowledge about the atomic vibrations can therefore be used to support experimental spectroscopy studies to identify elements in a structure.

Vibrations in solids can be considered at local and global level. At macroscopic level the periodic placed atoms are found to participate in collective vibrational motions. These collective lattice vibrations are called phonons, and are often treated as quasiparticles. Collective vibrations related to atomic motions in the direction parallel to the wavevector are called longitudinal modes, and collective vibrations related to atomic motions perpendicular to the wave vector are called

transverse modes. For an one-dimensional elemental solid, the phonon frequency is given by the dispersion relation (Kittel, 2005) p.92

$$\omega = \sqrt{\frac{4C}{m}}\left|\sin\left(\frac{ka}{2}\right)\right| \ ,$$

where C is the force between each neighboring atom, m is the elemental mass and k is the wavevector. The energy of a phonon mode is quantized and the energy for a phonon excited to quantum number n is $E_n = \left(n + \frac{1}{2}\right)\hbar\omega$, similar to the energy of a photon of an electromagnetic wave. The energy term $\frac{1}{2}\hbar\omega$ which are the energy at the ground state, are referred to as the zero-point energy. As lighter atoms vibrate faster due to its lower mass, the zero-point energy is of greater interest for lighter atoms. In calculations of relative energy differences, the zero-point energy can be neglected as it cancels out.

For crystals with two or more atoms per primitive basis the dispersion relation gives rise to two different solutions. In solids that consist of two different elements with two related masses, m_1 and m_2, the dispersion relation is given as (Kittel, 2005) p. 96.

$$\omega^2 = C\left(\frac{m_1 + m_2}{m_1 m_2}\right) \pm C\sqrt{\left(\frac{m_1 + m_2}{m_1 m_2}\right)^2 - \frac{4\sin^2(ka)}{m_1 m_2}}$$

This leads to two different types of modes, optical (plus sign) and acoustic (minus sign) vibrational modes as illustrated in figure 15. In a solid there are always three acoustic modes, one longitudinal and two transverse, and the rest of the modes are optical. From the dispersion relation, a gap occurs where no wave solutions are possible for any wavevectors, as the difference between the optical and acoustic mode at the edge of the first Brillouin zone.

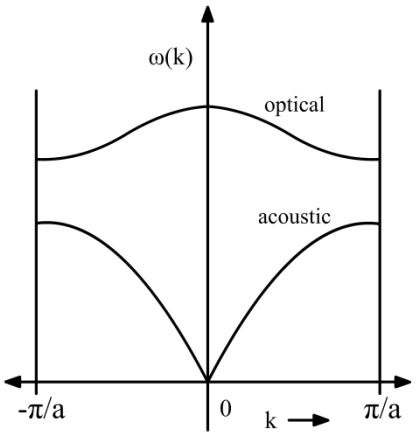

Figure 15: Schematic illustration of the two phonon modes, optical and acoustic, present in a crystal of two elements with different masses in the first Brillouin zone. Optical and acoustic modes differ in the way adjacent different atoms respond to each other. In an optical mode the different atoms move in opposite directions and in an acoustic mode the different atoms move together. For an elemental solid, only the acoustic modes are present.

First-principles calculations of the phonon frequencies are an advanced and difficult task. The most intuitive way to calculate the phonon frequencies is to let all the atoms in a very large supercell vibrate and then evaluate the frequencies associated with the overall motion. This is called the small

displacement method and calculates the force constants acting on each atom after a small displacement to calculate the frequencies. The force constants are found as the double derivative of the potential energy with respect to atomic positions evaluated at their equilibrium positions. The force constants are used to set up a so-called dynamical matrix, and the vibrational frequencies can be found as the eigenvalues of this dynamical matrix, defined as (Alfe, 2009)

$$D_{s,t}(\boldsymbol{k}) = \frac{1}{\sqrt{m_{\alpha'} m_{\alpha}}} \sum_{l} \Phi_{\alpha,\alpha',l,l'}\, exp\left[i\boldsymbol{k} \cdot \left(\boldsymbol{R}_{\alpha',l'} - \boldsymbol{R}_{\alpha,l}\right)\right] \quad .$$

$\boldsymbol{R}_{\alpha,l}$ is the equilibrium position of atom α in the primitive cell l. $\Phi_{\alpha,\alpha',l,l'}$ is the force-constant matrix element, and the summation are taken over all primitive cells in the crystal (Alfe, 2009). However, for supercells of somewhat larger size, this method will be time-consuming as all the atoms are treated separately and all the atoms need to move certain steps in all directions.

Another approach to obtain the dynamical matrix is from the linear response method (Baroni et al., 1987) which only implies calculations in the primitive cell. The linear response method uses density functional perturbation theory by the Hellmann-Feynman theorem to show that the second order change in energy implies a linear change in the charge density and potential. This linear change can be calculated by unperturbed wavefunctions and hence the calculations can be done for the ground state. The periodic perturbation can be associated with a wavevector $\boldsymbol{k}$ within the primitive cell and the linear density change can be used to determine the force constants (Alfe, 2009). By that, the dynamical matrix can be evaluated and the remaining task is reduced to an eigenvalue problem. Phonon frequencies are often treated in ab-initio simulation packages by the introduction of external codes and scripts to calculate the phonon frequencies. However, phonon calculations are not being calculated in this thesis, and the theory and issues behind the phonon calculations will not be further considered.

4.6.1 Local vibrational modes

At a local level one identifies the local vibrational modes, which are vibrations associated with atoms or small clusters of atoms in a structure. A molecule, or cluster, of N atoms has 3N degrees of freedom leading to 3N-6 vibrational modes (for $N > 2$) as three degrees of freedom are rotation about the three different axis and three degrees of freedom are translational motion. A linear molecule has 3N-5 modes, as rotation about the molecular axis is not observed (Landau and Lifshi͡ts, 1976) p.70. A diatomic molecule will therefore have only one vibrational mode, related to the stretching between the two atoms called stretch mode. For higher number of atoms several vibrational modes are possible, like bending, wagging, symmetrical and asymmetrical stretching, twisting etc.

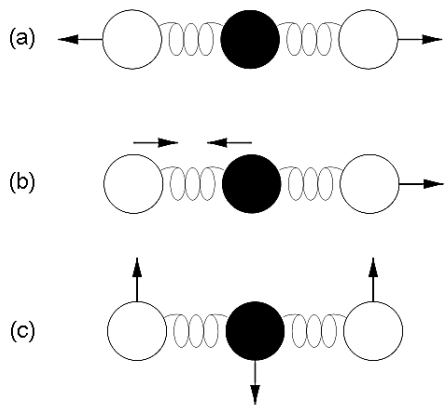

Figure 16: The local vibrational modes associated with a three-atomic linear molecule. For a three-atomic linear molecule there are four associated local vibrational modes. (a) shows symmetrical stretching, (b) shows assymmetrical stretching and (c) shows a bending mode. However, the bending mode is degenerated as the bend mode in plane and out of plane is equal. Figure adapted from ref. (UCAR, 2012)

The vibrations between atoms can be described in terms of a spring that constantly compresses and elongates. In the harmonic approximation between two vibrating atoms, the diatomic model, the spring force acting on one atom can be described by Hooke's law, $F = -kx$, where k is the spring constant. Inserted in Newton's second law of motion, $\mu\frac{\partial x(t)}{\partial t} + kx(t) = 0$, one obtains a second-order differential equation which can be solved to find an expression for the displacement of the vibrating atom. The solution implements that in the harmonic approximation the atomic motion is described by a simple harmonic function $x(t) = Acos(2\pi\nu t)$, where $x(t)$ is the displacement from equilibrium at the time t and A is the maximum amplitude from equilibrium. From this expression one finds a classical formula for the vibrational frequency

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \ .$$

Here, k is the spring constant from Hooke's law, which determines the stiffness of the spring and μ is the reduced mass between the two vibrating atoms. The reduced mass of the vibrating atoms A and B can be written as $\mu = \frac{m_A m_B}{m_A + m_B}$.

As an example relevant for this thesis, let us consider ZnO and the local vibrational modes of an interstitial impurity H atom. The H atom will bond to the nearest oxygen, to establish an $O-H$ bond. Assume that the hydrogen atom is vibrating harmonic about its equilibrium position. From a Taylor-expansion around the equilibrium, one can describe the energy as $E(x) = E_0 + \frac{1}{2}\alpha x^2$, where x is the distance from the equilibrium x = 0 and $\alpha = \frac{\partial^2 E}{\partial x^2}$ at x = 0. α is then a matrix-element of the Hessian matrix, which includes the second derivatives of all the atoms in all directions. The first-derivative term is zero, as the equilibrium position is a local minimum. By comparing the Taylor-expansion with the classical formula of the potential energy for a spring, α can be related to the spring constant between the two atoms. A higher α defines a stiffer string which means that the atoms are stronger bonded to each other. Inserted in the classical expression for vibrational frequency, one obtains the frequency of the $O-H$ stretch mode as $\nu = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}} = \frac{1}{2\pi c}\sqrt{\frac{\alpha}{\mu}} = \frac{1}{2\pi c}\sqrt{\alpha\frac{m_H + m_O}{m_H m_O}}$ (Sholl and Steckel,

2009) p. 115, with the notation cm^{-1}, assumed that the speed of light is given in cm/s. The reason for dividing by c is to obtain the common notation, cm^{-1}, used in scientific papers and experimental studies, i.e. light spectroscopy. The expression for the O − H stretch mode can be simplified by observing that $m_O = 16m_H$. In order to obtain the vibrational mode of the impurity hydrogen one assumes that hydrogen is the only moving atom, and that oxygen will remain at equilibrium at all time. This assumption is reasonable according to the larger mass of the O atom. By reducing the fraction under the square root, one obtains a simpler equation for the vibrational frequency, namely $\nu = \frac{1}{2\pi c}\sqrt{\frac{\alpha}{m_H}}$, where α is given in J/m^2, and m_H = 1.67*10^{-27} kg. This expression is nothing but the classical expression for vibration of a single object attached to a spring at a fixed position. The Hessian matrix element, α, might be negative as the curvature of the energy as a function of position might be negative. This will describe a situation where the vibrating object tends to move away from the recent position. In such a situation, the vibrational frequency will be imaginary implying an unstable configuration. Imaginary frequencies are important in the investigation of transition states described in section 4.7.

In typical DFT calculations, α is somewhat difficult to evaluate as an analytical expression as it is the second derivative of the energy with respect to the atomic positions. Instead, DFT-calculations solve α by the finite-difference approximation given, to first approximation, as

$$\alpha \cong \frac{E(x_0 + \Delta x) \; - \; 2E(x_0) \; + \; E(x_0 - \Delta x)}{(\Delta x)^2}$$

The atom is slightly moved from its equilibrium position x_0 by a distance Δx in each direction, and the energies are calculated after the displacement. For $\Delta x \to 0$, the finite-difference approximation will be exact. For numerical calculations, the displacement distance, Δx, can be chosen together with the number of displacements from equilibrium. The number of displacements from equilibrium is suggested to be even, to ensure the same number of displacements on each side of the equilibrium position. Although more displacements and smaller Δx are expected to give a better resolution and more accurate frequencies, the underlying assumption is the harmonic motion. A large number of displacements from equilibrium might lead to positions outside the harmonic region. Similarly, a large Δx might result in the same problem leading to inaccurate results according to the formula. Too small Δx can result in very small energy differences in the numerator which might be difficult to evaluate by the DFT calculation. As a rule of thumb, the energy differences should be in order of 0.01-0.10 eV. This corresponds to displacements in the finite-difference approximation at 0.01-0.10 Å which are required in order to obtain accurate values for the vibrational frequencies without further treatment and approximations (Sholl and Steckel, 2009) p. 116. For molecular vibrations, α is given in $eV/Å^2$ as the energies are given in eV and the length in Å. In VASP, the Hessian matrix can be calculated for all the atoms and directions and hence the vibrational frequencies are calculated explicitly.

For large vibrational amplitudes, the harmonic approximation for calculations of the vibrational frequency fails. However, the theory can be improved by the introduction of anharmonic effects outside the harmonic regions. Anharmonic effects are the higher order contributions to the potential energy, so $V(x) = \frac{1}{2}kx^2 + ax^3 + bx^4$. Such potentials can be solved by perturbation theory in order

to achieve an approximate analytical solution for the vibrational frequency including anharmonic effects (Landau and Lifshits, 1977).

$$\nu = \nu_H + \nu_{AH} = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} - 3\frac{1}{2\pi}\frac{\hbar}{\mu}\left(\frac{5a^2}{2k^2} - \frac{b}{k}\right)$$

However, anharmonic effects make the computational cost higher, and in this thesis the contribution from anharmonic effects will not be considered.

4.7 Transition state theory

Transition state theory (TST) is a theory to describe and calculate the reaction rates of chemical processes, and was developed in the 1930s. It is a classical theory developed from three main fields; thermodynamics, kinetics and statistical mechanics (Eyring, 1935, Laidler and King, 1983). From a chemist's point of view, the transition state theory describes a reaction from a reactant to a product. For the reaction to occur, the reactant has to pass an energy barrier obtained by atomic collisions or heating, called the activation energy. The state of the highest energy point of the reaction is called the transition state in physics, or activated complex in chemical reactions. The transition state is the least stable configuration through the reaction. When the reaction has passed the activation energy, the process continues spontaneously to the final state which is called the product.

However, the transition state theory can be used to describe several types of chemical and physical processes, not only chemical reactions. Therefore, TST can describe all types of diffusion mechanisms, like interstitial diffusion, vacancy diffusion, kick-out mechanism and so on. For diffusion processes, the activation energy is often called the diffusion barrier or the migration barrier. If one considers interstitial diffusion, the transition state will then be the least stable position of the interstitial atom on the diffusion pathway between the relaxed initial and final state. The activation energy will be the energy this atom needs in order to jump from the initial interstitial state to the final interstitial state.

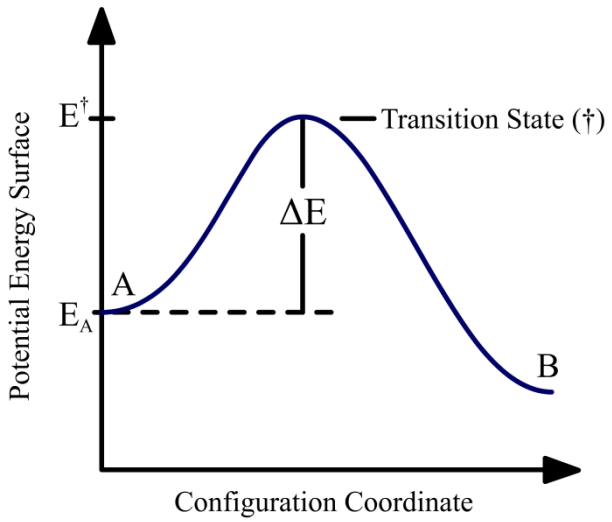

Figure 17: A schematic illustration of a diffusion process from the initial state A to the final state B through a transition state (†) described by the transition state theory. Transition state is the state with the highest energy during the process. The potential energy surface is plotted as a function of the configuration coordinate. E_A and $E^\dagger$ is the energy of the initial and transition state, respectively. ΔE is the activation energy of the process.

In a diffusion mechanism, the diffusing atom can follow a number of paths between two endpoints. One might expect that the diffusion path which gives lowest energy difference will be the most probable path. This path is called the minimum energy path (MEP). MEP is defined in the following way by Sholl and Steckel: "A path is an MEP only if the forces defined by any image along the path are oriented directly along the path" (Sholl and Steckel, 2009) p. 144. An image refers to an intermediate configuration step in the calculation between two endpoints. A MEP gives the highest diffusion rate as MEP is the path with the lowest possible energy barrier. The transition state in a MEP is a saddle point. The potential energy surface is lower in the direction of diffusion, but in all other directions there is an increase in the potential energy. From the theory of vibrational frequencies one knows that $\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$ where k is the spring constant, and m is the mass of the vibrating atom. At transition state, the spring constant will therefore be negative in the direction of diffusion due to the curvature of the potential energy surface, leading to one imaginary vibrational mode associated with the state. This leads to an important result; a transition state has one and only one imaginary vibrational mode. A way to search for true transition states is therefore to calculate the vibrational frequencies and see whether the state has one, and only one, imaginary vibrational mode.

The following derivation will follow the steps given by ref. (Sholl and Steckel, 2009) ch. 6. The transition state theory predicts that the rate for an atom to jump from site A to B is

$$k_{A\to B} = \frac{1}{2} \times (average\ thermal\ velocity\ of\ jumping\ atom) \times (probability\ of\ finding\ atom\ at\ the\ transition\ state) \quad .$$

As the diffusing atom is moving in the opposite direction half the time, statistically, one needs the pre-factor $1/2$. The probability to find an atom at any specific position is given by an exponential term dependent of temperature and the energy at the certain position. The probability of finding the atom at the transition state is therefore

$$P_{x=x^\dagger} = \frac{e^{-\frac{E^\dagger}{k_B T}}}{\int dx e^{-\frac{E(x)}{k_B T}}} \quad ,$$

where the integration is over all possible states associated with the initial position A. $E^\dagger$ is the energy associated with the transition state and $E(x)$ is the energy of a given position x. To simplify the integral in the denominator one can use Taylor expansion, $E(x) = E_A + k(x - x_A)^2/2$ around the minimum position A to approximate the energy. By the use of Taylor series of second order, the expression above can be simplified to

$$P_{x=x^\dagger} \cong \frac{e^{-\frac{E^\dagger}{k_B T}}}{\int dx e^{-\frac{E_A + k(x-x_A)^2/2}{k_B T}}} = \frac{e^{-\frac{E^\dagger}{k_B T}}}{e^{-\frac{E_A}{k_B T}} \int_{-\infty}^{+\infty} dx e^{-\frac{k(x-x_A)^2}{2k_B T}}} = \sqrt{\frac{k}{2\pi k_B T}} e^{-\frac{E^\dagger - E_A}{k_B T}}$$

To describe the average thermal velocity of a jumping atom one applies the Maxwell-Boltzmann distribution, assumed that the atom is in thermodynamic equilibrium.

$$Average\ thermal\ velocity\ of\ jumping\ atom = \sqrt{\frac{2k_BT}{\pi m}}$$

From the equations above, the following simple expression calculates the transition rate from a local/global minimum A to a local/global minimum B, in the one-dimensional case;

$$k_{A\to B} = \nu \exp\left(-\frac{E^{\dagger} - E_A}{k_BT}\right) = \nu \exp\left(-\frac{\Delta E}{k_BT}\right) \ .$$

Here, ν is the vibrational frequency of the atom in its potential minimum and ΔE is the energy difference between the transition state, $E^{\dagger}$, and the energy at the position A (local/global minimum), E_A. Due to the introduction of Taylor series of second order, the equation is only valid when the energy around the potential minimum can be treated as a harmonic motion. Therefore, this equation referred to as the harmonic transition state theory. Note that the equation of the transition rate is a variant of the famous Arrhenius' equation (Hanggi et al., 1990).

By further derivations similar to the one-dimensional case one can achieve an equation for the transition rate in the multi-dimensional case within the harmonic approximation,

$$k_{A\to B} = \frac{\nu_1 \times \nu_2 \times \cdots \times \nu_N}{\nu_1^{\dagger} \times \cdots \times \nu_{N-1}^{\dagger}} \exp\left(-\frac{\Delta E}{k_BT}\right) \ .$$

Here, ΔE is the same energy difference as defined in the one-dimensional case. ν_i are the vibrational frequencies in the position A (local/global minimum), and $\nu_j^{\dagger}$ are the real-valued vibrational frequencies at the transition state. There is one imaginary vibrational mode associated with the transition state, so there is one less real mode compared to the potential minimum as discussed above. The transition rate is the rate of sufficiently energetic jumps towards a neighboring site. It is the local vibrational frequency (Debye frequency) times a statistical factor, which considers the probability to pass the barrier. However, the transition rate does not consider the number of available sites in the surroundings and the fraction of available unoccupied sites in the crystal. The transition rate is often plotted as a linear function with the logarithm of the transition rate as a function of the inverse temperature, called a van't Hoff plot by chemists (Hanggi et al., 1990),

$$\ln k_{A\to B} = \ln A - \frac{\Delta E}{k_B}\frac{1}{T} \ ,$$

where A is the term related to the vibrational frequencies. In such plots, one can see that the energy barrier is the dominating term which decides the slope of the linear function, and a small change might lead to significant changes in the transition rate. The vibrational frequencies are typically in range 10^{12}-10^{13} Hz and shift the constant term. Therefore, the energy barrier must be calculated as precise as possible, while precise calculations the vibrational modes are less important.

Transition state theory has proven to be very successful in various areas both in physics and chemistry. However, due to its simplicity it has certain limitations. First, TST is based on classical mechanics and assumes that a certain energy barrier must be overcome in order to pass the barrier. However, from quantum mechanics there is a possibility that the barrier can be passed without the given energy, called tunneling. Tunneling is a pure quantum mechanical effect and is a function of

the barrier height. The tunneling probability decreases for larger and wider barriers. Therefore, for large barriers the tunneling effect is small and the TST should be a good approximation, but for smaller barriers, tunneling is more significant and limits the TST. Another limitation is associated with multi-barrier transitions, i.e. transitions including intermediate meta-stable states. Transition state theory is based on statistical distributions and assumes that the diffusing species are in thermal equilibrium. If the intermediate state is stable it can reach thermal equilibrium and the process can be seen as two separated transitions, well described by the TST. However, if the energy of the intermediate state is closer to the energy barrier and hereby less stable, it might not reach thermal equilibrium and the original TST fails. However, Giddings and Eyring (Giddings and Eyring, 1958) established a statistical theory for the treatment of multi-barrier transitions, discussed in the next section. As the TST assumes that the minimum energy path passes through a saddle point, problems might occur at higher temperatures. For higher temperatures the vibrational frequencies are higher and the motions of atoms are more complex. The thermal energy exceeds the energy barrier and the transition is assisted by diffusion through several transition states (Pechukas, 1981).

4.7.1 Multi-barrier transition state theory

In the presence of several transition states in a diffusion process, i.e. diffusion passing through an intermediate meta-stable state, the overall transition rate can be evaluated in terms of the single-barrier transition rates. The multi-barrier transition is divided into single-barrier transitions, denoted by k_i as the transition rate from i to i+1. From multi-dimensional TST the single barrier transition rate, k_i, is written as $k_i = A \exp\left(-\frac{\Delta E_i}{k_B T}\right)$, where A is the pre-factor related to the vibrational frequencies. Giddings and Eyring (Giddings and Eyring, 1958) showed that the overall transition rate for a multi-barrier diffusion process can be written as

$$k = \frac{1}{\frac{1}{k_0} + \frac{1}{k_1} + \cdots + \frac{1}{k_{n-1}}} = \left(\sum_0^{n-1} \frac{1}{k_i}\right)^{-1} ,$$

where k_i is the single-barrier transition rate from position i to i+1. As expected, and as the formula indicates, the largest energy barrier is the rate-determining step and gives the main contribution to the overall transition rate as it is the hardest barrier to pass. If the overall transition includes an unstable intermediate step, the contribution to the overall transition state will be smaller due to the smaller energy barrier. This is in thread with what is expected as such a potential energy surface profile approaches a single-barrier process. Figure 18 shows a schematic illustration of a three-barrier diffusion process where the overall transition rate can be calculated from the three single-barrier diffusion steps.

For a uniform multi-barrier process, all the single transition rates will be equal, $k_1 = k_2 = \cdots = k_{n-1} = k_i$. The overall transition rate can then be written as

$$k = \frac{1}{\frac{1}{k_0} + \frac{1}{k_1} + \cdots + \frac{1}{k_{n-1}}} = \frac{1}{\frac{1}{k_i} + \frac{1}{k_i} + \cdots + \frac{1}{k_i}} = \frac{k_i}{n} ,$$

where k_i is the uniform diffusion barrier for any of the single-barrier diffusions and n is the number of energy barriers in the overall multi-barrier transition process. The transition rate is then inversely proportional to the number of barriers.

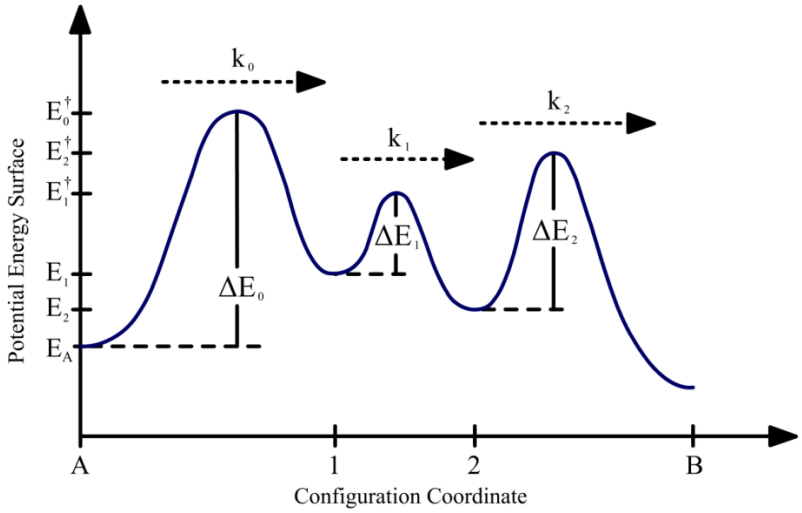

Figure 18: Schematic illustration of the concept behind the theory of multi-barrier transition rates. This illustration shows a three-barrier diffusion process which can be divided into three single transition processes, k_0, k_1 and k_2. The overall transition rate from A to B can therefore be calculated as $k_{A\to B} = \left(\sum_0^2 \frac{1}{k_i}\right)^{-1}$, where $k_i = A\exp\left(-\frac{\Delta E_i}{k_B T}\right)$ and A is related to the vibrational frequencies.

4.7.2 Random diffusion coefficient

The transition rate is closely related to diffusion and the diffusion coefficient. At microscopic level, diffusion is described by Fick´s laws. For a fixed concentration, the diffusion takes place according to Fick´s 1st law,

$$j = -D\frac{dc}{dx} \quad .$$

j is the flux density related to the number of particles crossing a unit area per second, $\frac{dc}{dx}$ is the concentration gradient which is considered as the ”driving force” for the diffusion and D is the diffusion coefficient with the dimension cm^2s^{-1}.

The following derivation of the diffusion coefficient follows the steps by ref. (Kofstad and Norby, 2012) ch. 5, and ref. (Balluffi et al., 2005) ch. 7 and 8. In the one-dimensional model at microscopic level, illustrated in figure 19, one considers two planes separated by the distance s. Let us assume that the number of particles is larger at plane 1, so $n_1 > n_2$. The actual jump frequency from plane 1 to plane 2 is denoted by Γ. The net flux density, j, of the particles jumping from plane 1 to 2 will then be

$$j = \frac{1}{2}(n_1 - n_2)\Gamma = \frac{1}{2}(c_1 - c_2)s\Gamma \quad ,$$

where c_i is the volume concentration of particles at plane i and the factor ½ is because the particle may jump in the opposite direction. In the model, the number of particles equals the volume concentration times an extension in the x-direction, i.e. the distance between the planes. As the concentration gradient between the two planes can be described as $\frac{c_2 - c_1}{s} = \frac{dc}{dx}$, one can compare the microscopic flux equation with Fick´s 1st law to obtain an expression for the diffusion coefficient,

$$D = \frac{1}{2}s^2\Gamma \quad .$$

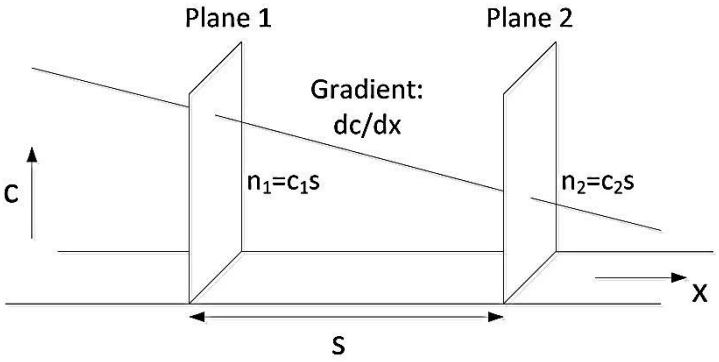

Figure 19: Schematic illustration of the one-dimensional diffusion model. Two planes are separated by the distance s and the number of particles at each plane is denoted by n_1 and n_2. In the present model, $n_1 > n_2$. The figure is inspired by ref. (Kofstad and Norby, 2012).

For the three-dimensional model, the diffusion coefficient is lowered by a factor 3 as only one third of the particles jump in the considered direction. For a three-dimensional diffusion process, the diffusion coefficient can be described by

$$D = \frac{1}{6} s^2 \Gamma \quad .$$

Γ is the jump frequency, and is closely related to the transition rate, k. The transition rate is the frequency of sufficiently energetic jumps toward an adjacent site. However, the transition rate does not say anything about whether the particles actually jump. The actual jump frequency is therefore dependent on the number of adjacent sites, Z, to which the particle can jump. This is often set to the number of nearest neighbor sites. In addition, jump frequency is dependent on the fraction of unoccupied sites, N_d, as the particle cannot jump to an occupied site. The expression for Γ will then be

$$\Gamma = N_d Z k \quad ,$$

and inserted in the expression for diffusion constant

$$D = \frac{1}{6} s^2 N_d Z k \quad .$$

For interstitial diffusion, N_d is set to unity as one expects all the interstitial atoms to be unoccupied. Hereby, the diffusion constant is related to the transition rate by the factor $\frac{1}{6} s^2 N_d Z$. The theory behind the derivation of the diffusion coefficient is based on a random walk approach. It is assumed that the diffusing particle is independent of surrounding particles, charge effects, previous steps etc., and that the movements are completely random, based solely on statistics. The diffusion coefficient based on this approach is therefore called the random diffusion coefficient. For non-isotropic systems, the transition rates differ in the different directions and hence the random diffusion coefficients are dependent on the directions.

4.7.3 Elastic band method

One of the most common methods for finding the minimum energy paths and the transition states for physical processes is by the nudged elastic band method (NEB). In order to understand the NEB, the elastic band method (EB) will first be briefly presented.

The elastic band method (Elber and Karplus, 1987) is a "chain-of-states" method based on a number of states which trace out a path between two stationary endpoints. These intermediate states are called images, and the EB searches to optimize the position and the energy of these images in order to find the MEP and the transition state. The minimum energy path is the path that connects the two endpoints with the lowest possible energy along the path. The elastic band method introduces an interaction between the neighboring images by a spring constant, k. This concept, together with evenly spaced images are treated by defining an objective function $M(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_P)$ with P images connecting the endpoints, described by (Sholl and Steckel, 2009) p. 144

$$M(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_P) = \sum_{i=1}^{P-1} E(\boldsymbol{r}_i) + \sum_{i=1}^{P} \frac{k_i}{2} (\boldsymbol{r}_i - \boldsymbol{r}_{i-1})^2 \quad ,$$

where $E(\boldsymbol{r}_i)$ is the energy of the *i*-th image. The energy of the endpoints 0 and P are not considered as they are energy minima and will be kept fixed during the calculation. k is the spring constant between image i-1 and i. The idea of the elastic band method is to minimize the objective function, $M(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_P)$, in order to move the images towards the minimum energy path. The elastic band method is highly dependent on how the initial images are chosen and how close they are to the real MEP. As the images are relaxed in the EB, the right choice of k is important as too weak spring constants might slide the images down the potential energy surface towards the minimum. This results in varying distances between the images and hence a poor approximation for the transition state (Wheeler, 2011) p. 90. However, if a too stiff spring constant is applied, one meets another challenge known as "corner cutting". During minimization the images might relax to a path which involves an overestimated transition state energy. This problem occurs when the true MEP follows a longer path than the path estimated by the EB (Sholl and Steckel, 2009, Jonsson et al., 1997).

4.7.4 Nudged elastic band method

To overcome the problems related to the elastic band method, the nudged elastic band method (NEB) includes some improvements to adjust images towards MEP more accurate and efficient. Both EB and NEB are "chain-of-states" methods and they are initialized by a set of images to connect the initial endpoints. The force acting on image i can be calculated by the derivative of the potential energy surface

$$\boldsymbol{F}_i = -\boldsymbol{\nabla} V(\boldsymbol{r}_i) \quad .$$

A path is a MEP if all the points on the path are attracted by a force only tangential to the path. If the only attraction is along the MEP, the points are at an energy minimum in all other directions, but in the direction of MEP (Sheppard et al., 2008).

For convenience one introduces a tangential unit vector, $\hat{\tau}_i$, as the vector between image i and i+1. By that, the path is a MEP if all the images have force components equal to zero in all directions except along the path (Sholl and Steckel, 2009) p. 145, (Wheeler, 2011) p. 91.

$$\boldsymbol{F}_{i,\perp} = -\boldsymbol{\nabla} V(\boldsymbol{r}_i)_\perp = -\boldsymbol{\nabla} V(\boldsymbol{r}_i) - \left(-\boldsymbol{\nabla} V(\boldsymbol{r}_i)\right)_\parallel = -\boldsymbol{\nabla} V(\boldsymbol{r}_i) + (\boldsymbol{\nabla} V(\boldsymbol{r}_i) \cdot \hat{\tau}_i)\hat{\tau}_i = 0 \quad .$$

The NEB therefore searches to relax the images along $\boldsymbol{F}_{i,\perp}$ until the component reaches zero. The parallel component of the force from the potential energy surface is projected out in order to keep even spacing between the images and is not updated in the NEB. However, from the objective

function there is also an additional force acting on image i due to the harmonic spring forces. The force acting on image i on the path can by the elastic band method be calculated by

$$\boldsymbol{F}_i = -\boldsymbol{\nabla} V(\boldsymbol{r}_i) + \boldsymbol{F}_i^S \quad ,$$

where the first term refers to the force experienced by the potential energy surface and $\boldsymbol{F}_i^S$ is the spring force between neighboring images. $\boldsymbol{F}_i^S$ can be defined as

$$\boldsymbol{F}_i^S \equiv k_{i+1}(\boldsymbol{R}_{i+1} - \boldsymbol{R}_i) - k_i(\boldsymbol{R}_i - \boldsymbol{R}_{i-1}) \quad ,$$

where k_i is the spring constant related to image i and to most extent $k_i = k_{i+1}$. To avoid "corner cutting", the spring force perpendicular to the path is projected out as it tends to move images away from the MEP due to forces from the neighboring images. Only the parallel component of the spring force is thus relevant, described by

$$\boldsymbol{F}_{i,\parallel}^S = \left(\boldsymbol{F}_{i,\parallel} \cdot \hat{\tau}_i\right)\hat{\tau}_i \quad .$$

The NEB differs from EB in these two simplifications; to avoid the problems with EB one projects out the parallel component of the potential energy surface and the perpendicular component of the harmonic spring force. This is referred to as "nudging" (Henkelman et al., 2000). When combining the equations above, the NEB method includes optimization of

$$\boldsymbol{F}_{i,NEB} = \boldsymbol{F}_{i,\perp} + \boldsymbol{F}_{i,\parallel}^S \quad ,$$

and when $\boldsymbol{F}_{i,NEB} = 0$, all images lie on the MEP. The NEB method is an iterative minimization method, meaning that it searches to minimize the forces after an initial guess. Initial images therefore need to be defined before a NEB calculation, and the closer the images are to the MEP, the faster the calculation will converge (Sholl and Steckel, 2009) p. 149. Although there are some exceptions, to most extent a linear interpolation between the endpoints is adequate. The number of images can be adjusted where an increased number of images will increase the computational demand. Even though a DFT calculation needs to be performed at each image, the calculations on each image are independent and can be calculated on parallel nodes to decrease the computer time. As mentioned above, an important parameter to define in the elastic band methods is the spring constant. For the EB method, the choice of k was quite sensitive in calculating the barrier height. For the NEB method, where only the spring force parallel to the path is considered, the choice of k does not have the same sensitivity on the barrier height. It has been shown (Jonsson et al., 1997) that the barrier height is converged to the same value with a spring constant ranging over three order of magnitudes (k = 0.01 to k = 20). The spring constant can also be set to zero in order to obtain faster calculations, but then the NEB method will not assure an even spacing between images, which may result in a poor resolution of the MEP and convergence problems. As a non-zero spring constant does not increase the computational time significantly, it is recommended to avoid a spring constant equal to zero (Jonsson et al., 1997).

Lately, there has been developed an improved NEB method, called the climbing nudged elastic band method (cNEB). The cNEB method is developed by Jónsson and Henkelman's group and is supposed to give a more precise transition state (Henkelman et al., 2000). In the NEB method, the transition state is not necessarily found, and is dependent on the number of images and the position of the initial images. However, in the cNEB method the spring force is removed, and the potential energy

force is inverted. The force is then maximized parallel to the band and minimized perpendicular to the band which makes the images climb towards the saddle point. Due to the movement along the band, the images are no longer evenly spaced, and the cNEB method should be obtained after a minimization of a regular NEB calculation. From that, the shape of the MEP is known, and the cNEB method can be applied to identify the accurate transition state. The cNEB method has been implemented in the VASP code by Jónsson and Henkelman's group (Henkelman et al., 2000, Sheppard et al., 2008).

In this thesis, the concepts of the transition state theory is used to calculate the diffusion barriers and the corresponding MEP by the NEB method, see chapter 4.

4.8 Defects and impurities

Introduction of defects and impurities in a material can give interesting new features, and are also responsible for certain properties in crystalline solids. For example, vacancies will lead to higher mobility of the ions due to increased number of available sites, hence increase the ionic conductivity and also ensure multiple pathways for diffusion. A defect is defined as a deviation or an imperfection from the ideal structure, and defects are present in all types of crystalline materials at any temperatures (Kofstad and Norby, 2012) p. 1-11. The appearance of defects in crystalline solids can be explained by a simple thermodynamic approach. The change in Gibbs free energy given by $\Delta G = \Delta H - T\Delta S$ where ΔH is the enthalpy change and ΔS is the entropy change at temperature T. In an ideal structure without any defects there are few degrees of freedom due to a close packed structure without vacancies. The introduction of defects will increase the entropy, and especially for the first defects. Even though the enthalpy increases by the number of defects, the entropy gain is larger so that the total free energy is lowered up to a certain number of defects, which is illustrated in figure 20.

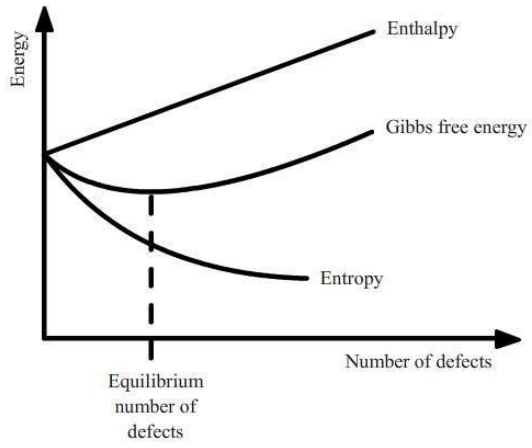

Figure 20: A schematic illustration of the Gibbs free energy as a function of the number of intrinsic defects in the material. The relaxed structure has a finite number of intrinsic defects due to the large gain of entropy for the first introduced defects.

One divides the defects into intrinsic and extrinsic defects. Intrinsic defects are defects created by the internal structure without any external influence. These defects occur naturally in all materials at $T > 0\ K$, and the number of intrinsic defects increases by an increased temperature due to vibrations and higher thermal energy in the system. Intrinsic defects can be internal point defects like a vacancy or a deviation from the stoichiometry by a displaced atom from its lattice position to an

interstitial position. Point defects are often formed in a stoichiometric ratio to preserve electroneutrality in ionic compounds (Schoonman, 2005). Typical intrinsic point defects are Schottky disorders, which are a double vacancy created by a cation and an anion, and Frenkel disorders, in which an atom displaces from its lattice position to an interstitial position leaving a vacancy behind. Intrinsic defects can also be higher-dimensional ordered defects like one-dimensional dislocations created by disordered atoms in certain directions, and two-dimensional defects like grain boundaries, stacking faults and internal interfaces.

Extrinsic defects are defects caused by dopants and impurities in the host structure, thus changes caused by external contributions. Extrinsic defects also include deviations from stoichiometry which are common in transition metal oxides. A dopant, or an impurity, can take different positions in the structure leading to different properties. Mainly, one divides into substitutional and interstitial impurities. Substitutional impurities take the position of a host atom by a kick-out mechanism or at an already empty structural site. Interstitial impurities are foreign elements positioned in different empty interstitial sites in the material. There are a specific number of interstitial sites in a structure with various sizes, and some are more favorable. An interstitial impurity will increase the number of atoms in the supercell compared to the original structure. Extrinsic defects can also be higher-dimensional, for instance the interface between two different materials as a two-dimensional defect.

In addition to structural defects, there are also electronic defects like defect electrons and electron holes. Electronic defects are important to determine electronic conductivity, optical properties, etc. Defect electrons and holes can occur due to extrinsic point defects which introduces different electronic configurations than the host material. In semiconductor science the electronic defects are of high importance. For instance, in silicon technologies, silicon can be acceptor-doped by boron which introduce electron holes to the silicon and donor-doped by phosphorus to introduce defect electrons. Electronic defects can also occur by internal excitation of valence electrons in the host material. The electronic defects are able to move through the crystal and can either be delocalized and smeared out through a large area of the crystal or localized at specific atoms.

To describe defects in this thesis there will be used a simplified Kröger-Vink notation (Kroger and Vink, 1956). The element considered is written by its chemical formula, with a subscript to denote the site where the element is placed. *i* means interstitial site and *v* denotes a vacancy site. Li_{Zn}, Li_i and v_{Zn} denote a substitutional Li atom at a Zn site, an interstitial Li atom and a Zn vacancy, respectively. In the original Kröger-Vink notation the effective charges are considered in order to preserve electroneutrality. A substitutional Li at a Zn site would then be denoted as $\mathrm{Li}'_{\mathrm{Zn}}$, where the prime denotes a negative effective charge since Li is lower-valent compared to the host Zn atom. In this thesis, only the real charges will be denoted, so $\mathrm{Li}'_{\mathrm{Zn}} = \mathrm{Li}_{\mathrm{Zn}}$, and a negatively ionized substitutional Li will be written as $\mathrm{Li}^{-}_{\mathrm{Zn}}$.

4.8.1 Formation energy

In defect physics, the formation energy is the energy required to create a defect. Some defects have low formation energy as they have a positive impact on the material i.e. through higher entropy by forming vacancies, and some defects are difficult to form and will therefore have a higher formation energy. The overall definition of formation energy of defects is given as

$$\Delta H_f = E_d - E_0 \quad ,$$

where E_d is the total energy of the system including the defect, and E_0 is the total energy before the defect was introduced, thus the reference energy. The energies E are the Gibbs free energies, which can be written as $E = E_{coh} - TS + PV$. E_{coh} is the cohesive energy of the system which is the energy required to break all the bonds in the structure, sometimes referred to as the binding energy. The cohesive energy is the energy calculated as the total free energy in VASP. S is the entropy of the system which may change by the introduction of defects. V is the volume of the system which may also change during the process. T and P have their usual meaning; temperature and pressure, respectively. Inserted in the formula for formation energy, and assuming a constant temperature and pressure one obtains

$$\Delta H_f = (E_{d.coh} - TS_d + PV_d) - (E_{0,coh} - TS_0 + PV_0)$$

As a reasonable approximation for large supercells or for small defect concentrations, the relative change in the entropy and the volume are so small that they can be set equal, namely $S_d = S_0$ and $V_d = V_0$. The formation energy can therefore be expressed as the change in the cohesive energy of the system,

$$\Delta H_f = E_{d,coh} - E_{0,coh} \quad .$$

However, this expression assumes that the number of atoms is equal before and after the formation of the defect, hence it is only valid for intrinsic defects. A defect in a system can also be external, like a vacancy or an interstitial foreign atom in the structure. By the definition given above for the cohesive energy, one cannot compare the energy between two systems with different number of atoms. When a vacancy is introduced the system will have one less atom compared to number of atoms before the defect was introduced. To find the formation energy one needs to take the number of atoms into account. In the case of introduction of a vacancy one needs to consider the removed atom in a reference state after the atom is removed. Indeed, the formation energy should be dependent on how easy it is to create a defect of that element in the given structure, and this is given as the chemical potential. The chemical potential of an atom is defined as the energy of a reservoir system containing only that specific atom, and can be directly related to the energies calculated from the first-principle methods (Van de Walle, 1997). Let us assume that a Zn vacancy is to be introduced in the ZnO structure. In order to calculate the formation energy of that defect, one needs to consider the chemical potential of Zn, and the energy would be given as (Van de Walle, 1997)

$$\Delta H_f = \left(E_{d,coh} + \mu_{Zn} N_{Zn}\right) - E_{0,coh} \ ,$$

where μ_{Zn} is the chemical potential of Zn and N_{Zn} is the number of defects. $N_{Zn} = 1$ in the case of formation of a single vacancy. $E_{d,coh}$ is the cohesive energy of the system including the Zn vacancy so the number of atoms considered is now equal before and after the defect formation. $E_{0,coh}$ is the energy of the system before the formation of the Zn vacancy.

The chemical potential is the energy required to remove the atom out of the material, or similar, how easy the atom will flow into the material. More general, one can say that the chemical potential is a parameter for the tendency of a substance to change. Therefore, a higher chemical potential means that the substance is more likely to change its state, i.e chemical composition, location, or aggregation state (Job and Herrmann, 2006). The chemical potential can be visualized as a reservoir

outside the material with a hill into the material. For a higher chemical potential, the hill is steeper so the atoms flow more easily into the structure. If the chemical potential is low, the hill is upward from the reservoir to the material, and the atoms are easily removed from the structure (see figure 21). This implies that atoms and molecules move from high chemical potential to low chemical potential. As a relevant example for this thesis, let us consider a substitutional Li at a Zn site, written as Li_{Zn}, and evaluate the formation energy of this defect. A schematic illustration of this process including the reservoirs can be seen in figure 21. The chemical potential decides the height of the reservoir, and hereby the ease of the element to be introduced in the material. In order to keep the number of atoms fixed before and after the process, the perfect ZnO structure plus an isolated Li must be calculated as the reference energy. It is not sufficient to have the same amount of atoms; the number of each element must be equal before and after the process. The total energy including the Li_{Zn} plus the chemical potentials of the removed Zn and the added Li must be calculated. The total formation energy is then given by

$$\Delta H_f = \left(E_{d,coh} + \mu_{Zn}N_{Zn} + \mu_{Li}N_{Li}\right) - E_{0,coh} = E_{d,coh} + \mu_{Zn} - \mu_{Li} - E_{0,coh} \quad ,$$

where in the last equation $N_{Zn} = 1$ and $N_{Li} = -1$ due to the removing of one Zn atom and the introduction of one Li_{Zn}. Here, $E_{d,coh}$ refers to the total cohesive energy of the system containing a Li_{Zn} defect.

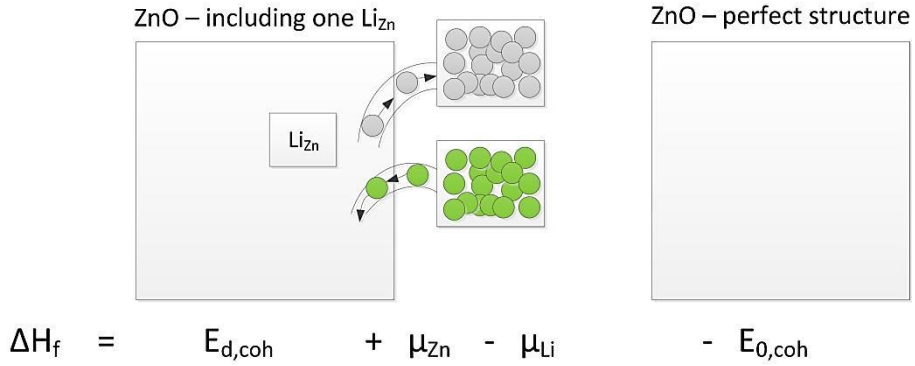

Figure 21: Schematic illustration of the concept behind calculation of the formation energy of a Li_{Zn} defect in ZnO. The left system includes the substitutional defect whereas the right system is the perfect structure without the introduced defect. The upper reservoir (grey balls) contains the Zn atoms and the lower reservoir contains (green balls) contains the Li atoms. The arrows show that a Zn atom is moved from the crystal into the reservoir, and a Li atom is placed on the empty Zn site in the structure. The formula for formation energy for the given system is written below the illustration as a reference.

4.8.2 Ionization energy

The link to ionization energy from the formation energy is simple. The ionization energy at atomic level is the energy required to remove an electron from the atom. For a larger system one can define the ionization energy as the energy required to remove (or add) an electron from the neutral system, or the reference system.

$$\Delta H_{ion} = E_{ion} - E_0 \quad ,$$

where E_{ion} is the energy of the ionized system and E_0 is the energy of the reference system.

Here, it is assumed that the two systems have equivalent number of atoms. However, the above equation is only valid for systems with the same number of electrons before and after the reaction, which for sure is not true for an ionization process. The ionization energy must therefore consider

the electrons which are removed from (or added to) the crystal. The energy of the ionized system consists of a different number of electrons which cannot be compared with the non-ionized system. Like chemical potentials for atoms, one assumes a reservoir where the electrons are moved into, and that the total energy of the ionized system is the energy of the ionized crystal plus the energy of the removed electron in the reservoir. The electron energy in the reservoir can be regarded in a similar way as for the atoms by the chemical potential of electrons, μ_e. The chemical potential of electrons is related to the thermodynamic work required to add one electron to a system, neglecting the work required to remove it from its current state. The chemical potential of electrons is often called the Fermi level, E_F, not to be confused with the Fermi energy. In terms of the electronic band structure, the Fermi level is a hypothetic energy level where the probability for this energy level to be occupied is 50 % at any given temperature, given by the Fermi-Dirac distribution

$$f(\varepsilon) = \frac{1}{e^{(\varepsilon - E_F)/k_B T} + 1} \ .$$

If there exists a state at the Fermi level, $\varepsilon = \mu_e = E_F$, this has a 50 % probability to be occupied for the given temperature according to the Fermi-Dirac distribution. However, the Fermi level is not always an existing energy state. For insulators, the Fermi level at normal temperatures lies in the bandgap and refers to a hypothetic energy state. For metals and semimetals, the Fermi level refers to a specific state, and for semiconductors the Fermi level is close to the bandgap edge, as some states might be thermally excited due to introduction of new energy levels as a cause of doping.

The ionization energy for a given system, where the number of electrons is taken into account, can therefore be written as

$$\Delta H_{ion} = (E_{ion} + E_F N_e) - E_0 \ ,$$

where N_e is the number of electrons that are removed from the system. $N_e = 1$ for a single donor ionization process, and $N_e = -1$ for an single acceptor ionization process. $\mu_e = E_F$ refers to the chemical potential of electrons and the remaining terms have their earlier meaning.

In the example given in the previous section, one can now take it a step further and calculate the formation energy of a substitutional single negatively charged Li atom at a Zn site, namely $\mathrm{Li}_{\mathrm{Zn}}^{-}$. In this case, both the chemical potential of the atoms in the reservoir must be considered in addition to the chemical potential of the electrons (Van de Walle, 1997). The total formation energy will therefore be given as

$$\Delta H_f = \left(E_{d,coh} + \mu_{Zn} N_{Zn} + \mu_{Li} N_{Li} + E_F N_e\right) - E_{0,coh} = E_{d,coh} + \mu_{Zn} - \mu_{Li} - E_F - E_{0,coh}$$

Here, $E_{d,coh}$ refers to the total cohesive energy of the system containing a single negatively charged $\mathrm{Li}_{\mathrm{Zn}}$ defect. If the substitutional atom was positively charged, the sign in front of E_F will change as the ionized system donates electrons to the crystal.

5. Results and discussions

In this chapter, the results will be presented and discussed. The discussion of the results is divided in two parts. In this chapter the present results will be discussed and evaluated when appropriate. However, the results in terms of the big picture will be brought to light in the next chapter.

At first there will be a focus on the host structure, ZnO. Before defects can be introduced one needs to relax the host structure, and check convergence parameters for the given structure. After convergence parameters are established, a geometrical optimization is performed to calculate the relaxed lattice parameters and the ionic positions. Next, dopants are introduced and hydrogen is the first element considered. The presence of hydrogen in ZnO has been described by earlier published reports. However, the calculations are included as a reference to compare. In addition, the same calculations are performed with a GGA potential to improve the existing values.

Next, lithium in ZnO will be comprehensively studied as the main research objective in this thesis. Substitutional and interstitial Li, as well as some Li-complexes are described to a large extent. Relaxation of the structure with introduced Li is performed first in order to find the most stable positions of Li, before the local vibrational frequency modes of Li are calculated. The diffusion profiles and the corresponding transition rates of diffusion for different Li configurations are of major interest in order to understand the mobility and trapping possibilities of Li in ZnO. The diffusion coefficients for the interstitial Li diffusion are calculated. The diffusion of interstitial Li is studied through the nudged elastic band method (NEB) for both the neutral and the charged states. Mobility of substitutional Li, which is suggested as a promising *p*-type dopant, will be considered through the kick-out mechanism of $Li_{Zn} - Li_i$ and $Li_{Zn} - Zn_i$ pairs.

Last, Ni in ZnO is studied to some extent. In this thesis only Ni_{Zn} will be considered and geometric optimizations and vibrational frequencies are calculated and evaluated. Investigation of Ni-doped ZnO co-doped with Li will be considered, and the structural relaxation and magnetic properties as a function of Li concentration will be presented.

All calculations in this thesis are based on DFT within the *Vienna Ab initio Simulation Package* (VASP). Different types of exchange-correlation functionals have been utilized, like LDA potentials and PBE potentials for the relaxation of the host structure and the H calculations in ZnO. For the Li and the Ni calculations in ZnO, the GGA with the PBE potentials have been utilized.

5.1 Relaxation of ZnO

Before one can introduce defects and impurities in the compound, the host structure must be relaxed, hence be in its ground state. One differs between electronic and ionic relaxation. Electronic relaxation is relaxation of the electrons and yields the electron distribution with the lowest possible energy. Ionic relaxation is to relax the ions in the supercell by moving atoms with respect to each other and obtain the optimized ionic configuration with the lowest possible energy. In an ionic relaxation, the lattice constants can also be relaxed with respect to energy. Ionic relaxation always involves an electronic relaxation.

In order to avoid interactions between defects in the supercell, a large supercell as possible is to prefer. Supercells are repeated in all directions to build the crystal. Introduction of a single defect in a small supercell, i.e. the primitive unit cell, would create the right host structure, but would lead to a large concentration of defects as the defect would be repeated in every unit cell. In this thesis, a supercell consisting of 3x3x3 unit cells was used to avoid this problem. Therefore, a single defect introduced in this large supercell, led to a defect concentration of 1/27 compared to the case if a primitive unit cell was used, and hereby a more realistic doping concentration. In treatment of a single-defect it is desirable to use as large supercell as possible, but compromise must be taken as larger supercells require longer computational time.

When the size of the supercell was decided, one needed to check the convergence of the total energy with respect to the energy cutoff and the number of k-points, as described in section 4.2. In order to reproduce and compare calculations with earlier published results it is good to use the same size of the supercell and the same convergence parameters, since different parameters lead to higher or lower accuracy relative to the real values. Convergence of the electronic relaxation with respect to the different convergence parameters is of high importance in order to check the accuracy of the calculated values. Higher energy cutoff and higher number of k-points will give higher accuracy and more reliable values, but there is a price to pay in terms of more computing power, so a compromise must be taken. The convergence tests in this thesis were performed within the GGA with PBE potentials.

The calculated total energy per atom with respect to energy cutoff can be seen in figure 22. As the figure indicates, the absolute convergence was converged within 10 meV per atom for an energy cutoff at 400 eV, and within 1 meV for an energy cutoff at about 500 eV. However, when comparing results one is interested in the relative energy difference between two structures. Since the same type of errors tend to occur in the same type of calculations, relative energy differences tend to converge much faster than the absolute value. The absolute convergence of an asymmetric structure is somewhat more difficult to converge than a symmetric supercell because the full BZ cannot be reduced into an irreducible BZ which simplifies the calculations by symmetric considerations. As a result, larger cutoff energies are required to achieve the same convergence for asymmetric structures. Therefore, one atom was slightly displaced to break the symmetry of the structure and then the total energy was calculated as a function of the energy cutoff. The relative energy difference between the perfect and the slightly displaced supercell were converged within 1 meV per atom already at an energy cutoff at 300 eV, as seen in figure 22. In this thesis an energy cutoff at 400 eV were used unless otherwise stated, and this value seemed to result in well-converged values for the relative energy differences, and for the total energy within 10 meV. Some of the references in this thesis also used 400 eV which confirmed the use of the given energy cutoff.

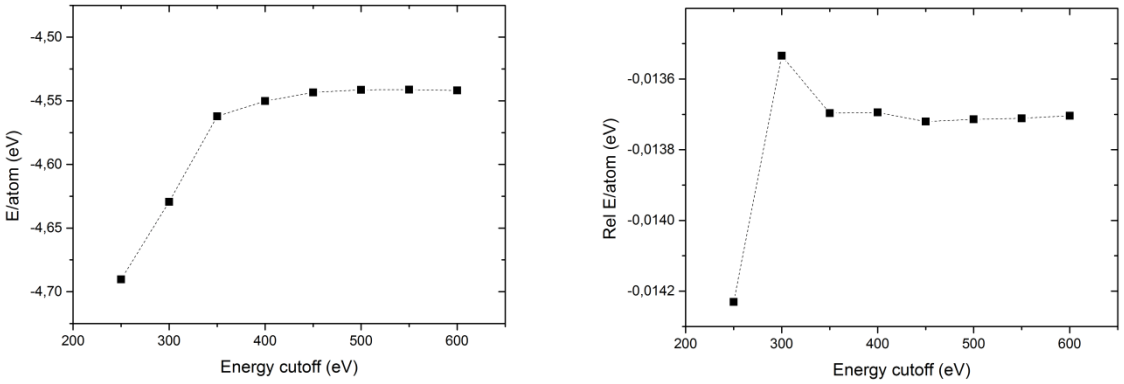

Figure 22: Left: Convergence of the total energy per atom with respect to the energy cutoff. At 400 eV, the total energy was converged within 10 meV, and at ~500 eV, the total energy was converged within 1 meV. Right: Convergence of the relative energy difference per atom between a perfect symmetric ZnO structure and a ZnO structure with a broken symmetry with respect to the energy cutoff. The relative energy difference converged at a lower energy cutoff compared to the absolute energy difference due to systematic errors in similar numerical calculations.

In figure 23, the total energy per atom has been calculated with respect to the number of k-points in each direction. Since the supercell is symmetric it was reasonable to use the same number of k-points in each direction. Already at a 2x2x2 k-point mesh, an adequate convergence was met within an accuracy of 1 meV for the total energy. While the relative energy differences converge at a lower number of k-points, 2x2x2 k-mesh seems like a useful convergence parameter. The reason for the quick convergence is due to the use of a large supercell. For larger supercells, fewer k-points in each direction are required due to the reciprocal relation between the real and the reciprocal space. A 2x2x2 k-point mesh in a 3x3x3 supercell corresponds to a 6x6x6 k-point mesh for a unit cell. This thesis has used a 2x2x2 k-point mesh unless otherwise stated, as it produced well-converged values and simultaneously saved some computational time compared to a higher number of k-points.

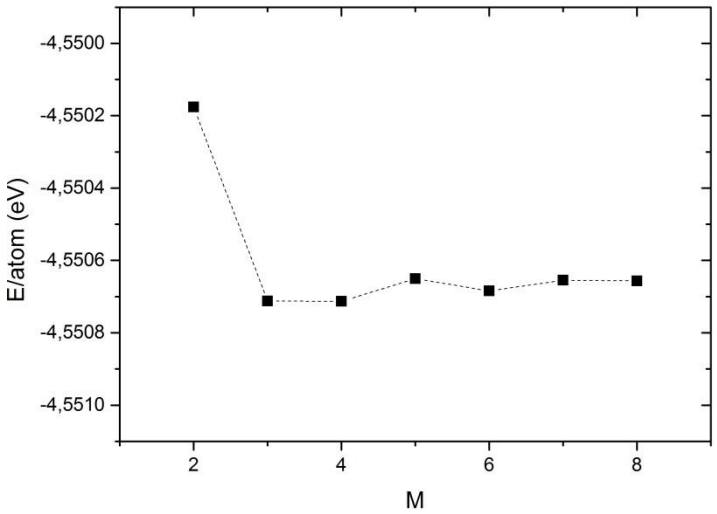

Figure 23: Convergence of the total energy with respect to the number of k-points M in each direction. The number of k-points was symmetric in each direction, resulting in a MxMxM k-mesh. High convergence was obtained at a small k-mesh due to the large supercell used in this thesis.

The initial supercell was built by three unit cells in each direction, creating a 3x3x3 supercell. Before the volume and the ionic relaxations of the supercell, initial lattice parameters based on the experimental lattice constants and the configuration coordinates, as presented for the unit cell given in section 2.6, were utilized. The lattice constants of the primitive cell were multiplied by three in order to obtain the lattice vectors of the 3x3x3 supercell. As the unit cell consisted of 4 atoms, the supercell used in this thesis consisted of 108 atoms. This increased the computational time significantly, but a large cell was required to avoid any defect-defect interactions between the supercells.

The ionic relaxation of the supercell was performed with an energy cutoff of 400 eV and a 2x2x2 k-point mesh in order to find the relaxed lattice parameters and the configuration coordinates of the ions. The relaxation used a quasi-Newton algorithm, and the tolerance of the ions set to 0.01 eV/Å with respect to forces, which means that the relaxation will terminate when all forces on the atoms are converged within 0.01 eV/Å. Another way to converge the ions is with a requirement on the total energy. The lattice constants were allowed to change together with the *c*-coordinate of the oxygen atoms, in order to find the parameter u. A quick calculation of the unit cell indicated that the oxygen atoms were positioned directly above the Zn atoms with only the *c*-coordinate changed, as expected from symmetrical considerations. As a result, the basal coordinates of the oxygen atoms, together with the position of the Zn atoms were held fixed in order to save computational time for the relaxation of the 3x3x3 supercell. The relaxation was performed within the GGA with the PBE potential as well as within the LDA.

The relaxed 3x3x3 supercell is depicted in figure 24, and the calculated lattice constants and the u parameter together with the internal bond lengths from the GGA and the LDA calculations are presented in table 2. Experimental values are given for comparison. The experimental bond lengths were found by evaluating the experimental lattice constants together with the u parameter, which determines the position of the oxygen atoms (Yao and Hong, 2009) p.5. As expected, the calculated lattice parameters, and hence also the internal bond lengths within the LDA approximation were underestimated by $\sim$2 % from the experimental constants. The GGA lattice constants were closer to the experimental values, although somewhat higher, but all the values were within 0.5 % from the experimental values. As seen from table 2, there was a difference between the Zn $-$ O bond length parallel to the *c*-axis and the other tetrahedral Zn $-$ O bond lengths. The u parameter was therefore larger than in the ideal hcp structure, i.e. c/a = 1.63 and $u \approx 3/8$. The elongation in *c*-direction is due to the interlayer distances in the *c*-direction. Seen from the oxygen, the distance from the overlying Zn layer is about three times smaller than the underlying Zn layer in *c*-direction (see figure 25). The overlying Zn layer can easily shorten the distance to the O layer by angle deformation between the Zn $-$ O bonds, whereas the underlying Zn layer must be moved directly towards the O layer in order to shorten the interlayer distance. The former is likely to happen and the Zn $-$ O bond is therefore slightly longer in *c*-direction.

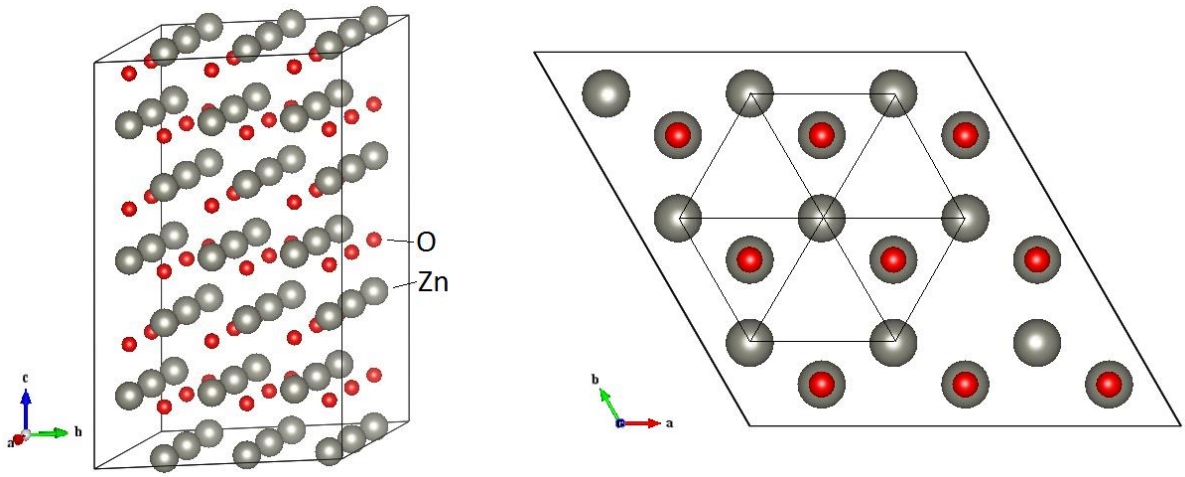

Figure 24: A schematic presentation of the relaxed 3x3x3 supercell of ZnO used in the thesis. In the right figure, the supercell is depicted along the c-direction.

Table 2: The calculated lattice parameters in the 3x3x3 supercell in comparison with experimental values by ref. (Karzel et al., 1996). The parameter u decides the position of the oxygen atoms. Also, the calculated Zn − O bond lengths are compared with the experimental bond lengths. The experimental bond lengths were obtained by a geometrical evaluation of the experimental lattice constants together with the u parameter (Yao and Hong, 2009) p. 5. As expected, the lattice parameters were somewhat underestimated within the LDA approximation, but the GGA approximation showed good consistency with the experimental values. A small difference in the bond lengths between the bond parallel to the *c*-axis, $\mathrm{d}_{\parallel}$, and the other nearest neighbor bond lengths, $\mathrm{d}_{\perp}$, was found.

	Lattice parameters				**Bond lengths {d_{Zn-O}}**	
	a [Å]	c [Å]	c/a	u	$\mathrm{d}_{\perp}$ [Å]	$\mathrm{d}_{\parallel}$ [Å]
Present – GGA	3.257	5.232	1.606	0.380	1.982	1.991
Present – LDA	3.180	5.124	1.611	0.379	1.937	1.943
Expt.[1]	3.250	5.204	1.601	0.382	1.974	1.988

[1] Ref. (Karzel et al., 1996)

The wurtzite structure of ZnO is based upon ABABAB… packing, which means that all the C positions are empty and available to be occupied. The C positions are also called octahedral positions (O). A schematic illustration of the ZnO structure is depicted in figure 25, with the relaxed interatomic distances calculated from the GGA calculation. An additional reference test was performed by relaxing the primitive unit cell with a 6x6x6 k-point mesh. The lattice parameters calculated from the unit cell was in accordance with the calculated lattice parameters from the supercell within 1 mÅ as was expected.

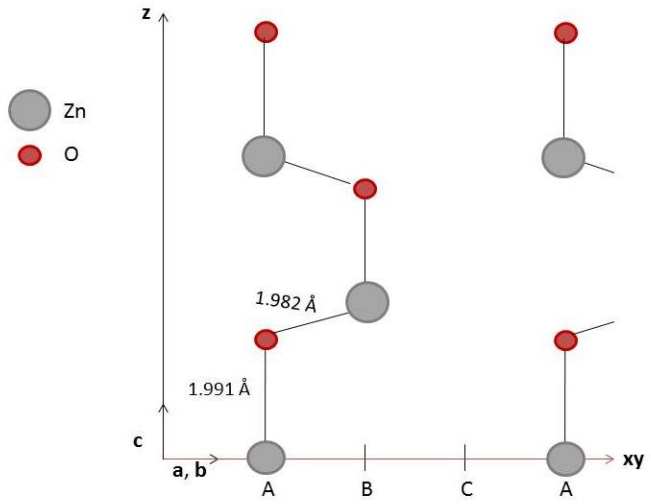

Figure 25: A schematic illustration of the ZnO structure with the relaxed Zn-O distances from the GGA calculation. The large (grey) circles represent Zn atoms, whereas the smaller (red) circles represent oxygen atoms. As seen from the figure, the C plane is empty and available to be occupied by interstitial atoms.

As a reference to the coming calculations, the internal local vibrational frequencies of ZnO were calculated. The stretch mode of the Zn − O bond was calculated by allowing both atoms to move towards and away from each other and evaluating the energy as a function of the displacement, as described within the finite-difference approximation. In this approximation, the displacement is assumed to be in the harmonic region. A stretch mode is related to the frequency in which the atoms vibrate against each other, and can give relevant information about the bond strength of the given bond. A large frequency indicates a strong bond as the vibrational frequency is related to the curvature of the energy as a function of displacement. If the energy is changed significantly for a small displacement, i.e. strongly bonded, the curvature is large which results in a high frequency. The stretch mode of the Zn − O bond in the host structure was calculated to $\nu = 441$ cm^{-1}. The stretch mode was primarily located at the oxygen atom, which was expected as the zinc mass is four times as large as the oxygen mass. The calculated frequency was in accordance with experimental results (Malaeru et al., 2012).

5.2 Hydrogen in ZnO

Hydrogen can be incorporated in ZnO during hydrothermal growth of the material or from other sources as hydrogen is present in most environments. Hydrogen is an active impurity in ZnO. It is found to diffuse easily into the material, and hereby increase the conductivity of ZnO. (Thomas and Lander, 1956, Mollwo, 1954).Thomas and Lander reported that hydrogen however will not easily diffuse through the interstitial holes itself, but rather bond with the oxygen ion to form an hydroxyl ion, OH^-, which makes the diffusion somewhat less probable.

In this thesis, only hydrogen at interstitial sites was considered. Hydrogen is small and can easily be placed in different interstitial sites in ZnO. Substitutional hydrogen was not considered due to the large size difference between hydrogen and the host Zn-atom. Van de Walle (Van de Walle, 2000) confirms this by predicting that interstitial hydrogen is the dominated type of H defect in ZnO. Hydrogen is an amphoteric impurity as hydrogen can act as both a donor, H^+, and as an acceptor, H^-. However, hydrogen is not amphoteric in ZnO, which means that it cannot obtain different charge

states. Hydrogen is positively charged and acts exclusively as a shallow donor in ZnO. ZnO is naturally an *n*-type semiconductor. The native *n*-type behavior is expected to occur by the presence of interstitial H in the structure (Van de Walle, 2000). It is not accepted as the only reason, but as the H is present in almost all growth methods of ZnO, it is accepted as the main reason for the *n*-type behavior of native ZnO.

A few reports have presented the relaxed structure and the local vibrational modes for H^+ in ZnO (Limpijumnong and Zhang, 2005, Van de Walle, 2000, Lavrov et al., 2002). The first part of this thesis focused on reproducing the data done by previous researchers, as well as improving the existing data by using improved potentials. For the calculations of hydrogen in ZnO, DFT was used within the LDA and GGA, and both the Vanderbilt-type ultrasoft pseudopotentials and projector augmented wave method (PAW) were utilized. All the approximations are implemented in the VASP code, which was the code used for the calculations. The relaxation was carried out in the 3x3x3 supercell, containing 109 atoms after hydrogen was introduced.

5.2.1 Relaxation of H in ZnO

Hydrogen prefers to bond with oxygen in ZnO. Therefore, the hydrogen was considered in four interstitial positions (Van de Walle, 2000, Limpijumnong and Zhang, 2005): $\mathrm{BC}_{\parallel}$, $\mathrm{BC}_{\perp}$, $\mathrm{AB}_{O,\parallel}$and $\mathrm{AB}_{O,\perp}$ where BC indicates a bonding site, and AB indicates an anti-bonding site. The subscripted parallel index, $\parallel$, indicates an interstitial H parallel to the Zn − O bond in the *c*-direction, and the normal index, $\perp$, refers to an interstitial H normal to the Zn − O bond. In addition, the subscripted O means that the hydrogen is related to oxygen as its nearest neighbor. The octahedral sites (C positions) were too large to be occupied by the small H_i, and the hydrogen preferred a closer bonding to the oxygen atoms. The four interstitial H positions and the relaxed interatomic O – H^+ distances are presented in figure 26. As the interstitial hydrogen acts as a donor, this section focused mainly on the positively charged H_i.

The relaxation of H^+ in ZnO was performed within both the LDA and the GGA, with an Γ-centered Monkhorst-Pack 2x2x2 k-point mesh. The energy cutoff for the plane-wave expansion was set to 400 eV. A quasi-Newton algorithm was used for the relaxation, and all the atoms relaxed with respect to the forces within an accuracy of 10 meV/Å. Some of the interstitial positions obtained even higher accuracy. The achieved accuracy was nevertheless acceptable for the purposes of the present study. These parameters were in compliance with the reference (Limpijumnong and Zhang, 2005) as one searched for a reproduction of the vibrational frequencies within the LDA.

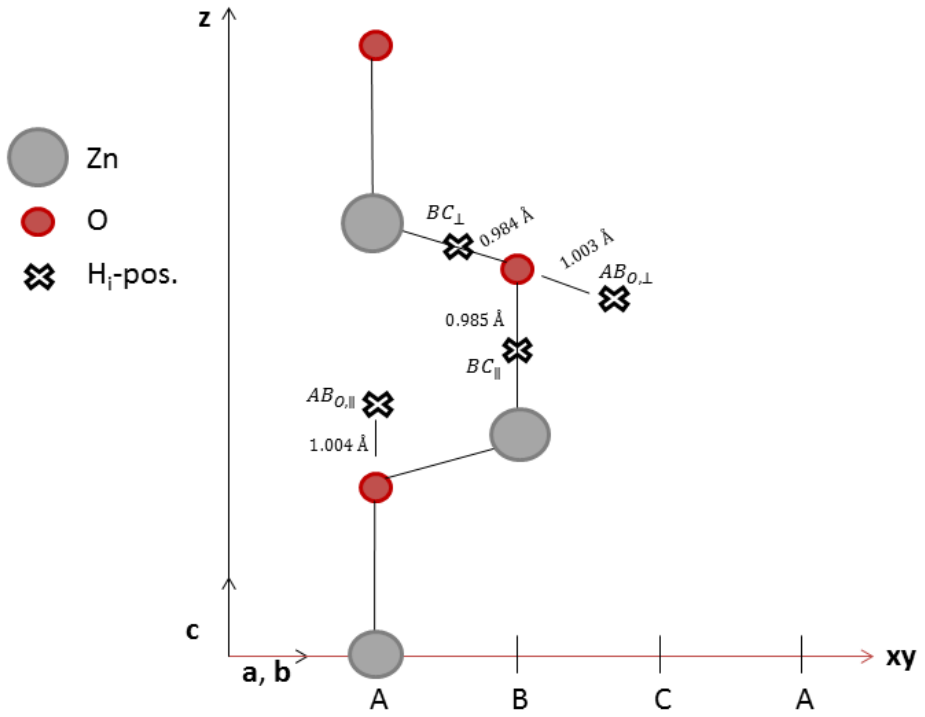

Figure 26: Schematic presentation of the four interstitial H sites in ZnO marked as non-filled crosses on the figure. The relaxed interatomic distances between H^+ and nearest O atom from the LDA calculations are also given. The large filled (grey) circles are Zn atoms, while the small filled (red) circles represent O atoms.

In table 3, the distances are summarized and related to the earlier published LDA calculations. The calculated $O-H^+$ distances for the GGA calculations were slightly smaller, ~1 %, than the calculated distances for the LDA. However, the trends were similar between the two approximations where the H_i^+ in $BC_\perp$ had the shortest $O-H^+$ distance. As the GGA relaxed supercell was larger, one should expect larger interatomic distances compared to the LDA, but the attraction between O and H^+ was larger for the GGA leading to shorter interatomic distances.

Table 3: The relaxed atomic distances, d_{O-H^+} (in Å), between the interstitial H^+ and the nearest O atom given in Å, for the different H_i positions in ZnO compared earlier LDA references. BC denotes H^+ positioned in a bond centered position, hence in the direction of a Zn-O bond, and AB denotes an anti-bonding site. The subscript shows whether the hydrogen was positioned parallel to *c*-axis or in a direction normal to the *c*-axis. The positions are defined in figure 26.

	d_{O-H^+} [Å]			
H_i^+ position	Present GGA	Present LDA	Ref.[1]	Ref.[2]
$BC_\parallel$	0.977	0.985	0.986	0.990
$BC_\perp$	0.976	0.984	0.982	---
$AB_{O,\parallel}$	0.991	1.004	1.003	1.006
$AB_{O,\perp}$	0.991	1.003	1.003	---

[1] Ref. (Limpijumnong and Zhang, 2005), [2] Ref. (Lavrov et al., 2002)

Table 4 shows the calculated relative formation energy with respect to $\mathrm{BC}_{\parallel}$ for the different interstitial H^+ positions in ZnO. Earlier LDA reports are included in the table as references. $\mathrm{BC}_{\parallel}$ was calculated as the most stable interstitial H^+ position in ZnO, in agreement with earlier results (Limpijumnong and Zhang, 2005, Lavrov et al., 2002). However, placing H^+ in $\mathrm{BC}_{\parallel}$ led to large local displacements of the host atoms, which broke the symmetry of the native ZnO. Since all the relaxed energies were within 0.2 eV, kinetic considerations might change the preferable position of H_i^+ in ZnO (Limpijumnong and Zhang, 2005). The calculated formation energies within the GGA are also presented, as it was supposed to give more accurate results compared to the experimental values. GGA showed no significant changes from the LDA and also concluded that $\mathrm{BC}_{\parallel}$ was the most stable position of the interstitial H^+ in ZnO.

The formation energy for an interstitial positively charged H is, in accordance with section 4.8.2., given by

$$\Delta H_f = E_{d,coh} + E_F - \mu_H - E_{0,coh} \quad .$$

Here, $E_{d,coh}$ is the total cohesive energy of the system after hydrogen is introduced and ionized, E_F is the chemical potential of electrons, μ_H is the chemical potential of hydrogen, and $E_{0,coh}$ is the total cohesive energy of the perfect structure before a hydrogen is introduced. As the relative formation energies is to be calculated, the three last terms will be equal for all interstitial positions, and the difference in formation energy will be given as the difference in the total cohesive energy. These energy differences are easily calculated. For example, the relative difference in formation energy between $\mathrm{BC}_{\perp}$ and $\mathrm{BC}_{\parallel}$ will therefore be given as

$$\begin{aligned}\Delta H_f^{rel} &= \Delta H_f(\mathrm{BC}_{\perp}) - \Delta H_f(\mathrm{BC}_{\parallel}) \\ &= E_{d,coh}(\mathrm{BC}_{\perp}) + E_F - \mu_H - E_{0,coh} - \left(E_{d,coh}(\mathrm{BC}_{\parallel}) + E_F - \mu_H - E_{0,coh}\right) \\ &= E_{d,coh}(\mathrm{BC}_{\perp}) - E_{d,coh}(\mathrm{BC}_{\parallel}) \quad .\end{aligned}$$

Table 4: The calculated relative formation energies, ΔH_f^{rel} (in eV), for an H_i^+ defect for the different H_i positions in ZnO with respect to $\mathrm{BC}_{\parallel}$. The calculations were performed within both the LDA and GGA and are compared with three LDA references. The interstitial positions of H are defined in figure 26.

	ΔH_f^{rel} (eV)				
H_i^+ position	Present GGA	Present LDA	Ref.[1]	Ref.[2]	Ref.[3]
$\mathrm{BC}_{\parallel}$	0.00	0.00	0.00	0.00	0.0
$\mathrm{BC}_{\perp}$	0.15	0.15	0.15	-0.02	0.1
$\mathrm{AB}_{O,\parallel}$	0.21	0.19	0.17	0.23	0.1
$\mathrm{AB}_{O,\perp}$	0.16	0.14	0.14	0.04	0.1

[1] Ref. (Limpijumnong and Zhang, 2005), [2] Ref. (Van de Walle, 2000) , [3] Ref. (Lavrov et al., 2002)

H acts exclusively as a donor in ZnO and is therefore easily ionized by donating its electron to the host structure. However, calculations of the relaxed structure with introduced H_i were performed for

neutral H as a reference. No significant differences were observed, and the neutral state was also most stable in the $BC_{\parallel}$ position. The relative energy differences with respect to $BC_{\parallel}$ were slightly smaller for the neutral case.

Even though the present work utilized the same given parameters as the reference (Limpijumnong and Zhang, 2005), the exact same energies and interatomic distances were not obtained. This is often a problem within numerical calculations. As one is not able to present all the employed calculation parameters in a paper, some approximations must be taken by the reproducer, leading to small deviations immediately. In addition, small deviations might have occurred from the ionic relaxations leading to deviations in the further calculations. Therefore, an exact reproduction of numerical calculations is difficult. However, the present study gave qualitatively the same results and the same trends as the earlier reports.

5.2.2 Vibration of H in ZnO

The local vibrational modes (LVM) of hydrogen has been carried out within the LDA in earlier publications (Limpijumnong and Zhang, 2005, Van de Walle, 2000, Lavrov et al., 2002), but in this work the vibrational modes were calculated within both the LDA and the GGA. The former was in order to reproduce the existing values and to get a feeling about the vibrational calculations, whereas the latter was to improve the frequencies within the GGA. In the vibrational calculations for the hydrogen, one frequency was significantly larger than the others, recognized as the stretch mode with the neighboring oxygen. By the fact that hydrogen is significantly lighter than the host atoms, the O-H frequency mode can be calculated by the expression $\nu = \frac{1}{2\pi c}\sqrt{\frac{\alpha}{m_H}}$, derivated in section 4.6.1.

The calculated vibrational frequencies of H^+ can be seen in table 5, and the frequencies calculated with the PBE potential had somewhat higher frequencies than the corresponding LDA values. Moreover, the GGA frequencies were closer to the measured values. The frequencies calculated by the reference (Limpijumnong and Zhang, 2005) included some systematic errors obtained by considering the stretch modes in the water molecule. As the present results also included some systematic errors, the reference values also included this error in order to compare the results more easily. The results of the LDA calculated frequencies were close to the references within ~4 %, which was an acceptable deviation. The deviations occurred from different relaxed structures and some systematic errors in the different calculations. However, small deviations in the vibrational frequencies were expected. As the interatomic distances obtained in the previous section deviated slightly from the references, small inaccuracies were introduced, which led to deviations in the calculations based on the relaxed ZnO structure.

Table 5: The calculated local vibrational frequencies (LVM), ν (in cm^{-1}), for H^+ in the different H_i positions in ZnO within both the LDA and GGA. Only the largest frequency mode (associated with $O-H^+$ stretch mode) is listed. The present values are compared with two LDA references and experimental values. The GGA frequencies were closer to the experimental value than the LDA frequencies. The positions of interstitial hydrogen are defined in figure 26.

	ν (cm^{-1})				
H_i^+ position	Present GGA	Present LDA	Ref.[1]	Ref.[2]	Measured
$BC_{\parallel}$	3522	3452	3377	3500	3611[2]
$BC_{\parallel}$	3538	3448	3421		
$AB_{O,\parallel}$	3286	3137	3018	3208	3326[3]
$AB_{O,\perp}$	3314	3166	3056		

[1] Ref. (Limpijumnong and Zhang, 2005), [2] Ref. (Lavrov et al., 2002), [3] Ref. (Van de Walle, 2000)

The calculated frequencies were also carried out for neutral hydrogen, as a reference. The resulting frequencies were slightly lower, ~1 %, for the neutral H, than the ionized H^+ atom for both the LDA and the GGA.

5.3 Lithium in ZnO

Lithium is a common impurity in ZnO and can be introduced by several means, both unintentionally and through doping. In the growth mechanism of hydrothermal ZnO, ZnO is carried out in an aqueous solvent containing LiOH and some of the solvent are therefore introduced unintentionally into ZnO. Therefore a study of Li in ZnO is of vital interest. However, the main reason why Li in ZnO has been studied the last years is due to the problem to make a stable *p*-type bulk ZnO. (Vines et al., 2010). There have been some reports that Li in ZnO is a candidate to make *p*-type ZnO (Schirmer, 1968, Wardle et al., 2005b, Meyer et al., 2004). While hydrogen acts exclusively as a donor in ZnO, Li has shown both donor- and acceptor-like behavior, dependent on the atomic configuration. The presence of interstitial Li will likely imply shallow donors, and donate the excessive electron to the host material (Wardle et al., 2005b). However, substitutional Li at a Zn site lacks one electron compared to the host Zn atom, and Li_{Zn} will therefore attract electrons resulting in an acceptor-like behavior. Li_{Zn} has proven to exhibit moderate *p*-type behavior, and the discussion whether it is a deep or a shallow acceptor is still investigated. Experiments show almost exclusively deep acceptor levels, but some theoretical calculations have indicated introduction of shallow levels through Li-doping (Park et al., 2002), providing a little hope that stable *p*-type doping of ZnO is achievable. One of the main problems to achieve *p*-type doping is from limiting effects due to self-compensation, for instance by interstitial H. In addition, Li is a small, mobile element which can easily occupy an interstitial site, compensating for the substitutional acceptors. ZnO is naturally an *n*-type semiconductor mainly through the presence of hydrogen donors, so both Li_i and H_i need to be suppressed in order to make stable *p*-type ZnO (Wardle et al., 2005b).

From chapter 4, the theory of vibrational modes states that $\nu \sim \sqrt{\frac{1}{m}}$. As Li has a similar valence state as H ($2s^1$ vs. $1s^1$), an effective-mass model to describe the Li atom from a modified H potential was considered. The atomic mass of hydrogen was set to 6.941 in terms of the atomic mass unit to equal the atomic mass of lithium. Several tests were performed, both for the ionic and electronic relaxation and also for the vibrational calculations. Due to the larger size of the Li atom, which was not taken into account in the H-like model, the displacements of the host atoms was more significant for the real Li potential than in the case of a heavy H potential during ionic relaxation. As a result, the compound relaxed in two different configurations, as depicted in figure 27. The total energy of the system calculated with a modified H potential was < 1 meV from the original H potential, which proved that the modification did not have any large effect upon relaxation. For the vibrational calculations, the differences were also significant. For the modified H potential, the calculated vibrational frequencies were the same as for hydrogen multiplied by $\sqrt{1.000/6.941}$, which were expected from theory. In the case of the true Li potential, the calculated vibrational modes were smaller, ~40 %, compared to the vibrational frequencies from the modified H potential. The Li atom was however placed somewhat further away from the host atoms from the relaxation compared to the modified H potential, which might explain the differences in the vibrational frequencies. Therefore, in order to check whether the theory was valid only for the vibrational calculations, an ionic relaxation was performed with the true Li potential. Thereafter, the local vibrational modes were obtained with the modified H potential and the true Li potential. In this case, the modified H potential did not give any reasonable frequencies. As the relaxed configuration was not a local minimum for the modified H potential, one frequency was imaginary and the other two frequencies were different from the real Li potential. As a summarized conclusion, the simplified description of using the H potential and adjust the atomic mass cannot be used to describe the Li atom, or any other H-like atoms. One of the reasons is expected to be that the size effects are not taken into account in the effective-mass model.

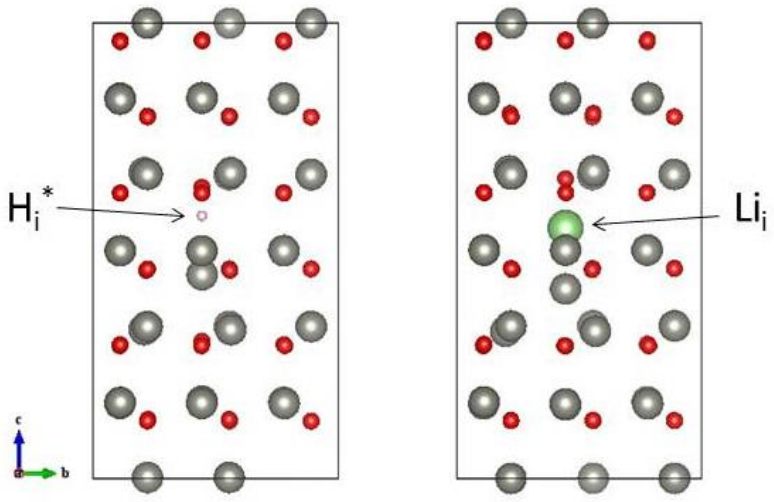

Figure 27: The ionic relaxation of the supercell containing one interstitial Li atom calculated with two different potentials. Left: Li described by a modified H potential. Right: Li described by a true Li potential. The two potentials led to different atomic distances due to the fact that the modified H potential has not taken the size of the Li atom into account.

5.3.1 Li_i in ZnO

From the discussion in the previous section, the self-compensation through the presence of H and Li, i.e. Li_i, Li_{Zn} and H_i is expected as the main reason why stable *p*-type doping is so difficult to obtain (Park et al., 2002). A clear knowledge about the mobility of Li_i and the possibility for trapping Li_i in ZnO is therefore of great interest in order to enhance *p*-type doping. In addition, interstitial Li is suggested to play an important role in establishing stable ferromagnetic ZnO (Awan et al., 2012).

As a continuation to the hydrogen calculations, a lithium atom was placed in the same four interstitial positions as hydrogen. While $\mathrm{BC}_{\parallel}$ was the position with the lowest energy for the hydrogen, this position was the least favorable in the case of lithium. The $\mathrm{BC}_{\perp}$ position had a somewhat lower energy, but both bond centered positions led to too large displacements in the structure, and were not energetically favored. The reason for the large distortions in the host structure is due to the size effects. As the covalent radius of Li (1.23 Å) is about three times larger than that of H (0.37 Å) (Spencer et al., 2012) p. 105, it is clear that the bond-centered positions are too short for the larger Li_i to occupy. Moreover, the interatomic Zn — O distance in ZnO was calculated ~2.0 Å. For the anti-bonding case, the $\mathrm{AB}_{O,\parallel}$ position was somewhat more energetically favored than the bond-centered positions as it is not as tightly bonded to the surrounding atoms. However, when relaxing the $\mathrm{AB}_{O,\perp}$ position the structure relaxed the Li_i into the octahedral position. This is also called the C position, referring to the atomic arrangement of different planes. The C positions are shown in figure 28. As seen from the figure, the $\mathrm{AB}_{O,\perp}$ position is somewhat closer to the available C position, explaining why the Li atom relaxed into the C position.

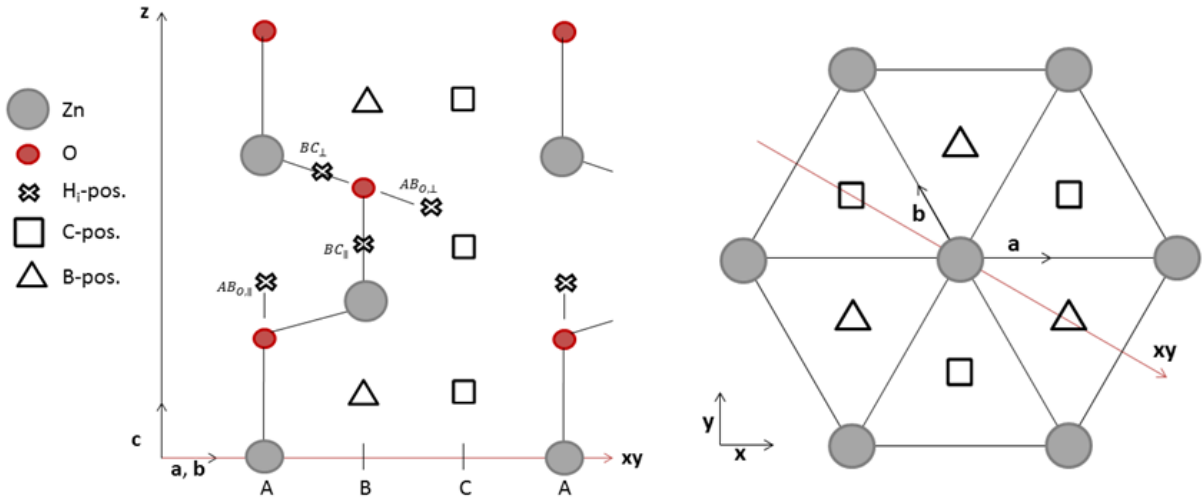

Figure 28: Schematic presentation of the possible Li_i positions in ZnO under consideration in this thesis depicted in two different planes. The large filled (grey) circles are Zn atoms, while the small filled (red) circles represent O atoms. The considered H_i positions (non-filled crosses) are included as a reference. From the wurtzite structure, the C plane is empty and involves sites available to be occupied by interstitial atoms marked as open squares. These positions are called octahedral positions (C positions). Tetrahedral sites (B positions) are also available to be occupied by Li_i, marked as open triangles. As described in the text, the bond-centered positions were highly unfavorable. The larger Li led to large displacements in the host structure, and hence a higher formation energy than for the case of H. The octahedral positions (C positions) were found to be the most energetically favored, 0.68 eV lower than the tetrahedral sites (B positions). The red xy-line relates the two figures. Note that the three different sites in the right figure do not lie within the same xy-plane, but are plotted like this for the sake of visualization.

It was therefore necessary to search for other positions when investigating interstitial Li. Due to the large displacements of the host structure, the initial positions from the H_i seemed unlikely to be occupied by interstitial Li. By analyzing the geometry of ZnO and the formation energy for Li_i at

different sites, two reasonable interstitial positions for Li_i were found. The most stable position was calculated as Li_i at the octahedral site, and the second most stable position was the tetrahedral site, also called the B-position as depicted in figure 28. The relative formation energy between the two sites was calculated to $\Delta H_f^{rel} = \Delta H_f^{Li_i^O} - \Delta H_f^{Li_i^T} = -0.68$ eV. The superscripts O and T denote the octahedral and tetrahedral site, respectively. The calculated value was in agreement with earlier results which achieved an energy difference at 0.62 eV between the two interstitial sites (Carvalho et al., 2009). The octahedral site was six-coordinated with the interstitial Li bonding to three neighboring oxygen atoms and three zinc atoms, while the tetrahedral site was four-coordinated with interstitial Li bonding to four neighboring oxygen atoms. The octahedral site led to small displacements of the host structure, whereas the Li atom at the tetrahedral site displaced the nearest zinc considerably from its equilibrium position, which is expected to be one of the main reasons for the energy difference.

As Li_i acts as a donor in ZnO, the ionic relaxation of Li_i^+ was calculated in the eight interstitial positions as presented in figure 29. As expected, the total energies were similar for Li_i^+ on the eight octahedral sites. The odd-numbered positions had exactly the same energy and similarly for the even-numbered positions as expected from the same atomic environment. The energies of the odd-numbered positions were slightly lower than the even-numbered by 20 meV. The energies were expected to be equal due to symmetry, but small deviations from the ionic relaxation of the host structure gave rise to small energy differences. The interatomic distances between Li and the nearest oxygen and zinc were calculated to be 1.93 Å and 2.42 Å, respectively. The covalent radii for Li (1.23 Å) and Zn (1.22 Å) are similar, and the covalent radius of O is 0.66 Å (Cordero et al., 2008). Due to charge effects, the Li_i^+ was expected to bond stronger to the oxygen atom, but also from size effects, the $Li_i - Zn$ distance was expected to be larger than the $Li_i - O$ distance. The interatomic distances were equal for the eight interstitial positions. The calculated distances were in agreement with earlier reports (Wardle et al., 2005b). As Li_i acts as a shallow donor it was relevant to treat the occurrence of Li_i^+ in the single-charge state. However, the calculations were performed for both the neutral and positively charged Li_i and any significant differences concerning the distances and energy differences were not registered.

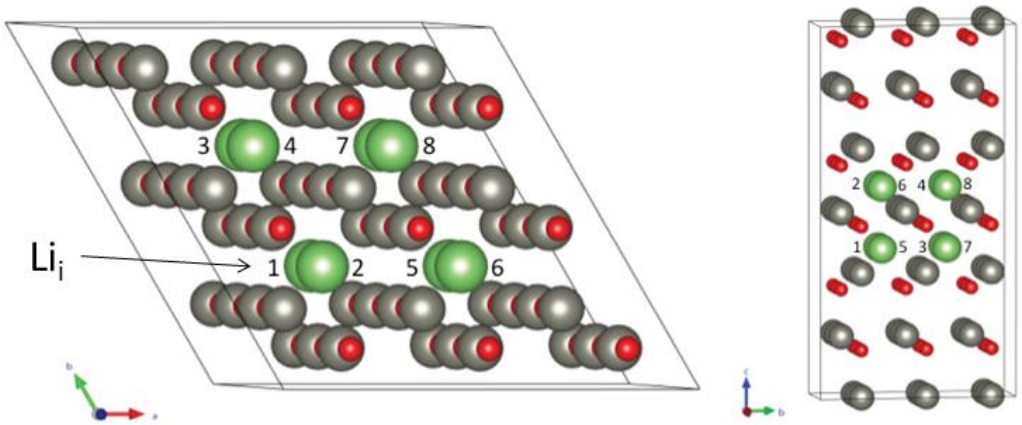

Figure 29: The octahedral positions (C positions) of interstitial Li in ZnO in the supercell considered in this thesis seen through the *c*-direction and the *a*-direction, respectively. The odd-numbered positions have the lowest *c*-value.

The vibrational frequencies of Li_i for the neutral and singly charged state were investigated for the eight different interstitial positions. The frequencies for Li_i in octahedral position 2 are presented in table 6. The other octahedral positions were similar by symmetry, and were expected to obtain the same frequencies. The calculated frequencies between the odd-numbered and the even-numbered positions were similar along the *c*-direction, but slightly differed in the *a*- and *b*-direction. The deviation is expected to be from the relaxation of the host structure. The calculated vibrational frequencies differed to some extent, but the trends were in good agreement with earlier calculations (Wardle et al., 2005b, Vidya et al., 2012) as can be seen from table 6. In the basal plane, the vibrational frequencies were associated with stretching of the $Li_i - O$ bond, and were primarily localized on the interstitial defect. The vibrational frequencies in the basal plane were degenerate in *a*- and *b*- direction because of the symmetric orientation around the vibrating Li_i. The vibrational frequency in the *c*-direction was significantly lower than in the basal plane, which indicates that the $Li_i - O$ bond is stronger in the basal plane. This is not surprising, as the *c*-channel is empty, leading to a weaker $Li_i - O$ bond in the axial direction. The vibrational modes did not show a significant dependence on the charge of Li_i, as the calculated vibrational frequencies for Li_i was equal to that of the charged Li_i^+. The calculated frequencies of the charged state were similar to the neutral state confirmed by Wardle. Although not so common in ZnO, negatively charged Li_i was also calculated as a reference. The calculation of the Li_i^- resulted in the same interatomic distances, and the same vibrational frequencies as the Li_i and Li_i^+. The effect of the isotopic mass on the vibrational frequencies will be discussed in the case of substitutional Li in section 5.3.2.1.

Table 6: The calculated local vibrational frequencies (LVM), ν (in cm^{-1}), for neutral and single-charged Li_i in octahedral position (O) 2 in ZnO compared with two references. The degenerate basal frequency, associated with a stretch mode of the $Li_i - O$ bond, was higher, as the *c*-channel is empty which led to a lower frequency in the axial direction. The vibrational frequencies were independent on the charge of Li_i. The present frequencies were somewhat higher than the references.

		ν (cm^{-1})		
Charge state, Li_i	**Vibrational mode**	Present GGA	Ref.[1]	Ref.[2]
Li_i (O)	A_1 (along *c* dir.)	399	390	323
	E (⊥ to *c* dir.)	498	423	413
Li_i^+ (O)	A_1 (along *c* dir.)	400	390	
	E (⊥ to *c* dir.)	498	425	

[1] Ref. (Wardle et al., 2005b), [2] Ref. (Vidya et al., 2012).

Regarding the experimental values, the Raman spectra of Li-doped ZnO thin films show peaks at 438 cm^{-1} and 579 cm^{-1} related to the vibrational modes of E_2 and E_1, respectively (Ni et al., 2001). The peak at 438 cm^{-1} can be assigned to the donor impurity, Li_i, which is expected to increase the conductivity. The peak at 438 cm^{-1} is about 12 % lower than the calculated E frequency mode in the present study, 498 cm^{-1}.

5.3.2 Li_{Zn} in ZnO

One of the challenges in ZnO-based technologies is to make stable bulk *p*-type ZnO. So far, it is only possible to make stable *n*-type ZnO. However, some reports (Park et al., 2002, Wardle et al., 2005b), claim that substituting Zn with Li can be a promising step in making *p*-type ZnO. Other substituted group-I elements are also found to be promising as acceptors in the material. Lithium has a high solubility in ZnO, so substitutional Li occupying Zn-sites are possible in the material. Experimental studies have been performed where up to 30 % of the Zn-sites are occupied by lithium (Onodera et al., 1996).

In the present work, substitutional Li was introduced in the host structure, making $Zn_{53/54}Li_{1/54}O$, resulting in a Li concentration of about 2 %. A relaxation of the substitutional Li_{Zn} in the neutral charge state resulted in small displacements of the $Li_{Zn} - O$ bonds compared to the ZnO host structure. This is because of the similar covalent radii between Li (1.23 Å) and Zn (1.22 Å) (Cordero et al., 2008). The tetrahedral coordinated $Li_{Zn} - O$ bonds were isotropic, with all the bond lengths calculated as 1.96 Å, in the single negatively charged case. For the neutral Li_{Zn} a small elongation was observed as the $Li_{Zn} - O$ bond lengths were calculated 1.96 Å and 1.97 Å for the axial and basal bonds respectively. These values were in agreement with several previous calculations (Park et al., 2002, Wardle et al., 2005b, Carvalho et al., 2009, Hu and Pan, 2008). However, in a recent paper, the geometric optimized structure of substitutional Li_{Zn} was calculated using a hybrid density functional (HDF) with a mixing parameter of $\alpha = 0.32$ (Carvalho et al., 2009). Their results were somewhat different from the achieved results both in this thesis and in other papers using PBE functionals. The relaxed Li_{Zn} was relaxed downward the z-axis and placed almost in plane with the three lower oxygen atoms, creating a stable off-center structure. This resulted in an increased axial $Li_{Zn} - O$ bond length by ~36 % and a localized unpaired electron level at the axial oxygen atom. In the present calculation the Li_{Zn} were relaxed in the center of the four surrounding O atoms making a delocalized unpaired electron pair distributed at the four O atoms. Compared with earlier experimental results, the present calculated bond lengths were not in agreement with the EPR observations done by Schirmer, reporting an off-centered structure with an elongation of the axial $Li_{Zn} - O$ bond length by ~40 % (Schirmer, 1968). As a conclusion, the use of different functionals can lead to different qualitative results. Therefore, it is always important to compare the calculated results with both earlier theoretical results and experimental results. Different functionals have different advantages, and the main task is to know when the given functional will give the right solution.

The local vibrational modes of Li_{Zn} were calculated and the results are summarized in table 7. The substitutional Li was bonded to four O atoms and the local vibrational frequencies were associated with the $Li_{Zn} - O$ bond strengths in each symmetry direction. Li_{Zn} was tetrahedral oriented with approximately the same distance to each O atom which was reflected by the similar frequencies in each direction. Table 7 shows that the vibrational frequencies were slightly higher than earlier reports. Agreed with the references (Vidya et al., 2012, Wardle et al., 2005b), the calculated frequencies were more independent of the direction than for the interstitial case indicating that the bond strengths between Li_{Zn} and the nearest O atoms were nearly isotropic. Unlike interstitial Li_i, one oxygen atom was now placed directly above the Li_{Zn} in *c*-direction whereas Li_i has an open channel in the *c*-direction which resulted in different frequencies for the two directions.

Table 7: The calculated local vibrational frequencies (LVM), ν (in cm^{-1}), for Li_{Zn} in ZnO in comparison with two references. The vibrational frequencies were degenerate in the basal plane, and slightly lower in the axial direction. The similar frequencies in each direction was a result of the tetrahedral oriented Li_{Zn}.

	ν (cm^{-1})		
Vibrational mode, Li_{Zn}	Present GGA	Ref.[1]	Ref.[2]
A_1 (along *c* dir.)	410	354	346
E (⊥ to *c* dir.)	421	368	339

[1] Ref. (Wardle et al., 2005b), [2] Ref. (Vidya et al., 2012).

5.3.2.1 $^6Li_{Zn}$ vs. $^7Li_{Zn}$ and different charge states of Li_{Zn}

The expression of the vibrational frequency of Li_{Zn} can be simplified into $\nu = \frac{1}{2\pi c}\sqrt{\frac{\alpha}{m_{Li}}}$. The expression implies that if the isotopic mass of Li decreases, the frequency should increase. In table 8 the calculated frequencies of the two most common Li-isotopes, ^6Li and ^7Li for the neutral and the negatively charged Li_{Zn} are presented. From the equation above, the frequency should be changed by a factor $\sqrt{7/6}$ due to the isotopic shift from ^7Li to ^6Li, which was in good agreement with the calculated frequencies listed in the table.

Table 8: The calculated local vibrational frequencies (LVM), ν (in cm^{-1}), for ^6Li and ^7Li in the neutral, single and double negatively charged state of Li_{Zn} in ZnO. The frequencies showed good agreement with the expression for the vibrational frequency by classical mechanics. For negatively charged Li_{Zn}, the vibrational frequencies increased.

	ν (cm^{-1})					
Vibrational mode, Li_{Zn}	Li_{Zn}^{0}		Li_{Zn}^{-}		Li_{Zn}^{2-}	
	^6Li	^7Li	^6Li	^7Li	^6Li	^7Li
A_1 (along *c* dir.)	443	410	455	421	457	423
E (⊥ to *c* dir.)	455	421	470	435	472	437

As can be seen from table 8, when the substitutional Li_{Zn} adds an extra electron, the vibrational frequency shifted to a higher frequency compared to the neutral Li_{Zn}, in agreement with earlier calculations (Wardle et al., 2005b). However, when another electron was introduced the frequency did not change. Li_{Zn} lacks one valence electron compared to the perfect host structure. When an electron is added to the cell it will be localized to the Li atom in order to compensate for this effect. This electron will then participate in and modify the vibrational frequencies. When another electron is added, there is no "need" for the defect to receive this electron, and the electron will therefore be smeared out throughout the crystal and will not affect the local vibrational frequencies of the substitutional Li_{Zn}.

5.3.3 Li-complexes in ZnO

Li-impurities in ZnO have the tendency to exhibit both donor- and acceptor-like behavior through Li_i and Li_{Zn}, respectively. This implies that Li can be self-compensating, meaning that both interstitial and substitutional Li occur in ZnO at the same time. Ref. (Awan et al., 2012) has shown that for relatively low (< 2 %) and for high (> 6 %) Li concentrations, the Li dopants mainly occupy the interstitial sites, whereas for medium doping concentrations (4 % to 6 %), substitutional Li is the main Li defect in ZnO. The presence of Li_{Zn} and Li_i can lead to the formation of a $Li_{Zn} - Li_i$ pair if the two impurities are located close to each other. Another pair that will be treated in this section is the presence of a $Li_{Zn} - Li_i$ pair. At first sight, this is expected be an unstable configuration as the Zn would prefer to be at a Zn site. However, this pair is important for the investigation of the kick-out mechanism for Li to obtain a Zn site.

5.3.3.1 $Li_{Zn} - Li_i$ pair

The $Li_{Zn} - Li_i$ pair was modeled with a substitutional Li positioned close to the center of the supercell, while the interstitial Li was placed in the eight octahedral positions (some positions were still equivalent) in the supercell. The interstitial positions were the same as without the substitutional defect, as seen in figure 30.

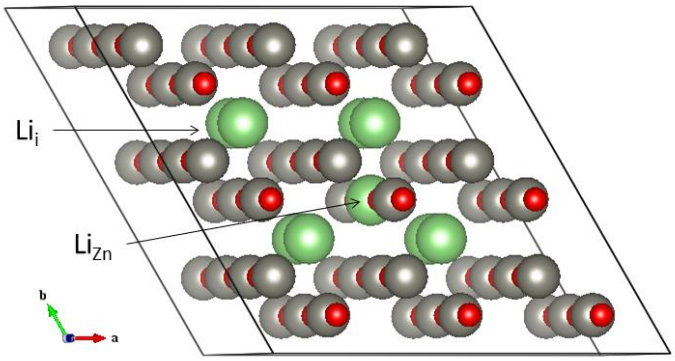

Figure 30: Supercell with the introduced $Li_{Zn} - Li_i$ pair for the eight considered interstitial octahedral Li_i positions presented. The Li_{Zn} was placed near the center of the supercell, whereas the Li_i was placed in the eight different interstitial positions. The $Li_{Zn} - Li_i$ pair with the Li_i positioned in octahedral position 2 (and position 6 and position 8 which was equal by symmetry) was the most stable as it obtained the shortest interatomic $Li_{Zn} - Li_i$ distance.

All the atoms were relaxed with a tolerance of the ions set to 0.02 eV/Å with respect to forces, which was an adequate accuracy for the calculations. After relaxation, for all the eight positions, the distance between the two lithium atoms decreased slightly relative to their initialized distance. The relaxed $Li_{Zn} - Li_i$ distances were 2.09 Å for interstitial position 2 (and by symmetry, also the equivalent positions 6 and 8), 2.29 Å for interstitial position 1 (and position 5 and 7), 3.95 Å for position 4, and 3.99 Å for position 3. For position 3 and 4, the Li_{Zn} did not move due to the small influence by the Li_i. Li_{Zn} lacked one electron compared to the perfect ZnO, and will therefore attract one electron from the Li_i. It was therefore reasonable that a configuration where the Li_i was positioned closer to the Li_{Zn} will be more stable. For the closer Li_i positions (1 and 2), the interstitial Li moved towards the substitutional Li in order to maintain the charge distribution of the host structure. As a result, the Li_{Zn} was slightly displaced away from the Li_i and the symmetry point. When placing a Li_i in an octahedral position, the charge center of the two Li atoms, i.e. Li_i and Li_{Zn}, will lie somewhere in between and therefore located significantly away from the symmetry point compared to the perfect structure. The small displacement of Li_{Zn} in the opposite direction is to limit this effect by

moving the center of charge towards the original Zn position, which was the center of charge for the host structure. This effect, together with the interatomic distances for the $Li_{Zn} - Li_i$ pair with Li_i positioned in octahedral position 2 can be seen in figure 31. As expected, the energy was lower for the interstitial Li-positions closer to the Li_{Zn} than for octahedral position 3 and 4 where the Li_i were placed far from Li_{Zn}. This is suggested to be due to conservation of charges from the host structure and from the attraction between the two Li atoms. Position 2 (and 6 and 8) was energetically the most stable, 0.12 eV lower than position 1 (and 5 and 7), 0.34 eV lower than position 3, and 0.41 eV lower than position 4. Li_i located on position 4 was thus the least favorable configuration. The reason why the $Li_{Zn} - Li_{i,2}$ can relax in a shorter distance and a more stable configuration is because there is no other atoms or atomic layers between the two Li atoms. In the case of a $Li_{Zn} - Li_{i,1}$ pair, an O layer is placed between the two Li atoms leading to a small repulsion and hence a larger interatomic distance.

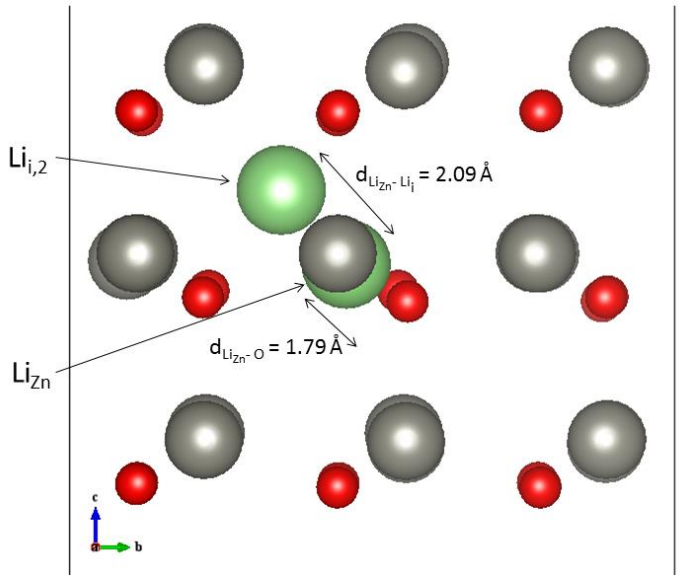

Figure 31: The relaxed structure of the $Li_{Zn} - Li_i$ pair with the interstitial Li located in octahedral position 2. The relaxed distances are also presented. As the Li_{Zn} relaxed towards Li_i due to charge conservation, the interstitial Li was slightly displaced leading to a shortening of the $Li_{Zn} - O$ bond.

It was interesting to compare the total energy of the system including the $Li_{Zn} - Li_i$ pair with two single-defect systems involving solely substitutional and interstitial Li, respectively. The formation energy for the $Li_{Zn} - Li_i$ pair can be calculated by

$$\Delta H_f^{Li_{Zn}-Li_i\,pair} = \left(E_{d,coh} + \mu_{Zn}N_{Zn} + \mu_{Li}N_{Li}\right) - E_{0,coh} = E_{d,coh}^{Li_{Zn}-Li_i\,pair} + \mu_{Zn} - 2\mu_{Li} - E_{0,coh} \ ,$$

and the introduction of single-defects for substitutional and interstitial Li can be calculated by

$$\Delta H_f^{Li_{Zn}} = \left(E_{d,coh} + \mu_{Zn}N_{Zn} + \mu_{Li}N_{Li}\right) - E_{0,coh} = E_{d,coh}^{Li_{Zn}} + \mu_{Zn} - \mu_{Li} - E_{0,coh} \ ,$$

$$\Delta H_f^{Li_i} = \left(E_{d,coh} + \mu_{Zn}N_{Zn} + \mu_{Li}N_{Li}\right) - E_{0,coh} = E_{d,coh}^{Li_i} - \mu_{Li} - E_{0,coh} \ ,$$

respectively. Combining these relations one can calculate the relative formation energy of a $Li_{Zn} - Li_i$ pair compared to the formation of the two defects separately, also referred to as the dissociation energy.

$$\Delta H_f^{rel} = \Delta H_f^{Li_{Zn}-Li_i\,pair} - \left(\Delta H_f^{Li_{Zn}} + \Delta H_f^{Li_i}\right) = E_{d,coh}^{Li_{Zn}-Li_i\,pair} + E_{0,coh} - \left(E_{d,coh}^{Li_{Zn}} + E_{d,coh}^{Li_i}\right)$$

In the present study, $\Delta H_f^{rel} = -1.62$ eV, indicating that the presence of a self-compensating defect pair was highly probable. Hereby, it was reasonable to suggest that the formation of a $Li_{Zn} - Li_i$ pair was more stable than to separate the two defects which was confirmed by the relaxation calculation earlier in this section.

The local vibrational modes were calculated for the $Li_{Zn} - Li_i$ pair with Li_i placed in the eight different octahedral positions. There were three prominent modes associated with this complex, listed in table 9. The highest mode was associated with the stretching mode between the two defects, Li_{Zn} and Li_i. This mode was calculated to be ν = 682 cm^{-1}, in accordance with previous reports (Wardle et al., 2005b). The stretch mode was situated primarily on the substitutional Li atom. From the formation energy, the formation of a $Li_{Zn} - Li_i$ pair was calculated to be highly favorable. Therefore, the $Li_{Zn} - Li_i$ bond was expected to be stable and strong, which was confirmed by the high frequency relative to the native $Zn - O$ stretch mode and the $Li_i - O$ stretch mode calculated in section 5.3.1. Another mode was associated with the stretching of the Li_{Zn} and the nearest O atom. The large frequency, ν = 657 cm^{-1}, was a result of the shorter $Li_i - O$ bond length compared with the single-defect configuration. As Li_i was slightly displaced in the opposite direction of Li_{Zn} during relaxation, the $Li_{Zn} - O$ was shortened. As a result the interatomic distance between Li_{Zn} and the nearest O was 1.79 Å. The covalent radii of Li and O are 1.23 Å and 0.66 Å (Cordero et al., 2008), respectively, indicating a strong interaction in the bond and a high frequency mode.The last mode was associated with the stretching between the interstitial Li atom and the nearest oxygen atom calculated to ν = 565 cm^{-1}. The O atom present in this mode was not the same as the O atom in the $Li_{Zn} - O$ stretch mode. The frequency was somewhat higher than the local vibrational modes of Li_i as a single-defect in ZnO calculated in section 5.3.1. Li_{Zn} hereby modified the $Li_i - O$ stretch mode to a slightly higher frequency. The $Li_{Zn} - O$ stretch mode was shared between the Li_i and the O, while the mode was more localized at the Li_i atom when it was not assisted by the Li_{Zn}. This present result may indicate that the $Li_{Zn} - Li_i$ pair can be seen as a distorted Li_2O unit cell inside the ZnO is probable. As the interatomic bonds are found to be small and strong, it is relevant to treat the $Li_{Zn} - Li_i$ pair together with a nearby O atom as a distorted Li_2O unit cell and compare with values of the Li_2O unit cell. This will be further discussed in section 6.1.3.

Table 9: The calculated local vibrational frequencies (LVM), ν (in cm^{-1}) associated with the $Li_{Zn} - Li_i$ pair in ZnO in comparison with a reference. The interstitial Li_i was placed in octahedral position 2 as indicated in figure 29. The high frequencies indicated a strong and stable $Li_{Zn} - Li_i$ pair and supported the formation of this pair. The O atoms in the stretch modes are not the same O atom, as they are the stretch modes between the defect and the nearest O atom.

	ν (cm^{-1})	
Vibrational mode	Present GGA	Ref.[1]
Li_{Zn} – Li_i stretch	682	718
Li_{Zn} – O stretch	657	
Li_i – O stretch	565	611

[1] Ref. (Wardle et al., 2005b).

5.3.3.2 $Li_{Zn} - Zn_i$ pair

In the discussion concerning diffusion of substitutional Li as a single defect, the Li_{Zn} can either diffuse to Zn vacancy sites, or through the kick-out mechanism of the host structure.

At first the structure with the $Li_{Zn} - Zn_i$ was relaxed in the same way as the $Li_{Zn} - Li_i$ pair, and the same initial positions were used as in the previous section. The relaxed distances were 2.23 Å for interstitial position 2 (and the positions 6 and 8), 2.45 Å for interstitial position 1 (and the positions 5 and 7), 3.86 Å for position 4, and 4.10 Å for position 3. The interatomic distances of the $Li_{Zn} - Zn_i$ were slightly longer than for the $Li_{Zn} - Li_i$ pair. Li_{Zn} lacked one electron and will attract one electron from the interstitial Zn atom. As Zn_i relaxed towards Li_{Zn}, the Li_{Zn} defect moved slightly off-center to try to imitate the same charge distribution as in the host structure as discussed in the previous section. However, this decreased the $Li_{Zn} - O$ bond leading to a stronger interaction between the Li_{Zn} and the nearest O atom.

The energy was lower when the Zn_i was placed closer to the substitutional Li compared to a longer $Li_{Zn} - Zn_i$ distance. This was because the charge distribution was easier fulfilled by this configuration. Position 2 was the most energetically stable, 0.37 eV more stable than position 1, 0.67 eV more stable than position 3, and 0.78 eV more stable than position 4. The calculated total energy for the system including the $Li_{Zn} - Zn_i$ pair with Zn_i in position 2 was 1.11 eV higher than the total energy of the system including a single Li_i defect in ZnO, which indicated that an interstitial Zn was not preferable.

However, there were some problems related to the relaxation of a closely bonded $Li_{Zn} - Zn_i$ pair. If Zn_i was placed in position 2 (and equivalent positions), the force relaxation was not met by the Li atom which experienced a drift of 0.12 eV/Å. This can have led to unreliable results for the $Li_{Zn} - Zn_i$ pair. Although the energy was lowest in position 2 due to the conservation of host charge, the position was quite unstable. In the other positions of Zn_i, the force relaxation was adequate with drifts below 0.01 eV/Å.

The local vibrational frequencies of the Li_{Zn} with an introduced Zn_i were calculated. The Zn_i located in the interstitial position 2 were considered, as the other interstitial Zn positions were less stable. The vibrational calculations indicated a weak bonding between the $Li_{Zn} - Zn_i$ as the stretch mode was calculated to $\nu = 179\ cm^{-1}$. The calculated frequency was primarily situated on the Zn atom which indicated that the Zn atom was not in a stable configuration. However, in identifying the local vibrational modes, the appearance of a high frequency mode was pointed out. From the evaluation of forces acting on the Li_{Zn}, the mode was identified as the $Li_{Zn} - O$ stretch mode in the basal direction with a frequency of $\nu = 739\ cm^{-1}$. The interatomic distance of the $Li_{Zn} - O$ was relaxed to 1.73 Å, indicating a tightly bonded $Li_{Zn} - O$. This mode was similar to the $Li_{Zn} - O$ stretch mode obtained in the investigation of the vibrational modes of the $Li_{Zn} - Li_i$ pair.

5.3.4 Diffusion of Li_i in ZnO

As a way of understanding Li in ZnO, different diffusion mechanisms of Li have been investigated. The diffusion barriers have been calculated for the different diffusion mechanisms, i.e. interstitial diffusion and kick-out mechanisms. In the case of interstitial diffusion, the transition rates and the corresponding random diffusion coefficients were calculated.

The calculations in this section were performed within the nudged elastic band (NEB) method as described in section 4.7.4. Restrictions concerning which directions the atoms were allowed to move were specified by the spring constant. The potential energy surface profile and the minimum energy path (MEP) were found by performing calculations with different images (intermediate steps) and spring constants. The number of images in the calculations varied from 2 to 9 images, and two spring constants were used, k = -5 and k = 0. The minus sign indicated nudging. The MEP was found to be somewhat independent on the spring constant and independent on the number of images. However, if too few images (< 3) were used, deviations occurred as the resolution was too low to generate a good estimate of the activation energies and the MEP. The initial paths were created by linear interpolations between the endpoints. The linear interpolations were done manually. The ionic relaxation for each image was performed with a quasi-Newton algorithm. This implemented algorithm is a good method when the initial positions are good guesses, as one expects that the minimum energy path was not too far from the initialized interpolated path.

The first diffusion mechanism considered was the movement of interstitial Li to other interstitial positions. The global minimum for a Li_i was found to be in the octahedral positions, and as a result, the diffusion processes between the octahedral positions were considered. The initial diffusion position was set to the octahedral position 1, as defined earlier, and two different diffusion paths were considered; parallel to the axial plane, *c*-axis, which was defined as position 2, and in the basal plane, which was considered in the *b*-direction, defined as position 3. The octahedral position 5 was also calculated as a reference to check that diffusion profile in *a*-direction was the same as in *b*-direction which by symmetry was expected to yield the same profile. In addition, the diffusion paths were carried out for single ionized interstitial Li atoms. Axial diffusion is described by

$$\mathrm{Li}_{\mathrm{i},1}^{\mathrm{O}} \rightarrow \mathrm{Li}_{\mathrm{i},2}^{\mathrm{O}} \ ,$$

and basal diffusion is described by

$$\mathrm{Li}_{\mathrm{i},1}^{\mathrm{O}} \rightarrow \mathrm{Li}_{\mathrm{i},3}^{\mathrm{O}} \ .$$

Figure 32 gives a schematic presentation of the two considered diffusion mechanisms for interstitial Li, and figure 33 presents the actual calculated MEPs from the present study. Figure 33 depicts that the MEP for axial diffusion follows a linear path between the two octahedral positions. The displacements of the host atoms during the diffusion processes are not shown. In the basal plane, the diffusion path of Li_i involved an intermediate state, the tetrahedral position.

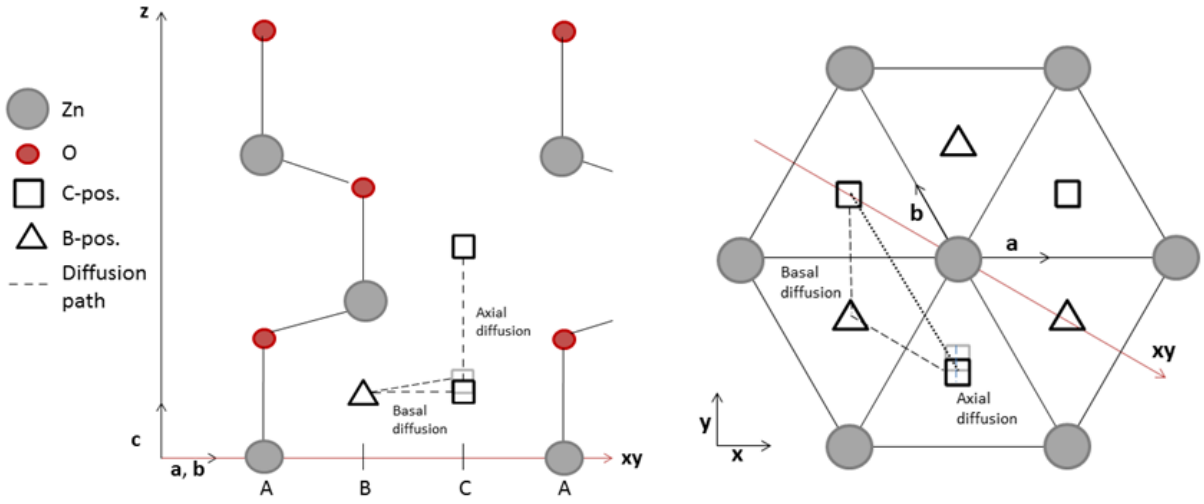

Figure 32: Schematic illustration of the diffusion paths of Li_i in ZnO, axial and basal. The large filled (grey) circles are Zn atoms, whereas the small filled (red) circles represent O atoms. Axial diffusion aligns parallel to the *c*-direction and basal diffusion occur normal to the *c*-direction. The basal diffusion was found to pass through an intermediate state, the B position. In the left figure the blurry square shows the end position for basal diffusion and the small displacement is due to visualization, as the end points have the same *c*-value. In the right figure, the dotted straight line shows the initial path for the basal diffusion and the dashed line shows the MEP obtained for the basal diffusion involving the intermediate position. The squares and the triangles are not in the same basal plane as the Zn-atoms, but are slightly lifted from the paper. However, they are plotted in the same plane and the oxygen atoms are not shown for simplicity. Here, the blurry square indicates the endpoint for the axial diffusion. The small displacement is for the sake of visualization, and the axial diffusion path is out of plane and not visible.

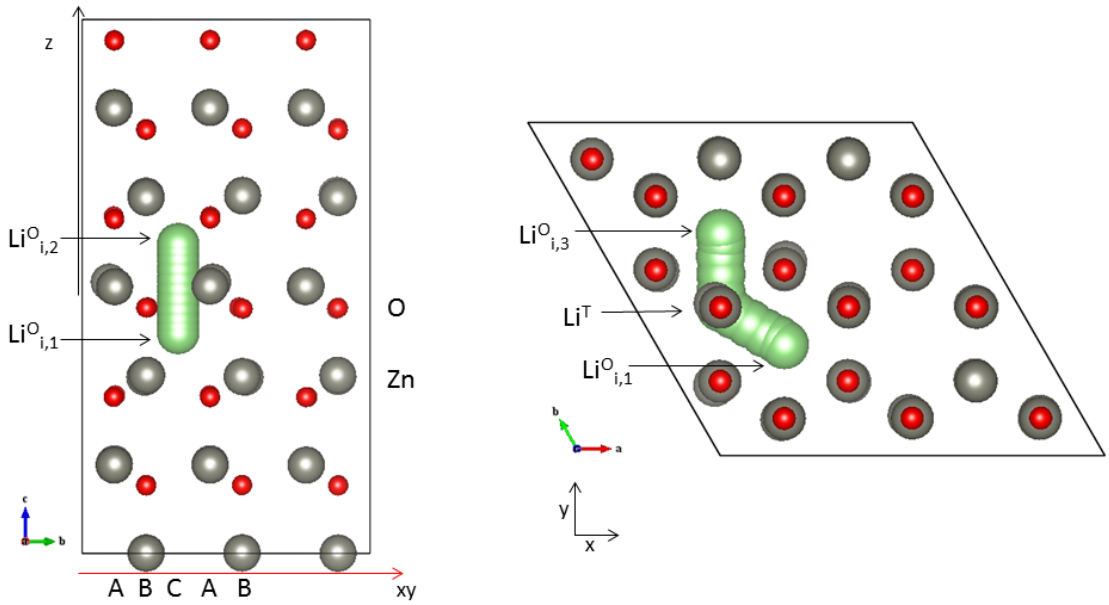

Figure 33: Left: The calculated MEP for axial diffusion (position 1 to position 2) for Li_i in ZnO depicted through the *a*-direction. The calculation used the NEB method with 9 images, and the displacements of the host atoms during the diffusion process are not shown. The MEP is aligned in the *c*-direction. Right: The calculated MEP for Li_i in ZnO in the basal plane along the *b*-direction (position 1 to position 3) depicted through the *c*-direction. As seen on the figure, the path was assisted by a relaxation into the intermediate tetrahedral position. The calculation used the NEB method with 9 images, and the displacements of the host atoms during the diffusion process are not showed.

In figure 34, the potential energy surface along the diffusion paths are shown as a function of the Li_i position. The filled symbols are the calculated image energies, and the solid line is plotted as an eye

guideline. Even though the calculations were performed with different spring constants and different number of images, the calculated potential energy surfaces showed impressive consistency as all the image energies lay on the calculated MEP. The calculated energy barriers were calculated to be $\Delta E_{a,\parallel} = 0.67$ eV and $\Delta E_{a,\perp} = 0.72$ eV for axial and basal diffusion, respectively, for the Li_i. The calculated barriers are somewhat larger, but in good accordance with earlier reports (Wardle et al., 2005b, Huang et al., 2009). The diffusion barriers of Li_i for three different charge states for the two different paths can be seen in table 10. The diffusion calculation from position 1 to position 5, which gave the diffusion in the *a*-direction, resulted in the same activation energy as in the *b*-direction which was predicted as a result of the equivalency between the two points. The diffusion mechanism to position 5 will not be further considered in this section, but calculations have been done in order to check the consistency with diffusion in the *a*-direction.

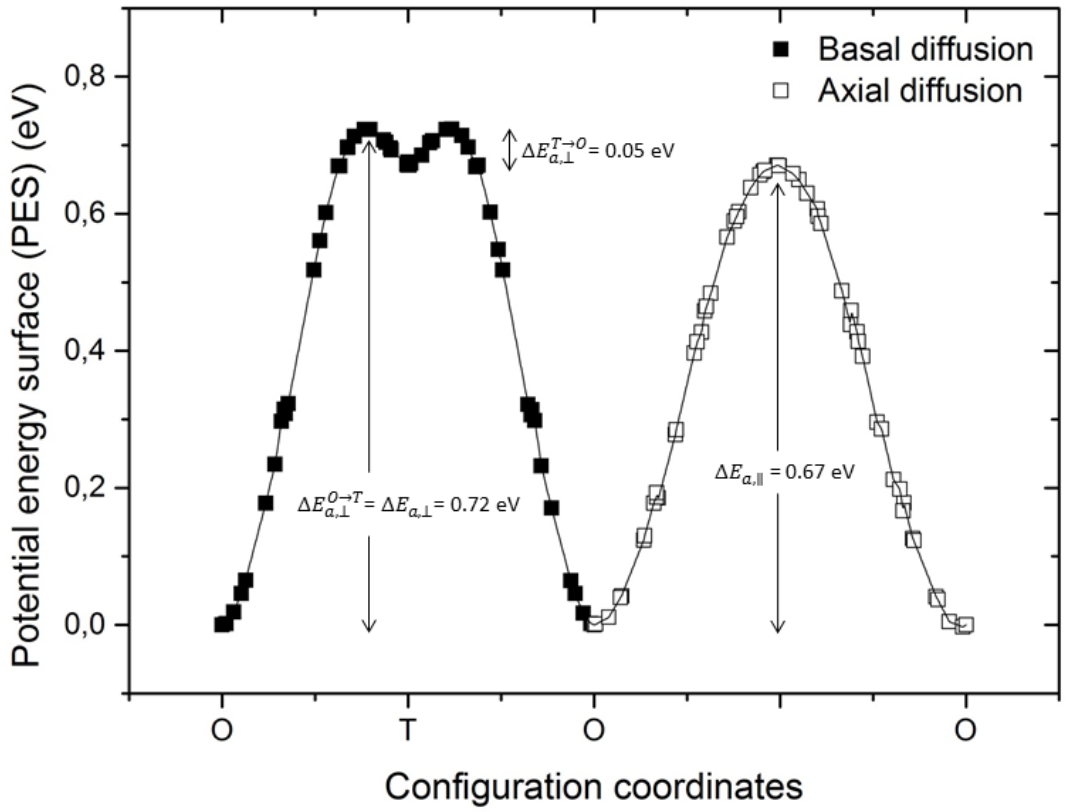

Figure 34: The diffusion profile for interstitial Li in ZnO between octahedral positions (O) in the basal and axial direction, calculated by the nudged elastic band (NEB) method. The points are associated with the calculated values obtained in the calculations. Filled squares show diffusion in the *b*-direction, whereas open squares show diffusion in the *c*-direction. The solid line is an interpolation to visualize the MEP. The diffusion barrier was calculated to be 0.72 eV in the basal plane and 0.67 eV in the axial direction. The basal diffusion included a metastable tetrahedral position (T) at an energy 0.67 eV higher than the octahedral position, and the diffusion profile was symmetric about the tetrahedral position. As the activation barrier from T→O was significant lower than the activation barrier from O→T, the barrier from O→T will be the rate-determining step as described in the text, and will decide the diffusion in the basal plane.

Table 10: The calculated diffusion barriers, ΔE_a (in eV), for the axial, ∥, and the basal, ⊥, diffusion for interstitial Li in ZnO, for three different charge states of Li. In the basal plane the diffusion was a two-step process. As the energy barrier between the octahedral position and the tetrahedral position was significantly larger, it was considered as the rate-determining step and determined the overall barrier to a large extent. The diffusion barrier was found to be somewhat independent on the charge state of Li_i, with a slightly lower barrier for the Li_i^- in the axial direction.

	ΔE_a (eV)	
Charge state, Li_i	Present – GGA	
	$\Delta E_{a,\parallel}$	$\Delta E_{a,\perp}$
Li_i	0.67	0.72
Li_i^+	0.68	0.73
Li_i^-	0.65	0.72

From the diffusion barriers and the vibrational frequencies one can evaluate the transition rates. Within the transition state theory, the transition rate for the axial diffusion can be written as

$$k_{\parallel} = \frac{\nu_1 \times \nu_2 \times \nu_3}{\nu_1^{\dagger} \times \nu_2^{\dagger}} \exp\left(-\frac{\Delta E_{a,\parallel}}{k_B T}\right) \ ,$$

where ν_i are the vibrational frequencies associated with the Li_i at the local energy minimum, and $\nu_i^{\dagger}$ are the vibrational frequencies associated with at Li_i at the transition state between the two local minima. In the case of basal diffusion from position 1 to position 3, the overall diffusion included an intermediate state between the barriers, resulting in a two-step diffusion process; from the octahedral to the tetrahedral position, and from the tetrahedral to the octahedral position. In multi-barrier transition state theory the overall transition rate can be evaluated from the single-barrier transition rates, k_i, from state i to i+1 according to the relation

$$k = \frac{1}{\frac{1}{k_0} + \frac{1}{k_1} + \cdots + \frac{1}{k_{n-1}}} = \left(\sum_0^{n-1} \frac{1}{k_i}\right)^{-1} \ .$$

For the basal diffusion, the overall transition rate are evaluated by the transition rates from the octahedral to the tetrahedral position, $k_{1\to T}$, and from the tetrahedral to the octahedral position, $k_{T\to 3}$. In the latter transition of the diffusion process the relaxed initial energy of the tetrahedral position was 0.62 eV higher than for the octahedral site resulting in a significantly smaller energy barrier (0.05 eV) compared to the first transition. Therefore, the latter transition will only give a small contribution to the overall transition rate. The overall transition rate for the basal diffusion from position 1 to position 3 can therefore be written as

$$k_{\perp} = \left(\frac{1}{k_{1\to T}} + \frac{1}{k_{T\to 3}}\right)^{-1} \ ,$$

where the single-barrier transition rates are calculated as normal. The transition rates, $k_{\parallel}$ and $k_{\perp}$, are plotted as a function of the inverse temperature in figure 35. The transition rates are found to be

highly dependent on the temperature. From the tables inside figure 35, Li_i is predicted to be quite mobile at room temperature. The transition rates are slightly anisotropic as a result from the different diffusion barriers. The anisotropic behavior decreases at higher temperature. A few calculations of the transition rate at different temperature were carried out in order to check if the $k_{T\to 3}$ term can be neglected. At room temperature, the transition rate was the same including the $k_{T\to 3}$ term or not. Even at a temperature of 40 K the $k_{T\to 3}$ term did not have any impact on the overall transition rate. In addition, the transition rate in this range was of order 10^{-80} s^{-1}, indicating that the $k_{T\to 3}$ term can be totally neglected from the overall transition rate, $k_{\perp}$. This is confirmed by the theory and that the diffusion from $1 \to T$ will be the rate-determining step. Moreover, this implies that the intermediate state is not occupied at any reasonable temperatures as the mobile Li will pass directly towards the octahedral site.

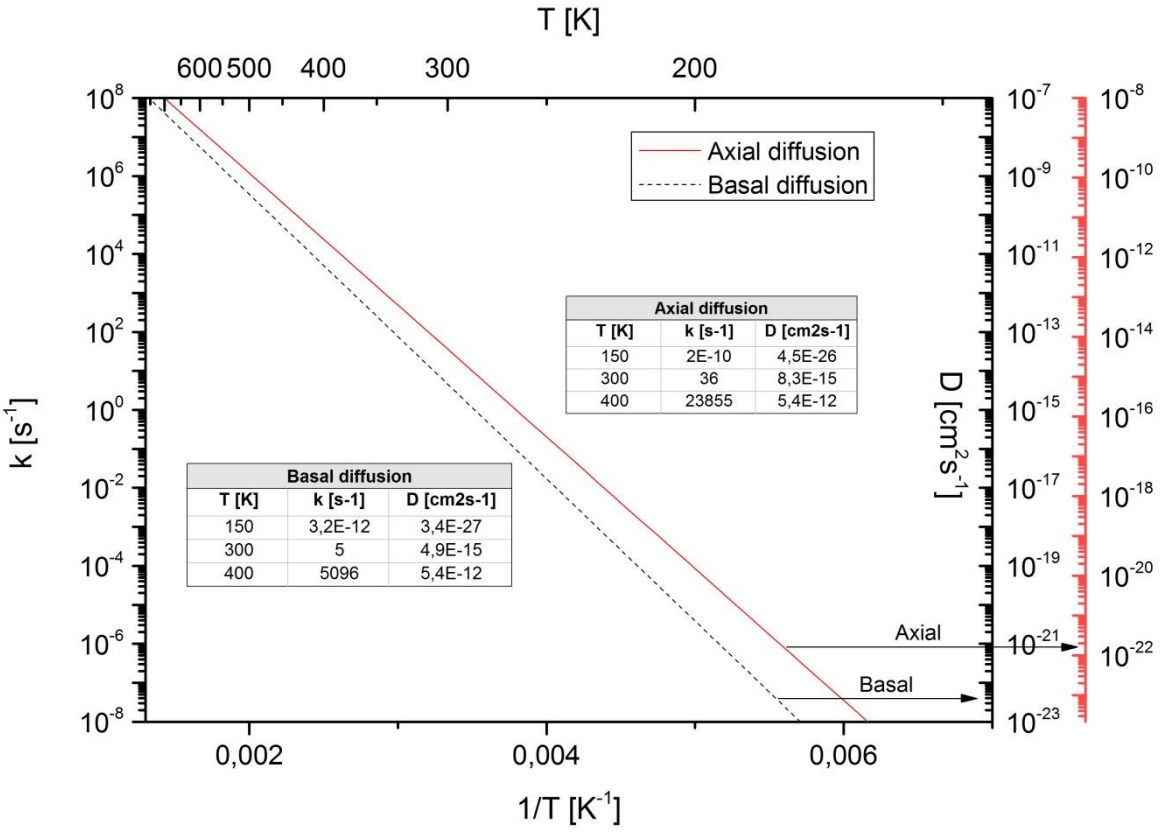

Figure 35: The calculated transition rates, $k_{\parallel}$ and $k_{\perp}$, and diffusion coefficients, $D_{\parallel}$ and $D_{\perp}$, as a function of temperature for axial and basal diffusion of Li_i in ZnO. The solid (red) line is related to the axial diffusion, and the dashed (black) line is related to basal diffusion. Some calculated values at appropriate temperatures are presented in the tables in the figure for the two diffusion paths. The relation between the transition rate and the diffusion coefficient is dependent on the direction so the left (black) D axis is related to the basal diffusion coefficient and the right (red) D axis is related to the axial diffusion coefficient.

In figure 35, the diffusion coefficients are plotted as a function of the inverse temperature together with the transition rates. By derivation of Fick's law one can obtain an expression for the diffusion constant. Assumed that the diffusion is random, the diffusion coefficient, as derived in section 4.7.2, can be calculated from

$$D = \frac{1}{6} s^2 Z k \quad .$$

Here, N_d is set to unity for interstitial diffusion, s is the jump distance, Z is the number of available sites around the diffusing atom and k is the transition rate calculated earlier in this section.

The expression for the diffusion coefficient can be used to evaluate the diffusion coefficients for the axial and basal diffusion, respectively. k is known from the calculations above, and the diffusion distances were measured to s = 2.61 Å and s = 3.26 Å for axial and basal diffusion, respectively. For axial diffusion $Z = 2$ as the diffusion can proceed in the positive or negative *c*-direction. For basal diffusion Z is set to 6 as the octahedral positions were surrounded by six octahedral positions equally spaced. The axial positions were also available in a shorter distance, but will not be considered in the diffusion coefficient for the basal diffusion. The transition rates can therefore be multiplied with a constant to obtain the diffusion coefficient, and the diffusion coefficients for the two diffusion paths can be read from figure 35. As the transition rates differ between the two directions, the diffusion coefficient will be dependent on the direction. For higher temperatures, the transition rates, and hence the diffusion coefficients, approach each other. Calculations considering diffusion coefficients are highly uncertain as a small change in the diffusion barrier will change the diffusion coefficient by orders of magnitudes. In addition, the factor that relates the transition rates and the diffusion coefficients must be evaluated after assumptions. However, this factor does not change the diffusion coefficients considerably as long as the jump distance is chosen within the right order of magnitude. As for the case of the transition rates, the diffusion barrier needs to be calculated as precise as possible in order to obtain a reasonable diffusion coefficient. Calculations of the transition rates and diffusion coefficients of interstitial Li in ZnO are not well described in the literature, but the diffusion coefficients are measured in a few experimental studies. As the diffusion barrier of Li in ZnO in the experimental studies $\geq$ 1 eV, the calculated diffusion coefficients are higher than the measured values (Lander, 1960, Knutsen et al., 2013). These results are discussed somewhat further in section 6.1.4.

5.3.5 Kick-out mechanisms

The possibility to diffuse Li-atoms through the kick-out mechanism in the host structure has been considered. Primarily, substitutional Li can diffuse through the presence of kick-out mechanisms. In this thesis, the diffusion paths for the two Li-complexes considered earlier, namely the $Li_{Zn} - Li_i$ pair and the $Li_{Zn} - Zn_i$ pair, have been calculated.

5.3.5.1 Kick-out mechanism of $Li_{Zn} - Li_i$ in ZnO

The initial position of the interstitial Li was set to position 1 as indicated in figure 29, and the kick-out mechanism consisted of moving the Li_i to the substitutional position by kicking out the Li_{Zn} to the interstitial position 6. The diffusion process is described as

$$\mathrm{Li_{i,1}} + \mathrm{Li_{Zn}} \rightarrow \mathrm{Li_{Zn}} + \mathrm{Li_{i,6}} \quad .$$

The MEP and the kick-out barrier were calculated from the nudged elastic band method as for the single-particle diffusion, with two different spring constants, $k =$ -5 and $k = 0$, and the number of images varying from 2 to 5. The initial images were obtained by a linear interpolation.

An interesting result of this kick-out mechanism was the presence of two different MEPs, depending on the initial images. If the images were initialized by a linear interpolation, where the two Li atoms displaced by the same distance for each image, a “linear-like” MEP was obtained with an energy barrier of ~0.37 eV. However, when the images were initialized from jumps with only one atom

jumping at a time, the kick-out followed a different path with a lower kick-out barrier, ~0.22 eV. This barrier was obtained from a step-wise initial path with three images and from two different spring constants. The two different calculated MEPs can be seen in figure 36, and the potential energy surface as a function of configuration coordinates of the kick-out process can be seen in figure 37. The low-barrier path followed a longer kick-out path than the linear initial images path. For the low-barrier path, the total distance increased by 19 %, and the distance for the interstitial Li increased by 31 %. The low barrier was also achieved when the calculations had two images (lower resolution), and when the calculations were performed with a conjugate-gradient algorithm, for some of the linear initialized kick-out paths. The conjugate-gradient algorithm is preferable for systems with difficult relaxation problems.

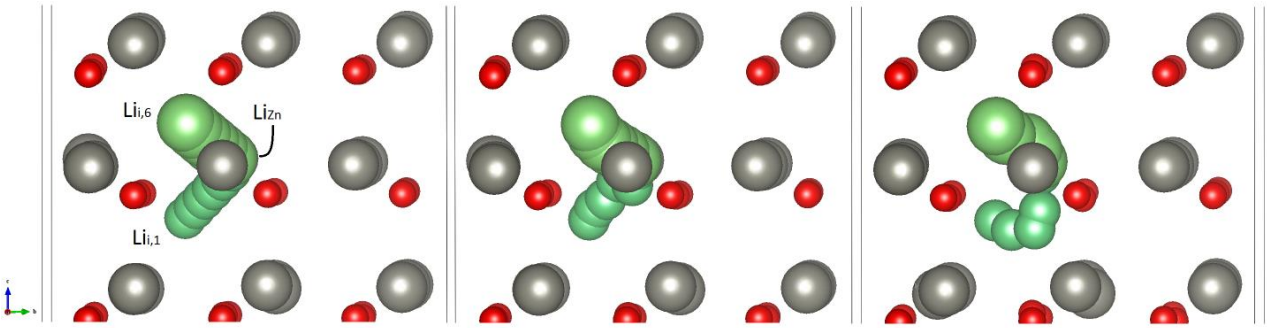

Figure 36: The calculated MEP of the kick-out mechanism of a $Li_{Zn} - Li_i$ pair, seen through the *a*-direction. Li atoms are shown in green. The smaller green balls denotes the atom moving from $Li_{i,1}$ to Li_{Zn}, and the larger green balls denotes the atom moving from Li_{Zn} to $Li_{i,6}$. Left: The linear initial kick-out path. Middle: The calculated MEP from linear initial images moving each Li atom the same distance for each image. The MEP was calculated with 4 images. Right: The calculated MEP from the linear path, but moving one atom at a time in a step-wise manner. The MEP was calculated with 3 images. The calculated MEP showed inconsistency related to the initially chosen kick-out path, leading to two different MEPs with two different kick-out barriers.

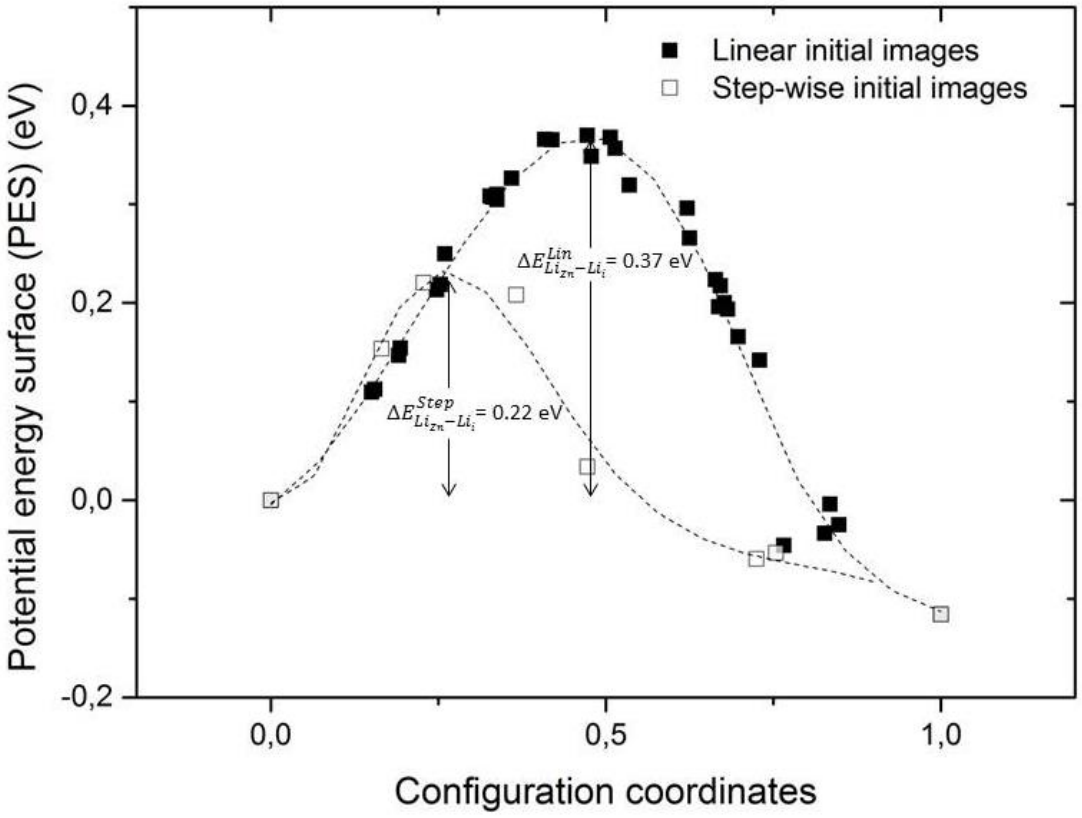

Figure 37: The kick-out mechanism profile for the $Li_{Zn} - Li_i$ pair in ZnO for the kick-out process $\mathrm{Li_{i,1}} + \mathrm{Li_{Zn}} \rightarrow \mathrm{Li_{Zn}} + \mathrm{Li_{i,6}}$ calculated by the NEB method. The squares are associated with the calculated values obtained in the calculations. Filled squares shows kick-out mechanism based on linear interpolation, while open squares shows kick-out mechanism based on the step-wise initial images where one atom jump at a time. The dashed lines are the interpolated guidelines related to the calculated values. The estimated kick-out barrier was 0.37eV, but the step-wise initial images seemed to yield a lower kick-out barrier.

After investigation of the kick-out mechanism from position 1 to position 6, the same kick-out mechanism was calculated from position 1 to position 8, in order to avoid any possible interatomic forces along the kick-out path. In this case, the calculated MEP followed an identical potential energy surface shape as the MEP obtained from the linear initial path from position 1 to position 6. This indicates that no significant repulsive forces are present in the intermediate steps. In addition, the kick-out mechanism was carried out from a step-wise initial path of 3 and 7 images for the kick-out process from position 1 to position 8. In both cases, the calculated MEP was the same path for the interstitial Li as obtained from the earlier step-wise initial path. Summarized, the calculated MEPs from a kick-out mechanism from position 1 to position 8 were nearly identical relative to the kick-out mechanism to position 6.

A suggested explanation for the occurrence of two different MEPs is that the first jump from the Li_i towards the Li_{Zn} leads to a large repulsive force between the Li-pair. The interatomic distance between the two Li atoms was 1.32 Å before ionic relaxation. As a result, the ionic relaxation may move the Li_i significantly to avoid the strong repulsive force, and the obtained configuration may occasionally fall into another potential minimum in the MEP relaxation. The relaxed configuration can therefore reach the global MEP, whereas the path obtained from the linear initialized images may reach a local MEP. In this respect, another step-wise initialized path with three images was calculated where the first jump was by the Li_{Zn} towards the endpoint, position 8. The next image was

initialized by a jump of the initial Li_i towards the empty Zn host site. By this jump mechanism one avoided the large repulsive force between the Li-pair in the images. This calculated MEP yielded a kick-out barrier of ~0.37 eV, confirming the statement that the large initialized repulsive force between the Li_{Zn} – Li_i pair is responsible for the two calculated MEPs. However, the present result is remarkable as the MEP is supposed to give an unambiguously path between two endpoints. This is further discussed in section 6.1.6.

However, in either case, the calculated energy barrier for the $Li_{Zn} - Li_i$ kick-out mechanism $\Delta E_{a,Li_{Zn}-Li_i} \leq 0.37$ eV is significant lower than for the single-defect diffusion of interstitial Li calculated in section 4.3.4. This indicates that the diffusion of Li_i can be trapped within the presence of Li_{Zn}. As the Li_i acts as a donor, and the Li_{Zn} acts as an acceptor in ZnO, there is an attractive force between the Li atoms from charge considerations. The present results predict that the Li_i prefers to kick out the Li_{Zn} rather than diffuse directly to the next octahedral position. From the calculated barriers, one can expect that the kick-out mechanism between Li_{Zn} and Li_i occurs to large extent at room temperature if the Li concentration is high and if the defects are present within the same region.

5.3.5.2 Kick-out mechanism of Li − Zn in ZnO

An important study is whether the interstitial Li can kick out a Zn host and locate a substitutional Zn site. The diffusion barrier and the MEP have been investigated for the kick-out mechanism of

$$\mathrm{Li_{i,1}} + \mathrm{Zn_{Zn}} \rightarrow \mathrm{Li_{Zn}} + \mathrm{Zn_{i,6}}$$

This is the same diffusion path as in the previous section, but now Li_{Zn} is substituted by a host Zn atom. The calculated potential energy surface for the kick-out mechanism of a $Li_{Zn} - Zn_i$ pair is shown in figure 38. The calculated MEP indicated a highly unstable endpoint with an increasing potential energy surface profile during the kick-out process. The results were as expected since the relaxed positions of the $Li_{Zn} - Zn_i$ pair seemed not stable, according to the calculations in section 4.3.3.2. To obtain a MEP from a NEB calculation under these circumstances are difficult. As NEB requires that both endpoints are local minima, the images tend to slightly move away from the endpoint due to its transition-state-like behavior as a result of too weak spring constants. From initial configurations where only one minimum is known, other methods, like the dimer method, is suggested in order to find saddle points on the MEP (Henkelman and Jonsson, 1999). From the calculated MEP profile in figure 38, a meta-stable configuration is indicated with energy 0.73 eV higher than the initial configuration. This configuration is depicted to the right in figure 38 and yielded a center of mass and charge close to the host Zn site which can explain the presence of a possible meta-stable configuration. However, higher resolution is required to investigate whether the meta-stable configuration is an achievable configuration. The presence of a highly unstable endpoint indicates that the Zn_i either reverse the kick-out process or diffuse further to kick out other host atoms and obtain a more stable position in the structure. The latter is studied by ref. (Huang et al., 2009). If the material exhibits a large Li-concentration, the kicked-out Zn_i can kick out another Li_{Zn} to obtain the stable Zn site, and hereby increase the mobility of Li_{Zn}.

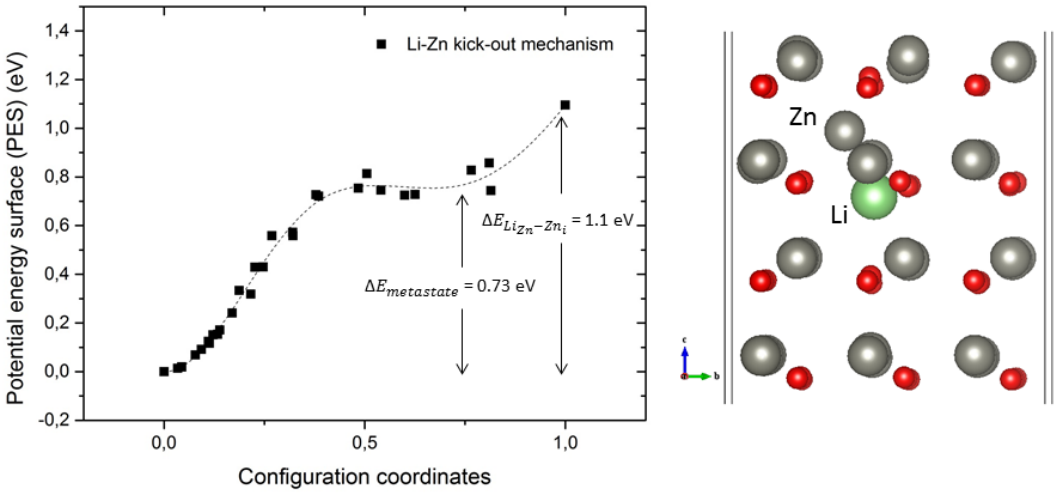

Figure 38: The kick-out mechanism profile for the Li – Zn pair in ZnO for the kick-out process $Li_{i,1} + Zn_{Zn} \rightarrow Li_{Zn} + Zn_{i,6}$ calculated by the NEB method. The filled squares are associated with the calculated values obtained in the calculations and the dashed line is an interpolated MEP based on the calculated values. The $Li_{Zn} - Zn_i$ pair was unstable, with an energy difference of 1.1 eV compared to an interstitial Li in the host structure. The calculated MEP indicated that the kick-out mechanism may include a meta-stable configuration ~0.73 eV higher than the initial position. To the right the meta-stable configuration is shown with a center of mass and charge between the two moving atoms, close to the host position of a Zn atom.

Related to the kick-out mechanism for the $Li_{Zn} - Li_i$ pair, the kick-out path calculated for the Li – Zn pair indicated a path similar to the step-wise initialized path from the $Li_{Zn} - Li_i$ pair, shown in figure 36. For the Li – Zn pair, the kick-out mechanism was only calculated by a linear movement on both atoms for each image, hence not a step-wise initial path. The present results support the presence of the initially step-wise calculated MEP from the $Li_{Zn} - Li_i$ kick-out mechanism as the global MEP.

Even though the activation barrier to kick out a Zn atom by a Li atom from its host position is not significant larger than other barriers calculated in this thesis, the relaxed configuration of $Li_{Zn} - Zn_i$ is so unstable that the kick-out mechanism will reverse easily as can depicted from figure 38. A $Li_{Zn} - Zn_i$ configuration is therefore highly unfavorable, and the kick-out process is expected to reverse easily to obtain the initial configuration.

5.4 Nickel in ZnO

An additional project in this thesis was the investigation of vibrational frequencies of Ni_{Zn} and Ni_{Zn} co-doped with Li, in ZnO. In addition, the magnetization of the system as a function of free-hole doping and Li_{Zn} doping was investigated. Nickel is a divalent element which can have several oxidation states. Compounds containing Ni^0, Ni^+, Ni^{2+}, Ni^{3+} and Ni^{4+} are known, but the most common is the divalent ion, namely Ni^{2+}. Like lithium, nickel is a well-known impurity in ZnO and happens to be unintentionally introduced in the structure during the growth process of ZnO (Lavrov et al., 2013). Ni_{Zn} is found to exhibit *n*-type conductivity in ZnO and is not a possible acceptor dopant for making *p*-type ZnO (Wardle et al., 2005a).

5.4.1 Ni_{Zn} in ZnO

Introduction of a substitutional Ni in the structure led to the chemical formula of $Zn_{53/54}Ni_{1/54}O$ and a Ni concentration at about 2 %. The relaxation was performed with a Γ-centered 4x4x4 k-point mesh,

and the energy cutoff was set to 400 eV. The relaxation was performed with respect to forces and all atoms were relaxed within a total force of 0.005 eV/Å. In the relaxation, the introduction of substitutional Ni_{Zn} led to only modest displacements of the host structure. As for lithium, the covalent radius of both Zn and Ni are similar, with a slightly longer covalent radius for Ni (1.24 Å) (Cordero et al., 2008) so one expects a similar Ni_{Zn} – O bond length compared to the previous Zn – O bond lengths. The distances for the Ni_{Zn} – O bonds were slightly shorter than the previous calculations, indicating a stronger bonding between the Ni_{Zn} and the surrounding oxygen atoms. The relaxed Ni_{Zn} — Zn distanceswere the same as for the host structure and the Ni_{Zn} was clearly tetrahedral located. An overview of the relaxed distances between the substituted Zn-site and the surrounding oxygen atoms for the substitutional defects considered in this thesis is presented in table 11.

Table 11: The interatomic distances between the substitutional defect and the nearest O-atoms compared to the host structure. The X – O distances calculated in this thesis were similar as a result of the similar covalent radii between the substitutional atoms.

	Bond length {$d_{X\text{-}O}$ (Å)}			
Direction	Zn_{Zn}	Li_{Zn}^{-}	Li_{Zn}	Ni_{Zn}
along *c* dir.	1.99	1.96	1.96	1.95
⊥ to *c* dir.	1.98	1.96	1.97	1.96

In the investigation of the local vibrational modes of Ni_{Zn}, the major vibrational mode was associated with the Ni_{Zn} – O stretch mode. The stretch mode was calculated by allowing both the Ni atom and the nearest O to move, as the mass difference is not significant and the vibrational mode is expected to be located at both atoms. The calculated frequency were ν = 458 cm^{-1} for the Ni_{Zn} – O stretch mode in accordance with experimental reports (Malaeru et al., 2012). The calculated frequency was 3.9 % larger than the calculated Zn – O stretch mode. The result is as expected; the lower mass of Ni (58.7 u) compared to Zn (65.3 u) and the shorter bond length of Ni_{Zn} – O indicates a strong interatomic bond which in turn yields a higher frequency. The vibrational mode was primarily located at the O atom.

5.4.2 Ni-doped ZnO co-doped with Li_{Zn}

Li_{Zn} has been reported to enhance the ferromagnetism of Ni-doped ZnO. Li_{Zn} act as an acceptor and thereby decrease the number of valence electrons of the system (Jayakumar et al., 2010). In this respect, the presence of Li_{Zn} has been investigated, both to find the atomic configuration and to see to what extent the Li_{Zn} modifies the Ni_{Zn}-O vibrational stretch mode. The original 3x3x3 supercell consists of 108 atoms, leading to the chemical formula $Zn_{53/54\text{-}X}Ni_{1/54}Li_XO$ after introduction of X Li_{Zn}. One substitutional Li will therefore result in ~2 % Li concentration. The relaxations in this section were performed by a 3x3x3 k-point grid and cutoff energies at 400eV.

The total energy was calculated to consider whether the substitutional Li will occupy a substitutional Zn site close to, or far from the Ni_{Zn}. After relaxation, Li_{Zn} was found to be most stable close the Ni_{Zn}-defect. The total energy was slightly lower, ~0.3 eV, than Li_{Zn} positioned far from Ni_{Zn}. The Li_{Zn} attracts an electron above the Fermi level from the ZnO host material. As Ni introduces spin-

polarized energy levels high above the Fermi level, Li_{Zn} as an acceptor is stable closer to atoms with high-energetic electrons. The presence of Li_{Zn} close to the Ni_{Zn} is also suggested to compensate the Ni_{Zn} – O bond (Persson et al., 2011). To investigate the configuration of a higher ratio of substitutional Li doping in the Ni-doped ZnO, another Li_{Zn} was introduced into the supercell. Based on the previous result, the first Li_{Zn} was held next to Ni, and the second Li atom was placed on another substitutional Zn site. After relaxation it was found that whether the second Li is placed next to, or further away from the Ni-defect is somewhat indifferent, with the configuration of two Li_{Zn} close to the Ni-defect slightly lower in energy, < 0.1 eV as obtained by earlier reports (Jayakumar et al., 2010).

Table 12 presents the calculated vibrational frequencies of the Ni_{Zn} – O stretch mode after the introduction of Li_{Zn} related to the doping concentration and the position of the Li_{Zn} relative to Ni_{Zn}. The Ni_{Zn} – O stretch mode was found to increase with the presence of Li_{Zn} in the structure. However, for increased fraction of Li_{Zn} in the material, the Ni_{Zn} – O stretch mode was slightly decreased. As Li_{Zn} steals one electron from the high-energetic levels of Ni_{Zn}, the Ni_{Zn} atom misses one electron and hereby bonds more strongly to the neighboring O atoms in order to conserve the charges. As a result, the vibrational frequency increases for the Ni_{Zn} – O stretch mode by introduction of the Li_{Zn}. For larger concentrations of Li_{Zn}, the effect is smaller and the frequencies are only slightly modified.

Table 12: The calculated vibrational frequencies, ν (cm^{-1}), of the Ni_{Zn} – O stretch mode as a function of Li_{Zn} concentration and location. For introduced Li_{Zn}, the Ni_{Zn} – O stretch frequency was modified to a higher frequency.

	ν (cm^{-1})				
Vibrational mode	0 % Li_{Zn}	2 % Li_{Zn}		4 % Li_{Zn}	
		Near	Random	Near	Random
Ni_{Zn} – O stretch	458	490	501	477	485

The magnetic moment in the supercell was calculated in terms of Bohr magneton, μ_B, as a function of doping concentration of Li_{Zn} around the substitutional Ni_{Zn}. The magnetic moment was calculated by Persson et. al, and the present calculation was to confirm and support the previous studies (Persson et al., 2011, Jayakumar et al., 2010). As mentioned earlier, Li_{Zn} will act as an acceptor and introduce hole dopants to the material. In addition, calculations were performed by removing electrons (adding holes) from the supercell, to compare the introduction of Li_{Zn} with free-hole doping directly into the $Zn_{53/54}Ni_{1/54}O$ in terms of the magnetic moment. The results can be seen in figure 39. For low Li-doping, Li_{Zn} can be described as a free-hole dopant, in accordance with the previous calculations. The effect is due to the introduced energy levels above the Fermi level from the magnetic Ni. The high-energetic energy levels from above the Fermi level are depopulated by holes, which are introduced by doping. This lowers the Fermi level and increases the magnetic moment for the compound at low hole concentrations. At higher acceptor concentrations, the ZnO host states are depopulated by holes indicating a Fermi level below the valence band maximum. This results in magnetically induced oxygen atoms, which stabilizes and decreases the magnetic moment. The induced oxygen states thus exhibit a paramagnetic phase rather than a ferromagnetic phase (Persson et al., 2011).

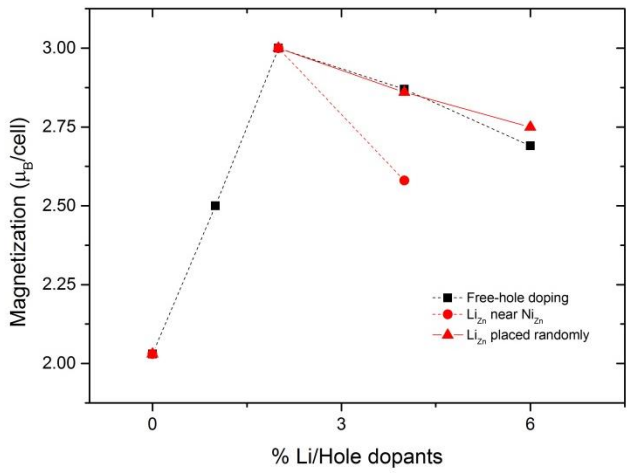

Figure 39: Total magnetization in terms of μ_B as a function of free-hole- and Li_{Zn}-doping in $Zn_{53/54}Ni_{1/54}O$. One Li_{Zn} will therefore give a doping concentration at ~2 %. The filled squares (black) show free hole doping. From relaxation the first Li_{Zn} preferred to be located near the Ni_{Zn}. The circles (red) shows the magnetization when Li_{Zn} was placed close to the Ni_{Zn} giving a significant decrease in total magnetization for higher Li-concentration. The triangles (red) shows the magnetization when the first Li was placed near the Ni-defect and for higher concentration of Li the Li was placed randomly (far away from Ni) in the supercell. In this case the Li_{Zn} acted as free-hole dopants to a larger Li concentration. These results are in agreement with earlier references (Jayakumar et al., 2010).

The magnetic moment was dependent on the position of the substitutional Li relative to the substitutional Ni. For a single-atom doping of Li_{Zn}, the position was irrelevant, but for higher doping ratio, the magnetic moment was related to the positions of Li_{Zn}. For ~4 % doping, if the two Li_{Zn} were placed close to substitutional Ni-defect, the decrease in magnetic moment was significant, down to 2.58 μ_B as shown in figure 39. However, if the second Li_{Zn} were situated far away from the Ni-atom, the magnetic moment slightly decreased to 2.86 μ_B similar to the case of free-hole doping. Therefore, for randomly distributed Li_{Zn} can be seen as free-hole dopants to a larger doping concentration, in line with earlier reports (Jayakumar et al., 2010).

6. Conclusions

In this chapter the discussions from the previous chapter are collected and seen in terms of the big picture. The major results will be further discussed to get an overview of the outcome of this thesis. Thereafter, there will be a short summary of the present work along with some thoughts and suggestions for further work to conclude this chapter.

6.1 Main results and final discussions

6.1.1 GGA vs. LDA

In this thesis, both the GGA and the LDA have been utilized in order to calculate the relaxed configurations and the vibrational frequencies of host ZnO and interstitial H in ZnO.

The GGA is an improved and more accurate approximation than the LDA. However, for some properties, the GGA might result in less accurate values compared to the experimental values. In the present study, the calculated lattice constants and interatomic distances for ZnO improved relative to the experimental values within the GGA. As the LDA tends to underestimate the lattice constants, the GGA is known to overestimate the lattice constants, however, by a smaller value. In this work, the lattice constants were underestimated by ~2 % within the LDA, and the GGA lattice constants were overestimated by ~0.5 %. In the study of the local vibrational modes of H in ZnO, the GGA improved the LDA frequencies relative to the experimental measured values. The GGA resulted in an error bar of < 3 %, compared to measured values. As a result, the present study indicates that calculations of the geometric relaxation of lattice constants and ionic configurations and the vibrational frequencies can be improved by the GGA from the LDA.

6.1.2 Effective-mass model

The effective-mass model, considered in section 4.3, to describe group I-elements by a modified H-like potential did not succeed. In the effective-mass model, an H-potential is modified by the atomic mass of the considered element. The idea of the model is relevant as the group-I elements have similar charge and valence states, s^1. From the present study, an interstitial Li atom was placed in a bond-centered position. Furthermore, the relaxed configurations and the vibrational frequencies were calculated with both a modified H potential, and a true Li potential. The results indicated a significant difference between the two potentials, both for the relaxed ionic positions and interatomic distances and for the vibrational frequencies. The presence of a Li atom in a bond-centered position led to large displacements of the neighboring host atoms, which was not obtained from the H-like potential. As the size of the Li atom is not taken into account in the H-like potential, one expects that the size effect is the main reason why the effective-mass model fails, and cannot be used to describe any H-like elements.

6.1.3 Vibrational frequencies of Li and Li-complexes in ZnO

The vibrational frequencies of Li_i in ZnO were calculated to be anisotropic with ~20 % larger frequency in the basal direction. As Li_i acts as a donor in ZnO, the Li_i is expected to be bonded closer to the O atoms, and the vibrational frequencies can be associated with the stretching of the Li_i – O bond. With Li_i positioned in an octahedral position (C position), no other atoms are located in the

axial direction above or below the Li atom. As a result, the interstitial Li are weakly bonded in the axial direction and hereby achieve lower vibrational frequencies. The calculated frequencies were $\nu = 399$ cm^{-1} and $\nu = 498$ cm^{-1} for the axial and basal direction of Li_i, respectively. The vibrational frequencies in the basal direction were degenerated as the Li_i was located with the same distance to the surrounding O atoms in the basal plane. For the substitutional Li, the calculated vibrational frequencies were nearly isotropic. From the ionic relaxation of Li_{Zn}, the interatomic $Li_{Zn} - O$ distances were calculated to be similar. As the local vibrational modes of Li_{Zn} are associated with the $Li_{Zn} - O$ stretch mode one expects Li_{Zn} to be bonded by equal strength to the surrounding oxygen atoms. The small difference between the axial and basal frequency mode is expected to appear from small deviations in the tetrahedral $Li_{Zn} - O$ bond lengths from the ionic relaxations. The substitutional Li frequencies were calculated to lie between the interstitial Li modes. The vibrational frequencies indicate that Li_{Zn} is more strongly bonded in the axial direction, than Li_i, and the opposite for the basal direction.

The calculated frequencies of Li_{Zn} were calculated for the isotope $^6Li_{Zn}$, and for negatively charged Li_{Zn}. The vibrational frequencies after the isotopic shift were proven to be multiplied by the factor $\approx \sqrt{7/6}$ almost exclusively, in accordance with the theory of vibration models. The negative charge increased the Li_{Zn} vibrational frequencies. For a single-negatively charged Li_{Zn}, an increase in the frequency can be explained in terms of localized electrons. The introduced electron will be localized at the acceptor-like Li_{Zn}, and participate in the local vibrational modes. For the double-negatively charged Li_{Zn}, no significant changes were observed. As the single-negatively Li_{Zn} exhibits the same charge as the host Zn atom, the second introduced electron become delocalized and do not affect the local vibrational modes.

The relevant vibrational modes associated with the $Li_{Zn} - Li_i$ pair, were the $Li_{Zn} - Li_i$, $Li_{Zn} - O$ and $Li_i - O$ stretch modes. The calculated frequencies were $\nu = 682$ cm^{-1}, $\nu = 657$ cm^{-1} and $\nu = 565$ cm^{-1}, respectively. A close located Li defect therefore significantly modified the $Li - O$ stretch mode to a higher frequency both for the interstitial and the substitutional Li. The present frequencies indicated a strong bonding as a result of shorter interatomic distances between the Li atoms and the nearest O atom. In this respect, it is relevant to investigate the Li-pair in its ground state together with their closest oxygen atom in the host structure and compare the distances and angles with that of a primitive Li_2O unit cell. The Li_i were more stable closer to the Li_{Zn}, rather than further away, which supports the formation of a stable Li_2O. The relaxed interatomic distances and the $Li_{Zn} - O - Li_i$ angle are depicted in figure 40. For the ionic relaxation of the Li_i in position 2, which was the most stable configuration, the calculated $Li_{Zn} - Li_i$ bond was 2.09 Å. This bond length was slightly different from the same bond length in a Li_2O unit cell (2.28 Å), calculated by ref. (Wardle et al., 2005b). The $Li - O$ bonds in the host structure were different between the Li_i and the Li_{Zn} measured to be 1.88 Å and 2.10 Å, respectively. For the Li_2O unit cell the $Li - O$ bond lengths are symmetric for Li_{Zn} and Li_i at a distance of 1.97 Å, indicating a distortion of the present Li_2O in order to fit the host structure. The $Li_{Zn} - O - Li_i$ angle was calculated to 62.9^o in the present study, not too far from the theoretical calculated angle in the original Li_2O, 70.5^o, as calculated by ref. (Wardle et al., 2005b). The present results suggest that the Li_{Zn}-Li_i pair can be seen in terms of a distorted Li_2O unit cell incorporated in ZnO.

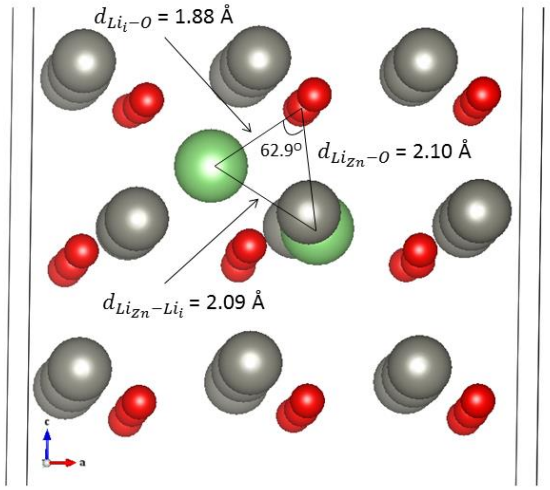

Figure 40: The relaxed structure of a $Li_{Zn} - Li_i$ in ZnO depicted in terms of a Li_2O unit cell where the Li-pair is associated with their nearest O atom to form Li_2O. The relaxed interatomic distances and the $Li_{Zn} - O - Li_i$ angle of the Li_2O are stated in the figure. The relaxed distances were slightly shorter and the cell was partly distorted compared to that of a true Li_2O unit cell calculated by ref. (Wardle et al., 2005b).

In investigation of the stability of a $Li_{Zn} - Zn_i$ pair, the pair was most stable positioned close to each other, with an interatomic distance of 2.23 Å. The stretch mode of $Li_{Zn} - Zn_i$ was calculated to $\nu = 179\ cm^{-1}$, indicating a weak bond of the $Li_{Zn} - Zn_i$ pair. However, during relaxation the Zn_i was slightly displaced towards its host position rather than stabilize in the symmetric octahedral position. This led to a repulsive force on the Li_{Zn} atom, which was displaced in the opposite direction of Zn_i, shortening the bond length of the $Li_{Zn} - O$ bond to 1.74 Å. With a covalent radius of 1.23 Å and 0.66 Å (Cordero et al., 2008) for the Li atom and the O atom, respectively, the present bond length leads to a large interaction of the $Li_{Zn} - O$ bond. As a result, the vibrational frequency of the $Li_{Zn} - O$ stretch mode was calculated to $\nu = 743\ cm^{-1}$. The Zn_i therefore modified the vibrational mode of the $Li_{Zn} - O$ stretch mode significantly. The $Li_{Zn} - O$ stretch mode was primarily located at the Li_{Zn}, indicating that the Li_{Zn} is primarily bonded to the closest O atom and is not as strongly affected and bonded to the other neighboring atoms.

6.1.4 Anisotropic transition rates of Li_i

The exact transition rates, and the corresponding diffusion coefficients, are somewhat difficult to predict, as the rate can change by order of magnitudes within a small change in the diffusion barrier, especially in the low-temperature range. The diffusion barrier defines the slope of the van't Hoff plot, and the vibrational frequencies determine the constant term. As a result, the diffusion barriers must be calculated precisely to yield the right transition rates. There are rarely a great variety in the vibrational frequencies, and frequencies are typically found in the range of $\nu = 10^{12} - 10^{13}$ THz $\approx (10^2 - 10^3\ cm^{-1})$. Therefore, the diffusion barrier is of crucial interest and the vibrational frequencies act as adjustment factors.

The present work resulted in anisotropic transition rates for Li_i diffusion. The transition rates were found to be an order of magnitude higher in the axial direction at room temperature, and the anisotropy was larger for lower temperatures. The *c*-channel is empty in the wurtzite ZnO structure, so the diffusion process is expected to be more probable in the *c*-direction. The basal diffusion was found to be assisted by an intermediate state, the tetrahedral position, 0.67 eV higher in energy than the octahedral position. The diffusion from octahedral position to tetrahedral was found to be the

rate-determining step, as the diffusion barrier $\Delta E_a^{O\to T} \gg \Delta E_a^{T\to O}$. From the literature (Huang et al., 2009, Wardle et al., 2005b, Carvalho et al., 2009), the axial diffusion barriers vary from 0.58 eV to 0.73 eV, where the latter is calculated with a HDF potential, and the basal diffusion barriers vary from 0.64 eV to 0.82 eV, where the latter is calculated with a HDF potential. The earlier reports for basal diffusion include the intermediate tetrahedral state for the interstitial Li. The present results hereby confirm the reported diffusion barriers and diffusion paths.

From the literature, there are few references regarding transition rates of Li in ZnO. However, there are a few experimental studies, evaluating the diffusion coefficients of Li_i in ZnO in the high temperature range (> 500 K) (Lander, 1960, Knutsen et al., 2013). Measured values for the diffusion barriers have been calculated somewhat higher than the theoretical barriers, $\Delta E_a = 0.98$ eV (Lander, 1960) and $\Delta E_a = 1.34$ eV (Knutsen et al., 2013). As a result, one expects deviations in the diffusion coefficients from the experimental values. The present study calculated the diffusion coefficients several orders of magnitudes larger than the experiments, as expected from a smaller diffusion barrier. However, Lander's work is over 50 years old, and applied a crude model for treatment of the theoretical data. In addition, the present work at UiO by Knutsen et al. performed SIMS measurements and a rather simple theoretical model to interpolate and estimate the activation energy. The diffusion barrier is highly sensitive to the measurements, and therefore small deviations can have major impacts on the value of the barrier. There are also some indications of a measurement error depicted in the paper by Knutsen et al. which can leads to a lower diffusion barrier, in the range of Lander's work. The larger barrier can be caused by interacting defects and competing reactions in the experimental measurements (Carvalho et al., 2009). As a result, the present study indicates that the diffusion of Li_i in ZnO is more mobile than predicted by the experimental studies.

6.1.5 Li_{Zn} as a trapping center for Li_i diffusion

The kick-out barrier for the $Li_{Zn} - Li_i$ has been calculated, where the NEB relaxed in two different MEPs. The appearance of two calculated MEPs is discussed in the next section. The calculated kick-out barrier, $\Delta E_{a,Li_{Zn}-Li_i} \leq 0.37$ eV, was almost halved compared to the lowest interstitial diffusion barrier, $\Delta E_{a,\parallel} = 0.67$ eV. In addition there were calculated MEPs for the kick-out mechanism of the $Li_{Zn} - Li_i$ pair indicating a barrier at an even lower value, ~0.22 eV. When Li_{Zn} is present in the compound, Li_i will more likely kick out a neighboring Li_{Zn} atom, than diffuse directly to the next octahedral site. In this respect, Li_{Zn} modifies the transition rate of Li_i diffusion. However, the Li_{Zn} can prevent further Li_i diffusion. After the kick-out process, if no neighboring Li_{Zn} is present, the kicked-out Li atom will prefer to reverse the kick-out process rather than diffuse further, as the given kick-out barrier is significantly lower than the barrier for interstitial diffusion. Hereby, Li_{Zn} can act as a trapping center for Li_i diffusion.

The kick-out mechanism of the $Li_{Zn} - Li_i$ pair was not symmetric about the transition state. In the present study, Li_i has been located in the octahedral position 1 as an initial configuration. After the kick-out process, the Li_{Zn} was kicked out to the octahedral position 6, $Li_{i,6}$. The configuration of a $Li_{Zn} - Li_{i,6}$ pair was slightly more stable than a $Li_{Zn} - Li_{i,1}$ pair by ~0.12 eV. The energy difference is caused by a shorter interatomic distance of the Li-Li pair for the interstitial position 6 resulting in a more energetically favorable configuration. In this respect, the kick-out barrier is related to the initial position of the interstitial Li atom. If the kick-out process of $Li_{i,1} + Li_{Zn} \rightarrow Li_{Zn} + Li_{i,6}$ is reversed, the kick-out barrier increases by 0.12 eV, and one obtains $\Delta E_{a,Li_{Zn}-Li_i} \leq 0.49$ eV. As a material always

searches the lowest possible energy configuration, Li_i is expected to be located in interstitial position 2 (or position 6 or position 8). However, if a Li_i is present in interstitial position 1, one expects that the kick-out process might occur rather than interstitial diffusion to position 2, so the previous presented barrier is still highly relevant. Earlier reports (Wardle et al., 2005b) states that the formation of a stable Li_2O unit cell inside ZnO is a limiting factor on Li_i diffusion. The present work supports this statement, but predicts that the formation of a stable Li_2O is due to the low barrier of the $Li_{Zn} - Li_i$ kick-out mechanism, and therefore the kick-out mechanism is the actual limiting factor of Li_i diffusion.

6.1.6 $Li_{Zn} - Li_i$ kick-out mechanism

The present work showed a remarkable result as the NEB calculation for the $Li_{Zn} – Li_i$ kick-out mechanism generated two different MEPs dependent on the initially chosen path. The calculation is supposed to give an unambiguously MEP, which is defined as the lowest-barrier path between two endpoints. The step-wise initial path led to small distances between the two Li atoms for some images and might give large movements during the relaxation due to the high repulsive force between the atoms. However, what is remarkable is that the step-wise path led to the lowest barrier as it was the path established from a non-stable initial path. Furthermore, the question to be asked is why the linear initial path did not achieve the same path, as it was a stable initialized path.

A possible explanation is that the NEB calculation has reached a local MEP. For a $Li_{Zn} - Li_i$ pair, the kick-out process is more complex than single-defect diffusions between two endpoints and several minima might occur when two atoms are relocated. The ionic relaxation from the linear path resulted in a MEP not too far from the initial guess. As described earlier, the step-wise initial path led to large distortions due to the small interatomic $Li_{Zn} - Li_i$ distances for some images. During ionic relaxation, the initial Li_i was significantly moved away from the initial path which often occurs in ionic relaxations when the forces between atoms are large. In terms of the potential energy surface for the kick-out process, the significantly moved Li_i atom might have reached another potential minimum with an even lower energy. The step-wise initialized path started with a jump from $Li_{i,1}$ towards Li_{Zn}. An additional step-wise NEB calculation starting with a jump from Li_{Zn} towards $Li_{i,8}$ was carried out, and resulted in a kick-out barrier of ~0.37 eV. This result confirms the explanation given above with large repulsive forces as the main reason for the low-barrier MEP. Another suggested method to investigate the result even further is to use a large displacement value during ionic relaxation, to see whether the linear initialized path can reach the global MEP.

The occurrence of two MEPs may also be related to the spring constant. However, the appropriate spring constant is used, together with another spring constant to confirm consistency. More values of the spring constants are preferable to check whether the MEP is dependent on the choice of the spring constant. In addition, the step-wise path was only tested for 2, 3 and 7 images. It is desirable to try even more images (higher resolution), to see at what step-distance the distinction between the two paths occur. There were observed no significant differences in the relaxed forces on the atoms for the images of the linear and the step-wise paths. Forces were relaxed in the 0.1 eV/Å range for the images, with exceptions of smaller forces for some images, as is expected since the relaxed images are not local/global minima, but saddle points with forces pointing toward the local/global endpoints.

In the discussion of the two calculated MEPs, the interatomic relaxed bond lengths between the two Li atoms have been evaluated along the MEP for the step-wise and the linear initialized paths from position 1 to position 6. These interatomic $Li_{Zn} - Li_i$ distances as a function of configuration coordinates along the kick-out path can be seen in figure 41. The $Li_{Zn} - Li_i$ distances were 2.29 Å and 2.09 Å for the initial and the final configurations, respectively. For the two different initialized paths, the interatomic distance decreased along the path, reaching a minimum at 2.00 Å and 1.93 Å for the linear and step-wise initialized path, respectively. This result indicates that a shorter interatomic distance leads to a more stable configuration for the $Li_{Zn} - Li_i$ pair.

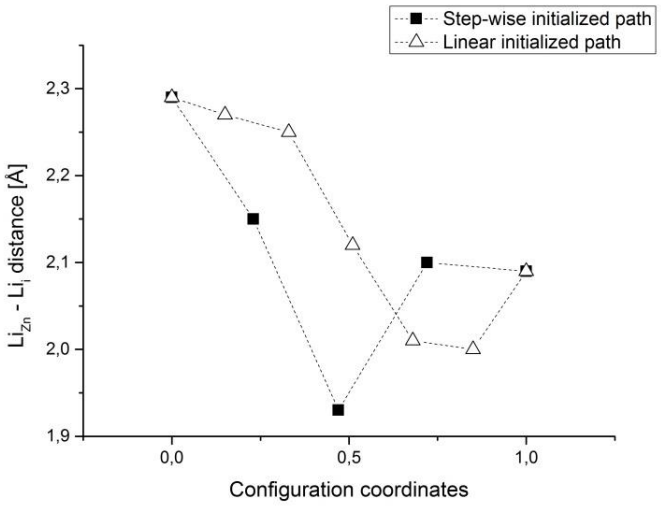

Figure 41: The calculated $Li_{Zn} - Li_i$ distances along the two calculated MEPs for the $Li_{Zn} - Li_i$ kick-out mechanism. The interatomic distances decreased along the paths, and for the step-wise initialized kick-out path the interatomic distance reached a minimum at 1.93 Å. As the step-wise initialized path had the lowest kick-out barrier this indicates that a short $Li_{Zn} - Li_i$ distance is preferred for the $Li_{Zn} - Li_i$ pair, predicting the $Li_{Zn} - Li_i$ pair are strongly attracted.

From the fact that one of the calculated MEPs has a significant increase in the kick-out path, it is relevant to discuss whether a MEP can have an unlimited length. A MEP is defined as the path that has the lowest energy barrier between two endpoints. However, it is reasonable to suggest that a diffusion/kick-out process will not be appropriate if it includes a significantly larger distance compared to a shorter path with a slightly higher barrier. To the most extreme; is it reasonable that a diffusing atom will pass through the whole crystal to diffuse to another nearby position just because the path has a slightly lower barrier? Of course, the question is not relevant for energies that are just above the MEP as MEP will then be the only possible path, but for energies somewhat larger it is reasonable to suggest that diffusion through a slightly higher barrier will be probable. To what extent the higher-barrier transition is present is definitely individual for each transition system and will be dependent on the temperature and the MEP distance.

6.1.7 $Li_{Zn} - Zn_i$ kick-out mechanism

From the relaxation, the total cohesive energy was calculated to be ~1.1 eV higher for the $Li_{Zn} - Zn_i$ configuration relative to the presence of an octahedral positioned Li_i in the host structure. This indicates that the Zn atom prefers to be located at its host site, as expected. The kick-out mechanism for an interstitial Li to kick out a host Zn atom has been investigated. For the calculated MEP, the potential energy surface increased towards the endpoint, with a possible meta-stable configuration along the MEP. The possible meta-stable state resulted in a configuration where the Li and Zn atom

had the same distance to the host Zn site. The state is expected to be slightly stable due to the center of mass and charge near the host Zn site by symmetry considerations. Whether the intermediate state is stable or not along the path is not clear to this point due to too small resolution in the calculations. The presence of the intermediate state was only predicted by an estimation of the MEP. Moreover, the state will not be stable for a wide energy range, and the transition rate are not expected to be modified by this state.

The ionic relaxations are expected to give energy minima. However, for the $Li_{Zn} - Zn_i$ pair with Zn positioned in octahedral position 2, the MEP indicated that the relaxed configuration was not an energy minimum. The ionic relaxation might therefore have reached a slightly stable energy minimum, which was confirmed by the ionic relaxation that did not reach the force tolerance of the ions and experienced a drift on the Li atom at 0.12 eV/Å. This indicates that the Zn atom is not stable in an octahedral interstitial position. During the ionic relaxation, the Zn atom was displaced in all directions in order to find energy minima. As the resulting configuration was only slightly stable, the interstitial Zn atom is expected to be stable far from the octahedral position. Moreover, the interstitial Zn is not certain to reach any stable configuration as it prefers a host site in the material. A suggestion is to increase the displacement distance during the ionic relaxation in order to search for more stable positions for the Zn_i.

6.2 Summary

In the present work there has been calculated several properties concerning hydrogen, lithium and nickel in ZnO. The following section will give a brief summary of the calculations performed, and the main results.

Hydrogen is a common interstitial impurity in ZnO. Due to its small radius, substitutional H_{Zn} is not appropriate. In this thesis the relaxed configurations of H_i in ZnO have been calculated and H_i was most stable in a bond-centered position, parallel to the axial Zn-O bond, $BC_{\parallel}$, in accordance with previous reports. In addition, the vibrational frequencies were carried out within the LDA with an adequate accuracy; with an error bar of $\sim$4 % from the earlier references. However, some systematic errors were present, together with a slightly different relaxed structure, which made it difficult to reproduce earlier calculations exactly. In addition, the vibrational modes were calculated for the interstitial H within the GGA by PBE potentials, and these frequencies were closer to the measured values, within an error bar of $\sim$0.5 eV.

Lithium can be introduced unintentionally in ZnO during hydrothermal growth, but can also be introduced intentionally in order to possibly obtain a shallow acceptor level in ZnO and hence give a stable *p*-type ZnO. However, no experimental data confirms the presence of a stable *p*-type ZnO by Li-doping. Li_i was found to be most stable in octahedral positions, about 0.68 eV more stable than Li_i in tetrahedral positions in ZnO. ZnO exhibit wurtzite structure where all the octahedral positions are empty, ready to be occupied. In the present study, the vibrational frequencies of interstitial Li have been carried out. The calculated frequencies of Li_i were directional dependent with a significant larger frequency ($\sim$20 %) in the basal direction, indicating a stronger $Li_i - O$ bond in the basal plane. The result is reasonable as the *c*-channel is empty indicating a weak axial bonding. For Li_{Zn}, the vibrational frequencies were nearly isotropic. The result is also reasonable as Li_{Zn} was surrounded by four oxygen atoms of equal lengths. Additional calculations of negatively charged substitutional Li

were performed, and an increase in the frequencies was observed. The vibrational frequencies of the $^{6}Li_{Zn}$ showed expected vibrational frequencies in accordance to Hooke's law for springs.

To investigate the diffusion barriers of Li_i, the nudged elastic band method was employed and evaluated diffusion between octahedral positions. The diffusion barrier in the axial direction was $\Delta E_{a,\parallel} = 0.68$ eV for the single-positively charged Li_i, and slightly smaller for the neutral and the negatively charged Li_i. The calculated diffusion barrier in the basal plane was slightly higher, $\Delta E_{a,\perp} = 0.73$ eV for the positively charged Li_i and $\Delta E_{a,\perp} = 0.72$ eV for the neutral Li_i. The diffusion in the basal direction involved an intermediate tetrahedral position. The intermediate state is not present at normal conditions as the barrier is too low to prevent diffusion towards the octahedral position. From the vibrational frequencies and the diffusion barriers, the transition rates were calculated according to transition state theory. From the fact that the diffusion barriers were directional dependent, the transition rates of Li in ZnO were anisotropic. In addition, the corresponding diffusion coefficients were carried out. In the present study, Li_i was found to be mobile under normal conditions. However, the calculation of transition rates and corresponding diffusion coefficients are highly dependent on the diffusion barrier and hence a small difference in diffusion carrier can give a large impact on the transition rates. Therefore it is more reasonable to compare the magnitudes of the transition rates rather than the specific values.

The calculated kick-out barrier of a $Li_{Zn} - Li_i$ pair is lower than the diffusion barrier of Li_i indicating that the kick-out mechanism will be present to large extent at room temperature. However, after the kick-out mechanism, the Li_i will rather reverse the kick-out process than continue to the next interstitial Li due to the lower activation barrier for the kick-out mechanism. In this respect, the Li_{Zn} can act as a trapping center and prevent further diffusion of Li_i. The effect is dependent of the fraction of substitutional Li atoms. The NEB calculation related to the kick-out mechanism of the Li_{Zn} – Li_i pair resulted in two different paths dependent on the initially chosen path. A reasonable explanation regarding the ionic relaxation was presented. When the Li_i jumps towards the Li_{Zn} tension is created between the atoms. As a result, the Li_i is significantly moved during ionic relaxation and can occasionally reach another potential well leading to another relaxed configuration.

A $Li_{Zn} - Zn_i$ pair was found to be unstable in the host structure. The calculated energy was 1.1 eV higher compared to the structure including a single interstitial Li. The kick-out mechanism of a Li – Zn pair has been studied to some extent. The kick-out configuration was found to spontaneously reverse and result in an interstitial Li in a ZnO structure for all normal conditions. However, the Zn_i might diffuse further into the material to other local minima. To this point, the diffusion of an interstitial Li atom to a Zn site by a kick-out mechanism is unfavorable.

Finally, the investigation of Ni_{Zn} in ZnO was considered. The calculated $Ni_{Zn} - O$ stretch mode frequency was found to be slightly higher than for the $Zn - O$ stretch mode. The relaxed interatomic distance of the $Ni_{Zn} - O$ bond was shorter than for the host structure, and together with the lower mass of Ni_{Zn} one can expect a higher frequency for the $Ni_{Zn} - O$ stretch mode. Ni_{Zn} in ZnO co-doped with Li_{Zn} indicated that the Li_{Zn} prefers to locate close to the Ni_{Zn} as the acceptor-like Li_{Zn} attracts an electron above the Fermi level from the high spin-polarized energy levels introduced by Ni_{Zn}. Li_{Zn} was found to modify the vibrational frequency of the $Ni_{Zn} - O$ stretch mode to a slightly higher frequency. As the Li_{Zn} attracts an electron from the high-energetic levels, the Ni_{Zn} loses one electron. As a result, Ni_{Zn} will be positively charged and hence more attracted to the negative O atom leading to higher

vibrational frequency. The magnetization of Ni_{Zn} was calculated as a function of Li-doping and free-hole doping. The results showed a decrease in the magnetization when Li_{Zn} was located near the Ni_{Zn} at Li-concentrations larger than 2 %. However, if the Li_{Zn} were placed randomly, the magnetization could be described in terms of free-hole doping up to 6 %, in accordance with previous reports.

6.3 Further work

The presented topics in this thesis have to a large extent been studied comprehensively. However, when working on a project within a limited time period, one is not able to perform calculations with all the available parameters in order to find the true ground states for all calculations. The one topic which is suggested to be considered further is the $Li_{Zn} - Li_i$ kick-out mechanism that resulted in an inconsistent result. The smooth linear interpolation resulted in a larger barrier compared to the initial path based on longer steps from one atom at a time. For further investigation it is suggested to use a larger displacement value for the ionic relaxation to check the appearance of a local and global MEP, and to use more images in order to obtain better resolution, both for the linear and step-wise kick-out path. In addition, one can change the spring parameter to other values, and use different ionic relaxation algorithms to uncover the inconsistency, whether there is a weakness related to the NEB method itself, or in the chosen parameters. Another interesting topic within the $Li_{Zn} - Li_i$ kick out mechanism is to investigate how the kick-out barrier and the MEP are affected by charge effects on the Li atoms.

The results of the present study, based on the calculated vibrational frequencies, diffusion and kick-out barriers and transition rates of Li in ZnO, is suggested to be published in a scientific article. Either the article will be published based solely on the present study or in collaboration with results obtained from colleagues at the material science group at UiO. The theoretical calculations are desired to be published together with some experimental data. The experimental local vibrational frequencies for Li in ZnO are not well described in the literature, and an FTIR measurement of Li in ZnO can provide these frequencies. The author has been in touch with the experimental semiconductor group, LENS, at UiO, to make this possible. However, the new FTIR instrument at UiO do not achieve high enough resolution in the low frequency range, which makes it difficult to identify the local vibrational modes of Li experimentally at this point.

7. References

ADAMO, C. & BARONE, V. 1999. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *Journal of Chemical Physics,* 110**,** 6158-6170.

ALFE, D. 2009. PHON: A program to calculate phonons using the small displacement method. *Computer Physics Communications,* 180**,** 2622-2633.

ANDERSEN, O. K. 1975. Linear Methods in Band Theory. *Physical Review B,* 12**,** 3060-3083.

AWAN, S. U., HASANAIN, S. K., BERTINO, M. F. & JAFFARI, G. H. 2012. Ferromagnetism in Li doped ZnO nanoparticles: The role of interstitial Li. *Journal of Applied Physics,* 112.

BACHELET, G. B., HAMANN, D. R. & SCHLUTER, M. 1982. Pseudopotentials That Work - from H to Pu. *Physical Review B,* 26**,** 4199-4228.

BALLUFFI, R. W., ALLEN, S. M., CARTER, W. C. & KEMPER, R. A. 2005. *Kinetics of materials,* Hoboken, N.J., J. Wiley & Sons.

BARONI, S., GIANNOZZI, P. & TESTA, A. 1987. Green-Function Approach to Linear Response in Solids. *Physical Review Letters,* 58**,** 1861-1864.

BECKE, A. D. 1988a. Correlation-Energy of an Inhomogeneous Electron-Gas - a Coordinate-Space Model. *Journal of Chemical Physics,* 88**,** 1053-1062.

BECKE, A. D. 1988b. Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Physical Review A,* 38**,** 3098-3100.

BECKE, A. D. 1993a. Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics,* 98**,** 5648-5652.

BECKE, A. D. 1993b. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *Journal of Chemical Physics,* 98**,** 1372-1377.

BECKE, A. D. 1997. Density-functional thermochemistry .5. Systematic optimization of exchange-correlation functionals. *Journal of Chemical Physics,* 107**,** 8554-8560.

BLOCH, F. 1929. About the Quantum mechanics of Electrons in Crystal lattices. *Zeitschrift Fur Physik,* 52**,** 555-600.

BLÖCHL, P. E. 1994. Projector Augmented-Wave Method. *Physical Review B,* 50**,** 17953-17979.

BORN, M. & OPPENHEIMER, R. 1927. Quantum theory of molecules. *Annalen Der Physik,* 84**,** 0457-0484.

CARVALHO, A., ALKAUSKAS, A., PASQUARELLO, A., TAGANTSEV, A. K. & SETTER, N. 2009. A hybrid density functional study of lithium in ZnO: Stability, ionization levels, and diffusion. *Physical Review B,* 80.

CCE-GROUP. *Zinc Oxide - The products* [Online]. Available: http://www.cn-zinc-oxide.com/product.htm [Accessed 01.11.13.

CHAKRAVORTY, S. J. & CLEMENTI, E. 1989. Soft Coulomb Hole for the Hartree-Fock Model to Estimate Atomic Correlation Energies. *Physical Review A,* 39**,** 2290-2296.

CHIANG, T. C. & HIMPSEL, F. J. *2.1.3 Si*, SpringerMaterials - The Landolt-Börnstein Database (www.springermaterials.com).

CORDERO, B., GOMEZ, V., PLATERO-PRATS, A. E., REVES, M., ECHEVERRIA, J., CREMADES, E., BARRAGAN, F. & ALVAREZ, S. 2008. Covalent radii revisited. *Dalton Transactions***,** 2832-2838.

COTTENIER, S. 2002. *Density Functional Theory and the family of (L)APW-methods: a step-by-step introduction,* (to be found at http://www.wien2k.at/reg_user/textbooks), Instituut voor Kern- en Stralingsfysica, K.U. Leuven, Belgium.

CRADDOCK, P. T. E. A. 1998. *2000 Years of Zinc and Brass*, British Museum.

DALCORSO, A., POSTERNAK, M., RESTA, R. & BALDERESCHI, A. 1994. Ab-Initio Study of Piezoelectricity and Spontaneous Polarization in Zno. *Physical Review B,* 50**,** 10715-10721.

DAVIS, U. C. *Cubic Lattices and Close Packing* [Online]. Available: http://chemwiki.ucdavis.edu/Under_Construction/chem1/States_of_Matter/Cubic_Lattices_and_Close_Packing [Accessed 10.11.2013.

DIETL, T., OHNO, H., MATSUKURA, F., CIBERT, J. & FERRAND, D. 2000. Zener model description of ferromagnetism in zinc-blende magnetic semiconductors. *Science,* 287**,** 1019-1022.

DILLI, Z. 2008. Intrinsic and Extrinsic Semiconductors, Fermi-Dirac Distribution Function, the Fermi level and carrier concentrations. *In:* DILLI, Z. (ed.).

DIRAC, P. A. M. 1928. The quantum theory of the electron. *Proceedings of the Royal Society of London Series a-Containing Papers of a Mathematical and Physical Character,* 117**,** 610-624.

ECKART, C. 1934. The kinetic energy of polyatomic molecules. *Physical Review,* 46**,** 0383-0387.

ELBER, R. & KARPLUS, M. 1987. A Method for Determining Reaction Paths in Large Molecules - Application to Myoglobin. *Chemical Physics Letters,* 139**,** 375-380.

EYRING, H. 1935. The activated complex and the absolute rate of chemical reactions. *Chemical Reviews,* 17**,** 65-77.

FERMI, E. 1928. A statistical Method for Determining some Properties of the Atoms and its Application to the Theory of the periodic Table of Elements. *Zeitschrift Fur Physik,* 48**,** 73-79.

FOCK, V. 1930. Approximation method for the solution of the quantum mechanical multibody problems. *Zeitschrift Fur Physik,* 61**,** 126-148.

FROESE FISCHER, C. 2003. *Douglas Rayner Hartree : his life in science and computing,* Singapore ; River Edge, NJ, World Scientific.

FUKUDA, R., KOTANI, T., SUZUKI, Y. & YOKOJIMA, S. 1994. Density-Functional Theory through Legendre Transformation. *Progress of Theoretical Physics,* 92**,** 833-862.

GIDDINGS, J. C. & EYRING, H. 1958. Multi-Barrier Kinetics - Nucleation. *Journal of Physical Chemistry,* 62**,** 305-308.

GOEDECKER, S. & MASCHKE, K. 1992. Transferability of Pseudopotentials. *Physical Review A,* 45**,** 88-93.

GRIFFITHS, D. J. 2005. *Introduction to quantum mechanics,* Upper Saddle River, NJ, Pearson Prentice Hall.

HAAS, P., TRAN, F. & BLAHA, P. 2009. Calculation of the lattice constant of solids with semilocal functionals. *Physical Review B,* 79.

HAFNER, J. 2008. Ab-initio simulations of materials using VASP: Density-functional theory and beyond. *Journal of Computational Chemistry,* 29**,** 2044-2078.

HAMANN, D. R., SCHLUTER, M. & CHIANG, C. 1979. Norm-Conserving Pseudopotentials. *Physical Review Letters,* 43**,** 1494-1497.

HAMMER, B., HANSEN, L. B. & NØRSKOV, J. K. 1999. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Physical Review B,* 59**,** 7413-7421.

HANDY, N. C. & COHEN, A. J. 2001. Left-right correlation energy. *Molecular Physics,* 99**,** 403-412.

HANGGI, P., TALKNER, P. & BORKOVEC, M. 1990. Reaction-Rate Theory - 50 Years after Kramers. *Reviews of Modern Physics,* 62**,** 251-341.

HELLMANN, H. 1935. A new approximation method in the problem of many electrons. *Journal of Chemical Physics,* 3**,** 61-61.

HENKELMAN, G. & JONSSON, H. 1999. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *Journal of Chemical Physics,* 111**,** 7010-7022.

HENKELMAN, G., UBERUAGA, B. P. & JONSSON, H. 2000. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *Journal of Chemical Physics,* 113**,** 9901-9904.

HEO, Y. W., TIEN, L. C., KWON, Y., NORTON, D. P., PEARTON, S. J., KANG, B. S. & REN, F. 2004. Depletion-mode ZnO nanowire field-effect transistor. *Applied Physics Letters,* 85**,** 2274-2276.

HEYD, J., SCUSERIA, G. E. & ERNZERHOF, M. 2003. Hybrid functionals based on a screened Coulomb potential. *Journal of Chemical Physics,* 118**,** 8207-8215.

HIMMETOGLU, B., FLORIS, A., DE GIRONCOLI, S. & COCOCCIONI, M. 2013. Hubbard-corrected DFT energy functionals: the LDA+U description of correlated systems.

HOHENBERG, P. & KOHN, W. 1964. Inhomogeneous Electron Gas. *Physical Review B,* 136**,** B864-&.

HU, J. & PAN, B. C. 2008. Electronic structures of defects in ZnO: Hybrid density functional studies. *Journal of Chemical Physics,* 129.

HUANG, G. Y., WANG, C. Y. & WANG, J. T. 2009. First-principles study of diffusion of Li, Na, K and Ag in ZnO. *Journal of Physics-Condensed Matter,* 21.

JAGADISH, C. & PEARTON, S. J. 2006. *Zinc oxide bulk, thin films and nanostructures : processing, properties and applications,* Amsterdam ; Boston, Elsevier.

JAYAKUMAR, O. D., SUDAKAR, C., PERSSON, C., SUDARSAN, V., NAIK, R. & TYAGI, A. K. 2010. Tunable Ferromagnetism accompanied by Morphology Control in Li-doped Zn0.97Ni0.03O. *Journal of Physical Chemistry C,* 114**,** 17428-17433.

JOB, G. & HERRMANN, F. 2006. Chemical potential - a quantity in search of recognition. *European Journal of Physics,* 27**,** 353-371.

JONSSON, H., MILLS, G. & JACOBSEN, K. W. 1997. Nudged elastic band method for finding minimum energy paths of transitions. *In:* BERNE, B. J. & CICCOTTI, G. (eds.) *Classical and Quantum Dynamics in Condensed Phase Simulations.*

KARZEL, H., POTZEL, W., KOFFERLEIN, M., SCHIESSL, W., STEINER, M., HILLER, U., KALVIUS, G. M., MITCHELL, D. W., DAS, T. P., BLAHA, P., SCHWARZ, K. & PASTERNAK, M. P. 1996. Lattice dynamics and hyperfine interactions in ZnO and ZnSe at high external pressures. *Physical Review B,* 53**,** 11425-11438.

KASAP, S. O. & CAPPER, P. 2006. *Springer handbook of electronic and photonic materials,* New York, Springer.

KISI, E. H. & ELCOMBE, M. M. 1989. Upsilon-Parameters for the Wurtzite Structure of Zns and Zno Using Powder Neutron-Diffraction. *Acta Crystallographica Section C-Crystal Structure Communications,* 45**,** 1867-1870.

KITCHIN, J. 2012. Modeling materials using density functional theory.

KITTEL, C. 2005. *Introduction to solid state physics,* Hoboken, NJ, Wiley.

KLEINMAN, L. & BYLANDER, D. M. 1982. Efficacious Form for Model Pseudopotentials. *Physical Review Letters,* 48**,** 1425-1428.

KLINGSHIRN, C. 2007. ZnO: Material, Physics and Applications. *ChemPhysChem,* 8**,** 782-803.

KLINGSHIRN, C. F. 2010. *Zinc Oxide : from fundamental properties towards novel applications,* Heidelberg ; London, Springer.

KNUTSEN, K. E., JOHANSEN, K. M., NEUVONEN, P. T., SVENSSON, B. G. & KUZNETSOV, A. Y. 2013. Diffusion and configuration of Li in ZnO. *Journal of Applied Physics,* 113.

KOBIAKOV, I. B. 1980. Elastic, Piezoelectric and Dielectric-Properties of Zno and Cds Single-Crystals in a Wide-Range of Temperatures. *Solid State Communications,* 35**,** 305-310.

KOELLING, D. D. & HARMON, B. N. 1977. Technique for Relativistic Spin-Polarized Calculations. *Journal of Physics C-Solid State Physics,* 10**,** 3107-3114.

KOFSTAD, P. & NORBY, T. 2012. *Defects and Transport in Crystalline Solids*.

KOHN, W. 1999. Nobel Lecture: Electronic structure of matter-wave functions and density functionals. *Reviews of Modern Physics,* 71**,** 1253-1266.

KOHN, W. & SHAM, L. J. 1965. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review,* 140**,** 1133-&.

KRESSE, G. & JOUBERT, D. 1999. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B,* 59**,** 1758-1775.

KROGER, F. A. & VINK, H. J. 1956. Relations between the Concentrations of Imperfections in Crystalline Solids. *Solid State Physics-Advances in Research and Applications,* 3**,** 307-435.

LAIDLER, K. J. & KING, M. C. 1983. The Development of Transition-State Theory. *Journal of Physical Chemistry,* 87**,** 2657-2664.

LALENA, J. N. & CLEARY, D. A. 2010. *Principles of inorganic materials design,* Hoboken, N.J., John Wiley.

LANDAU, L. D. & LIFSHITS, E. M. 1976. *Mechanics,* Oxford ; New York, Pergamon Press.

LANDAU, L. D. & LIFSHITS, E. M. 1977. *Quantum mechanics : non-relativistic theory,* Oxford ; New York, Pergamon Press.

LANDER, J. J. 1960. Reactions of Lithium as a Donor and an Acceptor in Zno. *Journal of Physics and Chemistry of Solids,* 15**,** 324-334.

LAVROV, E. V., HERKLOTZ, F. & KUTIN, Y. S. 2013. Effect of uniaxial stress on substitutional Ni in ZnO. *Solid State Communications,* 159**,** 36-40.

LAVROV, E. V., WEBER, J., BORRNERT, F., VAN DE WALLE, C. G. & HELBIG, R. 2002. Hydrogen-related defects in ZnO studied by infrared absorption spectroscopy. *Physical Review B,* 66.

LEVY, M. 1979. Universal Variational Functionals of Electron-Densities, 1st-Order Density-Matrices, and Natural Spin-Orbitals and Solution of the V-Representability Problem. *Proceedings of the National Academy of Sciences of the United States of America,* 76**,** 6062-6065.

LI, Y. B., BANDO, Y. & GOLBERG, D. 2004. ZnO nanoneedles with tip surface perturbations: Excellent field emitters. *Applied Physics Letters,* 84**,** 3603-3605.

LIMPIJUMNONG, S. & ZHANG, S. B. 2005. Resolving hydrogen binding sites by pressure - A first-principles prediction for ZnO. *Applied Physics Letters,* 86.

MALAERU, T., NEAMTU, J., MORARI, C. & SBARCEA, G. 2012. Structural and Magnetic Properties of Nanocrystalline Powders of Ni-Doped Zno Diluted Magnetic Semiconductors Synthesized by Sol-Gel Method. *Revue Roumaine De Chimie,* 57**,** 857-862.

MARQUES, M. A. L., MAITRA, N. T., NOGUEIRA, F. M. S., GROSS, E. K. U. & RUBIO, A. 2011a. *Fundamentals of time-dependent density functional theory,* New York, Springer.

MARQUES, M. A. L., VIDAL, J., OLIVEIRA, M. J. T., REINING, L. & BOTTI, S. 2011b. Density-based mixing parameter for hybrid functionals. *Physical Review B,* 83.

MARTIN, R. M. 2004. *Electronic structure : basic theory and practical methods,* Cambridge, UK ; New York, Cambridge University Press.

MERMIN, N. D. 1965. Thermal Properties of Inhomogeneous Electron Gas. *Physical Review,* 137**,** 1441-&.

MEYER, B. K., ALVES, H., HOFMANN, D. M., KRIEGSEIS, W., FORSTER, D., BERTRAM, F., CHRISTEN, J., HOFFMANN, A., STRASSBURG, M., DWORZAK, M., HABOECK, U. & RODINA, A. V. 2004. Bound exciton and donor-acceptor pair recombinations in ZnO. *Physica Status Solidi B-Basic Solid State Physics,* 241**,** 231-260.

MILLER, W. H. 1839. *A treatise on crystallography,* Cambridge,, For J. & J. J. Deighton; etc.

MOLLWO, E. 1954. Die Wirkung Von Wasserstoff Auf Die Leitfahigkeit Und Lumineszenz Von Zinkoxydkristallen. *Zeitschrift Fur Physik,* 138**,** 478-488.

MONKHORST, H. J. & PACK, J. D. 1976. Special Points for Brillouin-Zone Integrations. *Physical Review B,* 13**,** 5188-5192.

MORKOÇ, H. & ÖZGÜR, U. M. 2009. *Zinc oxide : fundamentals, materials and device technology,* Weinheim, Wiley-VCH.

NI, H. Q., LU, Y. F., LIU, Z. Y., QIU, H., WANG, W. J., REN, Z. M., CHOW, S. K. & JIE, Y. X. 2001. Investigation of Li-doped ferroelectric and piezoelectric ZnO films by electric force microscopy and Raman spectroscopy. *Applied Physics Letters,* 79**,** 812-814.

NOBELPRIZE.ORG. 1998. *Press release: The 1998 Nobel Prize in Chemistry* [Online]. Nobelprize.org: Nobel Media. Available: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998/press.html.

OLESON, J. P. 2008. *Oxford handbook of engineering and technology in the Classical world,* Oxford ; New York, Oxford University Press.

OLIVEIRA, L. N., GROSS, E. K. U. & KOHN, W. 1988. Density-Functional Theory for Superconductors. *Physical Review Letters,* 60**,** 2430-2433.

OLIVER, G. L. & PERDEW, J. P. 1979. Spin-Density Gradient Expansion for the Kinetic-Energy. *Physical Review A,* 20**,** 397-403.

ONODERA, A., TAMAKI, N., KAWAMURA, Y., SAWADA, T. & YAMASHITA, H. 1996. Dielectric activity and ferroelectricity in piezoelectric semiconductor Li-doped ZnO. *Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers,* 35**,** 5160-5162.

PAIER, J., MARSMAN, M., HUMMER, K., KRESSE, G., GERBER, I. C. & ANGYAN, J. G. 2006. Screened hybrid density functionals applied to solids. *Journal of Chemical Physics,* 124.

PARK, C. H., ZHANG, S. B. & WEI, S. H. 2002. Origin of p-type doping difficulty in ZnO: The impurity perspective. *Physical Review B,* 66.

PARTINGTON, J. R. 1989. *A short history of chemistry,* New York, Dover Publications.

PAYNE, M. C., TETER, M. P., ALLAN, D. C., ARIAS, T. A. & JOANNOPOULOS, J. D. 1992. Iterative Minimization Techniques for Abinitio Total-Energy Calculations - Molecular-Dynamics and Conjugate Gradients. *Reviews of Modern Physics,* 64**,** 1045-1097.

PEARTON, S. J., NORTON, D. P., IP, K., HEO, Y. W. & STEINER, T. 2005. Recent progress in processing and properties of ZnO. *Progress in Materials Science,* 50**,** 293-340.

PECHUKAS, P. 1981. Transition-State Theory. *Annual Review of Physical Chemistry,* 32**,** 159-177.

PERDEW, J. P., BURKE, K. & ERNZERHOF, M. 1996. Generalized gradient approximation made simple. *Physical Review Letters,* 77**,** 3865-3868.

PERDEW, J. P., CHEVARY, J. A., VOSKO, S. H., JACKSON, K. A., PEDERSON, M. R., SINGH, D. J. & FIOLHAIS, C. 1992. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B,* 46**,** 6671-6687.

PERDEW, J. P. & LEVY, M. 1983. Physical Content of the Exact Kohn-Sham Orbital Energies - Band-Gaps and Derivative Discontinuities. *Physical Review Letters,* 51**,** 1884-1887.

PERSSON, C. 2012. *Brief Introduction to the Density Funtional Theory*.

PERSSON, C., JAYAKUMAR, O. D., SUDAKAR, C., SUDARSAN, V. & TYAGI, A. K. 2011. Morphology and Magnetic Coupling in ZnO:Co and ZnO:Ni Co-Doped with Li. *Acta Physica Polonica A,* 119**,** 95-98.

RAJAGOPA.AK & CALLAWAY, J. 1973. Inhomogeneous Electron-Gas. *Physical Review B,* 7**,** 1912-1919.

RUNGE, E. & GROSS, E. K. U. 1984. Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters,* 52**,** 997-1000.

SCHIRMER, O. F. 1968. Structure of Paramagnetic Lithium Center in Zinc Oxide and Beryllium Oxide. *Journal of Physics and Chemistry of Solids,* 29**,** 1407-&.

SCHOONMAN, J. 2005. Nano-structured functional materials: The defect chemistry. *Nanostructured and Advanced Materials for Applications in Sensor, Optoelectronic and Photovoltaic Technology,* 204**,** 259-270.

SCHWERDTFEGER, P. 2011. The Pseudopotential Approximation in Electronic Structure Theory. *Chemphyschem,* 12**,** 3143-3155.

SHEPPARD, D., TERRELL, R. & HENKELMAN, G. 2008. Optimization methods for finding minimum energy paths. *Journal of Chemical Physics,* 128.

SHOLL, D. S. & STECKEL, J. A. 2009. *Density functional theory : a practical introduction,* Hoboken, N.J., Wiley.

SINGH, D. J. & NORDSTRÖM, L. 2006. *Planewaves, pseudopotentials, and the LAPW method,* New York, NY, Springer.

SITES.GOOGLE.COM. *History of Zinc Oxide* [Online]. Available: https://sites.google.com/site/zincoxidetco/project-definition [Accessed 01.09.2013.

SLATER, J. C. 1953. An Augmented Plane Wave Method for the Periodic Potential Problem. *Physical Review,* 92**,** 603-608.

SPENCER, J. N., BODNER, G. M. & RICKARD, L. H. 2012. *Chemistry : structure and dynamics,* Hoboken, N.J., Wiley.

SPROUL, A. B. & GREEN, M. A. 1991. Improved Value for the Silicon Intrinsic Carrier Concentration from 275-K to 375-K. *Journal of Applied Physics,* 70**,** 846-854.

STAROVEROV, V. N., SCUSERIA, G. E., TAO, J. M. & PERDEW, J. P. 2003. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *Journal of Chemical Physics,* 119**,** 12129-12137.

STARR, C., EVERS, C. & STARR, L. 2010. *Biology concepts and applications*.

STEPHENS, P. J., DEVLIN, F. J., CHABALOWSKI, C. F. & FRISCH, M. J. 1994. Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields. *Journal of Physical Chemistry,* 98**,** 11623-11627.

STREHLOW, W. H. & COOK, E. L. 1973. Compilation of Energy Band Gaps in Elemental and Binary Compound Semiconductors and Insulators. *J. Phys. Chem,* 2.

SUDARSHAN, E., RAU, J. & MATHEWS, P. M. 1961. Stochastic Dynamics of Quantum-Mechanical Systems. *Physical Review,* 121**,** 920-&.

TAKAHASHI, K., YOSHIKAWA, A. & SANDHU, A. 2007. *Wide bandgap semiconductors : fundamental properties and modern photonic and electronic devices,* Berlin ; New York, Springer.

TAO, J. M., PERDEW, J. P., STAROVEROV, V. N. & SCUSERIA, G. E. 2003. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Physical Review Letters,* 91.

THIJSSEN, J. M. 2007. *Computational physics,* Cambridge, UK ; New York, Cambridge University Press.

THOMAS, D. G. & LANDER, J. J. 1956. Hydrogen as a Donor in Zinc Oxide. *Journal of Chemical Physics,* 25**,** 1136-1142.

THOMAS, L. H. 1927. The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society,* 23**,** 542-548.

TOLBERT, L. M., OZPINECI, B., ISLAM, S. K. & CHINTHAVALI, M. S. 2003. Wide bandgap semiconductors for utility applications. *Power and Energy Systems, Proceedings***,** 317-321.

UCAR. 2012. *Molecules Vibrate* [Online]. Available: https://spark.ucar.edu/molecular-vibration-modes [Accessed 15.10.2013.

VAN DE WALLE, C. G. 1997. Doping of wide-band-gap II-VI compounds - Theory. *Ii-Vi Blue/Green Light Emitters : Device Physics and Epitaxial Growth,* 44**,** 121-162.

VAN DE WALLE, C. G. 2000. Hydrogen as a cause of doping in zinc oxide. *Physical Review Letters,* 85**,** 1012-1015.

VANDERBILT, D. 1990. Soft Self-Consistent Pseudopotentials in a Generalized Eigenvalue Formalism. *Physical Review B,* 41**,** 7892-7895.

VIDYA, R., RAVINDRAN, P. & FJELLVAG, H. 2012. Ab-initio studies on Li doping, Li-pairs, and complexes between Li and intrinsic defects in ZnO. *Journal of Applied Physics,* 111.

VINES, L., MONAKHOV, E. V., SCHIFANO, R., MTANGI, W., AURET, F. D. & SVENSSON, B. G. 2010. Lithium and electrical properties of ZnO. *Journal of Applied Physics,* 107.

VON BARTH, U. & GELATT, C. D. 1980. Validity of the Frozen-Core Approximation and Pseudopotential Theory for Cohesive Energy Calculations. *Physical Review B,* 21**,** 2222-2228.

WAKANO, T., FUJIMURA, N., MORINAGA, Y., ABE, N., ASHIDA, A. & ITO, T. 2001. Magnetic and magneto-transport properties of ZnO : Ni films. *Physica E,* 10**,** 260-264.

WARDLE, M. G., GOSS, J. P. & BRIDDON, P. R. 2005a. Theory of Fe, Co, Ni, Cu, and their complexes with hydrogen in ZnO. *Physical Review B,* 72.

WARDLE, M. G., GOSS, J. P. & BRIDDON, P. R. 2005b. Theory of Li in ZnO: A limitation for Li-based p-type doping. *Physical Review B,* 71.

WHEELER, R. A. 2011. *Annual reports in computational chemistry,* London ;, Elsevier Science.

WIKIPEDIA.ORG. *Pseudopotential* [Online]. Available: http://en.wikipedia.org/wiki/Pseudopotential [Accessed 15.09.2013.

WILSON, S. 2003. *Handbook of molecular physics and quantum chemistry,* Chichester, Wiley.

WU, Z. & COHEN, R. E. 2006. More accurate generalized gradient approximation for solids. *Physical Review B,* 73**,** 235116.

YAO, T. & HONG, S.-K. 2009. *Oxide and nitride semiconductors : processing, properties, and applications,* Berlin, Springer.

YIN, Z. G., CHEN, N. F., YANG, F., SONG, S. L., CHAI, C. L., ZHONG, J., QIAN, H. J. & IBRAHIM, K. 2005. Structural, magnetic properties and photoemission study of Ni-doped ZnO. *Solid State Communications,* 135**,** 430-433.

ZHEREBETSKYY, D. 2010. *Quantum Mechanical First Principles Calculations of the Electronic and Magnetic Structure of Fe-Bearing Rock-Forming Silicates*, Universal Publishers.

ZINC.ORG. 2011. Available: www.zinc.org [Accessed 21.10.2013.

8. Appendix

8.1 Proof of the Hohenberg and Kohn theorems

The derivations of the proofs of the Hohenberg and Kohn theorems are performed in accordance with ref. (Martin, 2004) ch. 6.

Theorem 1: The external potential $V_{en}(\boldsymbol{r})$, and hence the total energy, and all the ground state properties, is a unique functional of the electronic ground state density, $n_0(\boldsymbol{r})$.

Proof of theorem 1: Let us assume that two different external potential, $V_{en}(\boldsymbol{r})$ and $V'_{en}(\boldsymbol{r})$, give rise to the same ground state density, $n_0(\boldsymbol{r})$. From the external potentials one can find the associated Hamiltonians, H and H', respectively, which have two different corresponding ground state wavefunctions, $\psi_0(\boldsymbol{r})$ and $\psi'_0(\boldsymbol{r})$. By that, the eigenvalues (the ground state energies) of the two Hamiltonians can be found by the innerproducts and the variational principle.

$$E_0 = \langle \psi_0|H|\psi_0\rangle < \langle\psi'_0|H|\psi'_0\rangle$$

$$E'_0 = \langle\psi'_0|H'|\psi'_0\rangle < \langle \psi_0|H'|\psi_0\rangle$$

The electronic Hamiltonian can be written in three terms, $H = T + U_{ee} + U_{en}$, which represents the kinetic energy, and the potential energy between the electrons and the electrons and the nuclei, respectively. From this one can express E_0 by

$$\begin{aligned}
E_0 &= \langle \psi_0|H|\psi_0\rangle \\
&= \langle \psi_0|T + U_{ee} + U_{en}|\psi_0\rangle \\
&= \langle \psi_0|T + U_{ee}|\psi_0\rangle + \int \psi_0^*(\boldsymbol{r})V_{en}(\boldsymbol{r})\psi_0(\boldsymbol{r})d\boldsymbol{r} \\
&= \langle \psi_0|T + U_{ee}|\psi_0\rangle + \int V_{en}(\boldsymbol{r})n_0(\boldsymbol{r})d\boldsymbol{r} \\
&< \langle\psi'_0|H|\psi'_0\rangle \\
&= \langle\psi'_0|T + U_{ee} + U_{en}|\psi'_0\rangle \\
&= \langle\psi'_0|T + U_{ee} + U'_{en} + (U_{en} - U'_{en})|\psi'_0\rangle \\
&= \langle\psi'_0|T + U_{ee} + U'_{en}|\psi'_0\rangle + \int \big(V_{en}(\boldsymbol{r}) - V'_{en}(\boldsymbol{r})\big)n_0(\boldsymbol{r})d\boldsymbol{r} \\
&= \langle\psi'_0|H'|\psi'_0\rangle + \int \big(V_{en}(\boldsymbol{r}) - V'_{en}(\boldsymbol{r})\big)n_0(\boldsymbol{r})d\boldsymbol{r} \\
&= E'_0 + \int \big(V_{en}(\boldsymbol{r}) - V'_{en}(\boldsymbol{r})\big)n_0(\boldsymbol{r})d\boldsymbol{r} \quad .
\end{aligned}$$

Thus,

$$E_0 < E'_0 + \int \big(V_{en}(\boldsymbol{r}) - V'_{en}(\boldsymbol{r})\big)n_0(\boldsymbol{r})d\boldsymbol{r} \quad .$$

By doing the same derivation based on the equation for E'_0, one obtains

$$E'_0 < E_0 + \int \big(V'_{en}(\boldsymbol{r}) - V_{en}(\boldsymbol{r})\big)n_0(\boldsymbol{r})d\boldsymbol{r} \quad .$$

By adding the left and right hand side in the two previous equations one obtains

$$E_0 + E_0' < E_0' + E_0 \quad ,$$

which obviously is a contradiction since the two sides are equal. Therefore the external potential can uniquely (except an additive constant) be determined by the electronic ground state density, $n_0(\boldsymbol{r})$.

Theorem 2: The ground state energy can be obtained by the variational principle. The density that minimizes the total energy is the exact ground state energy, $n_0(\boldsymbol{r})$. If $n(\boldsymbol{r})$ is not the ground state density, then $E[n_0(\boldsymbol{r})] < E[n(\boldsymbol{r})]$.

Proof of theorem 2: Since the first theorem states that the external potential, and hence the total energy is uniquely determined by the ground state density of the system, one can write the energy terms as functionals of the density. Since the Hamiltonian consists of three terms, one can write the three energy terms as functionals of the density

$$E[n(\boldsymbol{r})] = T[n(\boldsymbol{r})] + U_{ee}[n(\boldsymbol{r})] + U_{en}[n(\boldsymbol{r})] \quad .$$

Here, $U_{en}[n(\boldsymbol{r})] = \int V_{en}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r}$, as shown earlier. Substituting the kinetic term and the electron-electron interactions with the Hohenberg-Kohn universal functional, $F[n(\boldsymbol{r})]$, one gets

$$E[n(\boldsymbol{r})] = F[n(\boldsymbol{r})] + \int V_{en}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r}$$

The first theorem states that the ground-state density, $n_0(\boldsymbol{r})$ gives the potential term, $V_{en}(\boldsymbol{r})$, which in turn can give the Hamiltonian, and also the corresponding ground-state wavefunction, $\psi_0(\boldsymbol{r})$. Another density corresponds to another wavefunction $\psi(\boldsymbol{r})$ which is not the ground state. The expectation value of the Hamiltonian in the state $\psi(\boldsymbol{r})$ will therefore give a higher energy corresponding to the variational principle,

$$E_0 = E[n_0(\boldsymbol{r})] = \langle \psi_0 | H | \psi_0 \rangle < \langle \psi | H | \psi \rangle = E[n(\boldsymbol{r})] \quad .$$

This proves the second theorem, so the ground-state density will give the lowest energy. From this the ground-state energy is found by minimizing the energy with respect to the density. There exist also other theories, i.e. Levy-Lieb theory, which finds the ground-state energy by minimizing the energy with respect to the many-electron wavefunctions.