

Dichtefunktionaltheoretische Untersuchungen zu Anti-Surfactants auf Galliumnitridoberflächen

vorgelegt von
Master of Science
Andréia Luísa da Rosa

von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
– Doctrix rerum naturalium –

genehmigte Dissertation

Promotionsausschuß:

Vorsitzender: Prof. Dr. Christian Thomsen

Berichter: Prof. Dr. Eckehard Schöll

Berichter: Dr. Jörg Neugebauer

Tag der wissenschaftliche Aussprache: 23. Januar 2003

Berlin 2003
D 83

Density-Functional Theory Calculations on Anti-Surfactants at Gallium Nitride Surfaces

Andréia Luísa da Rosa

Contents

Summary	iv
Zusammenfassung	vi
1 Introduction	1
1.1 Technological importance	1
1.2 Heteroepitaxy and growth	2
1.3 Dopants and surfactants	3
1.4 Aim of this work	4
1.5 Structure of this work	5
2 Theoretical Background	7
2.1 The Born-Oppenheimer Approximation	8
2.2 Density-functional Theory	10
2.2.1 The Hohenberg-Kohn theorems	10
2.2.2 The Kohn-Sham equations	11
2.2.3 The Local-Density Approximation	13
2.2.4 The Generalized Gradient Approximation	14
2.2.5 LDA/GGA limitations	15
2.3 Pseudopotential Method	16
2.3.1 Non-linear core correction	20
2.3.2 Reference atomic configuration	21
2.4 Plane-wave representation and k -points	22
2.5 Atomic forces	25
2.6 The fhiPP and fhi98md packages	25
3 Bulk Properties	27
3.1 Structural and thermodynamic properties	28
3.1.1 Structural and cohesive properties	29
3.1.2 Electronic properties	34
3.2 Ga-bulk	35
3.3 Si-bulk	37
3.4 Si ₃ N ₄ -bulk	37
3.5 N ₂ -molecule	40
3.6 Summary	41

4	Bare GaN (0001) surfaces	43
4.1	Surface reconstructions and relaxations	44
4.2	Surface energy	48
4.3	Electronic properties	53
4.3.1	Band structure	53
4.3.2	Ionization energy, electron affinity and work function	56
4.4	Summary	61
5	Si on GaN (0001) surfaces	63
5.1	Si doping in GaN-bulk	65
5.2	Si doping on GaN (0001) surfaces	67
5.2.1	MOCVD/MOVPE growth	67
5.2.2	MBE growth	68
5.3	Adsorption of Si at GaN (0001)	70
5.3.1	Si-adatom on Ga/N-terminated surfaces	70
5.3.2	Mixed surfaces	77
5.3.3	Construction of a surface phase diagram	79
5.3.4	Discussions and consequences for growth	82
5.4	Polarity inversion of GaN (0001) surfaces	84
5.4.1	Surface energy	85
5.4.2	Relaxations	86
5.4.3	Electronic structure	86
5.4.4	Electron counting model	88
5.5	Summary	89
6	Conclusions and outlook	91
A	Convergence Tests	95
A.1	GaN-bulk	95
A.2	Ga-bulk	96
A.3	N ₂ -molecule	97
A.4	Slab	98
	Acknowledgements	102
	Curriculum Vitae	103

Summary

Experimental results have shown that adsorption and incorporation of Si on GaN surfaces depend extremely on the growth technique and on the growth conditions. The aim of this work was to provide a deeper understanding of Si adsorption on the bare GaN (0001) surfaces and its effect on surface morphology employing density-functional theory within the pseudopotential approach. In addition, a systematic study of the structural and electronic properties of the most relevant bare GaN (0001) surfaces has been performed.

One important issue when theoretically studying bulk and surface properties is the choice of the appropriate methods and techniques to perform the calculations. In the beginning of this work, there was no systematic study of how the exchange-correlation functional (PBE, LDA) performs on the properties of GaN surfaces. We have shown that, as a rule, the structural properties are similarly described using both functionals. Also, surface states are not affected either qualitatively or quantitatively by the exchange-correlation functional.

Regarding the thermodynamic stability of the structures, under Ga-rich conditions the description using PBE is fully compatible with the LDA calculations. However, under N-rich conditions, PBE performs very differently from LDA. This behavior has been related to the severe underestimation of the GaN formation enthalpy when using PBE. Based on these results, we have chosen to apply LDA to investigate adsorption of Si on the GaN (0001) surfaces.

A systematic study involving more than 100 model structures showed that Si is stabilized at the surface only under extreme N-rich/Si-rich conditions. Under more Ga-rich/Si-rich conditions Si prefers to incorporate at subsurface sites. Based on these results, we have derived a surface phase diagram which includes all thermodynamically stable structures. The surface phase diagram allowed a direct comparison of our results with experiment.

The general trends are that under Ga-rich/Si-rich conditions surface segregation of Si does not occur and Si can be easily incorporated in GaN bulk. As a consequence for growth, we conclude that such surfaces are essentially free of Si in the top surface layer and topologically very similar to the bare GaN (0001) surface. Thus we expect no change in the growth morphology. Such conditions are typically realized in molecular beam epitaxy (MBE), where indeed Si (even at high concentrations) does not affect the surface morphology. On the other hand, under N-rich/Si-rich conditions (as realized in metalorganic chemical vapor deposition (MOCVD) and molecular vapor phase epitaxy (MOVPE) to achieve high Si-doping levels) our phase diagram predicts Si acts as an *anti-surfactant*, i.e., it enhanced the

roughness of the surface and promotes three-dimensional growth. This behavior is indeed observed in MOCVD.

In summary, our work allowed the interpretation of the puzzling behavior of Si adsorption on GaN (0001). We identified under which conditions Si acts as an anti-surfactant, when it has no effect on the surface morphology, when it can cause polarity inversion and how to improve the Si doping efficiency. Besides, our results gave contributions to a better understanding of the properties of the bare GaN (0001) surfaces.

Zusammenfassung

Experimentelle Untersuchungen haben gezeigt, dass die Adsorption und Inkorporation von Si auf GaN Oberflächen extrem stark von der Wachstumstechnik und den Wachstumsbedingungen abhängen. Ziel dieser Arbeit war das Verständnis der Si-Adsorption auf reinen GaN (0001) Oberflächen und der Einfluss auf die Oberflächen-Morphologie durch Anwendung der Dichte-Funktional-Theorie innerhalb des Pseudo-Potential-Ansatzes. Zusätzlich wurde eine systematische Studie der strukturellen und elektronischen Eigenschaften der wichtigsten reinen GaN (0001) Oberflächen durchgeführt.

Ein wichtiger Punkt bei einer theoretischen Betrachtung der Volumen- und Oberflächen- Eigenschaften ist die Wahl der geeigneten Methoden und Techniken zur Durchführung der Rechnungen. Zu Beginn dieser Arbeit gab es keine systematische Studie des Einflusses des Austausch-Korrelations-Funktionalen (PBE, LDA) auf die Eigenschaften von GaN-Oberflächen. Wir haben gezeigt, dass generell die strukturellen Eigenschaften mit beiden Funktionalen in gleicher Weise beschrieben werden. Desweiteren werden Oberflächen-Zustände weder qualitativ noch quantitativ vom Austausch-Korrelations-Funktional beeinflusst.

Im Hinblick auf thermodynamische Stabilität der Strukturen ist unter Ga-reichen Bedingungen die Beschreibung mit PBE vollständig kompatibel mit den LDA-Rechnungen. Unter N-reichen Bedingungen jedoch verhalten sich PBE und LDA sehr unterschiedlich. Dieses Verhalten konnte auf die starke Unterschätzung der GaN Formations-Enthalpie bei PBE zurückgeführt werden. Von diesen Ergebnissen ausgehend haben wir zur Untersuchung der Adsorption von Si auf GaN (0001) Oberflächen die Anwendung von LDA gewählt.

Eine systematische Studie unter Einbeziehung von mehr als 100 Modell-Strukturen hat gezeigt, dass Si auf der Oberfläche nur bei extrem N-reichen/Si-reichen Bedingungen stabilisiert ist. Unter mehr Ga-reichen/Si-reichen Bedingungen wird Si bevorzugt an Plätzen unterhalb der Oberfläche eingebaut. Basierend auf diesen Resultaten haben wir ein Oberflächen-Phasen-Diagramm entwickelt, dass alle thermodynamisch stabilen Strukturen enthält. Das Oberflächen-Phasen-Diagramm erlaubt einen direkten Vergleich unserer Ergebnisse mit dem Experiment.

Die generelle Tendenz ist, dass unter Ga-reichen/Si-reichen Bedingungen Oberflächen-Segregation von Si nicht auftritt und Si leicht in den Kristall eingebaut werden kann. Als Konsequenz für das Wachstum folgern wir, dass solche Oberflächen in der obersten Oberflächenlage im wesentlichen frei von Si sind und topologisch den reinen GaN-Oberflächen sehr ähneln. Demnach erwarten wir keine Änderung der Oberflächen-Morphologie. Solche Bedingungen sind typischerweise in der Molekular-

Strahl-Epitaxie (MBE) realisiert, wo tatsächlich Si (sogar in hohen Konzentrationen) die Oberflächen-Morphologie nicht beeinflusst. Andererseits sagt unser Oberflächen-Phasen-Diagramm voraus, dass unter N-reichen/Si-reichen Bedingungen (wie sie in metallorganischer Gas-Phasen-Abscheidung (MOCVD) und metallorganischer Gas-Phasen-Epitaxie (MOVPE) realisiert sind, um hohe Si-Dotierungen zu erreichen) Si als *anti-surfactant* agiert, d.h. es vergrößert die Rauigkeit der Oberfläche und fördert drei-dimensionales Wachstum. Dieser Effekt wird bei MOCVD/MOVPE tatsächlich beobachtet.

Zusammenfassend erlaubte unsere Arbeit die Interpretation des rätselhaften Verhaltens der Si-Adsorption auf GaN(0001). Wir stellten fest, unter welchen Bedingungen Si als *anti-surfactant* agiert, wann es keinen Einfluss auf die Oberflächen-Morphologie hat, wann es Polaritäts-Inversion verursachen kann und wie man die Effektivität der Si-Dotierung verbessern kann. Desweiteren trugen unsere Ergebnisse zu einem besseren Verständnis der Eigenschaften reiner GaN(0001) Oberflächen bei.

This work is dedicated
to my parents, Ciro and
Vanda, to my brothers Alan
and Alex and to my sister,
Alessandra.

Chapter 1

Introduction

1.1 Technological importance

The group-III nitrides (AlN, GaN and InN) represent an important trio of semiconductors because of their direct band gaps which span the range 1.9-6.2 eV, including the whole of the visible region and extending well out into the ultraviolet (UV) range. Most recent measurements have shown that the band gap of InN can be even smaller than 1.9 eV, as obtained in Ref. [1] (0.7 eV) and Ref. [2] (1.0 eV). The group-III nitrides form a complete series of ternary alloys which, in principle, makes available any band gap within this range and the fact that they also generate efficient luminescence has been the main driving force for their recent technological development. In particular, GaN has a band gap of 3.4 eV at room temperature (300 K) and has been by far the most intensively studied material among the group-III nitrides. A graph showing the dependence of the energy gap as a function of the lattice parameter for the group-III nitride and their alloys is shown in Fig. 1.1.

High brightness visible light-emitting diodes (LEDs) are now commercially available, a development which has transformed the market for LED-based full colour displays and which has opened the way to many other applications, such as in traffic lights and efficient low voltage, flat panel white light sources [3–5]. Continuously operating UV laser diodes have also been demonstrated in the laboratory, exciting tremendous interest for high-density optical storage systems, UV lithography and projection displays.

In a remarkably short space of time, the nitrides have therefore caught up with and, in some ways, surpassed the wide band gap II-VI (ZnS, ZnSe, MgSe) compounds as materials for short wavelength optoelectronic devices. The devices so far available are based on the binary compound GaN, together with the ternary alloys AlGaN and InGaN, containing modest amounts of Al or In.

The nitrides crystallize preferentially in the hexagonal wurtzite structure. However, the cubic, zinc blende form is known for all three compounds, and cubic GaN, which can be achieved when GaN is epitaxially grown on a substrate with cubic symmetry such as GaAs [6–8] or SiC [9, 10], has been the subject of sufficient work to merit a brief account in its own right. There is significant interest based on possible technological advantages, such as easier doping, easier cleaving (for laser

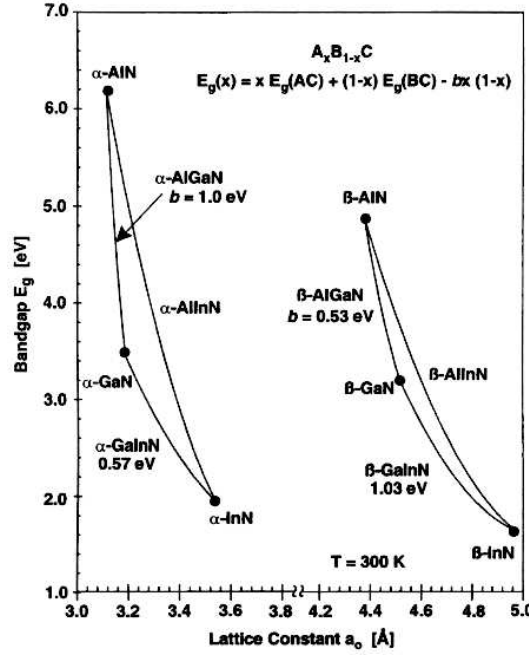


Figure 1.1: Band gap E_g and bowing parameters b of hexagonal (α -phase) and cubic (β -phase) InN, GaN, AlN and their alloys versus lattice constant a_0 . After O. Ambacher [13].

facets) and easier contacting. It also appears, at present, that the cubic form gives higher electron and hole mobilities than the hexagonal form. However, all the devices commercially produced until now were done using the hexagonal phase [5]. Therefore, in this thesis we will mainly focus on the hexagonal phase.

1.2 Heteroepitaxy and growth

The main problem in growing GaN thin films is the lack of suitable lattice matched substrates. Since the first synthesis of GaN has been successfully obtained in 1932 [11], much effort has been done in order to improve the quality of the grown material. However, at that time, neither metalorganic precursors containing In or Al with high purity nor substrate materials with reasonably good thermal properties and lattice matches to the nitrides were available. Nowadays MBE, MOCVD and MOVPE are the most used techniques.

In particular, single crystals of GaN for use as substrates for homoepitaxial growth have been produced [12], but they are still not available in large scale. Therefore the use of heteroepitaxial growth is unavoidable and the choice of a good substrate is essential.

The most popular substrates are sapphire (α - Al_2O_3) and SiC, due to their adequate thermal and chemical stability at growth temperatures, excellent structural and surface morphology. For device fabrication, sapphire has other advantages, such as the lower price and the availability of large crystals. SiC has attracted the inter-

est because of its good thermal conductivity and the ease of *n*- and *p*-type doping. However, SiC is still rather expensive.

The use of heteroepitaxial growth presents not only the disadvantages of lattice and thermal mismatch, but also the possibility of unintentional doping from the substrate during the growth. Measurements demonstrate that even stable compounds like Al₂O₃ and SiC are the source of large amounts of impurities such as O and Si [14, 15]. The contamination is enhanced by the presence of high concentrations of structural defects in GaN (such as dislocations, grain and twin boundaries) which provide efficient diffusion channels and gathering sites for the impurities [13, 16].

1.3 Dopants and surfactants

As-grown GaN shows a tendency to have *n*-type background carrier concentration, which can reach up to 10^{20} cm^{-3} . First studies invoked the spontaneous formation of N vacancies [17, 18] as source of the background concentration. Only recently Neugebauer and Van de Walle [19, 20] investigated the problem using first-principles calculations and showed that for *n*-type conditions the equilibrium concentration of N vacancies is too low to explain the unintentional doping. Thus, it has been concluded that the unintentional *n*-type conductivity must be attributed to the incorporation of donor impurities such as Si or O [19].

On the other hand, intentional doping of GaN is necessary to produce *n*- and *p*-type layers. Therefore an understanding of which elements can be efficiently used as dopants is crucial to obtain films and crystals with optimized properties. Atoms with tetrahedral radii close to the cation or anion would be most apt to substitute for these. However, a problem in doping a material with impurities is that compounds containing constituents of the host material and the dopant may be formed. Although this effect results in a high solubility of the impurity in the host material, the formation of such compounds will severely limit the doping efficiency.

Intentionally *n*-type doping of GaN-bulk has been always quite easy, although to achieve concentrations $>10^{19} \text{ cm}^{-3}$ is still a challenge. The impurity of choice for *n*-type doping is Si. For *p*-type doping Mg has been found to be the dopant with the best characteristics. However, highly efficient *p*-type doping has been difficult to achieve, since experimental investigations showed that the hole concentration is limited to 10^{18} cm^{-3} [21, 22]. The main doping limiting mechanism is the formation of Mg₂N₃, which is the solubility limiting phase. Alternative dopants such as Li, Na, K, Be, Zn and Ca have been theoretically investigated by Neugebauer and Van de Walle [23], but none of the candidates exhibited better characteristics than Mg. Only Be emerged as a potential alternative dopant, although it has limited solubility and the tendency to form interstitials under *p*-type conditions [24].

In order to fabricate devices with good performance, it is also important to understand the effect of doping on surfaces and interfaces. Recently, the use of surfactants has been used to modify the growth modes and consequently the surface properties of GaN. The basic idea of using the surfactant is the following: before growing a material A on a material B, a third material C, the surfactant, is deposited

on the substrate. Surfactants have a lower surface energy and floats on the surface without being incorporated into the subsurface layers. In particular, surfactant effects have been investigated for O [25], Mg [26,27], In [28] and Bi [29]. From the theoretical point of view, first-principles calculations have also been performed in order to predict or confirm how dopants affect the surface properties. For example, O [30], Mg [31] and In [32] have been investigated.

Surfactants do not necessarily need to be an impurity. The possibility of Ga itself acting as an *self-surfactant* on the GaN surfaces has also been recently explored [33–35]. Several works [33–36] have shown that by changing the Ga/N ratio and the substrate temperature it is possible to control the growth going from two-dimensional to three-dimensional mode and thus providing information under which conditions the material has a better morphology.

1.4 Aim of this work

Although Si is the most common impurity used as *n*-type dopant in GaN and its properties in GaN bulk have been extensively studied from both theoretical [20] and experimental [37, 38] point of view, little is known about its incorporation on GaN surfaces. From the experimental point of view, scanning tunneling microscopy (STM) measurements by Lee *et al.* [39, 40] demonstrated that the incorporation of Si leads to smooth surfaces under Ga-rich conditions.

On the other hand, it has been reported that Si-concentrations needed to achieve $10^{18} - 10^{19} \text{ cm}^{-3}$ carrier concentrations in GaN using MOCVD [41] and MOVPE [42] modify the GaN growth from a step-flow mode to a three-dimensional mode, i.e., Si acts as an *anti-surfactant*¹ and produces rough surfaces. This limits affects the achievable doping levels in GaN to $< 10^{19} \text{ cm}^{-3}$.

Indeed, MOCVD growth has shown that Si concentrations above $1 \times 10^{19} \text{ cm}^{-3}$ induce roughness [43] and crack formation [44]. Munkholm *et al.* [45] have reported that at high concentrations (above $2 \times 10^{19} \text{ cm}^{-3}$) Si segregates to the surface and changes the growth mode from step-flow to layer-by-layer over a large range of growth temperature (870-1170 K).

Thus, it is important to identify the mechanisms which cause the *anti-surfactant* behavior and under which conditions they occur. While the *anti-surfactant* behavior of Si is detrimental to achieving high *n*-type doping, it may be used to make structural modifications of the surface.

Therefore, in order to understand how Si affects microscopically the surface morphology and under which conditions the optimum growth can occur, we have performed density-functional theory (DFT) calculations for Si adsorbed on the GaN (0001) surfaces. In addition, to gain a better understanding of the bare GaN (0001) surfaces, we performed also an extensive study of the structural and electronic properties of the most relevant bare GaN (0001) surfaces.

¹An *anti-surfactant* has the opposite effect of a surfactant, i.e., it enhances the roughness of the surface, leading to three-dimensional growth.

1.5 Structure of this work

In Chapter 2 we start with a short introduction on DFT which allows us to accurately describe atomic geometry, electronic structure and energetics of the surfaces. Also, we describe the pseudopotential method, which will be used to solve the DFT equations.

The calculated structural and thermodynamic properties of both cubic (for a matter of comparison) and hexagonal phases of GaN, Si-bulk, Ga-bulk, Si₃N₄ and N₂-molecule are presented and discussed in Chapter 3 and compared with experimental and previous theoretical results.

In Chapter 4 we present and discuss results for structural properties of the most important bare GaN (0001) surfaces. To obtain a further understanding of these surfaces, a study of their electronic properties will also be performed.

In Chapter 5 the results concerning to adsorption of Si on GaN (0001) surfaces will be discussed. We will perform an analysis of Si adsorbed on the top layers and also at subsurface sites for various Si coverages in order to identify the thermodynamically most stable structures. Then, we make a connection between our results and the available experimental data, in order to understand the effect that Si adsorption has on the surface morphology of GaN (0001) surfaces.

Finally, mechanisms and consequences for growth of Si adsorption on GaN (0001) surfaces are summarized in Chapter 6.

Chapter 2

Theoretical Background

The systems we want to study, GaN bulk and surfaces, are many-body systems, i.e., involve many electrons and nuclei. In this chapter we will describe methods that allow us to model such complex systems. The Hamiltonian of a system consisting of N_e electrons and N_i nuclei is given by

$$\begin{aligned} H = & -\sum_{i=1}^{N_e} \frac{1}{2} \nabla_i^2 - \sum_{I=1}^{N_i} \frac{1}{2M_I} \nabla_I^2 - \sum_{i=1}^{N_e} \sum_{I=1}^{N_i} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \\ & + \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I=1}^{N_i} \sum_{J>I}^{N_i} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \end{aligned} \quad (2.1)$$

where M_I is the mass and Z_I is the atomic number of the nucleus I and \mathbf{r}_i and \mathbf{R}_I are the positions of the electron i and nucleus I , respectively. The first and second terms are the kinetic energy of the electrons and nuclei, respectively. The third term is the attractive Coulomb interaction between nuclei and electrons, the fourth term and fifth terms describe the repulsive Coulomb interaction between electrons and between nuclei, respectively. Here we use atomic units in all the equations, i.e. $\hbar = 1, m_e = 1, \frac{e^2}{4\pi\epsilon_0} = 1$.

The ground state of a system of N_e electrons and N_i nuclei is determined by solving the time-independent Schrödinger equation

$$H\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = E\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}), \quad (2.2)$$

where $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\})$ is the many-particle wave function of the system and E is the total energy of the system. Because of the complexity of the equation above, to make its solution feasible, some approximations must be done. One of the most used one in solid-state physics and also in atomic and molecular physics is the Born-Oppenheimer approximation, which decouples the electronic and nuclear movement. This approximation is described in the next section.

2.1 The Born-Oppenheimer Approximation

The Born-Oppenheimer approximation is based on the fact that the movement of the nuclei is much slower than the movement of the electrons, since the nuclei are much heavier than the electrons. Hence, a good assumption is that the N_e electrons move in the field of N_i fixed nuclei. Within this approximation, the second term of Eq. (2.1) can be neglected and the last term can be considered as a constant. Within the Born-Oppenheimer approximation, the wave function $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\})$ takes the form

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = \psi_e(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) \chi(\{\mathbf{R}_I\}), \quad (2.3)$$

where $\psi_e(\{\mathbf{r}_i\}; \{\mathbf{R}_I\})$ is the electronic wave function and $\chi(\{\mathbf{R}_I\})$ the nuclear wave function. The parametric dependence $\psi_e(\{\mathbf{r}_i\}; \{\mathbf{R}_I\})$ means that the nuclei are frozen in a single arrangement and the ground state of the electrons moving in this stationary potential is found. Thus, the Hamiltonian describing the motion of N_e electrons in the field of N_i nuclei is

$$H_e = - \sum_{i=1}^{N_e} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N_e} \sum_{I=1}^{N_i} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.4)$$

where the solution to a Schrödinger equation involving the electronic Hamiltonian H_e

$$H_e \psi_e(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) = E_e(\{\mathbf{R}_I\}) \psi_e(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) \quad (2.5)$$

is the electronic wave function $\psi_e(\{\mathbf{r}_i\}; \{\mathbf{R}_I\})$. The Hamiltonian of the nuclei moving in the average field of the electrons (term between brackets) is then

$$\begin{aligned} H_{\text{nuc}} &= - \sum_{I=1}^{N_i} \frac{1}{2M_I} \nabla_I^2 + \left\langle - \sum_{i=1}^{N_e} \nabla_i^2 - \sum_{i=1}^{N_e} \sum_{I=1}^{N_i} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right\rangle \\ &+ \sum_{I=1}^{N_i} \sum_{J>I}^{N_i} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &= - \sum_{I=1}^{N_i} \frac{1}{2M_I} \nabla_I^2 + E_e(\{\mathbf{R}_I\}) + \sum_{I=1}^{N_i} \sum_{J>I}^{N_i} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &= - \sum_{I=1}^{N_i} \frac{1}{2M_I} \nabla_I^2 + E_{\text{tot}}(\{\mathbf{R}_I\}). \end{aligned} \quad (2.6)$$

The total energy $E_{\text{tot}}(\{\mathbf{R}_I\})$ provides a curve for nuclear motion and also includes the nuclear repulsion

$$E_{\text{tot}}(\{\mathbf{R}_I\}) = E_e(\{\mathbf{R}_I\}) + \sum_{I=1}^{N_i} \sum_{J>I}^{N_i} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (2.7)$$

This function constitutes a potential energy surface. Thus the nuclei in the Born-Oppenheimer approximation move on a potential energy surface obtained by solving the electronic problem. Solutions to a nuclear Schrödinger equation describe the rotation, vibration and translation of the nuclei. The nuclear Schrödinger equation is given by

$$H_{\text{nuc}}\chi(\{\mathbf{R}_I\}) = E\chi(\{\mathbf{R}_I\}), \quad (2.8)$$

where E includes electronic, vibrational, rotational, and translational energy. In this work we do not consider the vibrational-rotational problem, but concentrate solely on the electronic problem.

In semiconductors, the use of the Born-Oppenheimer approximation is justified considering that the frequencies of ionic vibrations are typically less than 10^{13} s^{-1} . To estimate the electron response time, we note that the energy necessary to excite electrons in a semiconductor is given by its fundamental band gap. Taking as example a band gap of 1 eV, the frequencies of the electronic motion in semiconductors are of the order of 10^{15} s^{-1} [46]. As a result, electrons can respond to ionic motion almost instantaneously, which means that the ions are almost stationary to the electrons.

Although the Born-Oppenheimer approximation is broadly used, it is not always valid. A breakdown of the Born-Oppenheimer approximation occurs when the movement of the nuclei cannot be decoupled from the movement of the electrons. One example are highly excited rotational states of molecules, when the nuclear framework moves so fast that the electrons may be unable to follow this movement instantaneously [47].

The electronic structure described by the Eq. (2.5) is not easily solved in a reasonable computational time. Several different approaches have been used. For example, in quantum-chemistry a very common approach is the Hartree-Fock approximation. In the Hartree-Fock approximation the all-electron wave function is approximated by one Slater determinant that contains the single-electron wave functions. The Hartree-Fock method has the advantage of being variational and a trial wave function is assigned for the total wave function which minimizes the total energy. However, the Hartree-Fock method does not take into account the correlation between the electrons.

A very accurate method that accounts for the correlation is the Configuration-Interaction (CI) [48], which provides a systematic way of improving the Hartree-Fock wave function and energies by considering a combination of Slater determinants which includes the ground and excited states. However, due to the extremely large number of possible configurations, this method is very demanding and its applicability has been restricted to systems with only few electrons [49].

A statistical approach is the Quantum Monte Carlo method (QMC). This method can provide correlated wave functions, which provide rigorous upper bounds on the ground state energy. It is an explicit many-body method which takes electron correlation into account from the outset. It gives consistent, highly accurate results while at the same time exhibiting favourable scaling of computational cost with system size. This is in sharp contrast to the CI method. Although powerful, the use

of QMC has been greatly hampered over the last two decades by a combination of insufficient computer power and inefficient computer techniques. Only over the last few years many efforts have been done to speed up the method [50, 51].

In solid state physics, the state-of-the-art in first-principles calculations is the density-functional theory (DFT) [52, 53], which will be described in the next section. Since there are many text books and articles dedicated to this subject (see for example Refs. [49, 54, 55]), we will give here only a brief description.

2.2 Density-functional Theory

2.2.1 The Hohenberg-Kohn theorems

The history of the density-functional theory, which is mathematically based on two theorems of Hohenberg and Kohn [52], has its origins with the works of Thomas and Fermi in the 1920s [54, 56–59], which treats the many-body problem assuming the electron charge density $n(\mathbf{r})$ as basic variable. It has been originally developed for a system with non-degenerate ground-state, but was later extended for the degenerate case [55]. The two theorems of Hohenberg and Kohn are described in the following.

The first Hohenberg-Kohn theorem

Consider N_e electrons moving in an external potential $v(\mathbf{r})$ (for instance, electrons in a positive background of ions). The first Hohenberg-Kohn theorem [52] states that the external potential $v(\mathbf{r})$ is determined, within a trivial additive constant, by the electron density $n(\mathbf{r})$. Since $n(\mathbf{r})$ determines the number of electrons, it follows that $n(\mathbf{r})$ also determines the ground-state wave function and all other electronic properties of the ground state. The total energy functional $E_v[n]$ is then written as

$$E_v[n] = \int n(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{\text{HK}}[n], \quad (2.9)$$

where

$$F_{\text{HK}}[n] = T[n] + E_{\text{ee}}[n]. \quad (2.10)$$

Here $T[n]$ is the kinetic energy and $E_{\text{ee}}[n]$ is the energy due to the electron-electron interaction. Note that $F_{\text{HK}}[n]$ is defined independently of the external potential $v(\mathbf{r})$, which means that $F_{\text{HK}}[n]$ is an universal functional.

The second Hohenberg-Kohn theorem

The second Hohenberg-Kohn theorem states that the correct ground state electron density minimizes the total energy functional

$$E_v[n] = T[n] + E_{\text{ext}}[n] + E_{\text{ee}}[n] = F_{\text{HK}}[n] + E_{\text{ext}}[n], \quad (2.11)$$

where $E_{\text{ext}}[n]$ is the energy due to the external potential $v(\mathbf{r})$. If we denote $E_v[n_0]$ as the ground state energy and $n_0(\mathbf{r})$ as the ground state electron density, the second Hohenberg-Kohn theorem provides the variational principle. For a trial electron density $n(\mathbf{r})$, such that $n(\mathbf{r}) \geq 0$ and $\int n(\mathbf{r})d\mathbf{r} = N_e$, then

$$E_v[n] = \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + F_{\text{HK}}[n] \geq E_v[n_0] \quad (2.12)$$

with

$$\int v(\mathbf{r})n_0(\mathbf{r})d\mathbf{r} + F_{\text{HK}}[n_0(\mathbf{r})] = E_v[n_0]. \quad (2.13)$$

Assuming differentiability of $E_v[n]$, the variational principle requires that the ground-state density satisfies the stationary principle

$$\delta \left\{ E_v[n] - \mu \left[\int n(\mathbf{r})d\mathbf{r} - N_e \right] \right\} = 0 \quad (2.14)$$

which gives the Euler-Lagrange equation

$$\mu = \frac{\delta E_v[n]}{\delta n(\mathbf{r})} = v(\mathbf{r}) + \frac{\delta F_{\text{HK}}[n]}{\delta n(\mathbf{r})}, \quad (2.15)$$

where μ is the Lagrange multiplier.

If the exact form of the functional $F_{\text{HK}}[n]$ would be known, then Eq. (2.14) would be the exact equation for the ground-state electron density. Once we have an explicit form (approximate or accurate) for $F_{\text{HK}}[n]$, we can apply this method for any system. After the proposal of the density-functional theory, Kohn and Sham proposed an scheme to treat the variational problem of equation Eq. (2.14), which will be described in the next sections.

2.2.2 The Kohn-Sham equations

We have seen that the ground state energy of a many electron system can be obtained by minimizing the functional

$$E_v[n] = \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + F_{\text{HK}}[n] \quad (2.16)$$

with

$$F_{\text{HK}}[n] = T[n] + E_{\text{ee}}[n]. \quad (2.17)$$

In a trade for simplicity, Kohn and Sham [53] proposed an indirect approach to the kinetic functional $T[n]$. They proposed to introduce orbitals into the problem in a such a way that the kinetic energy can be computed simply to good accuracy. The functional $F_{\text{HK}}[n]$, which is unknown until this point, is rewritten as

$$F_{\text{HK}} = T_s[n] + J[n] + E_{\text{xc}}[n], \quad (2.18)$$

where $J[n]$ is the classical electrostatic repulsion (Hartree term)

$$J[n] = \frac{1}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (2.19)$$

$T_s[n]$ is the kinetic energy of a system of *non-interacting* electrons and $E_{xc}[n]$ is the so-called exchange-correlation energy. Thus, the energy functional $E_v[n]$ becomes

$$\begin{aligned} E_v[n] &= T_s[n] + J[n] + E_{xc}[n] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} \\ &= \sum_i^{N_{\text{occ}}} \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{r}) d\mathbf{r} + E_{xc}[n] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r}, \end{aligned} \quad (2.20)$$

where $\psi_i(\mathbf{r})$ are the Kohn-Sham orbitals and the sum runs over all the occupied states N_{occ} . According to the Kohn-Sham formalism, the Kohn-Sham orbitals should be orthonormal

$$\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}. \quad (2.21)$$

Now let us define the functional $\Omega[\{\psi_i\}]$ of N_{occ} orbitals

$$\Omega[\{\psi_i\}] = E_v[n] - \sum_i^{N_{\text{occ}}} \sum_j^{N_{\text{occ}}} \epsilon_{ij} \int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r}, \quad (2.22)$$

where ϵ_{ij} are the Lagrange multipliers for the constraints of Eq. (2.21). For $E_v[n]$ to be a minimum, it is necessary that

$$\delta\Omega[\{\psi_i\}] = 0, \quad (2.23)$$

which leads to the equations

$$h_{\text{eff}} \psi_i = \left[-\frac{1}{2} \nabla^2 + v_{\text{eff}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \sum_j^{N_{\text{occ}}} \epsilon_{ij} \psi_j(\mathbf{r}) \quad (2.24)$$

with $v_{\text{eff}}(\mathbf{r})$ determined through

$$\begin{aligned} v_{\text{eff}}(\mathbf{r}) &= v(\mathbf{r}) + \frac{\delta J[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \\ &= v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}) \\ &= v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}), \end{aligned} \quad (2.25)$$

where $v_{xc}(\mathbf{r})$ is the so-called exchange-correlation potential. Therefore, for a given $v_{\text{eff}}(\mathbf{r})$ one obtains $n(\mathbf{r})$ that satisfies the Eq. (2.15) by solving the one-electron equations

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (2.26)$$

and setting

$$n(\mathbf{r}) = \sum_i^{N_{\text{occ}}} |\psi_i(\mathbf{r})|^2. \quad (2.27)$$

The Eqs. (2.25), (2.26) and (2.27) are the so-called Kohn-Sham equations and need to be solved self-consistently. It follows then, that the total energy is calculated as

$$E_e = \sum_i^{N_{\text{occ}}} \epsilon_i - \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{\text{xc}}[n] - \int v_{\text{xc}}(\mathbf{r})n(\mathbf{r})d\mathbf{r}. \quad (2.28)$$

The Kohn-Sham equations are open for improvement with each better approximation to the exchange-correlation functional $F_{\text{HK}}[n]$, although there is no recipe to successively improve the functional $F_{\text{HK}}[n]$. As we will describe in the next sections, there are various approximations $F_{\text{HK}}[n]$.

2.2.3 The Local-Density Approximation

The Kohn-Sham equations, while exactly incorporating the kinetic energy $T_s[n]$, still leave the exchange-correlation functional $E_{\text{xc}}[n]$ unsettled. The most commonly used approximation to treat the exchange-correlation term within the density-functional theory is the local-density approximation (LDA). In this approximation, the exchange-correlation energy at each point of the space is replaced by that one for the homogeneous electron gas. The exchange-correlation energy in the local-density approximation is given by

$$E_{\text{xc}}[n] \approx E_{\text{xc}}^{\text{LDA}}[n] = \int n(\mathbf{r})\epsilon_{\text{xc}}^{\text{hom}}(n(\mathbf{r}))d\mathbf{r} \quad (2.29)$$

and therefore the potential by

$$v_{\text{xc}}^{\text{LDA}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}^{\text{LDA}}[n]}{\delta n(\mathbf{r})} = \epsilon_{\text{xc}}^{\text{hom}}(n(\mathbf{r})) + \frac{n(\mathbf{r})\partial\epsilon_{\text{xc}}^{\text{hom}}}{\partial n(\mathbf{r})}, \quad (2.30)$$

where $\epsilon_{\text{xc}}^{\text{hom}}n(\mathbf{r})$ is the exchange-correlation energy per particle of a homogeneous electron gas of density $n(\mathbf{r})$. Thus, the Kohn-Sham equations within the LDA takes the form

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{\text{xc}}^{\text{LDA}}(\mathbf{r}) \right] \psi_i = \epsilon_i \psi_i. \quad (2.31)$$

The function $\epsilon_{\text{xc}}^{\text{hom}}(n(\mathbf{r}))$ is then divided into exchange and correlation contributions

$$\epsilon_{\text{xc}}^{\text{hom}}(n(\mathbf{r})) = \epsilon_x^{\text{hom}}(n(\mathbf{r})) + \epsilon_c^{\text{hom}}(n(\mathbf{r})). \quad (2.32)$$

The exchange part is given by the exact Thomas-Fermi-Dirac exchange energy functional [55]

$$\epsilon_x^{\text{hom}}(n(\mathbf{r})) = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} n(\mathbf{r})^{1/3}. \quad (2.33)$$

Accurate values for the correlation part $\epsilon_c^{\text{hom}}(n(\mathbf{r}))$ are available, thanks to the Quantum Monte Carlo (QMC) calculation of Ceperley and Alder [60]. These values have been interpolated to provide an analytic form for $\epsilon_c^{\text{hom}}(n(\mathbf{r}))$. The most widely used has been constructed by Perdew and Zunger [61], which we will use in this work. The Perdew and Zunger [61] parameterization for spin-unpolarized systems¹ has the form

$$\begin{aligned} \epsilon_c^{\text{hom}} &= -0.1423(1 + 1.0529\sqrt{r_s} + 0.3334r_s)^{-1} & \text{if } r_s \geq 1 \\ &= -0.0480 + 0.0311\ln r_s - 0.0116r_s + 0.002r_s \ln r_s & \text{if } 0 \leq r_s \ll 1, \end{aligned} \quad (2.34)$$

where $r_s = (4\pi n(\mathbf{r})/3)^{-1/3}$ is the so-called Wigner-Seitz radius.

LDA has been the backbone of applications of density-functional theory. An approach to go beyond the LDA is to include gradient corrections in the electronic density, by making $E_{\text{xc}}[n]$ a functional of the density and its gradient, which is described in the next section.

2.2.4 The Generalized Gradient Approximation

A way of including the gradient in the expansion is to construct a functional that depends on the local-density $n(\mathbf{r})$ and on a gradient of the density $|\nabla n(\mathbf{r})|$. This approach is called generalized gradient approximation (GGA) and it has the general form

$$E_{\text{xc}}^{\text{GGA}}[n] = \int F_{\text{xc}}[n(\mathbf{r}), |\nabla n(\mathbf{r})|] d\mathbf{r}. \quad (2.35)$$

For the functional F_{xc} various forms have been proposed, such as Becke [62] (exchange part) and Perdew [63] (correlation part) (BP), Becke [62] (exchange part) and Lee-Yang-Parr [64] (BLYP), Perdew and Wang [65], Perdew and Vosko [66], and the revised PBE [67] (revPBE). Here we use the parameterization done by Perdew-Burke and Ernzerhof [68] (PBE).

In the GGA, the exchange-correlation energy is expressed in terms of an enhancement factor F_{xc} over the local exchange

$$E_{\text{xc}}^{\text{GGA}}[n] = \int n(\mathbf{r}) \epsilon_x^{\text{hom}} F_{\text{xc}}(r_s, s) d\mathbf{r}. \quad (2.36)$$

¹For spin-polarized systems (local spin-density density approximation (LSDA)) the charge density depends on the spin up $n_{\uparrow}(\mathbf{r})$ and spin down $n_{\downarrow}(\mathbf{r})$ densities, so that $n(\mathbf{r}) = n_{\uparrow}(\mathbf{r}) + n_{\downarrow}(\mathbf{r})$.

Here $s(\mathbf{r}) = \frac{|\nabla n|}{2k_F n}$ is a dimensionless density gradient, where $k_s = \sqrt{4k_F/\pi}$, $k_F = (3\pi^2 n)^{1/3}$. The exchange $F_x(s)$ term in the PBE parameterization is given by

$$F_x^{\text{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu s^2}{\kappa}}$$

$$\mu = \beta \left(\frac{\pi^2}{3} \right), \quad (2.37)$$

where $\kappa = 0.804$ and $\beta \simeq 0.066725$. The correlation part F_c is given by

$$F_c^{\text{PBE}}[n(\mathbf{r}), |\nabla n(\mathbf{r})|] = \int n(\mathbf{r}) [\epsilon_c^{\text{hom}}(n(\mathbf{r})) + H_c^{\text{PBE}}(r_s, t)] d\mathbf{r}, \text{ with} \quad (2.38)$$

$$H_c^{\text{PBE}}(r_s, t) = \gamma \ln \left[1 + \frac{\beta}{\gamma} t^2 \left(\frac{1 + At^2}{1 + At^2 + A^2 t^4} \right) \right], \quad (2.39)$$

$$A = \frac{\beta}{\gamma} \frac{1}{e^{-\epsilon_c^{\text{hom}}/\gamma} - 1}, \quad (2.40)$$

$$\gamma = \frac{1 - \ln 2}{\pi^2}, \quad (2.41)$$

and $t = \frac{|\nabla n|}{2k_s n}$ is another dimensionless density gradient.

2.2.5 LDA/GGA limitations

LDA is broadly used, mainly because of its ability to reproduce experimental ground state properties (bond lengths, cohesive energies) of many systems, although there is a tendency to overestimate binding energies and underestimate lattice constants of solids. LDA can successfully deal with atoms, molecules, clusters, surfaces and interfaces. Even for dynamical processes like the phonon dispersion, it has been shown to yield good results [69, 70].

However, in the course of time, many systems have been found that are incorrectly described by LDA. The most popular examples of this class are dielectric constants and related quantities, as well as weak bonds, in particular hydrogen bonds [71]. In the field of metals, the ground state structure of crystalline iron is predicted to be paramagnetic face-centered cubic instead of ferromagnetic body-centered cubic [72], which is correctly predicted using GGA.

For semiconductors it is well known that LDA/GGA underestimates the band gap. It has been claimed that this problem lies on the DFT itself, a ground-state theory. However, it is an open question if DFT would be solved exactly, it would

provide the correct band gap. Ways of remedy the band gap problem are exact-exchange and time-dependent density-functional theory (TD-DFT) (see for example Refs. [73] and [74]) and green-function methods (GW), but those methods are computationally expensive when treating solids.

In the past years several gradient-correction functionals were introduced and their properties were studied. In particular the description of the hydrogen bond is well achieved using BLYP [64, 75], as discussed in Ref. [76]. The general behavior of GGA is that the bond lengths in molecules and solids tend to be slightly too large compared to the experimental values and consequently the bulk moduli tend to be too small. The reaction barriers are better described with GGA [77, 78] and the surface energies are often lower than the ones described with LDA, probably due to the weaker bonds provided by GGA.

Until now, no systematic study of how LDA and GGA performs for GaN in the wurtzite structure and GaN surfaces was done. Therefore, we calculate various bulk and surface properties in order to test the performance of the LDA compared to GGA.

2.3 Pseudopotential Method

The Kohn Sham equations (2.25), (2.26) and (2.27) can be solved expanding the Kohn-Sham orbitals in a complete set of known basis functions (atomic orbitals, gaussians, plane-waves). Here we use plane waves, as it will be described in Sec. 2.4. When describing a periodic system, plane waves have invaluable numerical advantages, besides their conceptual simplicity. They allow a simple integration of the Poisson equation for the calculation of the Hartree potential and for the calculation of the kinetic energy expression.

However, due to the large oscillations of the core orbitals in the neighborhood of the atoms, plane waves cannot be used directly in the Kohn Sham formalism. These oscillations would require an enormous basis set size to be described with acceptable resolution.

However, the total energies associated with the core orbitals are several orders of magnitude larger than those of the valence band wavefunctions. Further, chemical reactions involve exclusively the valence electrons which are relatively far away from the nuclei. In contrast to this, the core electrons remain almost unaffected by the chemical bonding situation. They can be approximated to be frozen in their core configurations. This approximation considerably simplifies the task of solving the Kohn Sham equations, by eliminating all the degrees of freedom related to the core orbitals.

In the pseudopotential framework, the chemically inert core electrons are eliminated by considering them together with the nuclei as rigid ions. The electrostatic and quantum-mechanical interactions between the valence electrons and the ions and core electrons are described via pseudopotentials. Those pseudopotentials reproduce the true potential outside a defined core region, but are much smoother inside. The idea of the pseudopotential is to eliminate the nodes of the core wave functions,

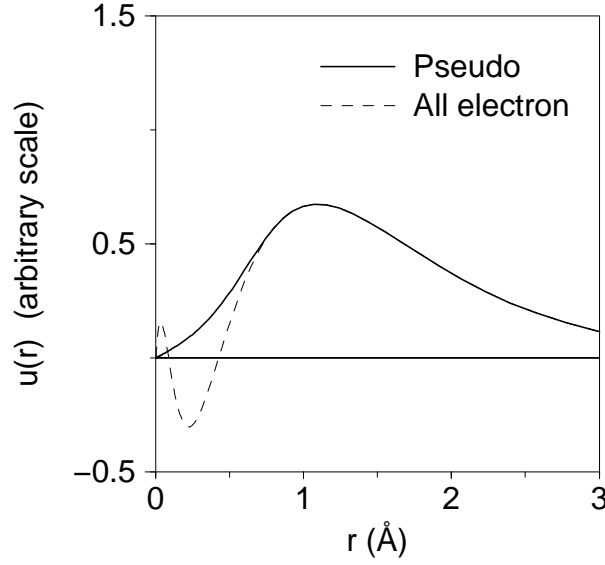


Figure 2.1: An example of the pseudo (solid line) and all electron (dashed line) wave functions for the 3s orbital of Al.

which strongly oscillates near the core region, and replace them by node-free wave functions, as shown in Fig. 2.1 for the 3s orbital of Al.

In the present work, we use pseudopotentials constructed within the Troullier-Martin scheme [79, 80]. Those pseudopotentials are constructed from atomic all-electron calculations. Within density-functional theory, the construction of such pseudopotentials is done by assuming a spherical screening approximation and solving self-consistently the radial Kohn-Sham equation

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + v_{\text{eff}}[n; r] \right] r R_{nl}(r) = \epsilon_{nl} r R_{nl}(r), \quad (2.42)$$

where $v_{\text{eff}}[n; r]$ is the self-consistent one-electron potential

$$v_{\text{eff}}[n; r] = -\frac{Z}{r} + v_{\text{H}}[n; r] + v_{\text{xc}}[n; r]. \quad (2.43)$$

Here $v_{\text{H}}[n; r]$ is the Hartree potential and $v_{\text{xc}}[n; r]$ is the exchange-correlation potential. $r R_{nl}(r) = u_l^{\text{ps}}$ is the radial part of the pseudopotential wavefunction. Most pseudopotentials are then constructed such that they satisfy four general conditions:

- (i) The valence pseudo wave functions generated from the pseudopotential should contain no nodes. This stems from the fact that we want to construct smooth pseudo wave functions and therefore nodes are not desirable.
- (ii) The normalized atomic radial pseudo wave function with angular momentum l is equal to the normalized radial all-electron wave function u_{nl}^{AE} beyond a chosen cutoff radius r_l^c , which divide the core ($r < r_l^c$) and valence ($r > r_l^c$) regions

$$u_l^{\text{ps}}(r) = u_{nl}^{\text{AE}}(r) \quad \text{for} \quad r > r_l^c. \quad (2.44)$$

This spatial separation of the region allows us to choose which electrons are included in the core region and which electrons are included in the valence region, controlling this way the number of plane waves used to described the system.

- (iii) To ensure the correct electrostatic description of the atom in various environments, the charge enclosed inside the core region should be the same for the all-electron and pseudopotential approximation. Thus, the following condition should be satisfied

$$\int_{r=0}^{r_l^c} |u_l^{\text{ps}}(r)|^2 dr = \int_{r=0}^{r_l^c} |u_{nl}^{\text{AE}}(r)|^2 dr \quad . \quad (2.45)$$

- (iv) The all-electron and pseudopotential eigenvalues must be equal

$$\epsilon_l^{\text{ps}} = \epsilon_{nl}^{\text{AE}}. \quad (2.46)$$

If a pseudopotential satisfies the conditions (i)-(iv), then it is called norm-conserving pseudopotential [81]. These conditions ensure that the electrostatic potential felt by the pseudo valence function outside the core region ($r > r_l^c$) is the same as the all-electron one. As we mentioned above, the pseudo wave function should not contain radial nodes. In order to obtain a continuous pseudopotential which does not diverge at the origin, it should satisfy $\lim_{r \rightarrow 0} u_l^{\text{ps}}(r) \propto r^{l+1}$. The pseudopotential components then correspond to the inversion of the Schrödinger equation for the respective wave function

$$V_l^{\text{ps,scr}}(r) = \epsilon_l^{\text{ps}} - \frac{l(l+1)}{2r^2} + \frac{1}{2u_l^{\text{ps}}(r)} \frac{d^2}{dr^2} u_l^{\text{ps}}(r). \quad (2.47)$$

The final ionic pseudopotentials are determined by subtracting from the screened pseudopotential $V_l^{\text{ps,scr}}(r)$ the electrostatic and exchange-correlation screening contributions due to the valence electrons

$$V_l^{\text{ps}}(r) = V_l^{\text{ps,scr}}(r) - v_{\text{H}}[n^{\text{ps}}; r] - v^{\text{xc}}[n^{\text{ps}}; r], \quad (2.48)$$

with the pseudo charge density

$$n^{\text{ps}}(r) = \frac{1}{4\pi} \sum_{l=0}^{l_{\text{max}}} f_l \left| \frac{u_l^{\text{ps}}(\epsilon_l^{\text{ps}}; r)}{r} \right|^2. \quad (2.49)$$

Here f_l are the occupancy numbers. Due to the spherical symmetry, states with the same quantum numbers $n_i = n$ and $l_i = l$, but different m_i are energetically degenerate, $\epsilon_i = \epsilon_{n_i l_i m_i} = \epsilon_{nl}$. Thus, it holds that $f_l = 2(2l+1)$ for $\epsilon_l < \epsilon_{N_e}$, $0 < f_l < 2(2l+1)$ for $\epsilon_l = \epsilon_{N_e}$ and $f_l = 0$ for $\epsilon_l > \epsilon_{N_e}$. All pseudopotential

components $V_l^{\text{ps}}(r)$ reduce to the ionic Coulomb potential at large distances r , i.e., they get independent from l . Therefore it is possible to express the pseudopotential as a local potential part V_{local} plus only a few l dependent components for $l \leq l_{\text{max}}$, of the type

$$\begin{aligned} \langle \mathbf{r} | \hat{V}_{\text{ps}} | \mathbf{r}' \rangle &= \langle \mathbf{r} | \hat{V}^{\text{local}} + \delta \hat{V}^{\text{semi-local}} | \mathbf{r}' \rangle \\ &= V_l^{\text{local}}(r) \delta(\mathbf{r} - \mathbf{r}') + \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^l Y_{lm}^*(\Omega_{\mathbf{r}}) \delta V_l^{\text{semi-local}}(r) \frac{\delta(r - r')}{r^2} Y_{lm}(\Omega_{\mathbf{r}'}) \end{aligned} \quad (2.50)$$

where the local potential is taken as one of the semi-local pseudopotential components and $\delta V_l^{\text{semi-local}}(r) = V_l^{\text{ps}}(r) - V_l^{\text{local}}(r)$ vanishes beyond r_l^c . The local potential can be in principle arbitrarily chosen. However, since the summation in l will be truncated at some l_{max} , l should be chosen in such a way to reproduce the scattering properties for all the higher angular momentum channels.

The computational cost can be dramatically reduced if the above form is transferred into the fully separable form of Kleinman-Bylander [82]:

$$\begin{aligned} \langle \mathbf{r} | \hat{V}_{\text{ps}} | \mathbf{r}' \rangle &= \langle \mathbf{r} | \hat{V}^{\text{local}} + \delta V^{\text{KB}} | \mathbf{r}' \rangle \\ &= V^{\text{local}}(r) \delta(\mathbf{r} - \mathbf{r}') + \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^l \langle \mathbf{r} | \chi_{lm}^{\text{KB}} \rangle E_l^{\text{KB}} \langle \chi_{lm}^{\text{KB}} | \mathbf{r}' \rangle. \end{aligned} \quad (2.51)$$

Here, the short-ranged second term is a fully nonlocal operator, where $\langle \mathbf{r} | \chi_{lm}^{\text{KB}} \rangle$ are the projector functions and E_l^{KB} is the Kleinman-Bylander energy [82, 83]. A problem with this form is that it may lead to additional (ghost) states that are unphysical [83]. Eigenstates below the real ground state can occur, if the choice of the local component of the pseudopotential is not done properly. The presence of *ghost* states affects the transferability of the pseudopotential when comparing the results with all-electron calculations.

In the reference configuration in which it was generated, the pseudopotential should accurately reproduce the all electron eigenvalues outside the core region compared to all electron calculations. In practice we want it to reproduce other all-electron calculations in different environments, that is we want it to be “transferable”. The logarithmic derivatives provides then a first test of the transferability of the screened pseudopotential.

A possible method to improve the transferability to the solid is to generate the pseudopotential using an atomic configuration that as closely as possible mimics the environment in which will be placed. This may require the use of non-bound eigenvalues, as suggested by Hamann [81]. In general, ionic pseudopotentials are insensitive to reasonable variations in the reference atomic configurations. Shirley *et al.* [84] proposed equating additional energy derivatives of the logarithmic derivatives of the wave function, but in general it has been found that the errors are negligible

[79]. The easiest way of improving the transferability is to decrease the cutoff radius r_l^c , reducing the difference between the all electron and pseudopotential results.

However, there are practical limits on how far r_l^c can be decreased, since it has to be larger than the outermost node of the all-electron wave function if we insist on having nodeless pseudo wave functions. Also, decreasing r_l^c makes the pseudopotential computationally more expensive. Therefore, according to the considerations above we use the ground state atomic configuration to generate all our pseudopotentials.

2.3.1 Non-linear core correction

As mentioned in the last section, the final ionic pseudopotentials are determined by subtracting from the screened pseudopotential above the electrostatic and exchange-correlation screening contributions due to the valence electrons

$$V_l^{\text{ps}}(r) = V_l^{\text{ps,scr}}(r) - v_{\text{H}}[n^{\text{ps}}; r] - v_{\text{xc}}[n^{\text{ps}}; r], \quad (2.52)$$

with

$$n^{\text{ps}}(r) = \frac{1}{4\pi} \sum_{l=0}^{l_{\text{max}}} f_l \left| \frac{u_l^{\text{ps}}(\epsilon_l^{\text{ps}}; r)}{r} \right|^2. \quad (2.53)$$

Here $v^{\text{xc}}[n^{\text{ps}}; r]$ refers to the exchange-correlation interaction between the valence electrons themselves. The exchange-correlation interaction between the valence and the core electrons is included in the pseudopotential, as a term that depends linearly on the valence charge density $n^{\text{ps}}(r)$. Although E_{xc} is a non-linear functional of the total electron density $n(r)$, the above linearization of its core-valence contribution is a usual and mostly used approximation within LDA and GGA [85].

However, explicit account of the non-linearity is sometimes required. This is done by restoring the non-linear dependence of $E_{\text{xc}}[n]$ on the total electron density. Rather than the full core density, it suffices to add a partial core density $\tilde{n}^{\text{core}}(r)$, as suggested by Louie *et al.* [86]. It reproduces the full core density $n^{\text{core}}(r)$ outside a chosen cutoff radius r_{nlc} , but it is a smoother function inside. In this work we use the same idea, but a different function from Ref. [86], which is given by

$$\begin{aligned} \tilde{n}^{\text{core}}(r) &= n^{\text{core}}(r) & \text{for} & & r \geq r_{\text{nlc}} \\ &= c_0 + \sum_{i=3}^6 c_i r^i & \text{for} & & r \leq r_{\text{nlc}}, \end{aligned} \quad (2.54)$$

where we take the coefficients c_i such that $\tilde{n}_0^{\text{core}}(r)$ has zero slope and curvature at the origin, decays monotonically, and joins the full core density up to the third derivative. The resulting nonlinear core-valence exchange-correlation scheme uses the redefined ionic pseudopotentials

$$V_l^{\text{ps}}(r) = V_l^{\text{ps,scr}}(r) - v_{\text{H}}[n^{\text{ps}}; r] - v_{\text{xc}}[n^{\text{ps}} + \tilde{n}^{\text{core}}; r]. \quad (2.55)$$

As a rule, the non-linear exchange-correlation plays a significant role for alkali metals and with increasing atomic numbers, i.e., the farther the upper core orbitals extend into the tails of the valence density (for example Zn, Cd). Of course if the uppermost semi-core states hybridize with true valence states they have to be treated as valence states, like $3d$, $4d$ and $5d$ states of Ga and In in GaN and InN, where these interact significantly with the N- $2s$ states [87].

2.3.2 Reference atomic configuration

In this section we describe the atomic configuration used to generate our pseudopotentials. The reference configuration and cutoff radii are shown in Table 2.1. Those pseudopotentials have been successfully used to describe bulk properties of cubic phases of group-III nitrides [87]. For H with fractional charge we use the $-Z/r$ potential, where $Z = 5/4e^-$ is the electronic charge of the hydrogen used to passivate the incomplete bond of the N atom on the back side of the slab used to model our surfaces (see Sec. 2.4). In principle, the H pseudopotential has no immediate physical meaning. However, it has been used as a powerful tool to perform plane wave calculations with a much smaller number of plane waves [88].

It works for H_2 reasonably well, provided the pseudopotential cutoff radius is chosen not too large. In our case we need a significantly large energy cutoff to calculate the GaN properties of the bulk and surfaces. Thus, we did not generate a H pseudopotential and decide to use the $-Z/r$ potential, which is well described with our optimum cutoff energy. We perform calculations for the NH_3 and H_2 molecules and negligible differences in the cohesive properties and bond lengths were found comparing $-Z/r$ and pseudopotentials [89].

As exemplified in Table 2.1, for the N_2 molecule, which has a short bond length of 1.06 \AA , such an overlap is significant ($2r_s = 1.58 \text{ \AA} > d_{N-N}$) and affects the total energy and consequently the properties of the molecule. For GaN-bulk, the overlap occurs only for the p component, and is negligible (0.05 \AA). For Si-bulk this overlap is slightly larger ($\approx 0.2 \text{ \AA}$) than the Si-Si bonds and occurs for the p and d components. For Ga-bulk, there is no overlap for the Ga- $3d$ pseudopotential, but a small overlap for the p -component of the Ga- $nlcc$ pseudopotential (0.1 \AA) and relatively large for the d -component (0.5 \AA). Decreasing the cutoff radii eliminates the overlap, but then the number of plane waves required increases, making the calculations computationally much more expensive. Besides, experimental bond lengths have an error bar of $0.1\text{-}0.2 \text{ \AA}$. Thus we consider the overlaps mentioned above as acceptable.

Fuchs *et al.* [87] have shown that the inclusion of Ga- f orbitals (when used as the local component) slightly improves the energetic properties of GaN. However this improvement does not compensate the computational effort demanded. Thus, we decided to have d as the highest angular momentum component. Their explanation why local f orbitals perform better is that the f component is less repulsive than the s component in the core region. Also, we also perform calculations using $nlcc$ to compare how the Ga- $nlcc$ pseudopotential compares to the Ga- $3d$ pseudopotential. The corresponding ionic pseudopotentials described above are shown in Fig. 2.2.

Table 2.1: Atomic reference configuration for Ga, N, and Si pseudo atoms used in the present work. r_s , r_p and r_d are the cutoff radii for the s , p and d pseudopotential components, r_{nlc} is the cutoff radii used in the $nlcc$ approximation and d_{N-N} , d_{Ga-Ga} , d_{Ga-N} , d_{Si-Si} are the experimental first neighbors distances for N_2 -molecule, Ga-bulk, GaN-bulk, and Si-bulk. All units are in Angstroms.

atom	reference configuration	r_s	r_p	r_d	r_{nlc}	d_{N-N}	d_{Ga-Ga}	d_{Ga-N}	d_{Si-Si}
						1.06	2.44	1.95	2.34
Ga 3d	$3d^{10}4p^24s^1$	1.10	1.21	1.10					
Ga <i>nlcc</i>	$4s^24p^24d^0$	1.21	1.27	1.48	0.68				
N	$2s^22p^33d^0$	0.79	0.79	0.79					
Si	$3s^23p^23d^0$	1.04	1.27	1.27					

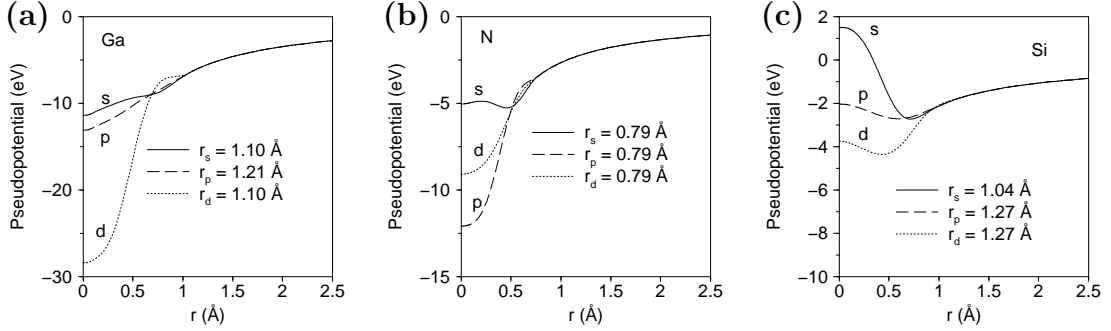


Figure 2.2: Ionic pseudopotentials for the elements (a) Ga (including the 3d electrons as valence electrons), (b) N and (c) Si using the configuration reference shown in Table 2.1.

2.4 Plane-wave representation and k-points

Density-functional theory allows the many-body problem to be replaced by a single particle problem. Nevertheless, it is impossible to solve a problem of 10^{23} particles moving in a certain effective potential. This problem can be overcome if we choose a minimum representation of the system we want to study, for instance the bulk unit cell for crystals and a slab to represent a surface.

The systems are treated periodically in all the three directions but perpendicular to the surface it is separated by a vacuum region to avoid interactions between the surfaces. For isolated systems, like atoms or molecules, we use large supercells to ensure that there is no interaction between isolated systems belonging to neighbouring super cells.

Since a plane wave is a periodic function, periodic boundary conditions [90] are a consequence. If there is an atom at the position \mathbf{R}_I , there are atoms also in $\mathbf{R}_I + n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$, where $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are the primitive lattice vectors of the supercell in the real space and n_1, n_2, n_3 are integers. Fig. 2.3 shows, for instance, a

repeated slab geometry which allows the description of surfaces. The Bloch theorem states that the solution for a one-electron Hamiltonian can be written as

$$\psi_{i\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{i\mathbf{k}}(\mathbf{r}), \quad (2.56)$$

where i is the band index, \mathbf{k} is a point in the first Brillouin Zone (BZ) and $u_{i\mathbf{k}}(\mathbf{r})$ is a function with the periodicity of the real space lattice so that

$$u_{i\mathbf{k}}(\mathbf{r} + \mathbf{R}_I) = u_{i\mathbf{k}}(\mathbf{r}). \quad (2.57)$$

As $u_{i\mathbf{k}}(\mathbf{r})$ is a periodic function, it can be expanded in Fourier components such that

$$u_{i\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{i,\mathbf{k}+\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (2.58)$$

where $V = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$. The set of \mathbf{G} vectors is limited to those which have the periodicity of the real space

$$\mathbf{G} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3, \quad (2.59)$$

with m_1, m_2, m_3 integers and the reciprocal lattice vectors are $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$. Thus, $\psi_{i\mathbf{k}}(\mathbf{r})$ can be written as

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{i,\mathbf{k}+\mathbf{G}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}. \quad (2.60)$$

Using the expression (2.60) for the Kohn-Sham orbitals, the Kohn-Sham equation (2.26) is written as

$$\sum_{\mathbf{G}'} \left[\frac{1}{2} |\mathbf{k} + \mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + v(\mathbf{G} - \mathbf{G}') + v_{\text{H}}(\mathbf{G} - \mathbf{G}') + v_{\text{xc}}(\mathbf{G} - \mathbf{G}') \right] C_{i,\mathbf{k}+\mathbf{G}'} = \epsilon_{i,\mathbf{k}} C_{i,\mathbf{k}+\mathbf{G}}. \quad (2.61)$$

In practice the expansion in plane-waves includes terms with kinetic energy up to a certain cutoff

$$E_{\text{cutoff}} \leq \frac{1}{2} |\mathbf{k} + \mathbf{G}|^2. \quad (2.62)$$

The quality of the basis set depends on how large E_{cutoff} is, i.e., on the number of plane waves. This number has to be checked for each system. Convergence tests with respect to the number of plane waves for the systems studied here can be found in Appendix A.

One consequence from the spatial periodicity within the supercell approach and the Bloch theorem is that all inequivalent values of the quasi-momentum \mathbf{k} lie within the BZ. In a trade for numerical simplicity, the integrals are replaced by weighted sums over a discrete set of \mathbf{k} -points, so that

$$V \sum_i \int_{\text{BZ}} f(\epsilon_{i\mathbf{k}} - \epsilon_{\text{F}}) |\psi_{i,\mathbf{k}}|^2 \frac{d\mathbf{k}}{(2\pi)^3} \rightarrow \frac{1}{N_{\mathbf{k}}} \sum_i \sum_{\mathbf{k}_i} w_{\mathbf{k}_i} f(\epsilon_{i,\mathbf{k}} - \epsilon_{\text{F}}) |\psi_{i,\mathbf{k}}|^2, \quad (2.63)$$

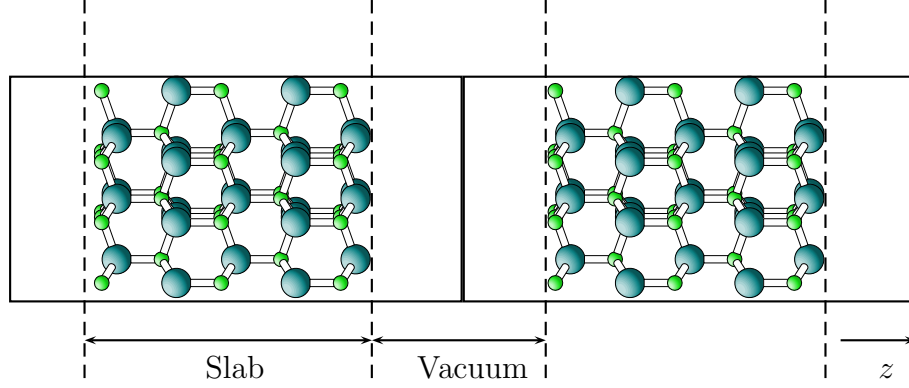


Figure 2.3: Schematic picture showing the slab model to study surfaces. The slab is periodic in all dimensions and separated along the z direction by a vacuum region.

where V is the volume of the super-cell, $w_{\mathbf{k}_i}$ is the weight of each \mathbf{k} -point which obeys $\sum_{\mathbf{k}} w_{\mathbf{k}_i} = 1$, and ϵ_F is the Fermi energy. The summation runs over all bands i and the occupancies have the step-like distribution

$$f(\epsilon_{i,\mathbf{k}} - \epsilon_F) = \frac{1}{e^{(\epsilon_{i,\mathbf{k}} - \epsilon_F)/k_B T} + 1}, \quad (2.64)$$

where k_B is the Boltzmann constant and T is an artificial temperature used to achieve a better convergence [91]. Choosing a sufficiently dense mesh of integration points is crucial for the convergence of the results, and is therefore one of the major objectives when performing convergence tests. Here it should be noted that there is no variational principle governing the convergence with respect to the \mathbf{k} -points mesh. This means that the total energy does not necessarily should decrease as the density of the \mathbf{k} -point mesh increases.

In this work we use the Monkhorst-Pack scheme [92] to construct the \mathbf{k} -point mesh inside the BZ. In practice one uses a few selected \mathbf{k} -points of the form

$$\mathbf{k}_k = k_1^{(k)} \mathbf{b}_1 + k_2^{(k)} \mathbf{b}_2 + k_3^{(k)} \mathbf{b}_3. \quad (2.65)$$

The three coefficients in Eq. (2.65) are given by

$$k_j^{(k)} = \frac{2p_j^{(k)} + 2q_j - l_j - 2}{2l_j}, \quad (2.66)$$

with $q_j = 1, 2, \dots, l_j$, $k = 1, 2, \dots, N_{\mathbf{k}}$.

The number of \mathbf{k} -points can be reduced if the crystal structure under investigation has additional point group symmetries (for instance rotations, mirror planes). In this case, the \mathbf{k} -points are chosen only to span the irreducible part of the Brillouin Zone (IBZ). As a result the $N_{\mathbf{k}}^r$ symmetry equivalent \mathbf{k} -points are only a subset of the $N_{\mathbf{k}}$ points. The sum is then taken only over the reduced set with corresponding weighting factors $w_{\mathbf{k}_i}$.

The number of \mathbf{k} -points used in the BZ depends on the system to be studied. In general systems with a gap need a few \mathbf{k} -points, while for metallic systems a

much denser \mathbf{k} -point mesh is required, to ensure a good description of the Fermi surface. The quality of the \mathbf{k} -point mesh should be tested for each system, so that a convergence of the total energy with respect to the number of \mathbf{k} -points is achieved. Such tests were performed in the present work and can be found in Appendix A.

2.5 Atomic forces

We are interest in achieving the lowest energy atomic configuration. This is in principle not known and we have to start from a certain guess geometry (arrangement of the atoms). During the atomic optimization, the forces \mathbf{F}_I on the atoms are calculated and use as a criterion to achieve the equilibrium configuration. In all structure relaxation runs performed in this work, geometries were considered converged if the residual forces were smaller than 0.025 eV/\AA . The Hellmann-Feynman theorem [93] states that the force on the ion I is minus the derivative of the total energy of the system with respect to the position of the ion

$$\mathbf{F}_I = -\frac{dE_{\text{tot}}}{d\mathbf{R}_I} = \sum_{J=1}^{N_i} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \frac{dE_e}{d\mathbf{R}_I}, \quad (2.67)$$

The electronic part consists of two terms, namely

$$\frac{dE_e}{d\mathbf{R}_I} = - \int \frac{\partial}{\partial \mathbf{R}_I} v(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + \int [v(\mathbf{r}) + v_H(\mathbf{r}) + v_{\text{xc}}(\mathbf{r})] \frac{\partial}{\partial \mathbf{R}_I} n(\mathbf{r}) d\mathbf{r}, \quad (2.68)$$

where the first term is called Hellmann-Feynman force [93] and the second term is the Pulay force [94]. It can be shown that once the Born-Oppenheimer surface is achieved, the Pulay forces are zero and only the Hellmann-Feynman force remains. As a criterion, all the structure relaxations in this work were performed until the forces on the atoms were smaller than $5 \times 10^{-4} \text{ Hartree/Bohr} \approx 0.025 \text{ eV/\AA}$.

2.6 The fhiPP and fhi98md packages

For all the calculations done in this work, we used the **fhiPP** and **fhi98md** packages, developed at the Fritz-Haber-Institute of the Max-Planck Society in Berlin. The **fhiPP** package was used to constructed the pseudopotentials used in this work. The **fhi98md** was used to perform the total energy calculations within the pseudopotential framework for all systems here considered. Informations about these programs and applications can be found in Refs. [91, 95, 96].

Chapter 3

Bulk Properties

Before starting surface calculations for any material, it is crucial to perform calculations for its bulk phase in order to determine the structural (e.g., lattice parameters, bulk modulus) and thermodynamic properties (e.g., cohesive energy, formation enthalpy). In particular, the lattice parameters of the bulk phase will be needed to perform the calculations of surfaces. Also, in the approach we use, the formation enthalpy is a quantity necessary to calculate the surface properties (surface energies). It is therefore assumed that the calculated properties are transferable and the properties determined using smaller systems like the bulk phase hold also for larger systems as surfaces.

As it was discussed in Chapter 2, there are different approximations for the exchange-correlation functional (LDA, GGA). In the recent years, several works were done to understand how these functionals perform for group-III nitride materials. Previous studies within the pseudopotential approach for group-III nitrides have been shown that the properties are sensitive to the choice of the pseudopotential [87, 97–99].

Two interpretations might come out from the pseudopotential calculations. One is that GGA might perform differently for group-III nitrides compared to the traditional semiconducting III-V materials (GaAs, AlAs, GaP). The other one is that the difference is not due to the exchange-correlation functional, but comes from the use of the pseudopotential approach.

To clarify this aspect, Fuchs *et al.* [87] carried out pseudopotential and all-electron Full-Potential Linearized Augmented Plane Waves (FP-LAPW) calculations (within LDA and GGA) for group-III nitrides. By comparing the FP-LAPW and pseudopotential results, it was possible to check the transferability of the pseudopotentials, and hence understand how good pseudopotentials should be constructed for the group-III nitrides.

However, this work focused only on the bulk properties of the cubic phase. Therefore, as we are interested in the wurtzite phase of GaN, we perform calculations to verify how the LDA and GGA (within the PBE parameterization described in Chapter 2) approach performs for GaN in this phase, which are necessary for our study on GaN surfaces. Based on those results then it is possible to decide which exchange-correlation functional is more accurate to describe the surface properties of GaN

surfaces.

In this chapter we present and discuss results for the structural and thermodynamic properties of the hexagonal and cubic phases of GaN-bulk, the N_2 molecule, Ga-bulk, Si-bulk and Si_3N_4 -bulk. Si-bulk and Si_3N_4 -bulk will be needed when studying Si adsorption on GaN surfaces.

3.1 Structural and thermodynamic properties

The equilibrium structure for GaN-bulk, Si-bulk and Si_3N_4 -bulk is obtained by minimization of the total energy with respect to the primitive unit cell volume V . For the zinc-blende case (GaN-bulk in cubic phase and Si-bulk) the volume is directly related to the lattice constant a ($V = a^3/4$), while for the wurtzite structure (GaN-bulk in hexagonal phase and Si_3N_4 -bulk) the volume is a function of the lattice parameters c and a ($V = \sqrt{3}a^2c/2$). Therefore, for the wurtzite structure, the minimization has to be performed by a two-step procedure: the ratio c/a and the internal parameter u have to be optimized for each given volume V . This procedure is repeated for other volumes near the experimental one. We chose for all cases a range from -9% to +9% around the experimental value.

The minimum of energy is calculated interpolating the total energy curve versus volume V using the Murnaghan equation of state [100]. Let us start with the definition of the bulk modulus B

$$B_0(T, P) = -V \left(\frac{\partial P}{\partial V} \right)_T. \quad (3.1)$$

It follows that the derivative of the bulk modulus B'_0 is then given by

$$B'_0 = -\frac{\partial}{\partial P} \left(V \frac{\partial P(T, V)}{\partial V} \right)_{T, P=0}. \quad (3.2)$$

In terms of B_0 and B'_0 the Murnaghan equation of state to determine the minimum of energy is given by

$$E(V) - E(V_0) = \frac{B_0 V}{B'_0} \left[\frac{(V_0/V)^{B'_0}}{B'_0 - 1} + 1 \right] - \frac{B_0 V_0}{B'_0 - 1}. \quad (3.3)$$

The equation above is solved self-consistently together with Eqs. (3.1) and (3.2) to determine the equilibrium volume $V_0 = V(P = 0)$, B_0 , B'_0 and $E(V_0)$.

The cohesive energy is by definition the energy necessary to separate the crystal into its constituent parts. This energy depends on what the constituent parts are considered to be. They are generally taken to be the individual atoms of the chemical elements out of which the solid is composed

$$E_{\text{coh}} = E_{\text{tot}}^{\text{AB-bulk}} - E_{\text{tot}}^{\text{A-atom}} - E_{\text{tot}}^{\text{B-atom}}, \quad (3.4)$$

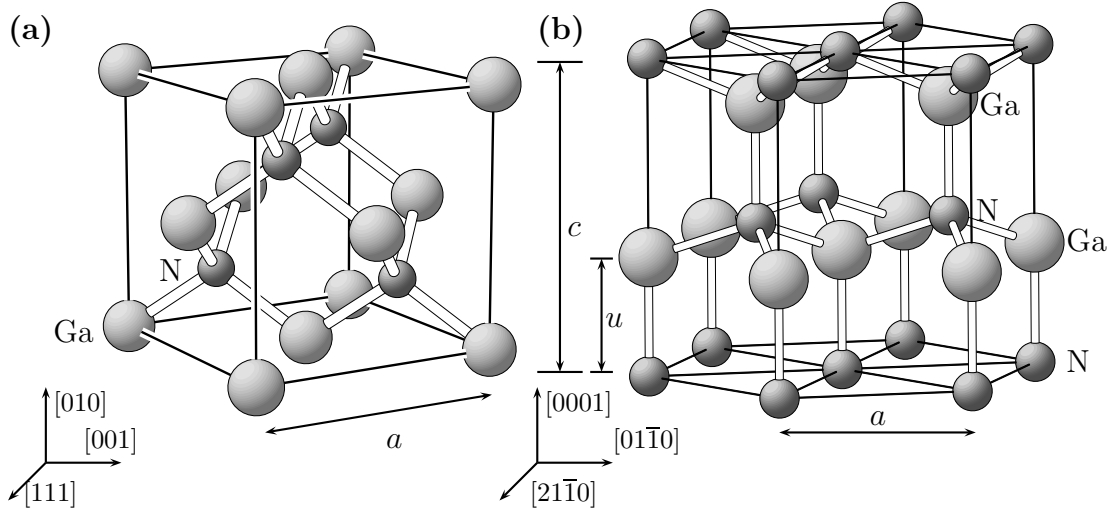


Figure 3.1: Conventional unit cells of (a) β -GaN (zinc-blende structure) and (b) α -GaN (wurtzite structure). a and c are the lattice parameters and u is the internal parameter. In the pictures the atoms belonging to the primitive unit cell are indicated.

where $E_{\text{tot}}^{\text{AB-bulk}}$, $E_{\text{tot}}^{\text{A-atom}}$, $E_{\text{tot}}^{\text{B-atom}}$ are the total energies of AB-bulk, A and B free atoms¹. In the calculation of the cohesive energy we include the so-called spin-polarization correction energy for the spin-saturated spherical free (pseudo) atom by adding the difference of the total energies of the spin-polarized and saturated all-electron atom within the respective exchange-correlation scheme.

When the constituent parts are taken to be the solid phases we have the formation enthalpy of the solid at zero temperature, defined as:

$$\Delta H_f^{\text{T}=0}(\text{AB}) = E_{\text{tot}}^{\text{AB-bulk}} - E_{\text{tot}}^{\text{A-bulk}} - E_{\text{tot}}^{\text{B-bulk}}, \quad (3.5)$$

where $E_{\text{tot}}^{\text{AB-bulk}}$, $E_{\text{tot}}^{\text{A-bulk}}$, $E_{\text{tot}}^{\text{B-bulk}}$ are the total energies of the bulk phases of the compounds AB, A, and B, respectively. Besides the comparison with experiment, the formation enthalpy will also be used in order to calculate the surface energies, as will be explained in the next chapter.

The primitive lattice vectors for the zinc-blende structure are $\mathbf{a}_1 = \frac{a}{2}\hat{i} + \frac{a}{2}\hat{k}$, $\mathbf{a}_2 = \frac{a}{2}\hat{i} + \frac{a}{2}\hat{j}$, $\mathbf{a}_3 = \frac{a}{2}\hat{j} + \frac{a}{2}\hat{k}$ and for the wurtzite structure are $\mathbf{a}_1 = \frac{a}{2}\hat{i} + a\frac{\sqrt{3}}{2}\hat{j}$, $\mathbf{a}_2 = \frac{a}{2}\hat{i} - a\frac{\sqrt{3}}{2}\hat{j}$, $\mathbf{a}_3 = c\hat{k}$, where \hat{i} , \hat{j} and \hat{k} are the unitary vectors along the x , y and z directions respectively.

3.1.1 Structural and cohesive properties

It has been argued that the explicit inclusion of the Ga-3d electrons are essential for a better description of the group-III nitrides structural and cohesive properties

¹We should note that in order to calculate the total energy of a free atom using periodic boundary conditions, some technical aspects should be considered. A large cubic supercell is needed to prevent that one atom from a certain unit cell interacts with an atom of a neighboring unit cell. Convergence tests to determine the optimum size for the supercell for all the species involved are presented in Appendix A.

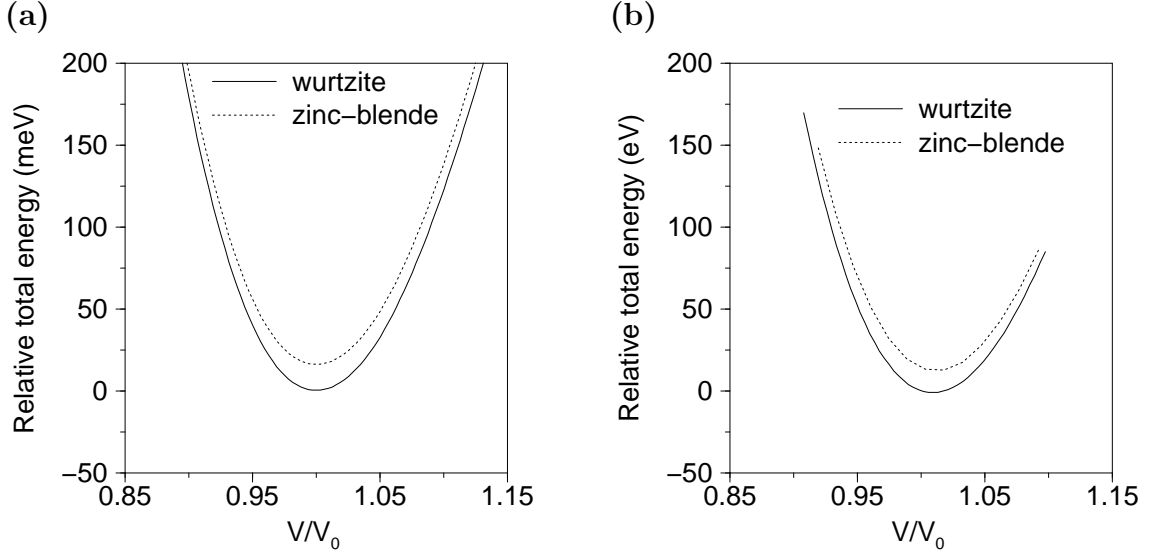


Figure 3.2: Calculated total energy difference of GaN-bulk per Ga-N pair versus normalized volume V/V_0 , where V_0 is the experimental volume for the wurtzite structure. Here we use a) LDA with the non-linear core correction for the exchange-correlation term and b) LDA including the Ga-3d electrons explicitly as valence electrons. The solid line refers to the hexagonal phase (wurtzite structure), while the dotted line refers to the cubic phase (zinc-blende structure). The energy difference in a) is 10 meV and in b) 16 meV. The energy reference is set on the lowest energy value for the wurtzite structure.

[20,101–104]. On the other hand, due to the localization of 3d functions, it is difficult to converge the wave functions in a plane-wave basis set since a high energy cutoff is required. Therefore, *nlcc*, which does not include the Ga-3d explicitly as valence electrons, has the advantage of being computationally cheaper, since it requires a smaller energy cutoff. Therefore, we carried out calculations including the Ga-3d electrons and within the *nlcc* approach to verify whether *nlcc* gives a good description of the structural and cohesive properties of GaN.

The total energy differences per Ga-N pair for GaN-bulk versus the normalized volume V/V_0 (V_0 is the experimental volume) are presented in Figs. 3.2(a) and (b). In Fig. 3.2(a) we treat the Ga-3d electrons as core, i.e., within the so-called *nlcc* approximation and in Fig. 3.2(b) we include the Ga-3d electrons explicitly.

Our results show that the *nlcc* approximation can predict correctly the stability of the GaN phases, where we find that the cubic phase is less stable than the hexagonal phase by 10 meV/pair. Including the Ga-3d electrons we find that this difference is 16 meV/pair. This difference is in good agreement with other theoretical results [97, 104, 105], where it was found that the difference in energy between the two structures less than 20 meV. Our results also agree with experiment, where the hexagonal phase of GaN is found to be the most stable phase [12].

Now we want to discuss the lattice parameters of GaN. The discussion concerns both cubic and hexagonal phases, otherwise noted. Tables 3.1 and 3.2 show the results for the calculated bulk properties: lattice constants a_0 and c_0/a_0 , bulk mod-

Table 3.1: Calculated and experimental structural and thermodynamic properties of β -GaN using LDA and PBE: equilibrium lattice constant a_0 , bulk modulus B_0 , bulk modulus derivative B'_0 , cohesive energy E_{coh} , and formation enthalpy $\Delta H_f^{T=0}$. A cutoff energy of 70 Ry and 28 \mathbf{k} -points in the IBZ were used. The *nlcc* calculations were done using a cutoff energy of 50 Ry. The cohesive energy was corrected by a spin-polarization correction of 0.145 eV (LDA) and 0.179 eV (PBE) for the Ga atom and 3.03 eV (LDA) and 3.12 eV (PBE) for the N atom. PP stands for pseudo-potential, FP-LMTO for full-potential linear muffin-tin orbitals, FP-LAPW for full-potential linearized augmented plane waves and USPP for ultra-soft pseudopotential. PP(*nlcc*) means that non-linear core correction was employed and PP(3d) that the Ga-3d electrons were explicitly included as valence electrons (see Chapter 2).

LDA						
Ref.	Method	a_0 (Å)	B_0 (Mbar)	B'_0	E_{coh} (eV)	$\Delta H_f^{T=0}$ (eV)
[99]	PP(no 3d)	4.33	2.38			-0.46
this work	PP(<i>nlcc</i>)	4.38	1.97	4.50	-11.01	-0.50
[106]	PP(<i>nlcc</i>)	4.45	1.95			
[99]	PP(<i>nlcc</i>)	4.44	1.93			-0.57
[104]	PP(<i>nlcc</i>)	4.45	2.07	3.94		
this work	PP(3d)	4.52	1.88	4.35	-10.38	-1.22
[97]	PP(3d)	4.52	1.91	4.14	-10.18	
[87]	PP(3d)	4.50	1.93		-10.42	-1.34
[102]	PP(3d)	4.46	1.87			
[107]	PP(3d)	4.52	2.06	3.70	-10.53	
[108]	PP(3d)	4.52	2.00	4.15	-10.54	
[109]	PP(3d)	4.30	2.51	2.76		
[99]	PP(3d)	4.48	2.18			-0.99
[110]	USPP	4.45			-10.98	-1.69
[87]	FP-LAPW	4.46		4.46	-10.80	-1.55
[101]	FP-LMTO	4.47	1.98		-10.88	
PBE						
[99]	PP(no 3d)	4.45	2.08			
this work	PP(<i>nlcc</i>)	4.43	2.13	4.41	-9.93	-1.28
this work	PP(3d)	4.60	1.66	4.12	-8.52	-0.62
[97]*	PP(3d)	4.59	1.56	4.25	-8.25	
[110]	USPP	4.54			-9.25	-1.10
[98]	FP-LMTO	4.57				
[99]	PP(<i>nlcc</i>)	4.45	2.08			
[87]	PP(3d)	4.60	1.67		-8.53	-0.59
[87]	FP-LAPW	4.55	1.72	4.01	-8.86	-0.90
[12, 111]	exp.	4.52	1.73	3.70	-8.96	-1.20

* Note: In this work the PWII [65] exchange-correlation potential was used. All other calculations were done using PBE [68].

ulus B_0 , first derivative of the bulk modulus B'_0 , cohesive energy E_{coh} , and formation enthalpy $\Delta H_f^{T=0}$.

We find that LDA underestimates the experimental lattice parameters (for the

Table 3.2: Calculated and experimental structural and thermodynamic properties of α -GaN using LDA and PBE: equilibrium lattice constants a_0 , c_0/a_0 ratio, internal parameter u , bulk modulus B_0 , bulk modulus derivative B'_0 , cohesive energy E_{coh} , and formation enthalpy $\Delta H_f^{T=0}$. A cutoff energy of 70 Ry and 33 \mathbf{k} -points in the IBZ were used. The *nlcc* calculations were done using a cutoff energy of 50 Ry. The cohesive energy was corrected by a spin-polarization correction of 0.145 eV (LDA) and 0.179 eV (PBE) for the Ga atom and 3.03 eV (LDA) and 3.12 eV (PBE) for the N atom. PP stands for pseudo-potential, FP-LMTO for full-potential linear muffin-tin orbitals, FP-LAPW for full-potential linearized augmented plane waves and USPP for ultra-soft pseudopotential. PP(*nlcc*) means that non-linear core correction was used and PP(3d) that the Ga-3d electrons were explicitly included as valence electrons.

LDA								
Ref.	Method	a_0 (Å)	c_0/a_0	u	B_0 (Mbar)	B'_0	E_{coh} (eV)	$\Delta H_f^{T=0}$ (eV)
[105]	PP(no 3d)	3.095	1.633	0.378				
[112]	PP(no 3d)	3.160	1.622	0.377	1.95			
[106]	PP(no 3d)	3.146	1.629	0.377	1.95			
this work	PP(<i>nlcc</i>)	3.133	1.633	0.375	2.17	4.23	-11.03	-1.80
[104]	PP(<i>nlcc</i>)	3.145	1.626	0.377	2.15	5.90		
this work	PP(3d)	3.196	1.631	0.375	1.87	5.39	-10.40	-1.25
[97]	PP(3d)	3.193	1.634	0.376			-10.19	
[110]	USPP	3.131	1.630	0.377			-10.99	-1.69
[98]	FP-LMTO	3.160	1.626	0.377	1.99			
[113]	FP-LMTO	3.170	1.626	0.379		4.5		
PBE								
this work	PP(<i>nlcc</i>)	3.191	1.633	0.375	1.84	4.79	-9.31	-1.15
this work	PP(3d)	3.252	1.629	0.376	1.62	4.06	-8.54	-0.64
[97]	PP(3d)	3.245	1.632	0.376	1.72	5.11	-8.27	
[110]	USPP	3.199	1.634	0.377			-9.27	-1.12
[12]	exp.	3.189	1.624	0.375	1.88	4.3	-9.06	-1.15

* Note: In this work the PWII [65] exchange-correlation potential was used. All other calculations were done using PBE [68].

cubic and hexagonal phases) by around 1% when the Ga-3d electrons are included explicitly and by 3% when *nlcc* is employed. We find that the cohesive energy using LDA is overestimated with respect to the experimental value by 17% and the formation enthalpy by 8%. The lattice parameter with PBE overcorrects and gives a 2% too large value (*nlcc*) and 3% (Ga-3d) compared to experiment.

The GaN cohesive energy is overestimated by 6% compared to the experimental result, which confirms the fact that PBE improves the binding energies in many solids [114–117]. However, the formation enthalpy (−0.64 eV) is underestimated by 44% compared to the experimental value (−1.15 eV). On the other hand, PBE (*nlcc*) gives a value (−1.28 eV) very close to the experimental value (overestimated by 3%). Compared to other theoretical results, we can see that the formation enthalpy is poorly described when using LDA-*nlcc* within the pseudopotential approach (less

than 50% of the experimental value). Compared to all-electron calculations, LDA-3d gives an error of 0.3 eV. Ultra-soft pseudopotential [110] results overestimate the formation enthalpy by 0.5 eV when LDA is used, but are in very close agreement when PBE is used (difference of 0.1 eV).

In order to understand this behavior, the individual contributions (N_2 -molecule, Ga-bulk and GaN-bulk) to calculate the formation enthalpy are analyzed. The formation enthalpy (Eq. (3.5)), involves the binding energies of solid phases of the constituent systems. According to the diagram of Fig. 3.3 we can see that LDA always overestimates the cohesive energy for the constituent species with respect to the experimental value, which means that the error is always positive. Thus, when one calculates the formation enthalpy according to Eq. (3.5), the difference among the individual terms brings the LDA value close to the experimental one.

However, for PBE, the error for the individual systems have different signs: for the N_2 molecule the error is positive, whereas for the GaN-bulk and Ga-bulk the error is negative. Therefore, the difference pushes the PBE value far from the experiment. On the other hand, from the Table 3.2 we can see that using PBE within the *nlcc* approximation gives very close results (difference of 0.03 eV) for the formation enthalpy as using the LDA including the Ga-3d electrons. However, the structural properties and cohesive energies (of the bulk materials) are not well described. Based on the results discussed above, we can draw the following conclusions:

- (i) the magnitude of the error of the structural properties are similar using LDA and PBE; however, in different directions: LDA underestimates the lattice parameters whereas PBE underestimates them;
- (ii) PBE gives better results than LDA for the structural properties and cohesive energy of all compounds (Ga-bulk, N_2 -molecule and GaN-bulk) if the Ga-3d electrons are included as valence electrons;
- (iii) PBE performs worse than LDA for the formation enthalpy, but we can explain it noticing that the description of the binding energies for the N_2 molecule and the bulk phases have errors with different signs (with respect to the experimental value) using LDA and PBE;
- (iv) LDA-*nlcc* performs worse than PBE-*nlcc* for structural properties and for energetics as well;
- (v) PBE-*nlcc* gives similar results to LDA for the structural properties, cohesive energy and formation enthalpy;

Therefore, to reduce computational efforts, combining PBE and *nlcc* might be a good approach for the exchange-correlation functional when calculating GaN bulk properties, since the energy cutoff can be reduced from 70 Ry to 50 Ry, which means that the number of plane waves is reduced. Now that we understand how LDA and PBE performs for GaN-bulk, let us discuss a possible way of improving the description of the energetics using the pseudopotential approach. The way of constructing

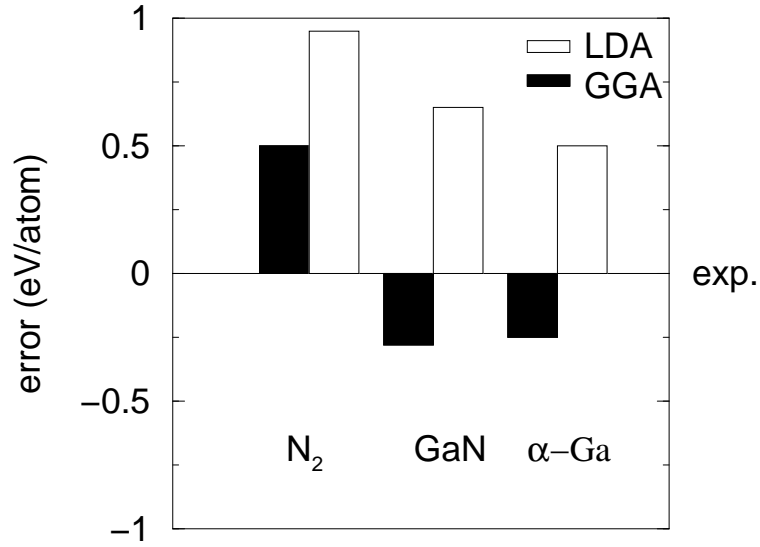


Figure 3.3: Relative error with respect to the experimental value for the binding energies of the N₂ molecule (LDA, 0.94 eV and PBE, 0.42 eV), α-Ga (LDA, 0.48 eV and PBE, 0.27 eV) and GaN-bulk (LDA, 0.67 eV and PBE, 0.26 eV). The black boxes refers to the PBE and the white boxes to the LDA.

more complete pseudopotentials consists in including more angular momentum components l for the pseudoatom. Fuchs *et al.* [87] have systematically studied such an effect for group-III nitrides materials. They found that the inclusion of the Ga- f orbitals in addition to Ga- d orbitals improves (leads values closer to the experimental ones) the formation enthalpy of cubic GaN by 0.19 eV using LDA and by 0.16 eV using PBE (see Table 3.1).

Therefore, trying to improve the pseudopotential leads to a better, description of the formation enthalpy compared to the experimental values. Since the inclusion of Ga-3 f orbitals is computationally more demanding when studying surfaces (even using f as the local component for the pseudopotential, we still have to calculate the d projectors). Since the improvement on the formation enthalpy is smaller than 0.2 eV, we consider the inclusion of the s, p and d orbitals only.

3.1.2 Electronic properties

Figs. 3.4(a)-(d) show results for the band structure calculations of GaN in the wurtzite and zinc-blende structures using LDA [(c) and (d)] and PBE [(a) and (b)]. Both phases have a direct band gap, with valence band maximum and conduction band minimum located at the Γ point of the Brillouin zone. The Brillouin zones are shown in Figs. 3.5(a) face-centered cubic lattice and (b) hexagonal lattice. The theoretical band gap is calculated using the energy difference between the highest occupied eigenvalue and the lowest unoccupied eigenvalue.

While there is a single conduction band with Γ_7 symmetry, there are three valence bands, which are non-degenerate. Due to the spin-orbit and crystal field splitting,

Table 3.3: Calculated and experimental band gap energies E_{gap}^{Γ} at the Γ point for β - and α -GaN bulk using LDA and PBE. PP stands for pseudo-potential and FP-LAPW for full-potential linearized augmented plane waves. PP(*nlcc*) means that non-linear core correction was used and PP(3d) that the Ga-3d electrons were explicitly included as valence electrons. All values are given in eV.

Ref.	Method	E_{gap}^{Γ} (eV)			
		β -GaN		α -GaN	
		LDA	PBE	LDA	PBE
[99]	PP(no 3d)	2.35	2.10		
this work	PP(<i>nlcc</i>)	2.20	1.90	2.35	2.01
[99]	PP(<i>nlcc</i>)	1.81	1.99		
this work	PP (3d)	1.60	1.27	1.76	1.39
[97]	PP(3d)	1.60	1.28	1.76	1.45
[97]	PP(3d)	1.89			
[99]	PP(3d)	1.48			
[87]	FP-LAPW	1.68			
[12]	exp.		3.45		3.41

the top of the valence band separates in three bands, two with Γ_7 symmetry (called light-hole and spin-orbit splitting band) and one with Γ_8 symmetry (called heavy-hole).

In this calculation, we do not include spin-orbit coupling. Therefore, the spin-orbit splitting is not seen in the band structure. However, the splitting due to the crystal-field is found to be 10 meV. The experimental values lie in the range of 10-25 meV [12].

The band structure was calculated at the optimized geometry for both structures. We can see that the band structures look almost identical for both exchange-correlation functionals. For the cubic phase we find a band gap of 1.60 eV using LDA (including the Ga-3d electrons) and 3.51 eV (*nlcc*), while using PBE we found 1.27 eV (including the Ga-3d electrons) and 3.20 eV (*nlcc*). For the hexagonal phase, we find a gap of 1.76 eV using LDA (including the Ga-3d electrons) and 3.30 eV (*nlcc*), while using PBE we found 1.39 eV (including the Ga-3d electrons) and 3.10 eV (*nlcc*). The good agreement with the measured gaps for α -GaN (3.4 eV) and β -GaN (3.2 eV) [12, 111] using *nlcc* comes from the fact that the lattice constant is significantly underestimated compared to the experimental values ($\approx 3\%$) and therefore the agreement is accidental.

3.2 Ga-bulk

Ga-bulk has different bulk phases (Ga-II, Ga-III [118], α [119], β [120], γ [121], δ [122] and ϵ [118]), depending on the pressure and temperature. Several theoretical

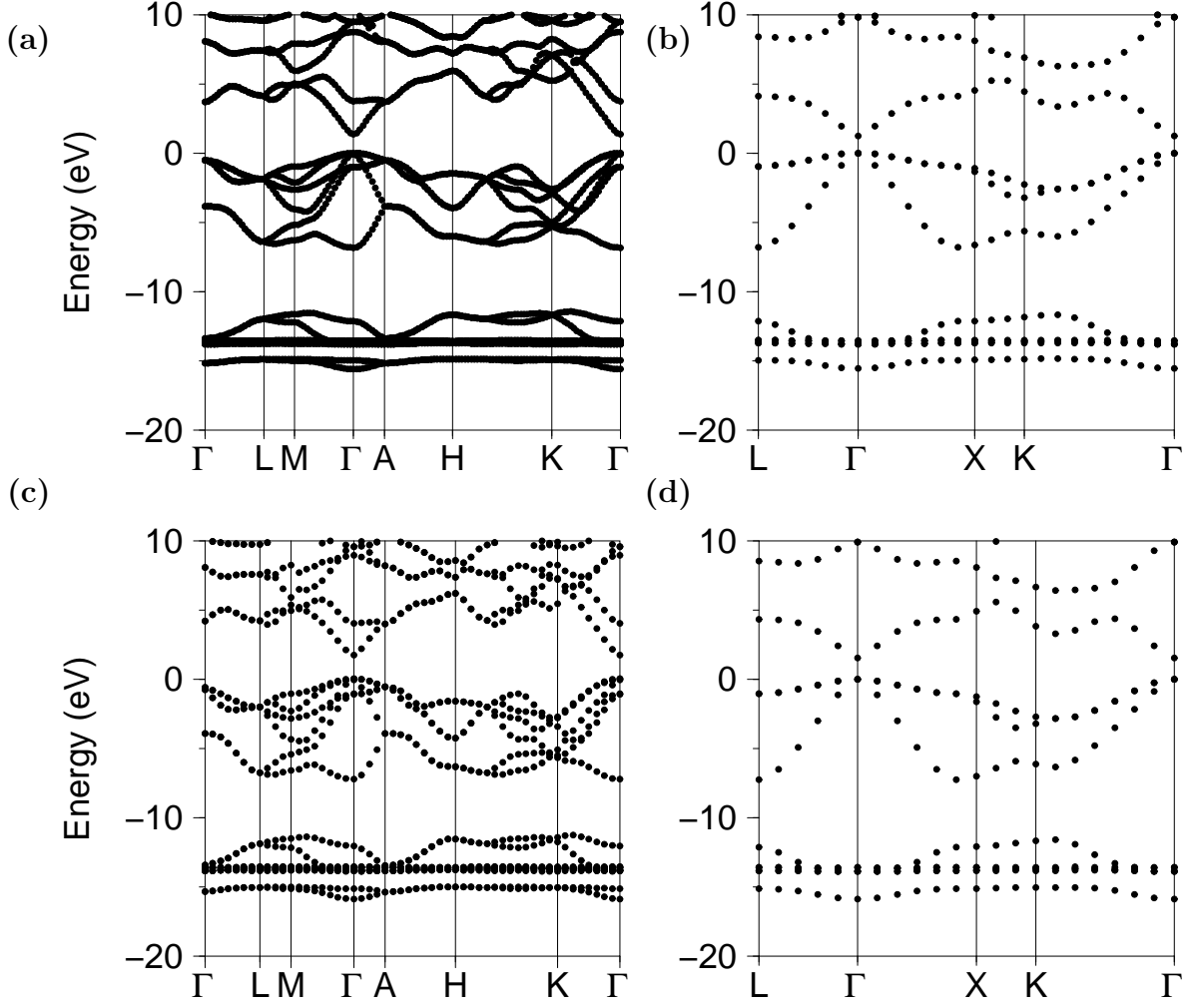


Figure 3.4: Bulk band structure of hexagonal (wurtzite) and cubic (zinc-blende) GaN. (a) Hexagonal phase using PBE, (b) cubic phase using PBE, (c) hexagonal phase using LDA, and (d) cubic phase using LDA. The optimized theoretical lattice constants were used for the calculations.

and experimental studies have shown that α -Ga is the stable phase at room temperature and low pressure (up to 16×10^3 atm) [123]. Therefore we concentrate ourselves only on the properties of α -Ga, because these conditions are similar to those in MBE experiments. Besides, the other phases are similar in energy and would not affect any of the conclusions drawn here. The α -phase has an orthorhombic structure with eight atoms per unit cell.

A peculiar feature is that each atom has only one nearest neighbor at a distance of 2.44 Å. The second, third and fourth shell each contain two atoms and are 0.27, 0.30 and 0.39 Å further away. The corresponding structure is shown in Fig. 3.6. The primitive lattice vectors are $\mathbf{a}_1 = a\hat{i}$, $\mathbf{a}_2 = b\hat{j}$ and $\mathbf{a}_3 = c\hat{k}$. The relative coordinates of the atoms of the basis in the unit cell are: $(0, \pm u, \pm v)$, $(\frac{1}{2}, \pm u, \frac{1}{2} \pm v)$, $(\frac{1}{2}, \frac{1}{2} \pm u, \pm v)$, $(0, \frac{1}{2} \pm u, \frac{1}{2} \pm v)$, where u and v are the internal parameters. We should

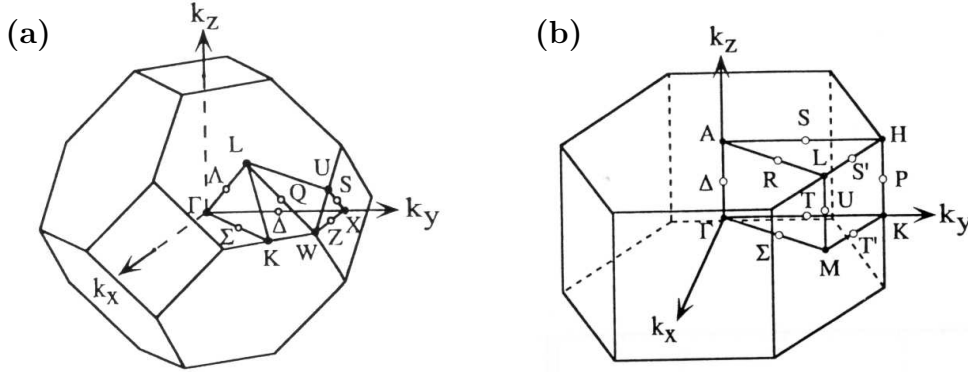


Figure 3.5: First Brillouin zones with the highest symmetry lines of (a) face-centered cubic (zinc-blende structure) and (b) hexagonal lattice.

mention that, for a sake of computational effort, all properties were calculated at the experimental geometry, i.e., the structure was not relaxed.

The results for the calculated and experimental properties are shown in Table 3.4. We can see that for LDA-*nlcc*, the cohesive energy is in good agreement with the all-electron calculations, although we should consider this agreement carefully, since the structure was not relaxed. We again emphasize that the inclusion of the Ga-3*d* electrons in principle should lead to a more accurate description of the bonds and consequently of the properties of the solid. The cohesive energy calculated including the Ga-3*d* electrons is overestimated by around 15% using LDA, while for PBE it is underestimated by 10%.

3.3 Si-bulk

We also calculated the structural and thermodynamical properties of Si-bulk, which will be needed when studying GaN surfaces. This element has a diamond structure at low pressure with 2 Si atoms in the primitive unit cell at the positions (0,0,0) and (1/4,1/4,1/4), as shown in Fig. 3.6(a). The calculated properties are shown in Table 3.5, where we compare our results with other calculations and experimental values. As expected, LDA predicts a slightly smaller lattice constant (−1%) and a larger cohesive energy (+10%) than the experimental value, while PBE gives a slightly larger lattice constant (+1%) and cohesive energies closer to the experimental value (+0.5%), as we have found for the other crystals.

3.4 Si₃N₄-bulk

Here we present results for the β -Si₃N₄, which might be the solubility limiting phase of Si in GaN structure. Such a structure has 14 atoms per unit cell (8 N and 6

Table 3.4: Calculated and experimental structural and thermodynamic properties of Ga-bulk in the α phase using LDA and PBE: equilibrium lattice constants a_0 , b_0/a_0 , c_0/a_0 , u and v and cohesive energy E_{coh} . The experimental lattice parameters were taken from Ref. [123]. In our calculations they were not optimized. Parameters used in the calculations: 70 Ry and 216 \mathbf{k} -points in the IBZ. The *nlcc* calculations were done using 50 Ry and 216 \mathbf{k} -points in the IBZ. The cohesive energy was corrected by a spin-polarization correction of 0.145 eV (LDA) and 0.179 eV (PBE). These values were extracted from Ref. [85]. PP stands for pseudopotential, FP-LAPW for full-potential linearized augmented plane waves and USPP for ultra-soft pseudopotential. PP(*nlcc*) means that the non-linear core correction was used and PP(3d) that the Ga-3d electrons were explicitly included as valence electrons.

LDA							
Ref.	Method	a_0 (Å)	b_0/a_0	c_0/a_0	u	v	E_{coh} (eV)
this work	PP(3d)	4.51	1.001	1.695	0.0785	0.1525	-3.29
this work	PP(<i>nlcc</i>)	4.51	1.001	1.695	0.0785	0.1525	-3.47
[87]	PP(<i>nlcc</i>)	4.51	1.001	1.695	0.0785	0.1525	-3.21
[87]	FP-LAPW	4.44	0.997	1.691	0.0801	0.1566	-3.46
[124]	PP(<i>nlcc</i>)	4.36	0.997	1.692	0.080	0.1560	-3.48
[110]	USPP	4.44	0.999	1.696	0.0816	0.1577	-3.48
[123]	PP(<i>nlcc</i>)	4.38	0.994	1.688	0.0803	0.1567	
PBE							
this work	PP(3d)	4.51	1.001	1.695	0.0785	0.1525	-2.54
this work	PP(<i>nlcc</i>)	4.51	1.001	1.695	0.0785	0.1525	-2.76
[87]	PP(3d)	4.51	0.997	1.691	0.0801	0.1566	-2.60
[87]	FP-LAPW	4.59	1.690	0.993	0.803	0.1567	-2.71
[124]*	PP(no 3d)	4.58	0.997	1.691	0.082	0.1530	-2.67
[110]	USPP(3d)	4.60	0.992	1.696	0.0834	0.1559	-2.80
[125]	exp.	4.511	1.001	1.695	0.0785	0.1525	-2.81

* Note: In this work the exchange-correlation functional used was PWII [65]. All other calculations were done using PBE [68].

Si) and belongs to the point group $P6_3/m$ (see Fig. 3.7). Each Si atom is tetrahedrally coordinated with one N atom. The N atoms are nearly planarly three-fold coordinated with Si.

The coordinates of the atoms of the basis in the primitive unit cell are: N atoms at $\pm(\frac{1}{3}, \frac{2}{3}, \frac{1}{4})$, $\pm(x_N, y_N, \frac{1}{4})$, $\pm(y_N - x_N, -x_N, \frac{1}{4})$ and $\pm(-y_N, x_N - y_N, \frac{1}{4})$ and Si atoms at $\pm(x_{\text{Si}}, y_{\text{Si}}, \frac{1}{4})$, $\pm(y_{\text{Si}} - x_{\text{Si}}, -x_{\text{Si}}, \frac{1}{4})$ and $\pm(-y_{\text{Si}}, x_{\text{Si}} - y_{\text{Si}}, \frac{1}{4})$, where the atomic positions x, y, z are meant to be multiplied by $\mathbf{a}_1, \mathbf{a}_2$ and \mathbf{a}_3 . The primitive lattice vectors are $\mathbf{a}_1 = \frac{a}{2}\hat{i} + a\frac{\sqrt{3}}{2}\hat{j}$, $\mathbf{a}_2 = \frac{a}{2}\hat{i} - a\frac{\sqrt{3}}{2}\hat{j}$, $\mathbf{a}_3 = c\hat{k}$.

Recent theoretical investigations have shown that this phase is the most stable one at $T = 0$ and $P = 0$. At $P = 42$ atm and $T = 1770$ K, the β phase transforms into

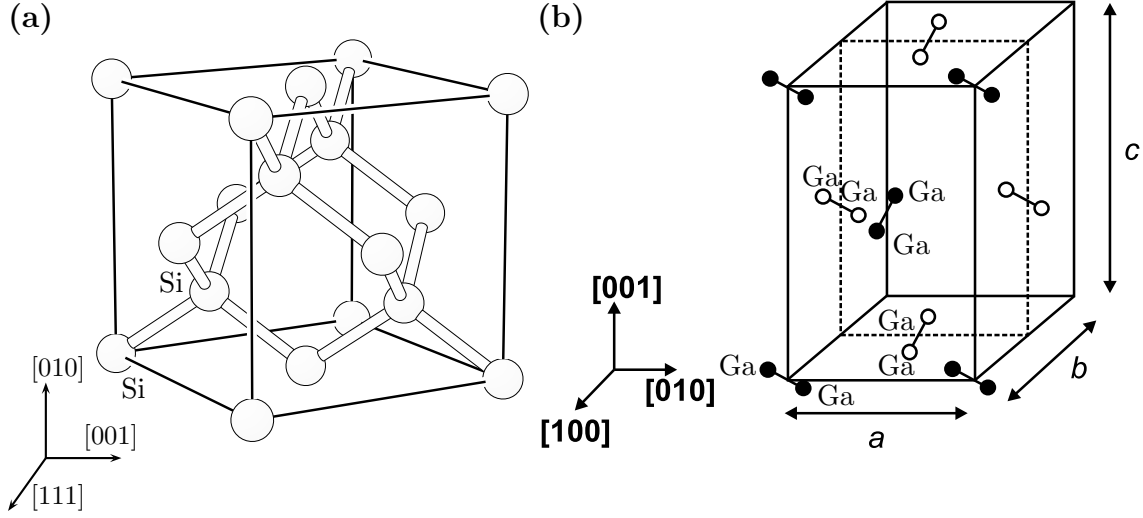


Figure 3.6: Atomic structures of (a) Si- bulk (diamond structure) and (b) α -Ga (orthorhombic structure). The two inequivalent Si atoms per primitive unit cell are indicated (a) and the eight inequivalent Ga atoms per primitive unit cell are indicated in (b).

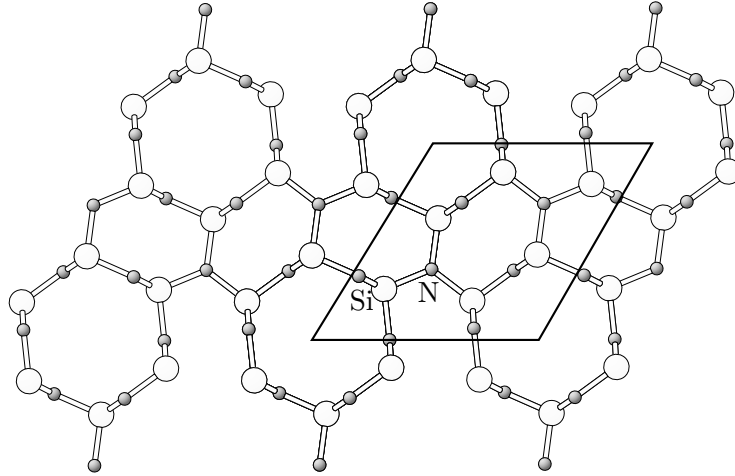


Figure 3.7: View along the [0001] plane of the hexagonal structure of β - Si_3N_4 contains 14 atoms per unit cell (8 N and 6 Si). The primitive unit cell is indicated.

the α phase, which has twice as many atoms per unit cell and a different stacking along the $\{c\}$ direction [129]. The optimized structure and the cohesive properties are shown in Table 3.6 and compared with the experimental and theoretical data available. Our results for the lattice parameters show very good agreement with experimental values for the structural properties and other theoretical calculations.

However, for the cohesive properties, our results differ drastically from the other

Table 3.5: Calculated structural and thermodynamic properties of Si-bulk using LDA and GGA: equilibrium lattice constant a_0 , bulk modulus B_0 , bulk modulus derivative B'_0 and cohesive energy E_{coh} . We use a energy cutoff of 50 Ry and 10 k-points in the IBZ. The spin-polarization correction for the Si atom is 0.66 eV (LDA) and 0.79 eV (PBE). These values were extracted from Ref. [85]. PP stands for pseudopotential, LAPW for linear augmented plane waves and PAW for projected augmented waves and USPP for ultra-soft pseudopotential. In brackets are the different parameterizations of the GGA exchange-correlation functional (see Sec. 2.2.4).

LDA				
Ref.	Method	a_0 (Å)	B_0 (MBar)	E_{coh} (eV)
this work	PP	5.38	0.95	-5.18
[85]	PP	5.38	0.94	-5.34
[126]	PAW	5.38	0.98	-6.03
[126]	LAPW	5.41	0.98	-5.92
[126]	PP	5.39	0.98	-5.99
[127]	USPP	5.40	0.95	-5.96
[127]	PAW	5.40	0.95	-5.96
GGA				
this work	PP(PBE)	5.47	0.85	-4.61
[85]	PP(BP)	5.47	0.85	-4.46
[85]	PP(PW)	5.46	0.87	-4.64
[128]	LAPW(BP)	5.54	0.80	
[128]	LAPW(PW)	5.50	0.83	
[125]	exp.	5.43	0.99	-4.63

pseudopotential calculations (-100%). We believe our pseudopotentials give more reliable results than the pseudopotential results of Ref. [112], since they give the typical LDA behavior ($\approx 20\%$ for the cohesive energy compared to experiment).

3.5 N₂-molecule

The results for the binding energy, bond length and vibration frequency for the N₂ molecule are shown in Table 3.7. The binding energy for the N₂ molecule is obtained as the energy difference between the N₂ molecule and the N atom total energies $E_b = E_{\text{tot}}^{\text{N}_2\text{-molecule}} - 2E_{\text{tot}}^{\text{N-atom}}$. The vibration frequency is calculated from the total energy versus N₂ bond length. The procedure was the same done for the GaN bulk. The bond length was varied around the experimental bond length (1.10 Å) ranging from -9% to 9%. Then, in the harmonic approximation the vibration frequency is calculated according to

Table 3.6: Calculated and experimental equilibrium lattice constants a_0 , c_0/a_0 , cohesive energy E_{coh} and formation enthalpy $\Delta H_f^{T=0}$ for the β - Si_3N_4 using LDA. In Refs. [130] and [112] the ratio c/a was not relaxed. We use a energy cutoff of 70 Ry and 63 \mathbf{k} -points in the IBZ. The spin-polarization correction for the Si atom is 0.79 eV and for the N atom is 3.03 eV. These values were extracted from Ref. [85]. PP stands for pseudopotential, HF for Hartree-Fock and OLCAO for orthogonalized linear combination of atomic orbitals.

Ref.	Method	a (Å)	c_0/a_0	B_0 (MBar)	E_{coh} (eV/cell)	$\Delta H_f^{T=0}$ (eV/atom)
this work	PP	7.600	0.383	2.56	-97.91	-9.95
[112]	PP	7.610	0.382	2.65	-48.96	
[130]	HF	7.610	0.382	2.97		
[129]	OLCAO	7.620	0.382	2.74	-74.3	
[131, 132]	exp.	7.608	0.382	2.56, 2.58, 2.73	-82.96	-7.71, -8.83

$$\mu = \frac{1}{2\pi} \sqrt{\frac{K}{\mu_r}}, \quad (3.6)$$

where μ_r is the reduced mass of the N_2 molecule and K is the force constant.

The present results are in good agreement with previous LDA and PBE calculations using the pseudopotential approach [87]. Compared to the LDA results, PBE leads to very similar results, but slightly longer bond lengths, lower frequencies and smaller binding energies which are closer to experiment. The binding energy is overestimated by 16% using LDA and 8% using PBE. From Table 3.7 we can see that our results compare quite well with others, indicating that our pseudopotential is reliable in describing properties of molecules as well.

3.6 Summary

After having examined the properties for the N_2 molecule, GaN-bulk, Ga-bulk, Si-bulk and Si_3N_4 -bulk described above, we can conclude that carefully constructed pseudopotentials accurately describe these properties. In particular, PBE performs better in describing the binding energy of solids and molecules. For the structural properties we did not find real improvement, since LDA always underestimates the experimental values, while PBE overestimate them by the same amount.

The formation enthalpy of GaN using PBE-3d is strongly underestimated, while LDA provides a value very close to the experiment. However, the use of PBE within *nlcc* gives a value close to the experimental one. Therefore, as the formation enthalpy is a key quantity to derive phase diagrams, we decided to use LDA in the following calculations instead of GGA. Based on the discussion above, we state that, if LDA is the chosen functional, the Ga-3d electrons should be treated as valence electrons.

Concerning to electronic properties we find a very good qualitative agreement

Table 3.7: Calculated and experimental bond length d , binding energy E_b and vibration frequency ω for an N_2 molecule using LDA and GGA. The binding energy was corrected by a spin-polarization correction of 3.03 eV (LDA) and 3.12 eV (PBE). These values were extracted from Ref. [87]. We use a cubic supercell with length $L = 20$ Bohr, 1 \mathbf{k} -point in the IBZ and 70 Ry cutoff energy. The zero point vibration energy of the molecule (0.153 eV) has been included.

Ref.	LDA				GGA		
	d (Å)	E_b (eV)	ω (cm $^{-1}$)	functional	d (Å)	E_b (eV)	ω (cm $^{-1}$)
this work	1.09	-11.71	2363	PBE	1.09	-10.66	2331
[97]	1.10	-11.59	2384				
[87]	1.08	-11.75	2385	PBE	1.09	-10.69	2325
[87]	1.10	-11.57	2398	PW	1.10	-10.41	2354
[110]	1.11	-11.33		PBE	1.11	-10.56	
[133]				PW	1.10	-9.87	2346
[68]				PBE	1.10	-10.49	
[75]				BP	1.10	-10.30	2330
[65]				PWII	1.10	-10.10	2320
[134](exp.)	1.10	-9.82	2360				

between LDA and PBE for the band structure. We conclude that the difference in the band gap is attributed only to different descriptions of the lattice parameters.

We have shown that the pseudopotential approach, when constructed in a careful manner, can provide very good description of bulk, atoms and molecules, compared to all-electrons calculations. Also, the necessity of optimizing the parameters that control the transferability of the pseudopotential, such as cutoff radii, number of the l components is crucial, as pointed out by Fuchs *et al.* [87].

Chapter 4

Bare GaN (0001) surfaces

The structural and electronic properties of GaN surfaces depend sensitively on the orientation of the surface (i.e., along which plane the crystal is cleaved), surface termination (i.e., which chemical species is in the top surface layer) and reconstruction. The most common growth direction of epitaxial hexagonal GaN is normal to the $\{0001\}$ basal plane¹.

In Fig. 4.1 we show a schematic representation of two possible directions $[0001]$ (also called Ga-polar or Ga-face) and $[000\bar{1}]$ (also called N-polar or N-face). It is important to note that the two possible directions are not equivalent, because the wurtzite structure has no inversion symmetry. Experimentally the orientation can be controlled by the choice of the substrate (surface orientation, surface preparation) and the growth process. A large number of studies have revealed that high quality GaN films deposited by MOCVD results in growth along $[0001]$ while MBE commonly results in growth along the $[000\bar{1}]$ direction [13].

In this work we will focus on the GaN (0001) surface, which has been demonstrated to have the better surface morphology and which is the relevant surface for all technological applications realized up to now [5]. The polar (0001) surface exhibits, depending on the growth conditions, a variety of structures, such as (1×1) (unreconstructed), and reconstructed (2×2) , (4×4) , (5×5) and (6×4) [32, 135–138].

Several theoretical and experimental studies were devoted to identifying the electronic structure of these surfaces. The (2×2) reconstruction has been explained as being a N-adatom structure under N-rich conditions and a Ga-adatom structure under Ga-rich conditions [139]. The (5×5) structure has been proposed to consist of Ga- and N-adatoms [136] and for the (4×4) and (6×4) the atomic structure has not been completely clarified yet.

Particular attention has been given to the (1×1) structure. Concerning the atomic structure, Sung *et al.* [140] concluded from their LEED (low energy electron

¹The planes in a crystal are usually denoted by Miller indices (hkl) where h, k, l are the integer reciprocal axes given by the intersections of the lattice planes with three crystallographic axes. In the case of trigonal and hexagonal lattices, four crystallographic axes are needed. The lattice planes are then characterized by four indices $(hkil)$ instead of three (hkl) . The first three are not independent of each other. It holds that $i = -h - k$. The $(hkil)$ are termed Bravais indices.

diffraction) studies that both GaN (0001) and (000 $\bar{1}$) surfaces are neither reconstructed nor relaxed.

The electronic structure of the (1×1) surface was addressed more recently by angle-resolved photoemission spectroscopy (ARPES) where Dhesi and *et al.* [141] investigated the bulk electronic band structure of thin GaN films grown by MBE. In addition to the bulk bands, they observed a dispersionless surface state band near the valence maximum as well.

Chao *et al.* [142] reported a more complete account on the electronic structure of the (1×1) surfaces using synchrotron-radiation-excited ARPES. However, as ARPES does not provide unambiguous information about the termination or polarity of the films, the observed states can be due either from the (0001) or (000 $\bar{1}$) surface. One band, close to the valence band maximum, was found to be weakly dispersive and very sensitive to the hydrogen adsorption, suggesting the surface has dangling bonds at the outermost layer. The other band was found to be very sensitive to the quality of the sample, as determined by LEED.

Wang *et al.* [143] performed density-functional theory calculations within the local-density approximation using pseudopotential method for the clean (0001) and (000 $\bar{1}$) surface in order to compare with the available experimental data. They calculated the atomic and electronic structure of several possible (1×1) structures. They found that the clean Ga-terminated surface has a very small relaxation. Concerning the electronic structure, none of the Ga-polar structures was consistent with the ARPES data by Chao *et al.* [142]. Instead, the best agreement with the experimental result was achieved when comparing the theoretical surface states of the N-terminated (000 $\bar{1}$) surface.

As we can see from the discussion above, although the atomic structure of the GaN (0001) surface has been largely investigated, [31, 36, 135, 136, 144, 145], concerning the electronic structure our understanding is only at the beginning. Also, which exchange-correlation functional (LDA, PBE) should be used to better describe the GaN surfaces properties has not yet been verified.

In this chapter we will therefore focus on the study of the atomic and electronic structure of the most relevant bare GaN (0001) surfaces employing LDA and PBE. The conclusions about the performance of the functionals will be an useful information in order to choose which functional will be used to study the adsorption of Si on GaN surfaces.

4.1 Surface reconstructions and relaxations

We will first focus on the unreconstructed surfaces. In the present work, the unreconstructed Ga and N-terminated (0001) surfaces are modeled using an (1×1) unit cell. A top view of these surfaces is shown in Fig. 4.2(a). A side view of these surfaces is shown in Fig. 4.2(b) (N-terminated) and Fig. 4.2(c) (Ga-terminated). Each (1×1) unit cell contains one atom per layer. Structures with a Ga-adatom and N-adatom are modelled using a (2×2) unit cell. Each (2×2) cell contains four atoms in each of the underlying layers and one additional atom per layer. A top view of the (2×2)

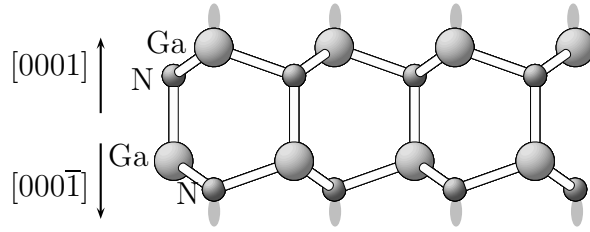


Figure 4.1: Schematic picture of the Ga- ([0001] direction) and N-polar ([000 $\bar{1}$] direction) GaN (0001) surfaces. The shaded gray ellipses mark the surface dangling bonds which are formed when cleaving the crystal. Here the (0001) surface is Ga-terminated and the (000 $\bar{1}$) is N-terminated.

unit cell is shown in Fig. 4.2(a) and a side view is shown in Figs. 4.2(d) and (e). There are many possible sites to adsorb atoms on the surface. Here we consider the highest symmetry adsorption sites, *hcp*, *fcc* and *ontop*, as shown in Fig. 4.2(a).

In addition to the clean and adatom structures, we also studied structures where additional Ga layers lay on the top layer of the clean Ga-terminated surface. The side view of a structure with an additional Ga layer (Ga-adlayer) on the top of the clean Ga-terminated surface is shown in Fig. 4.3 (b) and with two Ga layers (Ga-bilayer) on the top of the clean Ga-terminated surface is shown in Fig. 4.3 (c). Besides, we also studied structures where the outermost top layer was contracted. To model the laterally contracted adlayer and bilayer structures we employ a $(\sqrt{3} \times \sqrt{3})R30^\circ$ unit cell, as has been suggested in Ref. [36].

In these $(\sqrt{3} \times \sqrt{3})R30^\circ$ cells there are four atoms in the laterally contracted hexagonal overlayer for every three atoms in the (1×1) underlying hexagonal layer. The lattice vectors of the overlayer are rotated by 30 degrees with respect to those of the substrate. From now on we will call this cell $(\sqrt{3} \times \sqrt{3})$. Such a model allows the the Ga-Ga spacing to be close to the value which minimizes the formation energy of a free standing layer of Ga. Consequently, the $(\sqrt{3} \times \sqrt{3})$ cell should provide a very good upper bound for the energy of the optimal laterally contracted overlayer structure.

The laterally contracted Ga-Ga spacing is $a_c = (\sqrt{3}/2)a_{1 \times 1} = 2.75 \text{ \AA}$, where $a_{1 \times 1} = 3.19 \text{ \AA}$ is the in-plane spacing of Ga-atoms on the (1×1) unit cell, i.e., the Ga adatoms are compressed by around 14% compared to the full monolayer. The side view of the contracted Ga-bilayer structure is shown in Fig. 4.3(d) and of the contracted Ga-adlayer structure in Fig. 4.3(e).

For all of these structures the equilibrium geometry has been calculated. For the unreconstructed (1×1) surfaces the three topmost layers have been allowed to relax. For the adatom structures the two top layers in addition to the adatom were relaxed. For the contracted and non-contracted structures the four outermost layers were allowed to relax. The non-contracted structures will be considered later on. In Table 4.1 we show results for the surface relaxations of the structures described above.

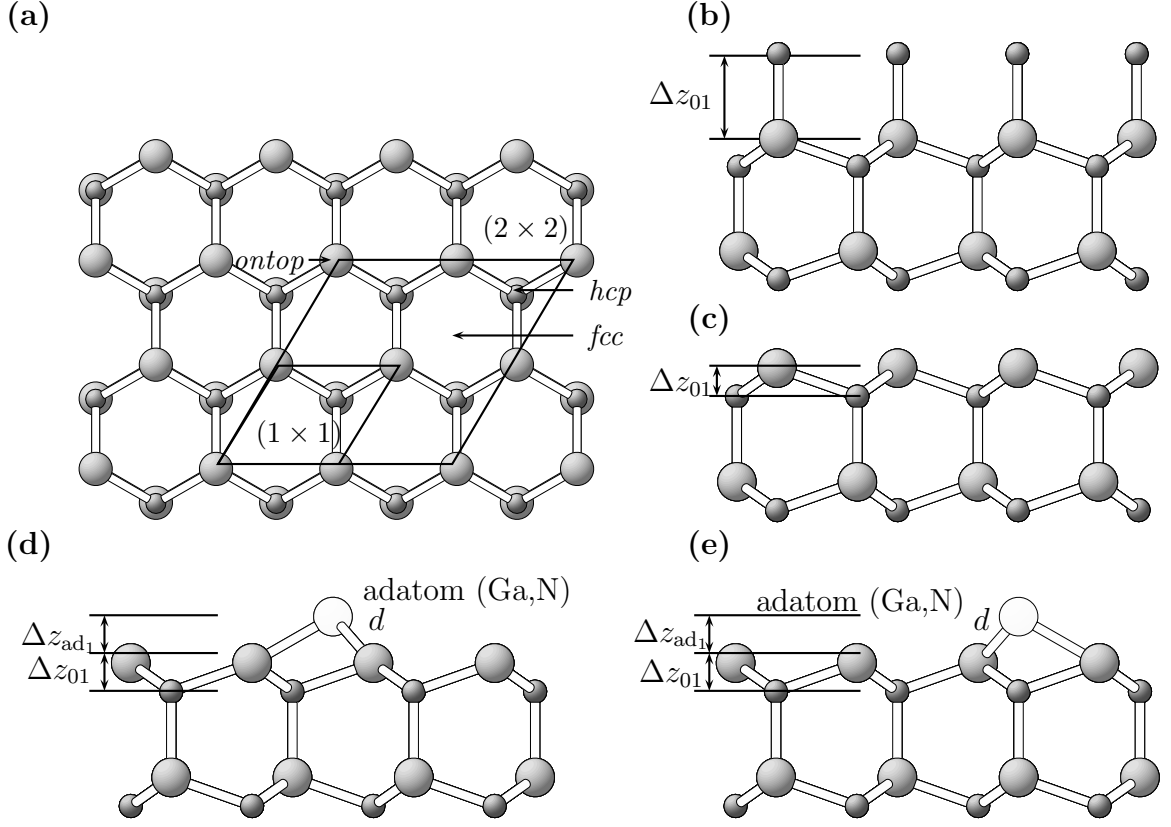


Figure 4.2: (a) Top view of the (1×1) and (2×2) unit cells used to model the GaN (0001) surfaces. The highest symmetry sites *fcc*, *hcp* and *ontop* are indicated. (b) Side view of the clean N-terminated, (c) clean Ga-terminated, (d) adatom-*fcc* structure, and (e) adatom-*hcp* structures. d is the bond length between the adatom (Ga or N) to the Ga atoms at the first plane, Δz_{ad1} is the distance between the adatom and the first plane and Δz_{01} the distance between the first and second plane. Small gray balls are N atoms, big gray balls mark the Ga atoms and white balls the (Ga,N) adatoms.

From the results listed in Table 4.1 we can see that the atomic geometries of GaN (0001) surfaces are quite similarly described by LDA and PBE. The main discrepancy is for the adatom structures where LDA gives slightly large values. However, the distance d is very similar in both cases. This is because the relative relaxation of the top layers is almost the same using both functionals. The LDA results are in good agreement with previous LDA studies [146], for the Ga- and N-adatom relaxations.

For the adatom structures, we can see that the hollow sites *fcc* and *hcp* show a different behavior when N is the adatom. For the *fcc* site, the distance between the N adatom and the first plane formed by Ga-atoms is much larger than for the *hcp* structure. This stems from the fact that N at the *hcp* site feels the presence of the N atom at the third layer directly below, while for the *fcc* site it does not occur, as pointed out in Refs. [144, 147]. For the Ga adatom, the relaxations for the *fcc* and *hcp* sites are almost identical. For the contracted Ga-bilayer structures the spacing between the Ga atoms between the first and the second layer (2.37 \AA)

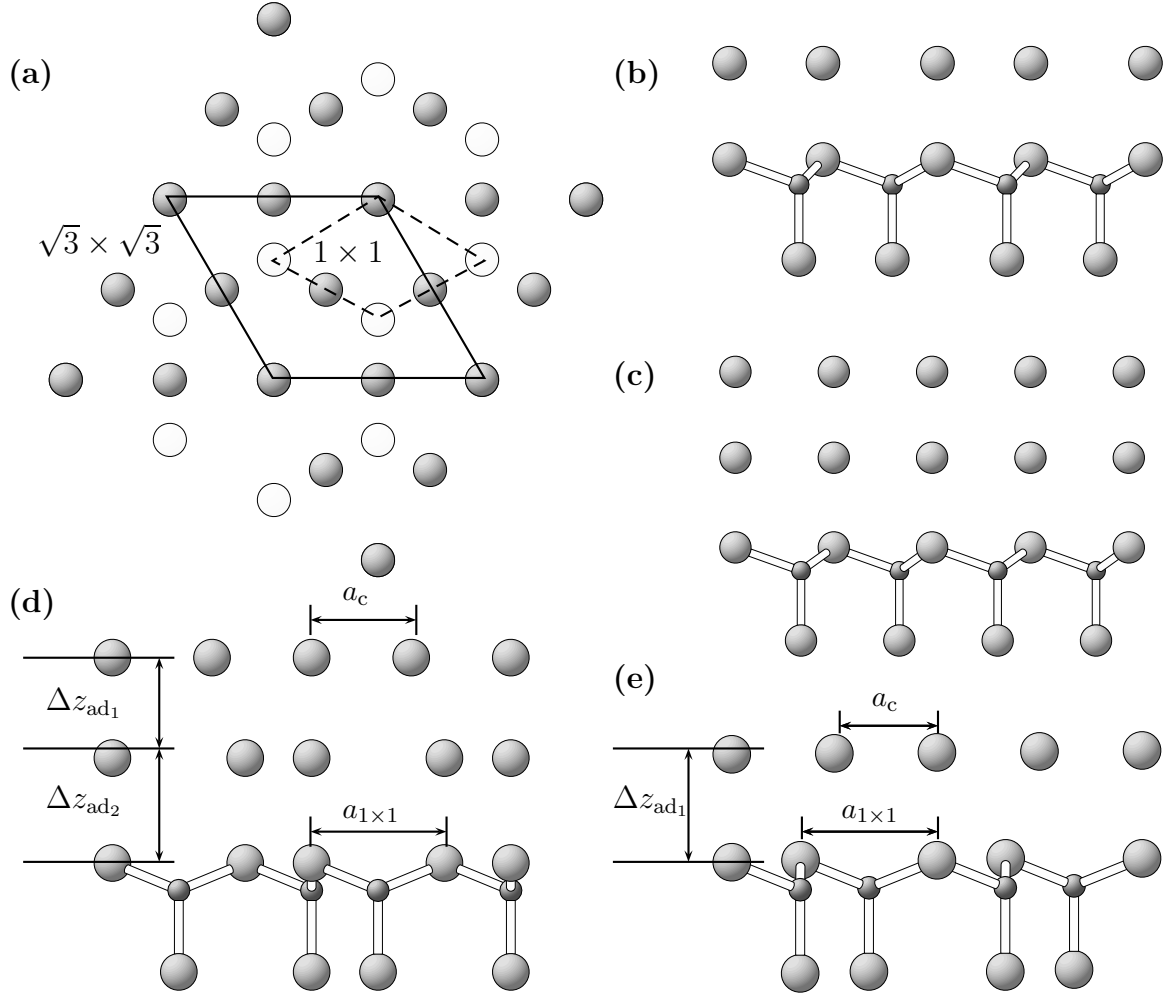


Figure 4.3: (a) Top view of the contracted Ga-adlayer and Ga-bilayer structures with the top layer (gray balls having a $(\sqrt{3} \times \sqrt{3})$ unit cell) and second layer (white balls having a (1×1) unit cell), (b) side view of the Ga-adlayer structure, (c) side view of the Ga-bilayer structure, (d) side view of the contracted Ga-bilayer structure and (e) side view of the contracted Ga-adlayer structure. Δz_{ad1} and Δz_{ad2} indicate the distance between the first and the second planes and between the second and third planes, respectively. In (c) and (e) the pictures were rotated by 30 degrees for ease of viewing.

and between the second and the third layer (2.50 \AA) are close to that of the nearest neighbor distance in the α -Ga bulk phase (2.44 \AA). For the contracted Ga-adlayer the distance between the Ga atoms in the first and the second layer is 2.47 \AA . These results are in very close agreement to the results of Ref. [36].

Our results using PBE are very similar to the results using LDA. For the contracted Ga-bilayer structures the spacing between the Ga atoms in the first and the second layer is 2.38 \AA and between the second and the third layer is 2.58 \AA . For the contracted Ga-adlayer the distance between the Ga atoms in the first and the second layer is 2.49 \AA .

Table 4.1: Calculated relaxations using LDA and PBE for the clean Ga- and N-terminated, Ga- and N-adatom structures, Ga-adlayer and Ga-bilayer surfaces as defined in Figs. 4.1(a)-(b), 4.2(b)-(e) and 4.3(a)-(b). d is the bond length between the adatom and the surface, Δz_{01} is the distance between the first and second plane for the bare surfaces, Δz_{ad1} is the distance between the adatom and the first plane for the adatom, Ga-adlayer and Ga-bilayer structures and Δz_{ad2} is the distance between the second and third planes for the Ga-bilayer structures. All values are given in Å.

LDA								
	unreconstructed		N-adatom		Ga-adatom		contracted	
	N-term	Ga-term	hcp	fcc	hcp	fcc	Ga-bilayer	Ga-adlayer
d			2.05	2.01	2.46	2.47		
Δz_{01}	1.92	0.66	0.62	0.70	0.65	0.80	0.68	0.68
Δz_{ad1}			1.13	1.51	1.57	1.57	2.37	2.47
Δz_{ad2}							2.50	
PBE								
d			2.09	2.05	2.50	2.53		
Δz_{01}	1.97	0.69	0.60	0.91	0.75	0.82	0.67	0.69
Δz_{ad1}			0.90	1.15	1.67	1.65	2.38	2.49
Δz_{ad2}							2.58	

4.2 Surface energy

To study the energetics of a surface it is important to note that the surface is typically not an isolated defect but that it is interacting with its environment. At usual growth temperatures adatoms/molecules can either adsorb to or desorb from the surface. Also, interaction is possible by diffusion from the surface to the bulk and vice-versa. If the surface is close to thermodynamic equilibrium (i.e. the growth rate is negligible) the surface energy can be expressed employing thermodynamic concepts [148]. Below we briefly describe the main aspects that relate surface energy and thermodynamic quantities.

Considering a system of m particles, being n_i the number of particles of type i , T the temperature and p the pressure. If any of these quantities change by a small amount during a process, then the change in the Gibbs free energy $G(T, p, n_1, n_2, \dots, n_m)$ during a process is given by

$$dG = \left(\frac{\partial G}{\partial T} \right)_{p, n_i} + \left(\frac{\partial G}{\partial p} \right)_{T, n_i} + \sum_i^m \left(\frac{\partial G}{\partial n_i} \right)_{p, T, n_j \neq n_i}. \quad (4.1)$$

From Maxwell's relations [149] we have

$$dG = -SdT + Vdp + \sum_i^m \left(\frac{\partial G}{\partial n_i} \right)_{p, T, n_j \neq n_i}. \quad (4.2)$$

The quantity $\left(\frac{\partial G}{\partial n_i}\right)_{T,p,n_j}$ is called chemical potential μ_i

$$\mu_i = \left(\frac{\partial G}{\partial n_i}\right)_{T,p,n_j}, \quad (4.3)$$

and for constant T and p we have

$$G = \sum_i^m n_i \mu_i. \quad (4.4)$$

Consider a system consisting of n_s atoms and surface energy γ_s with area \mathcal{A} in equilibrium with other phases, like atoms or molecules. Then, in thermodynamical equilibrium, we can write

$$dG = \sum_i^m \mu_i dn_i + d\gamma_s \mathcal{A} = 0. \quad (4.5)$$

Thus, for a surface of a crystal consisting of n_s particles A and B, we have

$$dG = \mu_s dn_s - \mu_A dn_A - \mu_B dn_B + d\gamma_s \mathcal{A} = 0, \quad (4.6)$$

where μ_s is the chemical potential of the crystal with surface. Thus, we obtain for the surface energy

$$\gamma_s \mathcal{A} = n_s \mu_s - n_A \mu_A - n_B \mu_B \quad (4.7)$$

In thermodynamic equilibrium it holds that the chemical potential of the crystal equals the chemical potential of the individual species. Thus, we have

$$G_{AB} = n_A \mu_A + n_B \mu_B. \quad (4.8)$$

It is important to note that the Gibbs free energy is dependent on the pressure and temperature. Therefore

$$G_{AB}(p, T) = n_A \mu_A(p, T) + n_B \mu_B(p, T). \quad (4.9)$$

For $p, T = 0$ the Gibbs free energy of the crystal is simply the free energy of the crystal. If zero point vibrations are also neglected the chemical potentials is the internal energy of the crystal $E_{\text{tot}}(\text{AB})$

$$\mu_{AB}(p, T) = E_{\text{tot}}(\text{AB}) \quad (4.10)$$

and thus the surface energy per area is given by

$$\gamma_s = \frac{E_{\text{tot}}(n_A, n_B) - \mu_A n_B - \mu_B n_B}{\mathcal{A}}. \quad (4.11)$$

The change of the entropy and volume are not explicitly included in the calculations. Typically, they are small and do not affect the results qualitatively at realistic growth conditions [150, 151]. The first part in Eq. (4.11) is solely determined by the

total energy calculations. The second part, however, depends on the choice of the chemical potentials. One condition is that the chemical potential for each species A, B must be smaller than the chemical potential of the corresponding elementary bulk/molecule as phase. Otherwise the system would be thermodynamically unstable. In particular, for GaN we have

$$\mu_{\text{Ga}} \leq \mu_{\text{Ga-bulk}} \quad \text{and} \quad \mu_{\text{N}} \leq \mu_{\text{N}_2\text{-molecule}}. \quad (4.12)$$

Combining Eqs. (4.12) and (3.5) (defined in Chapter 3) we obtain the range of thermodynamically allowed chemical potentials for N and Ga

$$\Delta H_f \leq \mu_{\text{N}} - \mu_{\text{N}_2\text{-molecule}} \leq 0 \quad (4.13)$$

and

$$\Delta H_f \leq \mu_{\text{Ga}} - \mu_{\text{Ga-bulk}} \leq 0. \quad (4.14)$$

To relate the formation enthalpy ΔH_f and the surface energy where T and p are essentially non-zero quantities, one needs to consider the corresponding temperature and pressure dependence of the chemical potentials. It can be shown that the experimentally relevant range of p and T the formation enthalpy changes by less than 0.05 eV [124]. Therefore, the temperature and pressure dependence will be disregarded.

The surface energy, by definition, is the energy necessary to create two equivalent surfaces by cleaving the crystal along a certain plane. However, for the GaN (0001) surface, it is worth noting that we cannot calculate the absolute surface energy since it is not possible to build up two equivalent surfaces cleaving the GaN crystal along the {0001} plane. Therefore, what we calculate is the relative surface energy [32]. For all further studies, we have chosen the clean Ga-terminated surface as reference.

Besides, to avoid interaction between the two surfaces of the slab one side of the slab is saturated with pseudo H atoms of fractional charge with $5/4e^-$. Doing so, we passivate the remaining dangling bonds of the N atoms on the back side of the slab. Thus, the N atom on the back side of the slab has three bonds with Ga atoms and one bond with pseudo H. Using the above considerations, the surface energy is expressed as

$$\gamma_s = E_{\text{tot}}(n_{\text{Ga}}, n_{\text{N}}) - \mu_{\text{Ga}} n_{\text{Ga}} - \mu_{\text{N}} n_{\text{N}} - \gamma_s^{\text{clean(Ga-term)}}(n_{\text{Ga}}, n_{\text{N}}). \quad (4.15)$$

As discussed above, we limit our calculations to $T = 0$, neglecting explicitly entropy and temperature effects, due to the fact that the entropy of crystals and surfaces is small compared to the gas phase. Implicitly, however, temperature and pressure are taken into account by the choice of the chemical potentials of the elementary phases. To estimate the dependence of μ_{N} on the temperature and pressure, let us take the translational, vibrational, and rotational contributions. The chemical potential of N is then given by

$$2\mu_{\text{N}} = E_{\text{tot}}^{\text{N}_2\text{-molecule}} + \frac{1}{2}h\nu + k_B T \ln \left(\frac{pV_Q}{k_B T} \right) - k_B T \ln Z_{\text{vib}} - k_B T \ln Z_{\text{rot}}, \quad (4.16)$$

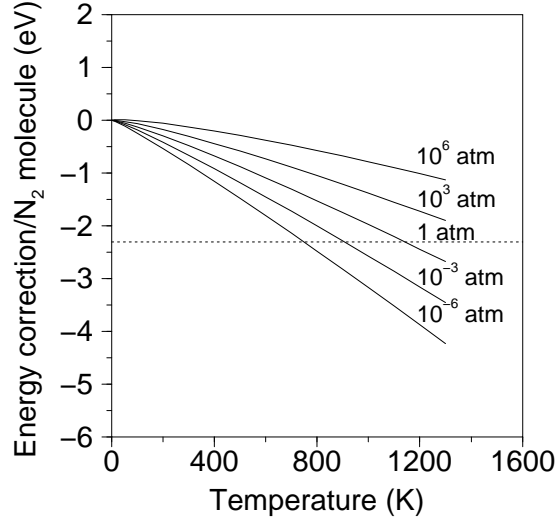


Figure 4.4: Chemical potential of a gas of N_2 -molecules as a function of temperature according to Eq. (4.16) for various pressures. The energy zero is set to the N_2 chemical potential at $p = 0, T = 0$. The dotted line represents the formation enthalpy of GaN multiplied by a factor of two.

where $E_{\text{tot}}^{N_2}$ is the total energy of the N_2 molecule at $p = 0$ and $T = 0$, ν is the zero point vibration frequency, $V_Q = (\hbar^2/2\pi m k_B T)^{3/2}$, k_B is the Boltzmann constant. Z_{vib} and Z_{rot} are the vibrational and rotational partition functions of the molecule given by

$$Z_{\text{vib}} = \frac{e^{\left(\frac{-\hbar\nu}{2k_B T}\right)}}{1 - e^{\left(\frac{-\hbar\nu}{k_B T}\right)}} \quad \text{and} \quad Z_{\text{rot}} = \frac{8\pi^2 \mu d^2 k_B T}{2\hbar^2}. \quad (4.17)$$

Here d is the bond length between the N atoms of the N_2 molecule. The vibration frequency ν can be calculated using the relation

$$\nu = \frac{1}{2\pi} \sqrt{\frac{K}{\mu}}, \quad (4.18)$$

where K is the force constant and μ is the reduced mass. The reduced mass μ was taken from the Ref. [134] and the vibration frequency and the force constant K were calculated theoretically from the curve $E_{\text{tot}}^{N_2-\text{molecule}}$ as a function of the N_2 -molecule bond length (see Table 3.7 in Chapter 3). The values used here are:

$$K = 23 \text{ N/m}, \quad \nu = 2363 \text{ cm}^{-1}, \quad m_N = 7 \text{ a.u.}$$

In Fig. 4.4 we plot the N_2 chemical potential as a function of the temperature (the typical growth temperature for MBE is 900 K and for MOCVD it is 1300 K),

including the last three terms in Eq. (4.16) for various pressures. This figure shows how it is possible to extract the N chemical potential for a certain temperature and pressure. Results reported by Gzregory and Porowski [12] have found that GaN is stable up to 1200 K at a pressure of 1 atm. According to the Fig. 4.4, at 1 atm we find that GaN decomposes at ~ 1140 K, in very good agreement with the experimental value.

Now we will analyze the surface energy for the bare GaN (0001) surfaces shown in Figs. 4.1 and 4.3. The optimization of the number of GaN layers and thickness of the vacuum region have shown that 9-11 layers and 11 \AA are sufficient².

In Figs. 4.5(a) and (b) we show the results for the relative surface energy of the clean Ga and N-terminated surfaces, Ga- and N-adatoms at *fcc* and *hcp* positions on the clean Ga-terminated, Ga-adlayer and Ga-bilayer structures. We compare PBE and LDA in order to verify whether they provide the same ordering (sequence) of surface energies. Our results for LDA are in excellent agreement with the LDA results from Refs. [31, 36]. This means that the energy difference between two particular structures, for example, agree within 10 meV/\AA^2 . Also, the energetical ordering of the structures are exactly the same as in Refs. [31, 36].

From Fig. 4.5(a), we see that under more Ga-rich conditions the Ga-adatom at the *hcp* position is the most stable structure. Under extreme Ga-rich conditions ($\mu_N = -1.25 \text{ eV}$) a structure consisting of a double layer on the top of the Ga-terminated surface (contracted Ga-bilayer) is the energetically favorable structure, in agreement with the theoretical results reported by Northrup *et al.* [36], who suggested that this structure might be the one observed experimentally under such growth conditions [32, 135, 140].

Under Ga-rich conditions the description using PBE is fully compatible with the LDA calculations, as we can see from Fig. 4.5(b), the sequence of structures is identical in LDA and PBE. Under N-rich conditions, however, LDA and PBE give qualitatively different results. While LDA predicts the N-adatom structure (on the *fcc* site) to be energetically most favorable, PBE predicts the Ga-adatom (on the *hcp* site) to be energetically preferred.

A closer look at Fig. 4.5 shows that the discrepancy between LDA and PBE is mainly due to the formation enthalpy, which defines the width of the phase diagram (the allowed range of chemical potentials). As has been pointed out the GaN formation enthalpy in PBE is significantly underestimated (-0.64 eV compared to experimental value of -1.15 eV).

This deficiency can be largely corrected if we use the experimental GaN formation enthalpy (rather than the PBE value). The allowed range extends then into the gray region in Fig. 4.5(b). Applying this shift the LDA results are reproduced. Thus, if we correctly shift the N-boundary to the experimental formation enthalpy, we can reproduce the correct ordering. In this sense, our results show that the only deficiency in PBE is the wrong description of the boundary (i.e. the binding energy

²The convergence of the surface energy with respect to the film thickness depends on how the inner layers of the film approach the GaN bulk. This is because we are interested in the surface energy itself and want to remove any contribution due to the bulk (see Eq. 4.11). Tests to determine the optimum number of layers and vacuum region thickness are presented in Appendix A.

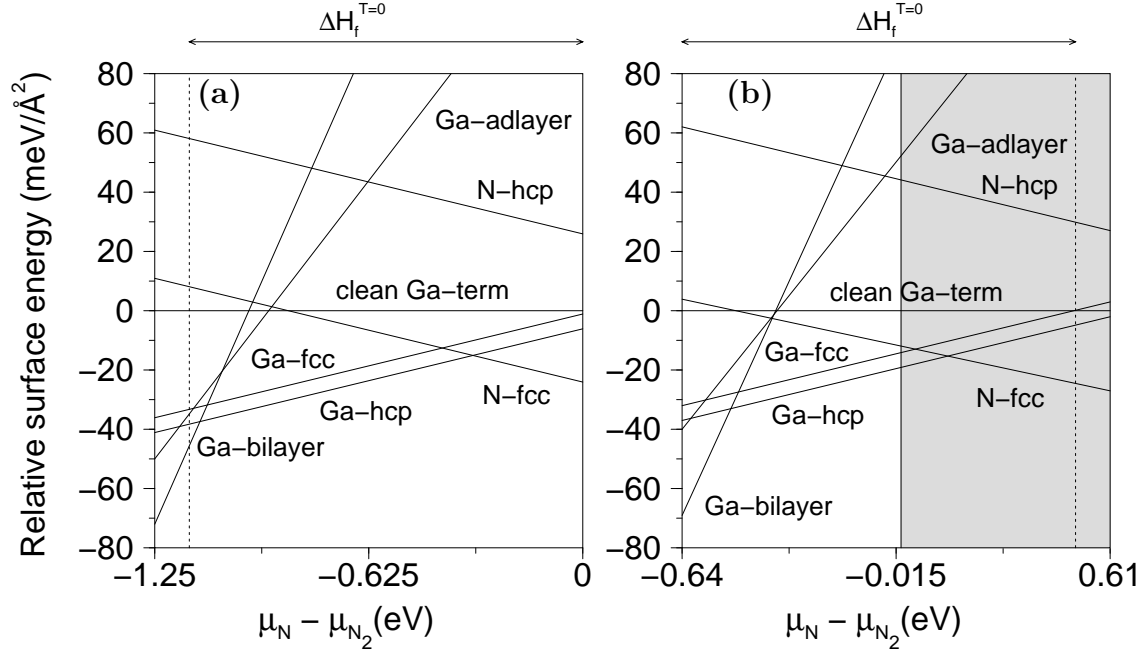


Figure 4.5: Relative surface energy per \AA^2 for bare GaN (0001) surfaces as a function of the N chemical potential μ_N . (a) using LDA and (b) using PBE. The labels have the following meaning: clean Ga-term refers to the clean Ga-terminated surface, N-*fcc*(*hcp*) refers to a structure with a N-adatom on the *fcc*(*hcp*) site of the clean Ga-terminated surface, Ga-*fcc*(*hcp*) refers to the a structure with a Ga-adatom on the *fcc*(*hcp*) site of the clean Ga-terminated surface, Ga-adlayer refers to the contracted Ga-adlayer surface and Ga-bilayer refers to the contracted Ga-bilayer surface (see Figs. 4.1 and 4.3). The energy zero is set to the clean Ga-terminated GaN surface. The surface area of the (1×1) unit cell is 8.84 \AA^2 for LDA and 9.16 \AA^2 for PBE. The experimental formation enthalpy of GaN ($\Delta H_f^{T=0}$) is indicated. The shaded region indicates the extrapolation to the experimental value of the theoretical PBE value of the formation enthalpy.

of the N_2 -molecule) under N-rich conditions.

4.3 Electronic properties

4.3.1 Band structure

Having calculated the equilibrium geometry for various bare GaN surfaces, we will now use these results to calculate the electronic structure of these surfaces. Let us first focus on the clean Ga-terminated surface (see Fig. 4.1(a)). Simple electron counting arguments show that this surface has one Ga dangling bond in the top layer, which is partially filled with $7/4$ of an electron. The existence of a partially filled state means that this state gives rise to a metallic surface with the Fermi energy crossing the surface state. In Figs. 4.6(a)-(b) we show calculated surface band structures.

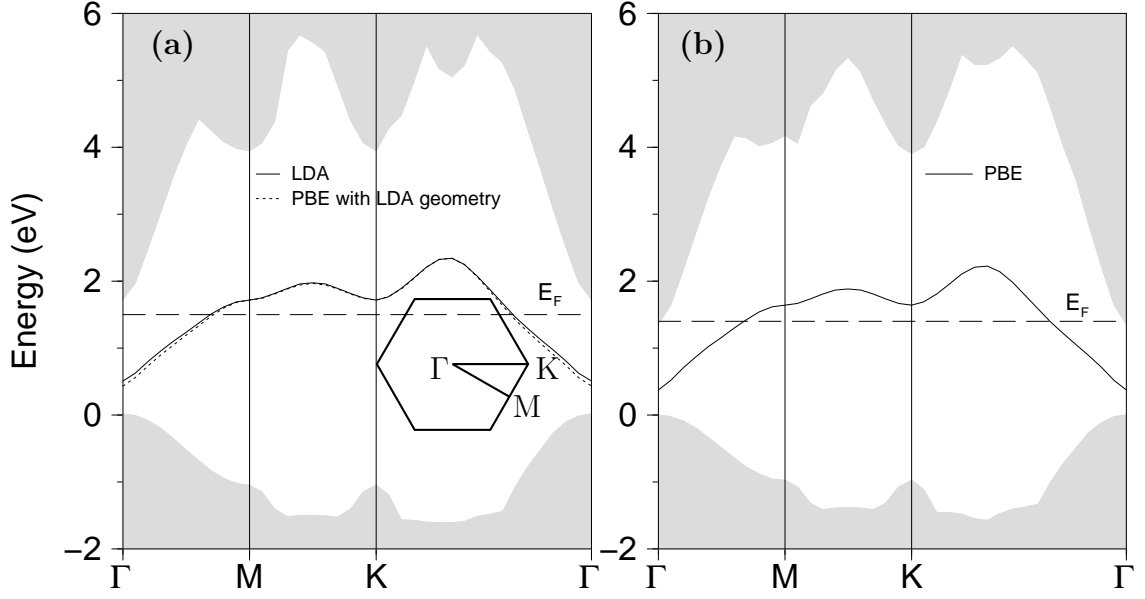


Figure 4.6: Band structure of the clean (unreconstructed) GaN (0001) Ga-terminated surface using (a) LDA (solid line) and PBE with LDA optimized geometry (dotted line). (b) shows the band structure using PBE calculation where the atomic geometry has been optimized using PBE. The inset in (a) shows the first surface Brillouin zone. The dashed lines indicates the position of the Fermi level E_F . The shaded region represents the projected bulk-band structure.

In Fig. 4.6(a) the calculated band structure using LDA and an optimized geometry (with respect to the lattice constant) is shown. Fig. 4.6(b) shows the equivalent result but using PBE. We can see that both exchange-correlation potentials give an almost identical dispersion for the surface state. Also, the position of the surface state above the top of the valence band of the projected bulk GaN band structure is the same for both calculations.

The main difference is that the band gap using LDA (1.7 eV) and PBE (1.4 eV) differs by 0.3 eV. The difference between LDA and PBE results may have two origins: *i*) electronic effects and *ii*) structural effects, since PBE leads typically to slightly larger bond lengths. To separate the two contributions we performed also a calculation within PBE but taking the geometry as optimized for LDA. Comparing this calculation with the fully optimized LDA allows to eliminate structural effects. As can be seen in Fig. 4.6 (a) the differences are smaller than 0.01 eV. We can therefore conclude that the differences (solid versus dotted line) in the electronic structure are primarily due to the effect of the exchange-correlation potential on the atomic structure.

Now we want to analyse the band structure of the clean Ga-terminated, Ga-adatom, contracted Ga-adlayer and contracted Ga-bilayer surfaces. Figs. 4.7(a)-(d) show the band structure of these surfaces. For the clean Ga-terminated surface, as we already mentioned, there is one partially occupied surface state in the band gap due to the Ga dangling bond. This result is in good agreement with what is found in

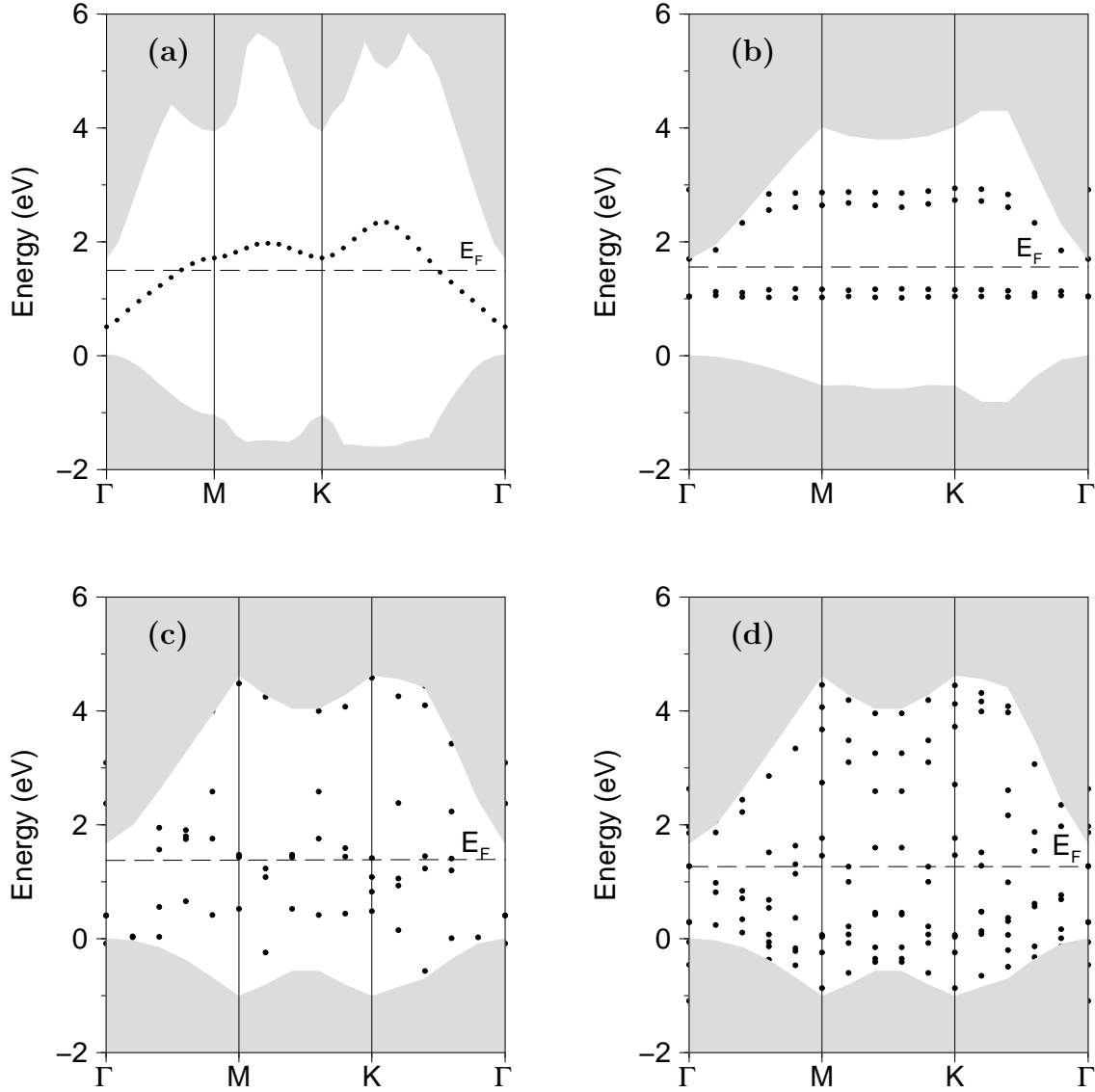


Figure 4.7: Band structure of the (a) clean Ga-terminated surface, (b) Ga-adatom, (c) contracted Ga-adlayer and (d) contracted Ga-bilayer structures. The shaded region shows the projected GaN bulk band structure. The surface states are the black dotted points. The dashed line indicates the Fermi level.

Ref. [143]. The Ga-adatom structure has four surface states in the band gap, as we can see in Fig. 4.7(b). The two empty states are due to the two empty Ga-dangling bonds at the surface. The two occupied states are due to the three back bonds of the Ga atoms in the first layer.

The contracted Ga-adlayer and Ga-bilayer structures have metallic character, as we can see in Figs. 4.7(c) and (d). The surface states are due to the Ga adlayers. From these band structures we can see that the metallization of the surface increases with increasing Ga coverage, as expected.

4.3.2 Ionization energy, electron affinity and work function

One of the measurable and calculable surface properties are the work function ϕ , ionization energy I (or photoelectric threshold) and electron affinity χ . Those quantities are relatively easy to measure and allow to monitor changes in the surface geometry. However, so far there is no relation between work function and surface structure for GaN surfaces available. Therefore it is worth to study how these quantities change with the stoichiometry of the surface.

In the following we use a procedure described in Ref. [152] to calculate the work function, electron affinity and ionization energy. We start with the definition of these quantities. The main idea is to combine bulk and slab calculations. In the bulk calculation, the top of the valence band E_v^{bulk} is fixed relative to the bulk potential $V_{\text{els}}^{\text{bulk}}$.

The slab calculation determines the bulk potential relative to the vacuum level. Once the slab is sufficiently thick that both vacuum and bulk regions are well described, the electrostatic potential for the central bulk-like layer in the slab $V_{\text{els}}^{\text{slab(bulk)}}$ is identical to the potential in the bulk calculation except by a constant shift (see Fig. 4.8). Using the results of the bulk calculation, which fixes the band energies relative to the bulk potential, we chose the potential so that the calculated top of the bulk valence band is at the energy zero.

The ionization energy is the energy difference between the vacuum level and the valence band maximum, i.e. it is the minimum energy which is necessary to lift one electron from the highest occupied state to the vacuum level. Based on it, the ionization energy is calculated as

$$I = V_{\text{els}}^{\text{vacuum}} - E_{v(\text{bulk})} - E_{\text{shift}} = V_{\text{els}}^{\text{vacuum}} - E_{v(\text{slab})}, \quad (4.19)$$

where $V_{\text{els}}^{\text{vacuum}}$ is the electrostatic potential in the vacuum region, $E_{v(\text{bulk})}$ is the energy of the top of the valence band in the bulk, $E_{v(\text{slab})}$ is the energy of the top of the valence band in the slab and $E_{\text{shift}} = V_{\text{els}}^{\text{slab(bulk region)}} - V_{\text{els}}^{\text{bulk}}$ is the difference between the electrostatic potential in the slab in the bulk region and the electrostatic potential of the bulk. A schematic picture of this procedure for the ionization energy is shown in Fig. 4.8.

For metals, the energy difference between the vacuum level and the Fermi level is defined as the work function. The work function of a metal is attributed to the atomic binding energies and the surface dipole [153]. In the simplest sense, the surface dipole is a quantum mechanical effect that is attributed to the fact that the wave function of the electrons extends beyond the positive ion background. This results in excess negative charge at the surface. Just below this excess negative charge will be an unbalanced positive charge due to the ion background. The two charge sheets will form a dipole at the surface. For a semiconductor the same process may be involved, but the directional bonding at the surface can lead to even larger effects (surface reconstructions and adsorbate layers, for instance). For example, filled dangling bond type surface states will contribute with a negative charge at the surface which is balanced by a positive charge nearby. The work function is defined as

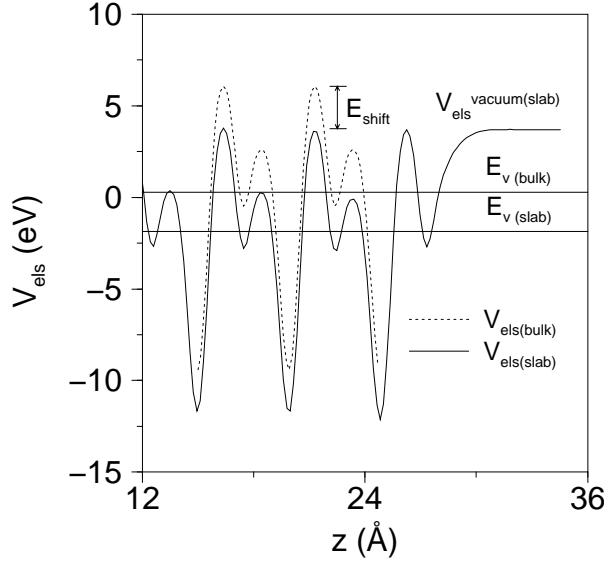


Figure 4.8: Averaged electrostatic potential for the clean surface and GaN-bulk parallel to the surface normal. $E_{v(\text{bulk})}$ is the top of the valence band in the bulk, $V_{\text{els}}^{\text{vacuum}}$ is the electrostatic potential in the vacuum region and E_{shift} is the shift of the top of the valence band in the slab with respect to the bulk. $E_{v(\text{slab})}$ is the calculated top of the valence band in the slab.

$$\phi = V_{\text{els}}^{\text{vacuum}} - E_{\text{F}}, \quad (4.20)$$

where E_{F} is the Fermi level.

The electron affinity relates the vacuum level to the conduction band minimum at the surface being calculated as

$$\chi = I - E_{\text{gap}}, \quad (4.21)$$

where E_{gap} is the bulk band gap.

A schematic picture showing the quantities defined in Eqs. (4.19), (4.20) and (4.21) is shown in Fig. 4.9. We calculated the quantities defined above for the clean Ga- and N-terminated, Ga-adatom, Ga-adlayer (contracted and non-contracted) and Ga-bilayer (contracted and non-contracted) surfaces. The results are presented in Table 4.2. We can see from our results that all the calculated properties change quite significantly depending on the stoichiometry of the surface. The main difference comes when going from the clean N-terminated surface to the clean Ga-terminated surface. So far, detailed surface investigations regarding surface reconstructions, termination, stoichiometry, impurities and dopants for the investigated surfaces are missing. Thus, only a very tentative comparison is possible. Experimentally, values for the electron affinity range between 2.1 and 4.1 eV. More recent experiments are in the range of 3.1 – 3.5 eV [25, 154, 155].

Grossner *et al.* [156] have performed first-principles calculations for the (111) face of the GaN in cubic phase [156]. Since the [111] direction corresponds to the

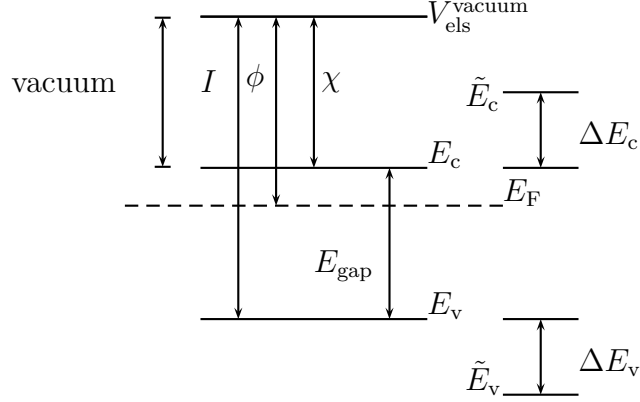


Figure 4.9: The definition of work function ϕ , electron affinity χ and ionization energy I . E_F is the Fermi energy and E_{gap} is the energy gap. $V_{\text{els}}^{\text{vacuum}}$ is the electrostatic potential in the vacuum, ΔE_c and ΔE_v the quasi-particle corrections for the conduction band minimum and for the valence band maximum [156]. E_c is the conduction band minimum, E_v the valence band maximum, \tilde{E}_c the quasi-particle corrected conduction band minimum and \tilde{E}_v the quasi-particle corrected valence band maximum.

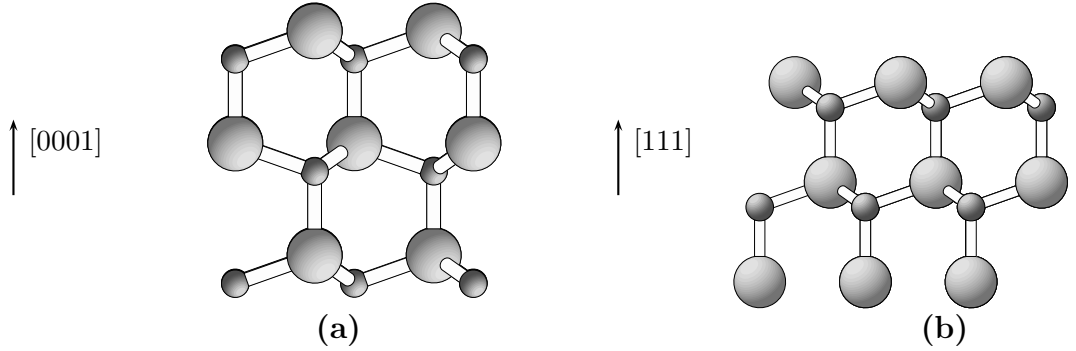


Figure 4.10: Schematic picture showing the stacking sequence of the wurtzite and zincblende GaN structures. a) along [0001] and b) along [111] directions.

$\{c\}$ -axis in the wurtzite structure parallel to the hexagonal [0001] direction, the (111) surfaces could provide a direct comparison with our results. However, we should keep in mind that our clean Ga-terminated surface should be compared with their clean Ga-terminated surface, our non contracted Ga-adlayer structure should be compared with their structure with 1 ML Ga, and our non contracted Ga-bilayer structure should be compared with their structure with 2 ML Ga (see Table 4.2).

Also, while we study the clean N-terminated surface, they have one layer of N adsorbed on the clean Ga-terminated surface, which leads to a slightly different surface, because wurtzite and zincblende structures have different stacking, as shown in Fig. 4.10. Thus, a direct comparison between their N terminated structure and ours is not straightforward.

Table 4.2: Calculated and experimental work function ϕ , electron affinity χ and ionization energy I for the bare GaN (0001) surfaces shown in Figs. 4.2(b)-(e) and 4.3(b)-(c). The values in brackets are calculated using the quasi-particle corrections for the valence band maximum (-0.74 eV) and conduction band minimum (0.44 eV) from Ref. [156]. All values are given in eV.

Ref.	Surface	χ (eV)	ϕ (eV)	I (eV)
theory				
this work	(0001) clean N-terminated	7.82 (7.38)	9.06	9.52 (10.26)
this work	(0001) clean Ga-terminated	4.47 (4.03)	4.42	6.17 (6.91)
this work	(0001) Ga-adatom hcp	4.17 (3.73)	4.20	5.87 (6.61)
this work	(0001) Ga-adlayer(non-contracted)	4.67 (4.23)	5.31	6.37 (7.11)
this work	(0001) Ga-adlayer(contracted)	3.20 (2.76)	3.40	4.72 (5.46)
this work	(0001) Ga-bilayer(non-contracted)	3.97 (3.53)	4.80	5.67 (6.41)
this work	(0001) Ga-bilayer(contracted)	2.42 (1.98)	3.10	4.12 (4.86)
[156]	(111) clean N-terminated	4.64		7.74
[156]	(111) clean Ga-terminated	2.49		5.59
[156]	(111) with 1 ML Ga	1.79		4.89
[156]	(111) with 2 ML Ga	2.02		5.12
[156]	(111) with 3 ML Ga	1.78		4.88
[157]	(0001) clean		4.5	
exp.				
[155]	(0001) <i>n</i> -type	3.50±0.10	4.30±0.10	6.90
[155]	(0001) <i>p</i> -type	3.50±0.10	4.30±0.10	6.90
[158]	<i>n</i> -type	4.10		
[158]	intrinsic	2.10		
[156]				6.80
[154]		3.10±0.20		
[159]	<i>n</i> -type		3.88	
[160]	Ga-face			
[161]			4.30±0.15	
[25]		3.2		
[162]	Ga-bulk		4.20	

It is expected that as the Ga coverage increases, there is a metallization of the surface: the ionization energy should converge to the work function of the elementary phase of bulk Ga (which is metallic). According to Table 4.2 we do not find a clear trend as the number of Ga atoms (or Ga coverage) increases at the surface. The experimental value of the Ga metal work function is 4.2 eV [Fomenko1981] and our results show a better agreement for the Ga-adatom structure work function.

However we should keep in mind that the melting point of Ga is 300 K and thus the additional Ga layers have a liquid-like behavior on the top of the Ga-terminated surface. Thus, a comparison of the calculated quantity and the experimental value

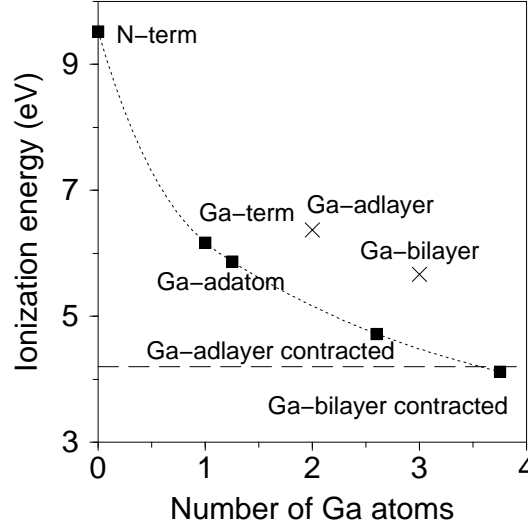


Figure 4.11: Calculated ionization energy as a function of the Ga coverage for the bare GaN (0001) surfaces shown in Figs. 4.1, 4.2, and 4.3. The dotted line is a guide for the eyes. The crossed indicate the thermodynamically not stable structures Ga-adlayer and Ga-bilayer and the dashed line the work function for Ga bulk.

for the work function of Ga bulk may not be direct. In fact, it is well known that the non contracted Ga-adlayer and Ga-bilayer structures are thermodynamically unstable [36].

This might explain the oscillatory behavior of the ionization energy and electron affinity as the number of Ga atoms increases, which is similar to the one found in Ref. [156]. Indeed, if we do not consider the thermodynamically not stable non-contracted Ga-adlayer and Ga-bilayer structures, we can see a clear trend: with increasing Ga coverage the ionization energy and work function decrease monotonically, as shown in Fig. 4.11.

Although we can find a qualitative trend for the calculated properties, we should mention that the main problem in determining the ionization energy and electron affinity is that DFT-LDA/GGA does not provide the correct band gap. Grossner *et al.* [156] included quasi-particle corrections for the top of the valence band ΔE_v (for the ionization energy) and the minimum of the conduction band ΔE_c (for the electron affinity) to correct the gap. Our DFT-LDA band gap is 1.7 eV. We use this value together with the quasi-particle corrections given in Ref. [156]. This gives us a band gap equal to 2.9 eV, which is closer to the experimental value (3.4 eV). With those values we re-calculated the properties listed in Table 4.2 (values in brackets). Generally that correction brings our results closer to the results of Ref. [156]. However, the agreement is not complete, which is not expected, since the structures are not completely equivalent.

4.4 Summary

With the results presented in this chapter we showed that LDA and PBE give very similar results for the atomic relaxations and electronic structure of the surfaces. We showed that PBE does not provide the same energetical ordering of structures, but that this is due to the underestimation of the GaN formation enthalpy. If the formation enthalpy is corrected, the LDA ordering is exactly reproduced. Also, our results clearly showed that the position of the surfaces states is not affected by the choice of the exchange-correlation functional. Therefore, LDA is our choice to study Si adsorption on GaN surfaces, as we will describe in the next chapter.

Concerning the band structure of the bare surfaces, we found a good agreement between our results and other theoretical results for the clean Ga-terminated surface. We also found that the surface becomes metallic as the number of Ga layers on the surface increases.

Concerning the ionization energy, electron affinity and work function, a direct comparison with experiment is not straightforward. This stems from the difficulty in determining experimentally the stoichiometry and termination of the surface, which strongly affect the surface properties. We also argue that the problem in determining correctly these properties lies on the wrong DFT/LDA-GGA description of the band gap, although we have shown that the quasi-particle corrections do not really bring the results close to the experimental values.

Chapter 5

Si on GaN (0001) surfaces

In this Chapter we will analyze and discuss our results for adsorption of Si on the GaN (0001) surface. Before the analysis we will discuss some aspects that will be helpful in order to interpret our results. We first briefly discuss some thermodynamic aspects of growth modes. Then we discuss the Si incorporation in GaN-bulk. We also give an overview about

According to the classification of Bauer [163], where thermodynamic equilibrium is assumed, there are three growth modes. Both lattice strain and surface free energy help to determine whether a film undergoes layer-by-layer growth followed by islanding (Stranski-Krastanov) [164], layer-by-layer growth (Frank-Van der Merwe) [165] or islanding (Volmer-Weber) [166]¹ Fig. 5.1 shows schematically these three growth modes. We should mention, however that experimentally, the terminology is often used in a pure phenomenological sense, since growth involves kinetic processes. Theoretical models of epitaxial growth suggest that the growth model is determined by the free energy of substrate surface γ_s , the interface free energy γ_i , and the surface free energy of the heteroepitaxial layer γ_f (neglecting the strain energy of the film).

- The inequality $\gamma_s > \gamma_i + \gamma_f$ sets the condition for the epitaxial film to wet the substrate. In this case Frank-Van der Merwe (FvdM) growth may occur (see Fig. 5.1(a)). The surface will remain flat, as it minimizes the total surface area. If adatom incorporation occurs either at step edges or kink sites, then layer-by-layer growth results, as shown in Fig. 5.2(b)-(c). Otherwise, if nucleation does not occur, step-flow growth takes place, as shown in Figs. 5.2(d).
- If the inequality has the opposite sign, i.e., $\gamma_s < \gamma_i + \gamma_f$, one usually obtains Volmer-Weber (VW) growth as shown in Fig. 5.1(c). The heteroepitaxial layers will thus not wet the substrate but form droplets or clusters on the surface.
- The Stranski-Krastanov (SK) growth generally occurs when there is wetting of the substrate but the overlayer strain is unfavorable. In such a case, layer-by-layer growth will take place for a few monolayers, followed by island formation as shown in Fig. 5.1(b).

¹A complete description of this subject can be found in Refs. [167, 168].

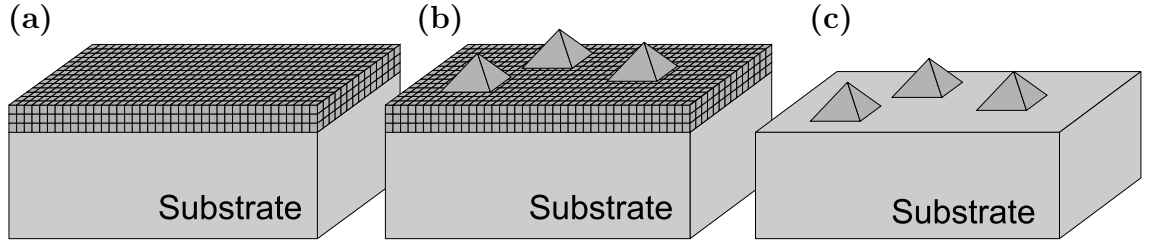


Figure 5.1: Schematic picture showing the growth modes a) Frank-van der Merwe (layer-by-layer mode), b) Stranski-Krastanov (layer-by-layer followed by island formation) and c) Volmer-Weber (island mode).

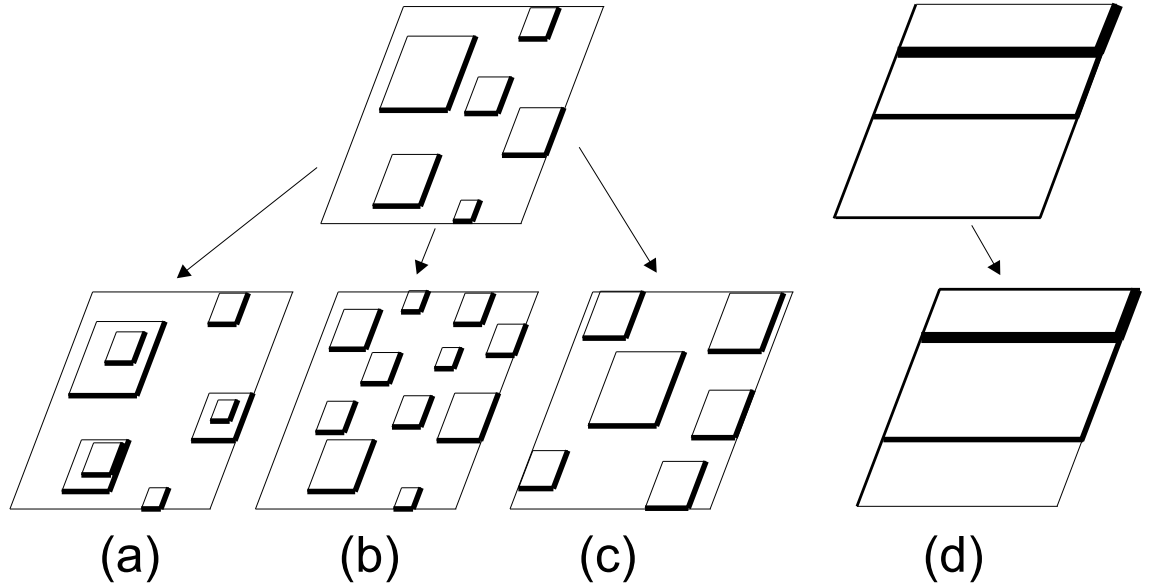


Figure 5.2: Schematic picture showing (a) multilayer growth, (b) layer-by-layer growth with high nucleation rate, (c) layer-by-layer-growth with low nucleation rate and (d) step-flow growth.

Although most experimental growth studies are made on high symmetry surfaces, where growth is dominated by nucleation of adatom clusters, it would seem that the best way to grow a good crystal is to use a substrate cut along a direction vicinal to a high symmetry orientation, i.e., a vicinal surface [169]. In this way, by choosing proper experimental conditions, the crystal will grow through a regular flow of the steps, avoiding all problems related to random nucleation. A schematic view of possible multilayer modes is shown in Fig. 5.2(a)-(d).

5.1 Si doping in GaN-bulk

Si can be incorporated in GaN-bulk at very high concentrations (10^{20}cm^{-3}) and behaves as a shallow donor, with binding energy of 30 meV [38]. Before starting the description of our results for Si adsorbed on GaN surfaces, we will describe some results for incorporation of Si in GaN-bulk. These results will be useful when analyzing the concentration of Si at the surface, subsurface and in bulk.

We therefore calculate Si incorporation in GaN-bulk, where Si replaces a Ga atom (Si_{Ga}). Configurations with Si at interstitial positions or at the N site have been found to be energetically unfavorable [20]. This can be understood considering that Si has an atomic radius (1.1 Å) very similar to Ga (1.26 Å). Si on the N site or interstitially causes a large strain, because Si has a much larger atomic radius than N (0.75 Å). The calculations were performed using a super-cell with 64 atoms containing 1 defect, i.e., 1 Si atom. A 64 atom supercell has been found to be sufficiently large to describe an isolated impurity.

To calculate the concentration of substitutional Si_{Ga} in GaN-bulk we have to take into account that the corresponding reaction is not stoichiometric: first a Ga atom has to be removed (i.e., a Ga vacancy is created) and then a Si is incorporated on the vacancy site. To be more specific, let us consider the chemical potential of a vacancy

$$\mu_{\text{vacancy}} = \frac{\partial G(N_A, N_B : V_A^C)}{\partial C}, \quad (5.1)$$

where C is the number of vacancies of type A and G is the Gibbs energy of an ideal crystal containing the N_A and N_B host atoms. At large defect-defect distances (i.e. $C \ll N$), the interaction between dopants can be neglected. Then the energy to create C vacancies is just the number of vacancies times the energy to create one isolated vacancy. For a substitutional impurity in GaN, the number of Ga or N sites is $4.4 \times 10^{22} \text{cm}^{-3}$. In thermodynamic equilibrium the systems are connected by the condition of constant chemical potential. If a difference between the chemical potential in two open subsystems exists, then an exchange of particles occurs, which reduces and eventually eliminates the chemical potential difference. In thermodynamic equilibrium the chemical potential of the vacancy and of the thermodynamic reservoir must be identical. Thus, the vacancy concentration is given by

$$C = N_{\text{sites}} e^{-E_f^{\text{D}}/k_B T} = 4.4 \times 10^{22} \text{cm}^{-3} e^{-E_f^{\text{D}}/k_B T}, \quad (5.2)$$

where k_B is the Boltzmann constant, T is the temperature. E_f^{D} denotes the defect formation energy and it is given by

$$E_f^{\text{D}} = \Delta G_f^{\text{D}} + \sum_i n_i \mu_i, \quad (5.3)$$

with i running over all chemical species n_i with chemical potential μ_i involved in the creation of the defect. The number of atoms of species i is positive if one atom is removed from the host and negative if it is added. In the case of Si_{Ga} (denotes one Si atom replacing substitutionally one Ga atom) we have

$$E_f^D = \Delta G_f^D + n_{\text{Ga}}\mu_{\text{Ga}} - n_{\text{Si}}\mu_{\text{Si}}. \quad (5.4)$$

At $p, T = 0$, ΔG_f^D becomes

$$\Delta G_f^D = E_{\text{tot}}^{\text{GaN-bulk}} - E_{\text{tot}}^{\text{GaN-bulk}}(\text{Si}_{\text{Ga}}), \quad (5.5)$$

where $E_{\text{tot}}^{\text{GaN-bulk}}$ and $E_{\text{tot}}^{\text{GaN-bulk}}(\text{Si}_{\text{Ga}})$ are the total energy of the GaN-bulk without and with a defect. Assuming thermodynamic equilibrium, the N and Ga chemical potentials are not independent, but related by $\mu_{\text{GaN}} = \mu_{\text{Ga}} + \mu_{\text{N}}$, where μ_{GaN} is the GaN chemical potential. For Si in GaN bulk, the following phases can limit the solubility: Si can form droplets on the surface or diffuse into the bulk forming Si/N and/or Si/Ga compounds. All these processes have to be considered, since the formation of such compounds is not desired. Thus, the Si chemical potential is imposed to obey the lower bound

$$\mu_{\text{Si}} \leq \mu_{\text{Si-bulk}}. \quad (5.6)$$

Other upper bounds may exist since Si may form parasitic phases with N or Ga. Si/Ga compounds are not known in nature. However, several Si/N compounds are known. Since Si_3N_4 is a very stable compound, we consider also the following bound

$$3\mu_{\text{Si}} + 4\mu_{\text{N}} \leq \mu_{\text{Si}_3\text{N}_4}. \quad (5.7)$$

Combining Eqs. (5.6) and (5.7), we obtain that the limit under N-rich conditions is

$$\mu_{\text{Si}} - \mu_{\text{Si-bulk}} = \frac{1}{3}\Delta H_f^{\text{Si}_3\text{N}_4} = -3.32 \text{ eV}, \quad (5.8)$$

and under Ga-rich conditions

$$\mu_{\text{Si}} - \mu_{\text{Si-bulk}} = \frac{1}{3}\Delta H_f^{\text{Si}_3\text{N}_4} - \frac{4}{3}\Delta H_f^{\text{GaN}} = -1.66 \text{ eV}. \quad (5.9)$$

In Fig. 5.3 we plot the Si concentration in GaN bulk as a function of the Si (μ_{Si}) and N (μ_{N}) chemical potentials for typical temperatures of MBE and MOCVD growth. The concentration increases going from N-poor (Ga-rich) to N-rich conditions if the Si chemical potential is kept constant. The solubility is limited by the formation of Si_3N_4 . The region where the system is unstable against Si_3N_4 is marked by the gray area in Figs. 5.3(a)-(b).

The maximum Si concentration which can be incorporated in GaN increases with temperature. However, with increasing temperature the Ga vacancy concentration increases and approaches the Si concentration, as shown by Neugebauer *et al.* [20]. Since the Ga vacancy is an acceptor, it will partially compensate the Si donors. It is interesting to note that the maximum solubility is achieved under N-poor conditions, because N-rich conditions promote the formation of Si_3N_4 .

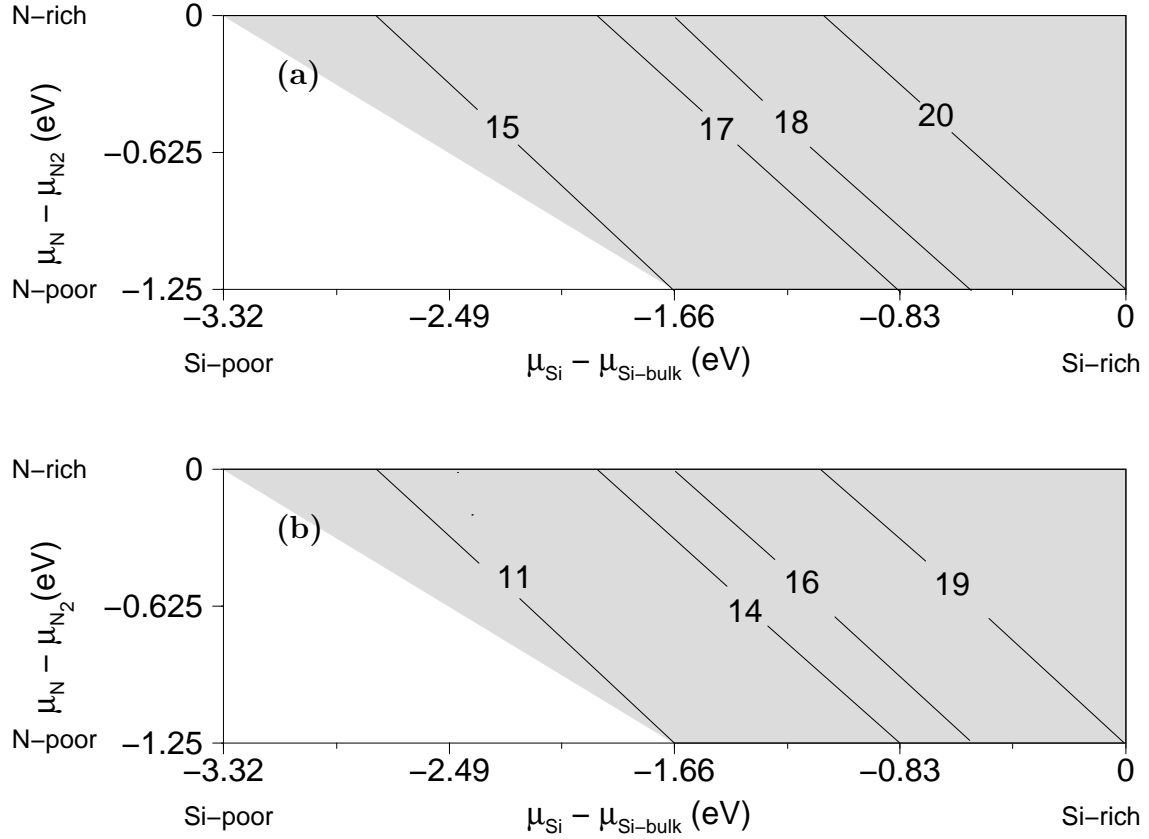


Figure 5.3: Si concentration as a function of the Si (μ_{Si}) and N (μ_{N}) chemical potentials for typical temperatures of (a) MOCVD (1300 K) and (b) MBE (900 K) growth. The solid lines represent the Si concentration given in $\log_{10} \text{cm}^{-3}$. The shaded region shows the allowed region where the formation of Si_3N_4 is thermodynamically stable.

5.2 Si doping on GaN (0001) surfaces

5.2.1 MOCVD/MOVPE growth

Experimentally, depending on the growth conditions (for instance, temperature and doping concentration) and on the growth techniques, Si doping affects in various ways the surface morphology, as we will describe in the following. MOCVD studies performed by Tanaka *et al.* [41] reported that Si concentrations in the range of 10^{18} - 10^{19}cm^{-3} during MOCVD modify the GaN growth from a step-flow mode to three-dimensional, suggesting that Si may act as an *anti-surfactant*. Under certain experimental conditions it might even lead to the formation of QDs. Such dots are shown in the Atomic Force Microscopy (AFM) images in Fig. 5.4.

In this experiment, a $\text{Al}_x\text{Ga}_{1-x}\text{N}$ ($x \approx 0.15$) buffer layer was grown on the Si face of the AlN/SiC (0001) substrate. The $\text{Al}_x\text{Ga}_{1-x}\text{N}$ surface was treated with tetraethyl-silicon [$\text{Si}(\text{C}_2\text{H}_5)_4$ (TESi)] prior deposition of an amount of GaN equiva-

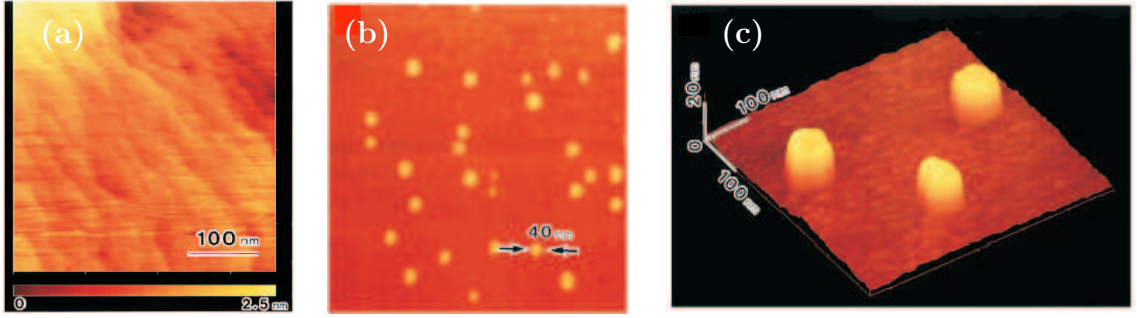


Figure 5.4: Atomic Force Microscopy images showing the effect of Si doping on GaN surfaces. a) GaN growth on the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ buffer layer (step-flow growth mode), b) QDs formed doping the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ buffer layer with TESI before the growth of GaN and c) zoom of the QDs shown in b). Such QDs have an average width of 400 Å and a height of 60 Å. The density of dots is $3 \times 10^9 \text{ cm}^{-2}$. After Tanaka *et al.* [170].

lent to an approximately 10 Å thick uniform layer. By adding a small amount of Si on the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ surface, the growth mode was changed from step-flow (smooth) to three-dimensional mode (rough).

In Fig. 5.4(a) we see a step-flow growth-mode, for GaN on a $\text{Al}_x\text{Ga}_{1-x}\text{N}$ buffer layer. The growth is smooth, with large terraces. In Fig. 5.4(b) we can see the formation of small dots once Si is used to dope the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ buffer layer before growing the GaN epilayer. Depending on the Si concentration, the surface shows different degrees of roughness. Once the Si doping concentration exceeds a critical value of 32 nmol (1 nmol of TESI corresponds to $\sim 0.001 \text{ ML}$), the roughness enhances and the spontaneous formation of QDs takes place. The dots of Fig. 5.4(b) and (c) have an average width of 400 Å and height of 60 Å. The density of dots was found to vary over the range of 10^7 - 10^{11} cm^{-2} and to depend on the Al, Ga and N contents, the Si doping and also the temperature.

Studies using MOVPE (which has a larger growth rate than MOCVD, but uses hydrides as precursors as well) confirm these results [42]. Indeed, MOCVD growth has shown that Si concentrations above $1 \times 10^{19} \text{ cm}^{-3}$ induce roughness [43] and crack formation [44]. However, at low temperature MOCVD growth ($\approx 1070 \text{ K}$), Si adsorption leads to smooth surfaces even for high Si concentrations ($3 \times 10^{19} \text{ cm}^{-3}$) [171]. In the latter case, the adsorption of Si appeared to change the surface mobility of the Ga species, resulting in a larger average terrace length. Munkholm *et al.* [45] have also reported that Si has a strong effect on the growth mode. They have found that at concentrations above $2 \times 10^{19} \text{ cm}^{-3}$ Si segregates to the surface and changes the growth mode from step-flow to layer-by-layer over a large range of growth temperatures (890-1220 K).

5.2.2 MBE growth

Si doping on GaN surfaces using MBE growth has been shown to lead to smooth surfaces. Recent STM measurements done by Lee *et al.* [39, 40] after MBE growth of GaN (0001) surfaces demonstrated that the incorporation of Si does not affect the

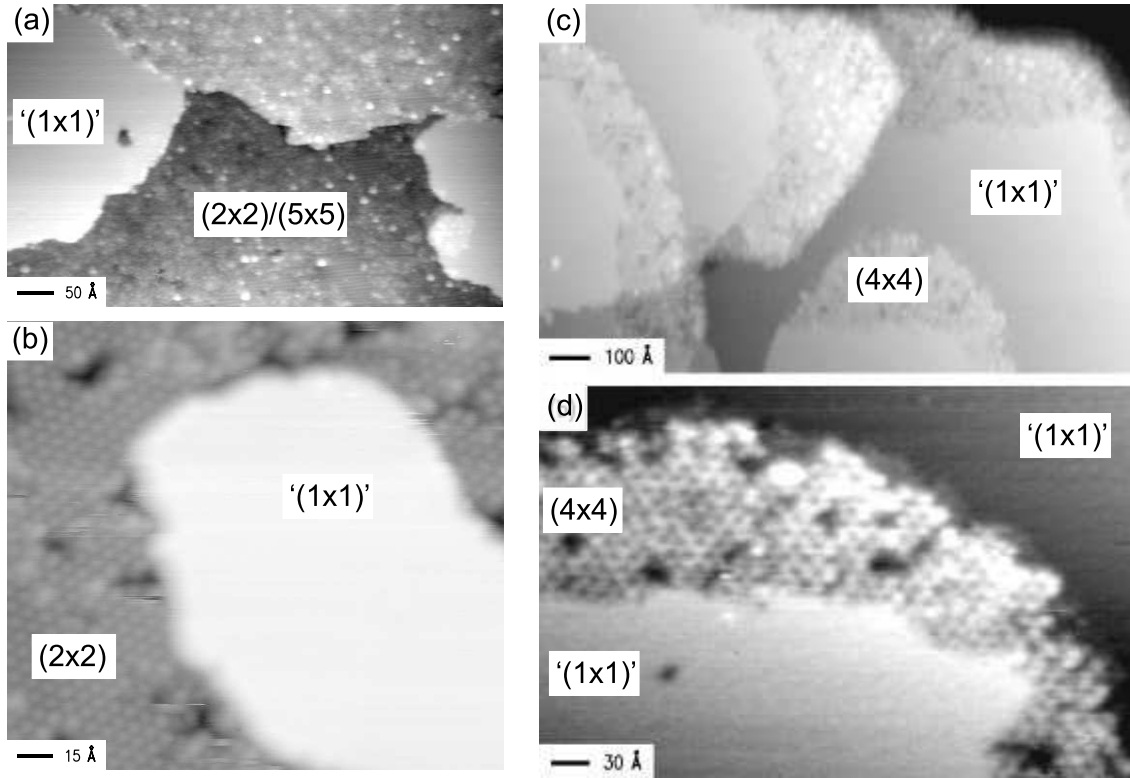


Figure 5.5: STM images of GaN (0001) surface following 0.5 ML Si exposure. (a) Surface displaying regions of mixed (2×2) and (5×5) reconstructions, together with islands of (1×1) structures. (b) High resolution view of (1×1) island surrounded by (2×2) structure. Images were acquired with sample bias voltages of -2.0 and -2.5 V, respectively, and are shown with gray-scale ranges of 4.9 and 5.3 Å, respectively. (c) Following ~ 1 ML Si exposure. Large image displaying terraces of (1×1) reconstruction with (4×4) structure seen at the terrace edges. (d) High resolution view of (4×4) structure near a terrace edge. Images were both acquired with a sample bias voltage of +2.0 V, and are shown with gray-scale ranges of 13 and 2.1 Å, respectively. After Lee *et al.* [39].

smooth surface morphology under Ga-rich conditions. Fig. 5.5 shows STM images of GaN (0001) exposed to ~ 0.5 -1 ML Si.

In this experiment Si exposure was performed on various GaN surface reconstructions at 570-620 K. At higher temperatures the Si induced reconstructions disappear. In fact, during growth and under Si exposure, the surface does not show any reconstructions except (1×1) . If Si is deposited on a Ga-rich (0001) surface displaying (1×1) reconstructions, no change in the surface structure is observed by reflection high-energy electron diffraction (RHEED). Si appears not to have modified the surface structure. If, alternatively, Si is deposited on an (5×5) reconstruction, a Si-induced (2×2) reconstructions results, as shown in Fig. 5.5(a)-(b). However, the formation of the (2×2) reconstruction is quite narrow. With increasing substrate

temperature the (2×2) disappears after it has formed, implying that the (2×2) structure is metastable.

Auger experiments estimate the Si coverage in Fig. 5.5(a) to be ~ 0.5 ML. The central part of the sample consists of $(2 \times 2)/(5 \times 5)$ reconstructions. This is probably due to incomplete Si coverage, where the (5×5) reconstruction is without Si on it, whereas the (2×2) is the changed reconstruction due to Si adsorption. The (1×1) reconstruction seems not to react with the Si atoms.

When additional ~ 0.5 ML of Si is deposited on the surface, small domains of well-ordered (2×2) reconstruction are seen on the surface, as well as numerous small islands with (1×1) reconstructions, as shown in Fig. 5.5(c). Relative to the initial Si exposure, the density of (1×1) domains increases.

Upon continuing the Si exposure up to ~ 1 ML at 570 K, the (2×2) pattern becomes weak and a (4×4) pattern appears. In addition, the featureless (1×1) region is dominant and the (4×4) region is seen only near step edges. This feature is shown in Fig. 5.5(d). With increasing annealing temperature, the (4×4) region disappears and at room temperature only the (1×1) reconstruction is seen. Auger experiments indicate that the whole surface is covered by ≈ 2 ML of Ga atoms.

In view of such an amount of experimental data, but no clear conclusion on what is the rule of Si on GaN surfaces, we performed density-functional theory calculations of Si adsorption on GaN (0001) surfaces. As we do not have information about the stoichiometry of the surface, we study various Si coverages Θ_{Si} and consider different adsorption sites (on the top layers and at subsurface sites).

5.3 Adsorption of Si at GaN (0001)

5.3.1 Si-adatom on Ga/N-terminated surfaces

For the sake of computational effort, we restricted our calculations to (1×1) , (2×2) and $(\sqrt{3} \times \sqrt{3})$ surface unit cells. To study the adsorption of Si at GaN (0001) surfaces we first deposited Si adatoms on the clean Ga- and N-terminated surfaces. The procedure is done as follows: Si atoms have been placed on the *fcc*, *hcp* and *ontop* positions on the Ga- and N-terminated surfaces, as shown in Fig. 5.6. Considering a (2×2) unit cell, the Si coverage thus can assume the values $\Theta_{\text{Si}} = 1/4, 1/2$ and $3/4$ ML. For $\Theta_{\text{Si}} = 1$ ML we use an (1×1) unit cell. For all reconstructions the atoms of the two top-layers of the slab are allowed to relax in addition to the adatoms. The relaxation of additional sublayers was found not be negligible.

In order to determine the stability of the different configurations we use the same procedure as in Chapter 4, Sec. 4.2. We calculate the relative surface energy as a function of the chemical potentials of the involved atomic species, Ga and N. However, now we have an additional variable, the Si chemical potential. Thus, the relative surface energy as a function of the N and Si chemical potentials is written as

$$\gamma_s(\mu_N, \mu_{\text{Si}}) = E_{\text{tot}}(n_{\text{Ga}}, n_{\text{N}}, n_{\text{Si}}) - \mu_{\text{N}}n_{\text{N}} - \mu_{\text{Ga}}n_{\text{Ga}} - \mu_{\text{Si}}n_{\text{Si}} - \gamma_s^{\text{clean(Ga-term)}}(n_{\text{Ga}}, n_{\text{N}}), \quad (5.10)$$

where $E_{\text{tot}}(n_{\text{Ga}}, n_{\text{N}}, n_{\text{Si}})$ is the total energy of the slab used to model the surface, μ_{α} is the atomic chemical potential of each species, n_{α} is the number of atoms of each species α , and $\gamma_s^{\text{clean(Ga-term)}}$ is the surface energy of the clean Ga-terminated surface, used as the reference energy.

Surface energy

In Fig. 5.7 we plot the relative surface energy as a function of the N chemical potential. The Si chemical potential is set to Si-rich conditions ($\mu_{\text{Si}} = \mu_{\text{Si-bulk}}$). From Fig. 5.7 we can see that all structures with Si on the clean Ga-terminated surface have a higher surface energy than the clean Ga-terminated surface, which means that they are energetically unstable against the formation of the clean Ga-terminated surface. This is found for all Si coverages, regardless of the adsorption site. The only exception is the hollow site for small coverages ($\Theta_{\text{Si}} = 1/4$), which is energetically more favorable than the clean Ga-terminated surface.

For the N-terminated surface we found that adsorption on the *ontop* site is energetically unfavorable, regardless of the Si coverage. The hollow *hcp* and *fcc* sites are degenerate in energy within the estimated accuracy and become more favorable as the Si coverage increases.

Based on the results of Fig. 5.7 we can see that Si strongly prefers the adsorption on N-terminated surfaces rather than on the Ga-terminated surface. Besides, Si prefers to maximize the formation of Si-N bonds. This can be seen noticing that the three-fold coordinated *hollow* site is preferred rather than the singly coordinated *ontop* site.

Relaxations

In order to understand why adsorbing Si on the Ga-terminated surface destabilizes the surface while adsorption of Si on the N-terminated surface stabilizes it for high Si coverages, we analyze in the following the bond lengths of Si-N and of Si-Ga as a function of the coverage for all adsorption sites *fcc*, *hcp* and *ontop*. As reference we model the strength of single Si-Ga and Si-N bonds, by calculating the Si-Ga and Si-N dimers. The calculations have been performed using a large cubic supercell of 20 Bohr to avoid the interaction of the dimer with another dimer of a neighboring cell.

The optimized bond lengths of these dimers are shown in Table 5.1. We can see that there is no significant difference between the bond length of a Si-Ga dimer (2.3 Å) and the bond lengths of a Si-Ga when Si is adsorbed on the *ontop* site of the Ga-terminated surface (maximum difference is smaller than 0.1 Å). For the *fcc* and *hcp* sites, the difference between the dimer bond length and the Si-Ga bond length on the surface lies in the range 0.1 – 0.3 Å. The general tendency is that the Si-Ga bond lengths tend to increase when the coverage increases.

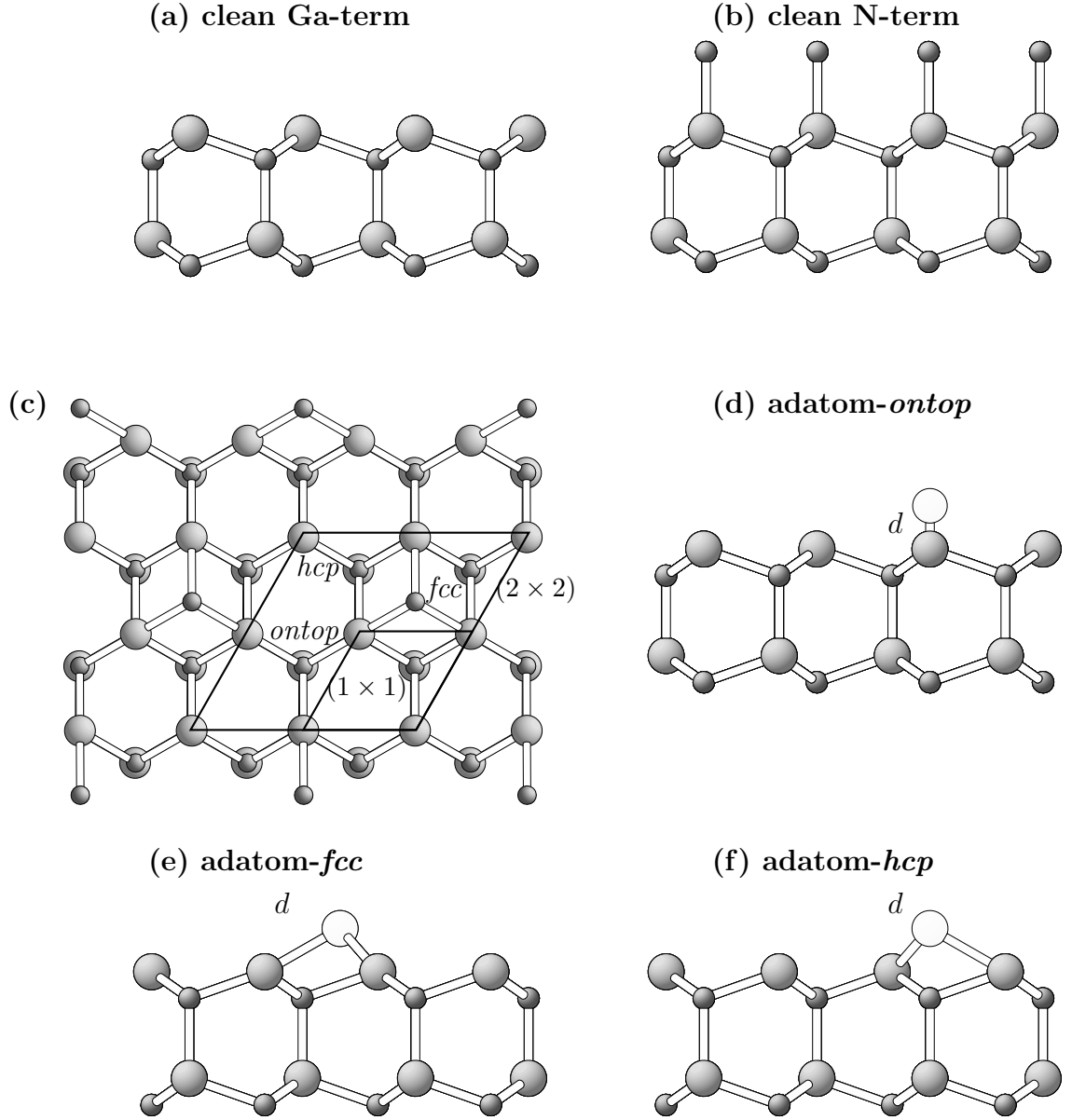


Figure 5.6: (a) Side view of the clean Ga-terminated and (b) N-terminated (0001) GaN surfaces. (c) Top view of the (1×1) and (2×2) unit cells of the clean GaN (0001) surface, indicating the highest symmetry adsorption sites *fcc*, *hcp* and *ontop*. (d) Side view of the Si-*ontop*, (e) Si-*fcc* and (f) Si-*hcp* structures. *d* is the bond length between the Si-adsorbate and the Ga atoms in the first plane. Also, the atoms in the first plane can be N.

For the adsorption on the N-terminated surface, we found that there is no significant difference between the Si-N dimer bond length and the Si-N bond lengths on the surface the *ontop* position (less than 0.1 \AA) for $1/4 \leq \Theta_{\text{Si}} = 3/4 \text{ ML}$. However, for $\Theta_{\text{Si}} = 1 \text{ ML}$ the difference is slightly larger (0.3 \AA). For the *hollow* sites, the tendency is that the bond length increases as the coverage increases. In particular, for a full adsorbed monolayer, the bond length is 0.4 \AA larger than the bond length

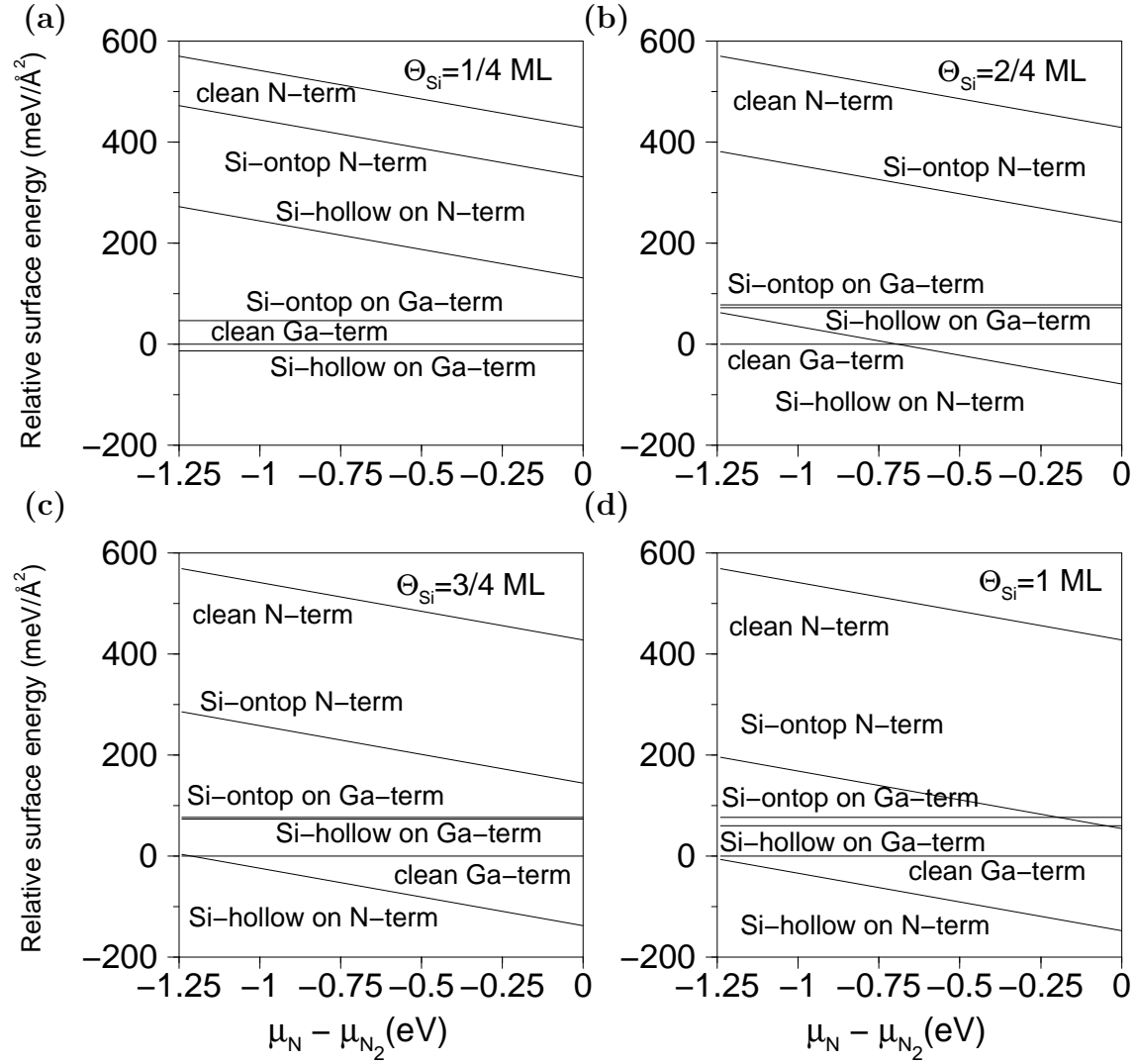


Figure 5.7: Relative surface energy per Å² for Si adsorbed on the clean Ga- and N-terminated GaN (0001) surface for a) $\Theta_{Si} = 1/4$, b) $\Theta_{Si} = 1/2$, c) $\Theta_{Si} = 3/4$ and d) $\Theta_{Si} = 1$ ML. Si-rich conditions ($\mu_{Si} = \mu_{Si-bulk}$) are assumed. Here LDA was employed.

of the dimer.

Besides, as the coordination of Si with the N atoms in the first layer increases, the bond length tends to the bond length of a Si-N bond in Si₃N₄-bulk (1.74 Å) for $\Theta_{Si} = 3/4$ ML. We can understand it by noting that in Si₃N₄-bulk Si is tetrahedrally coordinated with N (see Fig. 3.7). Thus the 3 Si atoms at the surface are arranged in a similar tetrahedral configuration as in Si₃N₄-bulk, except that on the surface the Si atoms are only three-fold coordinated. For $\Theta_{Si} = 1$ ML, however, the bond length is a bit larger (0.3 Å for the *hollow* sites and 0.5 Å for the *ontop* site).

Table 5.1: Bond lengths between the Si-adatoms and the substrate atoms (see Fig. 5.6) . $d_{\text{Si-N}}$ for adsorption on the N-terminated and $d_{\text{Si-Ga}}$ for adsorption on the Ga-terminated surface. Coverages in the range $1/4 \leq \Theta_{\text{Si}} \leq 1\text{ML}$ were considered. Additionally, the bond length for the Si-N and Si-Ga dimers are presented.

	$d_{\text{Si-N}} (\text{\AA})$				$d_{\text{Si-Ga}} (\text{\AA})$			
$\Theta_{\text{Si}} (\text{ML})$	<i>hcp</i>	<i>fcc</i>	<i>ontop</i>	dimer	<i>hcp</i>	<i>fcc</i>	<i>ontop</i>	dimer
				1.58				2.3
1/4	1.68	1.68	1.60		2.40	2.48	2.30	
2/4	1.70	1.77	1.65		2.46	2.46	2.31	
3/4	1.76	1.75	1.64		2.52	2.52	2.35	
4/4	2.01	2.02	1.86		2.54	2.58	2.37	

Adsorption energy

The adsorption energy per adatom E_{ad} of an adatom on a clean surface is written as

$$E_{\text{ad}} = E_{\text{tot}}^{\text{surf+adatom}} - E_{\text{tot}}^{\text{surf}} - E_{\text{tot}}^{\text{adatom}}. \quad (5.11)$$

where $E_{\text{tot}}^{\text{surf+adatom}}$ is the total energy of the surface with the adsorbed atoms, $E_{\text{tot}}^{\text{surf}}$ is the total energy of the surface without the adatom (clean surface) and $E_{\text{tot}}^{\text{adatom}}$ is the total energy of the adatom, i.e., of the free atom. For ease of viewing, Eq. (5.11) is referred with respect to the cohesive energy of Si-bulk. Therefore Eq. (5.11) is written as

$$E_{\text{ad}}^{\text{ref}} = E_{\text{tot}}^{\text{surf+adatom}} - E_{\text{tot}}^{\text{surf}} - E_{\text{tot}}^{\text{adatom}} - E_{\text{coh}}^{\text{Si-bulk}}, \quad (5.12)$$

where $E_{\text{ad}}^{\text{ref}}$ means that we have taken $E_{\text{coh}}^{\text{Si-bulk}}$ as the reference energy.

In Fig. 5.8 we plot the adsorption energy of Si on the Ga- and N-terminated surfaces as a function of the coverage. Here we set the zero of energy to the cohesive energy of the bulk phase of Si-bulk. If the adsorption energy lies above the zero of energy, then the adsorption on a specific site is favored, otherwise Si droplets may be formed. As we can see from Fig. 5.8(b), the adsorption energy of Si on the N-terminated surface is very large. This can be understood noticing that the experimental binding energy of a Si-N molecule is rather large (-4.8 eV) [134]. As the clean N-terminated surface has three dangling bonds, once one Si atom is adsorbed on this surface, it will form bonds with the N atoms on the surface, leading to a gain of energy of approximately 4.8 eV per bond.

However, clean N-terminated surfaces are found to be thermodynamically unstable [136] and it might be difficult to prepare a N-terminated surface in order to adsorb Si on it. Adsorption on the *hollow* (*fcc, hcp*) site is preferred instead of adsorption on the *ontop* site. We note that the adsorption energy for the *ontop*

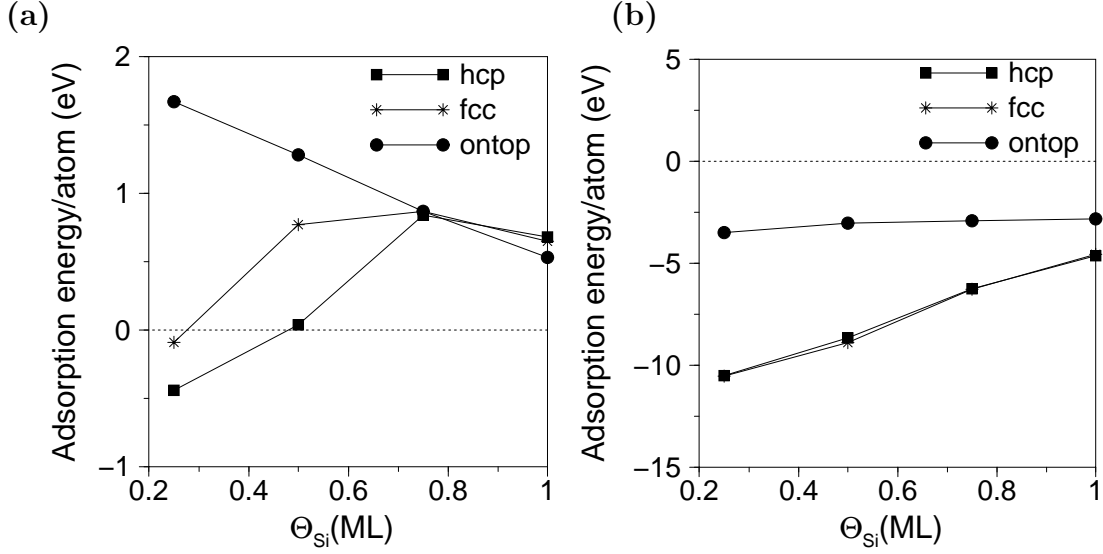


Figure 5.8: Adsorption energy calculated according to Eq. (5.12) of Si on the (a) Ga-terminated and (b) N-terminated GaN (0001) surface for $1/4 \leq \Theta_{\text{Si}} \leq 1$ ML for the *fcc*, *hcp* and *ontop* sites. The zero of energy is set to the cohesive energy of Si-bulk.

site is approximately the same for all coverages, while adsorption on the *hollow* site decreases with increasing Si-coverage.

Adsorption of Si on the Ga-terminated surfaces leads to a relatively small energy gain compared to the energy gain of Si on the N-terminated surfaces and is favored only for $\Theta_{\text{Si}} = 1/4$ ML at the *hollow* site.

Density of states

In order to have insight on the electronic structure (covalent/metallic character) we have calculated the total density of states (DOS) for the N- and Ga-terminated surfaces as a function of the Si coverage for the (2×2) and (1×1) reconstructions. The total density of states describes how many electrons are located in an infinitesimal region $d\epsilon$ around a given energy ϵ . In practice it is calculated from the Kohn-Sham eigenvalues $\epsilon_{i,\mathbf{k}}$ from the weighted sum

$$g(\epsilon) = \sum_{\mathbf{k}_i} w_{\mathbf{k}_i} \sum_i f(\epsilon - \epsilon_{i,\mathbf{k}}). \quad (5.13)$$

Here f is a broadening function necessary because the eigenvalues are discrete and $w_{\mathbf{k}_i}$ is the weight associated with each \mathbf{k} point with $\sum_{\mathbf{k}_i} w_{\mathbf{k}_i} = 1$. We have used Gaussian functions as broadening functions

$$f(\epsilon - \epsilon_{i,\mathbf{k}}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\epsilon - \epsilon_{i,\mathbf{k}})^2}{2\sigma^2}}, \quad (5.14)$$

where $\sigma = 0.1$ eV.

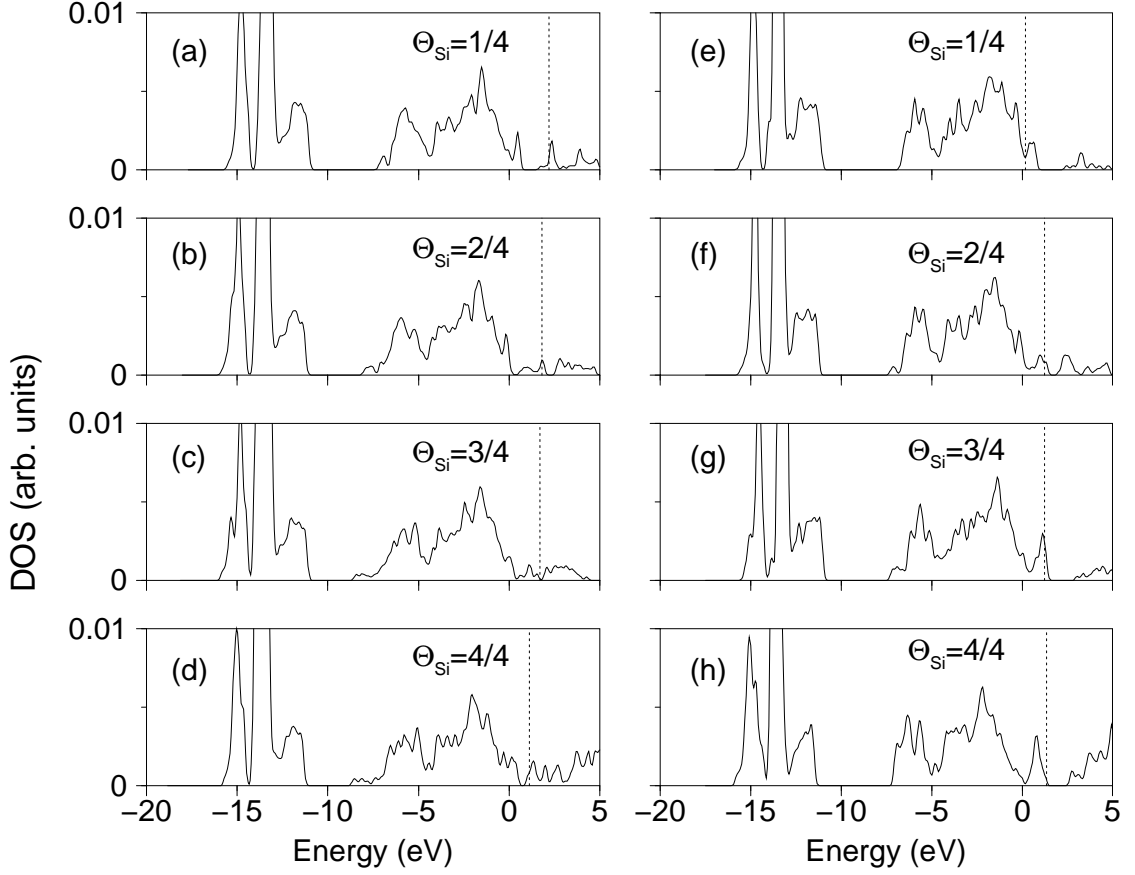


Figure 5.9: Total density of states for Si-adsorbed on the *hcp* site of a (a)-(d) Ga-terminated GaN (0001) surface and (e-h) N-terminated GaN (0001) surface. Silicon coverages in the range $1/4 \leq \Theta_{\text{Si}} \leq 1$ ML were considered. The dashed vertical line indicates the Fermi level. The zero of energy is set to the top the valence band of GaN-bulk. Here LDA was employed.

The total density of states of Si adsorbed on the *hollow* of the clean Ga- and N-terminated surfaces is shown in Figs. 5.9(a)-(d) and 5.9(e)-(h), respectively. Si coverages in the range $1/4 \leq \Theta_{\text{Si}} \leq 1$ ML were considered. In the case of Si adsorbed on Ga-terminated surfaces, we can see that the Fermi level almost does not change with the coverage, and the surfaces have metallic character.

On the other hand, for Si-adsorbed on the N-terminated surfaces the surface changes from a metallic behavior, with states inside the band gap, for $\Theta_{\text{Si}} = 1/4$ ML (Fig. 5.9(e)) to an almost semiconducting character for $\Theta_{\text{Si}} = 1$ ML (Fig. 5.9(h)). The Fermi level for $\Theta_{\text{Si}} = 1$ ML in Fig. 5.9(h) is almost pinned at the top of the valence band.

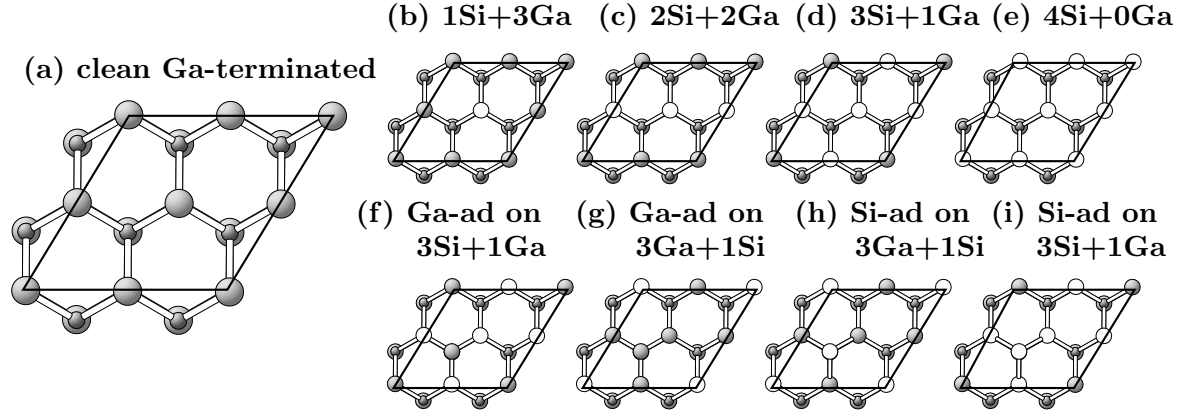


Figure 5.10: Top view of the mixed (2×2) surfaces. The starting point to obtain the mixed (2×2) structures is the clean Ga-terminated surface shown in (a). In (b)-(e) atoms on the first layer are systematically replaced by Si atoms, which leads to surfaces with coverages 1Si/3Ga (b), 2Ga/2Si (c), 3Si/1Ga (d), 4Si+0Ga (e) in the first layer. In (f)-(i) structures with adatoms on the mixed surfaces are shown. (f) Ga-adatom on 3Si/1Ga, (g) Ga-adatom on 3Ga/1Si, (h) Si-adatom on 3Ga/1Si and (i) Si-adatom on 3Si/1Ga. White balls are Si atoms, big grey balls Ga atoms and small grey balls N atoms.

5.3.2 Mixed surfaces

Si/Ga mixed surfaces

During the STM experiments described in Sec. 5.2.2, a (2×2) reconstruction is found as $\sim 0.5 - 1$ ML Si is adsorbed on the clean surface. Therefore, we systematically studied structures with (2×2) periodicity, containing Si and Ga atoms in the first layer. As the observed (2×2) structure was found to be semiconducting. We also studied structures that obey the electron counting rule (ECR) ².

It is worth noting, however, that Si and Ga atoms have very similar electronegativity. Therefore, in this case, the ECR should be taken with restriction.

At a first step, we studied structures that contain Ga and N atoms together with Si atoms in the first layer. For such a study we replace the Ga atoms in the outermost layer, as shown in Fig. 5.10(a) by Si atoms in the (2×2) unit cell, as shown in Fig. 5.10(b)-(e). Doing so, we obtain surfaces with Ga coverage in the range $0 \leq \Theta_{\text{Ga}} \leq 1$ ML and Si coverage in the range $0 \leq \Theta_{\text{Si}} \leq 1$ ML.

In addition, we study structures with Ga and Si adatoms on the mixed structures containing Si and Ga in the first layer, as shown in Fig. 5.10(f)-(i). Doing so, the Si

²It is widely accepted that the surfaces of polar III-V semiconductors should reconstruct such that all the dangling bonds on the electropositive surface atoms (III) are unoccupied and all those on the electronegative atoms (V) are doubly occupied, with the resulting surface band gap similar to that of the bulk. This guiding principle is usually referred as the electron counting rule. The most prominent application of the electron counting rule has been to narrow the possible structural models for the many reconstructions observed. It has been shown that it works well for the conventional III-V materials, but it is insufficient to explain surface stability in the case of clean GaN surfaces (see for example Ref. [147]).

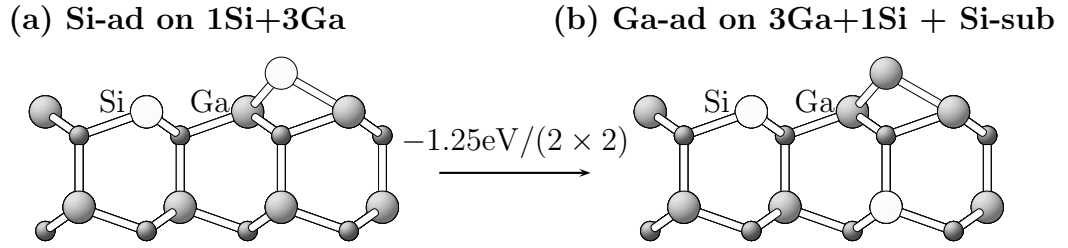


Figure 5.11: Schematic pictures showing how the energetics of the surfaces changes depending on the position of the Si atoms for (2×2) reconstructions. (a) Si-adatom on a surface with 3 Ga and 1 Si atoms in the first layer, and (b) Ga-adatom on a surface with 3 Ga and 1 Si atoms plus 1 Si in the third layer. The energy difference per unit cell between the two structures is also indicated.

coverage is in the range $1/4 \leq \Theta_{\text{Si}} \leq 1 \text{ ML}$.

Si/Ga mixed surfaces with Si-subsurface

We will now relax the condition of Si staying at the surface layers, and allow incorporation at subsurface sites. We study Ga terminated surfaces, with Si substitutionally incorporated in the first and third layers. As a first set of structures we considered a Si adatom on a Ga-terminated surface where 0 or 1 of the Ga atoms in the Ga surface layer have been replaced by Si atoms. The top view of these structures is the same as shown in Figs. 5.10(b)-(i). The only difference is that the structures contain additionally Si atoms in the third layer. One example of such structure is shown in Fig. 5.11(b).

In Fig. 5.12 we show the relative surface energy of the thermodynamically stable (2×2) structures we have studied. The (2×2) include structures where Si resides at the surface and at subsurface sites. Under N-rich conditions a full monolayer of Si on the hollow site of a N-terminated surface is found to be energetically favorable. This is expected, as we have seen from the last section, since Si prefers to form bonds with N atoms rather than with Ga atoms. Going to intermediate Ga-rich conditions, we find that the structure with the lowest energy consists of a configuration where the Ga-adatom is replaced by a Si-adatom and a Ga atom in the first layer is substituted by a Si atom ($\Theta_{\text{Si}} = 1/2 \text{ ML}$). Under this condition, the structure is $1.25 \text{ eV}/(2 \times 2)$ lower in energy than the original (2×2) adatom structure, as schematically shown in Fig. 5.11(a) and (b). Its low energy and topology makes this structure an ideal candidate to explain the (2×2) Si-induced structure observed in STM images (see Fig. 5.5)(a)-(b). Finally under extreme Ga-rich conditions a (2×2) structure with a Si-adatom on a surface with termination consisting of 3 Si atoms plus 1 Ga atom is found to be stable. As we will discuss later, none of these structures are found to be energetically stable against other structures with subsurface incorporation.

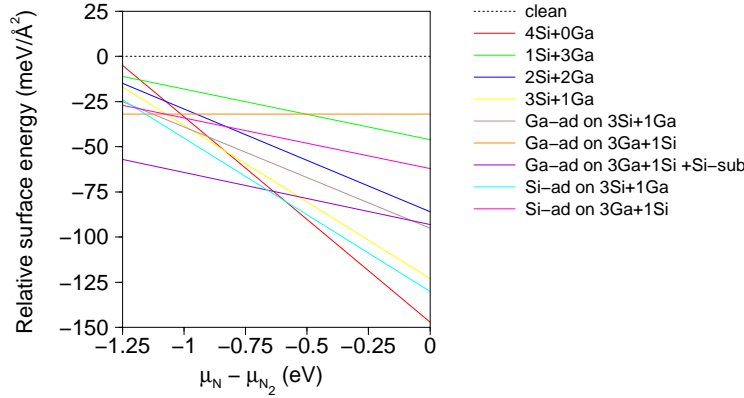


Figure 5.12: Relative surface energy of the (2×2) mixed surfaces shown in Fig. 5.10. Si-rich conditions are assumed. The energy zero was set to the clean Ga-terminated surface (dotted line).

Ga-bilayer with Si-subsurface

Concerning structures with Si residing at subsurface sites, we additionally studied Ga-bilayers on the top of Ga-terminated surfaces with Si in the third and fifth layers. This study is motivated considering that Auger experiments [39] done at the surface shown in Fig. 5.5 (c) have shown that this structure contains 2-3 ML of Ga atoms on the top layers. From a topological point of view, this structure is equivalent to the observed clean contracted Ga-bilayer structure [32,135,140]. In analogy to the clean contracted Ga-bilayer ‘ (1×1) ’ surface this structure is also denoted by ‘ (1×1) ’.

Therefore, we studied Ga-bilayer structures containing between $1/3$ and 1 ML of Si at subsurface sites, using a $(\sqrt{3} \times \sqrt{3})$ unit cell as described in Sec. 4.1. Under Si-rich/Ga-rich conditions, we find an energetically very favorable structure consisting of a double layer of Ga on a Ga-terminated surface with $\Theta_{\text{Si}} = 1/3$ ML (see Fig. 5.13(b)) to be energetically favorable. This structure is $1.15 \text{ eV}/(\sqrt{3} \times \sqrt{3})$ unit cell lower in energy than the clean contracted Ga-bilayer shown in Fig. 5.13(a).

It is worth mentioning that incorporation of Si at N sites, Si in the fifth layer or Si coverage higher than $1/3$ ML were found to be energetically unfavorable.

5.3.3 Construction of a surface phase diagram

In the previous analysis we have plotted the relative surface energy for various structures as a function of the N/Ga and Si chemical potentials. Here we want to combine all the results we have obtained in the last sections. This includes all of the (1×1) , (2×2) and $(\sqrt{3} \times \sqrt{3})$ structures described in the last sections. The results we have obtained in two dimensional plots, can be summarized by building up a three dimensional surface diagram $\gamma_s(\mu_N, \mu_{\text{Si}})$, i.e., $\gamma_s(\mu_N)$ for μ_{Si} fixed and $\gamma_s(\mu_{\text{Si}})$ for μ_N fixed. In this surface diagram, the surface energy is plotted as a function of the N and Si chemical potentials. The construction of such a phase diagram allows us to run over various possible values of the chemical potentials and compare our

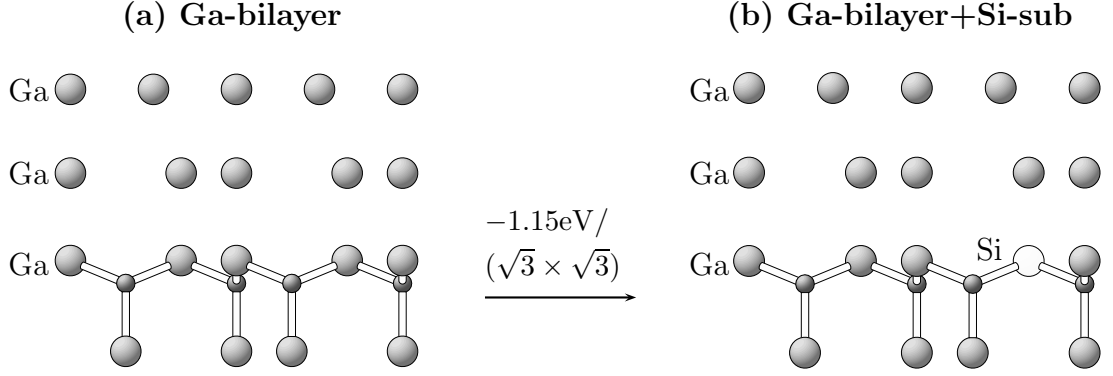


Figure 5.13: Atomic geometries of the (a) Ga-bilayer structure and (b) Ga-bilayer structure with Si in the third layer. The energy difference per unit cell between the two structures is also indicated ($-1.15\text{ eV}/(\sqrt{3} \times \sqrt{3})$).

results with experiment (see for example Ref. [146, 172]).

Before we discuss the surface diagram, we show the relative surface energy as a function of the Si chemical potential for the *most* stable structures we have found. This is shown in Figs. 5.14(a), where N-rich conditions are assumed and (b), where Ga-rich conditions are assumed. The corresponding structures are shown in Fig. 5.15.

Let us first discuss Fig. 5.14(a), where N-rich conditions are assumed. We can see that under Si-rich conditions, a (1×1) N-terminated structure with Si in the second layer is found to be thermodynamically stable. This structure is shown in Fig. 5.15(a).

Under Si-rich/N-rich conditions, a structure consisting of a full monolayer of Si on a N-terminated surface and 1 ML of Si in the third layer is found to be energetically stable. This structure is shown in Fig. 5.15(f). Going to less Si-rich conditions, structures without Si are found to be energetically favorable. We found the N-adatom at the *fcc* site on the clean Ga-terminated surface to be the most stable surface (see Fig. 5.15(e)).

Analyzing the behavior under Ga-rich/Si-rich conditions (Fig. 5.14(b)) a contracted Ga-bilayer containing $1/3$ ML of Si residing in the third layer is found to be favored. This structure is sketched in Fig. 5.15(c). Under less Si-rich conditions, the contracted Ga-bilayer structure is found to be energetically favorable, which means that under these conditions, incorporation of Si is not favored.

Now we want to summarize all these results in one phase diagram. For ease of viewing, we build up a two dimensional picture, where the relative surface energy is projected on a plane whose axes are the N(μ_N) and Si (μ_{Si}) chemical potentials. The range of the chemical potentials obeys the same boundaries as explained in Sec. 4.2.

Fig. 5.16 shows the surface diagram with the energetically most stable structures among all structures with (1×1) , (2×2) , $(\sqrt{3} \times \sqrt{3})$ periodicity discussed in the last sections. The shaded area shows the region where the structures are unstable against the formation of Si_3N_4 .

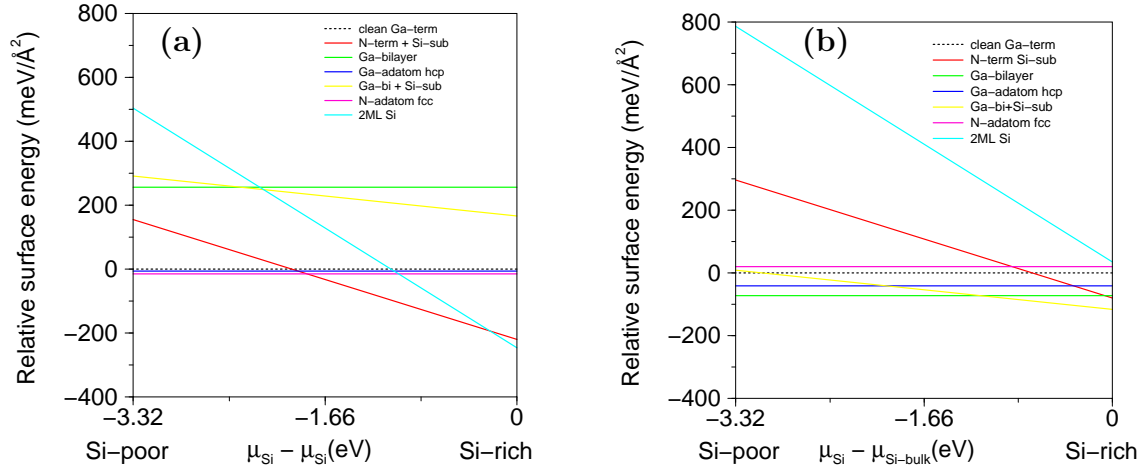


Figure 5.14: Relative surface energy of the thermodynamically stable structures. (a) N-rich conditions are assumed and (b) Ga-rich conditions are assumed. The zero of energy was set to the clean Ga-terminated surface. The atomic geometries of the corresponding structures are shown in Fig. 5.15.

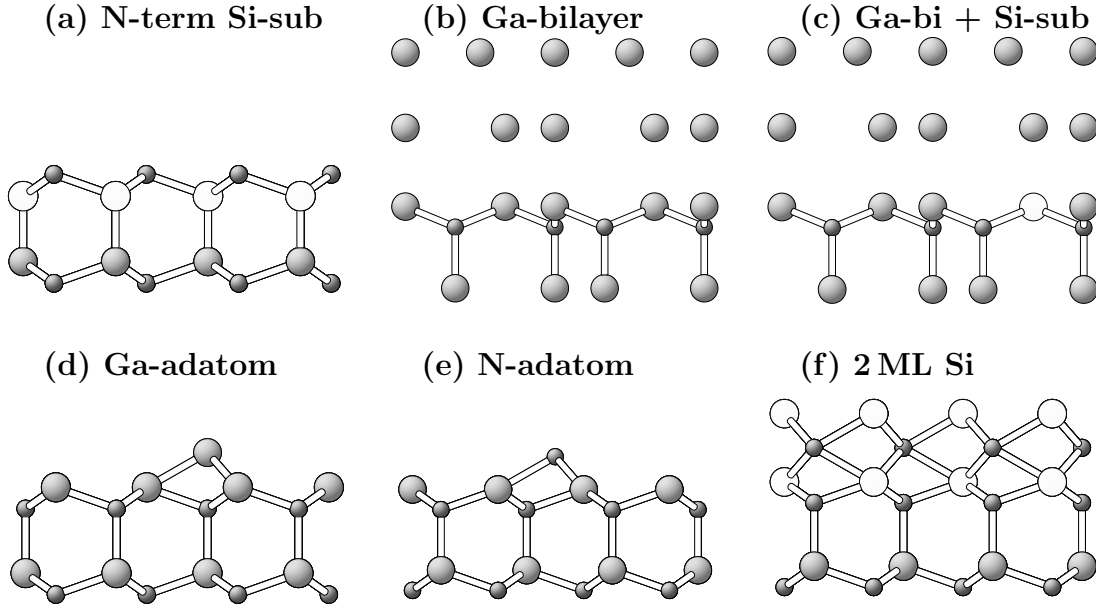


Figure 5.15: Atomic geometries of the thermodynamically stable structures. (a) N-terminated with Si in the second layer, (b) contracted Ga-bilayer, (c) contracted Ga-bilayer with Si at subsurface site in the third layer, (d) Ga-adatom at the *hcp* site on the clean Ga-terminated, (e) N-adatom at the *fcc* site on the Ga-terminated surface and (f) 1 ML of Si on the N-terminated surface with 1 ML of Si in the third layer. White balls are Si atoms, big grey balls are Ga atoms and small grey balls are N atoms

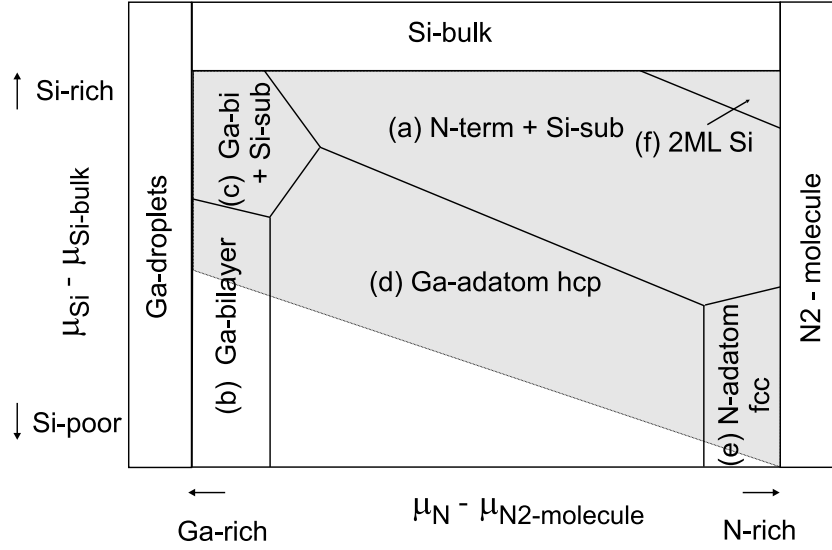


Figure 5.16: Phase diagram showing the energetically stable structures determined from first-principles calculations as function of both Si (μ_{Si}) and N (μ_{N}) chemical potentials. The shaded area shows the region where all structures are unstable against the formation of Si_3N_4 . The atomic geometry of these structures is shown in Fig. 5.15.

5.3.4 Discussions and consequences for growth

Now we want to use the surface phase diagram of Fig. 5.16 to compare our results with the experimental data in order to identify the mechanisms of Si adsorbed on the GaN (0001) surfaces.

Under more N-rich/Si-rich conditions the activation barrier to form Si_3N_4 is expected to be rather low. Thus Si_3N_4 islands/precipitates may be formed on the surface, since Si_3N_4 is well known to chemically passivate GaN surfaces and blocking growth [173]. Although all the structures we found are unstable against the formation of Si_3N_4 , the structures we have found, with a large number of Si-N bonds, might be considered as precursor states for Si_3N_4 formation. Thus, the presence of Si precipitates leads to three-dimensional growth, as observed in MOCVD growth (see Fig. 5.4).

On the other hand, under more Ga-rich conditions we observe a fundamentally different behavior, where Si prefers subsurface configurations rather than on-surface sites. Surface segregation of Si does not occur and Si can be easily incorporated in GaN-bulk. Adsorbing $1/4 - 1/2$ ML of Si on the clean (5×5) surface (we use the model structures shown in Fig. 5.17(a) changes the reconstruction to (2×2) (Fig. 5.17(e)) This would lead to the a metastable phase, which we identify as being the (2×2) surface shown Fig. 5.17(b).

With increasing Si coverage ($1/4 - 1/2$ ML of Si) more and more excess Ga atoms are created (because first a Ga vacancy should be created in order that Si

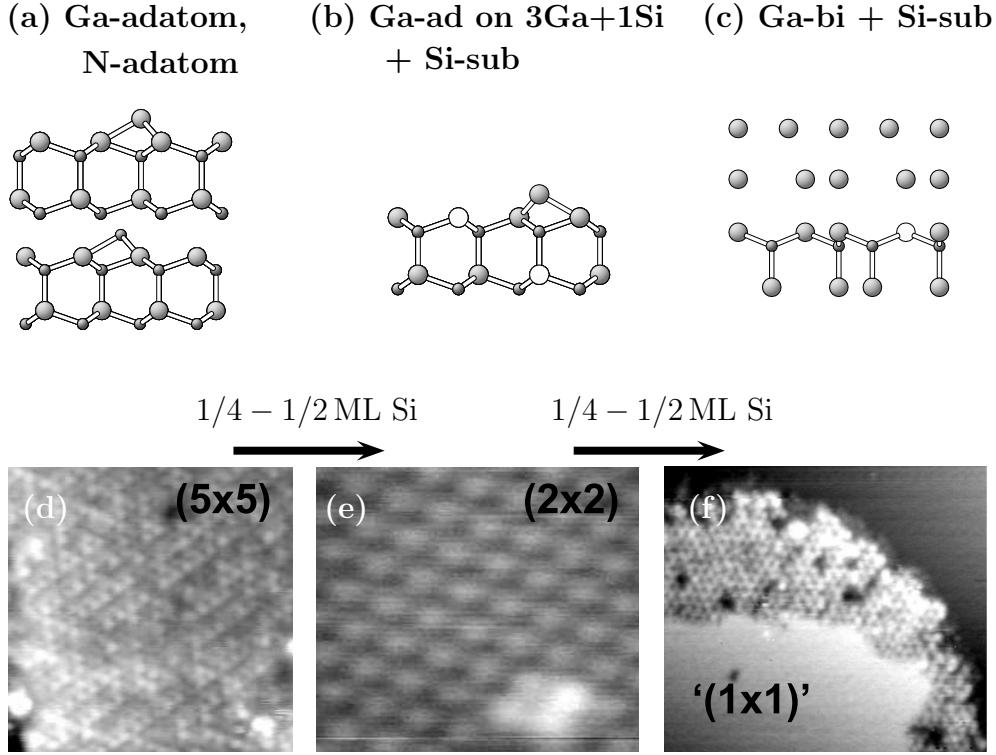


Figure 5.17: Comparison between our calculated structures using density-functional theory and the STM experiment of Ref. [39]. (a) Ga-atom and N-atom structures used to model the clean (5×5) reconstruction shown in (d). (b) Once $1/4 - 1/2$ ML of Si is adsorbed on the clean surface, the reconstruction changes to the metastable (2×2) reconstruction shown in (e), which we found to be a Ga-atom on a surface with 3 Ga and 1 Si atoms plus 1 Si in the third layer. (c) With additional $1/4 - 1/2$ ML of Si, the surface is almost completely covered by the $'(1 \times 1)'$ surface shown in (f) which we explain as being a contracted Ga-bilayer with Si at subsurface site in the third layer.

is incorporated in the Ga site) and the area covered by the Ga-bilayer will increase until eventually it covers the entire surface. Under these conditions, the Si-induced surfaces are essentially free of Si in the top surface layer and topologically very similar to the bare GaN surfaces, demonstrating efficient incorporation of Si (Fig. 5.17(e)). This is consistent with our results that predict the formation of a Ga-bilayer on the top of a Ga-terminated surface with Si being incorporated at subsurface sites (third layer), as shown in Fig. 5.17(c).

Therefore, from the topological point of view, no change in the growth mode occurs and the surface presents the same surface termination as the clean contracted bilayer surface observed under Ga-rich conditions. Thus, we can conclude that Ga-rich conditions are the optimum regime to incorporate Si in GaN, as observed in MBE growth.

5.4 Polarity inversion of GaN (0001) surfaces

A serious problem when going to high Mg doping levels is an instability of the GaN (0001) surface with respect to polarity inversion [27, 174]. Polarity inversion means that the starting sequence in the GaN double layer changes, i.e., the (0001) surface converts into a (000 $\bar{1}$) surface. From a crystallographic point of view, such a structure is regarded to be highly unstable since polarity inversion cannot be achieved by conserving the local tetrahedral environment of the atoms but only by local bond deformations. Thus, according to common wisdom a transformation of a (0001) into (000 $\bar{1}$) was considered to be impossible.

Theoretical models have found the formation of 1 ML of Mg-terminated (000 $\bar{1}$) surface atop an IDB to be more favorable than a Mg-terminated (0001) surface [27]. A consequence for growth is that the density of dislocations in the film increases dramatically (of the order of 10 times) as the polarity changes, since N-terminated surfaces are found to have poor morphology. It also results in a poor growth morphology. Since this mechanism limits the achievable doping efficiency it is crucial to understand its origin.

In this respect it is interesting to point out that one of the stable surface structures (Fig. 5.15(a)) we identified in the last chapter can be regarded as precursor for polarity inversion. In this structure the N atoms in the top surface layer have a single dangling bond pointing along (0001) and three bulk bonds pointing along (000 $\bar{1}$), i.e., equivalent to the configuration that N atoms have in GaN (000 $\bar{1}$).

The large stability of such a structure, as indicated by the large area in which this structure is stable (see Fig. 5.16) is a surprising result, since GaN structures with N-termination are highly unstable (see Sec. 4.2). Besides, Si is three-fold coordinated with N and forms only a very weak bond with the Ga-atom in the third layer, which would make such structure unstable against structures where Si is four-fold coordinated with N atoms (because Si-N bonds are much stronger than Si-Ga bonds).

In order to understand why and under what conditions polarity inversion takes place, we will analyze in the following the structural and electronic properties of relevant structures containing 1 ML of Si.

It is worth mentioning that recent experiments have found that inversion polarity of GaN surfaces can be caused by adsorption of Si [175] for concentrations above $7 \times 10^{20} \text{cm}^{-3}$. Specifically, the GaN films were grown using MBE, which commonly leads to GaN with (000 $\bar{1}$) polarity. As Si is adsorbed on the (000 $\bar{1}$) surface, the polarity is observed to change to (0001). The polarity inversion is observed in the change of RHEED patterns. However, RHEED provides only average information about the surface polarity. On the macroscopic scale there are many Ga-polar (N-polar) domains in the dominant N-polar (Ga-polar) sample. Therefore, a direct comparison with our theoretical results is at this stage not possible and further experimental studies are called for.

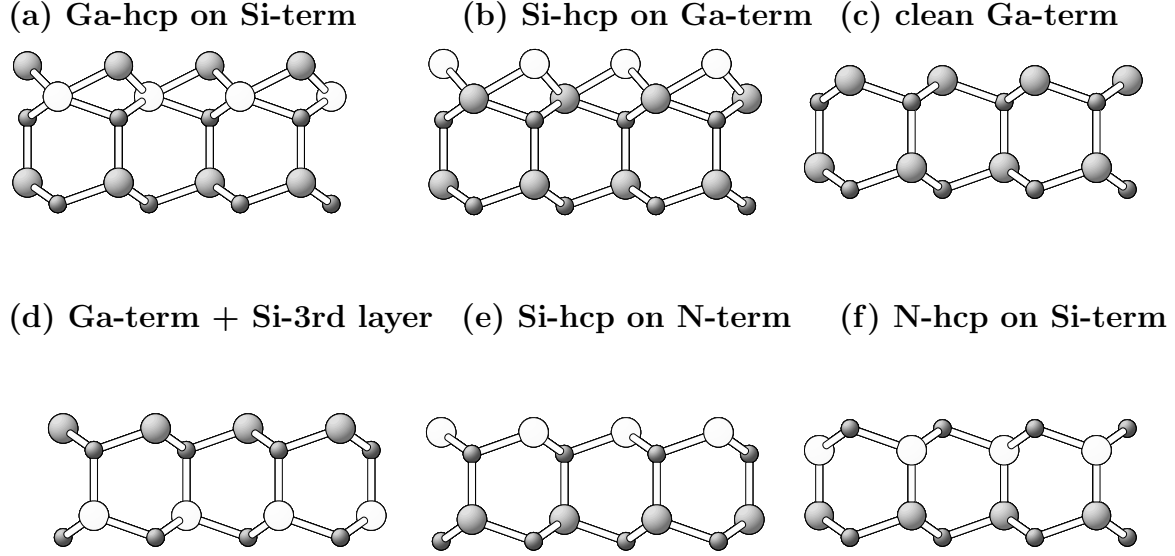


Figure 5.18: Atomic geometries of the energetically most stable structures with $\Theta_{\text{Si}} = 1\text{ML}$. Each surface has a dangling bond with $7/4 e^-$. All structures have a (1×1) surface unit cell. (a) Ga-adlayer (*hcp* site) on the Si-terminated surface, (b) Si-adlayer (*ontop* site) on the Ga-terminated surface, (c) clean Ga-terminated surface, (d) Ga-terminated surface with Si in the third layer, (e) Si-adlayer (*hcp* site) on the N-terminated surface and (f) N-adlayer (*hcp* site) on the Si-terminated surface. The big grey balls are Ga atoms, the small grey balls are N atoms and the white balls are Si atoms.

5.4.1 Surface energy

The structures we studied are schematically represented in Figs. 5.18 (a)-(f). For the following discussion we will assume that the surfaces can always reach the thermodynamic equilibrium configuration, i.e., kinetic effects (barriers for exchanging atoms) are not considered. In Fig. 5.19 we show the relative energy of such structures assuming Si-rich conditions, i.e., $\mu_{\text{Si}} = \mu_{\text{Si-bulk}}$. From Fig. 5.19 we can see that the exchange of Si atoms in the first layer with N or Ga atoms in the second layer is energetically favorable.

For example, a Ga monolayer on the *hcp* site of a Si-terminated surface, shown in Fig. 5.18(a) is energetically more favorable than a Si monolayer on the *hcp* of a Ga-terminated surface, shown in Fig. 5.18(b). Also, a Si monolayer adsorbed on the *hcp* on a N-terminated surface, shown in Fig. 5.18 (e) is energetically more favorable than and a N monolayer surface at the *hcp* site on a Si-terminated surface, shown in Fig. 5.18(f). Besides, exchange of surface Si atoms with Ga atoms in the third layer, as shown in Fig. 5.18(d) was found to lower the surface energy compared to structures with Si in the outermost layer.

To identify the mechanism that makes the inverted structure so stable, we will concentrate in the following discussion on structures with the same stoichiometry (structures shown in Figs. 5.18(d)-(f)). This has the advantage that the only differ-

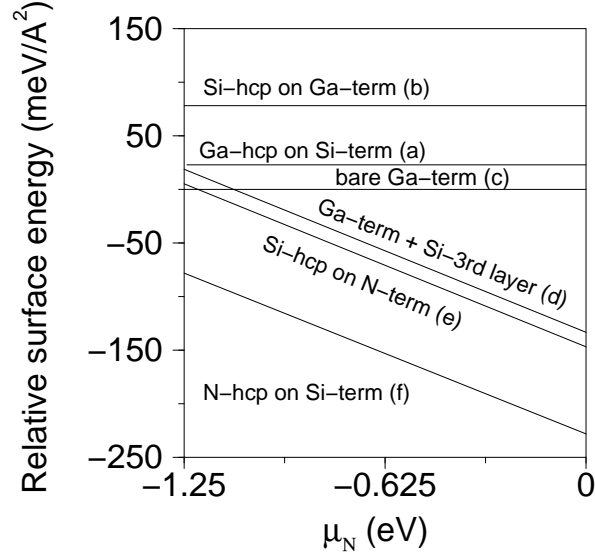


Figure 5.19: Relative surface energy for structures with $\Theta_{\text{Si}} = 1\text{ML}$. The labels (a)-(f) refer to the atomic geometries shown in Figs. 5.18 (a)-(f). Si-rich conditions are assumed ($\mu_{\text{Si}} = \mu_{\text{Si}_3\text{N}_4}$).

ence between the structures is the position of the Si, N and Ga atoms, i.e., effects due to different chemical potentials can be excluded.

5.4.2 Relaxations

Let us now analyze the atomic relaxations of the structures shown in Figs. 5.18(d)-(f). Note that the nearest distance between Ga and N in wurtzite GaN is 1.95 Å.

For the structure in Fig. 5.18 (d), the bond length between the Ga in the first layer and the N in the second layer is 2.11 Å, meaning that the surface relaxes inwards. On the other hand, we find that the Si-N bond is 1.63 Å, 0.1 Å shorter than Si-N bond in Si_3N_4 , where Si is fourfold coordinated with N.

For the Si-terminated structure with N in the second layer (Fig. 5.18(e)), we found that the surface relaxes towards, with the Si-N being 2.01 Å. The bond length between the N in the second layer and Ga in the third layer is 1.96 Å, slightly larger than in the ideal wurtzite structure.

For the N-terminated structure with Si in the second layer (Fig. 5.18(f)), we found a slightly shorter Si-N bond length (1.90 Å). The bond length between the Si-subsurface and the Ga in the third layer is found to be 2.45 Å, which is slightly larger than the bond length of a Si-Ga dimer (2.30 Å).

5.4.3 Electronic structure

All the structures shown in Figs. 5.18(d)-(f) have one dangling bond on the surface. Since the stoichiometry of all structures is identical the occupation of the dangling

Table 5.2: Calculated bond lengths for Si-N, Si-Ga and Ga-N bonds of the structures shown in Figs. 5.18 (d)-(f). All values are given in Å.

structure	Si-N	Ga-N	Si-Ga
Ga-term + Si-3rd layer	1.63	2.11	
Si- <i>hcp</i> on N-term	2.01	1.96	
N- <i>hcp</i> on Si-term	1.90		2.45

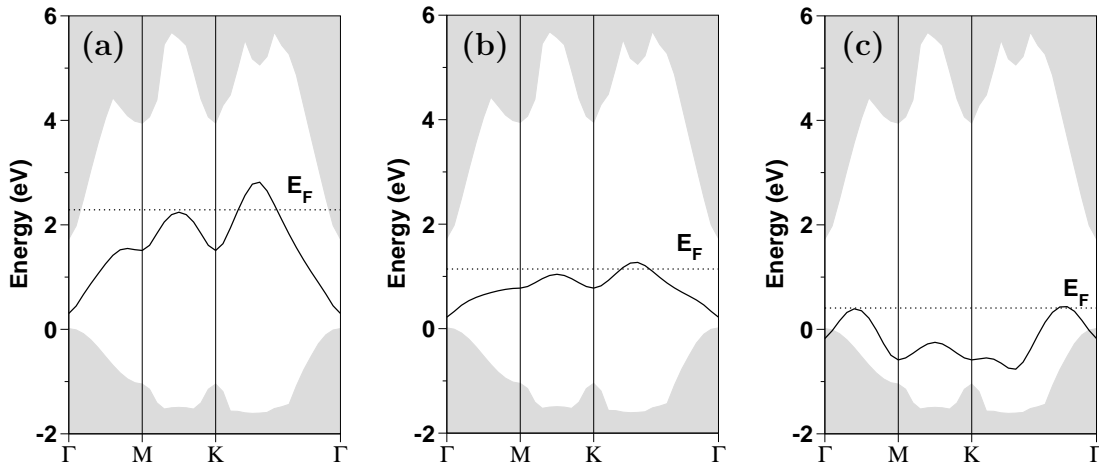


Figure 5.20: Surface states of the energetically most stable Si:GaN (1×1) surfaces (a) Ga-term with Si in the third layer (Fig. 5.18 (d)), (b) Si-*hcp* on N-term (Fig. 5.18 (e)) and (c) N-*hcp* on Si-term (Fig. 5.18 (f)) E_F is the Fermi level, the solid line represent the surface state and the shaded region is the projected band structure of GaN-bulk. The top of the valence band has been set to zero energy.

bond state must be identical. Taking into account that each Si atom substituting a Ga atom brings an additional electron, the dangling bond state is occupied with $7/4 e^-$. As the dangling bond is partially occupied, these surfaces become metallic.

In Fig 5.20 we show the corresponding band structures of the surfaces mentioned above. As expected from the simple electron counting analysis, the Fermi level crosses a surface state in the band gap. We further find a clear correlation between the position of the Fermi level and the stability of the surface: with decreasing Fermi level the structure becomes energetically more favorable. The Ga-terminated surface shown in Fig. 5.18(d) shows a larger dispersion compared to the structures of Figs. 5.18 (e) and (f).

The mechanism that stabilizes the inverted structure is understood considering that Si is a donor in GaN. Since N is much more electronegative than Si, charge is transferred from Si to the N dangling bond located at the surface. The Fermi level thus pins the surface state and so stabilizes the structure.

Table 5.3: Calculated and experimental binding energy E_b and cohesive energy per bond E_{coh} of Si-N, Si-Ga and Ga-N bonds using LDA. All values are given in eV.

	E_b			E_{coh}		
	Si-Ga	Si-N	Ga-N	Si-Ga	Si-N	Ga-N
theory	-1.64	-5.40	-4.07		-1.80	-2.60
exp. [131, 134]		-4.87	-2.18		-3.45	-2.24

5.4.4 Electron counting model

In order to discuss whether the stability of the inverted structures can be explained alone by this mechanism or whether it is important to include further processes we will derive a rather crude model, which will allow to estimate the stability of structures based on the dangling bond charge. The dangling bond energy E_{db} is estimated assuming that

$$E_{\text{db}} = E_{\text{bond}} - E_{\text{F}} \times q. \quad (5.15)$$

Here E_{bond} is the binding energy of the adsorbed layer adsorbed on the surface, E_{F} is the Fermi energy of the system and q is the occupation number of the dangling bond, i.e. for the surfaces discussed here $q = 7/4 e^-$. E_{bond} is defined as

$$E_{\text{bond}} = E_{\text{tot}} - E_{\text{tot}}^{\text{adsorbed layer}}, \quad (5.16)$$

where E_{tot} is the total energy of the system without the adsorbed layer and $E_{\text{tot}}^{\text{adsorbed layer}}$ is the total energy of the adsorbed layer. In this case, the adsorbed layer is composed by a double layer of Si-N, Si-Ga or Ga-N atoms.

To estimate the energy of the Si-Ga, Si-N and Ga-N bonds we follow two approaches. The first one is to take the binding energy of the Si-N, Si-Ga, Ga-N dimers. The other one is to extract it from the cohesive energy of a Si-N, Si-Ga and Ga-N compound. A problem with this approach is that in nature no Si-Ga compound exist. We therefore use as Si-Ga binding energy that of the Si-Ga dimer throughout.

For the bulk calculation, we use the cohesive energy per bond of GaN (Ga-N) and Si_3N_4 (Si-N). The binding energy of the dimers is calculated as $E_b^{\text{dimer}} = E_{\text{tot}}^{\text{dimer}} - \sum_i E_{\text{tot}}^i$, where $E_{\text{tot}}^{\text{dimer}}$ is the total energy of the dimer and E_{tot}^i is the total energy of the atoms that form the dimer. The corresponding energies are listed in Table 5.3.

We can see that the estimated binding energy using the cohesive energy of the bulk phases is higher than the cohesive energy of the dimer, as it was first pointed out by Pauling [176], who noted that the bonds are stronger in an environment with a smaller number of nearest neighbours (in this case the dimer). We note that our LDA values follow this trend for Si-N bonds.

With the values of Table 5.3 we calculated the dangling bond energy according to the Eq. 5.16.

Table 5.4: *Ab-initio* $E^{ab-initio}$ and model, E^{model} calculation of the dangling bond energy for the structures shown in Figs. 5.18(d)-(f). Here we shown the differences of energy, where the zero of energy was set to the most stable structure (Fig. 5.18 (f)). For $E_{\text{dimer}}^{\text{model}}$ the binding energy of the dimer was used and for $E_{\text{bulk}}^{\text{model}}$ was used. The energies are given in eV.

	Si-N	Si-Ga	Ga-N	$E_{\text{dimer}}^{\text{model}}$	$E_{\text{bulk}}^{\text{model}}$	$E^{ab-initio}$
N- <i>hcp</i> on Si-term	3	1	3	0	0	0
Si- <i>hcp</i> on N-term	4	0	3	-0.33	-0.62	-0.40
Ga-term + Si-3rd layer	3	0	4	-1.62	-1.20	-0.65

The results are shown in Table 5.4. We can see that the dangling bond counting model quantitatively predicts the stability of the surfaces. If we take the cohesive energy of the dimer, the energy differences using the model are closer to the *ab initio* calculations. It is worth noting that such simple model has its limitations and up to now has been applied only to surfaces with the same stoichiometry. To improve it, more sophisticated schemes to extract the binding energies and more refined total energy functionals (going beyond the single bond counting model) are needed.

5.5 Summary

In this chapter we studied the adsorption of Si on GaN (0001) surfaces. We found that for adsorption on the outermost layer, Si adsorption on the N-terminated surface is preferred rather than on Ga-terminated surfaces. Besides, Si prefers to form the maximum number of bonds with N atoms.

Relaxing the condition that Si stays on the surface, we found that adsorption at subsurface sites is energetically favorable compared to adsorption on the top layers. In particular, we find that under Ga-rich conditions a Ga-bilayer is formed on the top of a Ga-terminated surface with Si incorporated underneath. Under such conditions Si does not affect the surface morphology. Under N-rich/Si-rich conditions, a N-terminated structure with Si in the second layer is preferred. Under such conditions Si leads to rough surfaces

The structure which is stable under N-rich/Si-rich conditions is found to lead to polarity inversion of the GaN (0001) surface. The mechanism that stabilizes the polarity inverted structure is understood considering charge transfer from Si to the N dangling bond located at the surface.

Chapter 6

Conclusions and outlook

In the present work we have performed a systematic study on the adsorption and incorporation of Si atoms on GaN (0001) surfaces and their effect on surface morphology employing density-functional theory. Based on these results we could identify the microscopic mechanisms which affect the surface morphology, the doping efficiency of Si-donors and which determine the growth of quantum dots. Furthermore, the structural and electronic properties of the most relevant bare GaN (0001) surfaces have been calculated to provide further understanding on relevant properties of these surfaces. In the following, the key contributions of the present work will be summarized and discussed.

In the beginning of this work there was no clear picture of how different exchange-correlation functionals performed on describing GaN surfaces. Therefore, calculations for the most relevant bare GaN surfaces have been performed to understand how PBE performs compared to LDA. Our results show that atomic geometries (bond lengths and relaxations) are similarly described using both functionals. Concerning the electronic structure we have found that the surface states are affected neither qualitatively nor quantitatively by the exchange-correlation functional. The location of the states in the band gap are, for matter of comparison, practically identical. Besides, we concluded that the difference in the band gap energy is exclusively due to the difference in the atomic geometry.

Regarding the thermodynamic stability of the bare surfaces, we find that under Ga-rich conditions the description using PBE is fully compatible with the LDA calculations. However, due to the fact that the formation enthalpy is significantly underestimated under N-rich conditions we do not have the same ordering of structures as obtained for LDA results. Nevertheless, if we shift the N-boundary, such that it correctly describes the experimental GaN formation enthalpy, the correct ordering is reproduced. Based on these results, we concluded that LDA provides very similar results compared to PBE and it is suitable to describe surface properties. Since a good description of the formation enthalpy is needed in our approach to describe correctly the stability of the surfaces, we have chosen to apply LDA to investigate adsorption of Si on the GaN (0001) surfaces.

Besides, we performed a systematic study on the work function, electron affinity and ionization energy of the bare GaN (0001) surfaces. We found a qualitative trend

for the calculated properties as the Ga coverage increases. However, quantitative agreement of the ionization energy and electron affinity values with experimental results is limited by the incorrect description of the band gap using LDA/GGA.

The main difficulty we encountered when studying adsorption of Si on GaN bare surfaces concerned the lack of experimental data regarding the surface stoichiometry, i.e., regarding the precise number of each species in the top surface layers. Therefore, we had to study a large variety of structures, taking into account various possible Si coverages, adsorption sites and reconstructions. In a first step, we studied adsorption of Si on the outermost layer of the surface. For these surfaces, we found that under Si-medium conditions all Si-induced structures are thermodynamically unstable against bare GaN surfaces, meaning that the incorporation of Si at the surface is energetically unfavorable. Under Si-rich conditions, adsorption of Si on the Ga-terminated surfaces is even energetically unfavorable against the formation of the clean Ga-terminated surface.

In a second step, we investigated structures where Si is incorporated in subsurface sites. These structures were found to be energetically more stable than surfaces with Si on the outermost layer. Under Si-rich/N-rich conditions, a structure with N-termination and Si directly underneath is found to be stable in a large range of the chemical potential. This structure leads to polarity inversion of the GaN (0001) surface. Under Si-rich/Ga-rich conditions Si is buried in the third layer under a bilayer of Ga atoms.

All these results are summarized in a surface phase diagram which allows a direct comparison with experiment. Under Si-rich/Ga-rich conditions, Si prefers subsurface configurations rather than on surface sites. Under these conditions surface segregation of Si does not occur and Si can be easily incorporated in GaN bulk. The excess Ga atoms cluster in islands and form a Ga-bilayer with a pseudo (1×1) structure stabilized by Si atoms in the third layer. As a consequence for growth, we conclude that the Si-induced surfaces are essentially free of Si in the top surface layer and topologically very similar to the bare GaN surfaces. Thus we expect no change in the growth mode. These results are perfectly consistent with MBE growth, where Si is found to not affect the surface morphology.

On the other hand, under Si-rich/N-rich conditions Si is stabilized at the surface. Under such conditions the activation barrier to form Si_3N_4 is expected to be rather low. Thus Si_3N_4 islands/precipitates may be formed on the surface, since Si_3N_4 is well known to chemically passivate GaN surfaces and blocking growth. Therefore, the presence of Si precipitates leads then to three-dimensional growth and Si acts as an *anti-surfactant*.

One important conclusion that has been derived from the phase diagram is that all Si-induced structures are unstable against the formation of Si_3N_4 . Our calculations have shown that, in thermodynamic equilibrium, the formation of Si_3N_4 limits the solubility of Si in GaN to $10^{15}/10^{17} \text{ cm}^{-3}$ at characteristic MBE/MOCVD growth temperatures.

In order to identify chemical trends as described above, we assumed thermodynamic equilibrium throughout. The assumption of thermodynamic equilibrium works well for many impurities in semiconductors, but one should keep in mind that

kinetic effects might play a role, in particular at low growth temperatures. Under those conditions, the doping concentration may be determined by kinetic process rather than thermodynamic equilibrium. As we could see from our results, kinetic effects are involved, since assuming thermodynamic equilibrium, Si_3N_4 formation is expected, which is not observed experimentally. The explicit inclusion of kinetic effects would be an interesting task for the future. The main problem to include kinetic effects theoretically comes from the large number of possible paths for diffusion of the atoms between the surface and bulk, such as exchange of atoms or incorporation at interstitial sites.

Appendix A

Convergence Tests

Before starting the calculations of surface properties, it is important to determine the optimum parameters that control the accuracy of the results. The optimization of such parameters also helps to minimize the computing time and memory for the calculations. The calculations consists in determining the optimum number of \mathbf{k} -points in the irreducible part of the Brillouin zone (IBZ), cutoff energy for the bulk calculations. Below we present convergence tests for GaN-bulk in the wurtzite and zinc-blende structures, for the N_2 molecule, Ga-bulk, and Si_3N_4 -bulk. We also include optimization for the clean Ga-terminated GaN (0001) surface with respect to the thickness of the vacuum region and number of layers in the slab to model the substrate.

A.1 GaN-bulk

In Tables A.1 and A.2 we present tests for the structural properties of the β -GaN and α -GaN, respectively, for various sets of \mathbf{k} -points and cutoff energies. According to Table A.1 we can see that the lattice constant is converged for an energy cutoff equals to 70 Ry if 2 \mathbf{k} -points in the IBZ are used. However, the bulk modulus B_0 and the bulk modulus derivative B'_0 are still not converged. The cohesive energy E_{coh} changes within 100 meV as one goes from 70 to 80 Ry. The convergence is achieved when 90 Ry is used. The convergence of the lattice constant follows the same trend described above for 10 \mathbf{k} -points in the IBZ. However, the convergence of B_0 and B'_0 is better. The cohesive energy in this case changes within 30 meV as one goes from 70 to 80 Ry. The complete convergence is achieved when 28 \mathbf{k} -points in the IBZ and 70 Ry is used. All the structural properties do not change within our accuracy and the cohesive energy changes by 10 meV as one goes from 70 to 80 Ry. Therefore, in the case of β -GaN, we conclude that the converged parameters are 28 \mathbf{k} -points in the IBZ and cutoff energy equal to 70 Ry.

According to Table A.2 we can see that all parameters are quite well converged already for 10 \mathbf{k} -points in the IBZ and for a energy cutoff equals to 70 Ry. However, as the c/a ratio still oscillates a bit for this \mathbf{k} -points sampling, we decided to use a mesh consisting of 33 \mathbf{k} -points and cutoff energy equals to 70 Ry.

Table A.1: Structural and cohesive properties of β -GaN. Lattice constant a_0 , bulk modulus B_0 , bulk modulus derivative B'_0 , and cohesive energy E_{coh} for various number of \mathbf{k} -points in the IBZ and cutoff energies E_{cut} . Here LDA was employed.

$E_{\text{cut}}(\text{Ry})$	\mathbf{k} -points	$a_0(\text{\AA})$	$B_0(\text{Mbar})$	B'_0	$E_{\text{coh}}(\text{eV})$
40	2	4.33	1.56	13.53	-11.31
	10	4.40	2.08	9.11	-11.33
	28	4.49	2.11	3.91	-11.31
60	2	4.50	2.00	4.53	-10.38
	10	4.53	1.96	3.88	-10.48
	28	4.52	1.91	4.29	-10.45
70	2	4.52	1.85	4.35	-10.29
	10	4.53	1.87	4.38	-10.40
	28	4.52	1.88	4.34	-10.38
80	2	4.52	1.75	5.71	-10.22
	10	4.53	1.88	4.16	-10.40
	28	4.52	1.88	4.35	-10.38
90	2	4.52	1.89	4.19	-10.22
	10	4.53	1.88	4.35	-10.40
	28	4.52	1.88	4.35	-10.38
100	2	4.52	1.89	4.32	-10.22
	10	4.53	1.88	4.35	-10.40
	28	4.52	1.88	4.34	-10.38

Table A.2: Structural and cohesive properties of α -GaN. Lattice parameters a and c and ratio c/a for various cutoff energies E_{cut} and number of \mathbf{k} -points in the IBZ. Here LDA was employed.

$E_{\text{cut}}(\text{Ry})$	\mathbf{k} -points	$a_0(\text{\AA})$	$c_0(\text{\AA})$	c_0/a_0	$B_0(\text{MBar})$	B'_0	$E_{\text{coh}}(\text{eV})$
60	10	3.187	5.193	1.629	1.87	4.35	-10.42
	33	3.187	5.193	1.629	1.88	4.35	-10.38
	63	3.185	5.198	1.632	1.88	4.35	-10.38
70	10	3.187	5.193	1.629	1.88	4.35	-10.38
	33	3.186	5.195	1.630	1.88	4.35	-10.38
	63	3.186	5.196	1.631	1.88	4.35	-10.38
80	10	3.188	5.190	1.628	1.87	4.35	-10.38
	33	3.186	5.195	1.630	1.87	4.35	-10.38
	63	3.186	5.196	1.631	1.87	4.35	-10.38

A.2 Ga-bulk

In Table A.3 we present tests for the cohesive energy of α -Ga using LDA and PBE. In this case, the lattice parameters a , b , c , u and v were not optimized. We can

Table A.3: Cohesive energy of Ga-bulk for various cutoff energies E_{cut} using LDA and PBE. A $8 \times 8 \times 8$ \mathbf{k} -point mesh (128 \mathbf{k} -points in the IBZ) was used.

$E_{\text{cut}}(\text{Ry})$	$E_{\text{b}}(\text{eV})$	
	LDA	PBE
60	3.40	2.57
70	3.29	2.55
80	3.27	2.50

Table A.4: Equilibrium distance d and binding energy E_{b} as a function of the cutoff energy for the N_2 molecule using LDA for various supercell lengths $L(\text{\AA})$. The special \mathbf{k} -point $(1/4, 1/4, 1/4)$ was used.

$L(\text{\AA})$	12	15	17		20	
$E_{\text{cut}}(\text{Ry})$	$d(\text{\AA})$	$d(\text{\AA})$	$d(\text{\AA})$	$E_{\text{b}}(\text{eV})$	$d(\text{\AA})$	$E_{\text{b}}(\text{eV})$
40	1.074	1.080	1.080	12.14	1.080	12.14
60	1.076	1.076	1.076	12.21	1.075	12.21
70	1.076	1.076	1.076	11.71	1.076	11.71
80	1.076	1.076	1.076	11.70	1.076	11.70

see from these results that the convergence of the Ga-bulk cohesive energy changes 0.02 eV as one goes from 70 to 80 Ry for LDA and 0.05 eV for PBE. Therefore, we conclude that 70 Ry are enough to converge the cohesive energy.

A.3 N_2 -molecule

In Table A.4 we optimize the equilibrium distance d of a N_2 molecule with respect to the cutoff energy for the N_2 molecule using LDA. We can see from these results that a supercell size of 17 \AA and cutoff energy equals to 70 Ry are the parameters for convergence.

In Fig. A.1 we show the total energy difference of a N atom in a cubic supercell as a function of the unit cell size. The special \mathbf{k} -point $(1/4, 1/4, 1/4)$ was used. We can see that the total energy is not converged for a 40 Ry energy cutoff. Besides, the energy difference is rather large as one goes from 40 Ry to 60 Ry. The best convergence is achieved for a size of 9 \AA . The energy difference as one goes from 70 Ry to 80 Ry is smaller than 50 meV.

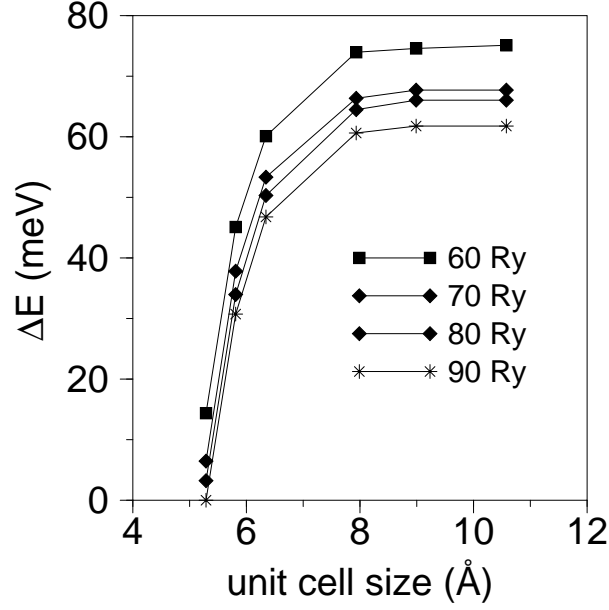


Figure A.1: Total energy difference as a function of the unit cell size for a N atom using different energy cutoffs. The zero of energy was set on the lowest energy value. Here LDA was employed.

A.4 Slab

Fig. A.2(a) shows the surface energy for the clean Ga-terminated GaN (0001) surface versus thickness of the vacuum region for three different approximations, LDA and PBE including the Ga-3d electrons as valence states, LDA within the nlcc approximation (Ga-3d electrons in the core region). We can see that the surface energy is converged for a vacuum thickness of 9 Å within an accuracy of less than 10 meV.

Fig. A.2(b) shows the surface energy for the clean Ga-terminated GaN (0001) surface versus number of GaN layers in the slab. We can see that the surface energy is converged within 5 meV already for 2 double layers. For 4 double layers the surface energy is converged.

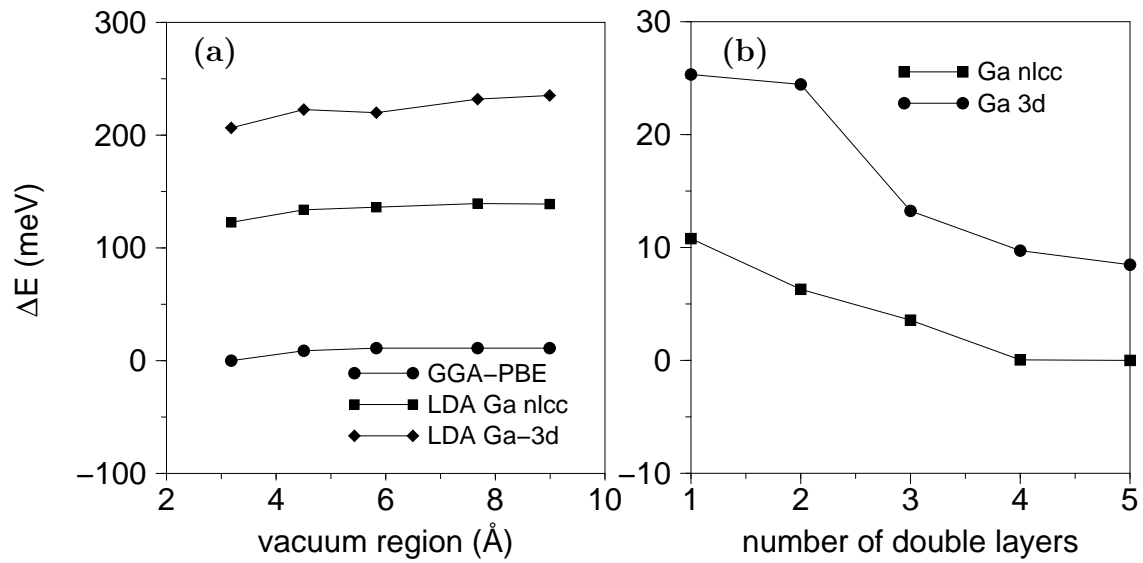


Figure A.2: (a) Surface energy for the clean Ga-terminated GaN (0001) surface versus thickness of the vacuum region using LDA, LDA within nlcc and PBE. (b) Surface energy for the clean Ga-terminated GaN (0001) surface versus number of GaN layers. The cutoff energy used for LDA and PBE was 70 Ry and for LDA within nlcc was 50 Ry. The \mathbf{k} -point mesh was $4 \times 4 \times 1$. The zero of energy is set on the lowest energy value.

Acknowledgements

First of all, I would like to thank my supervisor Dr. Jörg Neugebauer for his four years of supervision, and especially for his commitment to guiding me through my doctoral research, as well as for the time he has spent reading the various drafts of this thesis. His critical commentary on my work has played a major role in both the content and presentation of my discussion and arguments.

I also would like to express my gratitude to Prof. Dr. Matthias Scheffler for giving me the opportunity to work in the Theory group and Nachwuchsgruppe of the Fritz-Haber-Institut and to my Doktorvater at Technical University Berlin, Prof. Dr. Eckehard Schöll.

I would like to thank many helpful comments and suggestions that I have received over the last few years regarding my work. I am indebted to several people in this regard: Dr. Chris van de Walle, Dr. John Northrup, Dr. Randall Feenstra, Dr. Tosja Zywietz and Martin Fuchs.

I also extend my appreciation to the sources which provided financial support for me during my period in Germany: the Max-Planck-Gesellschaft and Deutsche Forschungsgemeinschaft.

I am very thankful to my former supervisors for my training in theoretical physics, which started many years ago, during my undergraduate course with Dr. Montserrat Freixa Pascual at Federal University of Goiás and continued during my graduation with Dr. Luisa Maria Ribeiro Scolfaro, Prof. Dr. Sylvio Roberto Accioly Canuto, Prof. Dr. José Roberto Leite and Prof. Dr. Rolf Enderlein (in memoriam) at University of São Paulo. I also would like to thank Prof. Dr. Adalberto Fazzio, for incentivate me in many aspects of my academic life.

Many thanks to people that are much more than colleagues, all my friends that helped and supported me: Dr. Juarez Lopes Ferreira da Silva, Dr. Evgeni Penev, Sixten Boeck, Dr. Cilaine Teixeira, Dr. Márcia Giacomini, Dr. Sladjana Stojkovic, Dr. Welch Cavalcanti, Mahbube Hortamani, Izabela Czekaj and Vera Stoyanova-Peneva.

My acknowledgement to the Theory department and Nachwuchsgruppe colleagues, for reading partially my thesis and making important suggestions and remarks: Liverios Lymperakis, Dr. Joel Ireta and Krzysztof Tatarczyk.

Also my thanks to Thomas Hammerschmidt for translating the abstract of my thesis to german.

And of course, I would like to thank all my colleagues of the Theory Department and Nachwuchsgruppe.

My infinite thanks to my parents, Ciro and Vanda, my sister Alessandra and

brothers Alex and Alan, so far on the other side of the Atlantic Ocean but always giving me strenght to reach my aims.

Curriculum Vitae

Name:	Andréia Luísa da Rosa
Birthdate:	11.07.1971
Nationality:	Brazilian
Address:	Fritz-Haber-Institut der Max-Planck-Gesellschaft Faradayweg 4-6 14195 Berlin, Germany
1991-1994	Bachelor degree Federal University of Goiás, Institut of Physics, Goiânia, Brazil
1995-1997	Master of Science degree University of São Paulo, Institut of Physics, Department of Materials Science, São Paulo, Brazil
1998-2002	Doctor degree Fritz-Haber-Institute, Theory Department/Technical University Berlin, Berlin, Germany

Bibliography

- [1] V. Y. Davydov, A. A. Klochikhin, R. P. Seisyan, V. V. Emtsev, S. V. Ivanov, F. Bechstedt, J. Furthmüller, H. Harima, A. V. Mudryi, J. Aderhold, O. Semchinova, and J. Graul, *Phys. Status Solidi B* **229**, R1 (2002).
- [2] T. Matsuoka, H. Okamoto, M. Nakao, H. Harima, and E. Kurimoto, *Appl. Phys. Lett.* **81**, 1246 (2002).
- [3] S. Nakamura, T. Mukai, and M. Senoh, *Appl. Phys. Lett.* **64**, 1687 (1994).
- [4] S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, *Jpn. J. Appl. Phys.* **35**, L74 (1996).
- [5] S. Nakamura and G. Fasol, *The Blue Laser Diode* (Springer-Verlag, 1997).
- [6] O. Brandt, H. Yang, B. Jenichen, Y. Suzuki, L. Däweritz, and K. H. Ploog, *Phys. Rev. B* **52**, R2253 (1995).
- [7] H. Yang, O. Brandt, M. Wassermeier, J. Behrend, H. P. Schönherr, and K. H. Ploog, *Appl. Phys. Lett.* **68**, 244 (1996).
- [8] D. Schikora, M. Hankeln, D. J. As, T. Litz, A. Waag, T. Buhrow, and F. Henneberger, *Phys. Rev. B* **54**, R8381 (1996).
- [9] Feuillet, H. Hamaguchi, K. Ohta, P. Hacke, H. Okumura, and S. Yoshida, *Appl. Phys. Lett* **70**, 1025 (1997).
- [10] D. J. As, T. Frey, D. Schikora, K. Lischka, V. Cimalla, J. Pezoldt, R. Goldhahn, S. Kaiser, and W. Gebhardt, *Appl. Phys. Lett* **76**, 1686 (2000).
- [11] W. C. Johnson, J. B. Parsons, and M. C. Crew, *J. Phys. Chem.* **36**, 2561 (1932).
- [12] J. H. Edgar, ed., *Properties of Group-III Nitrides and EMIS Datareviews* (IEE, London, 1994).
- [13] O. Ambacher, *J. Phys. D* **31**, 2653 (1998).
- [14] G. Popovici, W. Kim, A. Botchkarev, H. Tang, H. Morkoç, and J. Solomon, *Appl. Phys. Lett.* **71**, 3385 (1997).

- [15] W. Seifert, R. Franzheld, E. Butter, H. Sobotta, and V. Riede, *Crystal Res. and Technol.* **18**, 383 (1983).
- [16] S. Strite and H. Morkoç, *J. Vac. Sci. Technol* **10**, 1237 (1992).
- [17] H. P. Maruska and J. J. Tietjen, *Appl. Phys. Lett.* **15**, 327 (1969).
- [18] M. Ilegems and H. C. Montgomery, *J. Phys. Chem. Solids* **34**, 885 (1973).
- [19] J. Neugebauer and C. G. V. de Walle, in *Proceedings of the 22nd International Conference on the Physics of Semiconductors* (World Scientific, Singapore, 1994), vol. 3, p. 2327.
- [20] J. Neugebauer and C. G. V. de Walle (1996), vol. 35 of *Advances in Solid State Physics*, p. 25.
- [21] J. Myoung, K. Shim, C. Kim, O. Gluschenkov, K. Kim, S. Kim, D. Turnbull, and S. Bishop, *Appl. Phys. Lett.* **69**, 2722 (1996).
- [22] D. Bour, H. Chung, W. Götz, L. Romano, B. Krusor, D. Hofstetter, S. Rudaz, C. Kuo, F. Ponce, N. Johnson, M. Craford, and R. Bringans, *Mater. Res. Soc. Symp.* p. 449 (1997).
- [23] J. Neugebauer, *Ab initio analysis of group-III nitrides: Defects and Interfaces and Surfaces, Habilitation* (TU, Berlin, 2000).
- [24] C. G. V. de Walle, S. Limpijumnong, and J. Neugebauer, *Phys. Rev. B* **63**, 245205 (2001).
- [25] V. M. Bermudez, *Phys. Rev. B* **80**, 1190 (1996).
- [26] G. Mula, B. Daudin, C. Adelmann, and P. Peyla, *MRS Internet J. Nitride Semicond. Res.* **5S1**, W3.35 (2000).
- [27] L. T. Romano, J. E. Northrup, A. J. Ptak, and T. H. Myers, *Appl. Phys. Lett.* **77**, 2479 (2000).
- [28] B. Daudin, F. Widmann, G. Feuillet, Y. Samson, M. Arlery, and J. L. Rouvière, *Phys. Rev. B* **56**, R7069 (1997).
- [29] L. Zhang, H. F. Tang, J. Schieke, M. Mavrikakis, and T. F. Kuech, *Journal of Crystal Growth* **242**, 302 (2002).
- [30] T. Zywietz, J. Neugebauer, and M. Scheffler, *Appl. Phys. Lett.* **74**, 1695 (1999).
- [31] C. Bungaro, K. Rapcewicz, and J. Bernholc, *Phys. Rev. B* **59**, 9771 (1999).
- [32] H. Chen, R. M. Feenstra, J. E. Northrup, T. Zywietz, J. Neugebauer, and D. M. Greve, *J. Vac. Sci. Technol. B* **18**, 2284 (2000).

- [33] B. Heying, R. Averbeck, L. F. Chen, E. Haus, H. Riechert, and J. S. Speck, J. Appl. Phys. **88**, 1855 (2000).
- [34] G. Mula, C. Adelmann, S. Moehl, J. Oullier, and B. Daudin, Phys. Rev. B **64**, 195406 (2001).
- [35] L. X. Zheng, M. H. Xie, S. M. Seutter, S. H. Cheung, and S. Y. Tong, Phys. Rev. Lett. **85**, 2352 (2000).
- [36] J. E. Northrup, J. Neugebauer, R. M. Feenstra, and A. R. Smith, Phys. Rev. B **61**, 9932 (2000).
- [37] T. S. Cheng, C. T. Foxon, L. C. Jenkins, S. E. Hooper, D. E. Lacklison, J. W. Orton, B. Y. Ber, A. V. Merkulov, and S. V. Novikov, Semicond. Sci. Technol. **11**, 538 (1996).
- [38] W. Götz, N. M. Johnson, C. Chen, H. Liu, C. Kuo, and W. Imler, Appl. Phys. Lett. **68**, 3144 (1996).
- [39] C. D. Lee, R. M. Feenstra, A. L. Rosa, J. Neugebauer, and J. E. Northrup, J. Vac. Sci. Technol. B **19**, 1619 (2001).
- [40] A. L. Rosa, J. Neugebauer, J. E. Northrup, C. D. Lee, and R. M. Feenstra, Appl. Phys. Lett. **80**, 2008 (2002).
- [41] S. Tanaka, M. Takeuchi, and Y. Aoyagi, Jpn. J. Appl. Phys. **39**, L831 (2000).
- [42] P. Ramvall, P. Riblet, S. Nomura, Y. Aoyagi, and S. Tanaka, J. Appl. Phys. **87**, 3883 (2000).
- [43] A. Cremades, L. Görgens, O. Ambacher, M. Stutzmann, and F. Scholz, Phys. Rev. B **61**, 2812 (2000).
- [44] L. T. Romano, C. G. V. de Walle, J. W. A. III, W. Götz, and R. S. Kern, J. Appl. Phys. **87**, 7745 (2000).
- [45] A. Munkholm, C. Thompson, M. V. R. Murty, J. A. Eastman, O. Auciello, G. B. Stephenson, P. Fini, S. P. DenBaars, and J. S. Speck, Appl. Phys. Lett. **77**, 1626 (2000).
- [46] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin Heidelberg, 1996).
- [47] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (McGraw-Hill, New York, 1989).
- [48] P. Fulde, *Electron Correlations in Molecules and Solids* (Springer-Verlag, Berlin, 1991).
- [49] O. Gunnarson, Rev. Mod. Phys. **61**, 689 (1989).

- [50] C. Umrigar, Phys. Rev. Lett. **71**, 408 (1993a).
- [51] C. Umrigar, M. P. Nightingale, and K. J. Runge, Journal of Chemical Physics **99**, 2865 (1993).
- [52] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).
- [53] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).
- [54] S. Lundqvist and N. H. March, eds., *Theory of the Inhomogeneous Electron Gas* (Plenum Press, 1983).
- [55] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1989).
- [56] E. Fermi, Rend. Accad. **6**, 602 (1927).
- [57] L. H. Thomas, Proc. Cam. Phil. Soc. **23**, 542.
- [58] E. Fermi, Z. Phys. **48**, 73 (1928).
- [59] E. Fermi, Rend. Accad. **7**, 348 (1928).
- [60] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
- [61] J. P. Perdew and A. Zunger, Phys. Rev. Lett. **23**, 5048 (1981).
- [62] A. Becke, Phys. Rev. A **38**, 3098 (1988).
- [63] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
- [64] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
- [65] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
- [66] J. P. Perdew, J. A. Shevay, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
- [67] Y. Zhang and W. Yang, Phys. Rev. Lett. **80**, 890 (1998).
- [68] J. P. Perdew, K. Burke, and M. Ernzenhof, Phys. Rev. Lett. **77**, 3865 (1996).
- [69] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).
- [70] X. Gonze, D. C. Allan, and M. P. Teter, Phys. Rev. Lett. **68**, 3603 (1992).
- [71] K. Laasonen, F. Csajka, and M. Parrinello, Chemical Physics Letters **194**, 172 (1992).
- [72] C. S. Wang, B. M. Klein, and H. Krakauer, Phys. Rev. Lett. **54**, 1852 (1985).

- [73] E. K. U. Gross and K. Dreizler, *Density Functional Methods in Physics* (Plenum, New York, 1985), p. 81.
- [74] L. J. Bartolotti, J. Chem Phys. **90**, 5518 (1986).
- [75] A. D. Becke, J. Chem. Phys. **96**, 2155 (1992).
- [76] M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. (1996).
- [77] B. Hammer, M. Scheffler, K. W. Jacobsen, and J. K. Norskov, Phys. Rev. Lett. **73**, 1400 (1994).
- [78] B. G. Johnson, C. A. Gonzalez, P. M. W. Gill, and J. A. Pople, Chemical Physics Letters **221**, 100 (1994).
- [79] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
- [80] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 8861 (1991).
- [81] M. S. D. R. Hamann and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
- [82] L. Keinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
- [83] X. Gonze, R. Stumpf, and M. Scheffler, Phys. Rev. B **44**, 8503 (1991).
- [84] E. L. Shirley, Phys. Rev. B **40**, 3652 (1989).
- [85] M. Fuchs, M. Bockstedte, E. Pehlke, and M. Scheffler, Phys. Rev. B **57**, 2134 (1998).
- [86] S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).
- [87] M. Fuchs, J. L. F. D. Silva, C. Stampfl, J. Neugebauer, and M. Scheffler, Phys. Rev. B **65**, 245212 (2002).
- [88] E. Penev, P. Kratzer, and M. Scheffler, J. Chem. Phys. **110**, 3986 (1999).
- [89] A. L. Rosa and J. Neugebauer, unpublished.
- [90] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College Publishing, Philadelphia, 1976).
- [91] M. Bockstedte, A. Kley, J. Neugebauer, and M. Scheffler, Comput. Phys. Commun. **107**, 187 (1997).
- [92] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
- [93] R. P. Feynman, Phys. Rev. **56**, 340 (1939).
- [94] P. Pulay, Mol. Phys. **17**, 197 (1969).
- [95] M. Fuchs and M. Scheffler, Comp. Phys. Commun. **119**, 67 (1997).

- [96] see the URLs <http://www.fhi-berlin.mpg.de/th/fhimd> and <http://www.fhi-berlin.mpg.de/th/fhi98md/fhi98PP>.
- [97] C. Stampfl and C. G. V. de Walle, Phys. Rev. B **59**, 5521 (1999).
- [98] M. van Schilfagaarde, A. Sher, and A. B. Shen, J. Cryst. Growth **178**, 8 (1997).
- [99] R. Miotto, G. P. Srivastava, and A. C. Ferraz, Phys. Rev. B **59**, 3008 (1999).
- [100] F. D. Murnaghan, Proc. Nat. Acad. Sci. USA **30**, 244 (1944).
- [101] V. Fiorentini, M. Methfessel, and M. Scheffler, Phys. Rev. B **47**, 13353 (1993).
- [102] A. F. Wright and J. S. Nelson, Phys. Rev. B **51**, 7866 (1995).
- [103] K. Karch, F. Bechstedt, and T. Pletl, Phys. Rev. B **56**, 3560 (1997).
- [104] K. Karch, J. M. Wagner, and F. Bechstedt, Phys. Rev. B (1998).
- [105] C. Y. Yeh, Z. W. Lu, S. Froyen, and A. Zunger, Phys. Rev. B **46**, 10086 (1992).
- [106] K. Miwa and A. Fukumoto, Phys. Rev. B **46**, 10086 (1992).
- [107] V. Fiorentini, A. Satta, D. Vanderbilt, S. Massida, and F. Meloni, in *Proceedings of the 22nd ICPS*, edited by D. J. Lockwood (World Scientific and Singapore, 1995).
- [108] A. Satta, V. Fiorentini, A. Bosin, and F. Meloni, in *Gallium Nitride and Related Materials and MRS Symposia Proceedings*, edited by D. Dupuis, J. A. Edmond, F. A. Ponce, and S. Nakamura (MRS, Pittsburgh, 1996), vol. 395, p. 515.
- [109] S. J. Jenkins, G. P. Srivastava, and J. C. Inckson, J. Phys. C **6**, 8781 (1994).
- [110] A. Zoroddu, F. Bernardini, P. Ruggerone, and V. Fiorentini, Phys. Rev. B p. 045208 (2001).
- [111] R. C. Powell, N. E. Lee, Y. W. Kim, and J. E. Greene, J. Appl. Phys. **73**, 189 (1993).
- [112] A. Y. Liu and M. L. Cohen, Phys. Rev. B **41**, 10727 (1990).
- [113] K. Kim, W. R. Lambrecht, and B. Segall, Phys. Rev. B **53**, 16 (1996).
- [114] X. J. Kong, C. T. Chan, K. M. Ho, and Y. Y. Ye, Phys. Rev. B **42**, 5612 (1990).
- [115] M. Körling and J. Häglund, Phys. Rev. B **45**, 13293 (1992).
- [116] P. H. T. Philipsen and E. J. Baerends, Phys. Rev. B **54**, 5326 (1996).
- [117] I. H. Lee and R. M. Martin, Phys. Rev. B **56**, 7197 (1997).

- [118] L. Bosio, J. Chem. Phys. **68**, 1221 (1978).
- [119] R. W. G. Wyckoff, *Crystal Structures* (Wiley and New York, 1962), vol. 1, p. 22.
- [120] L. Bosio, A. Defrain, H. Curien, and A. Rimsky, Acta Crystalogr. B **25**, 995 (1969).
- [121] L. Bosio, H. Curien, M. Dupont, and A. Rimsky, Acta Crystalogr. B **28**, 1974 (1972).
- [122] L. Bosio, H. Curien, M. Dupont, and A. Rimsky, Acta Crystalogr. B **29**, 367 (1973).
- [123] M. Bernasconi, G. L. Chiarotti, and E. Tosatti, Phys. Rev. B **52**, 9988 (1995).
- [124] A. Kley, *Theoretische Untersuchungen zur Adatomdiffusion auf niederindizierten Oberflächen von GaAs*, PhD thesis (TU, 1997).
- [125] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1996).
- [126] N. A. W. Holzwarth, G. E. Matthews, R. B. Dunning, A. R. Tacket, and Y. Zeng, Phys. Rev. B **55**, 2005 (1997).
- [127] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
- [128] C. Filippi, D. J. Singh, and C. J. Umrigar, Phys. Rev. B **50**, 14947 (1994).
- [129] W. Y. Ching, L. Ouyang, and J. D. Gale, Phys. Rev. B **61**, 8696 (2000).
- [130] A. Reyes-Serrato, D. H. Galván, and I. Z. Garzón, Phys. Rev. B **52**, 6293 (1995).
- [131] O. Borgen and H. M. Seip, *Acta Chem. Scand.* (1961), vol. 15, p. 1789.
- [132] J. Liang, L. Topor, A. Navrotsky, and M. Mitomo, J. Mater. Res. **14**, 1959 (1999).
- [133] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
- [134] D. R. Lide, ed., *Handbook of Physics and Chemistry* (New York, New York, 1995), 76th ed.
- [135] A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, J. Vac. Sci. Technol. B **16**, 2242 (1998).
- [136] A. R. Smith, R. M. Feenstra, D. W. Greve, M. S. M.-S. Shin, J. Neugebauer, and J. E. Northrup, Surf. Sci. **423**, 70 (1999).
- [137] F. Widmann, B. Daudin, G. Feuillet, Y. Samson, J. L. Rouvière, and N. Pelekanos, Appl. Phys. Lett. **83**, 7618 (1998).

- [138] F. Widmann, B. Daudin, G. Feuillet, N. Pelekanos, and J. L. Rouvière, Appl.Phys. Lett. **73**, 2642 (1998).
- [139] Y. Chen, T. Takeuchi, H. Amano, I. Akasaki, N. Yamano, Y. Kaneko, and S. Wang, Appl. Phys. Lett. **72**, 710 (1998).
- [140] M. M. Sung, J. Ahn, V. Bykov, J. W. Rabalais, D. D. Koleske, and A. E. Wickenden, Phys. Rev. B **54**, 14652 (1996).
- [141] S. S. Dhesi, C. B. Stagarescu, K. E. Smith, D. Doppalapudi, R. Singh, and T. D. Moustakas, Phys. Rev. B **56**, 10271 (1997).
- [142] Y.-C. Chao, C. B. Stagarescu, J. E. Downes, P. Ryan, K. E. Smith, D. Hanser, M. D. Bremser, and R. F. Davis, Phys. Rev. B **59**, 15586 (1999).
- [143] Fu-He-Wang, P. Krüger, and J. Pollmann, Phys. Rev. B **64**, 035305 (2001).
- [144] K. Rapcewicz, M. B. Nardelli, and J. Bernholc, Phys. Rev. B **56**, 12725 (1997).
- [145] A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. E. Northrup, Phys. Rev. Lett. **79**, 3934 (1997).
- [146] T. K. Zywietz, *Thermodynamische und kinetische Eigenschaften von Galliumnitrid-Oberflächen, PhD Thesis* (Weissensee Verlag, Berlin, 2000).
- [147] T. Zywietz, J. Neugebauer, and M. Scheffler, Appl. Phys. Lett. **73**, 487 (1998).
- [148] J. W. Gibbs, *The collected works of J. W. Gibbs* (Longman, New York, 1928).
- [149] F. Reif, *Statistische Physik und Theorie der Wärme* (Walter de Gruyter and Co., Berlin, 1984).
- [150] U. Scherz and M. Scheffler, Semiconductors and Semimetals **38**, 1 (1993).
- [151] D. Laks, C. V. de Walle, G. Neumark, and P. Pantelides, Phys. Rev. B **45**, 10965 (1992).
- [152] G.-X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1988).
- [153] A. Zangwill, *Physics at Surfaces* (Cambridge University Press, Cambridge, 1988).
- [154] S. P. Grabowski, M. Schneider, H. Nienhaus, W. Mönch, R. Dimitrov, O. Ambacher, and M. Stutzmann, Appl. Phys. Lett. **78**, 2503 (2001).
- [155] C. I. Wu, A. Kahn, N. Taskar, D. Dorman, and D. Gallagher, J. Appl. Phys. **83**, 4249 (1998).
- [156] U. Grossner, J. Furthmüller, and F. Bechstedt, phys. stat. sol. (b) **216**, 675 (1999).

- [157] M. H. Tsai, O. F. Sankey, K. E. Schmidt, and I. S. Tsong, *Materials Science and Engineering B* **88**, 40 (2002).
- [158] J. I. Pankove and H. Schade, *Appl. Phys. Lett.* **25**, 53 (1974).
- [159] M. Eyckeler, W. Mönch, T. U. Kampen, R. Dimitrov, O. Ambacher, and M. Stutzman, *J. Vac. Sci. Technol. B* **16**, 2224 (1998).
- [160] G. Koley and M. G. Spencer, *J. Appl. Phys.* **90**, 337 (2001).
- [161] T. Valla, P. D. Johnson, S. S. Dhesi, K. E. Smith, D. Doppalapudi, T. D. Moustakas, and E. L. Shirley, *Phys. Rev. B* **59**, 5003 (1999).
- [162] V. S. Fomenko, *Emissionnye svoistva materialov* (Izd. Naukova Dumka, Kiev, 1981).
- [163] E. Bauer, *Z. Kristallogr.* **110**, 372 (1958).
- [164] I. N. Stranski and L. Krastanov, *Sitzungsber. Akad. Wiss. Wien and Math.-naturwiss. Kl. IIb* (1938), vol. 146, p. 797.
- [165] F. C. Frank and J. H. van der Merwe, *Proc. Roy. Soc. A* **198**, 205 (1950).
- [166] M. Volmer and A. Weber, *Z. Phys. Chem.* **119**, 277 (1926).
- [167] A. Pimpinelli and J. Villain, *Physics of Crystal Growth* (Cambridge University Press, Cambridge, 1998).
- [168] I. V. Markov, *Crystal Growth for Beginners, Fundamentals of Nucleation, Crystal Growth and Epitaxy* (World Scientific, Singapore, 1995).
- [169] W. K. Burton, N. Cabrera, and F. Frank, *Phil. Trans. Roy. Soc.* **243**, 299 (1951).
- [170] S. Tanaka, S. Iwai, and Y. Aoyagi, *Appl. Phys. Lett.* **69** (1996).
- [171] S. Keller, S. F. Chichibu, M. S. Minsky, E. Hu, U. K. Mishra, and S. P. DenBaars, *J. Cryst. Growth* **195**, 258 (1998).
- [172] C. G. V. de Walle and J. Neugebauer, *Phys. Rev. Lett.* **88**, 066103 (2002).
- [173] H. Lahrech, P. Vennèguès, B. Beaumont, and P. Gibart, *J. Crystal Growth* **205**, 245 (1999).
- [174] V. Ramachandran, R. M. Feenstra, W. L. Sarney, L. Salamanca-Riba, J. E. Northrup, L. T. Romano, and D. W. Greve, *Appl. Phys. Lett.* **75**, 808 (1999).
- [175] D. Wang, S. Yoshida, and M. Ichikawa, *Journal of Crystal Growth* **242**, 20 (2002).
- [176] L. Pauling, *J. Am. Chem. Soc.* **53**, 1367 (1931).