DISSERTATION

Dislocation Modeling in III-Nitrides

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von

Raffaele Alberto Coppeta

[redacted]

Wien, im June 2015

Kurzfassung

Galliumnitrid und seine Legierungen mit Aluminium und Indium sind direkte Halbleiter und weisen hohe Härte sowie hohe thermischer und mechanischer Stabilität auf. Die Anwendungsgebiete dieser Halbleiter sind breit gefächert. Eine typische Anwendung sind High-Electron-Mobility Transistoren welche die hohe Mobilität des zweidimensionalen Elektronengas ausnützen. Durch die Anwendung von Galliumnitridlegierungen weisen sie hohe Durchbruchsspannungen auf und sind deswegen für die Leistungselektronik von Interesse.

Durch das Fehlen von nativen Substraten mit passenden Gitterstrukturen werden Bauelemente auf Galliumnitridbasis auf fremdartige Substrate aufgewachsen. Dabei entsteht eine hohe Dichte von Defekten, welche die Leistungsfähigkeit diese elektronischen Bauelemente negativ beeinträchtigen. Eine dieser Klassen von Defekten sind Versetzungen, welche eine besonders negative Wirkung auf die Leistungsfähigkeit von High-Electron-Mobility Transistoren habe. Um die Dichte dieser Defekte zu verringern werden Mehrschichtstrukturen mit Schichten unterschiedlicher Dicken und Stoffzusammensetzungen verwendet.

Das Ziel dieser Arbeit ist die Erarbeitung von Entwurfsregeln um die Kristallqualität zu erhöhen oder mit anderen Worten die Defektdichte in Galliumnitrid basierten Bauelementen zu senken. Unter Verwendung der Kontinuumsmechanik zur Modellierung der Defekte im Rahmen der linearen Elastizität und der Theorie der Thermodynamik wurde ein tieferes Verständnis über das Verhalten von Versetzungen in diesen Strukturen erlangt.

Diese Arbeit ist in sechs Kapiteln gegliedert. Kapitel 1 begründet den weiten Anwendungsbereich von Galliumnitrid und dessen Legierungen in der Elektronik und zeigt das Versetzungen ein primärer Faktor für die Herabsetzung der Leistungsfähigkeit von Bauteilen auf Galliumnitridbasis darstellen. Kapitel 2 handelt von der Theorie der linearen Elastizität in der Modellierung von Versetzungen. In Kapitel 3 wird die allgemeine Elastizitätstheorie auf die Berechnung der Versetzungsenergie und der Gleichgewichtskonfiguration angewendet. In Kapitel 4 wird die Theorie zur Berechnung der kritischen Dicke auf Galliumnitridlegierungen erweitert. Dieser Ansatz basierend auf den Reaktionen zwischen Versetzungen wird in Kapitel 5 auf verschiedenen Mehrschichtstrukturen angewendet um die Dichte der Versetzungen in diesen zu bestimmen. Kapitel 6 fasst die Ergebnisse zusammen.

Abstract

Gallium nitride and its alloys with aluminium and indium are hard, thermally and mechanically stable direct band gap semiconductors used for a wide range of applications. A recent application is the high electron mobility transistor which is the result of the high two-dimensional electron gas density and the large breakdown field exhibited by using these materials.

Due to a lack of a suitable native lattice matched substrate, gallium nitride based devices are grown upon a foreign substrate causing the development of a high density of defects which damage the performance of the device. A particular class of defects called dislocations has a deleterious effect on high electron mobility transistors. In order to achieve desired specifications, it is necessary to reduce the dislocation density by using multilayered structures with varying geometry and composition.

The goal of this work is to define design rules to improve the crystalline quality, i.e., to reduce the dislocation density, of gallium nitrided based structures. Continuum theory of dislocations treated within the linear elasticity theory and the laws of thermodynamics are used for gaining understanding and modeling the dislocation development in these structures.

This work is structured in six chapters. Chapter 1 describes the reasons why GaN and its alloys are widely used in electronics, and further introduces the dislocations as a primary factor for damage of the device performance. Chapter 2 introduces the elements of linear elasticity theory useful to model dislocations. In Chapter 3, general elasticity theory is applied to evaluate the dislocation energy and their equilibrium configuration. In Chapter 4, theoretical studies of the critical thickness are extended for GaN based alloys. Subsequently (Chapter 5) the reaction-kinetic approach is used to evaluate the dislocation density in different multilayered structures. The main results and conclusions are summarized in Chapter 6.

Acknowledgment

My greatest gratitude belongs to Professors Tibor Grasser and Hajdin Ceric for their continuous patience, time, and support of my work through a countless number of discussions, suggestions, comments, and encouragements.

A huge thanks to Dr. David Holec for his kind collaboration and guidance. Calculations in Chapter 3 are based on his code, and the work in Chapter 4 is inspired and partially supported by him.

I would like to acknowledge Dipl.-Ing. Wolfhard H. Zisser for his continuous support through the entirety of my Ph.D. project, in particular his contribution to the discussion and implementation of the model described in Chapter 5.

During the course of my Ph.D. project I had the privilege of discussing my ideas and results with many people who consequently helped my research. For that I would like to express my sincere gratitude to Prof. M. Ciavarella, Prof. L. B. Freund, Prof. H. Gao, Prof. C. Humphreys, Dr. A. Köck, Prof. A. Krost, Prof. N. Neophytou, Prof. S. Raghavan, Prof. J. Vlassak, Prof. J. R. Willis, Dr. J. Baumgartl, Dr. T. Detzel, Dr. O. Haeberlen, Dr. B. Karunamurthy, Dr. G. Kravchenko, Dr. Lavanga, Dr. G. Pozzovivo, Dr. G. Prechtl, G. Galasso, and G. Vaccari.

I would like to thank my parents, Aurelia and Ferdinando, and my sister Sarida, for supporting me spiritually throughout my entire life. Without them I would never have reached such results. Grazie infinite!

I am grateful for Christ's provision of Life and Reason.

Contents

List of Chemical Formulae

AlN	Aluminium nitride
$Al_xGa_{1-x}N$	Aluminium gallium nitrogen alloy
Al_2O_3	Aluminium oxide
AlP	Aluminium phosphide
GaAs	Gallium arsenide
GaN	Gallium nitride
GaP	Gallium phosphide
Ge	Germanium
InN	Indium nitride
$In_xGa_{1-x}N$	Indium gallium nitrogen alloy
InP	Indium phosphide
MgO	Magnesium oxide
Si	Silicon
SiC	Silicon carbide
Si_xN_y	Silicon nitrogen alloy

List of Symbols

α	Angle between the [0001] direction and the dislocation line in polar spherical coordinate
β	Angle of the polar spherical coordinate
$\mathbf{a}_i$, a	Hexagonal lattice vectors and their length
$\mathbf{b}$, b	Burgers vector and its length
b_s, b_e	Screw and edge components of the Burgers vector
$\mathbf{c}$, c	Hexagonal lattice vector along [0001] and its length
γ	Angle between the dislocation line and the normal to the free surface
C	Stiffness tensor
c_{ij}	Elastic constants
Δ	Fractional change in volume
δ	Dislocation slip plane
$\mathrm{d}F$	Differential change in Helmholtz free energy
ε_{ij}	Strain components
$\boldsymbol{\varepsilon}^d$	Strain tensor associated to the dislocation
$\mathcal{E}$	Dislocation energy
$\mathcal{E}_\mathrm{d}$	Elastic part of the dislocation energy
$\mathcal{E}_\mathrm{core}$	Dislocation core energy
$\mathbf{f}$	Direction of the upper point of the threading dislocation
ζ_x, ζ_y	Shear angles
η_{ij}	Shear strain components
θ	Angle between the Burgers vector and the dislocation line
$h = h_\mathrm{f}$	Layer thickness
H	Bulk modulus
h_c	Critical thickness
ι	Angle of the polar spherical coordinate
K	Rotation matrix
K	Pre-logarithmic coefficient of the dislocation energy
$K^k_{G,ij}$	Reaction rate of two glissile threading dislocations of i and j types
$K^k_{GS,ij}$	Reaction rate of one glissile threading dislocation of i type and one glissile threading dislocation of j types
$K^k_{S,ij}$	Reaction rate of two sessile threading dislocations of i and j types
λ	Lamé constant

$\mathbf{l}$	Dislocation line direction
μ	Shear modulus or Lamé constant
$\mathbf{m}$	Direction of the misfit dislocation
ν	Poisson's ratio
$\mathbf{n}_\mathrm{f}$	Normal of the free surface
$\mathbf{n}_\mathrm{g}$	Normal of the slip plane
Ω, Ω_{ij}	Transformation tensor and components
p	Effective pressure acting on a volume element
$\rho = \rho_\mathrm{TD}$	Total threading dislocation density
ρ_G	Glissile threading dislocation density
ρ_i	Density of the i-type threading dislocations
ρ_MD	Misfit dislocation density
ρ_S	Sessile threading dislocation density
R	Outer cut-off radius
r_c	Radius of the dislocation core or inner cut-off radius
ς	Angle of the polar spherical coordinate
σ_{ij}	Stress components
σ_Y	Standard deviation of Young's modulus
$\boldsymbol{\sigma}^d$	Stress tensor associated with the dislocation
$\boldsymbol{\sigma}^m$	Tensor associated with the misfit stress
S	Compliance tensor
s_{ij}	Elastic compliances
u	Interatomic distance in the wurtzite structure
$\mathbf{u}$, u_{ij}	Displacement vector and components
ϕ	Angle between the slip plane and the normal to the film-substrate interface
$\mathbf{v}_\mathrm{G}$	Velocity of the upper point of one glissile TD along the free surface
$\mathbf{v}_\mathrm{S}$	Velocity of the upper point of one sessile TD along the free surface
$\mathcal{W}$	Work done by the misfit stress
w	Strain energy per unit of volume
x	Composition of an alloy
χ	Angle associated with pure tilt boundary
Y	Elastic or Young's modulus
$\langle Y \rangle$	Average elastic modulus
ζ	Angle associated with pure tilt boundary

List of Figures

List of Tables

1 Introduction

1.1 Introduction: gallium nitride and related alloys (III-nitrides)

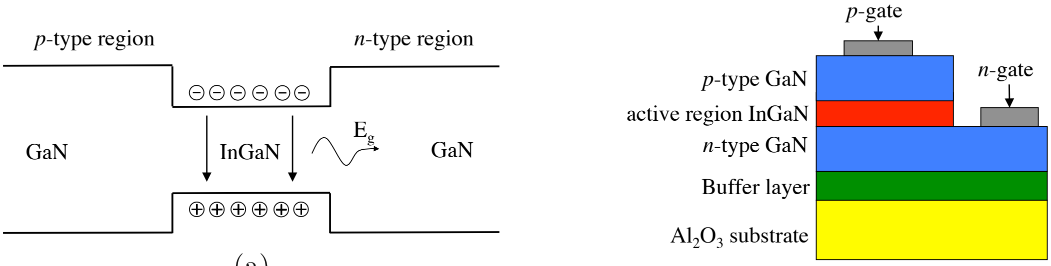
Gallium nitride (GaN) and its alloys with aluminum or indium are in their stable form wurtzite direct band gap semiconductors which have become the most important since silicon. The excitement over III-nitrides stems from their better material and electronic properties compared to silicon and other III-V compounds, like GaAs. These properties enable the use of GaN based devices in a broad range of applications in various fields, including the automotive, military and space industries as well as in high power amplifiers for wireless base stations, and high voltage electronics for power transmission lines.

In this chapter the most used GaN based devices - light emitting diodes and high electron mobility transistors - are briefly described (see Section 1.2) together with the physical properties (see Section 1.3) that make them well suited for industrial applications. A particular class of defects, the dislocations (see Paragraph 1.4.1), greatly affects the electrical performance of GaN devices. In order to reduce their impact, it is necessary to reduce their density. The goal of this work is therefore to define some design rules for dislocation filtering in devices based on III-nitrides (see Section 1.5).

1.2 GaN applications and characteristics

GaN and its alloys with aluminum or indium are usually grown at a high temperature (approximately 1100°C) by metal organic chemical vapor deposition or molecular beam epitaxy techniques on foreign substrates [59]. These substrates include silicon carbide for RF applications, silicon for power electronic applications, and sapphire for optical devices [54,67,73]. GaN and its alloys have a wide range of optoelectronic uses because the atoms are bonded by a very ionic gallium/aluminum/indium-nitrogen chemical bond [53]. This has two important consequences.

Foremost, AlN, GaN and InN have respective direct band gap energies of 6.2 eV, 3.4 eV and 0.7 eV at room temperature [1,11], which cover the entire visible spectrum from the ultraviolet to the infrared range (see Figure 1.2). This is in contrast to (Ga,Al)As based light emitting diodes (1.5-2.2 eV) and (Al,Ga,In)P based light emitting diodes (1.3-2.5 eV) which cover only the infrared to green region [25].



(a)

(b)

Figure 1.1: Schematic of a GaN based LED structure. The energy of the photon is equal to the energy gap E_g.

The second main consequence of the strong chemical bond is their stability and resistance to degradation under high electric currents and high temperatures. Therefore, III-nitrides are widely used in laser diodes and light emitting diodes [2, 56].

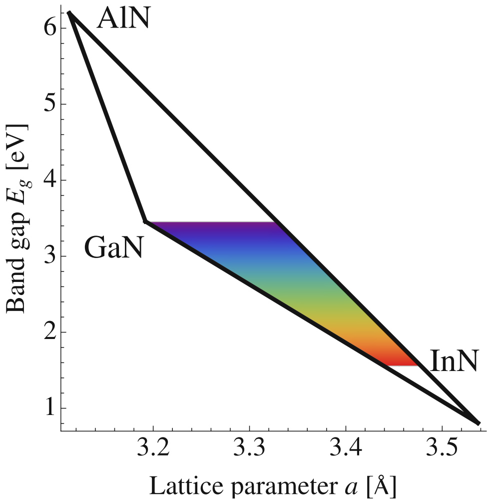
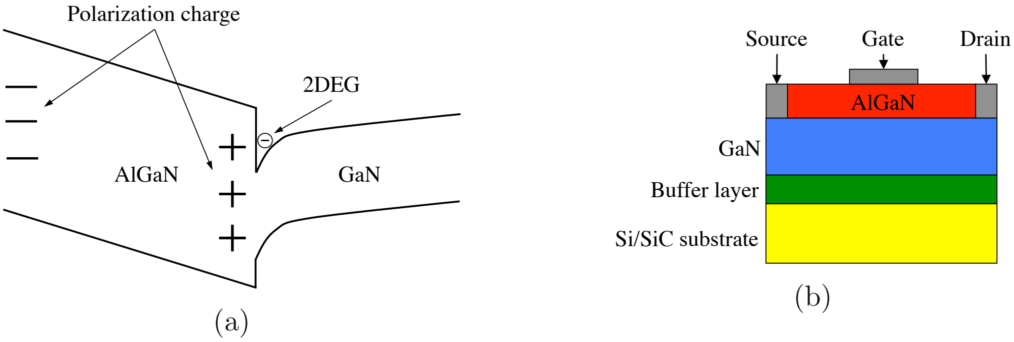


Figure 1.2: The band gap is shown as function of the lattice parameters. The range of visible spectrum is shown in colors.

GaN-based laser diodes enabled the development of Blu-ray technology which has become the standard replacement of DVDs. The previous GaAs-based red laser diodes (working at $\lambda = 650$ nm) have been replaced by GaN-based laser diodes emitting blue light ($\lambda = 405$ nm) in order to store more data on every single disk due to the shorter wavelength of the writing beam [25].

A second optoelectronic application of III-nitrides are the light emitting diodes. Compared to a traditional light source, light emitting diodes have many prominent advantages including their low energy usage, long lifetime and small size, making such a light source a strong competitor in the domestic and business lighting market. Because lighting accounts for about 20% of total electricity consumption, national



(a) (b)

Figure 1.3: Schema of GaN based HEMT structure. The bi-dimensional electron gas (2DEG) is along the interface between the AlGaN and GaN layers.

programs promoting light emitting diodes for lighting are underway in the US, Japan, and China. The main motivation is large-scale energy savings which would provide the benefits of reduced oil imports and lower greenhouse gas emissions [25]. Apart from optoelectronic applications, another important III-nitride-based device is the high-electron mobility transistor. Power amplifiers based on high electron mobility transistors have demonstrated high break-down voltage, high current density, and low thermal resistance [35, 83, 86]. The combination of these attributes makes GaN technology attractive for continual development of high power products for applications which include commercial and military radar, electronic warfare, radio communications, and data backhaul. The above mentioned characteristics are consequences of some particular physical properties of III-nitrides [52].

An important property of III-nitrides is the strength of their chemical bond, which implies a large energy band-gap. Therefore GaN-based devices have a high breakdown field, which allows them to operate at much higher voltages than other semiconductor devices. Under such conditions, the device is subjected to high enough electric fields such that the electrons in the semiconductor can acquire enough kinetic energy to break the chemical bond (a process called voltage breakdown). If impact ionization is not controlled, it can degrade the device. Since the chemical bond is strong, III-nitride-based devices can operate at higher voltages, which are typical of higher-power applications [52].

A second consequence of the strong chemical bond is a high saturation velocity. This is the velocity of electrons at very high electric fields. High saturation velocity means that GaN devices can deliver much higher current density. The radio frequency power output is the product of the voltage and the current swings, so a higher voltage and current density can produce higher radio frequency power in a practically sized transistor. Simply put, GaN devices can produce much higher power density [52].

Low thermal resistance arises in GaN devices grown upon SiC substrates. They exhibit outstanding thermal properties, due to the high thermal conductivity of SiC. In practical terms, this means that GaN-on-SiC devices do not get as hot as GaAs or Si devices when dissipating the same power. A colder device is more reliable [52].

Let us consider the structure of a high electron mobility transistor shown in Figure 1.3(b). The simple structure consists of a thin AlGaN layer which is deposited on top of a thick GaN layer in order to achieve a bi-dimensional electron or hole gas at the interface between the two materials [36, 68]. One of the most interesting things in GaN-based high electron mobility transistor is that a bidimensional electron or hole gas forms at the AlGaN/GaN interface without modulation doping [52]. This happens due to the spontaneous polarization and strain induced piezoelectric fields. GaN is piezoelectric because the gallium-nitrogen bond is ionic and the successive planes of gallium and nitrogen atoms are not at equal distance. When the atoms in a plane are displaced from their original position (by a mechanical pressure, for example), the planes of atoms above and below move to different positions, creating a net charge, an electric field, and a voltage. This bends the energy bands in a way that a bidimensional electron sheet is created at the interface of AlGaN/GaN or InGaN/GaN without the necessity of doping, like in Si-based transistors. Normally, the bi-dimensional carrier sheet density is around $10^{13}/\text{cm}^2$ and the mobility is around 1600 cm/Vs at room temperature [69], making GaN-based transistors fast and efficient.

1.3 Physical properties of III-nitrides

1.3.1 Crystal structure

Group III nitrides can have the following crystalline structures: the wurtzite and zinc blende. Under ambient conditions, the thermodynamically stable structure is wurtzite for bulk AlN, GaN, and InN. The zinc blende structure for GaN and InN has been stabilized by epitaxial growth of thin films on {011} crystal planes of cubic substrates such as Si, SiC, MgO, and GaAs. In these cases, the intrinsic tendency to form the wurtzite structure is overcome by topological compatibility [53].

The wurtzite and zincblende structures are somewhat similar and yet different. In both cases, each group of III atoms is coordinated by four nitrogen atoms. Conversely, each nitrogen atom is coordinated by four group III atoms. The main difference between these two structures lies in the stacking sequence of closest packed diatomic planes. The wurtzite structure consists of alternating biatomic close-packed (0001) planes of Ga and N pairs, thus the stacking sequence of the (0001) plane is AaBbAa in the (0001) direction. Another way to understand the wurtzite structure is to consider the anions (N^{3-}) forming a hexagonal close-packed structure in which the cations (Ga^{3+}) occupy half of the tetrahedral sites [53].

The zincblende structure occurs when the hexagonal double-layers are stacked in a periodic AaBbCcAaBbCc. For AlN, GaN and InN the zincblende structure is metastable while the wurtzite variant is stable and easier to grow [25, 53]. Therefore, scientific attention and industrial interest have been focused on the wurtzite form. This work considers consequently III-nitrides in their wurtzite structure.

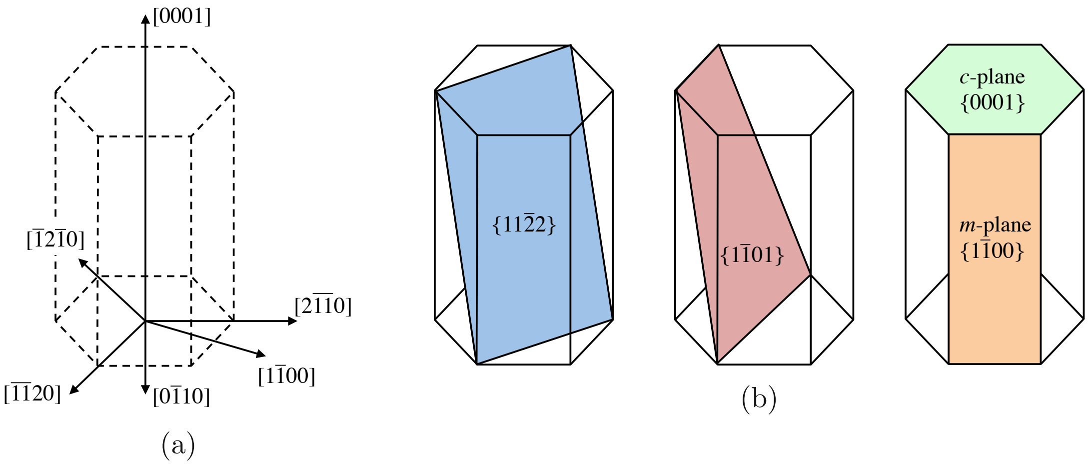


Figure 1.4: Directions and planes in a hexagonal lattice.

The wurtzite group III nitrides lack an inversion plane perpendicular to the c-axis; thus, nitride surfaces have either a group III element (Al, Ga, or In) polarity with a designation of (0001)A plane or a N-polarity with a designation of (0001)B plane [53]. The distinction between these two directions is essential in nitrides because of their implications for the polarity of the polarization charge.

Four surfaces are of special importance in nitrides: (0001), $(11\bar{2}2)$, $(1\bar{1}00)$, and $(1\bar{1}01)$ planes (see Figure 1.4(b)). Particular directions in hexagonal close-packed structures are $\langle 0001\rangle$, $\langle 2\bar{1}\bar{1}0\rangle$, and $\langle 1\bar{1}00\rangle$ as shown in Figure 1.4(b). The (0001), or the basal plane, is the most commonly used surface for growth. The three-dimensional arrangement of wurtzite nitrides is shown in Figure 1.5 where the red color represents nitrogen atoms while blue represents the group III atom sites (Al, Ga or In).

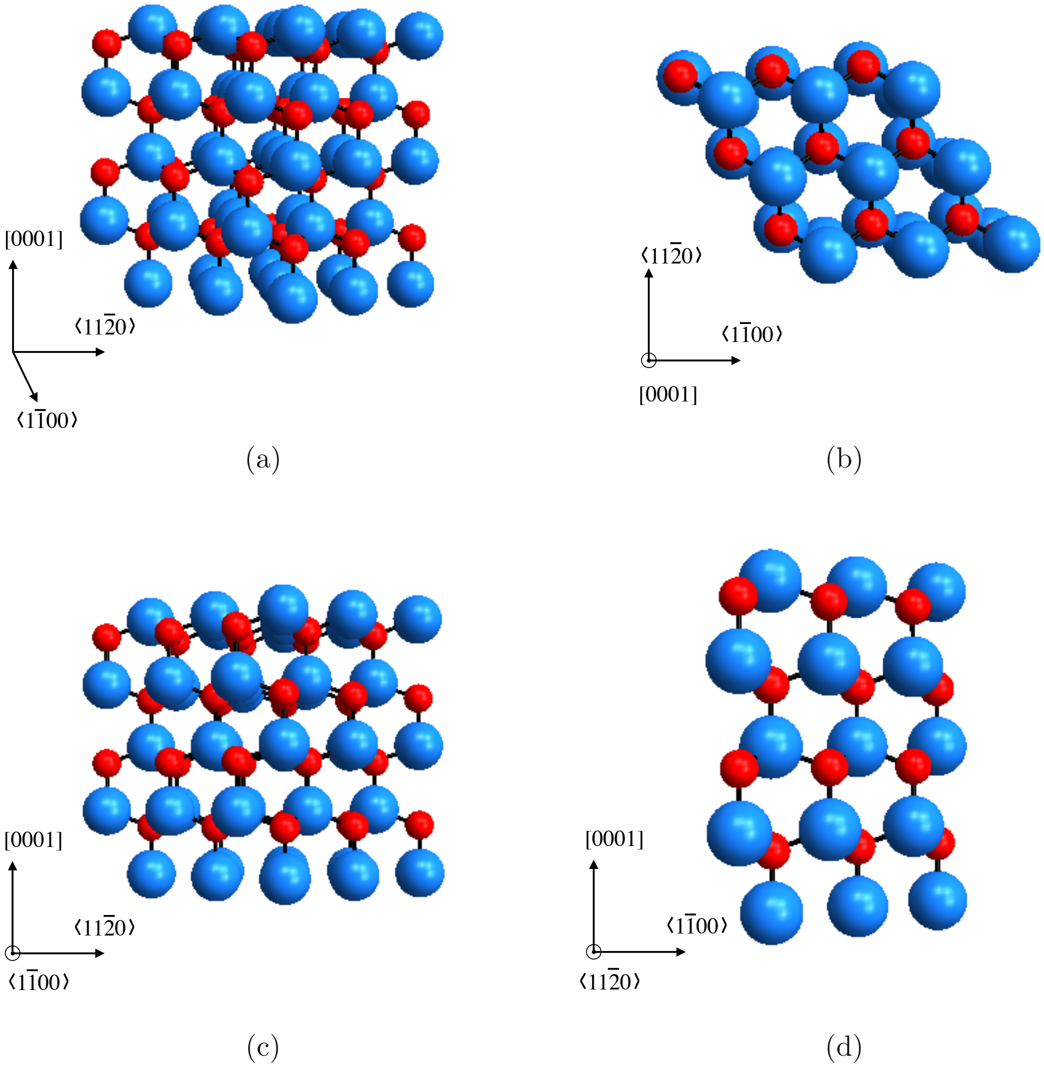
The wurtzite structure can be represented by lattice parameters a in the basal plane and c in the perpendicular direction, and the internal parameter u. The basal plane lattice parameter (the edge length of the basal plane hexagon) is universally depicted by a and the axial lattice parameter, perpendicular to the basal plane, is universally described by c. The c parameter depicts the unit cell height. The u parameter is defined as the anion-cation bond length (also the nearest neighbor distance or the interatomic distance) divided by the c lattice parameter. In an ideal wurtzite structure represented by four touching hard spheres, the values of the axial ratio and the internal parameter are $c/a = \sqrt{8/3} = 1.633$ and $u = 8/3 = 0.375$, respectively. The crystallographic vectors of wurtzite are $\mathbf{a}_1 = a(1/2, \sqrt{3}/2, 0)$, $\mathbf{a}_2 = a(1/2, -\sqrt{3}/2, 0)$, $\mathbf{c} = a(0, 0, c/a)$. In Cartesian coordinates, the basis atoms are $(0, 0, 0)$, $(0, 0, u \cdot c)$, $a(1/2, \sqrt{3}/6, c/2a)$, and $a(1/2, \sqrt{3}/6, [u + 1/2]c/a)$ [53].

The lattice parameters of the III-nitrides are summarized in Table 1.1. Whenever a material property of the alloy is needed, for example, the lattice constant of $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$, Vegard's law [76] is used in the following form:

$$a_{\mathrm{Al_xGa_{1-x}N}} = x\, a_{\mathrm{AlN}} + (1-x)\, a_{\mathrm{GaN}} \tag{1.1}$$

Table 1.1: The lattice constants (in Å) for GaN, AlN, InN [53], and for Si and Ge [19]. The lattice constants of $Al_xGa_{1-x}N/GaN$, $In_xGa_{1-x}N/GaN$ and $Si_{1-x}Ge_x/Si$ are calculated using Vegard's law with the data in the table.

	GaN	AlN	InN	Si(111)	Ge(111)
a	3.22	3.11	3.54	5.43	5.65
c	5.19	4.98	5.96		



(a) (b)

(c) (d)

Figure 1.5: A 3D picture of the wurtzite structure together with some important projections [12].

Vegard's law is valid only under specific conditions [76]. First principle calculations show [10] that Vegard's law holds for AlGaN (both lattice constants) and the *a*-lattice constant of InGaN and AlN. Some deviations from Vegard's law are predicted [10] for the *c*-lattice constant of the latter two alloys. However, since no better estimate of lattice and elastic constants is available for the whole range of compositions, Vegard's law is used throughout this work wherever necessary [25].

1.3.2 Mechanical properties

The elastic constants of GaN, AlN, and InN are given in Table 1.2. Vegard's law is used to calculate the constants of the ternary compounds $Al_xGa_{1-x}N$ and $In_xGa_{1-x}N$. Epitaxial layers composing GaN-based devices are monocrystals. Therefore elastic anisotropy must be considered in any attempt of modeling the mechanical behavior. In order to show the importance of anisotropy, the standard deviation σ_Y of the elastic modulus Y is calculated as a function of the x fraction for the two systems $Al_xGa_{1-x}N$ and $In_xGa_{1-x}N$:

$$\sigma_Y = \frac{2}{\pi}\sqrt{\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} \left(Y_i(x) - \langle Y\rangle\right)^2 d\iota\, d\varsigma}, \tag{1.2}$$

where Y_i is the elastic modulus function of x in a certain direction of the 3D space. ι and ς are the angles of the polar spherical coordinate system. $\langle Y\rangle$ is the average elastic modulus defined as follows:

$$\langle Y\rangle = \frac{4}{\pi^4}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} Y_i(x) d\iota\, d\varsigma. \tag{1.3}$$

The standard deviation of the elastic modulus quantifies the dispersion from the average value. For this reason, it is an index of the anisotropy of the compounds. The comparison of σ_Y for the three systems considered in this work is shown in Figure 1.6.

Table 1.2: The stiffness constants (in GPa) for GaN, AlN, InN [61], for Si [29], for Ge [19] . The stiffness constants of the systems $Al_xGa_{1-x}N/GaN$, $In_xGa_{1-x}N/GaN$ and $Si_{1-x}Ge_x/Si$ are calculated using Vegard's law with the data in the table.

	c_{11}	c_{12}	c_{13}	c_{33}	c_{44}	c_{66}
GaN	374.2	141.4	98.1	388.6	98.3	$(c_{11}-c_{12})/2$
AlN	410.5	148.5	98.9	388.5	124.6	$(c_{11}-c_{12})/2$
InN	223	115	92	224	48	$(c_{11}-c_{12})/2$
Si(111)	194.25	35.25	63.9	165.6	79.5	50.85
Ge(111)	155	21.6	48.2	128.4	66.7	40.1

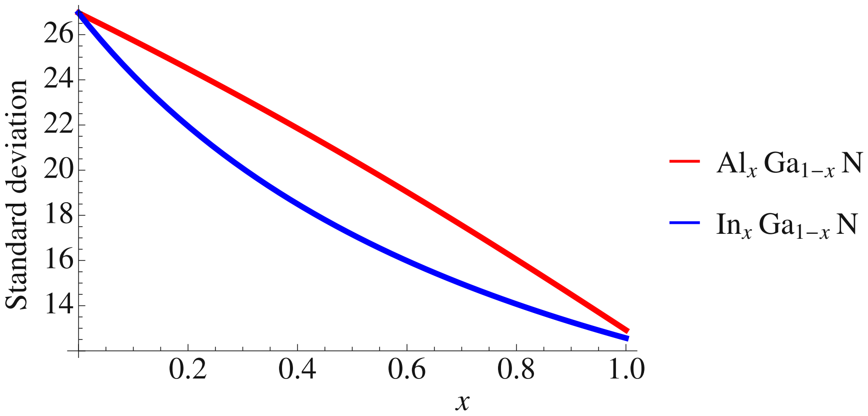


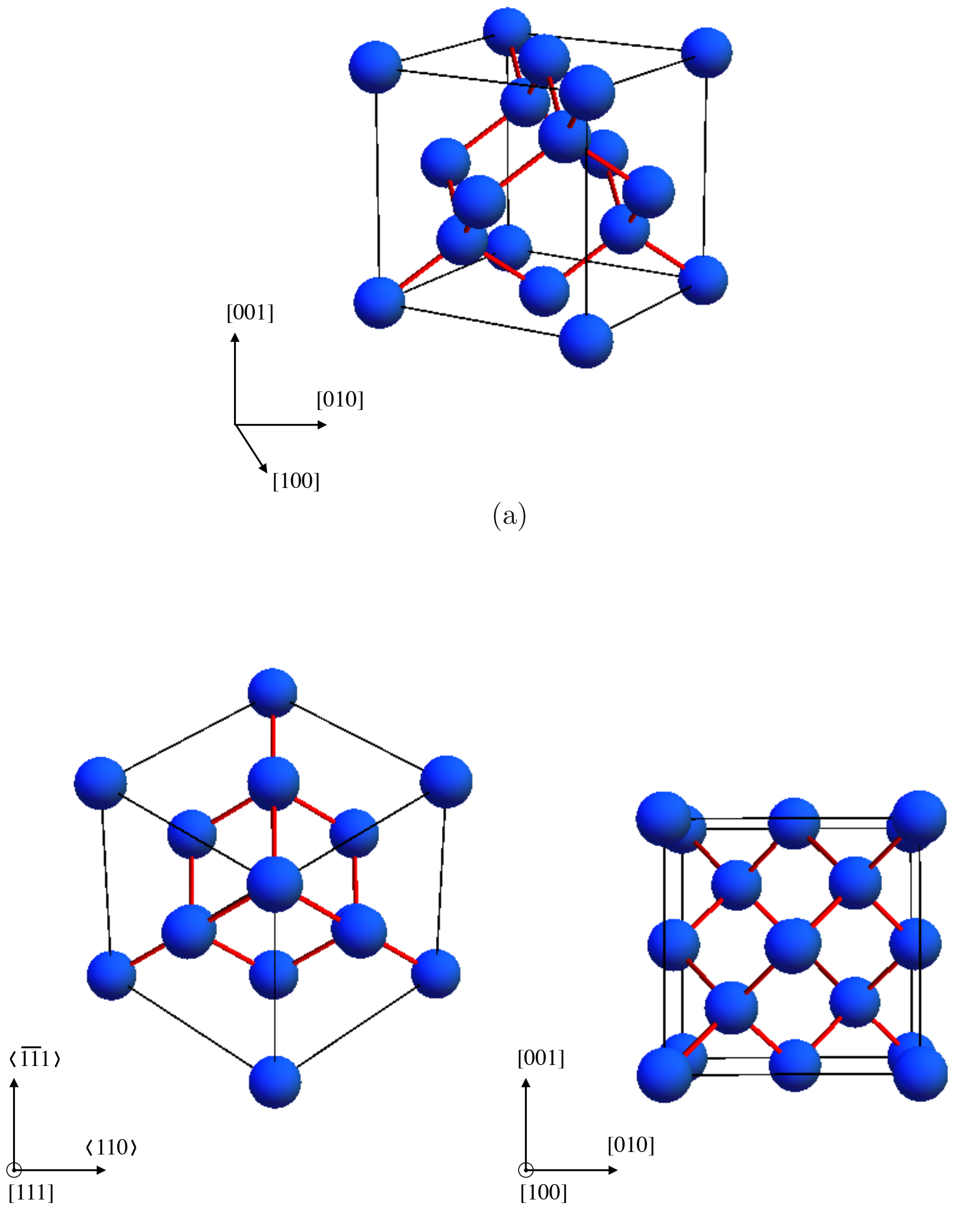
Figure 1.6: The standard deviation σ_Y as function of the fraction x for the three different systems.

Figure 1.6 shows that for all the systems the impact of the anisotropy decreases as x increases, meaning that small amounts of Al/In atoms increase strongly the effect of the anisotropy. For the wide range of composition, the standard deviation of the elastic modulus is always higher for the $Al_xGa_{1-x}N$ system with respect to the $In_xGa_{1-x}N$ system. As a consequence of the values of the standard deviation, elastic anisotropy is considered in the subsequent modeling.

1.3.3 Hetero-epitaxial relations with substrates

Different types of substrates are used to grow III-nitrides depending on the application. While sapphire is mainly used for optical devices and silicon carbide for military application, silicon is preferred for civil use. A silicon substrate causes the creation of a large amount of defects and thus the resulting performance of the final devices is lower than when SiC is used. However, Si substrates are much cheaper than SiC, and this is the main factor in requests for mass production for civil application.

For growth of III-nitrides on Si, one generally starts with an AlN buffer layer and hence an AlN/Si interface. Direct growth of GaN on Si is prevented by the formation of low-melting-point Ga-Si compounds and significant diffusion of Si to the film-vapor interface resulting in the formation of Si_xN_y therein [73]. The former results in the so-called melt-back etching, whereas the latter results in a rough interface. The presence of a defective Si_xN_y layer has been reported at the AlN/Si interface as well. However, it does not prevent epitaxy and it is still not quite clear if it forms prior to deposition of the AlN buffer layer or during subsequent exposure of the AlN/Si interface to high temperatures during growth [73].



(a)

(b)

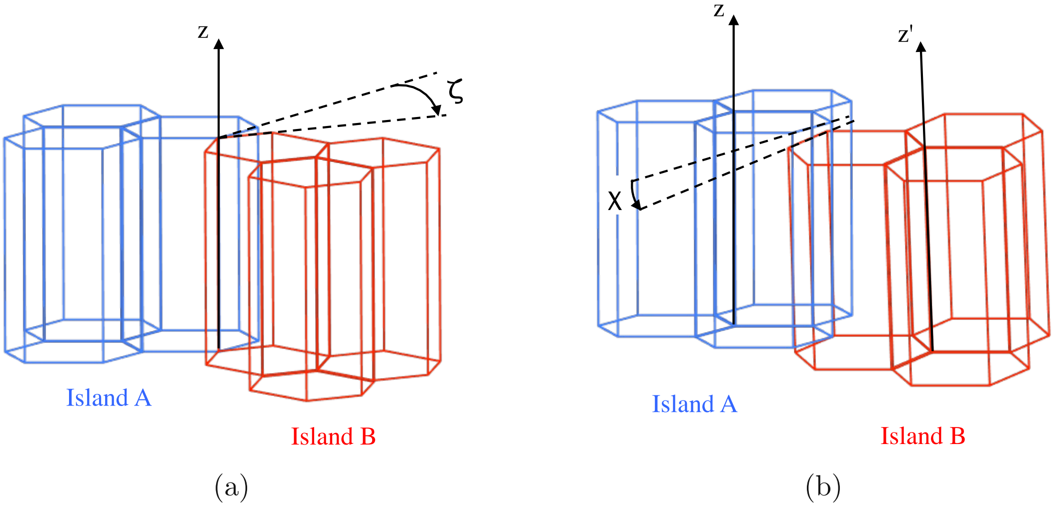
(c)

Figure 1.7: A 3D picture of the diamond structure together with some important projections [12]. Red lines represent the chemical bonds and the black lines are cell lines.

The (111) plane is the most commonly used surface of Si for growth of the III-nitrides. The epitaxial relationship is $\{111\}_{\mathrm{Si}}\langle 1\bar{1}0\rangle_{\mathrm{Si}} \parallel \{0001\}_{\mathrm{GaN}}\langle 11\bar{2}0\rangle_{\mathrm{GaN}}$. The sixfold symmetry of arrangement of atoms on the unreconstructed {111} surface (see Figure 1.7(b)) favors the growth of the wurtzite phase with hexagonal symmetry of Ga-polarity (Ga face up). However, the mismatch, defined as $(a_{\mathrm{substrate}} - a_{\mathrm{film}})/a_{\mathrm{film}}$, with a being the lattice parameter, is greater than 20% for both GaN and AlN, and of a nature that would result in tensile (or positive) stress in the film when grown directly on Si(111) [73]. Given the orientation relationship, it should be noted that for Si(111), $a_{\mathrm{substrate}}$ should be the interatomic distance along $[1 1 \bar{0}]$ and not the often quoted lattice parameter of Si, which is the interatomic distance along [100]. For such highly mismatched systems, strain is very quickly relaxed by the formation of islands, i.e., the Volmer-Weber growth mode takes place. A quick analysis of the various energies involved in the system then provides a deeper insight into the possible growth mode. Markov [48] determines the mechanism by which a thin film grows by considering the chemical potentials μ of the first few n deposited layers. The chemical potential is a function of the bond energies involved in the system. The island growth mode, i.e., the Volmer-Weber mode, takes place when the variation of the chemical potential within the first deposited layers $\mathrm{d}\mu/\mathrm{d}n$ is negative, which corresponds to the situation when the bond energy among the atoms of the deposited material is higher than the bond energy between the atoms of the deposited material and the atoms of the substrate. The Si-N bond energy of 4.5 eV [78] is lower than the Ga-N and Al-N bond energies of 9.12 and 1.17 eV, respectively [53]. Thus, there is no great incentive for GaN/AlN to spread out along the substrate or, in other words, for the two-dimensional growth mode. In addition, as discussed previously, III-nitride layers are strained, which would raise the chemical potential of the first deposited layer with respect to the layers composing the bulk even further.

In conclusion, the chemical potential of the first GaN/AlN monolayers is higher than that of the subsequent layers. Thermodynamics indicates that in such situations III-nitrides would prefer to form multilayer islands rather than spread out as a single layer on the substrate and for that matter even on other III-nitride layers. The resultant growth mode, based on observations at a thickness greater than a bilayer, is thus the Volmer-Weber mode, involving nucleation of three-dimensional (3D) islands that then grow laterally and coalesce. It is emphasized that this discussion has been on purely thermodynamic grounds and the growth mode actually observed could be subject to kinetic modifications.

After the coalescence, the islands become domains of the continuous film. Due to different crystallographic orientation, the domains can be twisted or tilted with respect to each other to varying extents (see Figure 1.8). The consequence is that the domains are separated by low angle grain boundaries, which are composed by sets of dislocations. The difference in tilt between two domains can be accommodated by pure edge dislocations, whereas difference in twist can be accommodated by pure screw dislocations [73]. When a combination of both a difference in twist and tilt between domains exist, then all three types of dislocations (edge, screw, and mixed dislocations) would be required. Dislocations created in this way during the growth of the first III-nitride layer on Si(111) propagate through the above grown layers, reaching the active part of the device and damaging its performance. When Si is used as substrate, an extremely high density of dislocations is detected in the active part in the range between 10^{10}-10^{14}cm^{-2}. X-ray diffraction (XRD) using a four-circle diffractometer [51], TEM [51], and atomic force microscopy (AFM) [58] are the three most commonly used techniques to characterize these dislocations.



(a)

(b)

Figure 1.8: The islands A and B separated by (a) a pure tilt boundary represented by the angle ζ and (b) a pure twist boundary represented by the angle χ.

1.4 Defects in III-nitrides

III-nitride films on Si are host to all major defect types including point defects, line defects or dislocations, plane defects, such as stacking faults, inversion domains and cracks, and 3D defects, such as V-shaped pits and pipes [53]. Despite defects affecting both device performance and lifetimes, III-nitride-on-Si devices have been commercialized in the last decade [73]. Extensive research has thus been devoted to understanding the origin, structure, and evolution of defects during growth.

Of the defects mentioned above, III-nitride devices stand apart because of their high density of dislocations. Dislocations propagate through the bulk of the heterostructure and reach the core part of the device. There they form new energy levels inside the energy band. Electrons can be trapped in these energy levels by reaching a lower energy state and thus damaging the electrical performance of the device. In addition, the evolution of dislocation configuration during growth alters the strain-stress fields of the structure. This is a second cause for the modification of the band structure.

The dislocation density in 1μm thick GaN layers grown on Si(111) is typically 10^{10}cm^{-2} at the surface of the film, and the corresponding number for films grown on sapphire and SiC is 10^{8}cm^{-2}. It is important to note the role of film thickness, as dislocation density is not a constant but varies across the film. Considering a structure composed by a GaN film grown upon an AlN layer on a Si(111) substrate, at the GaN/AlN interface the dislocation density can be in excess of 10^{14}cm^{-2} and its reduction with thickness depends on the process type and the device structure. A particular type of growth technique called lateral growth technique reduces the dislocation density to about 10^{6}cm^{-2} [21] but this is still about two orders of magnitude higher than the value found in GaAs-based devices.

1.4.1 Dislocations

A dislocation is a line defect in a crystal, which can greatly affect its material properties. A dislocation is usually represented by an oriented dislocation line, **l**, and characterized by its Burgers vector, **b**, describing displacements introduced in the crystal by the dislocation [25].

A dislocation can not end within a crystal and it is contained within a crystal only as a complete loop. Otherwise, the dislocation line can end at the material surface.

Based on the relationship of **l** and **b**, three dislocation types are distinguished (see Figure 1.9): an edge-type dislocation with $\mathbf{l} \perp \mathbf{b}$, a screw-type dislocation with $\mathbf{l} \parallel \mathbf{b}$, and a mixed-type dislocation ($\mathbf{b} \nparallel \mathbf{l} \not\perp \mathbf{b}$). The dislocation type can change as the dislocation line changes its direction in the crystal. This is because the Burgers vector **b** is a constant characteristic of a dislocation which remains unchanged along the whole single dislocation line. It is worth noting that the orientation of **b** in the standard definition (Burgers-Frank or a continuous elasticity definition [9]) depends on the actual orientation of the dislocation line **l**: an opposite dislocation line direction $-\mathbf{l}$ yields an opposite orientation of the Burgers vector $-\mathbf{b}$. However, the uncertainty in sign of **b** and **l** does not affect their mutual relation in terms of defining an edge, screw or mixed type dislocation.

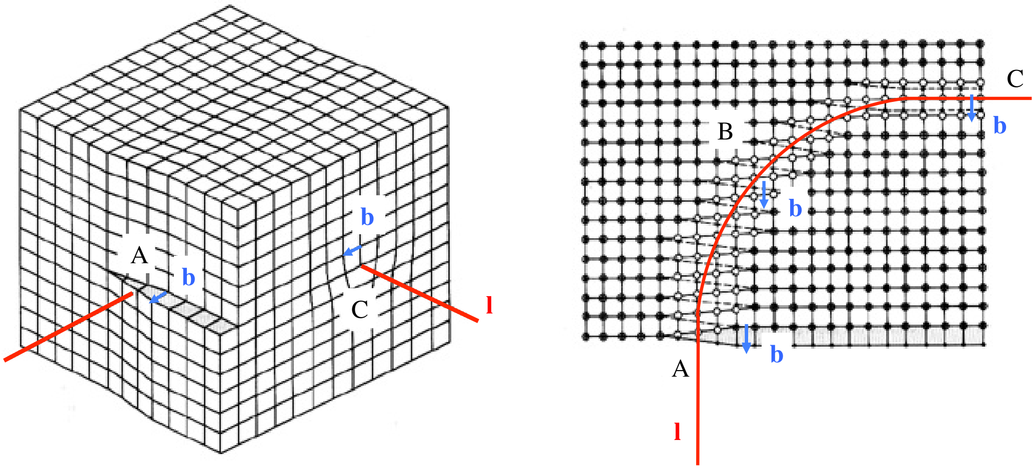


Figure 1.9: Changing dislocation type along a single dislocation line with Burgers vector **b**: screw-type (A), mixed-type (B) and edge-type (C) [80].

The Burgers vector of a perfect dislocation is a lattice vector. As such, there are three basic types of dislocations in III-nitrides according to the Burgers vector: a-type dislocations with $\mathbf{b} = 1/3\langle 11\bar{2}0\rangle$, c-type dislocations with $\mathbf{b} = \langle 0001\rangle$, and $(a + c)$-type dislocations with $\mathbf{l} = 1/3\langle 2\bar{1}\bar{1}3\rangle$. It is important to distinguish between these two nomenclatures: an a-type dislocation can be either an edge, screw or mixed dislocation depending on the actual dislocation line direction. Since the Burgers vector cannot change along the dislocation line, the latter terminology is used in this text for describing dislocation types.

1.5 Outline of the thesis

The goal of this work is to establish some design rules for dislocation filters in III-nitride based devices in order to improve their crystalline quality, i.e., their electrical performance. Linear elastic theory, thermodynamics and the *reaction–kinetics* approach [49] have been used to gain understanding and to develop models of the dislocation density in GaN based heterostructures.

This dissertation is divided into six chapters, including the current introductory chapter. Chapter 2 introduces the elements of linear elasticity theory useful to model dislocations. In Chapter 3 general elasticity theory is applied to evaluate the dislocation energy in hexagonal thin layers. In addition, dislocation energy is used to evaluate the equilibrium configuration of dislocations in III-nitrides. In Chapter 4, the previous results regarding the dislocation energy are used to extend the theoretical studies of the critical thickness for III-nitrides film. Subsequently (in Chapter 5) the reaction-kinetic approach is used to model the dislocation density in III-nitride bulk and multilayers with different geometries. Chapter 6 summarizes the work with the main achievements and presents an outlook for further studies.

2 Elements of elasticity theory

2.1 Introduction

The presence of one dislocation displaces the atoms in the crystal from their perfect lattice sites and the resulting distortion produces a stress field in the crystal around the dislocation. The dislocation is therefore a source of internal strain and stress fields which increase the internal energy of the crystal. For example, consider the edge dislocation in Figure 2.1. The region above the slip plane contains the extra half-plane forced between the normal lattice planes. Consequently, the region above the slip plane is in compression, while the region below is in tension. The stresses and strains produced by dislocations in the bulk of the crystal are sufficiently small for linear elasticity theory to be applied to quantify them. An understanding of elasticity theory is a necessary prerequisite for the development of a quantitative theory for dislocations.

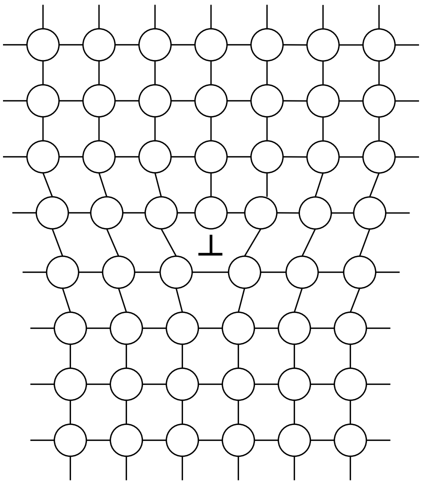


Figure 2.1: Edge dislocation in a cubic crystal.

In this chapter the parts of elasticity theory which are essential for later derivations are reviewed. Section 2.2 is mainly based on the second chapter of the book of Hirth and Lothe [23] and it is integrated with parts from the book of Hull and Bacon [32]. For the complete description of linear elasticity, the reader is referred to the text by Love [45] on which much of the reviews in the above mentioned books are based. The texts of Landau and Lifshitz [40], and Sokolnikoff [70] are useful for supplementary study. A good discussion of the limitations of linear elasticity is found in a review article by Eshelby [13]. In addition, the reader is referred to the text by Balluffi [4] for the application of elasticity theory to lattice defects. The procedure to rotate the stiffness and compliance tensors reported in Section 2.3 is derived from the text by Ting [74].

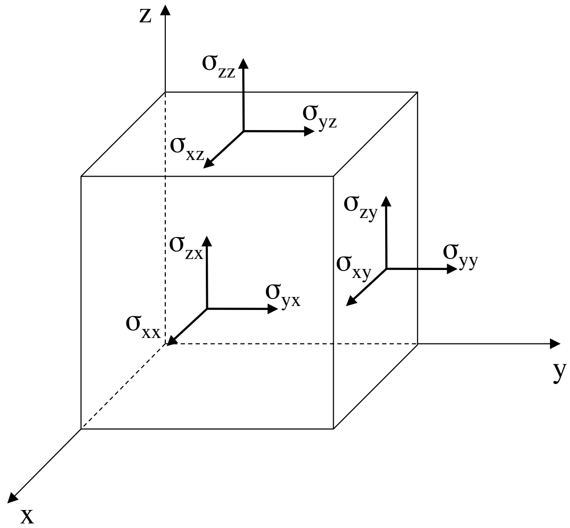


Figure 2.2: Stress distribution on an infinitesimal volume element.

2.2 Fundamental equations

In elasticity theory, an element of volume experiences forces via stresses applied to its surfaces by the surrounding material. Stress is the force per unit area of surface. A complete description of the acting stresses therefore requires not only specification of the magnitude and direction of the force but also of the orientation of the surface, for as the orientation changes so, in general, does the force. Consequently, nine components must be defined to specify the state of stress. They are shown with reference to an elemental cube aligned with the x, y, z axes in Figure 2.2. The component σ_{ij}, where i and j can be x, y or z, is defined as the force per unit area exerted in the $+i$ direction on a face with outward normal in the $+j$ direction by the material outside upon the material inside. For a face with outward normal in the $-j$ direction, σ_{ij} is the force per unit area exerted in the $-i$ direction. For example, σ_{yz} acts in the positive y direction on the top face and the negative y direction on the bottom face.

Each point in a strained body is displaced from its original position in the unstrained state. The point displacement is represented by the vector

$$\mathbf{u} = (u_x, u_y, u_z) \,. \tag{2.1}$$

Let us consider the body in the condition of mechanical equilibrium. In such situation, no net torques taken about x, y and z axes placed through the center of the cube are present in the element, so that

$$\sigma_{ij} = \sigma_{ji}. \tag{2.2}$$

Thus, the order in which subscripts i and j are written is immaterial. The three components σ_{xx}, σ_{yy}, σ_{zz} are the *normal* components. From the definition given above, a positive normal stress results in tension and a negative one in compression. The effective pressure acting on a volume element is therefore

$$p = -\frac{1}{3}\left(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}\right). \tag{2.3}$$

The six components with $i \neq j$ are the *shear* stresses.

As a consequence of equilibrium, no net force can act on the element, so that

$$\frac{\partial \sigma_{xi}}{\partial x} + \frac{\partial \sigma_{yi}}{\partial y} + \frac{\partial \sigma_{zi}}{\partial z} = 0, \quad i = x, y, z, \tag{2.4}$$

equations (2.4) are the so-called *equilibrium equations* of classic elasticity.

When acted upon by stresses, the body deforms. The displacement **u** has the components u_x, u_y, u_z representing its projections on the x, y, z axes, as shown in Figure 2.3.

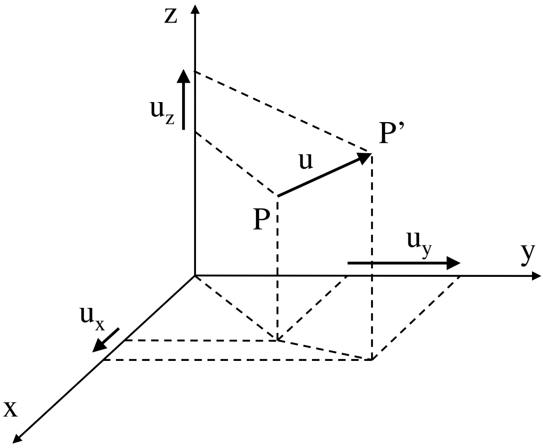


Figure 2.3: Displacement of P to P' by the displacement vector **u**.

Let us consider now the response to stress. In linear elasticity, the strains are defined in terms of the first derivatives of the displacement components

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial j} + \frac{\partial u_j}{\partial i}\right). \tag{2.5}$$

The nine components of strain are therefore

$$\begin{aligned}
\varepsilon_{xx} &= \frac{\partial u_x}{\partial x}, \qquad \varepsilon_{yz} = \varepsilon_{zy} = \frac{1}{2}\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right), \\
\varepsilon_{yy} &= \frac{\partial u_y}{\partial y}, \qquad \varepsilon_{zx} = \varepsilon_{xz} = \frac{1}{2}\left(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z}\right), \\
\varepsilon_{zz} &= \frac{\partial u_z}{\partial z}, \qquad \varepsilon_{xy} = \varepsilon_{yx} = \frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right). \qquad (2.6)
\end{aligned}$$

Partial differentials are used because in general each displacement component is a function of position (x, y, z).

The strains ε_{xx}, ε_{yy}, ε_{zz} defined in (2.6) are the *normal* strains. They represent the fractional change in length of elements parallel to the x, y and z axes, respectively, *e.g.* the length l_x of an element in the x direction is changed to $l_x(1 + \varepsilon_{xx})$.

The volume V of a small volume element is changed by strain to $(V + \Delta V) = V(1 + \varepsilon_{xx})(1 + \varepsilon_{yy})(1 + \varepsilon_{zz})$. In the linear approximation, the fractional change in volume Δ, known as the dilatation, is therefore

$$\Delta = \Delta V/V = (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}). \qquad (2.7)$$

Δ is independent of the orientation of the axes x, y, z.

Each component for $i \neq j$ is half of the shear strain η_{ij}, which is usually defined in engineering as

$$\eta_{ij} = 2\varepsilon_{ij}, \qquad i \neq j. \qquad (2.8)$$

The six components ε_{yz}, ε_{zy}, ε_{xz}, ε_{zx}, ε_{xy}, and ε_{yx} have a simple physical meaning. This is demonstrated by ε_{xy} in Figure 2.4, in which a small area element $ABCD$ in the xy plane has been strained to the shape $AB'C'D'$ without change of area. The angle between the sides AB and AD that was initially parallel to x and y, respectively, has decreased by $2\varepsilon_{xy}$. By rotating, but not deforming the element as in Figure 2.4, it is seen that the element has undergone a simple shear.

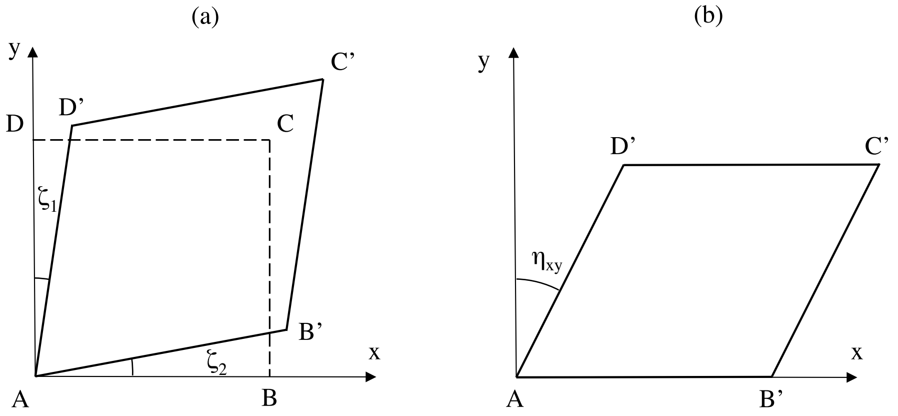


Figure 2.4: (a) Pure shear and (b) simple shear of an area element in the xy plane.

The shear strain η_{xy} is related to the angles of shear by the relation

$$\eta_{xy} = \zeta_x + \zeta_y = \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}. \tag{2.9}$$

The above expressions contain only first-order terms in derivatives of stress and displacement and are exact only in the limit that the stress and displacement approach zero, *i.e*, the magnitude of these components is $\ll$1. This procedure results in the linear theory of elasticity, which is a good approximation for small displacements. The inclusion of second-order terms makes the theory much more complicated, thus such second-order corrections are not considered here.

For small distortions ($\partial u_i/\partial x_j \ll 1$), each stress component is linearly proportional to each strain. The relationship between stresses and strains in linear approximation is called Hooke's law. Thus, the stresses depend only on the strains

$$\sigma_{ij} = c_{ijkl}\,\varepsilon_{kl}. \tag{2.10}$$

The coefficients c_{ijkl} are the elastic constants, also called stiffness constants. From (2.2) and the relation $\varepsilon_{kl} = \varepsilon_{lk}$, it follows directly that

$$c_{ijkl} = c_{jikl} = c_{jilk} = c_{jilk}. \tag{2.11}$$

The combination of the last two relations yields the expression

$$\sigma_{ij} = c_{ijkl}\,\frac{\partial u_k}{\partial x_l}. \tag{2.12}$$

The further combination of the last equation with (2.4) gives

$$c_{ijkl}\,\frac{\partial^2 u_k}{\partial x_j \partial x_l} = 0. \tag{2.13}$$

When a unit volume element deforms reversibly by differential strain increments $\mathrm{d}\varepsilon_{ij}$, the stresses work on the element by an amount

$$\mathrm{d}w = \sigma_{ij}\,\mathrm{d}\varepsilon_{ij} = c_{ijkl}\,\varepsilon_{ij}\,\mathrm{d}\,\varepsilon_{ij}. \tag{2.14}$$

If the deformation is reversible and isothermal, and if the work is restricted to that of elastic deformation, the differential work $\mathrm{d}w$ is equal to the differential change in Helmholtz free energy $\mathrm{d}F$ of the element

$$\mathrm{d}F = \mathrm{d}w = c_{ijkl}\,\varepsilon_{kl}\,\mathrm{d}\,\varepsilon_{ij}. \tag{2.15}$$

Thus

$$\frac{\partial^2 F}{\partial \varepsilon_{ij}\,\partial \varepsilon_{kl}} = c_{ijkl}. \tag{2.16}$$

The free energy is a state function and $\mathrm{d}F$ is an exact differential, so that the order of differentiation in the previous equation is immaterial. As a consequence it follows

$$c_{ijkl} = c_{klij}. \tag{2.17}$$

The strain-energy-density function (the strain energy per unit of volume) is determined by the integration of (2.15)

$$w = \frac{1}{2}\, c_{ijkl}\, \varepsilon_{ij}\, \varepsilon_{kl}. \tag{2.18}$$

A typical solution to an elasticity problem involves the determination of the displacements u_k from (2.13), given a set of boundary conditions. Stresses and strains then follow from (2.5) and (2.12), and the strain energy follows from (2.18).

2.2.1 Matrix notation

Hooke's law (2.10) written as an equation between matrices becomes

$$\{\sigma_{ij}\} = \{c_{ijkl}\}\{\varepsilon_{kl}\}. \tag{2.19}$$

The matrix c_{ijkl} is relating the nine elements σ_{ij} to the nine elements ε_{kl}. According to (2.17), the matrix $\{c_{ijkl}\}$ is symmetric. The elastic coefficients are often written in a contracted matrix notation as c_{mn} where m and n are each indices corresponding to a pair of indices ij or kl, according to the following reduction:

$$\begin{array}{ccccccccccccc} ij & \text{or} & kl & 11 & 22 & 33 & 23 & 31 & 12 & 32 & 13 & 21, \\ m & \text{or} & kl & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9. \end{array} \tag{2.20}$$

Thus, by definition,

$$\begin{array}{ll} c_{11} = c_{1111} & c_{12} = c_{1122} \\ c_{44} = c_{2323} & c_{46} = c_{2312} \\ \ldots & \ldots \end{array} \tag{2.21}$$

(2.11) and (2.17) indicate that due to symmetry there are only 21 independent elastic constants among the 81 in c_{mn}. Thus (2.19) reduces to

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \\ \sigma_{zy} \\ \sigma_{xz} \\ \sigma_{yx} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} & c_{46} & c_{56} & c_{66} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} & c_{45} & c_{55} & c_{56} \end{pmatrix} \begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \\ \varepsilon_{zy} \\ \varepsilon_{xz} \\ \varepsilon_{yx} \end{pmatrix}. \tag{2.22}$$

Because of the symmetry the previous relation is reduced to the following 6×6 representation

$$\begin{pmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \sigma_{yz}\\ \sigma_{zx}\\ \sigma_{xy}\end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16}\\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26}\\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36}\\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46}\\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56}\\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66}\end{pmatrix} \begin{pmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{xz}\\ 2\epsilon_{xy}\end{pmatrix}. \tag{2.23}$$

The c_{mn} in the 6×6 matrix are the same as the c_{mn} occurring in the 9×9 matrix. The 6×6 matrix is also symmetric. When the axes of reference are rotated, stresses, strains, and elastic constants must be transformed accordingly. Transformations are performed most conveniently in the complete 9×9 scheme. This procedure is described in Section 2.3. For most crystals, the number of independent elastic constants is reduced further from 21 because of crystal symmetry. As an example, only five constants are independent for hexagonal crystals:

$$\begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & c_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & c_{55} & 0\\ 0 & 0 & 0 & 0 & 0 & c_{66}\end{pmatrix}. \tag{2.24}$$

For cubic crystals, the independent constants are three:

$$\begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & c_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & c_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & c_{44}\end{pmatrix}. \tag{2.25}$$

For isotropic solids only two material properties – λ and μ, called Lamé constants – are required:

$$\begin{aligned} \sigma_{xx} &= 2\mu\varepsilon_{xx} + \lambda\left(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}\right)\,,\\ \sigma_{yy} &= 2\mu\varepsilon_{yy} + \lambda\left(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}\right)\,,\\ \sigma_{zz} &= 2\mu\varepsilon_{zz} + \lambda\left(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}\right)\,,\\ \sigma_{xy} &= 2\mu\varepsilon_{xy}\,,\quad \sigma_{yz} = 2\mu\varepsilon_{yz}\,,\quad \sigma_{zx} = 2\mu\varepsilon_{zx}\,. \end{aligned} \tag{2.26}$$

The *shear* modulus μ is calculated as [57]:

$$\mu = \frac{1}{2(s_{11}-s_{12})}. \tag{2.27}$$

Other elastic constants are frequently used, the most useful being *Young's* modulus Y, *Poisson's* ratio ν, and the *bulk* modulus H. Under uniaxial, normal loading in the longitudinal direction, Y is the ratio of longitudinal stress to longitudinal strain [57], and ν is minus the ratio of lateral strain to longitudinal strain [57], H is a function of the fractional change in volume Δ and the effective pressure p (see equations (2.3) and (2.7)):

$$\begin{aligned} Y &= 1/s_{11}, \\ \nu &= -s_{11}/s_{12}, \\ H &= 2p/\Delta. \end{aligned} \tag{2.28}$$

Since only two material parameters are required in Hooke's law, these constants are interrelated [57]. For example,

$$Y = 2\mu(1+\nu)\nu = \lambda/2(\lambda+\mu)H = \frac{Y}{1-2\nu}. \tag{2.29}$$

2.3 Transformations

The material constants c_{ij} or s_{ij} for a particular material are usually specified in a basis with coordinate axes aligned with particular symmetry planes (if any) in the material. When solving problems involving anisotropic materials it is frequently necessary to transform these values to a coordinate system that is oriented in some convenient way relative to the boundaries of the solid. The basis transformation formulas are listed below and are used for the calculation of the dislocation energy in Chapter 3. To this end, let's suppose that the components of the stiffness tensor are given in a basis, $\{\mathbf{e1}, \mathbf{e2}, \mathbf{e3}\}$, and we wish to determine its components in a second basis, $\{\mathbf{m1}, \mathbf{m2}, \mathbf{m3}\}$. We define the transformation tensor Ω with components $\Omega_{ij} = m_j \cdot e_j$, or in matrix form

$$\Omega = \begin{pmatrix} \mathbf{m}_1 \cdot \mathbf{e}_1 & \mathbf{m}_1 \cdot \mathbf{e}_2 & \mathbf{m}_1 \cdot \mathbf{e}_3 \\ \mathbf{m}_2 \cdot \mathbf{e}_1 & \mathbf{m}_2 \cdot \mathbf{e}_2 & \mathbf{m}_2 \cdot \mathbf{e}_3 \\ \mathbf{m}_3 \cdot \mathbf{e}_1 & \mathbf{m}_3 \cdot \mathbf{e}_2 & \mathbf{m}_3 \cdot \mathbf{e}_3 \end{pmatrix}. \tag{2.30}$$

This is a symmetric tensor satisfying $\Omega\Omega^{\mathrm{T}} = \Omega^{\mathrm{T}}\Omega = \mathrm{I}$. In practice, the transformation tensor can be computed in terms of the angles between the basis vectors. It is straightforward to show that stress, strain and elasticity tensors transform as

$$\begin{aligned} \sigma_{ij}^{(\mathrm{m})} &= \Omega_{ik}\, \sigma_{kl}^{(\mathrm{e})}\, \Omega_{jl}, \\ \mathrm{s}_{ij}^{(\mathrm{m})} &= \Omega_{ik}\, \sigma_{kl}^{(\mathrm{e})}\, \sigma_{jl}, \\ c_{ijkl}^{(\mathrm{m})} &= \Omega_{ip}\, \Omega_{jq}\, c_{pqrs}^{(\mathrm{e})}\, \Omega_{kr}\, \Omega_{ls}. \end{aligned} \tag{2.31}$$

The basis transformation formula for the stiffness tensor C is more conveniently expressed in matrix form as

$$\mathrm{C}^{(\mathrm{m})} = \mathrm{K}\, \mathrm{C}^{(\mathrm{e})}\, \mathrm{K}^{\mathrm{T}}, \tag{2.32}$$

where the rotation matrix K is computed as

$$\mathrm{K} = \begin{pmatrix} \mathrm{K}^1 & 2\mathrm{K}^2 \\ \mathrm{K}^3 & \mathrm{K}^4 \end{pmatrix}, \tag{2.33}$$

and

$$\begin{aligned} K^1_{ij} =& \Omega^2_{ij}, \\ K^2_{ij} =& \Omega_{i\,\mathrm{mod}(j+1,3)}\Omega_{i\,\mathrm{mod}(j+2,3)}, \\ K^3_{ij} =& \Omega_{\mathrm{mod}(i+1,3)\,j}\Omega_{\mathrm{mod}(i+2,3)\,j}, \\ K^4_{ij} =& \Omega_{\mathrm{mod}(i+1,3)\,\mathrm{mod}(j+1,3)}\Omega_{\mathrm{mod}(j+2,3)\,\mathrm{mod}(j+2,3)} + \\ & \Omega_{\mathrm{mod}(i+1,3)\,\mathrm{mod}(j+2,3)}\Omega_{\mathrm{mod}(i+2,3)\,\mathrm{mod}(j+1,3)}, \\ \text{with } & i,j = 1,2,3. \end{aligned} \tag{2.34}$$

The modulo function satisfies

$$\mathrm{mod}(i,3) = \begin{cases} i & i \leq 3 \\ i-3 & i > 3 \end{cases} . \tag{2.35}$$

The basis change for the compliance tensor (the inverse of the stiffness tensor) follows as

$$\mathrm{S}^{(\mathrm{m})} = \mathrm{K}^{-\mathrm{T}}\,\mathrm{S}^{(\mathrm{e})}\,\mathrm{K}^{-1}, \tag{2.36}$$

where

$$\mathrm{K}^{-\mathrm{T}} = \begin{pmatrix} \mathrm{K}^1 & \mathrm{K}^2 \\ 2\mathrm{K}^3 & \mathrm{K}^4 \end{pmatrix}. \tag{2.37}$$

The proof of these expressions is given in Ting [74]. For the particular case of rotation around an angle θ in a counterclockwise sense about the x, y, z axes, the rotation matrix reduces, respectively, to

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & c^2 & s^2 & 2cs & 0 & 0 \\ 0 & s^2 & c^2 & -2cs & 0 & 0 \\ 0 & -cs & cs & c^2-s^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & c & -s \\ 0 & 0 & 0 & 0 & -s & c \end{pmatrix} \begin{pmatrix} c^2 & 0 & s^2 & 0 & 2cs & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ s^2 & 0 & c^2 & 0 & -2cs & 0 \\ 0 & 0 & 0 & c & 0 & -s \\ -cs & 0 & cs & 0 & c^2-s^2 & 0 \\ 0 & 0 & 0 & s & 0 & c \end{pmatrix} \begin{pmatrix} c^2 & s^2 & 0 & 0 & 0 & 2cs \\ s^2 & c^2 & 0 & 0 & 0 & -2cs \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & c & s & 0 \\ 0 & 0 & 0 & -s & c & 0 \\ -cs & cs & 0 & 0 & 0 & c^2-s^2 \end{pmatrix}, \tag{2.38}$$

where $c = \cos\theta$ and $s = \sin\theta$. The inverse matrix K^{-1} can be obtained simply by changing the sign of the angle θ in each rotation matrix. Clearly, applying the three rotations successively can produce an arbitrary orientation change.

3 Dislocation energy

3.1 Introduction

The linear elasticity theory reviewed in Chapter 2 is applied here for the calculation of the dislocation energy within an isotropic (Section 3.3) and anisotropic (Section 3.4) framework. Both treatments are applied to calculate the pre-logarithmic coefficients (see Sections 3.5 and 3.6), both of which are important parameters used to evaluate the geometrical configurations of dislocations inside AlN, GaN and InN (Section 3.7). The importance of the dislocation configuration is made evident in the modeling of the dislocation density, as described in Chapter 5.

3.2 Dislocation energy

Dislocations are line imperfections in an otherwise perfect crystal. The atoms around a dislocation are displaced from their perfect lattice sites and the resulting distortion produces stress and strain fields around the dislocation line. Therefore, the dislocation is a source of internal stress in the crystal. The Burgers vector $\mathbf{b}$ represents the magnitude and the direction of the distortion in the crystal lattice. The stresses and the strains in the bulk of the crystal are sufficiently small to use the theory of linear elasticity to calculate them. This theory ceases to be valid around the dislocation line, where the stresses and the strains are infinite, and here non-linear theory should be applied. Consequently, it is necessary to exclude from the description a cylinder whose axis coincides with the dislocation line. The cylinder where linear elasticity breaks down is called *core* of the dislocation. The core is the plastic region of the system and the region around is the elastic one. The radius r_c of the dislocation core is called inner cut-off radius and its order of magnitude is $\mathbf{b}$. In this work, r_c is taken to be $b/2$ [19].

It is also necessary to limit externally the continuum under analysis with a second cylindrical surface having the dislocation line as the axis. The radius R of this second cylinder is called outer cut-off radius and its purpose is to keep expressions for the energy of the dislocation and the displacement field finite.

Linear elasticity is applied in the case of a straight infinite long dislocation to calculate its energy and the related stress and strain fields, which will be used in the next chapters. The subsequent treatment is developed for the geometry shown in Figure. 3.1. The *energy per unit length of the dislocation* $\mathrm{d}\mathcal{E}/\mathrm{d}y$ (for simplicity often called just the dislocation energy) may be divided into two parts:

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}y} = \frac{\mathrm{d}\mathcal{E}_{\mathrm{core}}}{\mathrm{d}y} + \frac{\mathrm{d}\mathcal{E}_{\mathrm{d}}}{\mathrm{d}y}, \tag{3.1}$$

where $\mathrm{d}\mathcal{E}_{\mathrm{core}}/\mathrm{d}y$ and $\mathrm{d}\mathcal{E}_{\mathrm{d}}/\mathrm{d}y$ account for the energy inside and outside the dislocation core region, respectively. For large film thickness h (see Figure. 3.1), the core energy constitutes only a minor contribution to the total dislocation energy and is thus neglected in the following calculation. Let $\boldsymbol{\sigma}^d$ be the stress tensor associated with the strain field $\boldsymbol{\varepsilon}^d$ caused by the straight dislocation. Then the strain energy of the dislocation is given by

$$\mathrm{d}\mathcal{E}_{\mathrm{d}} = \tfrac{1}{2}\int_V \sum_{ij} \sigma^d_{ij}\epsilon^d_{ij}\,\mathrm{d}V = \tfrac{1}{2}\sum_{ij}\int_V \tfrac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\sigma^d_{ij}\,\mathrm{d}V \tag{3.2}$$

and using the symmetry of the stress tensor $\sigma^d_{ij} = \sigma^d_{ji}$ we get

$$\tfrac{1}{2}\sum_{ij}\int_V \frac{\partial u_i}{\partial x_j}\sigma^d_{ij}\,\mathrm{d}V = \tfrac{1}{2}\sum_{ij}\int_V \frac{\partial}{\partial x_j}\left(u_i\sigma^d_{ij}\right)\mathrm{d}V - \tfrac{1}{2}\sum_{ij}\int_V u_i\frac{\partial \sigma^d_{ij}}{\partial x_j}\,\mathrm{d}V. \tag{3.3}$$

Recalling the equilibrium conditions for elastic media,

$$\sum_j \frac{\partial \sigma^d_{ij}}{\partial x_j} = 0, \tag{3.4}$$

one obtains

$$\mathrm{d}\mathcal{E}_{\mathrm{d}} = \tfrac{1}{2}\sum_{ij}\int_V \frac{\partial}{\partial x_j}\left(u_i\sigma^d_{ij}\right)\mathrm{d}V. \tag{3.5}$$

Using the Gauss-Ostrogradsky theorem to transform the volume integral into a surface integral over the surface S enclosing the volume V, we get:

$$\tfrac{1}{2}\sum_{ij}\int_V \frac{\partial}{\partial x_j}\left(u_i\sigma^d_{ij}\right)\mathrm{d}V = \tfrac{1}{2}\sum_{ij}\int_S u_i\sigma^d_{ij}n_j\,\mathrm{d}S. \tag{3.6}$$

The surface S can be divided into four parts (see Figure. 3.1) and the last term can be rewritten as

$$\tfrac{1}{2}\sum_{k=1}^{4}\left(\sum_{ij}\int_{S_k} u_i\sigma^d_{ij}n_j\,\mathrm{d}S_k\right). \tag{3.7}$$

Since S_1 is supposed to be infinitely far from the dislocation, the displacements vanish and the contribution of the surface is zero. The surface S_3 is a cylindrical surface whose axis is aligned along the y-axis. Since the dislocation is infinite along the y direction, $\boldsymbol{n} = (n_x, 0, n_z)$:

$$\tfrac{1}{2}\sum_{ij}\int_{S_3} u_i\sigma^d_{ij}n_j\,\mathrm{d}S_3 = \tfrac{1}{2}\sum_{i}\int_{S_3}\left(u_i\sigma^d_{ix}n_x + u_i\sigma^d_{iz}n_z\right)\mathrm{d}S_3. \tag{3.8}$$

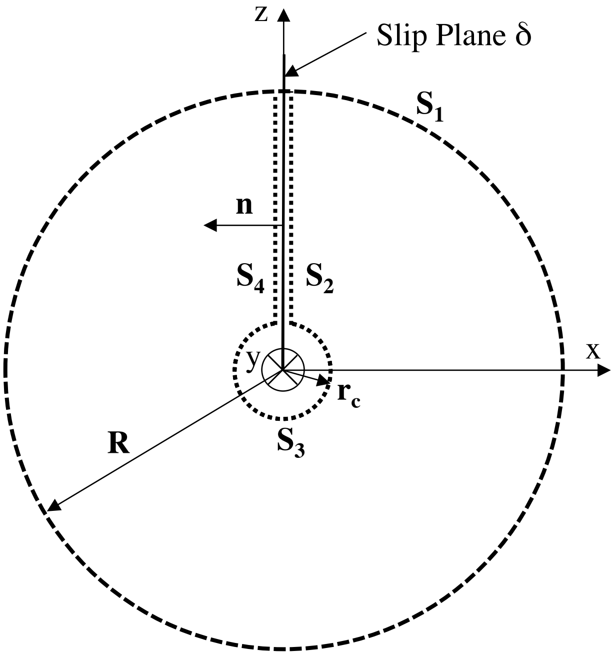


Figure 3.1: The dislocation is formed by an offset (defined by the Burgers vector **b**) of one side S_4 of the slip plane δ with respect to the other side S_2. The surface S_3 encloses the dislocation core region. h denotes the film thickness.

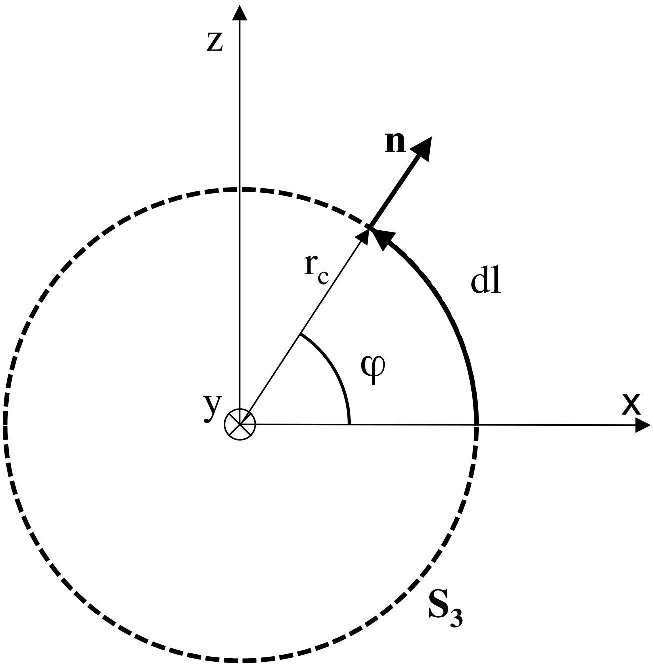


Figure 3.2: The core surface S_3 in a cylindrical coordinate system.

To simplify the integration, it is convenient to transform the Cartesian coordinates into cylindrical polar coordinates (see Figure 3.2) as

$$\begin{aligned} x &= r_\mathrm{c}\cos\varphi, \\ y &= t, \\ z &= r_\mathrm{c}\sin\varphi. \end{aligned} \tag{3.9}$$

Consequently,

$$\mathrm{d}S_3 = \mathrm{d}l\mathrm{d}y = r_\mathrm{c}\mathrm{d}\varphi\mathrm{d}y. \tag{3.10}$$

Considering that $\boldsymbol{n} = (\cos\varphi, 0, \sin\varphi)$, the dislocation energy expressed in (3.8) becomes

$$\frac{1}{2}\int_0^{2\pi}\sum_i u_i\left(\sigma^d_{ix}\cos\varphi + \sigma^d_{iz}\sin\varphi\right) r_\mathrm{c}\,\mathrm{d}\varphi, \qquad i = x,\ y,\ z. \tag{3.11}$$

Expanding the last equation yields

$$\begin{aligned} \tfrac{1}{2}r_\mathrm{c}\int_0^{2\pi}\Bigg(& \underbrace{u_x\left(\sigma^d_{xx}\cos\varphi + \sigma^d_{xz}\sin\varphi\right)}_{\text{edge}} + \underbrace{u_y\left(\sigma^d_{yx}\cos\varphi + \sigma^d_{yz}\sin\varphi\right)}_{\text{screw}} + \\ & \underbrace{u_z\left(\sigma^d_{zx}\cos\varphi + \sigma^d_{zz}\sin\varphi\right)}_{\text{edge}}\Bigg)\,\mathrm{d}\varphi. \end{aligned} \tag{3.12}$$

The substitution of the displacements and the stress components expressed in cylindrical polar coordinates allow for the evaluation of the impact of the core surface on the dislocation energy.

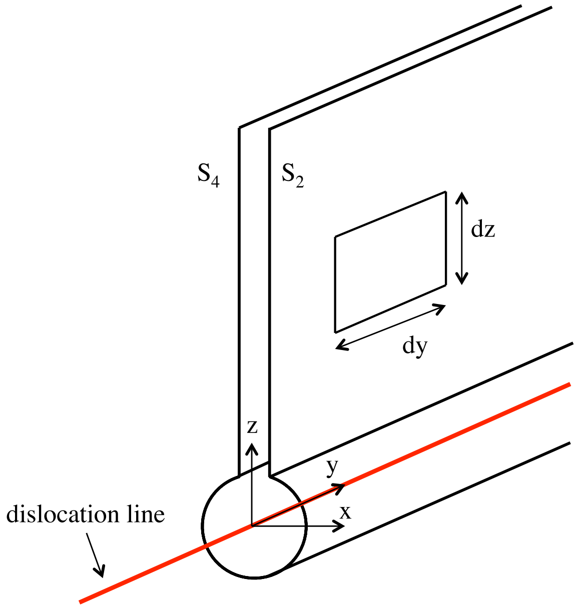


Figure 3.3: The dislocation lies along the y-axis.

When the impact of the core integral along S_3 is negligible, only the integrals along the surfaces S_2 and S_4 (see Figure. 3.1) have to be considered:

$$\mathrm{d}\mathcal{E}_\mathrm{d} = \tfrac{1}{2}\sum_{ij}\left(\int_{S_2} u_i\sigma^d_{ij}n_j\,\mathrm{d}S_2 + \int_{S_4} u_i\sigma^d_{ij}n_j\,\mathrm{d}S_4\right). \tag{3.13}$$

Considering Figure 4.2, $\mathrm{d}S_2 = \mathrm{d}S_4 = \mathrm{d}z\mathrm{d}y$ the energy per unit length of dislocation is

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \tfrac{1}{2}\sum_{ij}\left(\int_{r_c}^{R} u_i(x\to 0^+)\sigma^d_{ij}\,\mathrm{d}z + \int_{R}^{r_c} u_i(x\to 0^-)\sigma^d_{ij}\,\mathrm{d}z\right). \tag{3.14}$$

If the extremes of the second integral are inverted, one obtains

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \tfrac{1}{2}\sum_{i}\int_{r_c}^{R} b_i\sigma^d_{ij}\,\mathrm{d}z, \tag{3.15}$$

where $u_i(x\to 0^+) - u_i(x\to 0^-) = b_i$.

3.3 Dislocation in an isotropic continuum

When the impact of the core integral along S_3 is negligible, only the integrals along the surfaces S_2 and S_4 have to be considered. The evaluation of these two integrals for a straight dislocation inside an infinite isotropic medium yields the classic formula for the dislocation energy [19, 23, 32]:

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \frac{\mu b^2\left(1-\nu\cos^2\theta\right)}{4\pi\left(1-\nu\right)}\ln\left(\frac{R}{r_\mathrm{c}}\right), \tag{3.16}$$

where μ and ν are the shear modulus and the Poisson ratio, respectively, θ is the angle between the Burgers vector $\mathbf{b}$ and the dislocation line, and R is the outer cut-off radius.

3.4 Dislocation in an anisotropic continuum

Steeds [72] developed a treatment to derive the dislocation energy within the anisotropic elasticity. According to this treatment, the dislocation is considered straight and extended to infinity along the z-axis. This assumption simplifies the problem to a plane strain problem where no quantity depends on the z-coordinate, so

$$\frac{\partial}{\partial z} = 0. \tag{3.17}$$

The Burgers vector of the dislocation is $\mathbf{b}$. Displacements are given by the functions u_x, u_y and u_z. The displacements u_x and u_y correspond to the edge component of the considered dislocation whereas u_z corresponds to the screw component. The strain components are

$$\begin{aligned}
\varepsilon_{xx} &= \frac{\partial u_x}{\partial x}, & \varepsilon_{xz} &= \frac{1}{2}\frac{\partial u_z}{\partial x}, \\
\varepsilon_{zz} &= 0, & \varepsilon_{xy} &= \frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right), \\
\varepsilon_{yy} &= \frac{\partial u_y}{\partial y}, & \varepsilon_{yz} &= \frac{1}{2}\frac{\partial u_z}{\partial y}.
\end{aligned} \tag{3.18}$$

The previous equations relate the six strain components to the three components of the lattice displacement and this implies that relations exist between the various ε_{ij}. These relations are known as the compatibility equations [45]:

$$\frac{\partial^2 \varepsilon_{xx}}{\partial y^2} + \frac{\partial^2 \varepsilon_{yy}}{\partial x^2} = 2\frac{\partial^2 \varepsilon_{xy}}{\partial x \partial y}, \tag{3.19a}$$

$$\frac{\partial \varepsilon_{yz}}{\partial x} - \frac{\partial \varepsilon_{xz}}{\partial y} = 0. \tag{3.19b}$$

Hooke's law for an anisotropic material is written as:

$$\begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{xz} \\ 2\varepsilon_{xy} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & s_{15} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{pmatrix} \begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix}. \tag{3.20}$$

The infinite straight dislocation lies along the z-axis and therefore $\varepsilon_{zz} = 0$. This last relation yields an additional condition:

$$\begin{aligned}
&0 = \varepsilon_{zz} = s_{21}\sigma_{xx} + s_{22}\sigma_{zz} + s_{23}\sigma_{yy} + s_{24}\sigma_{yz} + s_{25}\sigma_{xy} + s_{26}\sigma_{xz}, \\
&\sigma_{zz} = \left(-s_{21}\sigma_{xx} - s_{23}\sigma_{yy} - s_{24}\sigma_{yz} - s_{25}\sigma xy - s_{26}\sigma_{xz}\right)/s_{22}.
\end{aligned} \tag{3.21}$$

Steeds [72] expressed the stresses as a function of the Airy functions, F and ϕ

$$\begin{aligned}
\sigma_{xx} &= \frac{\partial^2 F}{\partial y^2}, & \sigma_{xz} &= -\frac{\partial \phi}{\partial y}, \\
\sigma_{xy} &= -\frac{\partial^2 F}{\partial x \partial y}, & \sigma_{yz} &= -\frac{\partial \phi}{\partial x}, \\
\sigma_{yy} &= \frac{\partial^2 F}{\partial x^2}.
\end{aligned} \tag{3.22}$$

Substituting the expressions (3.21) and (3.22) into the compatibility equations (3.19) yields

$$\begin{aligned}
&(s_{11}+s_{22})\,\frac{\partial^4 F}{\partial x^2\partial y^2}+s_{12}\frac{\partial^4 F}{\partial y^4}+s_{21}\frac{\partial^4 F}{\partial x^4}+\\
&\frac{s_{63}}{s_{33}}\left(s_{31}\frac{\partial^4 F}{\partial x^3\partial y}+s_{32}\frac{\partial^4 F}{\partial x\partial y^3}-s_{36}\frac{\partial^4 F}{\partial x^2\partial y^2}-s_{34}\frac{\partial^4 \phi}{\partial x^2\partial y^1}+s_{35}\frac{\partial^4 \phi}{\partial x\partial y^3}+\right)+\\
&s_{66}\frac{\partial^4 F}{\partial x^2\partial y^2}+s_{15}\frac{\partial^4 \phi}{\partial y^4}+s_{25}\frac{\partial^4 \phi}{\partial x^2\partial y^2}+s_{64}\frac{\partial^4 \phi}{\partial x^2\partial y^1}=\\
&\frac{s_{13}}{s_{33}}\left(s_{31}\frac{\partial^4 F}{\partial x^2\partial y^2}+s_{32}\frac{\partial^4 F}{\partial y^4}-s_{36}\frac{\partial^4 F}{\partial x\partial y^3}-s_{34}\frac{\partial^3 \phi}{\partial x\partial y^2}+s_{35}\frac{\partial^4 \phi}{\partial y^4}\right)\\
&\frac{s_{23}}{s_{33}}\left(s_{31}\frac{\partial^4 F}{\partial x^4}+s_{32}\frac{\partial^4 F}{\partial x^2\partial y^2}-s_{36}\frac{\partial^4 F}{\partial x^3\partial y}-s_{34}\frac{\partial^3 \phi}{\partial x^3}+s_{35}\frac{\partial^4 \phi}{\partial x^2\partial y^2}\right)+\\
&s_{16}\frac{\partial^4 F}{\partial x\partial y^3}+s_{26}\frac{\partial^4 F}{\partial x^3\partial y}+s_{61}\frac{\partial^4 F}{\partial x^3\partial y}+s_{62}\frac{\partial^4 F}{\partial x\partial y}+\\
&s_{14}\frac{\partial^3 \phi}{\partial x\partial y^2}+s_{24}\frac{\partial^3 \phi}{\partial x^3}+s_{65}\frac{\partial^4 \phi}{\partial x\partial y^3},
\end{aligned}\tag{3.23a}$$

$$\begin{aligned}
&s_{11}\frac{\partial^4 F}{\partial x^2\partial y^2}+s_{12}\frac{\partial^4 F}{\partial y^4}+s_{21}\frac{\partial^4 F}{\partial x^4}+s_{22}\frac{\partial^2 F}{\partial y^2}+\\
&\frac{s_{63}}{s_{33}}\left(s_{31}\frac{\partial^4 F}{\partial x^3\partial y}+s_{32}\frac{\partial^4 F}{\partial x^1\partial y^3}-s_{36}\frac{\partial^4 F}{\partial x^2\partial y^2}-s_{34}\frac{\partial^3 \phi}{\partial x^2\partial y^1}+s_{35}\frac{\partial^4 \phi}{\partial x\partial y^3}\right)+\\
&s_{66}\frac{\partial^4 \phi}{\partial x^2\partial y^2}+s_{15}\frac{\partial^4 \phi}{\partial y^4}+s_{25}\frac{\partial^4 \phi}{\partial x^2\partial y^2}+s_{64}\frac{\partial^3 \phi}{\partial x^2\partial y}=\\
&\frac{s_{13}}{s_{33}}\left(s_{31}\frac{\partial^4 F}{\partial x^2\partial y^2}+s_{32}\frac{\partial^4 F}{\partial y^4}-s_{36}\frac{\partial^3 F}{\partial x\partial y^3}-s_{34}\frac{\partial^3 \phi}{\partial x\partial y^2}+s_{35}\frac{\partial^4 \phi}{\partial y^4}\right)+\\
&s_{16}\frac{\partial^4 F}{\partial x\partial y^3}+\frac{s_{23}}{s_{33}}\left(s_{31}\frac{\partial^4 F}{\partial x^4}+s_{32}\frac{\partial^4 F}{\partial x^2\partial y^2}-s_{36}\frac{\partial^4 F}{\partial x^3\partial y}-s_{34}\frac{\partial^3 \phi}{\partial x^3}+s_{35}\frac{\partial^4 \phi}{\partial x^2\partial y^2}+\right)+\\
&s_{26}\frac{\partial^4 F}{\partial x^3\partial y}+s_{61}\frac{\partial^4 F}{\partial x^3\partial y}+s_{62}\frac{\partial^4 F}{\partial x\partial y^3}+s_{14}\frac{\partial^3 \phi}{\partial x\partial y^2}+s_{24}\frac{\partial^3 \phi}{\partial x^3}+s_{65}\frac{\partial^3 \phi}{\partial x\partial y^3}.
\end{aligned}\tag{3.23b}$$

Solutions for the previous equations with cylindrical symmetry have the following form [72]:

$$F=B\,g\,(x+p\,y)\,,\qquad \phi=C\,f\,(x+p\,y)\,,\tag{3.24}$$

where g and f are functions of a linear combination of the coordinates x and y. Substituting (3.24) into (3.23) and eliminating B and C yields a sextic equation for the parameter p:

$$
\begin{aligned}
&p^6\left(S_{11}S_{55}-S_{15}^2\right)-\\
&2p^5\left(S_{11}S_{45}-S_{15}\left(S_{14}+S_{56}\right)+S_{16}S_{55}\right)+\\
&p^4\left(S_{11}S_{44}+S_{55}\left(2S_{12}+S_{66}\right)-\left(S_{14}+S_{56}\right)^2-2S_{15}\left(S_{25}+S_{46}\right)+4S_{16}S_{45}\right)-\\
&2p^3\left(S_{45}\left(2S_{12}+S_{66}\right)-\left(S_{14}+S_{56}\right)\left(S_{25}+S_{46}\right)-S_{15}S_{24}+S_{16}S_{44}+S_{26}S_{55}\right)+\\
&p^2\left(S_{44}\left(2S_{12}+S_{66}\right)-2S_{24}\left(S_{14}+S_{56}\right)+S_{22}S_{55}-\left(S_{25}+S_{46}\right)^2+4S_{26}S_{45}\right)-\\
&2p\left(S_{22}S_{45}-S_{24}\left(S_{25}+S_{46}\right)+S_{26}S_{44}\right)+\\
&S_{22}S_{44}-S_{24}^2=0 \qquad (3.25)
\end{aligned}
$$

where $S_{lm} = s_{lm} - \left(\left(s_{3l}s_{3m}\right)/s_{33}\right)$. Roots of this equation occur as pairs of complex conjugates p_1 and p_1^*, p_2 and p_2^*, and p_3 and p_3^*. Therefore the most general solution for the functions F and ϕ are

$$F = \sum_{n=1}^{3}\left(B_n\, g_n\left(\zeta_n\right) + B_n^*\, g_n\left(\zeta_n^*\right)\right), \qquad (3.26a)$$

$$\phi = \sum_{n=1}^{3}\left(C_n\, f_n\left(\zeta_n\right) + C_n^*\, f_n\left(\zeta_n^*\right)\right), \qquad (3.26b)$$

where $n = 1, 2, 3$, $\zeta_n = x + p_n\, y$, and g_n and f_n denote, respectively, the functions g and f belonging to the root p_n. In order to satisfy the physical expectations that the stress field decays as $1/r$, it is required that

$$\frac{\partial f_n}{\partial \zeta_n} = \frac{\partial^2 g_n}{\partial \zeta_n^2} = \frac{1}{\zeta}. \qquad (3.27)$$

Substituting (3.26) and (3.27) into (3.22) causes the stress components to become

$$
\begin{aligned}
\sigma_{xx} &= \sum_{n=1}^{3}\left(\frac{p_n^2 B_n}{\zeta_n} + \frac{p_n^{*2} B_n^*}{\zeta_n^*}\right),\\
\sigma_{xy} &= -\sum_{n=1}^{3}\left(\frac{p_n B_n}{\zeta_n} + \frac{p_n^* B_n^*}{\zeta_n^*}\right),\\
\sigma_{yy} &= \sum_{n=1}^{3}\left(\frac{B_n}{\zeta_n} + \frac{B_n^*}{\zeta_n^*}\right),\\
\sigma_{xz} &= \sum_{n=1}^{3}\left(\frac{p_n C_n}{\zeta_n} + \frac{p_n^* C_n^*}{\zeta_n^*}\right),\\
\sigma_{yz} &= -\sum_{n=1}^{3}\left(\frac{C_n}{\zeta_n} + \frac{C_n^*}{\zeta_n^*}\right), \qquad (3.28)
\end{aligned}
$$

The displacement components are obtained combining (3.18) and (3.26):

$$\begin{aligned} u_x = \int \varepsilon_{xx}\, dx = \\ \sum_{n=1}^{3} \Big\{ \big[\big(S_{11} p_n^2 - S_{16} p_n + S_{12} \big) B_n + \left(S_{15} p_n - S_{14} \right) C_n \big] \ln \zeta_n \\ \big[\big(S_{11} p_n^{*2} - S_{16} p_n^* + S_{12} \big) B_n^* + \left(S_{15} p_n^* - S_{14} \right) C_n^* \big] \ln \zeta_n^* \Big\} , \end{aligned} \tag{3.29a}$$

$$\begin{aligned} u_y = \int \varepsilon_{yy}\, dz = \\ \sum_{n=1}^{3} \Big\{ \big[\big(S_{12} p_n^2 - S_{26} p_n + S_{22} \big) B_n + \left(S_{25} p_n - S_{24} \right) C_n \big] \ln \zeta_n \\ \big[\big(S_{12} p_n^{*2} - S_{26} p_n^* + S_{22} \big) B_n^* + \left(S_{25} p_n^* - S_{24} \right) C_n^* \big] \ln \zeta_n^* \Big\} , \end{aligned} \tag{3.29b}$$

$$\begin{aligned} u_z = \int \varepsilon_{xz}\, dx = \\ \sum_{n=1}^{3} \Big\{ \big[\big(S_{15} p_n^2 - S_{56} p_n + S_{25} \big) B_n + \left(S_{55} p_n - S_{45} \right) C_n \big] \ln \zeta_n \\ \big(\big(S_{15} p_n^{*2} - S_{56} p_n + S_{25} \big) B_n^* + \left(S_{55} p_n^* - S_{45} \right) C_n^* \big) \ln \zeta_n^* \Big\} . \end{aligned} \tag{3.29c}$$

In order to evaluate the quantities B_n and C_n and their complex conjugates, six relationships are required.

The first set of boundary conditions is derived from the force equilibrium state of the media, i.e., the condition of zero net force on the dislocation [72], expressed as

$$\iint_S \sum_j \sigma_{ij} n_j \,\mathrm{d}S = 0, \tag{3.30}$$

for an arbitrary cylindrical surface S enclosing the dislocation line. n_j denotes components of the outer normal to the integration surface S. For an infinite straight dislocation along the z-axis, the previous equation becomes

$$\iint_S \sum_j \sigma_{ij} n_j \,\mathrm{d}S = 0, \qquad (j = 1, 2)\,. \tag{3.31}$$

It is assumed that the cylinder S has a height $\mathcal{L}$ with a circular base $\mathcal{C}$ on the xy-plane. Based on these assumptions, the previous equation yields for $i = x$

$$0 = \iint_S \left(\sigma_{xx} n_x + \sigma_{xy} n_y \right) \mathrm{d}S = \mathcal{L} \int_C \mathrm{d} \left(\frac{\partial F}{\partial y} \right) = \mathcal{L} \left[\frac{\partial F}{\partial y} \right] , \tag{3.32}$$

and when substituting the expression for F from (3.26) and using (3.27) one gets [72]

$$\left[\sum_{n=1}^{3}\left(B_n p_n \ln\zeta_n + B_n^* p_n^* \ln\zeta_n^*\right)\right]_{\mathcal{C}} = 0\ . \tag{3.33}$$

The start and the end point of the closed integration loop $\mathcal{C}$ differ by the argument $\Delta\theta_n = 2\pi$. Since the logarithm of a complex argument can be written as

$$\ln\zeta_n = \ln r_n + i\theta_n\ , \tag{3.34a}$$

$$\ln\zeta_n^* = \ln r_n^* - i\theta_n\ , \tag{3.34b}$$

equation (3.33) becomes

$$\sum_{n=1}^{3}\left[B_n p_n\left(\ln r_n + i2\pi - \ln r\right) + B_n^* p_n^*\left(\ln r_n^* - i2\pi - \ln r_n^*\right)\right] = 0. \tag{3.35}$$

Simplifying, one gets

$$\sum_{n=1}^{3}\left(B_n p_n - B_n^* p_n^*\right) = 0. \tag{3.36}$$

Analogously, another two conditions are obtained from equation (3.33) for $i = y$

$$\sum_{n=1}^{3}\left(B_n - B_n^*\right) = 0, \tag{3.37}$$

and for $i = z$

$$\sum_{n=1}^{3}\left(C_n - C_n^*\right) = 0. \tag{3.38}$$

The second set of boundary equations is provided by the displacement relations. The integral of the displacement acquisitions along the Burgers circuit encircling the dislocation line must be equal to the Burgers vector **b**:

$$\oint \mathrm{d}\mathbf{u} = \mathbf{b}. \tag{3.39}$$

Substituting equations (3.29) for the displacements components into the last relation yields a set of three equations of the form $b_i/2\pi i = Q_{ni}B_n - Q_{ni}^* B_n^*$. More explicitly the last three equations become

$$b_x = 2\pi i \sum_{n=1}^{3}\left[S_{11}\left(B_n p_n^2 - B_n^* p_n^{*2}\right) + S_{15}\left(C_n p_n - C_n^* p_n^*\right)\right], \tag{3.40a}$$

$$\begin{aligned} b_y = 2\pi i \sum_{n=1}^{3}\Big[&2S_{16}\left(B_n p_n^2 - B_n^* p_n^{*2}\right) - S_{11}\left(B_n p_n^3 - B_n^* p_n^{*3}\right) \\ &+ \left(S_{56}+S_{14}\right)\left(C_n p_n - C_n^* p_n^*\right) - S_{15}\left(C_n p_n^2 - C_n^* p_n^{*2}\right)\Big], \end{aligned} \tag{3.40b}$$

$$b_z = 2\pi i \sum_{n=1}^{3}\left[S_{15}\left(B_n p_n^2 - B_n^* p_n^{*2}\right) + S_{55}\left(C_n p_n - C_n^* p_n^*\right)\right]. \tag{3.40c}$$

The stress field of the dislocation is fully described by equations (3.28) where constants p_n are the roots of equation (3.25) and constants B_n are the solution of the system of linear equations (3.36), (3.37), (3.38) and (3.40).

The calculation of the dislocation energy is performed by substituting the stress components in equation (3.15) along the cut surface described by the coordinate r such that $x = r\sin\varphi$ and $y = r\cos\varphi$. The components of the outer normal $\mathbf{n}$ are $n_x = \cos\varphi$, $n_z = 0$ and $n_y = \sin\varphi$. The inner and outer cut-off radii are r_c and R, respectively. Then one obtains

$$\begin{aligned}\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} &= \frac{1}{2}\int_{r_\mathrm{c}}^{R}\sum_{n=1}^{3}\left[\frac{\left(b_xB_np_n + b_yC_n - b_zB_n\right)p_n\sin\varphi\ + \left(-b_xB_np_n - b_yC_n + b_zB_n\right)\cos\varphi}{cos\varphi + p_n\sin\varphi} + \right.\\ &\quad\left. + \frac{\left(b_xB_n^*p_n^* + b_yC_n^* - b_zB_n^*\right)p_n^*\sin\varphi\ + \left(-b_xB_n^*p_n^* - b_yC_n^* + b_zB_n^*\right)\cos\varphi}{cos\varphi + p_n^*\sin\varphi}\right]\frac{\mathrm{d}r}{r} = \\ &= \frac{1}{2}\sum_{n=1}^{3}\left[B_n\left(-b_xp_n + b_z\right) - C_nb_y + B_n^*\left(-b_xp_n^* + b_z\right) - C_n^*b_y\right]\ln\frac{R}{r_\mathrm{c}} = \\ &= K\ln\frac{R}{r_\mathrm{c}},\end{aligned} \qquad (3.41)$$

where K is the so called pre-logarithmic coefficient of the dislocation energy. K depends on the elastic constants, the Burgers vector $\mathbf{b}$ and the particular direction of the dislocation line with respect to the crystallographic axes.

3.5 Pre-logarithmic coefficients

Here the theory described in the last section is used to calculate the pre-logarithmic coefficient K for dislocations in AlN, GaN and InN bulk. The procedure has been introduced by Holec [24] to evaluate the dislocation configuration in GaN bulk.

Three different types of dislocations are present in hexagonal structures: c-type dislocation with Burgers vector $\mathbf{b} = [0001]$, a-type with Burgers vector $\mathbf{b} = 1/3\,[\bar{2}110]$, and $(a+c)$-type with Burgers vector $\mathbf{b} = 1/3\,[\bar{2}113]$. The coefficients K of these three types of dislocations are calculated for different directions of the dislocation line using the Steeds treatment [72]. This is done by rotating the elastic tensor according to the direction line. The calculations are performed numerically using the software Wolfram Mathematica. Hexagonal symmetry requires two angles, α and β, for the description of the direction of the dislocation line (see Figure. 3.4). As a consequence, the pre-logarithmic coefficient for one particular direction of the dislocation line is a function of the two angles α and β. The coefficients are shown in Figure 3.5 for AlN, GaN, and InN.

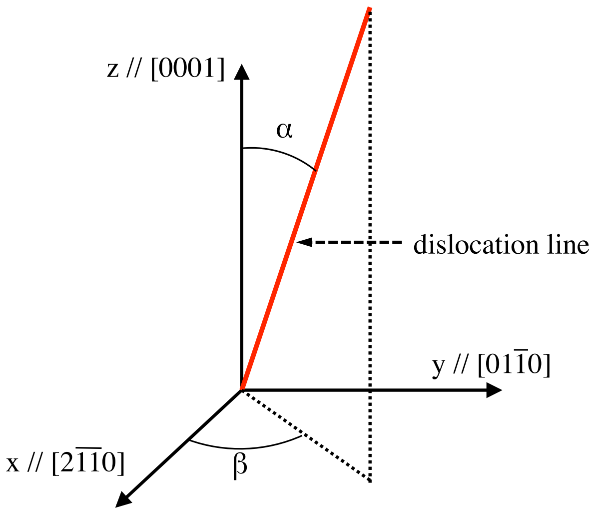


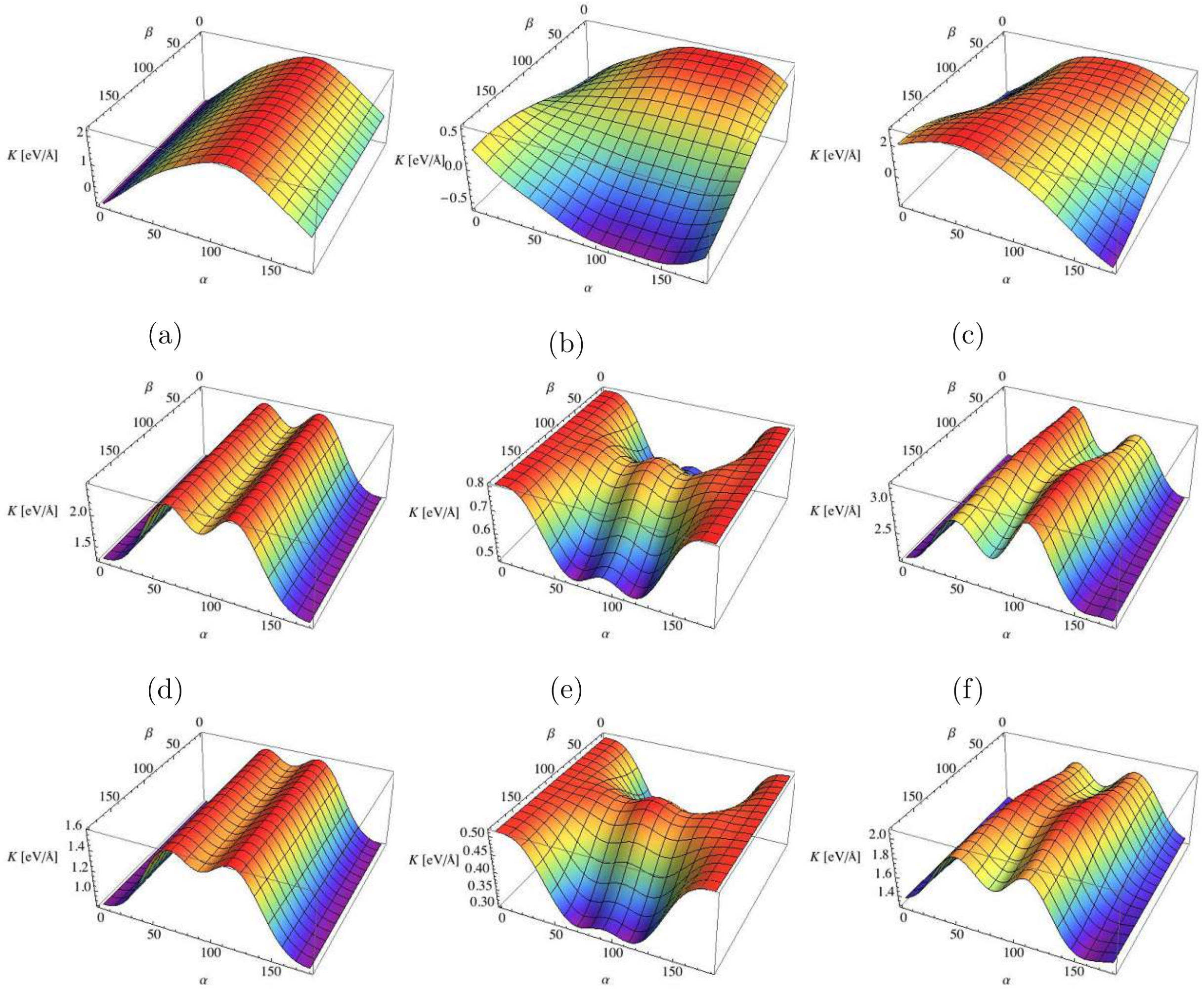
Figure 3.4: In the hexagonal (and more generally, in non-isotropic crystal) two angles, α and β, are needed for description of the dislocation line direction.

3.6 Comparison of pre-logarithmic terms based on a continuum and atomistic approach

In Section 3.5 the pre-logarithmic coefficient K was calculated as a function of the geometrical parameters α and β using an anisotropic elasticity. In this section, K is calculated using the anisotropic framework for a dislocation lying in the basal plane of $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$ and $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}$, and for a 60°-dislocation in $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ (see Figure 3.6). The K calculation is performed numerically according to equation (3.41) (dislocation in an infinitely large anisotropic medium, Steeds model [72]) using the software Wolfram Mathematica. Next the results are compared with values obtained with isotropic elasticity (dislocation in an infinitely large isotropic medium) and an atomistic scale computer simulation. The data obtained with atomistic simulations are taken from literature [5, 37, 42, 63]. Table 3.1 shows the comparison of results.

Table 3.1: Pre-logarithmic coefficients for some alloys according to the isotropic and anisotropic models. Data from atomistic simulations are included whenever available.

material	Si	$\mathrm{Si}_{0.5}\mathrm{Ge}_{0.5}$	$\mathrm{In}_{0.2}\mathrm{Ga}_{0.8}\mathrm{N}$	$\mathrm{Al}_{0.2}\mathrm{Ga}_{0.8}\mathrm{N}$	GaN
unit of measure	$\times10^{-9}$ J/m	$\times10^{-9}$ J/m	$\times10^{-9}$ J/m	$\times10^{-9}$ J/m	$\times10^{-9}$ J/m
dislocation	60° dislocation	60° dislocation	$\langle 11\bar{2}3\rangle\{1\bar{1}01\}$	$\langle 11\bar{2}3\rangle\{1\bar{1}01\}$	(a+c)-type
Isotropic elasticity	1.7	1.7	5.55	5.44	5.51
Anisotropic elasticity	0.56	0.52	3.53	3.95	3.94
Atomistic simulation		0.64 [63]			3.44 [5]



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 3.5: Pre-logarithmic coefficient K calculated as a function of the angles α and β for c-type dislocations (Burgers vector $b = [0001]$) in AlN (a), in GaN (d), in InN (g), for a-type dislocations (Burgers vector $\mathbf{b} = 1/3\,[11\bar{2}0]$) in AlN (b), in GaN (e), in InN (h), for $(a+c)$-type dislocations (Burgers vector $\mathbf{b} = 1/3\,[2\bar{1}\bar{1}3]$) in AlN (c), in GaN (f), in InN (i).

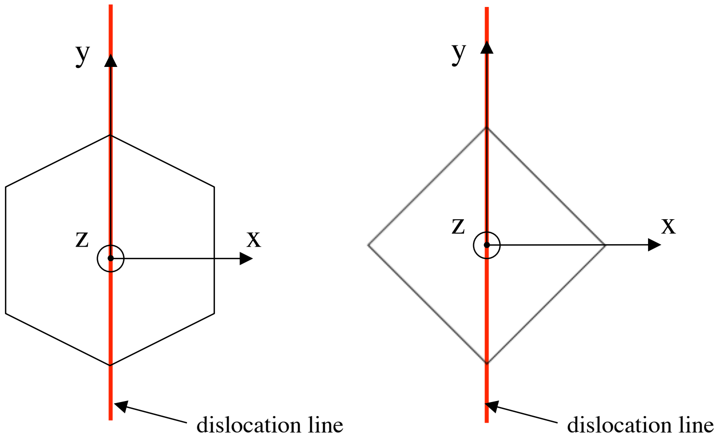


Figure 3.6: The z-axis is perpendicular to the c-plane for the hexagonal symmetry and to the closed packed plane for the cubic one. The dislocation line lies along the y-axis in both cases.

Table 3.1 shows a good agreement between Steeds' model [72] and atomistic simulations whenever the data are available in literature. This result is understandable as the atomistic simulations consider a dislocation in an infinite anisotropic structure, which is the same assumption as in Steeds' model.

The calculations in Section 4.6 require the total energy of the dislocation. To express the total energy of a dislocation, the dislocation core energy needs to be added to the continuum formula for the bulk energy (see equation (3.1)). Based on the above comparison, the continuum approach, including the elastic anisotropy, seems to be compatible with the atomistic evaluation of the dislocation core energy and radius. Consequently, the values of the dislocation core energy obtained with atomistic simulations (listed in Table 3.2) are used to calculate the total dislocation energy in Section 4.6.

Since the core energy values are available only for certain compositions, the dis-

Table 3.2: Parameters of the dislocation cores from atomistic simulations. 0.43 eV/Å is equal to 0.64 nJ/m.

material	dislocation	$\mathrm{d}\mathcal{E}_{\mathrm{core}}/\mathrm{d}y$ [eV/$\AA$]	Pre-logarithmic term [eV/$\AA$]	Reference
Si	60°	0.43		[42]
$Si_{0.5}Ge_{0.5}$	60°	0.59	0.43	[63]
GaN	(a+c)-type	3.12	2.15	[5]
GaN	a-type	1.61	0.81	[37]
AlN	a-type	1.71	0.90	[37]
InN	a-type	1.66	0.41	[37]

location core energy for each composition of the alloys considered in Section 4.6 is calculated through line interpolation (Vegard-like behavior) of the values summarized in Table 3.2. For example, in the case of $\mathrm{Si}_{1-x}\mathrm{Ge}_x$, one gets

$$\mathcal{E}_{\mathrm{core}}\Big|_{\mathrm{Si}_{1-x}\mathrm{Ge}_x} = x\ \mathcal{E}_{\mathrm{core}}\Big|_{\mathrm{Ge}} + (1-x)\ \mathcal{E}_{\mathrm{core}}\Big|_{\mathrm{Si}}. \tag{3.42}$$

The core energy $\mathcal{E}_{\mathrm{core}}$ for Ge is calculated by linearly extrapolating the core energies of Si and $\mathrm{Si}_{0.5}\mathrm{Ge}_{0.5}$. Analogous expressions are also used for $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$ and $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}$.

3.7 Dislocation configurations

The geometrical configuration of dislocations in real hexagonal single crystals depends on thermodynamic and kinetic factors. The theory described in the previous sections is used here to find the geometrical arrangement assumed by dislocations in the crystal structure in order to minimize their energy.

Sections 3.3 and 3.4 describe the procedures for the calculation of the dislocation energy in isotropic and hexagonal materials, respectively. This section shows the energetically most favorable configurations of dislocations in AlN, GaN and InN single crystals, using the above calculated pre-logarithmic coefficients (see Section 3.5). The dislocation disposition in three-dimensional space is given in terms of α and β, as shown in Figure 3.4. The dislocation arrangement is at first evaluated (Section 3.7.1) far from the free surface of the material and then (Section 3.7.2) close to it. The results are compared with those obtained within the isotropic approximation.

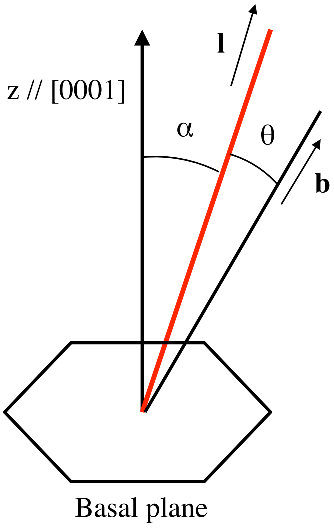


Figure 3.7: Dislocation geometry in a hexagonal crystal.

3.7.1 Dislocation types in the bulk

Assuming dislocations in the bulk means neglecting the effect of the material free surface. In the bulk, dislocations are supposed to be inside an infinite medium, and the dislocation energy per unit length of dislocation needs to be minimized in order to find the most favorable configuration. $\mathrm{d}\mathcal{E}/\mathrm{d}y$ calculated within the anisotropic approximation is directly proportional to the pre-logarithmic coefficient K. As a consequence, the most favorable configuration corresponds to the angles α and β which minimize K and $\mathrm{d}\mathcal{E}/\mathrm{d}y$.

Isotropic approximation

Holec [25] has already shown that in an infinite isotropic continuum the dislocation energy is minimized when the dislocation line is aligned with the direction of the Burgers vector (see equation (3.16)). Therefore the screw-type dislocation always has the lowest energy within the isotropic approximation.

Anisotropic approximation

In order to understand how anisotropy affects the energetically most favorable dislocation configuration, it is sufficient to evaluate the pre-logarithmic coefficient of Section 3.5. The pre-logarithmic coefficient of the c-type dislocation ($b = [0001]$) as a function of the two angles α and β is plotted in Figures 3.5(a), 3.5(d) and 3.5(g) for monocrystalline AlN, GaN and InN, respectively. In all cases, the coefficient does not depend on the angle β because the hexagonal c-plane is isotropic. Therefore all dislocation lines with different angles β are equivalent. It follows that the lowest energy configuration is obtained for $\alpha = 0°$ or $\alpha = 180°$. The Burgers vector $\mathbf{b} = 1/3\,[0001]$ is characterized by $\alpha = 0°$ and consequently its direction is coincident with the energetically most favorable dislocation line. It follows that the c-dislocation is a screw-dislocation. The situation is different for the a-type ($\mathbf{b} = 1/3\,[\bar{2}110]$) and the $(a + c)$-type ($\mathbf{b} = 1/3\,[\bar{2}113]$) dislocations.

Directions of the a-type dislocations with the lowest K and energy have $\beta = 180°$ and $\alpha \approx 130°$ for AlN, $\beta = 0°$ or $\beta = 180°$ and $\alpha \approx 65°$ or $\alpha \approx 115°$ for GaN (like [25]), $\beta = 0°$ or $\beta = 180°$ and $\alpha \approx 70°$ or $\alpha \approx 115°$ for InN. As the Burgers vector $\mathbf{b} = 1/3\,[\bar{2}110]$ is characterized by $\beta = 0°$ and $\alpha = 90°$, the Burgers vector and the dislocation line are coplanar but not orthogonal nor coincident. Therefore, the energetically most favorable configuration for the a-type dislocation is a mixed-type dislocation, unlike the screw-type dislocation predicted by the isotropic approximation.

Directions of the $(a + c)$-type dislocations with the lowest K and energy are $\beta = 0°$ and $\alpha = 0°$ or $\beta = 180°$ and $\alpha = 180°$ for AlN, $\alpha = 0°$ or $\alpha = 180°$ for GaN (like [24]), and $\alpha = 0°$ or $\alpha = 180°$ for InN. As the Burgers vector $\mathbf{b} = 1/3\,[\bar{2}113]$ is characterized by $\beta = 0°$ and $\alpha \approx 31.6°$, the Burgers vector and the dislocation line are coplanar but not orthogonal nor coincident. Therefore, the energetically most favorable configuration for the a-type dislocation is a mixed-type dislocation, which is again different from the prediction of the isotropic theory.

3.7.2 Dislocation types near the free surface

In real finite crystals, dislocations which are not closed loops have to terminate at the free surface of the crystal and, consequently, they must have a finite length. Therefore, the energetically most favorable dislocation direction is not only a function of K but also of the actual length of the dislocation. As a consequence, during the growth of a monocrystalline layer over a surface where one dislocation terminates, the dislocation elongates following the shortest path to the new free surface. This is possible because the distance between the dislocation and the newly forming free surface is so small that it can not be neglected like it is assumed for dislocations in the bulk, which are far from the free surface.

In order to consider the effect of the free surface on the dislocation configuration, γ is defined as the angle between the dislocation line and direction orthogonal to the free surface. If the film thickness grows by Δh, then the dislocation length increases by $\Delta h / \cos\alpha$. The formulas for the dislocation energy that must be used within the isotropic and the anisotropic approximation are, respectively,

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \frac{\mu b^2 \left(1 - \nu \cos^2\theta\right)}{4\pi \left(1 - \nu\right)} \frac{\Delta h}{\cos\gamma} \ln\left(\frac{R}{r_\mathrm{c}}\right), \tag{3.43}$$

and

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = K \frac{\Delta h}{\cos\gamma} \ln\left(\frac{R}{r_\mathrm{c}}\right). \tag{3.44}$$

Isotropic approximation

The angle α associated to the lowest dislocation energy is calculated minimizing the expression (3.43) for the dislocation energy with respect to α. θ depends on the position of the Burgers vector of each kind of dislocation (see Figure 3.7). The results for the angle α are summarized in Table 3.3 for monocrystalline AlN, GaN and InN. Regarding GaN, the calculated angles match with the results of Mathis [49] and Holec [25].

Table 3.3: Angle α minimizing the dislocation energy for each kind of dislocation within isotropic elasticity.

$\mathbf{b}$	AlN	GaN	InN
$\mathbf{c}$	0°	0°	0°
$\mathbf{a}$	0°	0°	0°
$(\mathbf{a}+\mathbf{c})$	14.5°	14°	10°

Anisotropic approximation

The angle α associated to the lowest dislocation energy is calculated minimizing the expression (3.44) for the dislocation energy with respect to α. This has been done for (0001)-oriented monocrystalline AlN, GaN, and InN. Different crystallographic planes have been considered as free surface.

When the monocrystal is grown in the planar growth mode (Frank-van der Merwe growth), the free surface is the (0001) plane (the polar plane). In this case the angle α between the dislocation line and the [0001] direction, which minimizes the dislocation energy, is shown in Table 3.4. The calculation reveals that for the (0001) surface, dislocation lines along the [0001] direction constitute the lowest energy configuration for all three dislocation types (a-, (a+c)- and c-type) in AlN and GaN. This contradicts the results obtained within the isotropic approximation. Regarding InN, both a- and ($a+c$)-type dislocations are inclined with respect to the [0001] direction with different angles with respect to what is predicted with the isotropic approximation.

When the monocrystal is grown with the island growth mode (Volmer-Weber growth), the free surface of an island can be the 1st order pyramidal plane (the $(1\bar{1}00)$ plane) or the 2nd order pyramidal plane (the $(11\bar{2}2)$ plane, also called the semipolar plane). For the first case, the angle α minimizing $\mathcal{E}_\mathrm{d}$ is shown in Table 3.5, while for the second case it is shown in Table 3.6. Regarding GaN, the calculated angles matches with the results of Holec [25].

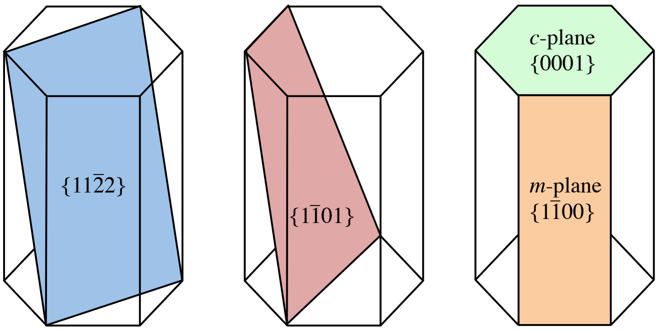


Figure 3.8: Some important crystallographic planes of the hexagonal system.

Table 3.4: Angle α minimizing the dislocation energy for each kind of dislocation in monocrystals (0001) oriented with (0001) free surface (planar growth mode).

$\mathbf{b}$	AlN	GaN	InN
$\mathbf{c}$	0°	0°	0°
$\mathbf{a}$	0°	0°	69°
$(\mathbf{a}+\mathbf{c})$	0°	0°	24°

Table 3.5: Angle α minimizing the dislocation energy for each kind of dislocation in monocrystals (0001) oriented with $(1\bar{1}00)$ free surface (island growth mode).

$\mathbf{b}$	AlN	GaN	InN
$\mathbf{c}$	0°	24°	0°
$\mathbf{a}$	0°	65°	67°
$(\mathbf{a}+\mathbf{c})$	0°	35°	0°

Table 3.6: Angle α minimizing the dislocation energy for each kind of dislocation in monocrystals (0001) oriented with $(11\bar{2}2)$ free surface (island growth mode).

$\mathbf{b}$	AlN	GaN	InN
$\mathbf{c}$	0°	23°	0°
$\mathbf{a}$	0°	64°	67°
$(\mathbf{a}+\mathbf{c})$	0°	34°	0°

Experimental TEM observations in III-nitrides regarding the dislocation inclination with respect to [0001], i.e., the angle α, are quite contradictory. Regarding GaN, Mathis [49] reports an inclination of 12°, while Holec [25] suggests that dislocations are usually vertical rather than inclined straight lines. Other authors [22, 42] mention an effective inclination of the dislocations with respect to [0001] direction without giving quantitative details.

It must be remarked that the inclination angles α here calculated are equilibrium values, *i.e.* they are obtained minimizing the dislocation energy. The real dislocation inclination could be affected by many kinetic factors, such as temperature and growth rate, which are not considered in this work.

3.8 Summary

The derivation of the dislocation energy is described in this chapter assuming an isotropic and anisotropic elasticity. In order to show the different results of the two models, they were used to calculate the equilibrium configuration for dislocations in AlN, GaN and InN compounds, all [0001] oriented. Two cases were considered.

At first, dislocations were supposed to be in bulk, far from the free surface of the material. In this case, the two theories give different results: in monocrystals the a and $(a + c)$ dislocations are not screw type - as predicted in the isotropic framework - but instead mixed type dislocations.

Secondly, dislocations were supposed to be close to the free surface, *i.e.*, the actual length of the dislocation must be considered. Different crystallographic planes were considered as free surface.

When the free surface is the (0001) plane, elastic anisotropic theory predicts that the dislocation line along the [0001] direction possesses the lowest energy configuration regardless of the dislocation type in both AlN and GaN. In InN only the c-type dislocations are screw dislocations, while the other two types are mixed dislocations.

In the the case of the inclined facets {1122} and {1101}, the results differ depending on the compound. Regarding AlN, the preferred dislocation line is the [0001] direction independent from the dislocation type and crystallographic plane as free surface. The a-type dislocations in GaN and InN propagate almost perpendicular to the {1122} and {1101} facets. Regarding the other types, in GaN the dislocations are inclined by 20°- 30°, while in InN they are aligned along the [0001] direction.

The inclined facets {1122} and {1101} are free surfaces when the island-growth mode is favored during the deposition of the compounds. The conclusion is that the island-growth mode favors the inclination of the dislocations. The importance of the inclination angle of dislocations will be shown in more detail in Chapter 5 in relation to the reduction of the dislocation density in real heterostructures.

In order to include the dislocation core energy in the evaluation of the equilibrium critical thickness (see Chapter 4), the pre-logarithmic terms of the analytical models are compared with the corresponding values obtained by atomistic simulations. Table 3.1 shows a good agreement between the continuum predictions based on anisotropic elasticity.

4 Critical Thickness

4.1 Introduction

A serious issue for the reliability and performance of heterostructure devices is the deteriorating impact of dislocations, mostly originating from epitaxy, when a thin layer is grown on a substrate with significantly different lattice parameters. Below a certain layer thickness, called the *critical thickness*, the layer is grown pseudomorphically on a substrate, i.e., the layer is grown with the same lattice parameter as the substrate. Consequently, the layer is strained, leading to large strain energy. When the critical thickness is reached, a relaxation of the strain occurs via plastic flow. The most common mechanism of plastic relaxation is an introduction of misfit dislocations along the interface between the film and the substrate.

Several models have been proposed to calculate the critical thickness. In 1974 Matthews and Blakeslee [50] compared the force exerted by strain on the extension of the misfit dislocation line with the tension in the dislocation line acting against its elongation. Their model assumes that both the film and the substrate have the same isotropic elastic properties, the film has a finite thickness (however, neglecting free surface effects) and the substrate is semi-infinite. Starting from the same hypotheses, Freund [17,19] compared the energy necessary to create a misfit dislocations with the energy inside the fully strained thin film. He arrived at the same critical thickness formula as Matthews and Blakeslee. However, the two models differ conceptually: Matthews and Blakeslee assumed a pre-existing dislocation in the substrate which creates an misfit dislocation segment along the heteroepitaxial interface due to the action of the misfit stress; Freund assumed a freshly generated dislocation at the film surface which glides into the interface between the film and the substrate. The impact of the free surface of the film and the difference between the elastic constants of the film and the substrate on the critical thickness was discussed by Willis, Jain and Bullough [82], yet they still worked in the framework of isotropic elasticity.

The anisotropy of the heterostructure can be considered in the critical thickness calculation using the methodology developed by Steeds [72] for the energy of an infinitely long straight dislocation inside an anisotropic medium. This has been done in the works by Holec *et al.* [26–28], however, the free surfaces and differences in elastic response of the film and the substrate were ignored.

Here, the gap is closed by evaluating the impact of different effects – elastic anisotropy (Section 4.3), difference between the elastic constants of the film and the substrate and the free surface of the film (Section 4.4) – on the critical thickness values. After reviewing the methodology in Sections 4.2-4.4, this method is applied to several material systems, namely, $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}/\mathrm{GaN}$, $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}/\mathrm{GaN}$, and $\mathrm{Si}_{1-x}\mathrm{Ge}_x/\mathrm{Si}$, in Section 4.5, and compared to available experimental data from literature in Section 4.6.

4.2 The critical thickness criterion

The system shown in Figure 4.1 is composed of a thin film with a finite thickness in the region $0 < z < h$, and a semi-infinite substrate filling the half space $z < 0$. The dislocation is along the interface $z = 0$ aligned along the y-axis. The critical thickness criterion compares the dislocation energy per unit length, $\mathrm{d}\mathcal{E}/\mathrm{d}y$, with the work, $\mathrm{d}\mathcal{W}/\mathrm{d}y$, done by the misfit stress, σ_{ij}^m, in forming the unit length of a dislocation along the film–substrate interface. As a consequence, the critical thickness criterion is given by

$$\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}y} = \frac{\mathrm{d}\mathcal{E}}{\mathrm{d}y}. \tag{4.1}$$

4.2.1 Work done by the misfit strain

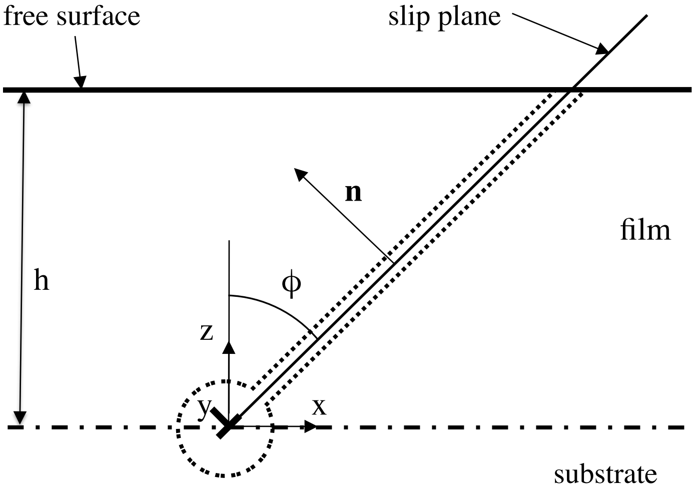


Figure 4.1: The straight infinitely long dislocation at the film–substrate interface. The film thickness is h, the slip plane is tilted by an angle ϕ from the normal to the interface.

Since the film is much thinner than the semi-infinite substrate, the strain in the substrate is assumed to be zero while the complete lattice mismatch is accommodated in the film (i.e., any mismatch strain relaxation by substrate bending is neglected). According to the geometry shown in Figure 4.1, the top surface $z = h$ of the film is traction free and therefore $\sigma_{xz}^m(x,h) = \sigma_{zz}^m(x,h) = \sigma_{yz}^m(x,h) = 0$. Assuming all directions within the xy-plane (the hexagonal plane) are equivalent, a planar stress state is obtained and thus the mismatch strain components are $\varepsilon_{xx}^m = \varepsilon_{yy}^m = \varepsilon_m$, where ε_m is the misfit strain, defined as

$$\varepsilon_m = \frac{a^* - a}{a}, \tag{4.2}$$

considering that the thin film's in-plane lattice constant a adjusts to the rigid substrate's lattice constant a^*.

Hooke's law (2.10) for isotropic structures reads

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{44} \end{pmatrix} \begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{xz} \\ 2\varepsilon_{xy} \end{pmatrix}, \tag{4.3}$$

where $c_{44} = (c_{11} - c_{12})/2$. In the case of cubic and hexagonal symmetries, the shape of the stiffness matrix c_{ij} changes to

$$\begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{44} \end{pmatrix}, \text{and} \begin{pmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix}, \tag{4.4}$$

respectively, where $c_{66} = (c_{11} - c_{12})/2$ for the hexagonal symmetry (the right matrix above).

Considering for example the hexagonal symmetry, Hooke's law gives

$$0 = \sigma_{zz}^m = c_{13}\varepsilon_{xx}^m + c_{13}\varepsilon_{yy}^m + c_{33}\varepsilon_{zz}^m \tag{4.5}$$

which results in

$$\varepsilon_{zz}^m = -2\frac{c_{13}}{c_{33}}\varepsilon_m. \tag{4.6}$$

Finally, Hooke's law yields for the mismatch stress

$$\sigma_m = \sigma_{xx}^m = \sigma_{yy}^m = \frac{(c_{11} + c_{12})c_{33} - 2c_{13}^2}{c_{33}}\varepsilon_m. \tag{4.7}$$

In the chosen Cartesian coordinate system, the Burgers vector $\mathbf{b}$, the tensor of the misfit stress $\boldsymbol{\sigma}^m$, and the outer normal to the cut surface $\Gamma = S_2 + S_4$ (see Figure 3.1) denoted by $\mathbf{n}$ take the forms

$$\mathbf{b} = \begin{pmatrix} -b\sin\theta\sin\phi \\ b\cos\theta \\ b\sin\theta\cos\phi \end{pmatrix}, \quad \boldsymbol{\sigma}^{\mathrm{m}} = \begin{pmatrix} \sigma_m & 0 & 0 \\ 0 & \sigma_m & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{n} = \begin{pmatrix} -\cos\phi \\ 0 \\ \sin\phi \end{pmatrix}, \tag{4.8}$$

where θ is the angle between the dislocation line and the Burgers vector, and ϕ the angle between the slip plane and the normal to the film-substrate interface (see Figure 4.1). The *work per unit length of the dislocation*, $\mathrm{d}\mathcal{W}/\mathrm{d}y$, done by the misfit stress is calculated according to [19]

$$\begin{aligned}\frac{d\mathcal{W}}{dy} &= \int_0^{h/\cos\phi} \sum_{i,j} b_i \sigma_{ij}^m n_j \, dz = \int_0^{h/\cos\phi} \sum_i \left(b_i \sigma_{ix} n_x + b_i \sigma_{iz} n_z\right) \, dz \\ &= b\sigma_m h \sin\theta \sin\phi. \end{aligned} \tag{4.9}$$

In a straightforward manner, analogous expressions for the other symmetries of the stiffness tensor (isotropic or cubic) can also be obtained.

4.2.2 Dislocation energy

The formula for the dislocation energy $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ for a straight dislocation inside an infinite isotropic medium reads [19, 23, 32]:

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \frac{\mu b^2 \left(1 - \nu \cos^2\varphi\right)}{4\pi \left(1 - \nu\right)} \ln\left(\frac{R}{r_\mathrm{c}}\right), \tag{4.10}$$

where μ and ν are the shear modulus and the Poisson ratio, respectively, φ is the angle between the Burgers vector $\mathbf{b}$ and the dislocation line, and R is the outer cut-off radius. Here it is taken to be equal to the film thickness.

Chapter 3.2 describes how to calculate the dislocation energy $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ inside an infinite anisotropic medium according to the procedure developed by Steeds [72] and described in Section 3.4:

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = K \ln\left(\frac{\mathrm{R}}{r_\mathrm{c}}\right), \tag{4.11}$$

where K is the pre-logarithmic factor and it is a function of the crystal structure and the elastic properties (see Section 3.4). Holec [25] derived a simplified expression of K from the Steeds procedure specifically for a dislocation along the y-axis inside a hexagonal crystal. The expression is easily adapted to the cubic symmetry by simply modifying the stiffness tensor.

If we want to calculate the dislocation energy at the interface between the film and the substrate of a heteroepitaxial structure, it is necessary to evaluate the effect of the free surface of the film and the different elastic properties of the two materials. Willis, Jain and Bullough [82] have evaluated both effects on the dislocation energy within isotropic elasticity.

The adaptation of the Willis, Jain and Bullough model to the hexagonal symmetry can be found in the works of Holec [24] and Coppeta [8]. This procedure is reported in details in the next Paragraph 4.2.3 for a straight dislocation at the interface between a finite hexagonal film and a semi-infinite hexagonal substrate with different elastic constants.

In conclusion, four approaches (summarized in Table 4.1) for calculating the dislocation energy were defined above, which will be used to evaluate the impact of different approximations.

4.2.3 A straight dislocation at the interface of anisotropic materials

The evaluation of the integrals along the surfaces S_2 and S_4 (see Figure 3.1) deals with a straight dislocation at the interface between a finite anisotropic film and a semi-infinite anisotropic substrate with different elastic properties:

$$\mathrm{d}\mathcal{E}_{\mathrm{d}} = \tfrac{1}{2}\sum_{ij}\left(\int_{S_2} u_i\sigma^d_{ij}n_j\,\mathrm{d}S_2 + \int_{S_4} u_i\sigma^d_{ij}n_j\,\mathrm{d}S_4\right). \tag{4.12}$$

Considering Figure 4.2, the previous equation becomes

$$\tfrac{1}{2}\sum_{i}\left(\int_{r_c}^{h} u_i(x\to 0^+)\sigma^d_{ix}\,\mathrm{d}z + \int_{h}^{r_c} u_i(x\to 0^-)\sigma^d_{ix}\,\mathrm{d}z\right). \tag{4.13}$$

If the extremes of the second integral are inverted, one obtains

$$\frac{\mathrm{d}\mathcal{E}_{\mathrm{d}}}{\mathrm{d}y} = \tfrac{1}{2}\sum_{i}\int_{r_c}^{h} b_i\sigma^d_{ix}\,\mathrm{d}z \tag{4.14}$$

where $u_i(x\to 0^+) - u_i(x\to 0^-) = b_i$.

The evaluation of the stress components in (4.14) is performed through the treatment suggested by the Willis, Jain and Bullough model [82] for hexagonal and cubic symmetries following the Steeds procedure [72].

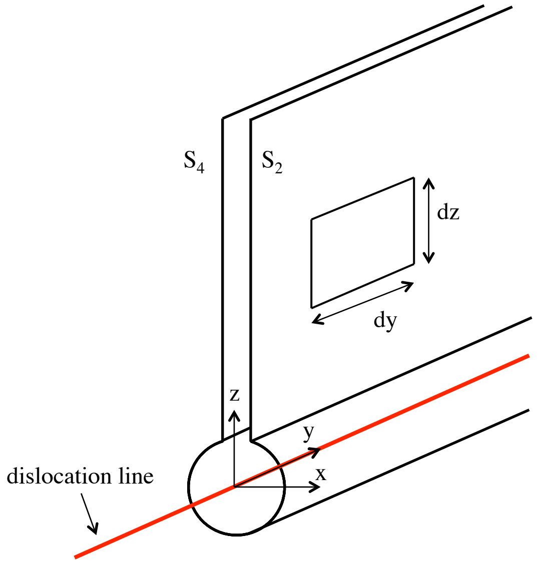


Figure 4.2: The z-axis is perpendicular to the c-plane for the hexagonal symmetry and to the closed packed plane for the cubic one. The dislocation line lies along the y-axis in both cases.

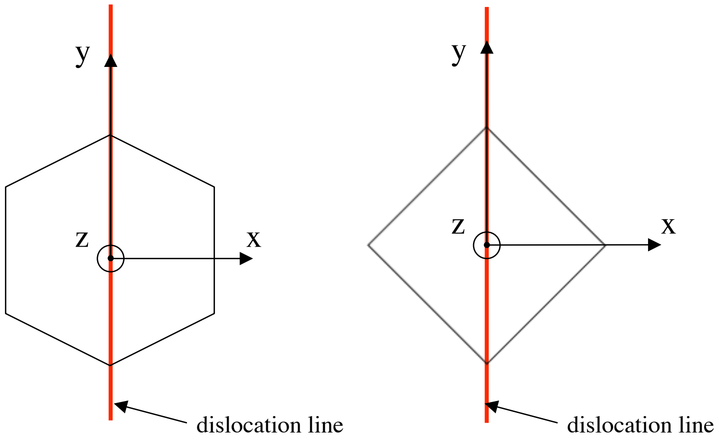


Figure 4.3: The z-axis is perpendicular to the c-plane for the hexagonal symmetry and to the closed packed plane for the cubic one. The dislocation line lies along the y-axis in both cases.

The dislocation of interest is a misfit dislocation along the c-plane of the wurtzite structure or the closed packed plane of the diamond structure. The z-axis is perpendicular to the c-plane or to the closed packed plane, respectively (See Figure 4.3). The dislocation is considered straight and extended to infinity along the y-axis. This assumption simplifies the problem to a plane strain problem where no quantity depends on the y-coordinate, so

$$\frac{\partial}{\partial y} = 0, \tag{4.15}$$

The Burgers vector of the dislocation is **b**. Displacements are given by the functions u_x, u_y and u_z. The displacements u_x and u_z correspond to the edge component of the considered dislocation whereas u_y corresponds to the screw component. The strain components are

$$\begin{aligned} \varepsilon_{xx} &= \frac{\partial u_x}{\partial x}, & \varepsilon_{xy} &= \frac{1}{2}\frac{\partial u_y}{\partial x}, \\ \varepsilon_{yy} &= 0, & \varepsilon_{xz} &= \frac{1}{2}\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right), \\ \varepsilon_{zz} &= \frac{\partial u_z}{\partial z}, & \varepsilon_{yz} &= \frac{1}{2}\frac{\partial u_y}{\partial z}. \end{aligned} \tag{4.16}$$

The compatibility equations [72] provide two relations:

$$\frac{\partial^2 \varepsilon_{xx}}{\partial z^2} + \frac{\partial^2 \varepsilon_{zz}}{\partial x^2} = 2\frac{\partial \varepsilon_{xz}}{\partial x \partial z}, \tag{4.17a}$$

$$\frac{\partial \varepsilon_{yz}}{\partial x} - \frac{\partial \varepsilon_{xy}}{\partial z} = 0. \tag{4.17b}$$

The fact that $\varepsilon_{yy} = 0$ yields a relation between particular stress components:

$$0 = \varepsilon_{yy} = s_{12}\sigma_{xx} + s_{11}\sigma_{yy} + s_{13}\sigma_{zz} \quad \Rightarrow \quad \sigma_{yy} = -\frac{s_{12}}{s_{11}}\sigma_{xx} - \frac{s_{13}}{s_{11}}\sigma_{zz}. \tag{4.18}$$

The compliances reflecting the proper hexagonal symmetry have been used to obtain:

$$\begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{xz} \\ 2\varepsilon_{xy} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{13} & 0 & 0 & 0 \\ s_{13} & s_{13} & s_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{66} \end{pmatrix} \begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix}, \tag{4.19}$$

where

$$s_{66} = 2\left(s_{11} - s_{12}\right). \tag{4.20}$$

The shape of the tensor for cubic symmetry (in both Si(100) and Si(110)) is the same as for the hexagonal one. The subsequent treatment is valid for the cubic symmetry with $s_{66} = s_{44}$ and $s_{13} = s_{12}$.

According to the Willis, Jain and Bullough model, it is useful now to have the jumps in displacements occur over a surface which is perpendicular to the free surface instead of across the slip plane. All quantities in the substrate have a $*$ superscript to distinguish them from quantities related to the thin film. The Fourier transform of a function f is denoted by $\widetilde{f}$. It is possible to split the problem into two independent parts, resolving the edge and the screw components separately.

Edge component

The displacement field in the thin film can be decomposed into

$$\boldsymbol{u} = \frac{1}{2}\boldsymbol{b_e}\mathrm{sgn}(x) + \boldsymbol{v}, \tag{4.21}$$

where $\boldsymbol{b_e} = (b_x, 0, b_z)$ and the function $\boldsymbol{v}$ is continuous for all x. The strain field (and thus also the stress field) inside the thin layer is determined only by the $\boldsymbol{v}$ part as (except at the cut surface $x = 0$) the derivatives of the sgn(x) function are equal to zero everywhere. To solve this problem, Fourier-transformed variables are employed:

$$\widetilde{f}(\xi, z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{i\xi x} f(x, z) dx. \tag{4.22}$$

The Fourier transform of the $\partial/\partial x$ operator is $-i\xi$. Using (4.21) and (4.16), the Fourier components of the strain tensor are:

$$\widetilde{\epsilon}_{xx}(\xi, z) = -i\xi\widetilde{v}_x(\xi, z), \tag{4.23a}$$

$$\widetilde{\epsilon}_{zz}(\xi, z) = \frac{\partial\widetilde{v}_z}{\partial z}(\xi, z), \tag{4.23b}$$

$$\widetilde{\epsilon}_{xz}(\xi, z) = \frac{1}{2}\left(\frac{\partial\widetilde{v}_x}{\partial z}(\xi, z) - i\xi\widetilde{v}_z(\xi, z)\right). \tag{4.23c}$$

The equilibrium conditions [72] take the form

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} = 0, \tag{4.24a}$$

$$\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} = 0. \tag{4.24b}$$

Transforming these equations into their Fourier equivalents and using Hooke's law with the stiffness tensor gives the following system of partial differential equations for the Fourier-transformed components of the displacements $\widetilde{v}_x$ and $\widetilde{v}_z$:

$$\begin{pmatrix} -c_{11}\xi^2 + c_{44}\dfrac{\partial^2}{\partial z^2} & -i\xi\left(c_{13}+c_{44}\right)\dfrac{\partial}{\partial z} \\ -i\xi\left(c_{13}+c_{44}\right)\dfrac{\partial}{\partial z} & -c_{44}\xi^2 + c_{33}\dfrac{\partial^2}{\partial z^2} \end{pmatrix} \begin{pmatrix} \widetilde{v}_x \\ \widetilde{v}_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{4.25}$$

In order to simplify the notation, the quantities $A = -c_{11}$, $B = c_{44}$, $C = -i(c_{13}+c_{44})$, $D = -c_{44}$, $E = c_{33}$ are introduced. From the second equation of the system one obtains

$$\xi C \frac{\partial \widetilde{v}_x}{\partial z} = -D\xi^2 \widetilde{v}_z - E\frac{\partial^2 \widetilde{v}_z}{\partial z^2}. \tag{4.26}$$

Using this equation and its twice differentiated form with respect to z and substituting them into the first equation of the system (4.25), which is once differentiated with respect to z, yields

$$EB\frac{\partial^4 \widetilde{v}_z}{\partial z^4} + \left(EA + DB - C^2\right)\xi^2 \frac{\partial^2 \widetilde{v}_z}{\partial z^2} + AD\xi^4 \widetilde{v}_z = 0. \tag{4.27}$$

The solution has the exponential form $\mathrm{e}^{\lambda|\xi|z}$. The corresponding proper λs are solutions of the characteristic equation

$$EB\lambda^4 + \left(EA + DB - C^2\right)\lambda^2 + AD = 0. \tag{4.28}$$

The solutions are $\lambda_1 = \kappa_1$, $\lambda_2 = -\kappa_1$, $\lambda_3 = \kappa_2$ and $\lambda_4 = \kappa_2$, where

$$\kappa_{1,2} = \sqrt{\frac{-\left(EA + DB - C^2\right) \pm \sqrt{\left(EA + DB - C^2\right)^2 - 4ABDE}}{2EB}}. \tag{4.29}$$

Therefore, the general solution to (4.27) has the form [72]

$$\widetilde{v}_z = \sum_{i=1}^{4} A_i \lambda_i \left|\xi\right| \mathrm{e}^{\lambda_i|\xi|z}, \tag{4.30}$$

where A_i are functions of ξ to be determined. The component $\widetilde{v}_x$ can then be obtained from (4.26). Its general solution is

$$\widetilde{v}_x = -\sum_{i=1}^{4} A_i \xi \frac{D + E\lambda_i^2}{C} \mathrm{e}^{\lambda_i|\xi|z}. \tag{4.31}$$

Evidently

$$\frac{D+E\lambda_i^2}{C} = -i\frac{c_{44}-c_{33}\lambda_i^2}{c_{13}+c_{44}}. \tag{4.32}$$

The Fourier-transformed components of the strain tensor are now easy to obtain from combining Equations (4.30), (4.31), and (4.23):

$$\widetilde{\varepsilon}_{xx} = \sum_{i=1}^{4} A_i \xi^2 \frac{c_{44}-c_{33}\lambda_i^2}{c_{13}+c_{44}} \mathrm{e}^{\lambda_i |\xi| z}, \tag{4.33a}$$

$$\widetilde{\varepsilon}_{xz} = -\frac{1}{2}\sum_{i=1}^{4} iA_i \left|\xi\right| \xi \lambda_i \frac{c_{13}+c_{33}\lambda_i^2}{c_{13}+c_{44}} \mathrm{e}^{\lambda_i |\xi| z}, \tag{4.33b}$$

$$\widetilde{\varepsilon}_{zz} = \sum_{i=1}^{4} A_i \xi^2 \lambda_i^2 \mathrm{e}^{\lambda_i |\xi| z}. \tag{4.33c}$$

Using Hooke's law, which has the same form when Fourier transformed, one obtains:

$$\widetilde{\sigma}_{xx} = c_{11}\widetilde{\varepsilon}_{xx} + c_{13}\widetilde{\varepsilon}_{zz} = -\sum_{i=1}^{4} A_i \xi^2 \lambda_i^2 c_{44} \frac{c_{13}+c_{33}\lambda_i^2}{c_{13}+c_{44}} \mathrm{e}^{\lambda_i |\xi| z}, \tag{4.34a}$$

$$\widetilde{\sigma}_{xz} = 2c_{44}\widetilde{\varepsilon}_{xz} = -\sum_{i=1}^{4} iA_i \left|\xi\right| \xi \lambda_i c_{44} \frac{c_{13}+c_{33}\lambda_i^2}{c_{13}+c_{44}} \mathrm{e}^{\lambda_i |\xi| z}, \tag{4.34b}$$

$$\widetilde{\sigma}_{zz} = c_{13}\widetilde{\varepsilon}_{xx} + c_{33}\widetilde{\varepsilon}_{zz} = \sum_{i=1}^{4} A_i \xi^2 c_{44} \frac{c_{13}+c_{33}\lambda_i^2}{c_{13}+c_{44}} \mathrm{e}^{\lambda_i |\xi| z}, \tag{4.34c}$$

where the identity $c_{11}c_{44} + \left(c_{13}^2 + 2c_{13}c_{44} - c_{11}c_{33}\right)\lambda_i^2 + c_{44}c_{33}\lambda_i^4 = 0$ was used in the expression for $\widetilde{\sigma}_{xx}$.

The general solution for the substrate region takes the same form (except that all variables have a star superscript). It is assumed that the substrate is not influenced by the thin layer as $z \to -\infty$. To fulfill this condition, the constants A_2^* and A_4^* must be identically zero. Boundary conditions must be employed now in order to determine the constants A_1, A_2, A_3, A_4, A_1^*, and A_3^*.

The Fourier transform of $1/2\ \mathrm{sgn}(x)$ is $\mathrm{i}/\left(\sqrt{2\pi}\xi\right)$. The continuity of displacements across the interface $z=0$ is expressed by the following equations:

$$\widetilde{v}_x\left(\xi,0\right) + \frac{ib_x}{\sqrt{2\pi}\xi} = \widetilde{v}_x^*\left(\xi,0\right), \tag{4.35a}$$

$$\widetilde{v}_z\left(\xi,0\right) + \frac{ib_z}{\sqrt{2\pi}\xi} = \widetilde{v}_z^*\left(\xi,0\right), \tag{4.35b}$$

which result in the equations

$$\left(A_1+A_2\right)\Gamma_1 + \left(A_3+A_4\right)\Gamma_2 + \frac{b_x}{\sqrt{2\pi}\xi^2} = A_1^*\Gamma_1^* + A_3^*\Gamma_2^*, \tag{4.36a}$$

$$\left(A_1-A_2\right)\kappa_1 + \left(A_3-A_4\right)\kappa_2 + \frac{i\,b_z}{\sqrt{2\pi}\left|\xi\right|\xi} = A_1^*\kappa_1^* + A_3^*\kappa_2^*, \tag{4.36b}$$

where the notation

$$\Gamma_i = \frac{c_{44} - c_{33}\kappa_i^2}{c_{13} + c_{44}}, \qquad \Gamma_i^* = \frac{c_{44}^* - c_{33}^*\kappa_i^{*2}}{c_{13}^* + c_{44}^*}, \qquad i = 1, 2. \tag{4.37}$$

has been introduced.

The second set of equations is obtained from the requirement that tractions must be continuous across the interface $z = 0$:

$$\widetilde{\sigma}_{xz}\left(\xi, 0\right) = \widetilde{\sigma}_{xz}^*\left(\xi, 0\right), \tag{4.38a}$$

$$\widetilde{\sigma}_{zz}\left(\xi, 0\right) = \widetilde{\sigma}_{zz}^*\left(\xi, 0\right). \tag{4.38b}$$

Combining Equations (4.34) and (4.38) gives

$$\left(A_1 - A_2\right)\kappa_1\Lambda_1 + \left(A_3 - A_4\right)\kappa_2\Lambda_2 = A_1^*\kappa_1^*\Lambda_1^* + A_3^*\kappa_2^*\Lambda_2^* \tag{4.39a}$$

$$\left(A_1 + A_2\right)\Lambda_1 + \left(A_3 + A_4\right)\Lambda_2 = A_1^*\Lambda_1^* + A_3^*\Lambda_2^*, \tag{4.39b}$$

where the simplifying notation

$$\Lambda_i = \frac{c_{44}\left(c_{13} + c_{33}\kappa_i^2\right)}{c_{13} + c_{44}}, \qquad \Lambda_i^* = \frac{c_{44}^*\left(c_{13}^* + c_{33}^*\kappa_i^2\right)}{c_{13}^* + c_{44}^*}, \qquad i = 1, 2, \tag{4.40}$$

has been introduced. The last set of equations arises from the requirement that the free surface is at $z = h$.

The condition $\widetilde{\sigma}_{xz}\left(\xi, h\right) = 0$ is expressed as

$$\left(A_1 \mathrm{e}^{\kappa_1|\xi|h} - A_2 \mathrm{e}^{-\kappa_1|\xi|h}\right)\kappa_1\Lambda_1 + \left(A_3 \mathrm{e}^{\kappa_2|\xi|h} - A_4 \mathrm{e}^{-\kappa_2|\xi|h}\right)\kappa_2\Lambda_2 = 0 \tag{4.41}$$

and the requirements $\widetilde{\sigma}_{zz}\left(\xi, h\right) = 0$ yields

$$\left(A_1 \mathrm{e}^{\kappa_1|\xi|h} + A_2 \mathrm{e}^{-\kappa_1|\xi|h}\right)\Lambda_1 + \left(A_3 \mathrm{e}^{\kappa_2|\xi|h} - A_4 \mathrm{e}^{-\kappa_2|\xi|h}\right)\Lambda_2 = 0. \tag{4.42}$$

Equations (4.36), (4.39), (4.41) and (4.42) constitute a linear system of equations for the unknown variables $A_1, ..., A_3^*$. Obviously, the solution of this system provides $A_1, ..., A_3^*$ as functions of ξ. To obtain the stress components σ_{xx}, σ_{xz} and σ_{zz}, it is necessary to perform the inverse Fourier transforms of $\widetilde{\sigma}_{xx}$, $\widetilde{\sigma}_{xz}$ and $\widetilde{\sigma}_{zz}$, respectively, with the substituted resolved constants $A_1, ..., A_4$.
An attempt to obtain an analytical formula for the stress components and energy for the edge type dislocation would be unreasonably complicated. Therefore the calculations were performed numerically.

Screw component

The procedure for the screw component is similar but simpler. The Burgers vector of a screw dislocation has only one non-zero component, b_y, and thus the only non-zero component of the displacement is u_y. Moreover, it is again a plane strain problem, i.e., $\partial/\partial y = 0$. As a consequence, the only non-zero strain components are ε_{xy} and

ε_{yz}. Similarly to the case of the edge dislocation, it is more convenient to deal with the Fourier-transformed component $\widetilde{u}_y$ rather than with u_y itself. The Fourier-transformed components of the strain are

$$\widetilde{\varepsilon}_{xy} = -\frac{1}{2} i\xi \widetilde{v}_y, \tag{4.43a}$$

$$\widetilde{\varepsilon}_{yz} = \frac{1}{2}\frac{\partial \widetilde{v}_y}{\partial z}, \tag{4.43b}$$

and the FT stress components obtained from Hooke's law are

$$\widetilde{\sigma}_{xy} = 2c_{66}\widetilde{\varepsilon}_{xy} = -i\xi c_{66}\widetilde{v}_y, \tag{4.44a}$$

$$\widetilde{\sigma}_{yz} = 2c_{44}\widetilde{\varepsilon}_{yz} = c_{44}\frac{\partial \widetilde{v}_y}{\partial z}. \tag{4.44b}$$

The equilibrium condition [72]

$$\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial z} = 0 \tag{4.45}$$

provides the second-order differential equation for $\widetilde{v}_y$:

$$-\xi^2 c_{66}\widetilde{v}_y\left(\xi, z\right) + c_{44}\frac{\partial^2 \widetilde{v}_y}{\partial z^2}\left(\xi, z\right) = 0, \tag{4.46}$$

the general solution of which is

$$\widetilde{v}_y = A_1 \mathrm{e}^{\sqrt{\frac{c_{66}}{c_{44}}}|\xi| z} + A_2 \mathrm{e}^{-\sqrt{\frac{c_{66}}{C_{44}}}|\xi| z}. \tag{4.47}$$

The general solution in the substrate has exactly the same form except for the fact that all variables have a superscript. Since the substrate is assumed to be semi-infinite, one requires all quantities to vanish as $z \to -\infty$ and therefore $A_2^* = 0$. Applying the same boundary conditions as in the case of the edge dislocation (continuity of displacements and tractions over the interface $z = 0$ and the free surface at $z = h$) yields

$$A_1 + A_2 + \frac{i b_y}{\sqrt{2\pi}\xi} = A_1^*, \tag{4.48a}$$

$$\left(A_1 - A_2\right)\sqrt{c_{66}c_{44}} = A_1^*\sqrt{c_{66}^* c_{44}^*}, \tag{4.48b}$$

$$A_1 \mathrm{e}^{\sqrt{\frac{c_{66}}{c_{44}}}|\xi| h} - A_2 \mathrm{e}^{-\sqrt{\frac{c_{66}}{c_{44}}}|\xi| h} = 0. \tag{4.48c}$$

Values of A_1 and A_2 fulfilling these equations are

$$A_1 = \frac{i b_y}{\sqrt{2\pi}\xi}\frac{\sqrt{c_{66}^* c_{44}^*}\,\mathrm{e}^{-2\sqrt{\frac{c_{66}}{c_{44}}}|\xi| h}}{\mathrm{e}^{-2\sqrt{\frac{c_{66}}{c_{44}}}|\xi| h}\left(\sqrt{c_{66}c_{44}} - \sqrt{c_{66}^* c_{44}^*}\right) - \left(\sqrt{c_{66}c_{44}} + \sqrt{c_{66}^* c_{44}^*}\right)}, \tag{4.49a}$$

$$A_2 = \frac{i b_y}{\sqrt{2\pi}\xi}\frac{\sqrt{c_{66}^* c_{44}^*}}{\mathrm{e}^{-2\sqrt{\frac{c_{66}}{c_{44}}}|\xi| h}\left(\sqrt{c_{66}c_{44}} - \sqrt{c_{66}^* c_{44}^*}\right) - \left(\sqrt{c_{66}c_{44}} + \sqrt{c_{66}^* c_{44}^*}\right)}. \tag{4.49b}$$

The stress components can be obtained by the inverse Fourier transform of $\widetilde{\sigma}_{xy}$ and $\widetilde{\sigma}_{xz}$ with the substituted A_1 and A_2 from the last two equations.

After that, the components of the stress and the Burgers vector can be substituted in (4.14) to calculate the dislocation energy.

4.2.4 The critical thickness models

According to equation (4.1), equating the work done by the misfit strain with each of the four formulas for dislocation energy yields four different models with which to calculate the equilibrium critical thickness . These models are summarized with the respective hypotheses in Table 4.1 and their results are discussed in the subsequent Paragraphs. In particular, the critical thickness model developed by Freund [17], which neglects the differences between film and substrate, considers only isotropic elasticity, and in its simplified (often used) formula also neglects the effect of the free surface. The Freund model is corrected for the elastic anisotropy using the Steeds treatment for the dislocation energy [72]. The fully isotropic model, but with the free surface and different elastic constants in the film and the substrate, employs the Willis, Jain and Bullough formula [82]. Finally, a model addressing all effects [8, 24] is reported in Paragraph 4.2.3.

Table 4.1: An overview of different assumptions for evaluating misfit dislocation energy, and equilibrium critical thickness .

	Freund [17, 19]	Steeds [72]	Willis *et al.* [82]	Steeds + Willis *et al.* [8, 24]
Anisotropy	no	yes	no	yes
Different elastic properties of the film and the substrate	no	no	yes	yes
Free surface of the film	no	no	yes	yes

4.3 Impact of the anisotropy on the critical thickness criterion

Steeds developed a procedure to calculate the energy of one infinite straight dislocation in the bulk of an infinite material [72], which is considered anisotropic with a certain crystallographic symmetry. The result for the hexagonal symmetry is briefly described in [28].

Here, this treatment is adopted in order to calculate the energy (with and without the contribution of the dislocation core, see Chapter 3.2) for a dislocation whose slip system is $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ or $1/3\langle 11\bar{2}3\rangle\{1\bar{1}01\}$ in $\mathrm{Al_{0.2}Ga_{0.8}N}$ and in $\mathrm{In_{0.2}Ga_{0.8}N}$ (see Figure 4.4). The same treatment with the appropriate elastic constants is used to calculate the energy for the so called 60° dislocation slip system $\langle 110\rangle\{111\}$ in $\mathrm{Si_{0.8}Ge_{0.2}}$ (see Figure 4.5).

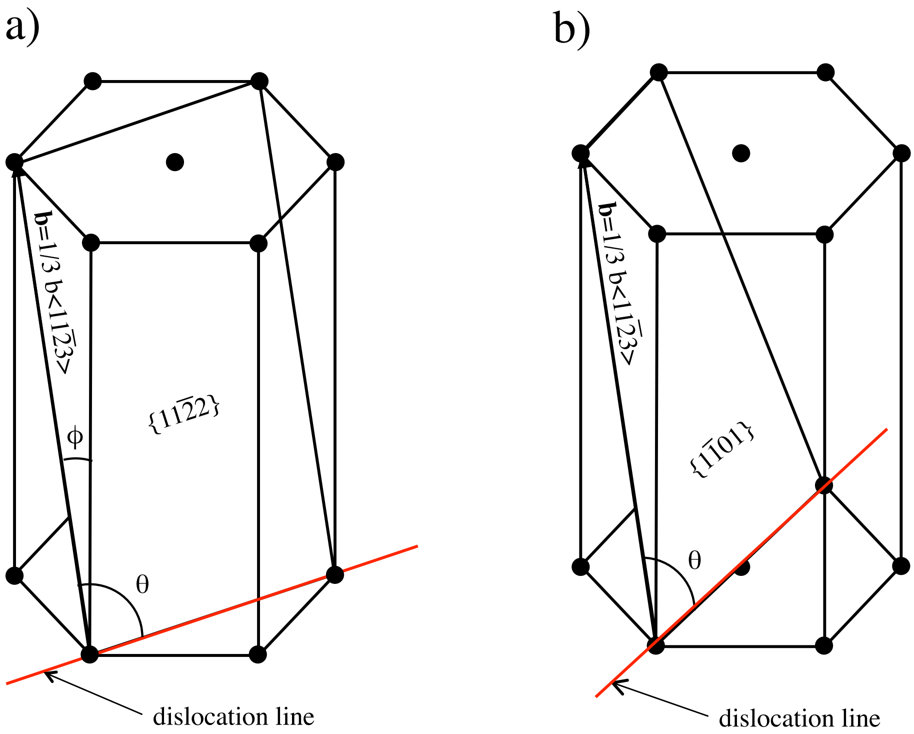


Figure 4.4: The two most favorable slip systems in the systems $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}/\mathrm{GaN}$ and $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}/\mathrm{GaN}$: a) $\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ observed by Srinivasan [71] and b) $\langle 11\bar{2}3\rangle\{1\bar{1}01\}$ determined by Jahnen [33].

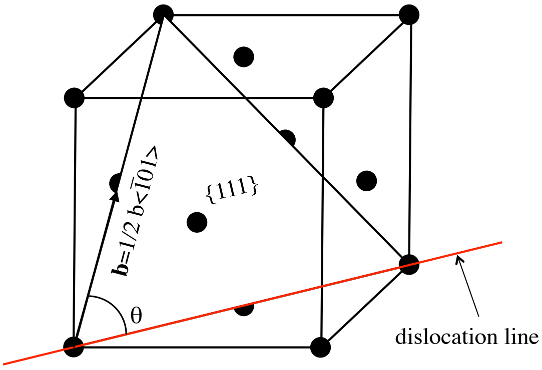


Figure 4.5: The slip system $\langle 110\rangle\{111\}$ of the 60° dislocation shown in the FCC structure.

In the following, the difference between the dislocation energy $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ and the work $\mathrm{d}\mathcal{W}/\mathrm{d}y$ done by the misfit stress field for each alloy is discussed. When the resulting value is positive, fully coherent accommodation of the misfit strain is energetically preferred, while for negative values introduction of misfit dislocations becomes favored.

The highest film thickness yielding the difference of 0 indicates the equilibrium critical thickness . The resulting values for the three different alloys as a function of the film thickness are plotted in Figures 4.6, 4.7, and 4.8. Each figure has two sets of curves, one for the Freund,(isotropic elasticity) and one for the Steeds (anisotropic elasticity) procedure. For all systems investigated here, the anisotropy lowers the critical thickness . Additionally, it turns out that the inclusion of the integral along the core surface S_3 (values labelled "with $\mathrm{E_{cs}}$" in the figures. For the description of $\mathrm{E_{cs}}$ see Section 3.2) has in all cases only a negligible impact (or at least on order of magnitude smaller effect than the correct crystal symmetry) on the predicted critical thickness (see Table 4.2).

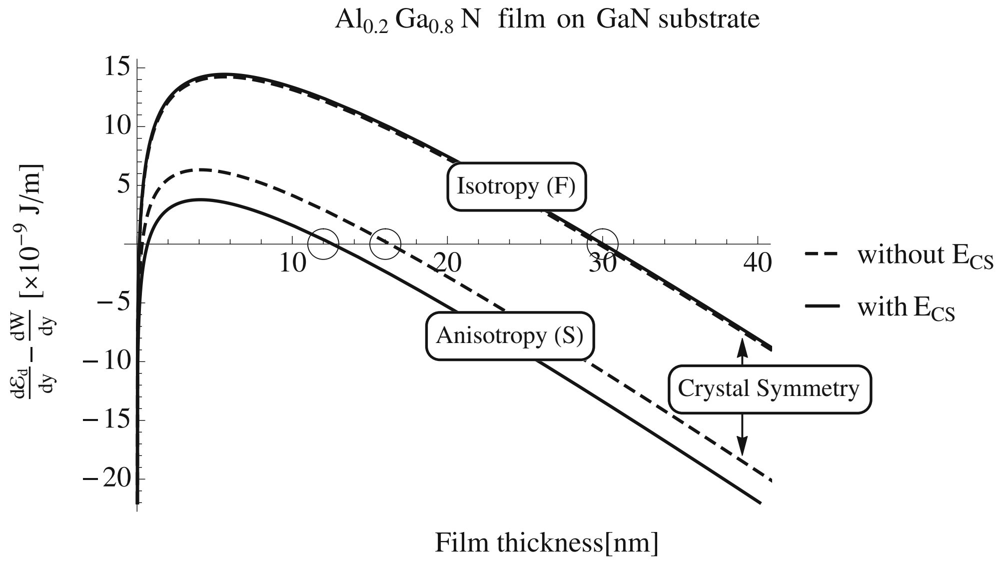


Figure 4.6: $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y - \mathrm{d}\mathcal{W}/\mathrm{d}y$ as function of the $\mathrm{Al_{0.2}Ga_{0.8}N}$ film thickness. $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ is calculated assuming isotropic (Freund model (F)) and anisotropic (Steeds model (S)) elasticity, with or without the evaluation of the integral along the core surface $\mathrm{E_{cs}}$. $\mathrm{d}\mathcal{W}/\mathrm{d}y$ is calculated according to equation (4.9). The critical thickness values of an $\mathrm{Al_{0.2}Ga_{0.8}N}$ film grown on a GaN substrate are indicated by a circle.

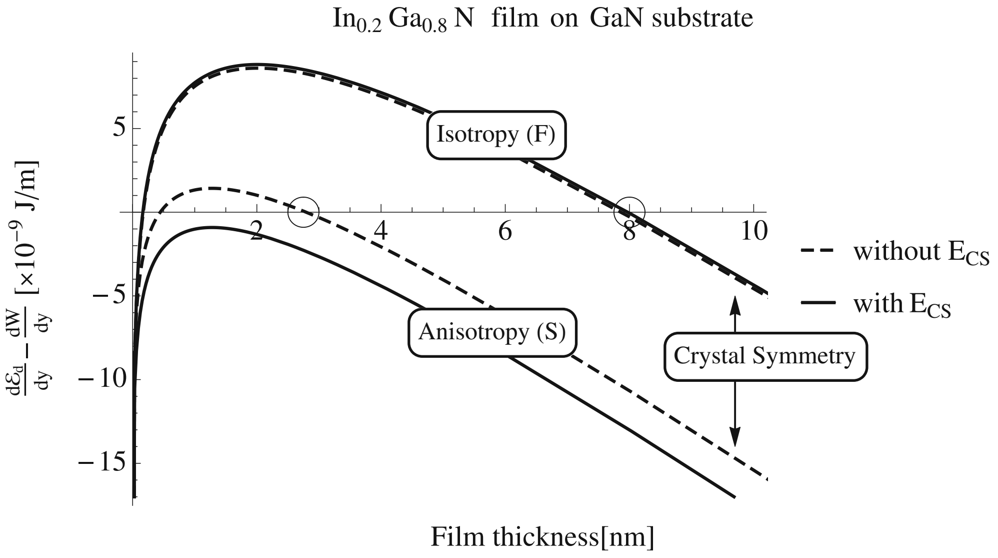


Figure 4.7: Same as in Figure 4.6 but for an $In_{0.2}Ga_{0.8}N$ film grown on a GaN substrate.

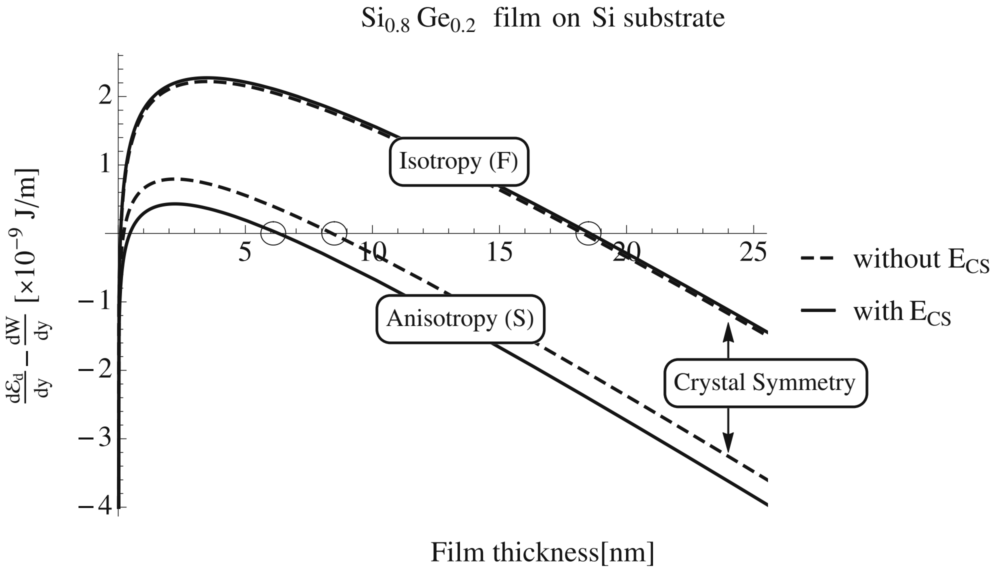


Figure 4.8: Same as in Figure 4.6 but for a $Si_{0.8}Ge_{0.2}N$ film grown on a Si substrate.

4.4 The critical thickness according to Willis, Jain and Bullough

Willis, Jain, and Bullough [82] derived a procedure to calculate the energy $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ of a misfit dislocation at the interface between a film with a finite thickness and a semi-infinite substrate. The film and the substrate are both supposed to be isotropic but with different elastic properties.

Figure 4.9 shows two sets of curves representing $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ as a function of the film thickness for three different systems: $\mathrm{Al_{0.2}Ga_{0.8}N}$ film on a GaN substrate, $\mathrm{In_{0.2}Ga_{0.8}N}$ on GaN, and a GaN film grown on a GaN substrate. One set of the $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ curves is calculated with the Willis *et al.* model, while the second uses the Freund treatment. Therefore, the difference stems from including (Willis *et al.*) or neglecting (F) the free surface effects and the difference in the elastic constants of the film and the substrate. In both cases the isotropic elasticity framework is used. The curves within each set are very close to each other, meaning that the impact of different elastic properties of the film and the substrate is negligible for these material systems. Therefore it can be concluded that the difference between the two sets originates predominantly from the impact of free surface. This factor significantly increases the $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ term and, as a consequence, also the critical thickness value.

Figure 4.10 shows a similar analysis for two different systems: a $\mathrm{Si_{0.2}Ge_{0.8}}$ film on a Si substrate and a Si film grown on a Si substrate. As in the case of the III-nitrides, the variation caused by the different elastic properties of the film and the substrate is negligible. On the other hand, the difference between the Willis *et al.* and Freund formalism, which is now related predominantly to the inclusion of the free surface, is significant. Therefore, also for silicon as a substrate the free surface increases $\mathrm{d}\mathcal{E}_\mathrm{d}/\mathrm{d}y$ and, as a consequence, the critical thickness value.

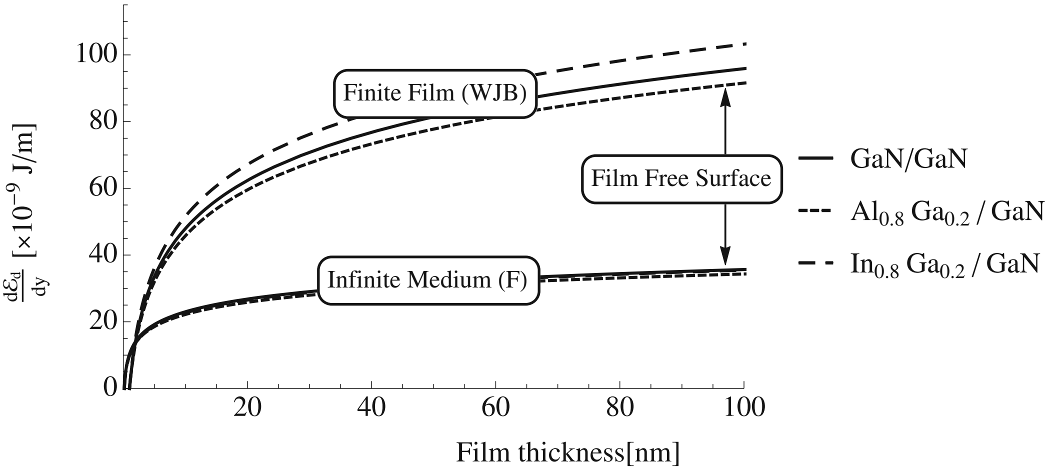


Figure 4.9: $d\mathcal{E}_d/dy$ is a function of the film thickness for the material systems $Al_xGa_{1-x}N/GaN$ and $In_xGa_{1-x}N/GaN$. The $\langle 11\bar{2}3\rangle\{1\bar{1}01\}$ slip system is considered. The two sets of curves are calculated through the Willis *et al.* (WJB) and Freund (F) procedures, respectively.

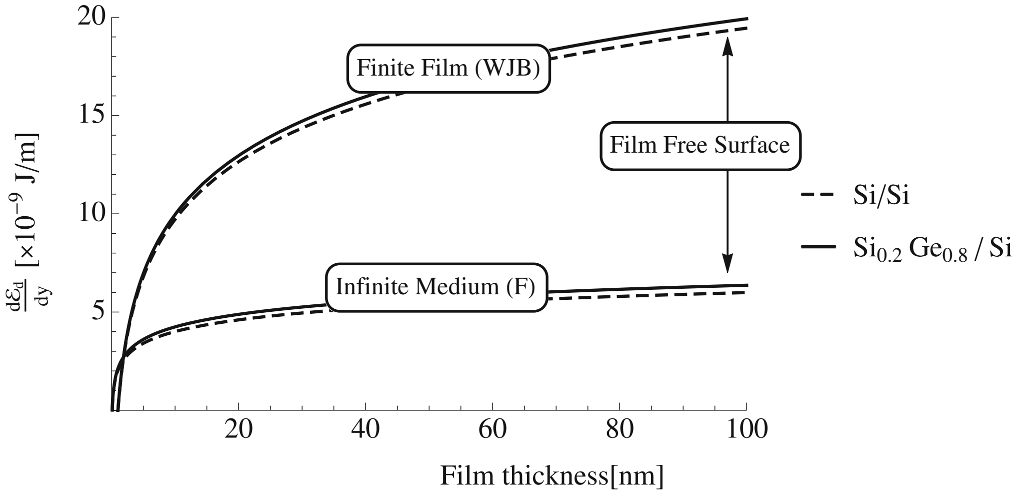


Figure 4.10: $d\mathcal{E}_d/dy$ is a function of the film thickness for $Si_{1-x}Ge_x/Si$. The $\langle 110\rangle\{111\}$ slip system of a 60° dislocation is considered. The two sets of curves are calculated through the Willis *et al.* (WJB) and Freund (F) procedures, respectively.

4.5 Overall comparison of various effects on the predicted critical thickness

In order to consider the anisotropy, the difference between the elastic properties of the film and the substrate as well as the impact of the free surface, the Steeds and the Willis *et al*. approaches are combined. Based on the conclusions of Section 4.3, in the following the contribution of the integral along the core surface S_3 is neglected. The detailed mathematical derivation of this treatment (indicated with Steeds+Willis *et al*.) is given in 4.2.3. The model Steeds+Willis *et al*. is used to compute $d\mathcal{E}_d/dy$ as a function of the film thickness for three different alloys: an $Al_{0.2}Ga_{0.8}N$ film on a GaN substrate, $In_{0.2}Ga_{0.8}N$ on GaN and a $Si_{0.8}Ge_{0.2}$ film on a Si substrate. The results are shown and compared with corresponding Freund, Steeds, and Willis *et al*. results in Figures 4.11, 4.12, and 4.13. The critical thickness values of the systems are listed in Table 4.2. Considering all the three systems, the Willis *et al*. model increases the critical thickness by ≈ 300% with respect to the Freund model. The critical thickness value of the theoretically most complete scheme, Steeds+Willis *et al*., is always between the Steeds and the Freund values. In particular, the Steeds+Willis *et al*. critical thickness is lower by ≈ 50% relative to the Freund critical thickness for $Al_{0.2}Ga_{0.8}N$ and $In_{0.2}Ga_{0.8}N$, while it is the same in the case of $Si_{0.8}Ge_{0.2}$.

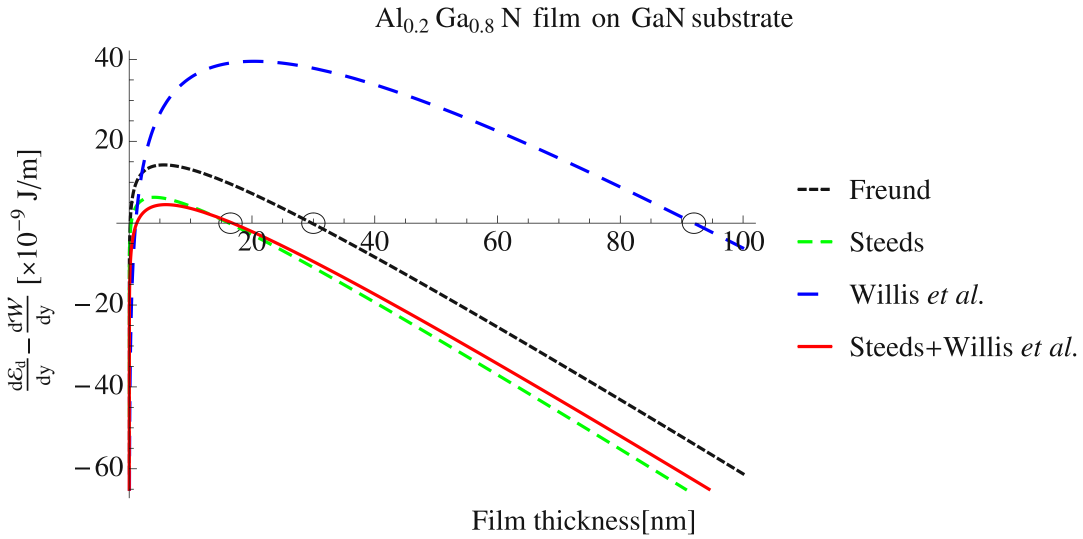


Figure 4.11: $d\mathcal{E}_d/dy - d\mathcal{W}/dy$ is a function of the film thickness. $d\mathcal{E}_d/dy$ is calculated according to the Freund, Steeds, Willis *et al*., and Steeds+Willis *et al*. approaches. $d\mathcal{W}/dy$ is calculated according to equation (4.9). The critical thickness values an $Al_{0.2}Ga_{0.8}N$ film grown on a GaN substrate are indicated by a circle.

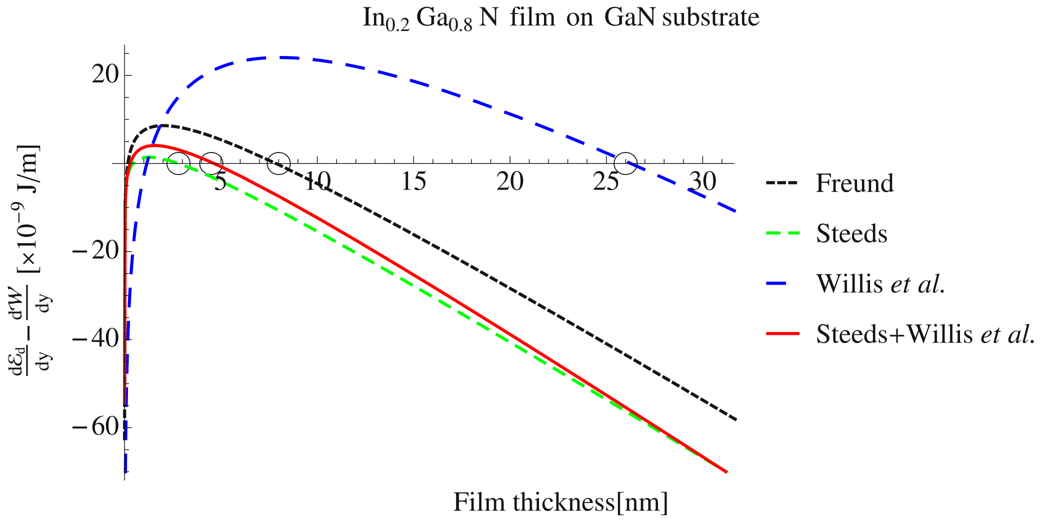


Figure 4.12: Same as in Figure 4.11 but for an $In_{0.2}Ga_{0.8}N$ film grown on a GaN substrate.

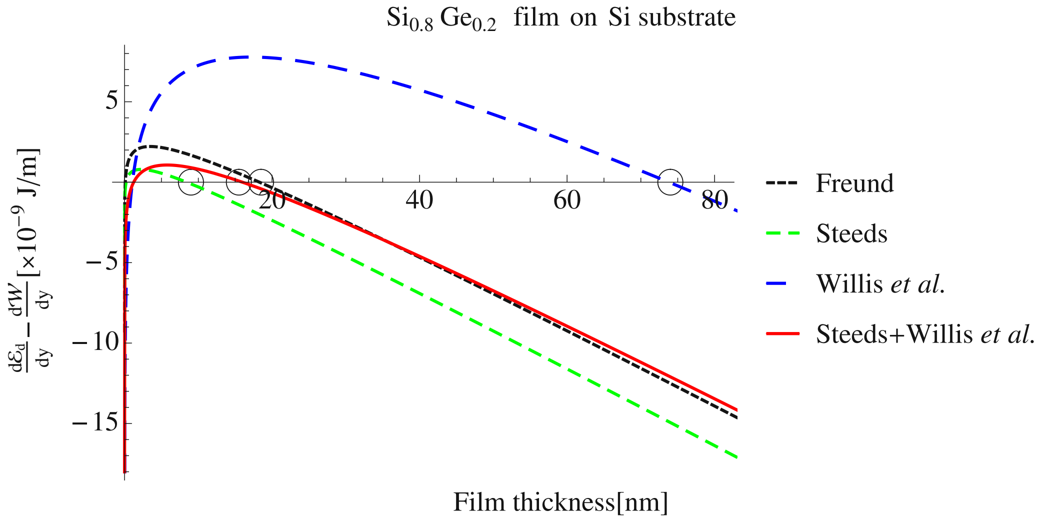


Figure 4.13: Same as in Figure 4.11 but for a $Si_{0.8}Ge_{0.2}$ film grown on a Si substrate

Table 4.2: Critical thickness values (in nm) of the studied systems.

	F	F+E_{cs}	S	S+E_{cs}	WJB	S+WJB
$Al_{0.2}Ga_{0.8}N$ film on GaN substrate	30	30	16	12	93	17
$In_{0.2}Ga_{0.8}N$ film on GaN substrate	8	8	3	0	26	5
$Si_{0.2}Ge_{0.8}$ film on Si substrate	18	18	8	6	74	16

4.6 Comparison between theoretical and experimental critical thicknesses

After illustrating the differences among various treatments of dislocation energy and their impact on the equilibrium critical thickness , the Freund and Steeds+Willis *et al*. models, including the dislocation core energies estimated in Section 3.6, are used to calculate the equilibrium critical thickness as a function of composition, x, for the three different alloys. The two models, Freund and Steeds+Willis *et al*., are chosen as they are based on opposite hypotheses (see Table 4.1).

The $1/3\langle 11\bar{2}3\rangle\{1\bar{1}01\}$ and $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ slip systems for $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$ film on GaN substrate and $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}$ film on GaN substrate, and the 60° misfit dislocation with the $\langle 110\rangle\{111\}$ slip system for $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ film on Si substrate are considered. The calculated critical thickness results obtained from the Freund and Steeds+Willis *et al*. models are compared with experimental observations and data available in literature (Figures 4.14, 4.15, and 4.16). The difference between the Freund and Steeds+Willis *et al*. results is small, with the Steeds+Willis *et al*. model yielding lower values as compared to F. In all cases the experimental data are close to the theoretical curves, suggesting that the experimental critical thickness values were obtained from epitaxial depositions close to the thermodynamical equilibrium.

Regarding the $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$/GaN and $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}$/GaN systems, the Steeds+Willis *et al*. model provides a more severe condition for the onset of the misfit dislocation at the interface than the Freund model. It is important to realize that the critical thickness reported here is the so-called *equilibrium* critical thickness . That means that it corresponds to the configuration where it for the first time becomes energetically favorable to relieve the misfit strain by introducing misfit dislocations. However, any mechanism for the creation of the misfit dislocations, which may require certain extra activation energy, is not considered in the model. Similarly, parameters influencing the kinetics of the epitaxial deposition, such as temperature and deposition rate, are also not considered by the equilibrium critical thickness models. Finally, the current experimental techniques are unable to detect the exact onset of the appearance of misfit dislocations. It therefore follows that no misfit dislocations are expected *below* the predicted critical thickness values, however, the detection of misfit dislocations may be (sometimes significantly) higher than the theoretical critical thickness.

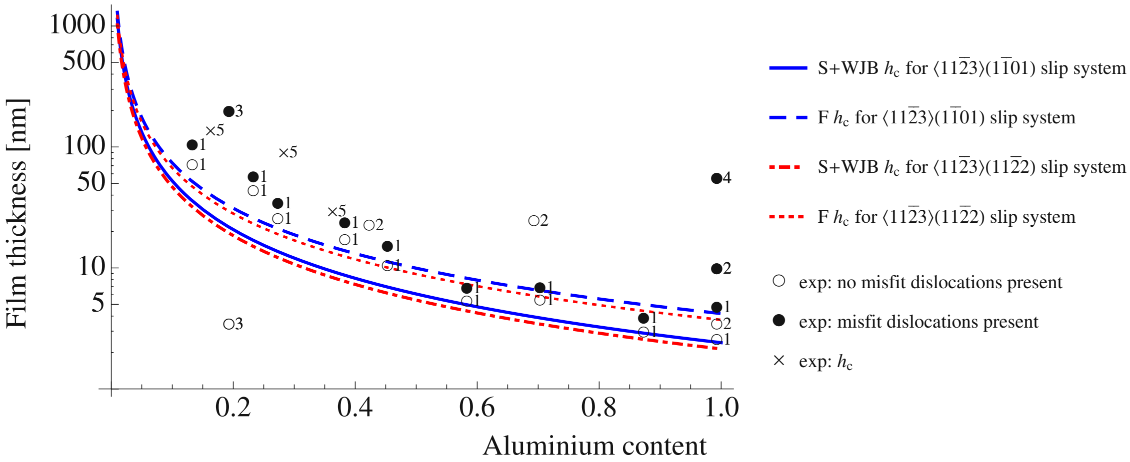


Figure 4.14: The equilibrium critical thickness h_c as a function of the AlN mole fraction x calculated through Freund and Steeds+Willis *et al.* (S+WJB) models including the core energy for two different slip systems. The theoretical curves are compared to experimental data, 1- [41], 2- [77], 3- [6], 4- [20], 5- [15]. Empty and filled circles indicate the absence and the presence of misfit dislocations respectively. Crosses indicate the experimental value of the critical thickness h_c.

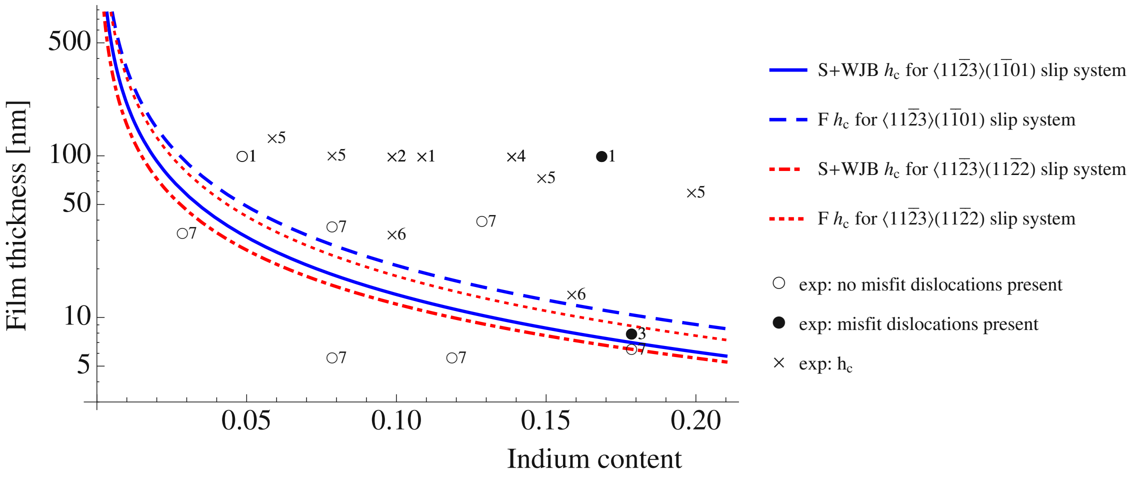


Figure 4.15: The equilibrium critical thickness h_c as a function of the InN fraction x calculated through Freund and Steeds+Willis *et al.* (S+WJB) models including the core energy for two different slip systems. The theoretical curves are compared to experimental data, 1 - [44], 2- [71], 3- [46], 4- [33], 5- [60], 6- [62], 7- [28]. Empty and filled circles indicate the absence and the presence of misfit dislocations respectively. Crosses indicate the experimental value of the critical thickness h_c.

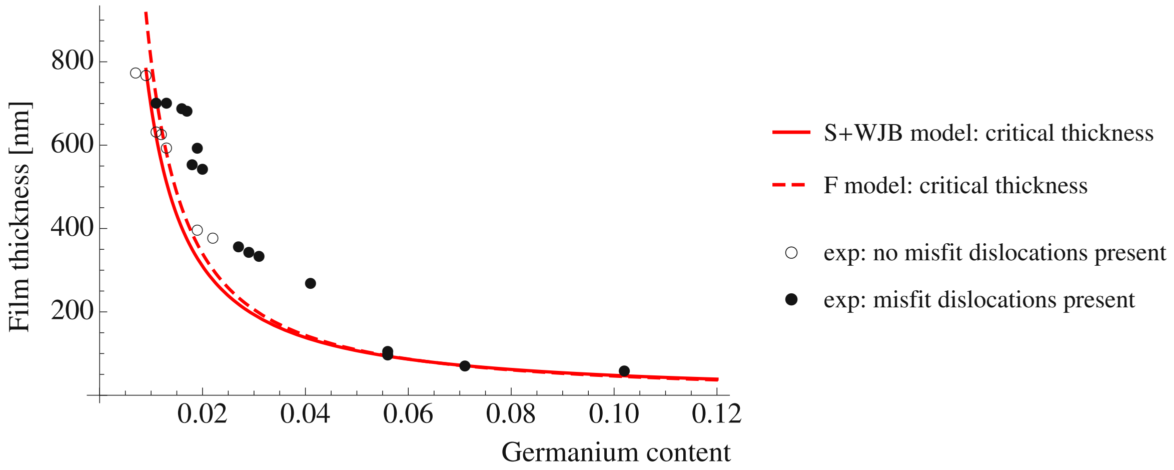


Figure 4.16: The equilibrium critical thickness h_c as a function of the Ge fraction x calculated with the Freund (F – dashed line) and Steeds+Willis *et al.* (S+WJB – solid line) models including the core energy. The 60° misfit dislocation with the $\langle 110 \rangle \{111\}$ slip system is considered. The theoretical curves are compared with experimental data from [30]. Empty and filled circles indicate the absence and the presence of misfit dislocations respectively.

A closer inspection of Figures 4.14-4.16 reveals that the here refined Steeds+Willis *et al.* model fulfills this criterion (unlike the Freund model). Indeed for all the studied systems, the misfit dislocations are always experimentally detected above the critical thickness values predicted by the model Steeds+Willis *et al.* unlike the values given by the Freund model.

4.7 Summary

In this chapter, the different continuum-based approaches for calculating the energy of a straight infinitely long dislocation in an elastic medium has been revisited. Motivated by the misfit dislocations in a heteroepitaxial interface, the influence of (i) the free surface, (ii) different elastic constants in the film and substrate, and (iii) elastic anisotropy have been evaluated separately. The results suggest that starting from a homogeneous infinite isotropic medium, the inclusion of a free surface increases the dislocation energy, and the difference in elastic constants of the film and substrate does not play any significant role (because it is typically an order of magnitude smaller than the impact of, e.g., the free surface), while the inclusion of elastic anisotropy decreases the dislocation energy.

Finally, the equilibrium critical thickness was calculated for three important heteroepitaxial material systems, namely an $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$ film on a GaN substrate, an $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}$ film on a GaN substrate, and a $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ film on a Si substrate. The models for the equilibrium critical thickness, described in Section 4.2, provide a condition when it first becomes energetically favorable to start the relaxation of the mismatch strain via plastic flow. The results suggest that the model [8, 24], reported in Paragraph 4.2.3, including the elastic anisotropy of the film and the substrate, the difference of their elastic constants and the impact of the film free surface, yields an excellent agreement with the available experimental data in the sense that no misfit dislocations are detected below the here predicted threshold.

5 Reduction of the threading dislocation density

5.1 Introduction

A serious issue for the reliability and performance of heterostructure devices is the deteriorating influence of dislocations, which originate during the crystal growth of the epitaxial layers composing the heterostructure. For a film grown by the Volmer-Weber mechanism (island growth mode), dislocations may result from surface half-loop generation or from island coalescence. These dislocations are called threading dislocations. The dislocation line can either be straight and perpendicular or inclined to the interface. As a consequence, heterostructures have lines which propagate through their thickness, reaching the active part of the device and degrading its performance. It is important to note that the threading dislocation density is not constant along the film thickness. Growth techniques and dislocation configuration can reduce the density of threading dislocations. In particular, the density is reduced by reactions among dislocations when they, by some process, come in contact with each other. The driving force for this process is the minimization of the internal energy of the system. Reactions among inclined threading dislocations in bulk GaN were modeled by Mathis *et al*. [49]. The model supposes the inclination of threading dislocations as the only possible source for their reactions. An application of this model to semipolar GaN is given in Paragraph 5.2.1. The limit of Mathis' model is that reactions among dislocations happen only if they are inclined with respect to the growth direction.

Real devices are based on heterostructures composed of many lattice mismatched layers. For such structures, regardless of the growth mechanisms – such as those by Frank-van der Merwe (layer-by-layer), Stranski-Krastanov (initial wetting followed by islanding), or Volmer-Weber (incoherent islanding) – increasing film thickness ultimately leads to the glide of the threading dislocations with the concomitant generation of so called misfit dislocations. In particular, below a certain layer thickness, called the *critical thickness* (CT), each film is grown pseudomorphically on the lower one, i.e., the film is grown with the same lattice parameter as the substrate. Consequently, the layer is strained, leading to large strain energy. When the critical thickness is reached, a relaxation of the strain occurs via plastic flow. The most common mechanism of plastic relaxation is the glide of the threading dislocations with the introduction of misfit dislocations along the interface between the two materials [17, 28, 82]. Gliding dislocations are called *glissile* dislocations, instead they are defined *sessile* when they can not glide below the critical thickness. Misfit dislocations change the stress gradient in the structure, causing its relaxation. The threading dislocation glide along the interface increases the probability of contact among dislocations. As a consequence, the probability of reactions increases in the presence of a glide, i.e., in presence of an interface between two mismatched materials. threading dislocation reactions and their density decrease due to dislocation glide along a bilayer interface, as modeled by Romanov *et al.* [66]. Romanov calculated the threading dislocation density in multilayered structures but his model differs from Mathis' because threading dislocation glide is the only possible source for threading dislocation reactions. Threading dislocation inclination with respect to the growth direction, which is the source of reactions in Mathis model, is not considered. In addition, the different threading dislocation types present in hexagonal symmetry are neglected in Romanov's work. Mathis and Romanov's models can be considered complementary to each other, since each is based on a hypothesis not consider by the other.

In 2014 Ward and co-authors [79] evaluated the threading dislocation density reduction as a function of the misfit stress associated with two mismatched layers and the threading dislocation inclination together, using simple numerical modeling. Their result was that the key to significantly reduce threading dislocation density is the movement of threading dislocations resulting from misfit stress. The most important parameter affecting threading dislocation reduction is the amount of strain relaxed by misfit dislocations in the structure. The simple numerical model developed by Ward and coauthors can roughy describe the behavior of dislocations on average, since the different threading dislocation types and their specific reactions are not considered.

Real heterostructures are instead composed of many lattice mismatched film materials, where the reactions of the several threading dislocation types are simultaneously affected by their inclination and their glide along the interfaces. All these phenomena are not evaluated together by any of the above mentioned models. In this chapter, this gap is closed by proposing a new description of threading dislocation density in GaN-based multilayers considering the different threading dislocation types with their specific reactions and inclinations. Further, to go beyond Mathis' efforts, the effect of the dislocation glide on the threading dislocation density is also evaluated. The new formulation also considers the impact of the hexagonal symmetry on threading dislocation glide, using the treatments proposed by Steeds [72] and Holec [24] (see Section 5.3). After the description of Mathis' model in Section 5.2, the new treatment is described in Section 5.4. It is then applied to a simple case, the AlN/GaN bilayer, in order to separately evaluate the impact of the model parameters (see Paragraph 5.4.5). After that, the model is used to calculate the threading dislocation density in an $Al_{1-x}Ga_xN$ step-graded layer and in an $(AlN/GaN)_{10}$ superlattice. The results are compared with experimental data from literature in Paragraph 5.4.6. General rules to decrease threading dislocation density in heterostructures are deduced and listed at the end of this chapter, including a proposal for further improvement of the model.

5.2 Mathis Model

Mathis [49] described the threading dislocation density in GaN bulk using a series of previous models developed by Romanov and co-workers in 1996 [64,65]. These models are based on reactions of inclined dislocations which change the threading dislocation density. The reactions among threading dislocations are caused by the minimization of the internal energy of the system. The energy per unit length of a dislocation with the Burgers vector $\mathbf{b}$ is within the isotropic approximation given by

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \frac{\mu\left|\mathbf{b}\right|^2(1-\nu\cos^2\theta)}{4\pi(1-\nu)}\ln\left(\frac{R}{r_\mathrm{c}}\right). \tag{5.1}$$

Since the energy of a dislocation is proportional to b^2, the above expression can be written as

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{d}}{\mathrm{d}y} = \mathrm{const}\cdot\left|\mathbf{b}\right|^2. \tag{5.2}$$

If two dislocations with Burgers vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ become contiguous by some process, a reaction between them happens only if the reaction product is a new dislocation with Burgers vector $\mathbf{b}_3$ whose energy is lower than the sum of the previous dislocation energies (see Figure 5.3), i.e., in terms of the respective Burgers vector:

$$\left|\mathbf{b}_1\right|^2 + \left|\mathbf{b}_2\right|^2 > \left|\mathbf{b}_3\right|^2 \qquad \text{(Frank's rule)}\ . \tag{5.3}$$

The above formula is called *Frank's rule* [55]. Different reactions occur dependent upon the dislocation geometry. If $\mathbf{b}_1 = -\mathbf{b}_2$, i.e., two dislocations with opposite Burgers vectors meet, they *annihilate*. Other possible reactions are *fusion* ($\mathbf{b}_1+\mathbf{b}_2 \rightarrow \mathbf{b}_3$) and *scattering* ($\mathbf{b}_1+\mathbf{b}_2 \rightarrow \mathbf{b}_3+\mathbf{b}_4$, Frank's rule in this case is analogous to 5.3), or no reaction at all.

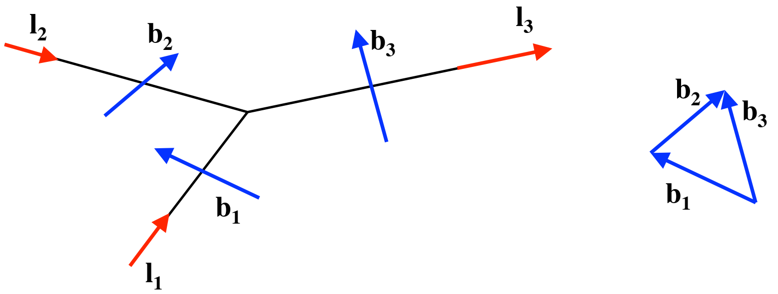


Figure 5.1: An example of a dislocation reaction (fusion) $\mathbf{b}_1+\mathbf{b}_2 \rightarrow \mathbf{b}_3$. Vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are the respective dislocation line directions.

Twenty different types of threading dislocations exist in a crystal with hexagonal symmetry (see Figure 5.2): two different types of screw dislocations with $\mathbf{b} = \pm\mathbf{c}$, six different types of edge dislocations with the Burgers vectors $\mathbf{b} = \pm\mathbf{a}_i$ $(i = 1, 2, 3)$, and 12 different mixed threading dislocations with $\mathbf{b} = \pm\mathbf{c} \pm \mathbf{a_i}$ $(i = 1, 2, 3)$. Table 5.2 shows the products of reaction among these threading dislocations according to Frank's rule.

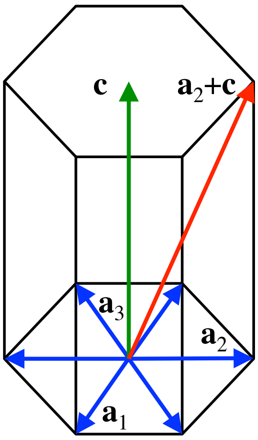


Figure 5.2: Hexagonal lattice. Burgers vectors of the type $1/3\langle 11\bar{2}0\rangle$ are blue. Burgers vector of the type $\langle 0001\rangle$ is green. Burgers vector of the type $1/3\langle 11\bar{2}3\rangle$ is red.

In Romanov's [64, 65] and Mathis' [49] models, threading dislocations do not move physically (i.e. do not glide or climb), instead movements result from growth. Suppose the situation sketched in Figure 5.3.

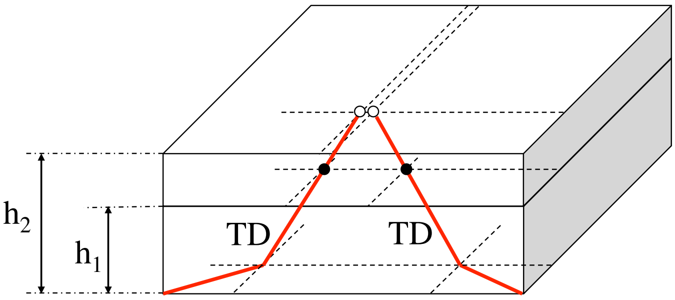


Figure 5.3: Relative motion of inclined TDs as a result of film growth.

The threading dislocation segments are inclined with respect to the growth direction and as a consequence the terminating points at the upper surface move laterally and can eventually come in contact with each other. The threading dislocation intersections with the free surface will move until the point where they come within a distance r such that the interaction forces are sufficient to initiate an additional motion of dislocations. At this point they start to react among each other. The growth surface is assumed to be planar, which corresponds to a Frank-van der Merwe's layer-by-layer growth mode. Mathis considered threading dislocations in bulk GaN, supposing the dislocations are lying on planes which are inclined with respect to the [0001] growth direction and thus during the growth may become close enough to react. The inclination of (some) dislocations was obtained by considering the dislocation energy within the isotropic theory. In Chapter 3, the optimal dislocation line direction for c- and a-type dislocations in elastically isotropic GaN is vertical whereas for $(a+c)$-type dislocations it is approximately 14.5° inclined from the vertical direction, which is similar to Mathis' result of 15.6°. Mathis also presented TEM evidence for some dislocations being inclined and claimed agreement with the above mentioned inclination angle. Section 3 shows that within the anisotropic theory all dislocations are parallel to the [0001] direction. In an absence of factors which can incline the dislocations, such as the island growth mode or impurities, the vertical threading dislocations are not able to come close enough to react, so the model can not be applied. A simple application of Mathis' model is given in the next Paragraph, where the threading dislocation density is calculated in a GaN film and where the growth direction is the $[11\bar{2}2]$ direction (see Figure 5.4). In this situation, Mathis' treatment is used to calculate the threading dislocation density because all threading dislocation types are inclined with respect to the $[11\bar{2}2]$ direction, as shown in Table 3.6. The discussion of the results will provide a first general idea about which model parameters affect the threading dislocation density more.

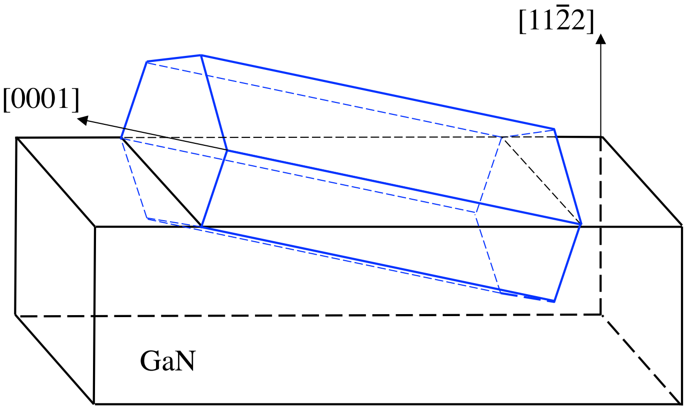


Figure 5.4: Perspective view schematic of a semipolar GaN film with growth direction $[11\bar{2}2]$.

5.2.1 Application to semipolar GaN

Chapter 3 shows that in GaN all dislocations close to the $(11\bar{2}2)$ facet are inclined with respect to the [0001] direction. The related angles α of inclination are listed in Table 3.6. Mathis' model is used to calculate the threading dislocation density in semipolar GaN since all dislocations are inclined with respect to the $[11\bar{2}2]$ growth

Table 5.1: Reaction table [49] for threading dislocations (TDs) in GaN with hexagonal symmetry. Reactions which are not possible either due to the Frank's criterion or for geometric reasons are indicated with a "—" while reactions which are possible are indicated with their designated number. Reactions producing two dislocations are denoted with two numbers corresponding to their products. Annihilation reactions are denoted by "A".

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		a_1	$-a_1$	a_2	$-a_2$	a_3	$-a_3$	a_1 $+c$	a_1 $-c$	$-a_1$ $+c$	$-a_1$ $-c$	a_2 $+c$	a_2 -c	$-a_2$ $+c$	$-a_2$ $-c$	a_3 $+c$	a_3 $-c$	$-a_3$ $+c$	$-a_3$ $-c$	c	$-c$
1	a_1	–	—	—	—	—	—	—	—	19	20	17	18	—	—	13	14	—	—	—	—
2	$-a_1$	—	—	—	—	—	—	19	20	—	—	—	—	15	16	—	—	11	12	—	—
3	a_2	—	—	—	—	—	—	17	18	—	—	—	—	19	20	9	10	—	—	—	—
4	$-a_2$	—	—	—	—	—	—	—	—	15	16	19	20	—	—	—	—	7	8	—	—
5	a_3	—	—	—	—	—	—	13	14	—	—	9	10	—	—	—	—	19	20	—	—
6	$-a_3$	—	—	—	—	—	—	—	—	11	12	—	—	7	8	19	20	—	—	—	—
7	a_1+c	—	19	17	—	13	—	—	1&1	—	A	—	6	—	1&4	—	4	—	1&6	—	—
8	a_1-c	—	20	18	—	14	—	1&1	—	A	—	6	—	1&4	—	4	—	1&6	—	1	—
9	$-a_1+c$	19	—	—	15	—	11	—	A	—	2&2	—	2&3	—	5	—	2&5	—	3	—	2
10	$-a_1-c$	20	—	—	16	—	12	A	—	2&2	—	2&3	—	5	—	2&5	—	3	—	2	—
11	a_2+c	17	—	—	19	9	—	—	6	—	2&3	—	3&3	—	A	—	2	—	3&6	—	3
12	a_2-c	18	—	—	20	10	—	6	—	2&3	—	3&3	—	A	—	2	—	3&6	—	3	—
13	$-a_2+c$	—	15	19	—	—	7	—	1&4	—	5	—	A	—	4&4	—	4&5	—	1	—	4
14	$-a_2-c$	—	16	20	—	—	8	1&4	—	5	—	A	—	4&4	—	4&5	—	1	—	4	—
15	a_3+c	13	—	9	—	—	19	—	4	—	2&5	—	2	—	4&5	—	5&5	—	A	—	5
16	a_3-c	14	—	10	—	—	20	4	—	2&5	—	2	—	4&5	—	5&5	—	A	—	5	—
17	$-a_3+c$	—	11	—	7	19	—	—	1&6	—	3	—	3&6	—	1	—	A	—	6&6	—	6
18	$-a_3-c$	—	12	—	8	20	—	1&6	—	3	—	3&6	—	1	—	A	—	6&6	—	6	—
19	c	—	—	—	—	—	—	—	1	—	2	—	3	—	4	—	5	—	6	—	—
20	$-c$	—	—	—	—	—	—	1	—	2	—	3	—	4	—	5	—	6	—	—	—

direction. The inclination angle with respect to the $[11\bar{2}2]$ direction is denoted γ in Figures 5.5.

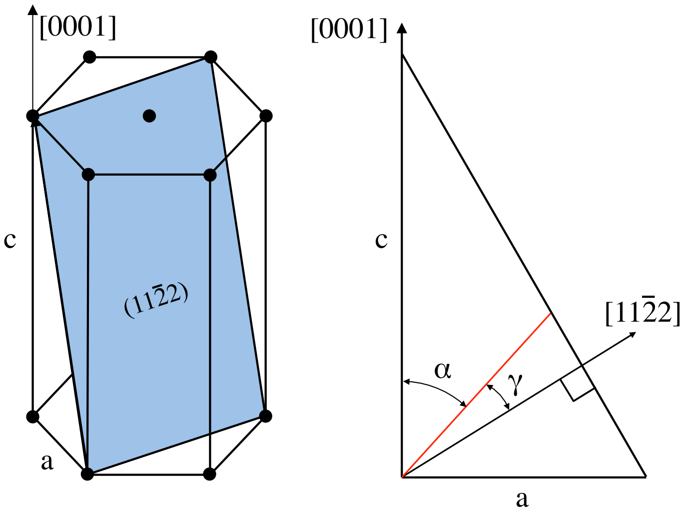


Figure 5.5: Geometry of the hexagonal lattice with the $(11\bar{2}2)$ plane indicated in blue.

The angle γ for the three threading dislocation types are calculated using the inclination angle θ listed in Table 5.2 and the GaN lattice parameters a and c.

Table 5.2: Inclination angle γ of threading dislocations (TDs) with respect to the $[11\bar{2}2]$ direction.

TD type	α	γ
c	23°	35°
a	64°	6°
$(a+c)$	34°	24°

Using these values within Mathis' model allows for the calculation of the threading dislocation density in semipolar GaN. The equations of Mathis' model are solved supposing that all types of dislocations can react with each other and be produced by reactions between other dislocation pairs. Eight different initial conditions are defined changing the proportion among the dislocation families (see Table 5.3).

Table 5.3: Ratios of the three threading dislocation (TD) types (in %) used as initial conditions.

	c-type TD	a-type TD	(a+c)-type TD
1	0	70	30
2	0	30	70
3	10	70	20
4	10	20	70
5	20	70	10
6	20	10	70
7	30	70	0
8	30	0	70

The interaction radius is set to 40 nm [79]. The initial value for threading dislocation density is set to 10^{10}cm^{-2}, which is a typical value for III-nitride grown upon non native substrates [38, 56, 81].

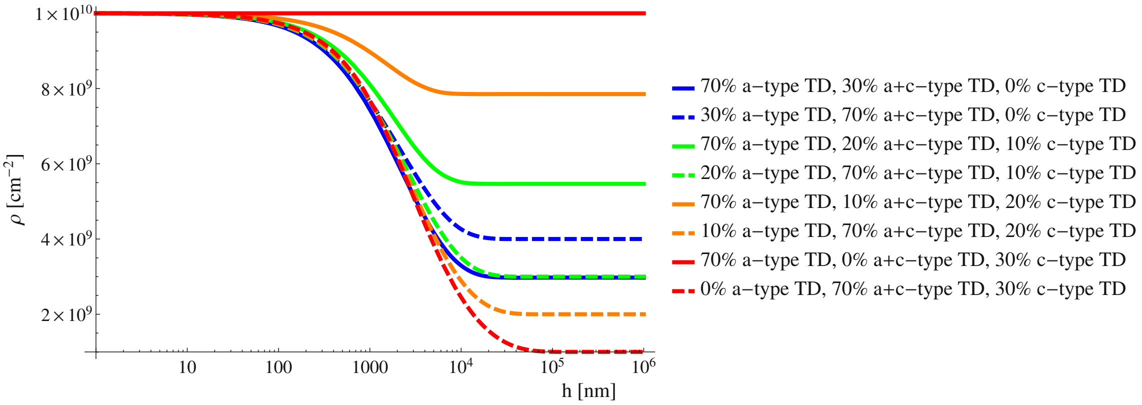


Figure 5.6: The total threading dislocation density, called ρ, as a function of the GaN thickness h for different initial conditions.

Figure 5.6 shows the total threading dislocation density (i.e., the sum of the densities of the twenty dislocation families) as a function of the film thickness for the different initial conditions listed in Table 5.3. For all cases, the general behavior can be divided into three stages. The first stage covers the initial 100nm of the film thickness, where the density is almost constant, i.e., the number of occurring reactions is small because the inclination angle is small. In the second stage, between 100nm and 10μm, the density decreases from the initial value until a saturation point. The reduction is caused by reactions among dislocations. After 10μm, the density becomes constant, meaning that reactions are no longer taking place. The saturation of the threading dislocation density was observed experimentally by several authors [3, 34, 47]. Consider now the different curves of Figure 5.6. Keeping constant the starting value of c-type threading dislocations, an initial amount of $(a+c)$-type threading dislocations, which is bigger than a-type threading dislocations, implies a lower final asymptotical value of the density. This is explained by the fact that $(a+c)$-type threading dislocations have an inclination angle γ bigger than the corresponding one for a-type threading dislocations. Despite the influence of the inclination angle, it is not always straightforward to predict the results only based on considerations of the inclination angle.

For example, increasing the initial amount of c-type threading dislocations when the initial amount of a-type threading dislocations is bigger than the initial amount of $(a+c)$-type threading dislocations, causes the final threading dislocation density to increase. This is explained considering the particular case where there are no initial $(a+c)$-type threading dislocations (see the red solid line in Figure 5.6). Here, the threading dislocation density does not change from the initial value because no reactions can take place among a-type threading dislocations themselves, among c-type threading dislocations themselves, and between a-type threading dislocations and c-type threading dislocations, as shown by Table 5.2. Instead, increasing the initial amount of c-type threading dislocations when the initial amount of a-type threading dislocations is lower than the initial amount of $(a+c)$-type threading dislocations, causes the final value of threading dislocation density to decrease. In all cases, after 10μm, the overall decrease from the initial density is less than one order of magnitude, independently of the initial conditions. This decrease is not enough to improve the performance of real devices. Therefore, a better method to reduce the threading dislocation density is the use of multilayered structures. In this situation, Mathis' model [49] is no longer sufficient to calculate the threading dislocation density due to the presence of interfaces among lattice-mismatched films. Under such circumstances, a threading dislocation glide along the interface must be considered as an additional source of reactions. In the next Paragraph, a threading dislocation glide along the interface is evaluated together with the evolution of the misfit dislocation density. The model of misfit dislocation density will be incorporated into an improved treatment capable of evaluating the threading dislocation density in heterostructures (see Section 5.4).

5.3 Misfit dislocation density

After the critical thickness, threading dislocations glide along the interface of a lattice mismatched bilayer with the concomitant formation of misfit segments. The misfit segment is called misfit dislocation because its formation is caused by the relaxation of the misfit strain associated with the lattice mismatch. At the critical thickness, the formation of misfit dislocations is energetically favorable as they relax the internal energy. Increasing the film thickness after the critical value, the evolution of the misfit dislocation density was modeled by Hu [31]. The model is based on balancing the elastic energy from a partially relaxed misfit strain and the energy of the non-interacting misfit dislocation arrays. The main hypothesis is that all misfit dislocations appear suddenly when the film thickness reaches the critical value. There is no reason to assume that this happens but this assumption is retained to keep the equations simple, thus avoiding the inclusion of any kinetic parameter into the treatment. Hu's expression [31] for the misfit dislocation density ρ_{MD} after the critical thickness reads as follows:

$$\rho_{\mathrm{MD}} = \frac{\epsilon_m}{b \sin\theta \sin\phi}\left(1 - \frac{h_{\mathrm{c}}}{h}\right). \tag{5.4}$$

The elongation of the misfit dislocations causes the relaxation of the initial misfit strain ϵ_m. Therefore the partially relaxed strain ϵ is [66]

$$|\epsilon| = ||\epsilon_m| - \rho\, b \sin\theta \sin\phi|, \tag{5.5}$$

where θ is the angle between the dislocation line and the Burgers vector, and ϕ is the angle between the slip plane and the interface. The expression has been deduced within the elastically isotropic framework.

A similar formula of misfit dislocation density was derived by Holec [24] for the hexagonal symmetry:

$$\rho_{\mathrm{MD}} = \frac{2}{3}\frac{\epsilon_m}{b \sin\theta \sin\phi}\left(1 - \frac{h_{\mathrm{c}}}{h}\frac{\ln\frac{h}{r_c}}{\ln\frac{h_{\mathrm{c}}}{r_c}}\right). \tag{5.6}$$

The relaxation of the misfit strain as a consequence of misfit dislocation elongation is calculated as [24]

$$|\epsilon| = ||\epsilon_m| - |\tfrac{3}{2}\rho\, b \sin\theta \sin\phi||. \tag{5.7}$$

The models of both Hu [31] and Holec [24] are used to calculate the misfit dislocation density and the strain relaxation after the critical thickness for a GaN film grown upon an AlN substrate. The critical thickness is calculated according to Steeds [72] when Holec's model [24] is used. Otherwise, when Hu's model [31] is used, the critical thickness is calculated following the Freund procedure [17].

Figure 5.7 shows that the dislocation density approaches the limit for an infinitely thick layer quite rapidly, indicated by the dashed lines. The figure also shows the difference between the two models. The misfit dislocation density increases faster when Hu's model [31] is used. The difference between the final values of the density after the saturation point is around 25%.

Regarding the stress relaxation, Hu's model [31] predicts a slightly lower residual strain compared to Holec's model [24], but the difference between the two is negligible. The two models have the same asymptotical value corresponding to the same in-plane strain, i.e., the complete relaxation.

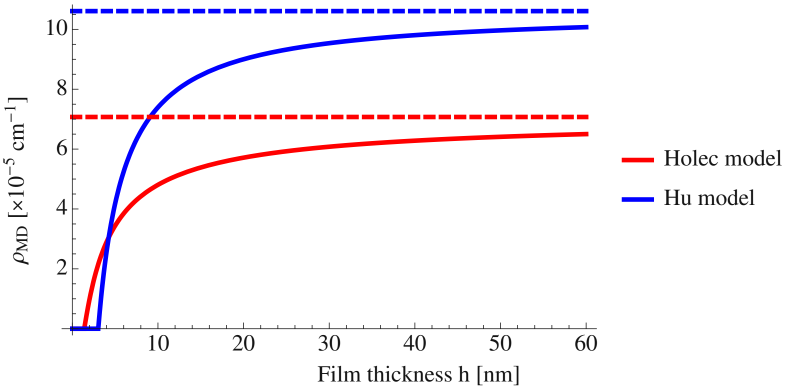


Figure 5.7: The misfit dislocation density ρ_{MD}, calculated within the isotropic and anisotropic frameworks, as function of the film thickness h.

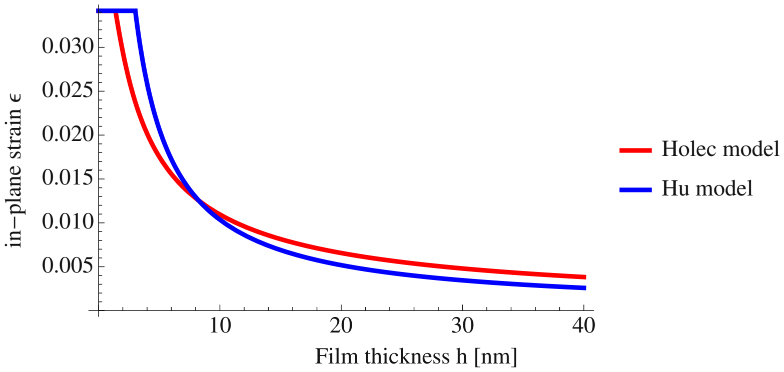


Figure 5.8: The in-plane strain ε, calculated within the isotropic and anisotropic frameworks, as function of the film thickness h.

5.3.1 Application to AlGaN film on AlN substrate

Holec model [24] is used to calculate the misfit dislocation density and the strain relaxation in $\mathrm{Al}_{1-x}\mathrm{Ga}_x\mathrm{N}$ film grown upon an AlN substrate for x=25%, 50%, 75%, and 100%. The critical thickness is evaluated according to Steeds [72]. Figure 5.9 shows that in increasing the lattice mismatch between film and substrate, the misfit dislocation density increases as well as the corresponding asymptotical value. In particular, in increasing the lattice mismatch, the increase of the misfit dislocation density is faster in the early stage of the growth and then slows down.

Figure 5.9 shows also that the saturation of the misfit dislocation density is almost reached at 50nm. Temperature can delay the development of the misfit dislocation density. Kinetic parameters are not considered in this work but the delay of the misfit dislocation elongation is negligible at typical processing temperatures of III-nitrides, which is around 1000°C [11]. A relaxation of the misfit stress results from

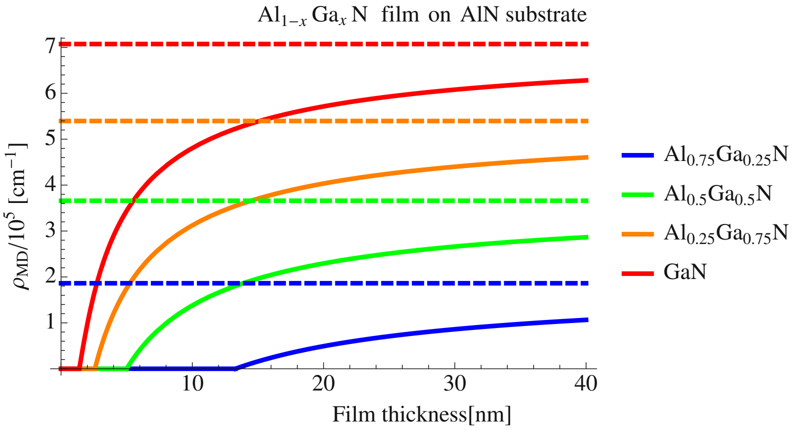


Figure 5.9: The misfit dislocation density ρ_{MD}, calculated within the anisotropic framework, as a function of the film thickness h for different chemical compositions of the film alloy.

the elongation of the misfit dislocations. The stress relaxation has been evaluated for an Al$_{1-x}$Ga$_x$N film grown upon an AlN substrate for x=25%, 50%, 75%, and 100%. The results are shown in Figure 5.10, where the strain decrease is a function of the film thickness. In particular, the decrease is faster for a high lattice mismatched film than for low ones. Another observation of Figure 5.10 is that the curves overlap after a certain point and they have the same asymptotical value which corresponds to the complete relaxation of the structure. It is important to observe that the in-plane strain is still not relaxed after 40nm. In particular, it is still 1/2 and 1/6 of the initial misfit strain for Al$_{0.75}$Ga$_{0.25}$N and GaN films respectively.

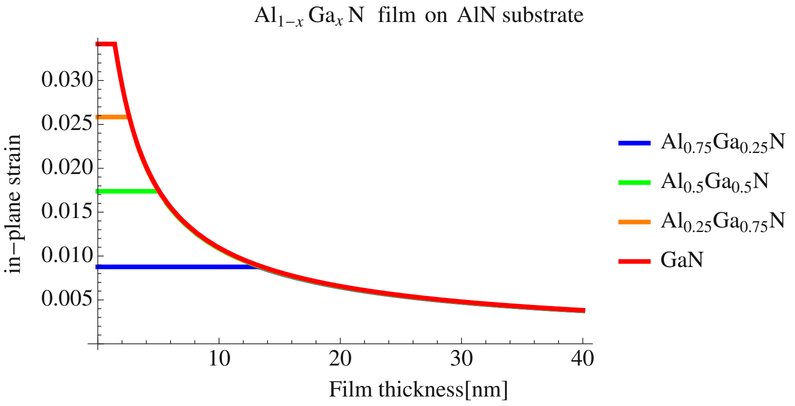
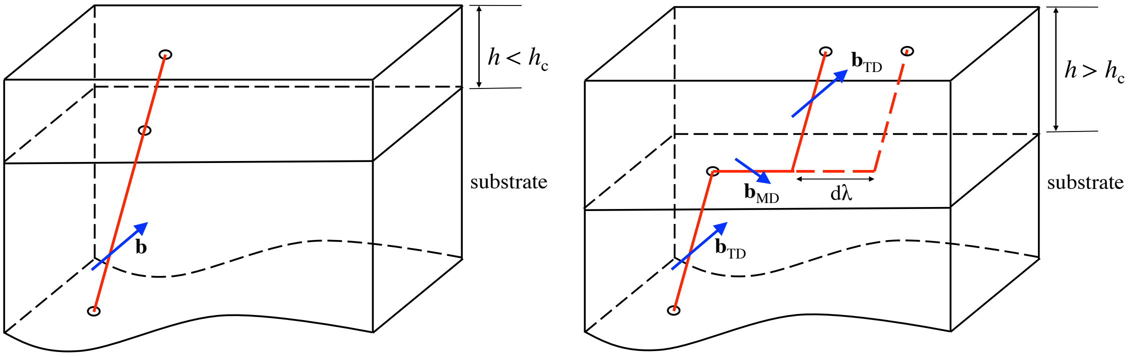


Figure 5.10: The in-plane strain as a function of the film thickness h for different initial conditions.

5.4 Threading dislocation density in heterostructures

In this section a model for the evolution of threading dislocations in GaN-based multilayers is derived by integrating Mathis' model [49] with the model of Romanov [66]. The impact of the hexagonal symmetry is considered with an appropriate expression for the misfit dislocation density [24, 31]. In the first Paragraph 5.4.1 of this section, appropriate equations describing the motion of one isolated dislocation are derived. Then in Paragraph 5.4.2, the *reaction-kinetic* approach introduced by Romanov [64] is used to quantify the probability that two threading dislocations come in contact with each other. A model describing the threading dislocation density with increasing film thickness in multilayers is described through a set of differential equations (see Paragraph 5.4.3). These are then applied to a bilayer (Paragraph 5.4.5) in order to evaluate which parameters have the strongest impact on the results. After this, the threading dislocation density is calculated in a step-graded layer and superlattice to understand which structure is better for minimizing the threading dislocation population (Paragraph 5.4.6). Concluding the section, various design rules for dislocation filter are deduced and listed.



(a) (b)

Figure 5.11: Basic processes of threading dislocation motion in a strained epitaxial film. 5.11(a) An isolated threading dislocation for $h < h_c$ for which no motion is possible. 5.11(b) Threading dislocation – misfit dislocation system for $h > h_c$: increasing film thickness leads to an increasing configurational force on the threading dislocation which leads to threading dislocation motion and generation of additional misfit dislocation segment length.

5.4.1 Motion of an isolated TD

An isolated threading dislocation is depicted in Figure 5.12.

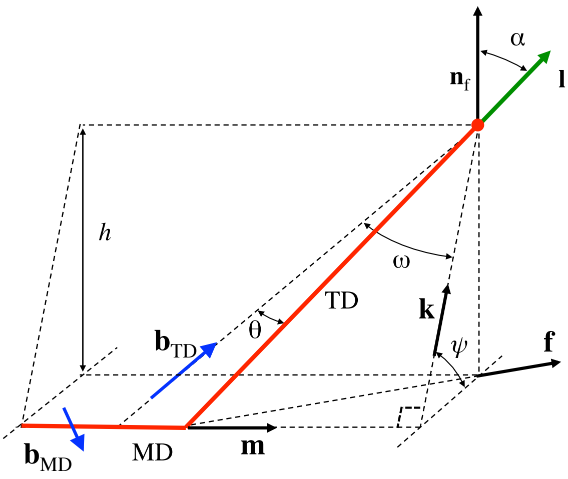


Figure 5.12: Geometry of an isolated dislocation composed of a threading arm (TD) and a misfit segment (MD).

An uniformly strained layer of thickness h_f is epitaxially bonded to a substrate of relatively larger thickness. The normal to the film-substrate interface is denoted by $\mathbf{n}_\mathrm{f}$. The equibiaxial mismatch strain until critical thickness is denoted by ϵ_m. A threading dislocation is assumed to exist along a particular glide plane in the presence of the background mismatch strain field. The threading dislocation line extends from the free surface of the film to the substrate. The tangent vector of the threading dislocation is denoted $\mathbf{l}$. During the film growth, it is assumed that the threading dislocation geometry remains self-similar. This means that the threading dislocation remains in the position of minimum energy with constant line direction $\mathbf{l}$. The Burgers vector of the threading dislocation is denoted by $\mathbf{b}$. The slip plane of the dislocation (the plane containing the dislocation line and the Burgers vector) is defined by its normal $\mathbf{n}_\mathrm{g}$. The angle α between the dislocation line $\mathbf{l}$ and the normal $\mathbf{n}_\mathrm{f}$ of the free surface is calculated according to the minimization of the dislocation energy in Chapter 3. Therefore $\mathbf{l}$ can be derived from the relation:

$$\cos\alpha = \mathbf{l}\cdot\mathbf{n}_\mathrm{f}. \tag{5.8}$$

ψ is defined as the angle between $\mathbf{k}$ and the projection of $\mathbf{k}$ to the film/substrate interface [65]:

$$\psi = \cos^{-1}\left(\mathbf{n}_\mathrm{g}\cdot\mathbf{n}_\mathrm{f}\right), \tag{5.9}$$

where $\mathbf{k}$ is defined as [65]

$$\mathbf{k} = \frac{\mathbf{n}_\mathrm{g}\times\left(\mathbf{n}_\mathrm{f}\times\mathbf{n}_\mathrm{g}\right)}{\sin\psi}. \tag{5.10}$$

The angle ω is defined as [65]

$$\omega = \cos^{-1}\left(\frac{\mathbf{b}\cdot\mathbf{n}_\mathrm{f}}{|\mathbf{b}|\sin\psi}\right). \tag{5.11}$$

After the vector $\mathbf{k}$ and the angle ω are calculated using the previous relations, the angle between $\mathbf{b}$ and the threading dislocation line $\mathbf{l}$ is denoted as θ, and can derived by [65]

$$\mathbf{l} = \frac{\sin(\omega-\theta)}{\sin\omega}\frac{\mathbf{b}}{|\mathbf{b}|} + \frac{\sin\theta}{\sin\omega}\mathbf{k}. \tag{5.12}$$

During film growth, if film thickness is below the critical value, the threading dislocation is sessile, i.e., it is not able to glide. As a consequence, the threading dislocation does not move physically but the threading dislocation upper point moves along the free surface due to the inclination of the threading dislocation with respect to the growth direction. According to Figure 5.12, the point of intersection of the threading dislocation with the free surface of the film moves with direction $\mathbf{f}$ along the free surface. The velocity $\mathbf{v}_\mathrm{S}$ of the upper point of one sessile threading dislocation along the free surface is calculated as [65]:

$$\mathbf{v}_\mathrm{S} = \tan\alpha\frac{\mathrm{d}h}{\mathrm{d}t}\,\mathbf{f}. \tag{5.13}$$

As the film thickness increases after the critical thickness, the relaxation of the misfit stress occurs via the glide of the threading dislocation, which leaves behind one misfit segment along the interface. This misfit segment is generally called misfit dislocation and its length is denoted λ in this work. The misfit dislocation elongates until the complete relaxation of the misfit strain and the direction $\mathbf{m}$ of the elongation is the direction of the gliding threading dislocation [65]:

$$\mathbf{m} = \frac{\mathbf{l}-(\mathbf{l}\cdot\mathbf{n}_\mathrm{f})\,\mathbf{n}_\mathrm{f}}{|\mathbf{l}-(\mathbf{l}\cdot\mathbf{n}_\mathrm{f})\,\mathbf{n}_\mathrm{f}|}\;. \tag{5.14}$$

In this work, the portion of the dislocation spanning the film thickness is still called the threading dislocation segment, or simply the threading dislocation. During the film growth after the critical thickness, λ increases as a result of relaxation of the initial misfit strain ε_m, causing a further glide of the threading segment. Concurrently, the threading dislocation length increases due to the increase of the film thickness. Thus after the critical thickness is reached, the threading dislocation intersection with the free surface will move not only due to the (eventual) inclination with respect to the free surface but also as a result of the threading dislocation glide, i.e., the misfit dislocation elongation along the interface. In other words, the upper point of the threading dislocation moves due to the inclination of the threading dislocation with respect to the free surface but also moves physically as a result of the threading dislocation glide along the interface. Therefore, the upper point of a glissile threading dislocation (i.e., a dislocation able to glide) has a velocity defined of

$$\mathbf{v}_\mathrm{G} = \tan\alpha\frac{\mathrm{d}h}{\mathrm{d}t}\,\mathbf{f} + \frac{\mathrm{d}\lambda}{\mathrm{d}t}\,\mathbf{m}, \tag{5.15}$$

where the first term takes the motion due to the inclination with respect to the film free surface into account, and the second term the motion due to the elongation of the misfit segment.

The elongation of the misfit segment is related to the relaxation of the structure caused by the film growth after the critical value. In order to establish a relation between the misfit dislocation elongation $\mathrm{d}\lambda$ and the increase of the film thickness $\mathrm{d}h$, it is necessary to consider the expression for the misfit dislocation density (denoted as ρ_{MD}) reported in Section 5.3.

For the isotropic system, the differentiated form of expression (5.4) with respect to h is

$$\mathrm{d}\rho_{\mathrm{MD}} = \frac{\epsilon_m h_\mathrm{c} \csc\theta \csc\phi}{b} \frac{1}{h^2} \mathrm{d}h = \mathrm{C}_{\mathrm{iso}} \frac{\mathrm{d}h}{h^2}. \tag{5.16}$$

For the hexagonal system, the differentiated form of expression (5.6) with respect to h is

$$\mathrm{d}\rho_{\mathrm{MD}} = \frac{2\epsilon_m h_\mathrm{c} \csc\theta \csc\phi}{3b \ln \frac{h_\mathrm{c}}{r_c}} \frac{\ln \frac{h_\mathrm{c}}{r_c} - 1}{h^2} \mathrm{d}h = \mathrm{C}_{\mathrm{hex}} \frac{\mathrm{d}h}{h^2}. \tag{5.17}$$

Misfit dislocations are generated through the motion of glissile threading dislocations, so it is possible to give a second definition of ρ_{MD} as a function of the total density ρ_{G} of the glissile threading dislocations:

$$\mathrm{d}\rho_{\mathrm{MD}} = \tfrac{1}{3} \rho_{\mathrm{G}} \mathrm{d}\lambda, \tag{5.18}$$

where the factor 1/3 accounts for the use of a linear misfit dislocation density in a hexagonal grid array of misfit dislocations. When combining expression (5.16) or expression (5.17) with expression (5.18), and solving for $\mathrm{d}\lambda$, the result is that $\mathrm{d}\lambda = \mathrm{C}\mathrm{d}h/(h^2\rho_{\mathrm{G}})$, with $\mathrm{C} = 3\mathrm{C}_{\mathrm{iso}}$ or $\mathrm{C} = 3\mathrm{C}_{\mathrm{hex}}$. Substituting the last relation in expression (5.15), the velocity of the glissile threading dislocation is expressed as

$$\mathbf{v}_{\mathrm{G}} = \left(\tan\alpha \, \mathbf{f} + \frac{\mathrm{C}}{h^2 \rho_{\mathrm{G}}} \, \mathbf{m} \right) \frac{\mathrm{d}h}{\mathrm{d}t}. \tag{5.19}$$

5.4.2 Reactions among dislocations

Two threading dislocations with different Burgers vectors $\mathbf{b}_i$ and $\mathbf{b}_j$ and line directions $\mathbf{l}_i$ and $\mathbf{l}_j$ are considered. The two threading dislocations can be either glissile or sessile. Supposing that the film thickness is increasing, the upper points of the two threading dislocations move along the film free surface until the point they come within a distance r such that the interaction forces are sufficient to initiate additional motion of dislocations. They can now start to react with each other. The driving force for the interaction between dislocations is the minimization of the internal energy. Three possible reactions (annihilation, fusion, and scattering) can occur depending on Frank's criterion (see equation (5.3)). The interaction radius r within which all these reactions occur is intended to be the same value in this work [49]. The probability of reaction between these threading dislocations at any point on the free surface is supposed to be uniform and occur with a reaction rate of V_{ij}. This is defined as a function of the difference between the velocities $\mathbf{v}_i$ and $\mathbf{v}_j$ of the threading dislocations [49]:

$$V_{ij} = 2r\ \left|\mathbf{v}_i - \mathbf{v}_j\right|. \tag{5.20}$$

The velocities are calculated properly for sessile and glissile threading dislocations through expressions (5.13) and (5.19), respectively.

Using 5.20 as well as the reaction kinetic approach introduced by Romanov [64], it is possible to describe how the threading dislocation density changes during epitaxial growth. The densities of the i-th and j-th threading dislocation families are indicated with ρ_i and ρ_j respectively. If annihilation or fusion occur between dislocations of i-th and j-th kinds, then the diminishing density for the i-th kind can be calculated by [49]

$$\frac{\mathrm{d}\rho_i}{\mathrm{d}t} = -V_{ij}\rho_i\rho_j. \tag{5.21}$$

and a similar expression can be written for the density of the j-th kind threading dislocations. When multiplying both the members of the previous equation by $\mathrm{d}t$ and combining with expressions (5.19) and (5.20), (5.21) becomes

$$\frac{\mathrm{d}\rho_i}{\mathrm{d}h} = -R_{ij}\rho_i\rho_j. \tag{5.22}$$

where

$$R_{ij} = 2\mathrm{r}\left|\left(\tan\alpha_i\ \mathbf{m}_i + \frac{\mathrm{C}}{\rho_{\mathrm{G}}h^2}\ \mathbf{f}_i\right) - \left(\tan\alpha_j\ \mathbf{m}_j + \frac{\mathrm{C}}{\rho_{\mathrm{G}}h^2}\ \mathbf{f}_j\right)\right|. \tag{5.23}$$

Fusion reactions between the i-th and j-th threading dislocations also have the effect of producing a new threading dislocation of the k-family; therefore a production term is taken into account for each kind of threading dislocation:

$$\frac{\mathrm{d}\rho_k}{\mathrm{d}h} = R_{ij}\rho_i\rho_j. \tag{5.24}$$

Scattering reactions between the i-th and j-th threading dislocations can produce two k-th threading dislocations (see the reaction Table 5.1); therefore the previous equation is corrected as

$$\frac{\mathrm{d}\rho_k}{\mathrm{d}h} = 2R_{ij}\rho_i\rho_j. \tag{5.25}$$

In general, the evolution of density ρ_k of dislocations belonging to the k-th family with the film thickness is described by

$$\frac{\mathrm{d}\rho_k}{\mathrm{d}h} = M_{ij}^k R_{ij}\rho_i\rho_j = K_{ij}^k\rho_i\rho_j, \tag{5.26}$$

where $M_{ij}^k = 0$ in case of annihilation, $M_{ij}^k = 1$ ($k \neq i$ and $k \neq j$) for fusion and scattering reactions, and $M_{ij}^k = 2$ ($k \neq i$ and $k \neq j$) for scattering reactions that produce two threading dislocations belonging to the same ρ_k family (see the reaction Table 5.1).

5.4.3 Sessile and glissile dislocations

Until now, the evolution of the dislocation density was described independently from the ability of the dislocation to glide or not. As discussed previously, when the film thickness is lower than its critical thickness, all threading dislocations are unable to glide along the substrate-film interface. Therefore, all threading dislocations are sessile before reaching the critical thickness. In hexagonal structures there are 20 threading dislocation families (see Section 5.2). Since each threading dislocation can be glissile or sessile above the critical thickness depending on the above mentioned conditions, there are 20 families of glissile threading dislocations and 20 families of sessile threading dislocations. The densities of these two groups are evaluated separately in the subsequent treatment. In particular, the rate R_{ij} is denoted $R_{\mathrm{G},ij}$ in the case where threading dislocations are glissile, $R_{\mathrm{S},ij}$ both the threading dislocations are sessile or $R_{\mathrm{GS},ij}$ when one threading dislocation is glissile and the other sessile. Similarly, K_{ij}^{k} is denoted $K_{\mathrm{G},ij}^{k}$ in the case where both threading dislocations are glissile, $K_{\mathrm{S},ij}^{k}$ when both threading dislocations are sessile or $K_{\mathrm{GS},ij}^{k}$ when one threading dislocation is glissile and the other sessile.

The total density of threading dislocations, i.e., the number of threading dislocations per unit area of the free surface, is indicated with ρ_{TD}, the total density of sessile threading dislocations and the total density of glissile threading dislocations are indicated with ρ_{S} and ρ_{G} respectively. Therefore, below the critical thickness we have

$$\rho_{\mathrm{TD}} = \rho_{\mathrm{S}} = \sum_{k=1}^{20} \rho_{\mathrm{S},k}. \tag{5.27}$$

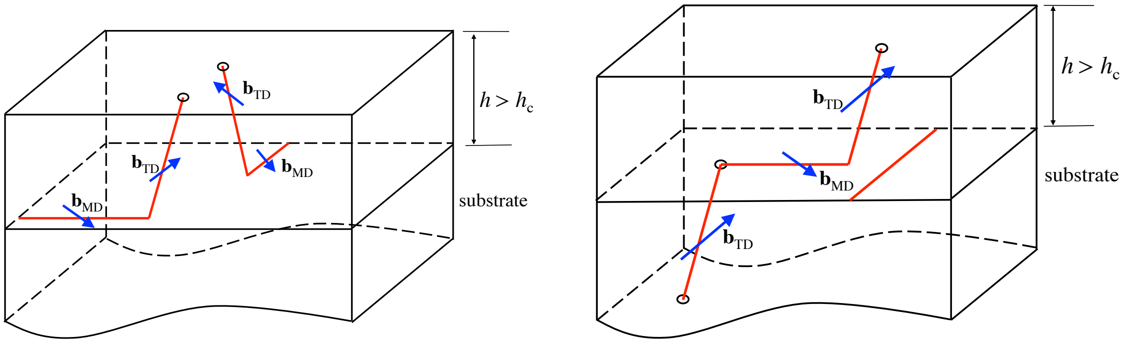
Above the critical thickness, it is assumed that each threading dislocation can glide, becoming glissile and leaving a misfit segment behind. Consequently, the total density of threading dislocations is the sum of sessile and glissile threading dislocation densities. If ρ_{G} indicates the total density of glissile threading dislocations, above the critical thickness this results in

$$\rho_{\mathrm{TD}} = \rho_{\mathrm{G}} + \rho_{\mathrm{S}} = \sum_{k=1}^{20} \rho_{\mathrm{G},k} + \sum_{k=1}^{20} \rho_{\mathrm{S},k}. \tag{5.28}$$

Changing the film thickness by $\mathrm{d}h$, the previous expression becomes

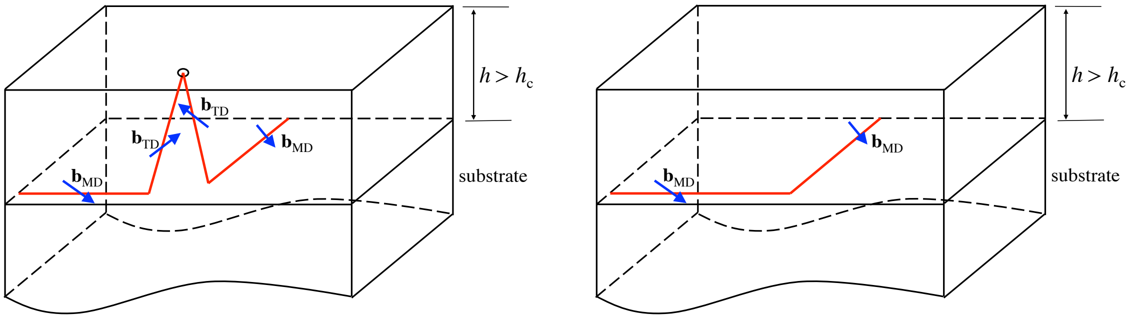
$$\frac{\mathrm{d}\rho_{\mathrm{TD}}}{\mathrm{d}h} = \sum_{k=1}^{20} \frac{\mathrm{d}\rho_{\mathrm{G},k}}{\mathrm{d}h} + \sum_{k=1}^{20} \frac{\mathrm{d}\rho_{\mathrm{S},k}}{\mathrm{d}h}. \tag{5.29}$$

The complete relaxation of the structure stops the threading dislocation glide, and glissile threading dislocations become sessile again. Considering equation (5.19), the complete relaxation is reached when $\mathbf{v}_{\mathrm{G}}$ reduces to $\mathbf{v}_{\mathrm{S}}$ for $h \to \infty$.



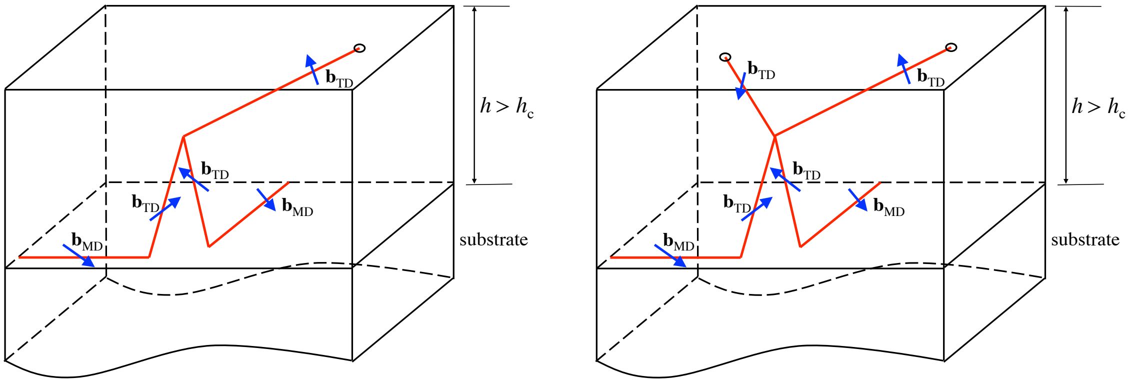
(a) Initial configuration with two glissile threading dislocations in strained film of thickness $h_f > h_c$.

(b) Blocking of threading dislocation motion by a misfit dislocation in the slip path of the threading dislocation.



(c) Intermediate configuration with threading dislocations at the distance of the annihilation radius r.

(d) Annihilation of the two threading dislocations: final configuration with misfit dislocations only (the threading dislocations are annihilated).



(e) Fusion of the two threading dislocations: final configuration with one new threading dislocation.

(f) Scattering of the two threading dislocations: final configuration with two new threading dislocations.

Figure 5.13: Movement and reactions of the threading dislocations

In parallel, other factors can stop the dislocation glide above the critical thickness and before the saturation point. During the motion, one glissile threading dislocation may be blocked by a misfit segment lying in the path of another moving threading dislocation, as shown schematically in Figure 5.13(b). The problem of threading dislocation blocking was treated in detail by Freund [18]. In the present treatment, once a threading dislocation has been blocked by a crossing misfit dislocation, its motion will be considered arrested for all further film growth. The chance of intersection of a randomly located glissile threading dislocation with a misfit dislocation is supposed to be $\rho_{\mathrm{G},k}\rho_{\mathrm{MD}}\mathrm{d}h$. The result is a decrease of the density of k-type glissile threading dislocations by

$$\mathrm{d}\rho_{\mathrm{G},k} = -\rho_{\mathrm{G},k}\rho_{\mathrm{MD}}\mathrm{d}h, \tag{5.30}$$

and consequently the density of the k-type sessile threading dislocations increases by

$$\mathrm{d}\rho_{\mathrm{S},k} = \rho_{\mathrm{G},k}\rho_{\mathrm{MD}}\mathrm{d}h. \tag{5.31}$$

Also the reactions among threading dislocations can affect the ratio between glissile and sessile dislocations. The number of glissile dislocations decreases when two glissile threading dislocations annihilate each other and therefore the density of k-type glissile threading dislocations is decreased by the following:

$$\mathrm{d}\rho_{\mathrm{G},k} = -\left(\sum_{j} R_{\mathrm{G},kj}\rho_{\mathrm{G},k}\rho_{\mathrm{G},j} + \sum_{j} R_{\mathrm{GS},kj}\rho_{\mathrm{G},k}\rho_{\mathrm{S},j}\right)\mathrm{d}h. \tag{5.32}$$

Annihilation is not the only reaction which decreases the density of glissile dislocations. Few experimental observations of the interaction between a glissile and a sessile dislocations are available in literature. Thus it is assumed in this work that the motion of one glissile dislocation ends when it encounters another threading dislocation (either glissile or sessile). At this point, the two threading dislocations react with each other by fusion or scattering and the resulting dislocation/dislocations remains/remain sessile. The density decreases by

$$\mathrm{d}\rho_{\mathrm{G},k} = -\left(\sum_{ij} K^{k}_{\mathrm{G},ij}\rho_{\mathrm{G},i}\rho_{\mathrm{G},j} + \sum_{ij} K^{k}_{\mathrm{GS},ij}\rho_{\mathrm{G},i}\rho_{\mathrm{S},j}\right)\mathrm{d}h. \tag{5.33}$$

In parallel, the density of k-type sessile threading dislocations increases by

$$\mathrm{d}\rho_{\mathrm{S},k} = \left(\sum_{ij} K^{k}_{\mathrm{G},ij}\rho_{\mathrm{G},i}\rho_{\mathrm{G},j} + \sum_{ij} K^{k}_{\mathrm{GS},ij}\rho_{\mathrm{G},i}\rho_{\mathrm{S},j}\right)\mathrm{d}h. \tag{5.34}$$

Based on equations (5.30), (5.32), and (5.33), by increasing the film thickness by $\mathrm{d}h$, the total variation of the density of k-type glissile threading dislocations is

$$\begin{aligned}\frac{\mathrm{d}\rho_{\mathrm{G},k}}{\mathrm{d}h} =& -\rho_{\mathrm{G},k}\ \rho_{\mathrm{MD}} \\ & -\sum_j\ R_{\mathrm{G},kj}\ \rho_{\mathrm{G},k}\ \rho_{\mathrm{G},j} - \sum_j\ R_{\mathrm{GS},kj}\ \rho_{\mathrm{G},k}\ \rho_{\mathrm{S},j} \\ & -\sum_{ij}\ K^k_{\mathrm{G},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{G},j} - \sum_{ij}\ K^k_{\mathrm{GS},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{S},j}. \qquad (5.35)\end{aligned}$$

The number of sessile threading dislocations is increased by the decrease of the glissile threading dislocation density (see equations (5.31) and (5.34)). In addition the reactions among sessile threading dislocations affect the density of the k-type sessile threading dislocation:

$$\mathrm{d}\rho_{\mathrm{S},k} = \left(\sum_{ij} K^k_{\mathrm{S},ij}\rho_{\mathrm{S},i}\rho_{\mathrm{S},j} - \sum_j R_{\mathrm{S},kj}\rho_{\mathrm{S},k}\rho_{\mathrm{S},j}\right)\mathrm{d}h. \qquad (5.36)$$

The k-type sessile threading dislocations can annihilate with glissile dislocations of an another family, therefore

$$\mathrm{d}\rho_{\mathrm{S},k} = -\left(\sum_j R_{\mathrm{GS},kj}\rho_{\mathrm{S},k}\rho_{\mathrm{G},j}\right)\mathrm{d}h. \qquad (5.37)$$

Based on equations (5.31), (5.34), (5.36), and (5.37), by increasing the film thickness by $\mathrm{d}h$, the total variation of the density of k-type sessile threading dislocations is

$$\begin{aligned}\frac{\mathrm{d}\rho_{\mathrm{S},k}}{\mathrm{d}h} =& +\rho_{\mathrm{G},k}\ \rho_{\mathrm{MD}} \\ & +\sum_{ij} K^k_{\mathrm{G},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{G},j} + \sum_{ij} K^k_{\mathrm{GS},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{S},j} \\ & +\sum_{ij} K^k_{\mathrm{S},ij}\ \rho_{\mathrm{S},i}\ \rho_{\mathrm{S},j} - \sum_j R_{\mathrm{S},kj}\ \rho_{\mathrm{S},k}\ \rho_{\mathrm{S},j} \\ & -\sum_j R_{\mathrm{GS},kj}\ \rho_{\mathrm{S},k}\ \rho_{\mathrm{G},j}. \qquad (5.38)\end{aligned}$$

The variation of ρ_{TD} with the film thickness $\mathrm{d}h$ is obtained by combining expression (5.29) with expressions (5.17), (5.35) and (5.38):

$$\frac{\mathrm{d}\rho_{\mathrm{TD}}}{\mathrm{d}h} = \sum_{k=1}^{20}\frac{\mathrm{d}\rho_{\mathrm{G},k}}{\mathrm{d}h} + \sum_{k=1}^{20}\frac{\mathrm{d}\rho_{\mathrm{S},k}}{\mathrm{d}h}, \qquad (5.39)$$

$$\begin{aligned}\frac{\mathrm{d}\rho_{\mathrm{G},k}}{\mathrm{d}h} =& -\rho_{\mathrm{G},k}\ \rho_{\mathrm{MD}} \\ & -\sum_j\ R_{\mathrm{G},kj}\ \rho_{\mathrm{G},k}\ \rho_{\mathrm{G},j} - \sum_j\ R_{\mathrm{GS},kj}\ \rho_{\mathrm{G},k}\ \rho_{\mathrm{S},j} \\ & -\sum_{ij}\ K^k_{\mathrm{G},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{G},j} - \sum_{ij}\ K^k_{\mathrm{GS},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{S},j}, \qquad (5.40)\end{aligned}$$

$$
\begin{aligned}
\frac{\mathrm{d}\rho_{\mathrm{S},k}}{\mathrm{d}h} = &+ \rho_{\mathrm{G},k}\ \rho_{\mathrm{MD}} \\
&+ \sum_{ij} K^{k}_{\mathrm{G},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{G},j} + \sum_{ij} K^{k}_{\mathrm{GS},ij}\ \rho_{\mathrm{G},i}\ \rho_{\mathrm{S},j} \\
&+ \sum_{ij} K^{k}_{\mathrm{S},ij}\ \rho_{\mathrm{S},i}\ \rho_{\mathrm{S},j} - \sum_{j} R_{\mathrm{S},kj}\ \rho_{\mathrm{S},k}\ \rho_{\mathrm{S},j} \\
&- \sum_{j} R_{\mathrm{GS},kj}\ \rho_{\mathrm{S},k}\ \rho_{\mathrm{G},j},
\end{aligned}
\tag{5.41}
$$

$$
\frac{\mathrm{d}\rho_{\mathrm{MD}}}{\mathrm{d}h} = \frac{2\epsilon_m h_c \csc\theta \csc\phi}{3b \ln\frac{h_c}{r_c}} \frac{\ln\frac{h_c}{r_c} - 1}{h^2} \mathrm{d}h = \frac{\mathrm{C}_{\mathrm{hex}}}{h^2}. \tag{5.42}
$$

The calculation of the total density of the threading dislocation is performed with the software Wolfram Mathematica for the heterostructures considered in the next Paragraphs.

5.4.4 Initial and boundary conditions

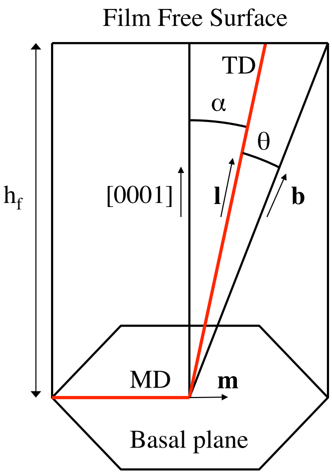


Figure 5.14: Geometry of a mixed dislocation in GaN.

In this section the initial and boundary conditions used to solve the system of equations defined by equations (5.29), (5.17), (5.35) and (5.38) are listed.

- A single-crystal structure of $\mathrm{Al}_{1-x}\mathrm{Ga}_x\mathrm{N}$ oriented along the [0001] direction is considered. The values of the lattice parameters of GaN and AlN are $\mathrm{a}_{\mathrm{GaN}}$=3.22Å, $\mathrm{c}_{\mathrm{GaN}}$=5.19Å, $\mathrm{a}_{\mathrm{AlN}}$=3.11Å, and $\mathrm{c}_{\mathrm{AlN}}$=4.98Å respectively. The values for the $\mathrm{Al}_{1-x}\mathrm{Ga}_x\mathrm{N}$ alloy are calculated using Vegard's law [76].

- The core radius of each threading dislocation is supposed to be half of its Burgers vector.

- Anisotropic elasticity is used to calculate the critical thickness and the misfit dislocation density; consequently, C=C$_{\text{hex}}$ (see equation (5.17)).

- The interaction radius is 40nm [79].

- The importance of the inclination angle for the calculation of the dislocation density was discussed in Section 5.2.1. The equilibrium values for the inclination angle were calculated in Section 3.7. According to this calculation, in [0001] GaN all types of threading dislocation are parallel to the [0001] direction. In reality, experimental observations show that this angle is influenced by many factors not considered in this work, such as temperature, island growth, doping, masking, and misfit stress relaxation. Mathis [49] and other authors [22] observed that a- and c-type threading dislocations are parallel to the [0001] direction but the $(a+c)$ type threading dislocations are inclined by 12°. Holec [25] showed how the island-growth mode can bend threading dislocations in the proximity of inclined facets. Cantu [7] grew Si doped $Al_{0.49}Ga_{0.51}N$ films upon $Al_{0.62}Ga_{0.38}N$ and observed inclined dislocations in the film. The inclination was suggested to be caused by the Si doping. Li [43] arrived at the same conclusion by doping a GaN-based film with Mg. In contrast to their explanations, Follstaedt [16] and coauthors realized that in similar structures, dislocations are bent before the introduction of any dopant. According to him, the relaxation of the misfit stress causes the bending. In this work, the screw and edge threading dislocations in GaN are supposed to be aligned with the [0001] direction, and the $(a+c)$-type threading dislocations inclined with respect to it (like assumed by Mathis [49]). Based on these experimental findings, the inclination angle is assumed to be very small (5°). The reason for this comes from evaluating the threading dislocation density in an unfavorable condition, i.e., the reaction probability is minimum.

- It is also assumed that the planar growth mode prevents threading dislocation inclination due to the vicinity with the inclined facets.

- The initial amount of threading dislocations is set at 10^{10}cm^{-2} as reported by different authors [38,56,81]. The initial threading dislocation density is composed of 70% a-type threading dislocation and 30% $(a+c)$-type threading dislocation. Different ratios are evaluated in Section 5.4.5.

- Several authors [39, 49] show that c-type dislocations are not very mobile and not capable of reacting with other dislocations. Different explanations for this behavior have been proposed. The c-type dislocation oriented along the [0001] direction is a screw dislocation. It might happen, for example, that screw dislocations exist at the center of a growth hillock or at the center of a nanopipe. In the case of a growth hillock, mixed dislocations must elongate and move across the hillock to reach the screw dislocation, making reactions less likely.

- A major assumption of the present model is a simplification of the geometry of the dislocation glide. It is assumed that a misfit segment and the upper point of a threading dislocation have the same direction, *i.e.*, they have the same velocity $\mathbf{f} = \mathbf{m}$ (see Figure 5.14), and equation (5.19) for glissile dislocations reduces to

$$\mathbf{v}_{\mathrm{G}} = \left(\tan\alpha + \frac{\mathrm{C}}{h^2 \rho_{\mathrm{G}}} \right) \frac{\mathrm{d}h}{\mathrm{d}t}\, \mathbf{m}. \tag{5.43}$$

As a consequence, it is assumed that only the $(a+c)$-type threading dislocations have the possibility to move, being that they are the only ones tilted with respect to the growth direction, the only ones to execute lateral motion toward each other and the only ones to react with an increase in the layer thickness. Despite this simplification, the model brings some insight and, although not quantitatively, at least qualitatively accounts for the contribution of glide to the threading dislocation density reduction.

5.4.5 GaN/AlN bilayer

The model described in Section 5.4 is used to calculate the threading dislocation density as a function of the film thickness h_{f} in a [0001] GaN film grown upon a AlN substrate. The results (red line in Figure 5.15) are compared with those from Mathis' model [49] (blue line in Figure 5.15). The initial and boundary conditions are listed in Paragraph 5.4.4.

The development of the threading dislocation density according to Mathis can be divided into three stages, like as with semipolar bulk GaN (see Section 5.2.1). Initially, the density remains quite constant because no reactions occur due to the big distance among dislocations. After 1μm, the density decreases due to threading dislocation reactions. When all dislocations having reacted, their density assumes an asymptotical value, reaching the saturation point at 1mm.

The threading dislocation density according to the model proposed in this work is represented by the red solid line in Figure 5.15. This can be divided into five stages.

1. Initially the density remains constant because dislocations are too far from each other to react.

2. After the critical thickness is reached, which is 3μm according to Steeds [72], dislocations glide along the interface and this causes the reactions. Therefore, the threading dislocation density decreases in the 10nm above the critical thickness. As a result of the threading dislocation glide, the density is 15% less with respect to what was predicted by Mathis.

3. After $h_{\mathrm{f}} = 15$nm, the density is quite constant because dislocations not longer glide (since the misfit stress is totally relaxed) but reactions driven by threading dislocation inclination do not take place until 1μm.

4. After $h_{\mathrm{f}} = 1\mu$m, reactions caused by threading dislocation inclination occur, which decreases the density until the point the model reduces to the Mathis model at $h_f = 10\mu$m.

5. For $h_\mathrm{f} > 10\mu$m, the behavior is equal to that of the Mathis model.

Experimental observations (from private discussions with Dr. Baumgartl) confirm that the presence of the interface and the subsequent glide of the dislocations decreases the threading dislocation density. Therefore, the improved model agrees better with the experimental observations respect to the Mathis' model.

The red dotted line in Figure 5.15 represents the threading dislocation density – according to the model described in this work – without threading dislocation blocking caused by the presence of misfit dislocations along the interface. The red solid and dotted lines are close to each other meaning that threading dislocation blocking is negligible. This means that the first term from the left sides of both equations (5.35) and (5.38) is negligible. In the following, the threading dislocation density for the same bilayer – GaN film upon AlN substrate – is calculated, thus the initial conditions are changed.

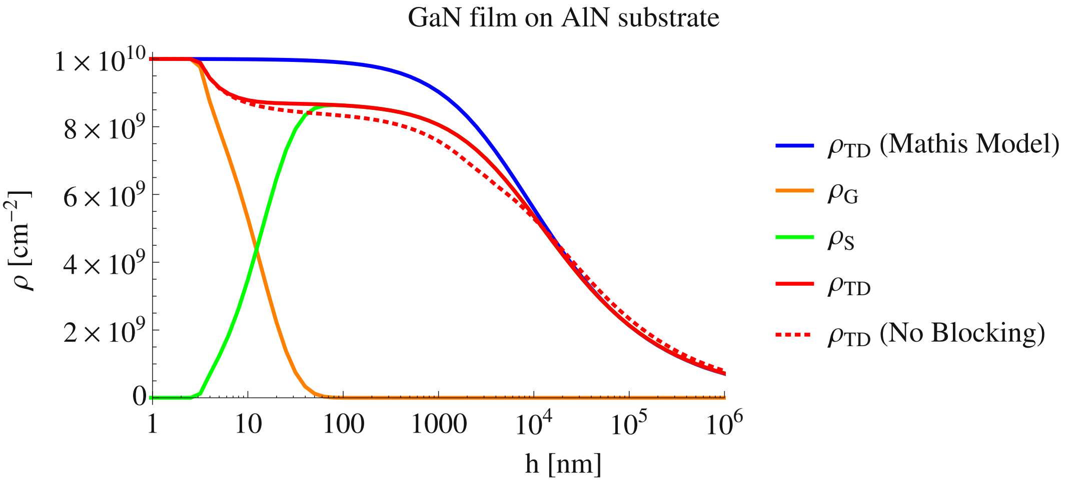


Figure 5.15: The red solid line shows the threading dislocation density, calculated according to the model described in Section 5.4, as a function of the GaN film thickness h. The GaN film is grown upon an AlN thicker layer. The green and orange lines show the sessile and glissile threading dislocation densities, respectively. The red dotted line shows the threading dislocation density, according to the same model, without threading dislocation blocking. The blue line shows the threading dislocation density, calculated according to Mathis' model [49], as a function of the GaN film thickness h.

One assumption is that only $(a+c)$-type threading dislocations are capable of gliding. In order to increase the impact of the glide, the initial amount of $(a+c)$-type threading dislocations is raised from 30% up to 60% and 90%. The related results are shown in Figure 5.16. When the initial threading dislocation density is composed of 90% of $(a+c)$-type threading dislocations, the final density is reduced by 30% with respect to the value given by Mathis' model for $h_\mathrm{f} < 2\mu\mathrm{m}$. Instead, for $h_\mathrm{f} > 2\mu\mathrm{m}$, the variation of the initial amount of $(a+c)$-type threading dislocations has a negligible impact.

The effect of the inclination angle α is evaluated keeping the other initial and boundary conditions equal to those reported in Section 5.4.4. Both Mathis' model and the improved one are used to calculate the threading dislocation density for different values of α. Figure 5.17 shows that the variation of α becomes remarkable in both models at higher thicknesses. At $h_\mathrm{f} = 1\mu\mathrm{m}$, the threading dislocation density with $\alpha = 60°$ is halved with respect to the case with $\alpha = 5°$. The models confirm that increasing the inclination angle – by island growth, doping, and strain relaxation – drastically improves the quality of the crystalline structure.

In conclusion, the crystalline quality of GaN film grown upon an AlN thicker layer is improved by increasing the inclination angle rather than other factors. The dislocation glide decreases the threading dislocation density only by 15% with respect to what was predicted by Mathis' model for bulk structures. The presence of one interface is ineffective in reducing the dislocation density. The main goal now is to evaluate the impact of the glide in a more complex heterostructure, with many interfaces along which threading dislocation glide occurs.

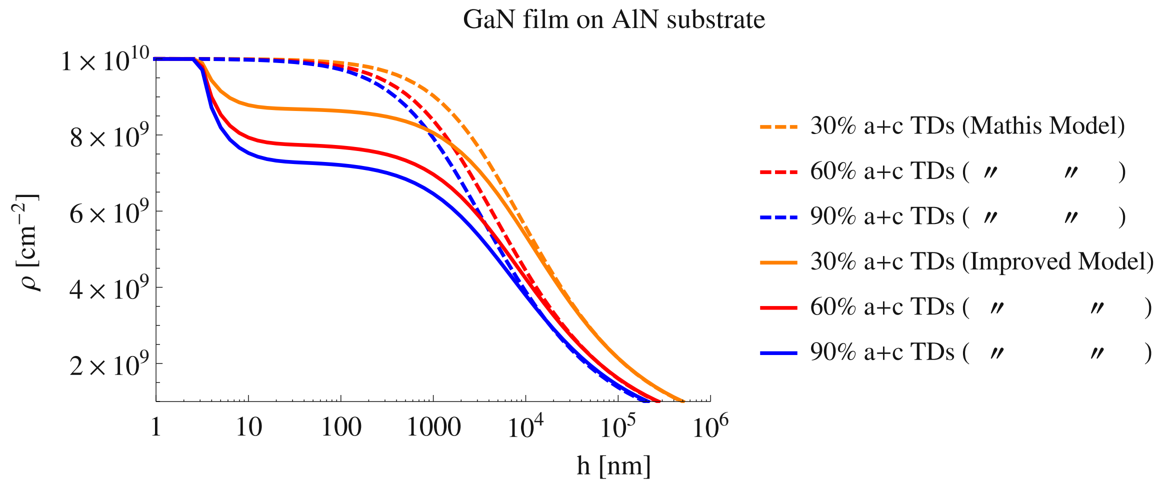


Figure 5.16: Threading dislocation density as function of the GaN film thickness h calculated according to the Mathis model [49] and the treatment described in Section 5.4 with different initial amounts of $(a+c)$-type threading dislocations.

5.4.6 GaN-based multilayer

The improved model described in Section 5.4 is used to calculate the threading dislocation density as a function of the multilayer thickness grown upon an AlN substrate. Two types of multilayer are considered. The first is a step-graded layer composed of $\mathrm{Al_{0.75}Ga_{0.25}N}$ (200nm) / $\mathrm{Al_{0.5}Ga_{0.5}N}$ (200nm) / $\mathrm{Al_{0.25}Ga_{0.75}N}$ (200nm) / GaN. The

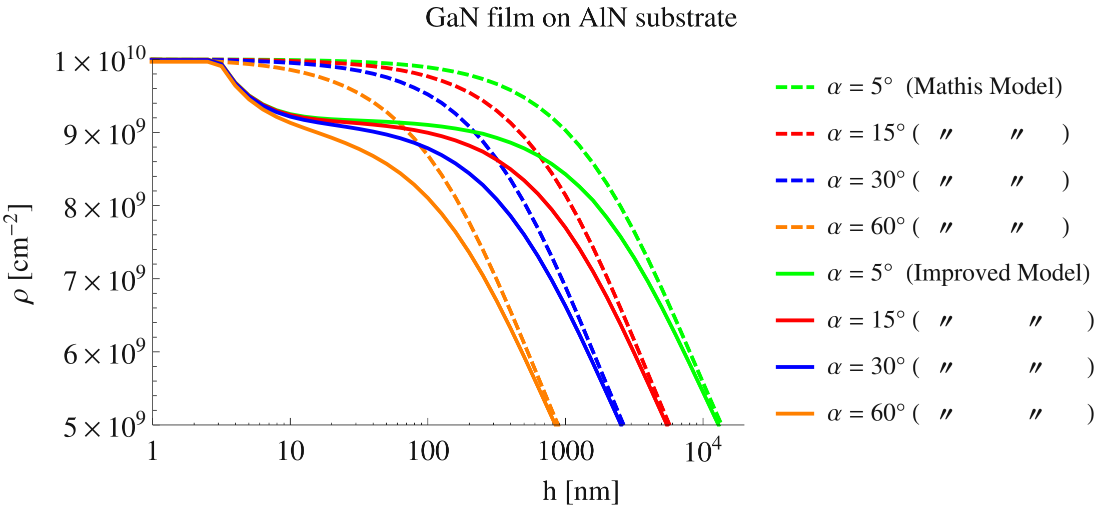


Figure 5.17: Threading dislocation density as a function of the GaN film thickness h calculated according to the Mathis model [49] and the treatment described in Section 5.4 using different values for the inclination angle of the $(a+c)$-type threading dislocations.

results are shown in Figure 5.18. The second structure is a $(\mathrm{AlN/GaN})_{10}$ superlattice, and the related evolution of the threading dislocation density is shown in Figure 5.19. In both cases the initial and boundary conditions are the same as reported in Paragraph 5.4.4.

Considering the red line in Figure 5.18, the threading dislocation density after four interfaces at 1μm is 25% less than the initial value. A decrease of the threading dislocation density of technological relevance should be at least of one order of magnitude. Therefore this step graded layer seems to not be effective for improving the crystalline quality of the device.

As can be observed in Figure 5.18, the threading dislocation density quickly decreases suddenly after the critical thickness associated with each interface. This decrease is caused by threading dislocation glide and is proportional to the lattice mismatch associated with each interface.

The other curves of Figure 5.18 are related to multilayers with a lower number of layers with respect to the one which is represented by the red line. The results indicate that a lower number of layers decreases the threading dislocation density less.

Figure 5.19 shows the development of threading dislocation density in a $(\mathrm{AlN/GaN})_{10}$ superlattice. This is composed of 10 periods of AlN/GaN. Each layer is 20nm thick. The superlattice is grown upon the AlN substrate and is h = 400nm high. A GaN layer extends for $h >$ 400nm.

The red line shows the threading dislocation density, calculated with the model of Section 5.4, across the superlattice. Other lines show the threading dislocation density with a lower number of layers in the superlattice. The black line corresponds to the threading dislocation density according to Mathis [49] for a GaN layer grown upon an AlN substrate. All curves have the same behavior at high thickness, $h > 10^5$nm. However, differences among the cases are evident in the first 400nm. In particular, the superlattice reduces the threading dislocation density by 70% with respect to the initial value, resulting in a more efficient method than the step-graded layer to decrease the dislocation density. Again, a lower number of layers corresponds to a lower value of the final threading dislocation density. A quick decrease of the threading dislocation density is visible after the critical thickness of each interface is reached. This is caused by threading dislocation glide. However, the density decreases more at interfaces placed at lower thicknesses than at higher thicknesses. For example, the threading dislocation density decreases by 15% at the first interface while it decreases by 1% at the last one. This is caused by the incomplete relaxation of the misfit stress of each interface. The smaller the layer thickness, the smaller the relaxation of the misfit stress. This implies that in the last layers of a superlattice, the lattice parameter tends to be constant through the thickness and equal to a value between the AlN and GaN lattice parameters. The consequence it that a negligible lattice mismatch is present at high thicknesses, i.e., the misfit stress is small, i.e., the glide is negligible. Therefore, increasing the number of layers after this threshold does not improve the crystalline quality. One possible method to complete the stress relaxation is growth of a thick layer, GaN or AlN, to recover the intrinsic value of the lattice parameter. Then a new superlattice can be grown to further reduce the threading dislocation density. This has been observed by several authors [14, 85]. Their experimental data (points 3 and 4 in Figure 5.19) show a threading dislocation density which is slightly smaller than the values predicted by the improved model because the experimental structures are composed of several superlattices separated by thick layers. Other experimental data [75, 84] tend to confirm the prediction of the threading dislocation density calculated according to the model proposed in this work (see Section 5.4).

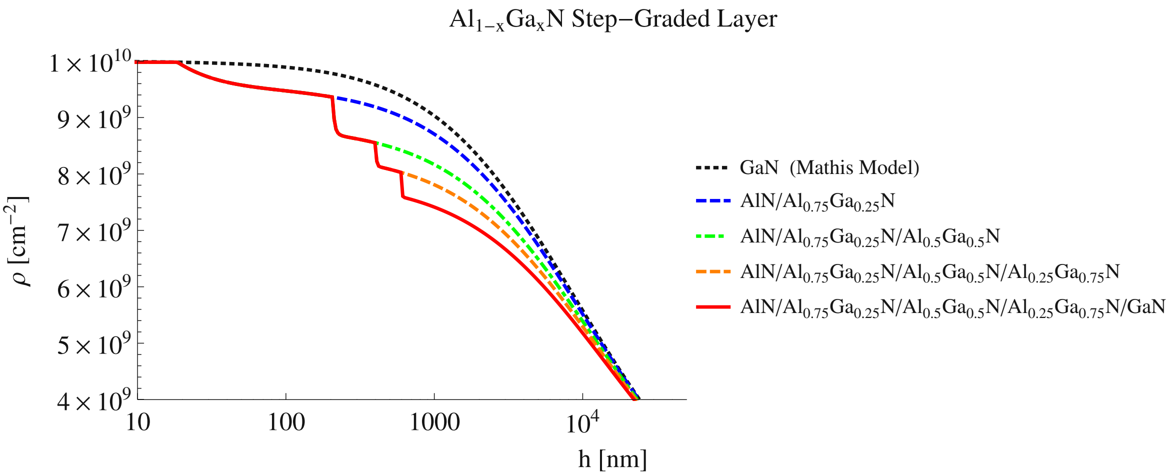


Figure 5.18: The red line shows the threading dislocation density, calculated using the model of Section 5.4, as a function of the thickness h of the $\mathrm{Al}_{1-}\mathrm{Ga}_\mathrm{f}\mathrm{N}$ step-graded layer grown upon an AlN substrate. Dashed lines show the threading dislocation density for structures with a lower number of layers. The black dotted line shows the threading dislocation density according to Mathis [49] for a GaN film grown upon the AlN substrate.

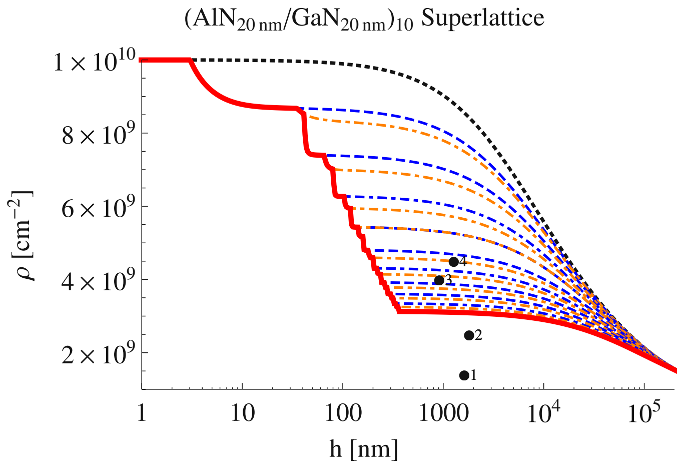


Figure 5.19: The red line shows the threading dislocation density, calculated using the model of Section 5.4, as a function of the thickness h of the $(\mathrm{AlN}/\mathrm{GaN})_{10}$ superlattice grown upon an AlN substrate. The dashed lines show the threading dislocation density using a superlattice with a lower number of layers. The black dotted line shows the threading dislocation density according to Mathis' model for a GaN film grown upon an AlN substrate. The theoretical results are compared to experimental data, 1- [85], 2- [14], 3- [75], 4- [84].

5.5 Summary

In this chapter Mathis' model [49] is improved with regard to the dislocation glide along interfaces and its effect on the dislocation density. The new treatment assumes a simplified bidimensional geometry of the dislocation. Several kinetic factors affecting the dislocation movement, like temperature and speed of growth, are neglected. The results of the calculations show the following

1. The glide of the dislocations alone one interface happen in a short range on tens of nanometer after the critical thickness. The glide increases the probability of reactions among dislocations, reducing their density. The reduction is generally small for a bilayer structure but can be magnified in multilayer structures;

2. The effect of the inclination angle on the dislocation density is evident at a thickness much higher (on the order of hundreds of nanometers) than the critical thickness. The higher the inclination angle, the lower the final dislocation density. In particular, the island-growth mode seems to favor the inclination of the dislocations (see Chapter 3 and [25]);

3. Superlattice structures reduce the dislocation density more efficiently than the step-graded layer. In the last case, the glide of the dislocations can be neglected;

4. Superlattice structures reduce the dislocation density by nearly one order of magnitude when the thickness of the structure is half μm, hence yielding good agreement with experimental data. This also demonstrates that the improved model better predicts the final density with respect to Mathis' model. The efficiency of the superlattice is due to the high lattice mismatch between layers. The higher the lattice mismatch, the more intense the dislocation glide.

The model presented in this chapter can be improved by the following:

1. Consideration of the real 3D geometry of the dislocation;

2. Consideration of the actual dislocation-dislocation interactions. This would involve calculation of dislocation stress fields and, in fact, the employment of some methods of dislocation dynamics;

3. Evaluation of the impact of the growth mode. Considering the island-growth mode would mean considering the different facets of each island and the different growth rates of the facets;

4. Choosing of the inclination angle in relation to the growth mode (see Chapter 3 and [25]). This would involve the calculation of the stochastic distribution of the angles – which dislocations assume during the layer growth – and in fact, the use of some statistical methods.

6 Conclusions and outlook

Gallium nitride (GaN) and its alloys with aluminum or indium are the most important semiconductors because they have better material and electronic properties compared to silicon and other III-V compounds, like GaAs. These properties enable the use of III-nitride based devices in a broad range of applications, from electronic to optical devices.

Mass production of III-nitrides is already a reality, but despite this, many efforts are still placed on the further development of their technology and the overcoming of challenges posed by the lack of a native substrate, such as the large amount of dislocations.

This thesis provides a theoretical contribution to the definition of design rules for improving the crystalline quality of III-nitrides, in particular, the reduction of the dislocation density. Continuum theory of dislocations treated within the linear elasticity theory and the laws of thermodynamics are used to gain understanding and model the dislocation development in these materials.

Chapter 3 summarizes the calculation of the dislocation energy within the isotropic and anisotropic elasticity. The models were used to calculate the equilibrium configuration for dislocations in AlN, GaN, and InN compounds, all oriented along the [0001] direction.

In the first part of this work, dislocations were assumed to be in bulk, far from the free surface of the material. In this case, the two theories give different results: in monocrystals the a and $(a + c)$ dislocations are not of the screw type – as predicted in the isotropic framework – but instead mixed type dislocations.

Secondly, dislocations were assumed to be close to the free surface. When the free surface is the (0001) plane, elastic anisotropic theory predicts that the dislocation line along the [0001] direction possesses the lowest energy configuration regardless of the dislocation type in both AlN and GaN. In InN only the c-type dislocations are screw dislocations, while the other two types are mixed dislocations.

When the free surfaces are inclined facets {1122} and {1101}, elastic anisotropic theory predicts different results depending on the compound. Regarding AlN, the preferred dislocation line is along the [0001] direction independent from the dislocation type and crystallographic plane as a free surface. The a-type dislocations in GaN and InN propagate almost perpendicular to the {1122} and {1101} facets. Regarding the other types, in GaN the dislocations are inclined by 20°- 30°, while in InN they are aligned along the [0001] direction.

The inclined facets {1122} and {1101} are the free surfaces when the island-growth mode is favored during the deposition of the compounds. The conclusion is that the island-growth mode increases in general the inclination of the dislocations with respect to the growth direction. The same theoretical result was obtained by Holec [25] for dislocations in GaN single crystals. However, Mathis [49] and other authors [22] measured experimentally other values of the inclination angle. In addition, experimental observations show that the inclination angle is influenced not only by the anisotropic structure of the crystal but also by factors not evaluated in this work. For example, Cantu [7] and Li [43] demonstrated experimentally that the doping of the crystal can influence the inclination angle. The wide scatter of the experimental data of the inclination angle suggests also that an important source of the dislocation inclination is the stress/strain gradient of the structure, as shown by Follstaedt [16]. In order to calculate the stress/strain gradient, it would be necessary to evaluate the numerous factors which influence the mechanical stability of the epilayers. This would address primarily the thermal behavior of the III-nitrides.

The pre-logarithmic terms of the analytical models have been compared with the corresponding values obtained by atomistic simulations. Good agreement has been found with the continuum predictions based on anisotropic elasticity. Therefore both approaches have been combined into a kind of a multiscale approach for predicting the onset of misfit dislocations in thin film, as shown in Section 4.6.

In Chapter 4, different continuum-based approaches for calculating the energy of a straight infinitely long dislocation in an elastic medium have been evaluated. Motivated by the misfit dislocations in a heteroepitaxial interface, we evaluated separately the influence of (i) free surface, (ii) different elastic constants in the film and substrate, and (iii) elastic anisotropy. The results suggest that starting from a homogeneous infinite isotropic medium, the inclusion of a free surface increases the dislocation energy, and the difference in elastic constants of the film and substrate does not play any significant role (because it is typically an order of magnitude smaller than the impact of, e.g., the free surface), while the inclusion of elastic anisotropy decreases the dislocation energy.

Finally, the equilibrium critical thickness was calculated for three important heteroepitaxial material systems, namely, an $Al_xGa_{1-x}N$ film on a GaN substrate, an $In_xGa_{1-x}N$ film on a GaN substrate, and a $Si_{1-x}Ge_x$ film on a Si substrate. A new model including elastic anisotropy of the film and the substrate, the difference of their elastic constants and the impact of the film free surface has been proposed. Recalling that this is a model for the equilibrium critical thickness, i.e., it provides a condition when it first becomes energetically favorable to start the relaxation of the mismatch strain via plastic flow, the refined model yields excellent agreement with the available experimental data in the sense that no misfit dislocations are detected below the here predicted threshold.

The calculations of the critical thickness are of a great interest for technologists. The wide scatter of experimental data of the critical thickness suggests that an important role for the generation of the misfit dislocations is played by the growth technique. Evaluating the impact of the growth technique would be helpful to explain the scatter of the experimental data, nevertheless, the theory of dislocations would not be sufficient alone to reach this goal.

Another possible way to improve the current models is by accounting for the dislocation generation mechanism(s) and thus providing a theory which goes beyond the equilibrium critical thickness as presented in this chapter.

In Chapter 5, a new treatment assuming a simplified 2D geometry of the dislocation describes the dislocation density in heterostructures based on III-nitrides. The new model considers both different threading dislocation types with their specific reactions and inclinations and also their glide along the interfaces. The results of the calculations show that (i) the glide increases the probability of reactions among dislocations, reducing their density; (ii) the glide of the dislocations along one interface happen in the first ten nanometers after the critical thickness while the effect of the inclination angle is more evident after hundreds of nanometers after the critical thickness; and (iii) $(\mathrm{AlN/GaN})_x$ superlattice structures reduce the dislocation density more efficiently than the $\mathrm{Al}_{1-x}\mathrm{Ga}_x\mathrm{N}$ step-graded layer because the first ones reduce the dislocation density by nearly one order of magnitude when the thickness of the structure is half μm high, yielding a good agreement with experimental data.

The model of the dislocation density can be improved in several ways. In order to have a model capable of better quantitative prediction of the experimental data of the dislocation density, it would be necessary to consider the real 3D geometry of the threading dislocations. This would imply the assumption that the direction of the misfit dislocation is different from the direction of the upper point of the threading dislocation, i.e. $\mathbf{f} \neq \mathbf{m}$ (see Paragraph 5.4.1).

Another significant improvement toward realism would be to consider dislocation-dislocation interactions. This would involve calculation of dislocation stress fields and, in fact, employing some methods of dislocation dynamics.

Bibliography

[1] Adachi, S. *Properties of Semiconductor Alloys: Group-IV, III-V and II-VI Semiconductors*. John Wiley & Sons, 2009.

[2] Akasaki, I., and Amano, H. Breakthroughs in improving crystal quality of GaN and invention of the pn junction blue-light-emitting diode. *Jpn. J. Appl. Phys.* (2006), 9001.

[3] Albrecht, M., Nikitina, I., Nikolaev, A., Melnik, Y., Dmitriev, V., and Strunk, H. Dislocation reduction in AlN and GaN bulk crystals grown by hvpe. *Phys. Stat. Sol. 176*, 1 (1999), 453458.

[4] Balluffi, R. W. *Introduction to Elasticity Theory for Crystal Defects*. cambridge university press, 2012.

[5] Belabbas, I., Bere, A., Chen, J., A., S. P. M., Belkhir, Ruterana, P., and Nouet, G. Atomistic modeling of the (a+c)-mixed dislocation core in wurtzite GaN. *Phys. Rev. B: Condens. Matter 75* (2007), 115201.

[6] Bethoux, J.-M., and Venngus, P. Ductile relaxation in cracked metal-organic chemical-vapor-deposition-grown AlGaN films on GaN. *J. Appl. Phys. 97* (2005), 123504.

[7] Cantu, P., Wu, F., Waltereit, P., Keller, S., Romanov, A., Debars, S., and Speck, J. Role of inclined threading dislocations in stress relaxation in mismatched layers. *J. Appl. Phys. 97*, 10 (2005), 103535.

[8] Coppeta, R., Holec, D., Ceric, H., and Grasser, T. Evaluation of dislocation energy in thin films. *Philos. Mag. A* (2015).

[9] Cottrell, A. H. *The Mechanical Properties of Matter*. John Wiley & Sons, 1964.

[10] Dridi, Z., Bouhafs, B., , and Ruterana, P. First-principles calculation of structural and electronic properties of wurtzite $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}$, $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}$, and $\mathrm{In}_x\mathrm{Al}_{1-x}\mathrm{N}$ random alloys. *phys. stat. sol. 1* (2002), 315–319.

[11] Edgar, J. H. *Properties of Group III Nitrides*. INSPEC, the institution of electrical engineers, 1994.

[12] Eifert, B. Some representative crystal structures. `http://demonstrations.wolfram.com/SomeRepresentativeCrystalStructures/`, 2011.

[13] Eshelby, J. D. The continuum theory of lattice defects. *Solid State Phys. 3* (1956), 79–303.

[14] Feltin, E., Beaumont, B., Lagt, M., de Mierry, P., Venngus, P., Leroux, M., and Gibart, P. Crack-free thick GaN layers on silicon (111) by metalorganic vapor phase epitaxy. *Phys. Stat. Sol. (a) 188*, 2 (2001), 531535.

[15] Floro, J. A., Follstaedt, D. M., Provencio, P., Hearne, S. J., and Lee, S. R. Misfit dislocation formation in the AlGaN/GaN heterointerface. *Appl. Phys. Lett. 94* (2003), 1565.

[16] Follstaedt, D., Lee, S., P.P. Provence, A. A., Floro, J., and Crawford, M. Relaxation of compressively-strained algan by inclined threading dislocations. *J. Appl. Phys. 87*, 12 (2014), 121112.

[17] Freund, L. B. The stability of a dislocation threading a strained layer on a substrate. *J. Appl. Mech. 54* (1987), 554.

[18] Freund, L. B. A criterion for arrest of a threading dislocation in a strained epitaxial layer due to an interface misfit dislocation in its path. *J. Appl. Phys. 68* (1990), 2073.

[19] Freund, L. B., and Suresh, S. *Thin Film Materials*. Cambridge University Press, 2003.

[20] Gherasimova, M., Cui, G., Ren, Z., Su, J., Wang, X.-L., Han, J., Higashimine, K., and Otsuka, N. Heteroepitaxial evolution of AlN on GaN grown by metal-organic chemical vapor deposition. *J. Appl. Phys. 95* (2004), 2921.

[21] Gibarth, P. MOVPE of GaN and lateral overgrowth. *Rep. Prog. Phys.* (2004).

[22] Harame, D., Boquet, J., Ostling, M., Yee, Y., Masini, G., Caymax, M., Krishnamohan, T., Tillack, B., Bedell, S., Miyazaki, S., Reznicek, A., and Koester, S. *SiGe, Ge, and Related Compounds 4: Materials, Processing, and Devices*, vol. 33 of *6*. The Electrochemical Society, 2003.

[23] Hirth, J. P., and Lothe, J. *Theory of Dislocations*. Krieger Publishing Company, 1982.

[24] Holec, D. Critical thickness calculations for InGaN/GaN systems, june 2006.

[25] Holec, D. Multi-scale modelling of III-nitrides: from dislocations to the electronic structure, july 2008.

[26] Holec, D., Costa, P. M. F. J., Kappers, M. J., and Humphreys, C. J. Critical thickness calculations for InGaN/GaN. *J. Cryst. Growth 303*, 1 (May 2007), 314–317.

[27] Holec, D., and Humphreys, C. J. Calculations of equilibrium critical thickness for non-polar wurtzite InGaN/GaN systems. *Mater. Sci. Forum.*

[28] Holec, D., Zhang, Y., Rao, D. V. S., Kappers, M., McAleese, C. J., and Humphreys, C. J. Equilibrium critical thickness for misfit dislocations in III-nitrides. *J. Appl. Phys. 104*, 12 (2008), 123514.

[29] Hopcroft, M. A., Nix, W., and Kenny, T. What is the youngs modulus of silicon?. *J. Microelectromech. S. 19*, 2 (2010), 229.

[30] Houghton, D. C., Gibbings, C. J., Tuppen, C. G., Lyons, M. H., and Halliwell, M. A. G. Equilibrium critical thickness for SiGe strained layers on (100)Si. *Appl. Phys. Lett. 56* (1990), 460.

[31] Hu, S. Misfit dislocations and critical thickness of heteroepitaxy. *J. Appl. Phys. 69*, 11 (1991), 7901–7903.

[32] Hull, D., and Bacon, D. J. *Introduction to Dislocations.* Butterworth-Heinemann, 2011.

[33] Jahnen, B., Albrecht, M., Dorsch, W., Christiansen, S., Strunk, H. P., Hanser, D., and Davis, R. F. Pinholes, dislocations and strain relaxation in InGaN. *MRS Internet J. Nitride Semicond. Res. 3* (1998), 39.

[34] Jasinski, J., and Liliental-Weber, Z. Extended defects and polarity of hydride vapor phase epitaxy GaN. *J. Electron. Mater. 31*, 5 (2002), 429436.

[35] K. Krishnamurthy, T. Driver, R. V. J. M. 100W GaN HEMT power amplifier module with >60% efficiency over 100-1000 MHz bandwidth. *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International* (2010), 940 – 943.

[36] Kaiser, F., Jakob, M., Zweck, J., Gebhardt, W., Ambacher, O., Dimitrov, R., Schremer, A. T., Smart, J. A., and Shealy, J. R. High-electron-mobility AlGaN/GaN heterostructures grown on Si(111) by molecular-beam epitaxy. *Appl. Phys. Lett 18* (2000), 733.

[37] Kioseoglou, J., Komninou, P., and Karakostas, T. Core models of a-edge threading dislocations in wurtzite III(Al,Ga,In)-nitrides. *Phys. Status Solidi A 206*, 8 (2009), 1931.

[38] Koblmüller, G., Chu, R. M., Raman, A., Mishra, U. K., and Speck, J. S. High-temperature molecular beam epitaxial growth of AlGaN/GaN on GaN templates with reduced interface impurity levels. *J. Appl. Phys. 107* (2010), 043527.

[39] Kravchuk, K. S., Mezhennyi, M. V., and Yugova, T. G. Determination of the types and densities of dislocations in gan epitaxial layers of different thicknesses by optical and atomic force microscopy. *Crystallogr. Rep. 57*, 2 (2012), 277–282.

[40] Landau, L. D., and Lifshitz, E. M. *Theory of elasticity.* Pergamon Press, 1959.

[41] Lee, S. R., Koleske, D. D., Cross, K. C., Floro, J. A., Waldrip, K. E., Wise, A. T., and Mahajan, S. In situ measurements of the critical thickness for strain relaxation in AlGaN/GaN heterostructures. *J. Appl. Phys. 85* (2004), 6164.

[42] Li, G., Meng, Q., Yang, L., and Li, C. Molecular dynamics study on the 60° dislocation in silicon. *J. Atom. Molec. Phys.* (2006), 71–74.

[43] Li, L., Li, D., Wei, Q., Chen, W., Yang, Z., Zhang, G., and Hu, X. Inclined dislocation generation in compressive-strain-enhanced Mg-doped GaN/$Al_{0.15}Ga_{0.85}N$ superlattice with AlN interlayer. *Appl. Phys. Express 6* (2013), 061002.

[44] Liu, R., Mei, J., Srinivasan, S., Ponce, F. A., Omiya, H., Narukawa, Y., and Mukai, T. Generation of misfit dislocations by basal-plane slip in InGaN/GaN heterostructures. *Appl. Phys. Lett.*, 89 (2006), 201911.

[45] Love, A. E. H. *A treatise on the mathematical theory of elasticity.* Cambridge University Press, 1904.

[46] Lü, W., Li, D. B., Li, C. R., and Zhang, Z. Generation and behavior of pure-edge threading misfit dislocations in $In_xGa_{1-x}N$/GaN multiple quantum wells. *J. Appl. Phys.*, 96 (2004), 5267.

[47] Manuel, J., Morales, F., Garcia, R., Aidam, R., Kirste, L., and Ambacher, O. Threading dislocation propagation in AlGaN/GaN based hemt structures grown on Si(111) by plasma assisted molecular beam epitaxy. *J. Electron. Mater. 357* (2012), 3541.

[48] Markov, I. V. *Crystal Growth for Beginners: Fundamentals of Nucleation, Crystal Growth, and Epitaxy.* Singapore: World Scientific, 1995.

[49] Mathis, S., Romanov, A., Chen, L., Beltz, G., Pompe, W., , and Speck, J. Modeling of threading dislocation reduction in growing GaN layers. *J. Cryst. Growth 231*, 3 (2001), 371390.

[50] Matthews, J. W., and Blakeslee, A. E. Defects in epitaxial multilayers, 1. misfit dislocations. *J. Cryst. Growth 27* (1974), 118.

[51] Metzger, T., Hopler, R., Born, E., Ambacher, O., and Stutzmann, M. Defectstructure of epitaxial GaN films determined by transmission electron microscopy and triple-axis x-ray diffractometry. *Phil. Mag.*, A77 (1998), 1013.

[52] Moore, A., and Jimenez, J. *GaN RF for technology.* John Wiley & Sons, 2014.

[53] Morkoc, H. *Handbook of Nitride - Semiconductors and Devices*, vol. 1. WILEY-VCH Verlag, 2008.

[54] Morkoc, H., Strite, S., Gao, G. B., Lin, M. E., Sverdlov, B., and Burns, M. Large-band-gap SiC, III-V nitride, and II-VI ZnSe-based semiconductor device technologies. *J. Appl. Phys. 76*, 3 (1994), 1363.

[55] Nabarro, F. R. N. *Dislocations in Solids: Dislocations in crystals.* North-Holland Publishing Company, 1979.

[56] Nakamura, S., and Krames, M. R. History of gallium-nitride-based light-emitting diodes for illumination. *Proceedings of the IEEE* (2013), 0018–9219.

[57] Nye, J. F. *Physical Properties of Crystals – Their Representation by Tensors and Matrices.* Clarendon Press Oxford, 1985.

[58] Oliver, R. A., Kappers, M. J., Sumner, J., Datta, R., and Humphreys, C. J. Highlighting threading dislocations in movpe grown GaN using an in-situ treatment with silane and ammonia. *J. Cryst. Growth* (2006), 289.

[59] Pankove, J. I., and Moustakas, T. D. *Gallium Nitride (GaN) II, semiconductors and semimetals*, vol. 57. INSPEC, the institution of electrical engineers, 1999.

[60] Parker, C. A., Roberts, J. C., Bedair, S. M., Reed, M. J., Liu, S. X., and El-Masry, N. A. Determination of the critical layer thickness in the InGaN/GaN heterostructures. *Appl. Phys. Lett. 75* (1999), 2776.

[61] Reeber, R. R., and Wang, K. High temperature elastic constant prediction of some group iii- nitrides. *MRS Internet J. Nitride Semicond. Res. 6* (2001), 1–5.

[62] Reed, M. J., and El-Masry, N. A. Critical layer thickness determination of GaN/InGaN/GaN double heterostructures. *Appl. Phys. Lett. 77* (2000), 4121.

[63] Remediakis, I. N., Jesson, D. J., and Kelires, P. C. Probing the structure and energetics of dislocation cores in SiGe alloys through Monte Carlo Simulations. *Phys. Rev. Lett. 97* (2006), 255502.

[64] Romanov, A. E., Pompe, W., Beltz, G. E., and Speck, J. S. An approach to threading dislocation reaction kinetics. *J. Appl. Phys. 69* (1996), 3342.

[65] Romanov, A. E., Pompe, W., Beltz, G. E., and Speck, J. S. Modeling of threading dislocation density reduction in heteroepitaxial layers. *Phys. Stat. Sol. (b) 198* (1996), 599.

[66] Romanov, A. E., Pompe, W., Mathis, S., Beltz, G. E., and Speck, J. S. Threading dislocation reduction in strained layers. *J. Appl. Phys. 85*, 1 (1999), 182–191.

[67] Runton, D. W., Trabert, B., Shealy, J. B., and Vetury, R. History of GaN. *Microwave Magazine, IEEE 14* (2013), 82 – 93.

[68] Schremer, A. T., Smart, J. A., Wang, Y., Ambacher, O., MacDonald, N. C., , and Shealy, J. R. High electron mobility AlGaN/GaN heterostructure on (111)Si. *Appl. Phys. Lett 76*, 6 (2000), 733.

[69] Semond, F., Lorenzini, P., Grandjean, N., and Massies, J. High-electron-mobility AlGaN/GaN heterostructures grown on Si(111) by molecular-beam epitaxy. *Appl. Phys. Lett 78* (2001), 335.

[70] Sokolnikoff, I. S. *Mathematical theory of elasticity.* McGraw-Hill, 1956.

[71] Srinivasan, S., Geng, L., Liu, R., Ponce, F. A., Narukawa, Y., and Tanaka, S. Slip systems and misfit dislocations in InGaN epilayers. *Appl. Phys. Lett. 83* (2003), 5187.

[72] Steeds, J. W. *Introduction to anisotropic elasticity theory of dislocations.* Clarendon Press, 1973.

[73] T. Li, M. Mastro, A. D. *III-V compound semiconductor – integration with silicon-based microelectronics.* CRC Press, 2011.

[74] Ting, T. C. T. *Anisotropic Elasticity, Theory and Applications.* Oxford university press, 1996.

[75] Ubukataa, A., Ikenaga, K., Akutsua, N., Yamaguchia, A., Matsumoto, K., Yamazakia, T., and Egawab, T. GaN growth on 150-mm-diameter (111)Si substrates. *Journal of Crystal Growth 298* (2007), 198201.

[76] Vegard, L. Die konstitution der mischkristalle und die raumfllung der atome. *Zeitschrift fr Physik 5* (1921), 17–26.

[77] Vennéguès, P., Bougrioua, Z., Bethoux, J. M., Azize, M., and Tottereau, O. Relaxation mechanisms in metal-organic vapor phase epitaxy grown Al-rich (Al,Ga)N/GaN heterostructures. *J. Appl. Phys. 97* (2005), 024912.

[78] Wang, C., Pan, X., and Ruhle, M. Silicon nitride crystal structure and observations of lattice defects. *J. Mater. Sci. 31* (1996), 5281.

[79] Ward, T., Snchez, A., Tang, M., Wu, J., Liu, H., Dunstan, D. J., and Beanland, R. Design rules for dislocation filters. *J. Appl. Phys. 116* (2014), 063508.

[80] Warnes, W. Changing dislocation type along a single dislocation line. `http://oregonstate.edu/instruct/engr322/Exams/Previous/S98/ENGR322MT2.html`, 1998.

[81] Webb, J. B., Tang, H., Rolfe, S., and Bardwell, J. A. Semi-insulating C-doped GaN and high-mobility AlGaN/GaN heterostructures grown by ammonia molecular beam epitaxy. *Appl. Phys. Lett. 75* (1999), 953–955.

[82] Willis, J. R., Jain, S., and Bullough, R. The energy of an array of dislocations - implications for strain relaxation in semiconductor heterostructures. *Philos. Mag. A A 62* (1990), 115.

[83] Wu, Y.-F., Saxler, A., Moore, M., Wisleder, T., Mishra, U., and Parikh, P. Field-plated GaN HEMTs and amplifiers. *Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC 2005.* (2005), 170 – 172.

[84] Yoshida, S., Katoh, S., Takehara, H., Takehara, Y., Li, J., Ikeda, N., Hataya, K., and Sasaki, H. Investigation of buffer structures for the growth of a high quality AlGaN/GaN heterostructure with a high power operation FET on Si substrate using MOCVD. *Phys. Stat. Sol. (a) 203*, 7 (2006), 1739 1743.

[85] Zhe, L., Xiao-Liang, W., Jun-Xi, W., Guo-Xin, H., Lun-Chun, G., and Jin-Min, L. The influence of AlN/GaN superlattice intermediate layer on the properties of GaN grown on Si(111) substrates. *Chin. Phys. 16*, 5 (2007), 1467–05.

[86] Zolper, J. C. Wide bandgap semiconductor microwave technologies: From promise to practice. *International Electron Devices Meeting, 1999. IEDM 1999* (1999), 389 – 392.

List of Publications

Publications in Scientific Journals

[1] **R. A. Coppeta**, D. Holec, H. Ceric, T. Grasser: "Evaluation of dislocation energy in thin films"; Philosophical Magazine, Vol. 95, Iss. 2, January 2015, pp. 186-209, DOI: 10.1080/14786435.2014.994573.

Publications in Conference Proceedings

[1] **R. A. Coppeta**, H. Ceric, D. Holec, T. Grasser: "Critical thickness for GaN thin film on AlN substrate". IEEE International Integrated Reliability Workshop (IIRW), Stanford Sierra Conference Center, South Lake Tahoe, CA, USA; 2013-10-13 - 2013-10-17; in: "Proceedings of the International Integrated Reliability Workshop 2013", (2013), pp. 133-136.

[2] **R. A. Coppeta**, H. Ceric, T. Grasser: "Epitaxial Volmer-Weber Growth Modelling". Talk at the International Conference on Simulation of Semiconductor Processes and Devices (SISPAD); 2013-09-05 - 2013-09-05; in: "Proceedings of the International Conference on Simulation of Semiconductor Processes and Devices 2013", (2013), pp. 45-48.

Curriculum Vitae

March 22$^{\text{th}}$ 1986

Born in Torino, Italy

July 2005

Institute ”Aldo Moro”, Rivarolo Canacese, Italy
Scientific Gymnasium graduation

September 2005 – March 2009

Polytechnic of Torino, Torino, Italy
Bachelor of Material Engineering
Bachelor Thesis: “Electromigration in gold wires”

July 2008 – August 2009

χ-Lab Materials and Microsystems Laboratory, Chivasso, Italy
Part-time job
Modeling and experimental analysis of electromigration in gold nanowires

March 2009 – December 2011

Polytechnic of Torino, Torino, Italy
Master of Material Engineering
Master Thesis: “The relaxation of a strained heteroepitaxial structure”

February 2011 – November 2011

MEMC Electronic Materials, Novara, Italy
Internship
Mathematical modeling and experimental validation of mechanical behavior of silicon epitaxial structures

March 2012 – April 2015

Vienna University of Technology, Vienna, Austria
PhD program at the Institute for Microelectronics